University of Cape Town

Department of Mathematics & Applied Mathematics

ERRORS IN SOUTH AFRICAN SECONDARY SCHOOL MATHEMATICS TEXTBOOKS

By

Carol Harris

A thesis prepared under the supervision of Professor J H Webb, in fulfilment of the requirements of the degree of Master of Science in Mathematics.

CONTENTS

ACKNOWLEDGEMENTS

There are several people who have played major roles in helping me complete this project. There are no suitable words that can express my appreciation to my supervisor, Professor John Webb. His expertise and guidance have been my fuel in getting this project in motion and keeping me on track. I am especially grateful for his patience and understanding of all the dynamic variables in my life that have at times added to the challenges of completing this thesis.

My thanks go to the University of Cape Town Mathematics Education Project for their assistance in providing source material. I am also grateful to many in the Department of Mathematics and Applied Mathematics at U.C.T. for valuable advice and assistance. My thanks also go to the Department of Statistical Sciences at U.C.T for the use of their computers.

I am grateful for the support of my family and of my husband for their motivation and encouragement.

Carol Harris February 1998

INTRODUCTION

Since 1960 various new topics were introduced into the South African mathematics syllabus for high schools. This was generally described as “New Maths”.

These topics were characterised by an emphasis on mathematical structure (set theory, groups, rings and fields, the construction of the real number system, relations and functions, vectors and mathematical induction).

Textbooks that were published in South Africa during this time have been found to contain numerous significant mathematical errors, especially in connection with these new topics. This project has entailed the collection, classification and commentary on these problems. It has involved a study of a range of texts from Standards 6 to 10.

The methodology utilised entailed reading as many maths textbooks as we could find in libraries such as Education libraries. Other sources of textbooks were from the library of the Mathematics Education Project (MEP)(of the University of Cape Town), private collections and second hand bookshops.

Some misconceptions arise across a range of texts indicating that either a general misunderstanding has occurred or that authors have used one another's work in their research.

We conducted a search for mathematical errors and not minor misprints, arithmetical slips, algebraic errors, mistakes in answers at the back of the book, spelling or grammatical mistakes.

Nor did we seek out issues of weak didactics, poor wording, poor style, gender or social bias. These could be matters of opinion. We have concentrated on issues that unambiguously demonstrate fundamental lack of understanding of important mathematical concepts on the part of the authors.

We know of no comparable study in South Africa or overseas. Errors in overseas texts seem to be rare in comparison with their South African counterparts. Overseas publishers have a more rigorous selection, refereeing and editorial policy. Overseas journals for mathematics teachers (e.g. Mathematical Gazette) have extensive book reviews. The main journal for mathematics teachers in South Africa (Pythagoras) publishes relatively few reviews, usually not of textbooks.

This work has been divided into 9 chapters, each discussing a particular area in mathematics. Photocopied extracts from the textbooks are included and are presented inside rectangular "boxes". Abbreviations and page numbers are used to show the source of the extract. The explanations of these abbreviations are included in the first part of the Bibliography. Any mathematics not in rectangular boxes is used to explain concepts correctly and to supply correct proofs. Where these have been based on works found in published literature, the number appearing at the end of the excerpt refers to the reference as documented in the second part of the Bibliography.

1. LANGUAGE, LOGIC AND PROOF

1. Introduction

One often hears mathematics being referred to as a language, even though not as a natural one such as English or French.

In the book "Language in Mathematical Education" David Pimm writes an article in which he says, "One way of describing the relation between mathematics and a natural language such as English is in terms of the linguistic notion of register." [1]

Pimm states that linguist Michael Halliday considers "register" as "a set of meanings that is appropriate to a particular **function** of language, together with the words and structures which express these meanings."

Frequently register in mathematics differs from register in the English language. The mathematical meanings of words get confused with their English counterparts. In this section we take a look at meanings of words as used in mathematics as opposed to the meanings in usual English terms.

Any work in mathematics requires correct definitions, correct use of language and correct logical deductions. This will ensure effective communication and a successful system of proof in mathematics.

In this section the issues that are studied are:

- misuse of logical connectives,
- mathematical versus English use of terms,
- logically invalid inference.

2. Logical connectives

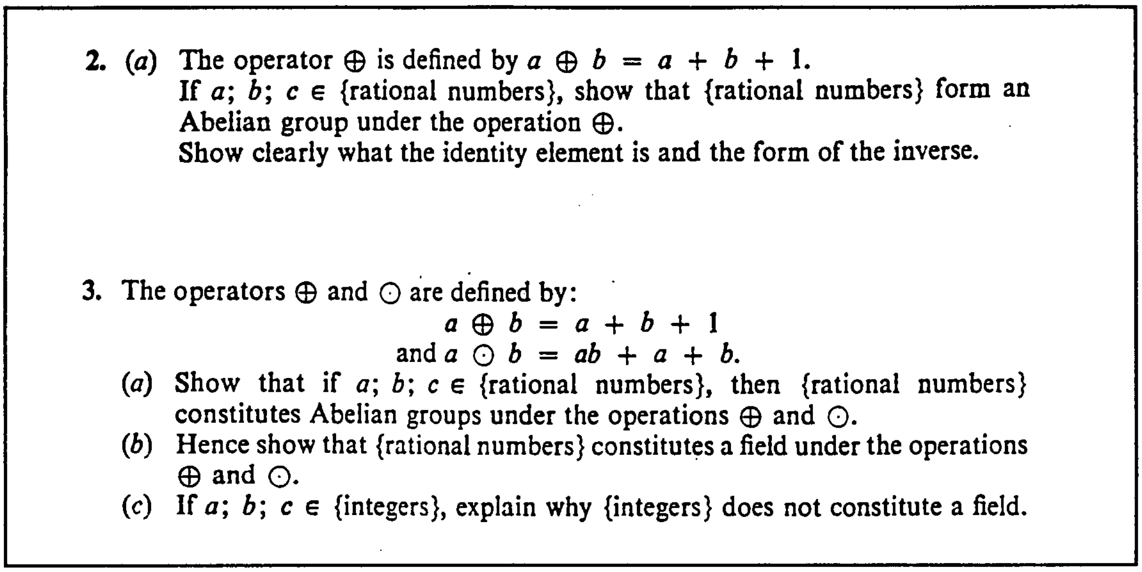
The word "if " has a precise logical use which is often at variance with common English usage of the word.

NAM 8, pg. 5, writes the following:

> **Note**
> If the rational number $\frac{22}{7}$ represents the irrational π, it is still true that $\pi \neq \frac{22}{7}$.

In this text the writer seems to be saying "if $\pi = \frac{22}{7}$, then $\pi \neq \frac{22}{7}$". What is meant is, "While the rational number $\frac{22}{7}$ is often used to represent π, it is still true that $\pi \neq \frac{22}{7}$."

MGM 9&10, pg. 173 and 209, respectively, provides two examples of the misuse of "if".



2. (*a*) The operator $\oplus$ is defined by $a \oplus b = a + b + 1$.
If a; b; $c \in$ {rational numbers}, show that {rational numbers} form an Abelian group under the operation $\oplus$.
Show clearly what the identity element is and the form of the inverse.

3. The operators $\oplus$ and $\odot$ are defined by:

$$a \oplus b = a + b + 1$$
$$\text{and } a \odot b = ab + a + b.$$

(*a*) Show that if a; b; $c \in$ {rational numbers}, then {rational numbers} constitutes Abelian groups under the operations $\oplus$ and $\odot$.
(*b*) Hence show that {rational numbers} constitutes a field under the operations $\oplus$ and $\odot$.
(*c*) If a; b; $c \in$ {integers}, explain why {integers} does not constitute a field.

The fact that the rational numbers together with these operations form an abelian group is not a consequence of $a, b, c \in \{\text{rational numbers}\}$. In other words it is not the case that $a, b, c \in \{\text{rational numbers}\}$ implies that **Q** together with these operations form an abelian group.

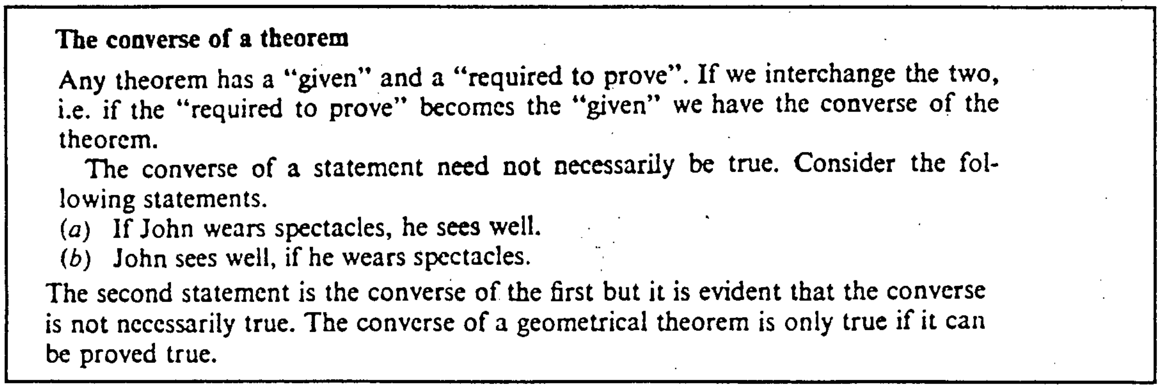
What has been confused here is what is meant by "if" and "then". It would be correct to say:

"Show that if $a, b, c \in \mathbf{Q}$ then $a \oplus b \in \mathbf{Q}$ and $a \odot b \in \mathbf{Q}$."

A concept that is closely related to the "if-then" concept is that of converse. The most common approach is: the converse of an "if-then" statement is given by interchanging the clauses that follow "if" and "then" respectively.

The concept of converse appears to be a confusing issue in the following text. The extract from this text shows that the concept has been misunderstood and has been expressed in a way that will guarantee confusion.

MGM 8, pg.213, says the following about converses:



The converse of a theorem

Any theorem has a "given" and a "required to prove". If we interchange the two, i.e. if the "required to prove" becomes the "given" we have the converse of the theorem.

The converse of a statement need not necessarily be true. Consider the following statements.

(*a*) If John wears spectacles, he sees well.

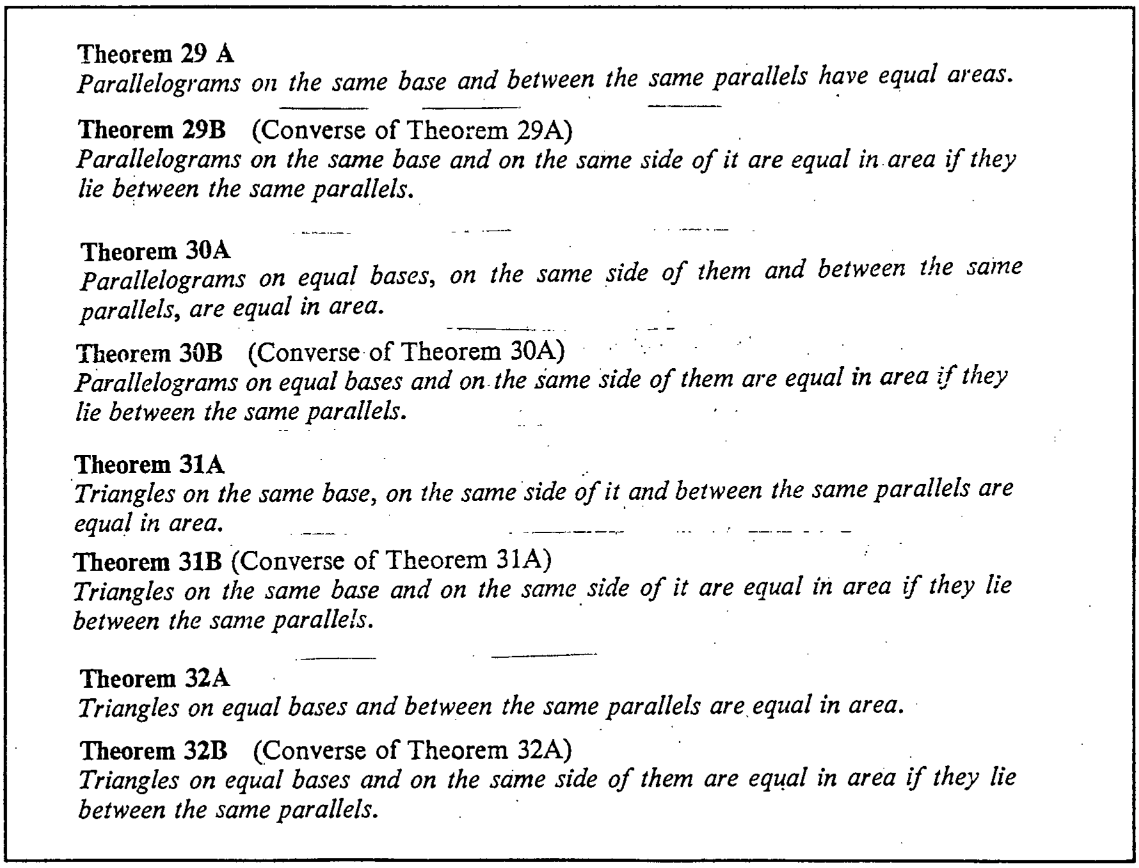
(*b*) John sees well, if he wears spectacles.

The second statement is the converse of the first but it is evident that the converse is not necessarily true. The converse of a geometrical theorem is only true if it can be proved true.

The second statement is not a converse of the first statement. The two statements are logically equivalent. It is not "evident" that the converse is not true.

In an "if-then" statement there are conditions and conclusions. If a new statement is made with the original conclusions becoming the new conditions and the original conditions becoming the new conclusions, then the new statement is the converse of the original statement.

MGM 9&10, pp. 384 to 388, states the following:



Theorem 29 A
Parallelograms on the same base and between the same parallels have equal areas.

Theorem 29B (Converse of Theorem 29A)
Parallelograms on the same base and on the same side of it are equal in area if they lie between the same parallels.

Theorem 30A
Parallelograms on equal bases, on the same side of them and between the same parallels, are equal in area.

Theorem 30B (Converse of Theorem 30A)
Parallelograms on equal bases and on the same side of them are equal in area if they lie between the same parallels.

Theorem 31A
Triangles on the same base, on the same side of it and between the same parallels are equal in area.

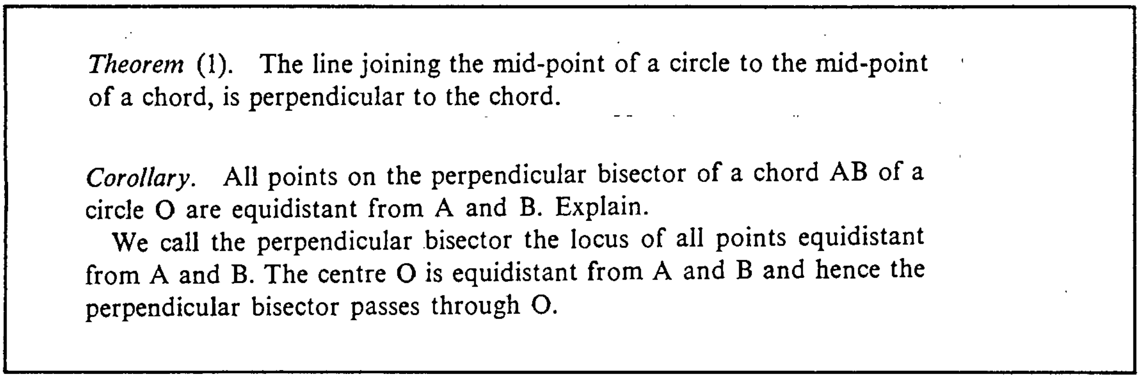
Theorem 31B (Converse of Theorem 31A)
Triangles on the same base and on the same side of it are equal in area if they lie between the same parallels.

Theorem 32A
Triangles on equal bases and between the same parallels are equal in area.

Theorem 32B (Converse of Theorem 32A)
Triangles on equal bases and on the same side of them are equal in area if they lie between the same parallels.

These are not "converses", but the actual theorems stated in a different way. (The proofs which the authors give are correct proofs of the true converses.)

MM 9&10, pg. 180, confuses the concept of converse and corollary:



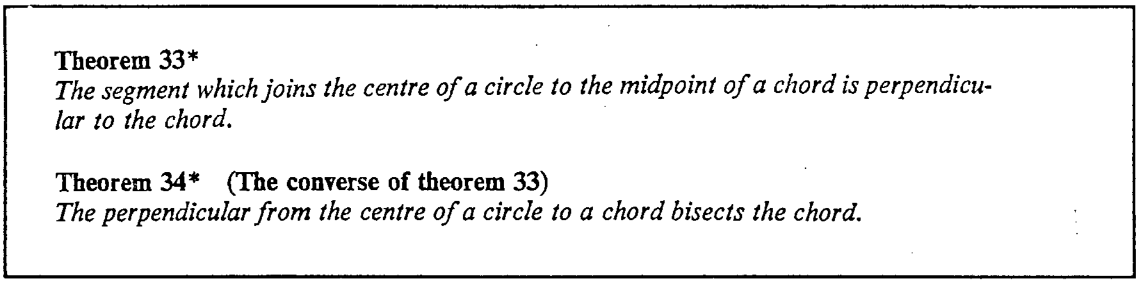
Theorem (1). The line joining the mid-point of a circle to the mid-point of a chord, is perpendicular to the chord.

Corollary. All points on the perpendicular bisector of a chord AB of a circle O are equidistant from A and B. Explain.

We call the perpendicular bisector the locus of all points equidistant from A and B. The centre O is equidistant from A and B and hence the perpendicular bisector passes through O.

This is not a corollary, but a more general statement from which the converse follows, namely: the perpendicular bisector of a chord passes through the midpoint of a circle.

MGMSC, pg. 344, says the following:



Theorem 33*

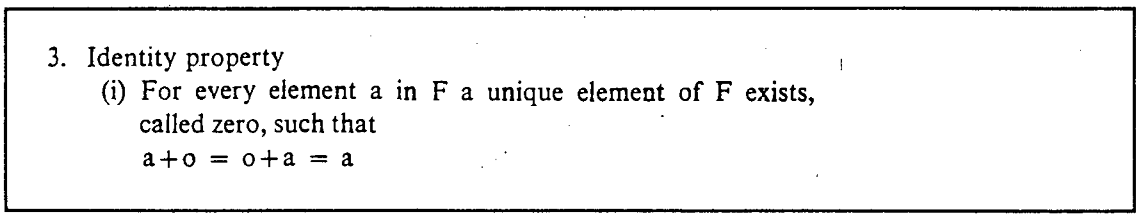
The segment which joins the centre of a circle to the midpoint of a chord is perpendicular to the chord.

Theorem 34* **(The converse of theorem 33)**

The perpendicular from the centre of a circle to a chord bisects the chord.

This example is included to emphasize that there could be two converses to a theorem. Theorem 34, above, is not **the** converse of theorem 33, as seen in the extract prior to this one. In terms of logical connectives the statement "$A \& B \Rightarrow C$" has two converses: "$C \& A \Rightarrow B$" and "$C \& B \Rightarrow A$".

The placing of a logical quantifier in a sentence is important. In logical terms and symbols what is meant by $(\forall a)(\exists b)P(a,b)$ is different from $(\exists b)(\forall a)P(a,b)$. MBM 9&10, pg. 23 ,says the following:



3. Identity property
   (i) For every element a in F a unique element of F exists, called zero, such that
   a+o = o+a = a

> (ii) For every element a in F a unique element of F exists, called unity, such that
> $$a \times 1 = 1 \times a = a$$

This implies that each element has its own zero or unity, which may vary with each different element. This is not the case and the correct statement would be: "A unique element $0 \in F$ ($1 \in F$) exists, called zero (unity), such that for every element a in F $a + 0 = 0 + a = a$ ($a \times 1 = 1 \times a = a$)."

MGM 9&10, pg. 205, says the following:

> $F = \{a; b; c; \ldots\}$ **constitutes a field if:**
>
> 1. The *closure property* holds for *addition* and *multiplication*,
>   i.e. if $(a \oplus b) \in F$ and $a \otimes b \in F$
> 2. *The operators are associative in F,*
>   i.e. if $(a \oplus b) \oplus c = a \oplus (b \oplus c)$
>   and $a \otimes (b \otimes c) = (a \otimes b) \otimes c.$
> 3. *There exist identity elements i and e* such that
> $a \oplus i = a$
>   and $a \otimes e = a$
>   Very often $i = 0$ and $e = 1$.
> 4. *Inverse elements exist for each* $a \in F$
>   i.e. $a \oplus \tilde{a} = i$ and very often $\tilde{a} = -a$
>   and $a \otimes \acute{a} = e$ and very often $a' = \frac{1}{a}$
>   The notations $\tilde{a}$ and a' denote inverses of a.
> 5. *The commutative property holds for both operations,*
>   i.e. $a \oplus b = b \oplus a$
>   and $a \otimes b = b \otimes a$
>
> We note that so far a field constitutes Abelian groups for both $\oplus$ and $\otimes$.
>
> 6. *The distributive property of the second operator over the first is satisfied,*
>   i.e. $a \otimes (b \oplus c) = (a \otimes b) \oplus (a \otimes c).$

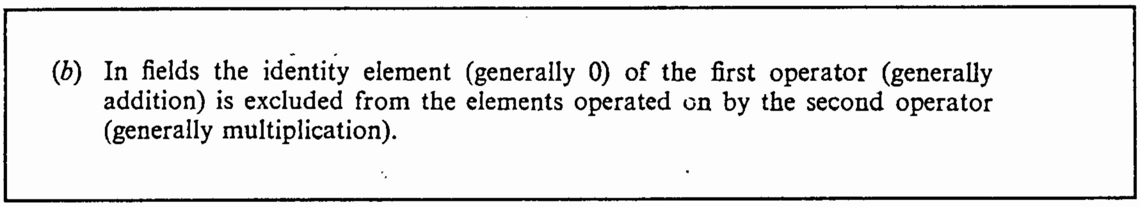
Listing F in set notation could lead the reader to believe that F is countable, which is not a necessary criterion. Logical connectives have been used incorrectly or quantifiers have been omitted. In part **1** the statement after the "i.e." is incomplete. It should read: "For all a and b in F $(a \oplus b) \in F$ and $(a \otimes b) \in F$ ". Similarly **2** should read as "For all $a, b, c \in$ F, $(a \oplus b) \oplus c = a \oplus (b \oplus c)$ and $(a \otimes b) \otimes c = a \otimes (b \otimes c)$." In part **3** there is also an omission of a quantifier. The statement is corrected as follows, "There exits identity elements i and e such that for all $a \in$ F $a \oplus i = a$ and $a \otimes e = a$."

The statement that "**very often** $i = 0$ and $e = 1$" could be better expressed as "usually i is **represented** by 0 and e is **represented** by 1." The impression that is created here is that i and e take on the meaning 0 and 1 and all the associated characteristics. Similarly in **4** the statements “very often $\tilde{a} = -a$” and “very often $a' = \frac{1}{a}$” should rather be stated as “usually $\tilde{a}$ is represented by $-a$” and “usually a' is represented by $\frac{1}{a}$” respectively. Finally it should be noted that the additive identity does not have a multiplicative inverse. (This will be discussed in more detail in chapter 9.)

3. Language

The difference between the English and the mathematical use of words is subtle, but important. As already seen with the use of the word “if”, confusion can arise when it is not used in its correct mathematical context. There are several terms which have one connotation in mathematics and another in the usual English sense of the word, for example the word “unique”. The usual meaning of the word is “sole: without a like” [2]. For example, “she has a unique car.” This word is important when dealing with functions. A function is such that for one input value x there is a **unique** output value y. This means that there is **only one** output value. This does not mean that there cannot be two of the same output values (i.e. for two different input values of x there can be the same output value) whereas in “she has a unique car”, “**unique**” implies there is no other car like it . This, however, does not restrict her from having a second (non-unique) car.

MGM 9&10, pg. 207, writes the following:

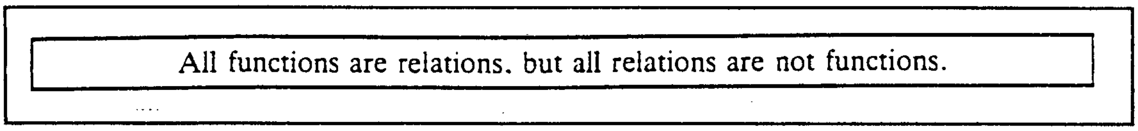


(*b*) In fields the identity element (generally 0) of the first operator (generally addition) is excluded from the elements operated on by the second operator (generally multiplication).

The common English meaning of the word “generally” is “usually”, whereas mathematically it means “universally” or “always”. The above extract should have used the term “usually” rather than “generally”, as it is not universal that the first operation is addition and that the identity element is zero. Strictly speaking zero, addition and multiplication are concepts that are associated with a **particular** field. For the general theory of fields, it is for convenience that these operations are called addition and multiplication and that the identity elements are denoted by the symbols used for “normal” addition and multiplication (i.e. that which is associated with the field of real numbers).

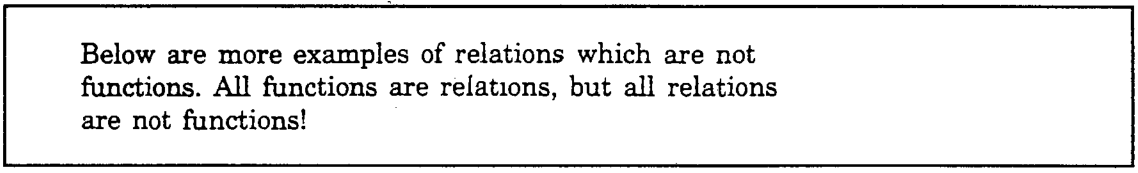
The operations are binary and therefore one cannot refer to elements being “operated on by the second operator.” This statement also implies something that is false. The second operation is defined for all elements including the identity of the first operation. The inverse with respect to the second operation is not defined for the identity of the first operation.

NMM 8, pg.75, says:



All functions are relations, but all relations are not functions.

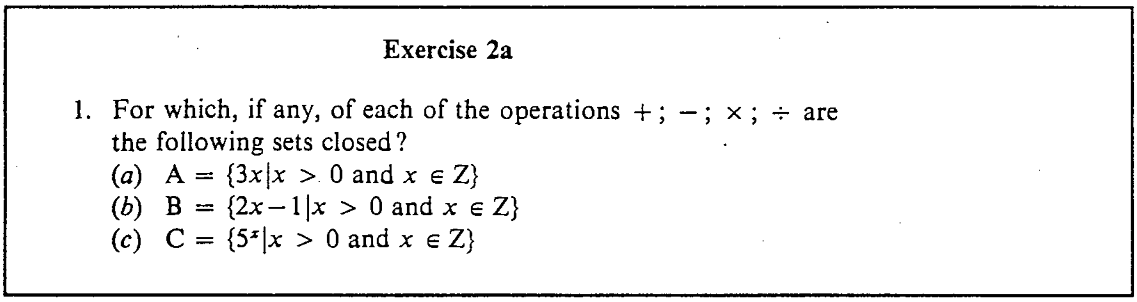
MP 8, pg. 101, has a similar statement:



Below are more examples of relations which are not functions. All functions are relations, but all relations are not functions!

In these extracts it is merely the case that the word "not" has been misplaced but this has resulted in confusing what is meant by a relation versus a function.

MBM 9&10, pg. 16, says the following:

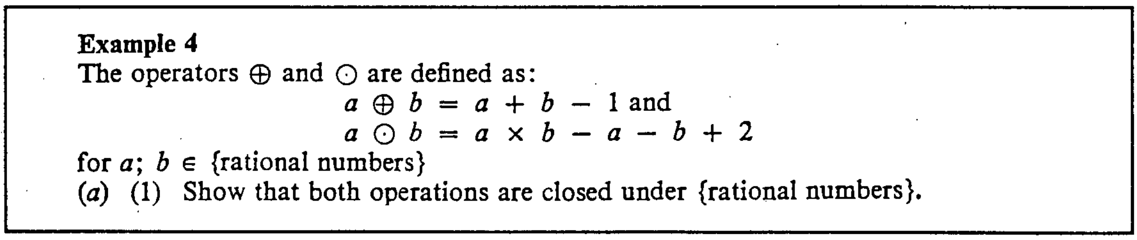


Exercise 2a

1. For which, if any, of each of the operations $+$; $-$; $\times$; $\div$ are the following sets closed?
   (*a*) $A = \{3x | x > 0 \text{ and } x \in Z\}$
   (*b*) $B = \{2x-1 | x > 0 \text{ and } x \in Z\}$
   (*c*) $C = \{5^x | x > 0 \text{ and } x \in Z\}$

In mathematical terms we do not say "sets are closed **for** addition, subtraction, multiplication or division, but rather "sets are closed **under** certain operations."

MGM 9&10, pg. 207, says:



Example 4
The operators $\oplus$ and $\odot$ are defined as:

$$a \oplus b = a + b - 1 \text{ and}$$
$$a \odot b = a \times b - a - b + 2$$

for a; $b \in$ {rational numbers}
(*a*) (1) Show that both operations are closed under {rational numbers}.

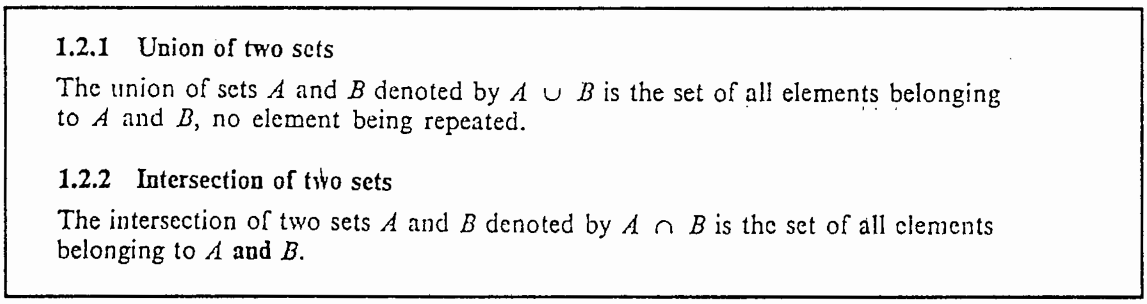
Operations are not “closed under” sets, but sets may be shown to be “closed under” certain operations

The words “**or**” and “**and**” are often misunderstood. Deciding when to use these two grammatical connectives in mathematics poses a problem. In the English language the word “or” is used in an exclusive context. The statement “A or B” means “either A otherwise B, but not both”. Mathematically the statement “A or B” is true if one or both A and B are true.

Students encounter difficulties in understanding the difference between the words “and” and “or”. For example, students often write the inequalities $x < -2$ or $x > 2$ as $-2 > x > 2$. This is incorrect since x cannot be greater than 2 **and** less than -2. However the answer to the question “What are the solutions to $(x-2)(x+2) > 0$?” is “$x < -2$ **and** $x > 2$”.

In simpler terms: The problem “Solve for x: $(x-1)(x+2) = 0$”, has answer “$x = 1$ **or** $x = -2$”. The answer to the question “what are the solutions to the equation $(x-1)(x+2) = 0$?”, has answer “$x = -2$ **and** $x = 1$ ”.

MGM 8, pg. 5 and 6, shows how the word “and” is used when “or” should be used:



1.2.1 Union of two sets

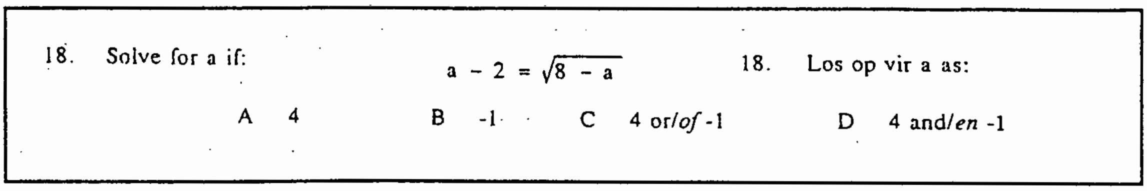
The union of sets A and B denoted by $A \cup B$ is the set of all elements belonging to A and B, no element being repeated.

1.2.2 Intersection of two sets

The intersection of two sets A and B denoted by $A \cap B$ is the set of all elements belonging to A **and** B.

The word “and” is used in the definitions of both the **union** and **intersection** of two sets. The latter use of the word “and” implies that the elements which are **common** to both sets are being considered, whereas in the first definition the meaning that is intended is that the elements of the one set **together with** elements of the other set form the union. In the second definition “and” should be replaced with the word “or”. The inclusion of the phrase ‘no element being repeated’ demonstrates a common misconception about sets.

The following extract is taken from the 1994 National Senior Certificate supplementary exam (The Answer series 1994, pg. 53):



18. Solve for a if: $a - 2 = \sqrt{8 - a}$ 18. Los op vir a as:

A 4 B -1 C 4 or/*of* -1 D 4 and/*en* -1

The desired answer was A, since -1 is not a solution. But if "4 is a solution" is correct, then " either 4 or -1 is a solution" is also correct, so that both A and C are correct responses.

As notation is a large part of the language of mathematics it is important to note that misuse of notation results in a lack of understanding of concepts.

MGM 9&10, pg.173, writes the following:

> 3. Consider the set of number pairs $\{x; y\}$ with $x; y \in Q$ and the operator defined by $(a; b) \oplus (c; d) = (a + c; b + d)$.
> Does Q = {rational numbers} constitute an Abelian group under the operation $\oplus$?

Set notation $\{x; y\}$ should not be used for an ordered pair (x, y). The operation $\oplus$ is not defined on **Q**, but on $\mathbf{Q} \times \mathbf{Q}$.

MGM 9&10, pg. 209, says:

> 2. (*a*) Show that $a - \sqrt{2}b$ constitutes a field if $a; b \in Q$, where Q represents {rational numbers}.
> (*b*) Why does $a - \sqrt{2}b$ not constitute a field if $a; b \in I$?

A field is a set together with operations and the criterion that certain conditions regarding these operations must be met. The expression "$a - \sqrt{2b}$" is not a set, so one cannot refer to it as a field. It is a general form of the elements that belong to a field, and hence can be used in describing a field.

4. Proofs and logical inferences

A frequent error in proving theorems is that the particular is used to prove the general. Understandably when proofs are too involved for high school level a brief outline of the proof and an example can be sufficient in substantiating any claims.

The following is a case where the particular is used to imply that the idea holds in general. MGM, pg. 207, says:

> We already know that Geometry is mainly concerned with figures and their properties. We examine a particular property of a figure in a practical way by construction and measurement in order to arrive at a general conclusion. As soon as we have ascertained by construction or in another way that a large number of figures have this characteristic, we may assume that all similar figures have this same characteristic. This fact may now be used in any other calculations, constructions and problems to discover new facts about figures.

This text is saying that the particular is sufficient to prove the general and this is not true.

MM 9&10, pg. 234, has the following proof of a trigonometric identity:

> 17.1 *DERIVING AND VERIFYING IDENTITIES* **(Higher Grade)**
>
> The open sentence $\sin x \cot x = \cos x$ implies:
>
> $$\sin x \frac{\cos x}{\sin x} = \cos x$$
>
> $$\Rightarrow \quad \cos x = \cos x$$
>
> Hence the sentence is true for all angles for which both sides are defined and is called an *identity*.

This example illustrates incorrect logical argument to prove an identity. This kind of argument is a common mistake amongst high school pupils. Often the identity which needs proof is written down as an equation and then each side of the equation is simplified. When each side of the equation has been simplified to the same expression then the proof is considered complete. These steps are not logical inferences because the first line is what is requiring proof. A logical proof begins with a sentence or statement that is true. A logical inference is a step that uses this true sentence with other established facts and then comes to a conclusion. One cannot start a proof with the statement that is requiring proof.

The 1976 the Joint Matriculation Board set a final std.10 mathematics examination with the following question:

> 6. (a) Prove that $\sec A \cdot \cos A - \dfrac{1}{\operatorname{cosec}^2 A} = \cos^2 A.$
>
> (Ignore the values of A for which the identity is not true.) (4)

The instruction is to ignore all values of A for which the identity is not true. Therefore the values of A under consideration are exactly those that make the statement true. Hence no further proof is required.

2. SET THEORY

1. Introduction

Set theory was introduced into South African schools (both primary and secondary) during the "New Maths" reforms of the 1960's. In the spirit of the axiomatic approach to mathematics, epitomised by the Bourbaki movement, it was promoted as the central theme of mathematics, and therefore school mathematics had to be built around set theory.

Set theory, in the Zermelo-Fraenkel sense, was not taught in schools. Instead, some basic terms (set, element, subset) were introduced, and simple manipulations were described (union, intersection), illustrated by Venn diagrams. Cardinality was brought in, but infinite sets were only briefly, and never satisfactorily, discussed.

Set theory was later phased out of the syllabus. In texts that were published at that time definitions given for the concept of a set and related topics were unclear and confusing.

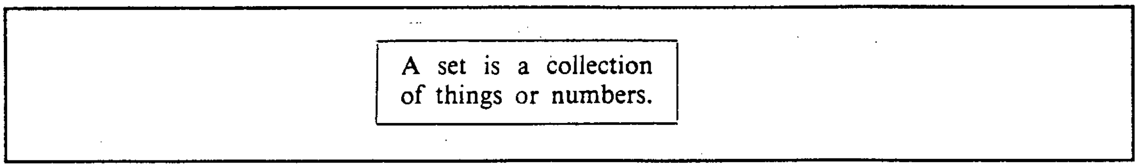
In this section the handling by texts of the basic concept of sets and related topics such as subsets, union and intersection, set notation and infinite sets will be investigated.

2. Definitions

It appears that defining the concept of a set has been a difficult task. Several text books that discuss set theory give contradicting opinions about explaining what is meant by a "set". Some texts say that it "cannot be defined" while others say that it is a collection of "objects", "things" or "numbers".

The following are references and the corresponding extracts which give definitions of "sets".

1. ASC, pg. 107:



A set is a collection of things or numbers.

2. MASAS, pg. 1:

Definition. A set is a group of things which have some quality in common.

These two are of the type "a set is a collection of objects", which was commonplace at the time (neither "collection" nor "object" being defined).

3. MGM 8, pg. 1:

1.1.1 Description of a set; elements; set-builder notation

(*a*) A set cannot be defined. Any group of objects or ideas having some common feature constitutes a set.

4. MGM 9&10, pg. 3:

We cannot define a set but we are able to recognize one. In Mathematics, set is the word used for any class or group of well-defined and distinguishable objects or concepts which are of such a nature that we can tell at once whether any given object or concept belongs to a particular group or not.

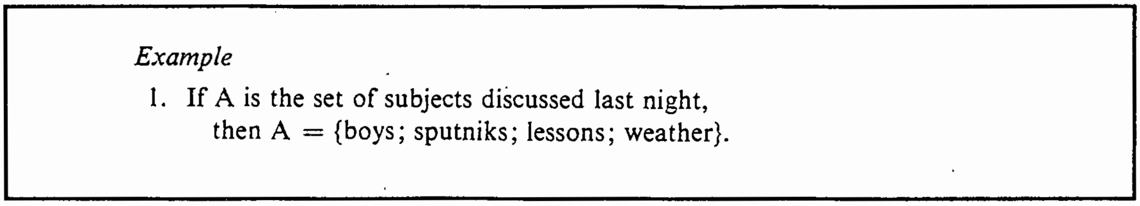
At the level of high school mathematics one can concede that the concept of a set is difficult to define in rigorous terms, but a general statement saying that it cannot be defined is not true.

5. MBM 8, pg. 1:

1. An **element** is anything at all provided it can be *clearly* and *uniquely* specified.

A SET is a collection of elements regarded as a whole.
$A = \{a; b; c\}$ means A is the set of the 3 elements a, b, and c.
$a \in A$ means *a is an element of the set* A.
$k \notin A$ means *k is not an element of the set* A.

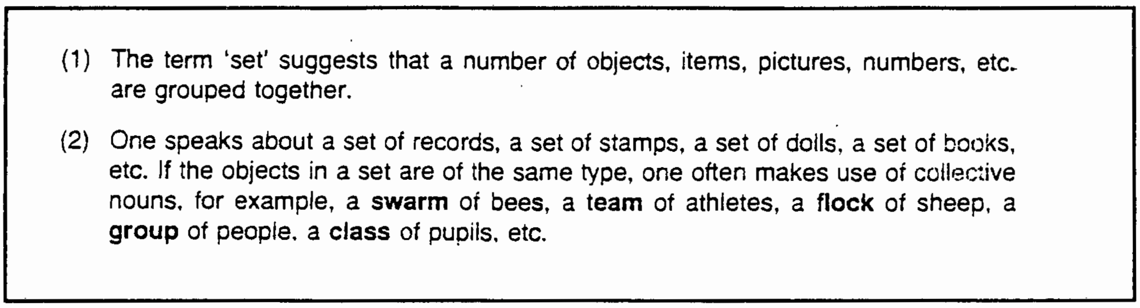
Generally the elements of a set are of the same kind (coins, numbers) but it is possible to have sets of dissimilar elements.



Example

1. If A is the set of subjects discussed last night,
then A = {boys; sputniks; lessons; weather}.

6. PM 5, pg. 1:



(1) The term 'set' suggests that a number of objects, items, pictures, numbers, etc. are grouped together.

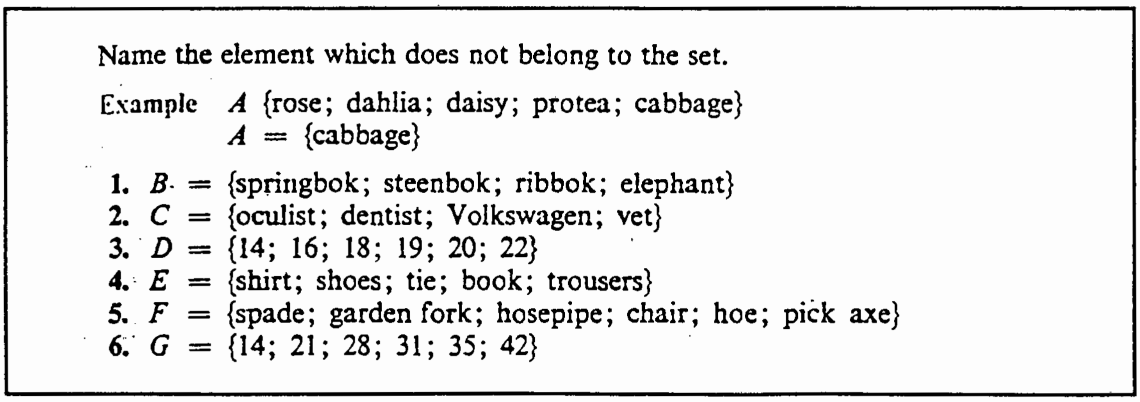
(2) One speaks about a set of records, a set of stamps, a set of dolls, a set of books, etc. If the objects in a set are of the same type, one often makes use of collective nouns, for example, a **swarm** of bees, a **team** of athletes, a **flock** of sheep, a **group** of people, a **class** of pupils, etc.

There seems to be contrasting opinions of what a set is. One view is that any collection of objects may form a set while another view is that objects must have a common characteristic in order to be a part of a set.

To give meaning to the word "set" is quite a difficult task. A collection of unrelated objects does form a set, but this may be too broad because we need to be able to decide whether a given object is a member of the set or not. This raises the question of what a well-defined set is.

From the texts above there is an overall opinion that a well-defined set is a collection of objects having some common characteristic or quality so that given any particular object one can decide whether or not it is a member of the set. How do we decide what the common quality is and hence what is meant by "well-defined"?

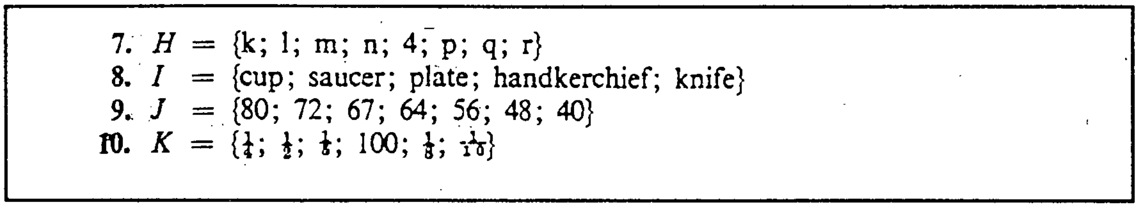
M 3, pg. 81, lists several sets and then asks which element does **not** belong to the set. This does not make sense because once the set has been listed each object in the list is an element of that particular set.



Name the element which does not belong to the set.

Example *A* {rose; dahlia; daisy; protea; cabbage}
A = {cabbage}

1. *B* = {springbok; steenbok; ribbok; elephant}
2. *C* = {oculist; dentist; Volkswagen; vet}
3. *D* = {14; 16; 18; 19; 20; 22}
4. *E* = {shirt; shoes; tie; book; trousers}
5. *F* = {spade; garden fork; hosepipe; chair; hoe; pick axe}
6. *G* = {14; 21; 28; 31; 35; 42}

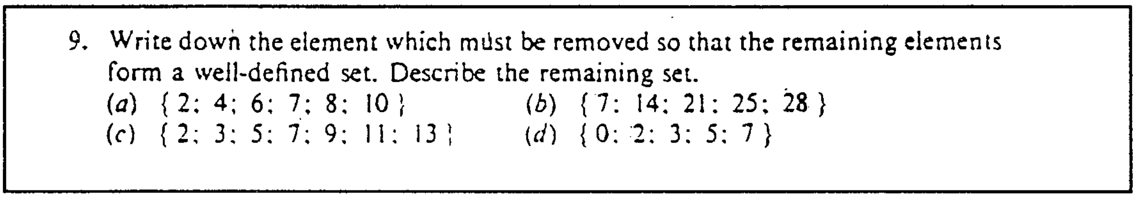


7. H = {k; l; m; n; 4; p; q; r}
8. I = {cup; saucer; plate; handkerchief; knife}
9. J = {80; 72; 67; 64; 56; 48; 40}
10. K = $\{\frac{1}{4}; \frac{1}{2}; \frac{1}{6}; 100; \frac{1}{8}; \frac{1}{16}\}$

There is a misuse of notation in the example that is given in the above exercise. Assuming that a printing error has resulted in the equality sign having been omitted, **A** is the five element set described in the first line of the example, but in the very next line **A** = {cabbage}. This example shows the misuse of notation as well as a misconception of the notion of a set. Once a set is defined by means of listing its elements, the question "which element does not belong to the set" cannot be asked.

If one followed the textbook slavishly, one would be required to write out solutions as follows:
1. B = {elephant}
2. C = {volkswagen} etc.

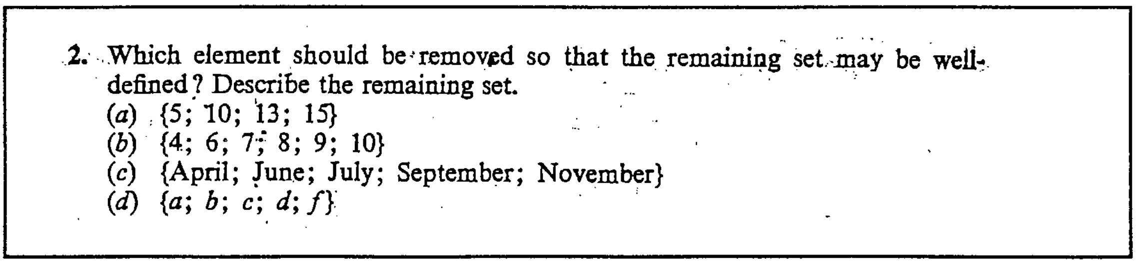
MGM 6, pg. 6, has the following exercise:



9. Write down the element which must be removed so that the remaining elements form a well-defined set. Describe the remaining set.
(*a*) { 2; 4; 6; 7; 8; 10 }
(*b*) { 7; 14; 21; 25; 28 }
(*c*) { 2; 3; 5; 7; 9; 11; 13 }
(*d*) { 0; 2; 3; 5; 7 }

In part (c) "9" could be left out and the remaining elements make up the "well-defined" set described as "the first six prime numbers", or "2" can be omitted to form the "set of the first six odd numbers greater than 1".

MGM 8, pg. 3, has a similar exercise, but nowhere in the text up to this point has the term "well-defined" been mentioned or explained.



2. Which element should be removed so that the remaining set may be well-defined? Describe the remaining set.
(*a*) {5; 10; 13; 15}
(*b*) {4; 6; 7; 8; 9; 10}
(*c*) {April; June; July; September; November}
(*d*) {*a*; *b*; *c*; *d*; *f*}

In part (d) omitting the letter "*f*" gives us the set consisting of the first four letters of the alphabet, which is a "well-defined" set, while leaving out the letter "*a*" gives the "well-defined" set of the first four consonants.

FMF III, pg. 21, gives this task, to distinguish between a "set" and a "well-defined set", as an initial exercise for the reader, without any previous discussion in the text. The debate is whether or not such a distinction does exist.

1. Try and explain in your own words what is meant by the following:
(*a*) a set (*b*) a well-defined set

Richard Skemp in his book "The Psychology of Learning Mathematics" (pg.134) [3] brings up the question of what a set is and in particular what is meant by a "well defined" set. He refers to "agreement on rational grounds" to avoid having to decide according to our own judgement and therefore risking the fact that others may not agree.

He gives an example of listing the elements of the set of all good-looking girls in the room. Skemp then writes "A difficulty presents itself: whom do we include ? And supposing that we are able to decide according to our own judgement, will others necessarily agree? We are already in danger of losing, at the outset, one of the essential bases of our mathematics - agreement on rational grounds."

For mathematical purposes we need to avoid difficulties that arise from situations like the one described above. To do this Skemp states that we need to concentrate on **"well-defined"** collections of objects. A "well-defined" collection of objects is such that given any object we can say whether or not it is in the collection. This collection is then called a **set**.

The question still remains: "How do we 'well-define' our collection of objects?" Some adjectival phrase giving a brief description may give us the discerning ability to decide whether a given object is in the collection. If the object fits the description then it is in the collection, and vice-versa. Such a defining phrase or property is called the **characteristic property** of that set.

Skemp writes that another way of defining a set is by listing its elements. This allows for unrelated objects to constitute a set. "Having thus defined a set, we could go on to say that it has a characteristic property - that the names of its elements appear on the list." Skemp validates this by saying that a characteristic property of this sort does not lead to new ideas so therefore may be uninteresting to mathematicians, but it nevertheless is a perfectly good way of defining sets. He also points out that these two ways of defining sets give us two directions. We can first decide on a criterion and then collect objects which satisfy that criterion into our set. Or we can first form a collection of objects and then try to find what (if any) characteristic property may exist.

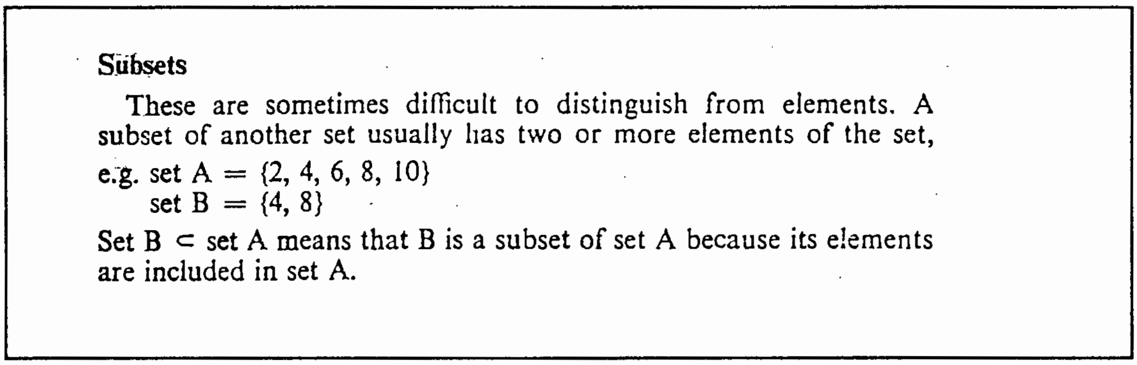
As they stand the sets given in the exercises in MGM 6 and MGM 8 are perfectly "well-defined" and hence the question would thus be better phrased by asking "Which element should be removed so that the remaining set has a characteristic property that can be described in words." Note that there still may be two opinions, as shown earlier in the commentary after each extract.

3. Subsets

The notion of subset is an important one. There are difficulties with regard to showing that the empty set is a subset of every set. Hence the concept of a subset needs to be carefully established.

A look will be taken at definitions of subsets and various errors that indicate that the notion of a subset is not understood. The examples that are given lead to a great deal of confusion regarding subsets.

MASAS,pg.2 says the following about **subsets**:



Subsets

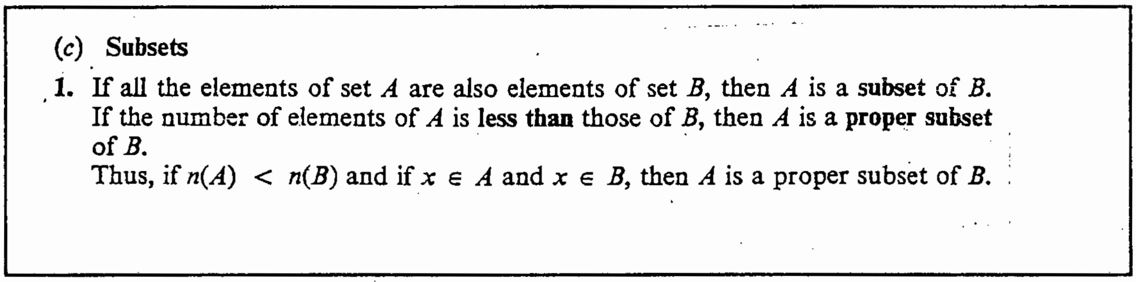
These are sometimes difficult to distinguish from elements. A subset of another set usually has two or more elements of the set,

e.g. set A $= \{2, 4, 6, 8, 10\}$
set B $= \{4, 8\}$

Set B $\subset$ set A means that B is a subset of set A because its elements are included in set A.

This gives no clarity to the definition of subsets. It is particularly misleading when the text says that "a subset of another set usually has two or more elements of the set". A subset of a set need not consist of two or more elements of the set. The statement that subsets are "sometimes difficult to distinguish from elements" probably refers to the distinction between an element "*a*" of a set "A" and the one element subset $\{a\}$ of A. However, these two statements appear to be contradictory. From this text it is difficult to define the empty set and to conceptualise that the empty set is a subset of every set.

MGM 9&10, pg.4, defines a subset in the following way :

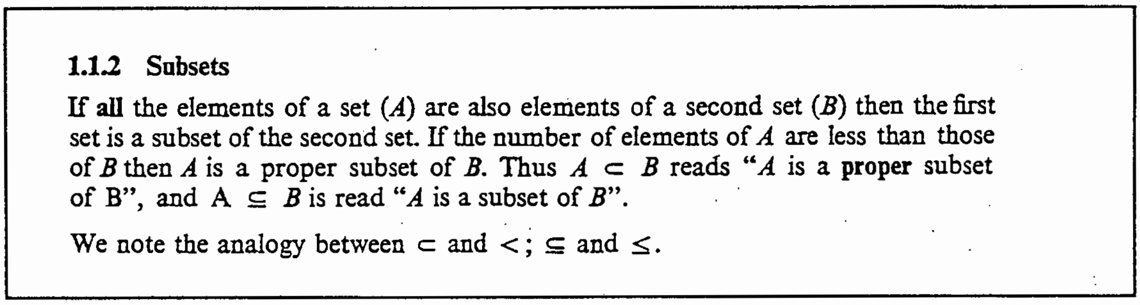


(*c*) **Subsets**

1. If all the elements of set A are also elements of set B, then A is a **subset** of B. If the number of elements of A is **less than** those of B, then A is a **proper subset** of B.
   Thus, if $n(A) < n(B)$ and if $x \in A$ and $x \in B$, then A is a proper subset of B.

This text does not provide a solid idea of what a subset is. The condition "if $x \in \mathbf{A}$ and $x \in \mathbf{B}$" should be "$x \in \mathbf{B}$ when $x \in \mathbf{A}$". This will then lead to the conclusion that **A** is a proper subset of **B**. This extract is defining the concept of subset and proper subset and there is a danger that "if the number of elements of **A** is less than the number of elements of **B**" will be seen as a necessary condition for **A** to be a proper subset of **B**. This is true for finite sets, but not in general. If $\mathbf{A} \subseteq \mathbf{B}$, a sufficient condition that **A** is a proper subset of **B** is $n(\mathbf{A}) < n(\mathbf{B})$. The symbol $n(A)$ is one that several texts use for cardinality. While this text does not go deeply into the notion of cardinality, other than it being the number of elements in a set, several contemporaries of this text book do.
It is possible to have two sets with the same cardinality but with one being a proper subset of the other, e.g. $\mathbf{A} = \{1,2,3,...\}$ and $\mathbf{B} = \{0,1,2,3...\}$. The cardinalities of **A** and **B** are the same, i.e. $n(\mathbf{A}) = n(\mathbf{B})$ but $\mathbf{A} \subset \mathbf{B}$, a proper subset.

A similar example is found in MGM 8, pg.1:

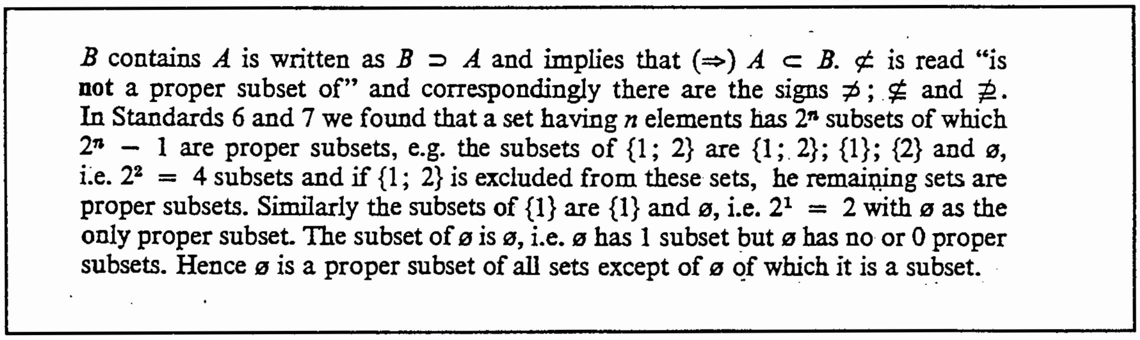


1.1.2 Subsets

If all the elements of a set (A) are also elements of a second set (B) then the first set is a subset of the second set. If the number of elements of A are less than those of B then A is a proper subset of B. Thus $A \subset B$ reads "A is a **proper** subset of B", and $A \subseteq B$ is read "A is a subset of B".

We note the analogy between $\subset$ and $<$; $\subseteq$ and $\leq$.

MGM 8, pg.1, has also has the following confusing paragraph about the number of subsets of a set with **n** elements.



B contains A is written as $B \supset A$ and implies that ($\Rightarrow$) $A \subset B$. $\not\subset$ is read "is **not** a proper subset of" and correspondingly there are the signs $\not\supset$; $\not\subseteq$ and $\not\supseteq$. In Standards 6 and 7 we found that a set having n elements has 2^n subsets of which $2^n - 1$ are proper subsets, e.g. the subsets of $\{1;\ 2\}$ are $\{1;\ 2\}$; $\{1\}$; $\{2\}$ and $\varnothing$, i.e. $2^2 = 4$ subsets and if $\{1;\ 2\}$ is excluded from these sets, he remaining sets are proper subsets. Similarly the subsets of $\{1\}$ are $\{1\}$ and $\varnothing$, i.e. $2^1 = 2$ with $\varnothing$ as the only proper subset. The subset of $\varnothing$ is $\varnothing$, i.e. $\varnothing$ has 1 subset but $\varnothing$ has no or 0 proper subsets. Hence $\varnothing$ is a proper subset of all sets except of $\varnothing$ of which it is a subset.

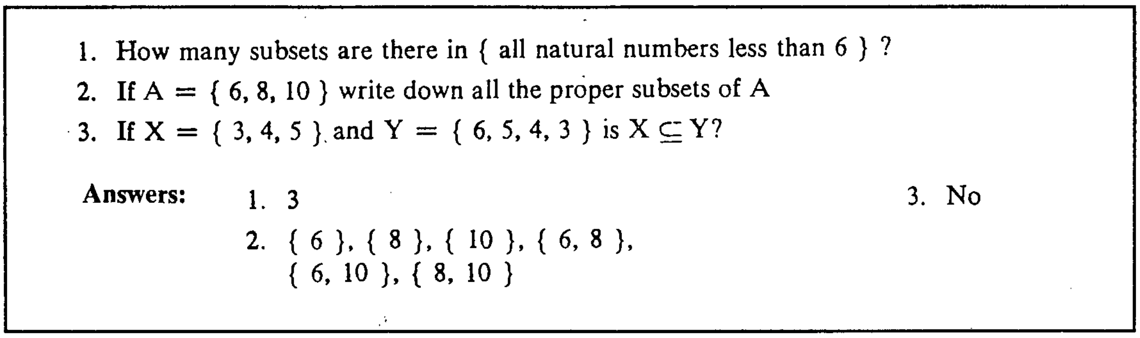
In this extract a conclusion is drawn that the empty set is a subset of every set. From what has previously been written in the paragraph no logical conclusion that the empty set is a subset of every set can be made, and therefore the last sentence, beginning with "hence", is incorrect.

To prove that the empty set is a subset of every set is a difficult task and involves proof by contradiction or by referring to the definition of a subset: "**A** is a subset of **B**" means "for every $x \in \mathbf{A}$, we have that $x \in \mathbf{B}$." This is vacuously true when **A** is the empty set ϕ. Hence $\phi \subset \mathbf{B}$.

However, this argument would probably not be understood at school level, and we have not found it in any of the textbooks that have been examined.

A definition that might alleviate this difficulty is " a subset of a set **A** is a set consisting of all, part or none of the elements of **A**."

EX.M F3, pg. 8, gives an exercise and the corresponding answers, as follows:



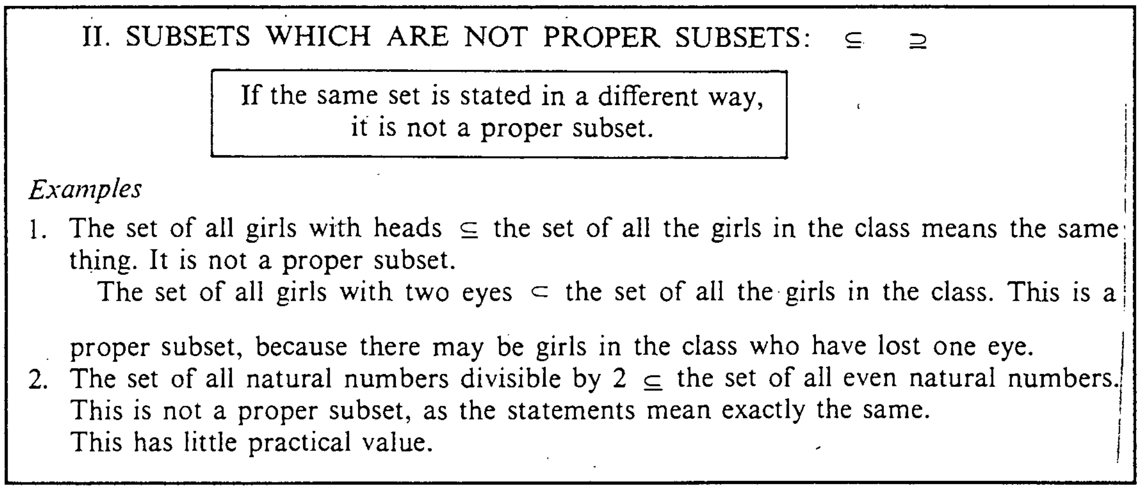
1. How many subsets are there in { all natural numbers less than 6 } ?
2. If A = { 6, 8, 10 } write down all the proper subsets of A
3. If X = { 3, 4, 5 } and Y = { 6, 5, 4, 3 } is $X \subseteq Y$?

Answers: 1. 3 3. No

2. { 6 }, { 8 }, { 10 }, { 6, 8 },
{ 6, 10 }, { 8, 10 }

The first answer could be a misprint for "32". In the second, the authors seem to assume that the empty set is not a proper subset. In the third, they seem to expect that the answer should be no because $X \subset Y$, falling into the same error as students who assert that "$3 \leq 4$ is false because $3 < 4$".

JMAF 1, pg. 7, says the following about sets which are not proper subsets:



II. SUBSETS WHICH ARE NOT PROPER SUBSETS: $\subseteq$ $\supseteq$

If the same set is stated in a different way,
it is not a proper subset.

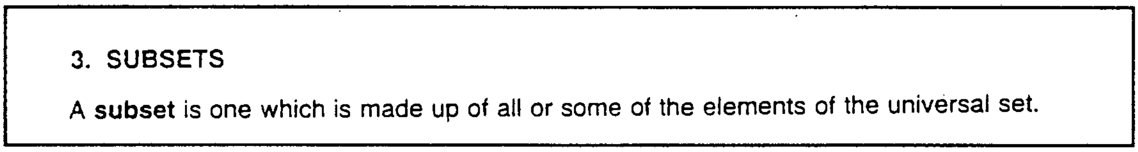
Examples

1. The set of all girls with heads $\subseteq$ the set of all the girls in the class means the same thing. It is not a proper subset.
   The set of all girls with two eyes $\subset$ the set of all the girls in the class. This is a proper subset, because there may be girls in the class who have lost one eye.
2. The set of all natural numbers divisible by 2 $\subseteq$ the set of all even natural numbers. This is not a proper subset, as the statements mean exactly the same.
   This has little practical value.

This explanation of a proper subset is confusing and there is an attempt to substantiate it with examples that are incorrect. The notion of a proper subset has nothing to do with the way in which the set is stated. Like the analogous concept of "less than or equal to" with regards to real numbers, the notion "subset of " establishes a way in which two sets are related. One has to say "**A** is a subset of **B**" and thus there must be a reference to two sets. **A** is not a proper subset of **B** if **A** = **B**.

The "set of all girls with heads" is not the same thing as the "set of all girls **in the class.**" The first line of example one does not make sense, neither set theoretically nor grammatically. If we let **G** denote the set of girls in the class and **E** the set of all girls in the class with one eye, then $\mathbf{E} \subseteq \mathbf{G}$. For certain **G**, $\mathbf{E} \subset \mathbf{G}$ and for other **G**, $\mathbf{E} = \mathbf{G}$. (The example is insensitive and should have been changed at an early stage of the editing process.)

PM 5, pg. 8, says the following about subsets:

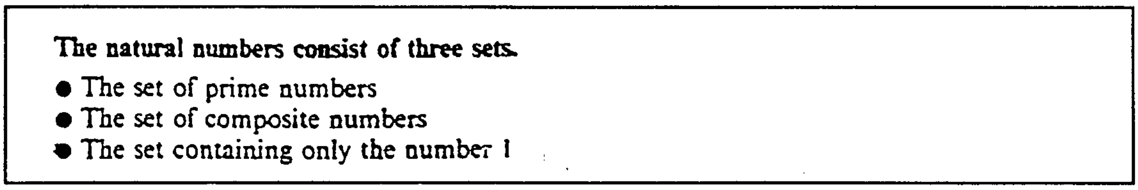


3. SUBSETS

A **subset** is one which is made up of all or some of the elements of the universal set.

This gives no clarification on the notion of subsets. From this definition how would one explain what is meant by "a set **A** is a **subset** of a set **B**"? The reference to the universal set is irrelevant.

MGM 6, pg.4, says

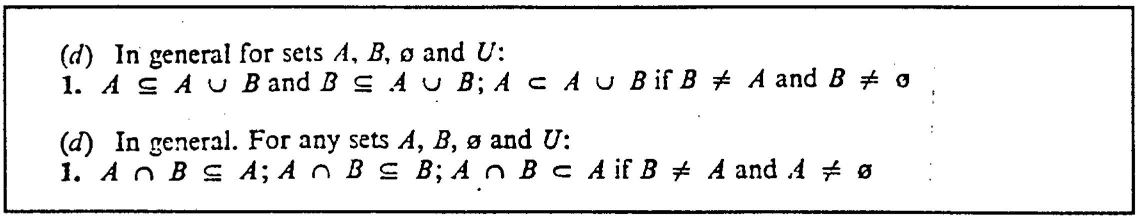


The natural numbers consist of three sets.

- The set of prime numbers
- The set of composite numbers
- The set containing only the number 1

The natural numbers are not limited to consisting of just these three sets. The natural numbers **can** be split into these interesting subsets, but it can also "consist" of many other subsets.

MGM 8, pg. 5, makes the following generalisation:



(*d*) In general for sets A, B, $\varnothing$ and U:
1. $A \subseteq A \cup B$ and $B \subseteq A \cup B$; $A \subset A \cup B$ if $B \neq A$ and $B \neq \varnothing$

(*d*) In general. For any sets A, B, $\varnothing$ and U:
1. $A \cap B \subseteq A$; $A \cap B \subseteq B$; $A \cap B \subset A$ if $B \neq A$ and $A \neq \varnothing$

In both these statements the condition "if $\mathbf{B} \neq \mathbf{A}$ and $\mathbf{B} \neq \phi$ " in the first part and "if $\mathbf{B} \neq \mathbf{A}$ and $\mathbf{A} \neq \phi$ " in the second part are incorrect. For the first part, if **B** is a non-empty proper subset of **A** then it is not true that $\mathbf{A} \subset \mathbf{A} \cup \mathbf{B}$. For the second part, if **A** is a non-empty proper subset of **B** then it is also not true that $\mathbf{A} \cap \mathbf{B} \subset \mathbf{A}$. These will not be strict inclusions, but in fact equalities.

4. Unions and intersections

MGM 8, pg. 5 and 6, has the following definitions of the union and intersection of two sets.

> **1.2.1 Union of two sets**
>
> The union of sets A and B denoted by $A \cup B$ is the set of all elements belonging to A and B, no element being repeated.
>
> **1.2.2 Intersection of two sets**
>
> The intersection of two sets A and B denoted by $A \cap B$ is the set of all elements belonging to A **and** B.

As discussed in chapter 1, this shows a language misunderstanding which makes the concepts of union and intersection very confusing. The intersection of two sets **A** and **B** is the set consisting of the elements that are common to **A** and **B**. The union of two sets **A** and **B** is the set consisting of elements that belong to **A or B**, i.e. the set of all the elements of each set together.

It is important to note that the union and intersection of sets is a binary operation on sets.

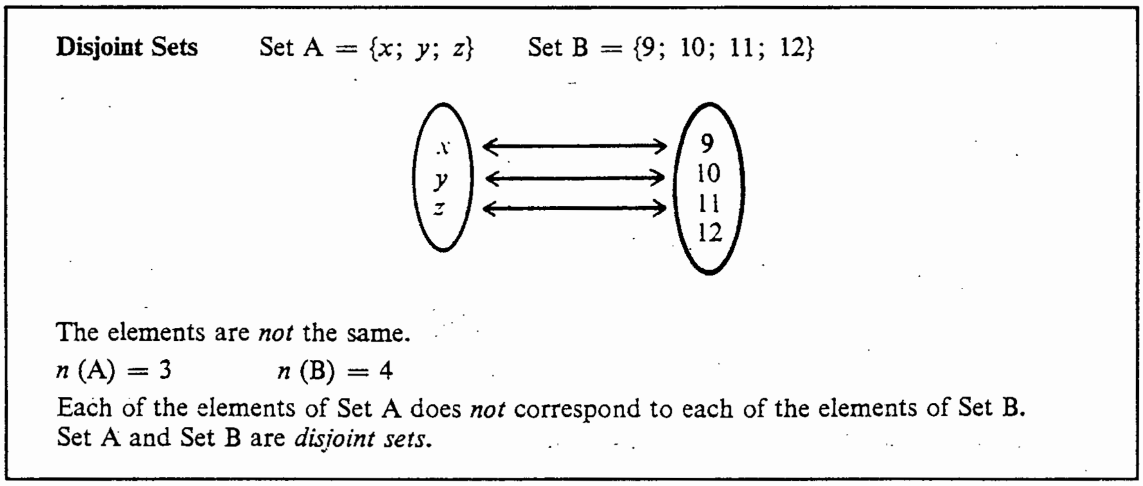
MGM 8, pg. 15, says the following:

> **1.4 The distributive property of sets and property of absorption**
>
> **1.4.1 The distributive property of sets**
>
> Remember, the distributive property states that multiplication of real numbers distributes over addition.
>
> Consider again sets A, B and D of the preceding section if $A = \{1;\ 2;\ 3;\ 4\}$, $B = \{2;\ 3;\ 5;\ 6\}$ and $D = \{3;\ 4;\ 6;\ 7\}$
>
> (a) We have
>
> $$\begin{aligned} A \cup (B \cap D) &= \{1;\ 2;\ 3;\ 4\} \cup \{3;\ 6\} = \{1;\ 2;\ 3;\ 4;\ 6\} \\ \text{also} \quad (A \cup B) \cap (A \cup D) &= \{1;\ 2;\ 3;\ 4;\ 5;\ 6\} \cap \{1;\ 2;\ 3;\ 4;\ 6;\ 7\} \\ &= \{1;\ 2;\ 3;\ 4;\ 6\} \end{aligned}$$
>
> Therefore $A \cup (B \cap D) = (A \cup B) \cap (A \cup D)$
>
> Compare this result with $a \times (b + d) = (a \times b) + (a \times d)$ in Algebra. This shows that the distributive property holds, i.e. union distributes over intersection. This result may be illustrated by Venn diagrams, as follows:

This text implies that the property of absorption and distributivity is a property of sets, when it is a property of the operations of union and intersection. Using the distributivity of multiplication over addition does not show the property holds for union over intersection. This text omits to show that intersection also distributes over union.

Two sets are **disjoint** if they have no common elements, in other words if the intersection of the two sets is empty.

FM 3, pg. 19, says the following about disjoint sets:

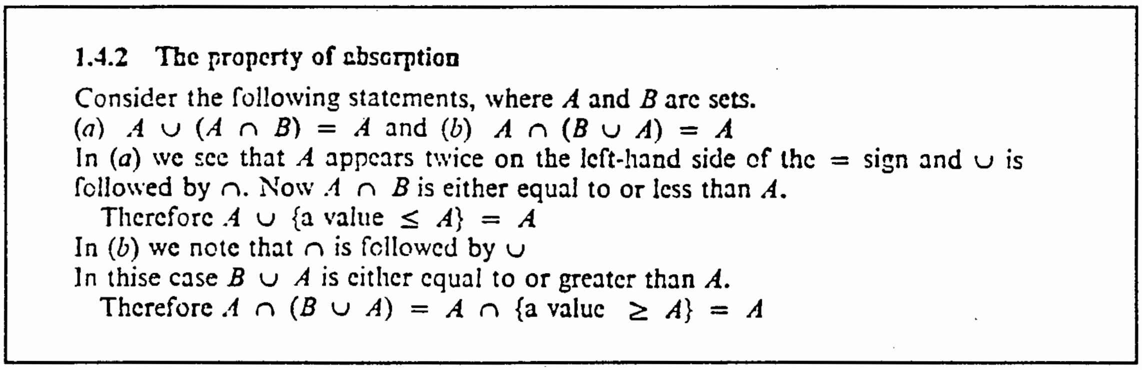


Disjoint Sets Set A = $\{x;\ y;\ z\}$ Set B = $\{9;\ 10;\ 11;\ 12\}$

The elements are *not* the same.
n (A) = 3 n (B) = 4
Each of the elements of Set A does *not* correspond to each of the elements of Set B.
Set A and Set B are *disjoint sets.*

This text is implying that sets that do not have a one-to-one correspondence are disjoint sets, which is clearly incorrect. The sets **A** = {1, 2, 3, 4} and **B** = {1, 2, 3} do not have a one-to-one correspondence and they are not disjoint.

The following extract comes from MGM 8, pg. 16. The analogy between $\leq$ and $\subseteq$ is made but the text gets mixed up and confuses the two concepts.



1.4.2 The property of absorption

Consider the following statements, where A and B are sets.
(a) $A \cup (A \cap B) = A$ and (b) $A \cap (B \cup A) = A$
In (a) we see that A appears twice on the left-hand side of the = sign and $\cup$ is followed by $\cap$. Now $A \cap B$ is either equal to or less than A.
Therefore $A \cup$ {a value $\leq A$} $= A$
In (b) we note that $\cap$ is followed by $\cup$
In thise case $B \cup A$ is either equal to or greater than A.
Therefore $A \cap (B \cup A) = A \cap$ {a value $\geq A$} $= A$

MGM 8, pg. 26, writes the following:

> (2) If $a = b$ then $b = a$ but if $a < b$ then $b \not< a$ and also if $b > a$ then $a \not> b$. The $=$ and $\subseteq$ signs are symmetrical but not the signs $>$ and $<$, $\subset$ and $\supset$.

The relation of "subset of" is not symmetrical. If $\mathbf{A} \subseteq \mathbf{B}$ then it is not necessarily the case that $\mathbf{B} \subseteq \mathbf{A}$.

5. Notation

The notation that is used most frequently is listing or set-builder notation of the form $\{ x \mid x \text{ satisfies some condition} \}$, but there are several instances that show the misuse of set notation. Set builder notation is introduced in PM 5, pg.1, by means of the following example:

> 3. Set-builder notation can be used:
> B = {x/x is a multiple of 2 smaller than 20}

It then goes on to describe sets using the following form of notation:

$$\mathbf{A} = \{ x \mid x \in \mathbf{B} \}$$

PM 5, pg. 2, has an example using the notation in this form:

> A = {x/$x \in$ months of the year that have 30 days}

The same awkward notation is used in this text to give examples of sets that are subsets of another set.

> ***Example of subsets:***
> {x/$x \in$ children} $\subset$ {x/$x \in$ people}
> {x/$x \in$ fractions} $\subset$ {x/$x \in$ numbers}

This notation is clumsy and could lead to the following description:

$$\mathbf{A} = \{\, x \mid x \in \mathbf{A} \,\} = \{\, x \mid x \in \{\, x \mid x \in \mathbf{A} \,\} = \{\, x \mid x \in \{\, x \mid x \in \{....\}..\}$$

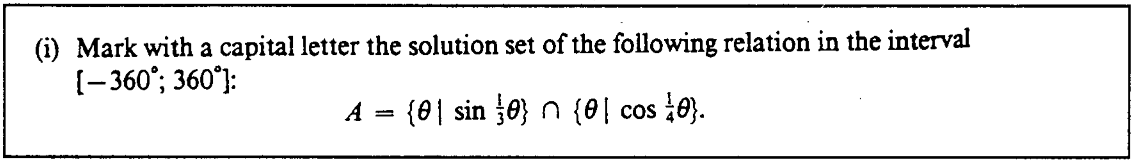
JM 9&10, pg. 321, gives as an answer to a range and domain question the following:

(b) $Df^{-1} = \{x : x \in \mathbf{R}\}$

$Rf^{-1} = \{y : y \in \mathbf{R}\}$

This text is saying that $\mathbf{R} = \{\, x \mid x \in \mathbf{R} \,\}$, which is of the form as described above.

UM 9&10, pg. 198, in an exercise to find the points of intersection of two graphs, says the following:

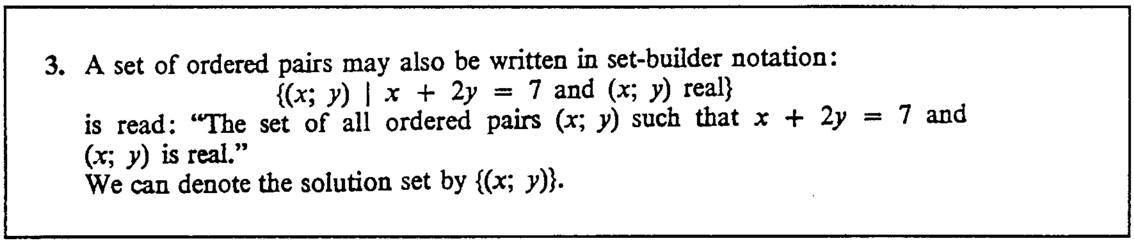


(i) Mark with a capital letter the solution set of the following relation in the interval $[-360^\circ; 360^\circ]$:

$$A = \{\theta \mid \sin \tfrac{1}{3}\theta\} \cap \{\theta \mid \cos \tfrac{1}{4}\theta\}.$$

A relation is already a set, so cannot have a solution set. Here set notation has been incorrectly used to ask the question: "Mark with a capital letter the point(s) of intersection of the graphs $y = \sin\frac{1}{3}\theta$ and $y = \cos\frac{1}{4}\theta$, for $-360^o \le \theta \le 360^o$."

MGMSC, pg.40, writes the following:



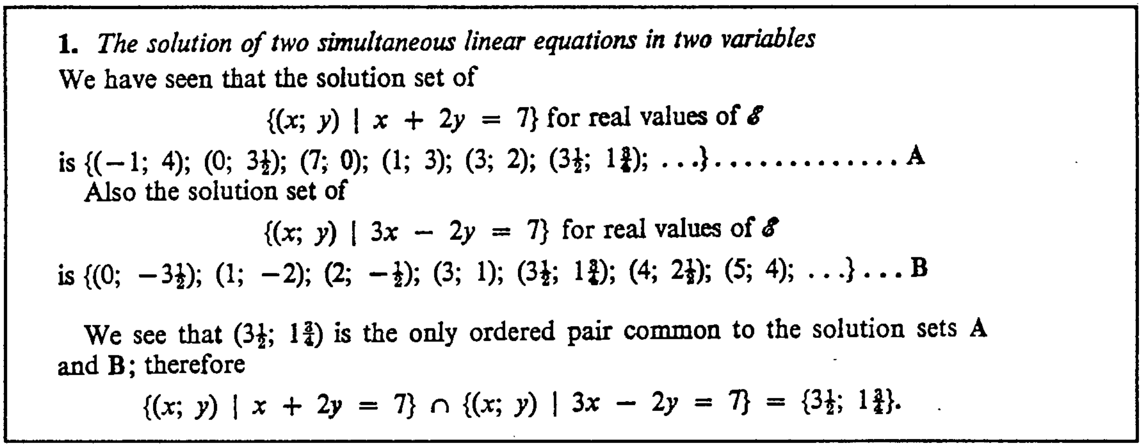
3. A set of ordered pairs may also be written in set-builder notation:

$$\{(x;\, y) \mid x + 2y = 7 \text{ and } (x;\, y) \text{ real}\}$$

is read: "The set of all ordered pairs $(x;\, y)$ such that $x + 2y = 7$ and $(x;\, y)$ is real."
We can denote the solution set by $\{(x;\, y)\}$.

Though x and y are real numbers, $(\, x; y \,)$ is not a real number.

This text continues, on the same page, and refers to the solution set of a set:



1. *The solution of two simultaneous linear equations in two variables*
We have seen that the solution set of

$$\{(x;\ y)\ |\ x + 2y = 7\} \text{ for real values of } \mathscr{E}$$

is $\{(-1;\ 4);\ (0;\ 3\frac{1}{2});\ (7;\ 0);\ (1;\ 3);\ (3;\ 2);\ (3\frac{1}{2};\ 1\frac{3}{4});\ \ldots\}$ .... **A**

Also the solution set of

$$\{(x;\ y)\ |\ 3x - 2y = 7\} \text{ for real values of } \mathscr{E}$$

is $\{(0;\ -3\frac{1}{2});\ (1;\ -2);\ (2;\ -\frac{1}{2});\ (3;\ 1);\ (3\frac{1}{2};\ 1\frac{3}{4});\ (4;\ 2\frac{1}{2});\ (5;\ 4);\ \ldots\}$ ... **B**

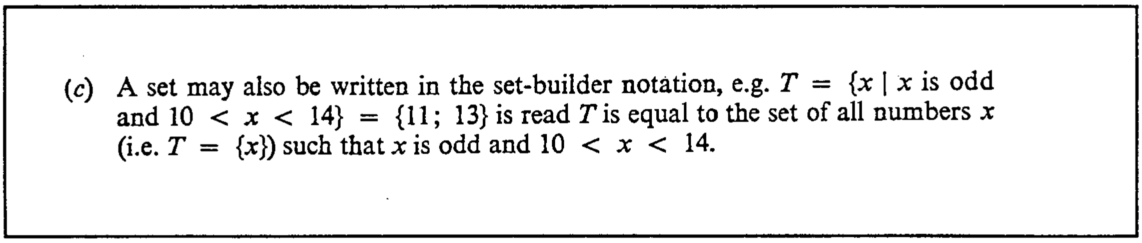
We see that $(3\frac{1}{2};\ 1\frac{3}{4})$ is the only ordered pair common to the solution sets **A** and **B**; therefore

$$\{(x;\ y)\ |\ x + 2y = 7\} \cap \{(x;\ y)\ |\ 3x - 2y = 7\} = \{3\frac{1}{2};\ 1\frac{3}{4}\}.$$

The purpose (and meaning) of the phrase "for real values of $\mathscr{E}$" is unclear. It should rather say "for real values of x and y". The listing of the solution set implies countability of the real numbers, which is false. An equation or inequality has a solution set. A set cannot have a solution set.

The solution of equations, and systems of equations, had always been a part of the school syllabus. With the introduction of set theory, the concept of a "solution set" of an equation was introduced. The solution set of an equation $f(x) = 0$ is the set $\{x : f(x) = 0\}$. Many writers used this terminology incorrectly, requiring pupils to "find the solution set of $\{x : f(x) = 0\}$".

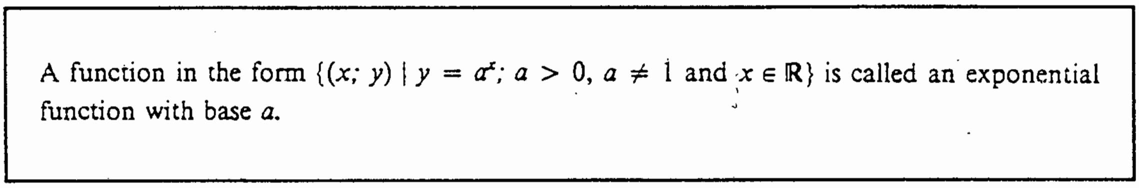
MGM 8, pg. 1, writes the following:



(*c*) A set may also be written in the set-builder notation, e.g. $T = \{x\ |\ x \text{ is odd and } 10 < x < 14\} = \{11;\ 13\}$ is read T is equal to the set of all numbers x (i.e. $T = \{x\}$) such that x is odd and $10 < x < 14$.

The set $\mathbf{T} = \{x\}$ is a single element set, consisting of a letter, and is not equal to the set of all numbers x such that x is odd and $10 < x < 14$. Not including the characteristic property inside the set brackets results in a different set from the set that the authors are trying to describe.

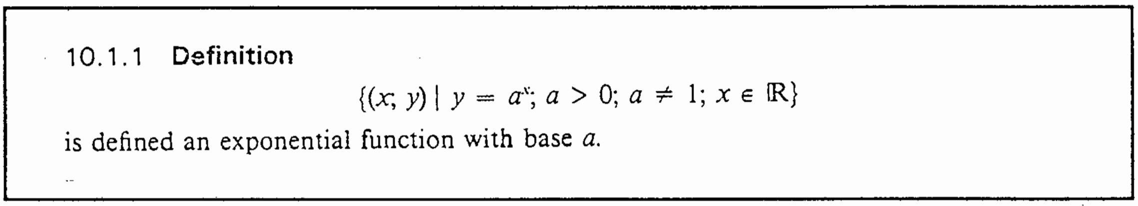
EM 10, pg.2, misuses set notation in the following example:



A function in the form $\{(x; y) \mid y = a^x; a > 0, a \neq 1$ and $x \in \mathbb{R}\}$ is called an exponential function with base a.

Note that the use of set notation is incorrect. The conditions $a > 0, a \neq 1$ must be written outside the set brackets. When these conditions are written inside the brackets the set is not describing a function but in fact the upper half of the real plane, i.e. $\{ (x,y) \in \mathbf{R} \times \mathbf{R} \mid y > 0 \}$.

UM 9&10, pg. 90, make the same error:

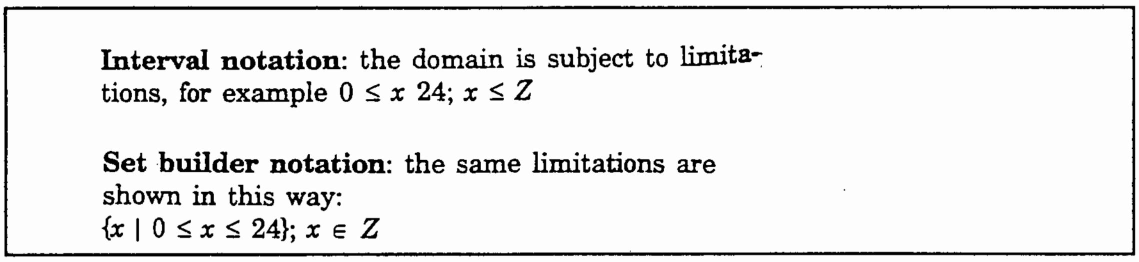


10.1.1 Definition

$$\{(x; y) \mid y = a^x; a > 0; a \neq 1; x \in \mathbb{R}\}$$

is defined an exponential function with base a.

MP 8, pg. 106 and 107, says the following:



Interval notation: the domain is subject to limitations, for example $0 \leq x$ 24; $x \leq Z$

Set builder notation: the same limitations are shown in this way:
$\{x \mid 0 \leq x \leq 24\}; x \in Z$

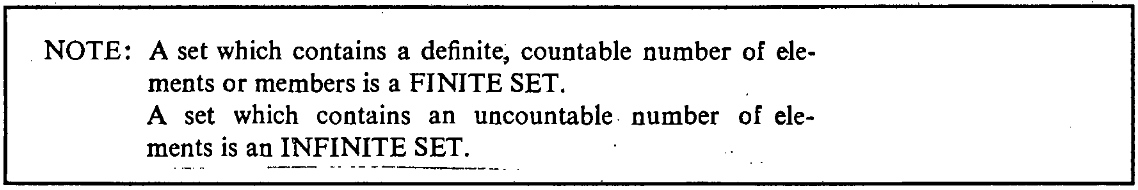
The second line contains two misprints that could cause confusion, and should be replaced with "$0 \leq x \leq 24$; $x \in \mathbf{Z}$". In the last line $x \in \mathbf{Z}$ should be written inside the brackets.

6. Finite and Infinite sets

Text books that cover the concept of finite versus infinite sets usually use the notion that a finite set can be counted and that the counting terminates, while for an infinite set the counting process will not terminate.

While the idea behind this makes sense, it is not the strict definition of an infinite set. An infinite set is characterised by the fact that there exists a one-to-one correspondence between the set and a **proper** subset of itself.

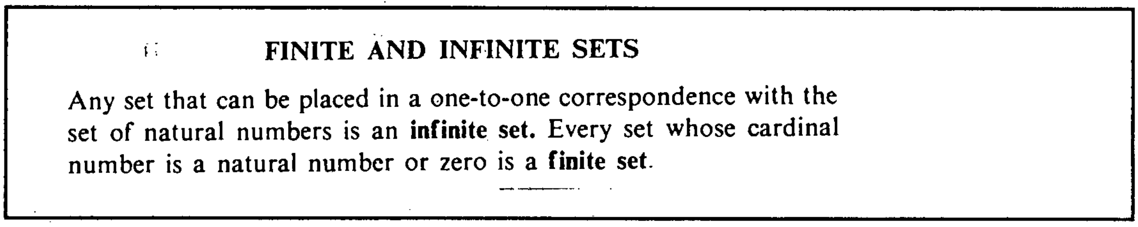
IM Vol. 1, pg. 208 says the following:



NOTE: A set which contains a definite, countable number of elements or members is a FINITE SET.
A set which contains an uncountable number of elements is an INFINITE SET.

While an "uncountable" set in strict mathematical terms is infinite, it is not a necessary condition for a set to be infinite. The term "uncountable" is reserved for infinite sets that cannot be put into a one-to-one correspondence with the natural numbers, whereas "countable" sets are exactly those that can be put into a one-to-one correspondence with the natural numbers.

ES.M 7, pg.15, says the following:

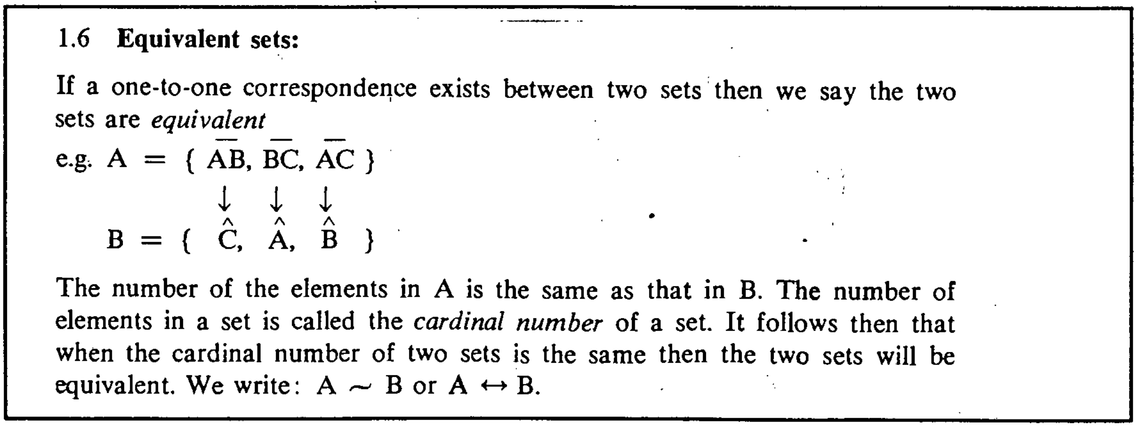


FINITE AND INFINITE SETS

Any set that can be placed in a one-to-one correspondence with the set of natural numbers is an **infinite set.** Every set whose cardinal number is a natural number or zero is a **finite set**.

Again this is a sufficient condition for an infinite set, but not a necessary condition. For example the real numbers cannot be put into a one-to-one correspondence with the natural numbers, but it is an infinite set.

The following extract also deals with the concept of correspondence between sets. The extract is taken from EX.M F3, pg. 5:



1.6 **Equivalent sets:**

If a one-to-one correspondence exists between two sets then we say the two sets are *equivalent*

e.g.
$$\begin{array}{rcl} A & = & \{\ \overline{AB},\ \overline{BC},\ \overline{AC}\ \} \\ & & \quad\ \downarrow\quad\ \downarrow\quad\ \downarrow \\ B & = & \{\ \ \hat{C},\ \ \ \hat{A},\ \ \ \hat{B}\ \ \} \end{array}$$

The number of the elements in A is the same as that in B. The number of elements in a set is called the *cardinal number* of a set. It follows then that when the cardinal number of two sets is the same then the two sets will be equivalent. We write: $A \sim B$ or $A \leftrightarrow B$.

EX.M F3, pg. 5, then says the following:

> *Note*:
>
> i. If two sets are equal then they are also equivalent, however, when two sets are equivalent they are not necessarily equal.
>
> ii. All infinite sets are equivalent, i.e. A = { all natural numbers }
> and B = { all even numbers }
> then A $\sim$ B.

Note that "i.e." should be "e.g.". This text defines two sets to be equivalent if there exists a one-to-one correspondence between the two sets. An easy counter-example to the claim that all infinite sets are equivalent, as made in the extract above, would again be the real numbers, which cannot be put into a one-to-one correspondence with the natural numbers.

These extracts suggest that the authors were unacquainted with the result that the natural numbers are countable and the real numbers uncountable.

MP 8, pg.106, says the following:

> An interval is a finite subset. For example, $-3 < x < 3$ is an interval of the real numbers with the endpoints excluded. It can also be written as (–3; 3).

The concept of a **bounded** set and a **finite** set is mixed up. This is also a frequent misconception amongst high school pupils, i.e. that there are only finitely many numbers between two integers.

3. FUNCTIONS

1. Introduction

During the 1960's and 1970's the topic of relations was a part of the high school syllabus. Textbooks usually discussed relations and functions in one chapter. This was necessary as a foundation to the discussion of functions via the **ordered** pair approach, which was popular at that time.

In the 1980's a syllabus change resulted in the topic of relations being omitted. This necessitated the **rule/machine** approach to functions. Errors in the more recent textbooks have occurred because of overlap of these two approaches. A partial understanding of the concept of relations adds to the reasons for the errors that occur in texts when dealing with functions.

The two approaches to functions are as follows:

1. Rule/machine approach:
"A function f is a rule that assigns to each element x in a set **A** exactly one element, called $f(x)$, in a set **B**." [4]

2. Ordered pair approach:
"A function from a non-empty set **A** to a non-empty set **B** is a non-empty subset f of $\mathbf{A} \times \mathbf{B}$ such that for each $a \in \mathbf{A}$ there exists exactly one $b \in \mathbf{B}$ for which $(a,b) \in f$." [5]

These are two formal definitions of the two approaches and are equivalent .They have three components

1. the **domain A**, i.e. the set of "independent" variables
2. the **codomain B**, in which the values of the function (the "dependent" variables) are located
3. the **rule**. In the first definition it is explicitly stated as such. In the second, the rule is implicit, since the definition of a set requires an underlying rule to decide whether an element belongs to the set or not.

The second definition (ordered pair approach) was used in textbooks of the early "New Maths" era, when "set theory" featured prominently in school text books, and hence relations were regarded as a subset of a Cartesian product. Functions were defined as special types of relations, and hence also as subsets of a Cartesian product. In the 1980's relations were omitted from the syllabus, and the "rule" definition came into use. However, many textbooks still refer to relations without explaining what they are.

Difficulties in South African textbooks arise with both these approaches. Definitions are sometimes incorrect or incomplete and conceptual errors arise as a consequence.

Other errors associated with functions are those related to subsidiary concepts such as range, domain, injectivity, surjectivity and continuity.

2. Definitions

With the ordered pair approach of the New Math era, the understanding of the concept of relations was essential in order to develop the concept of a function.

MGM 9&10, pg. 23, confuse the idea of a relation with that of an operation.

> The relation $\div$ is not symmetrical because $\frac{a}{b} \neq \frac{b}{a}$ unless $a = b \neq 0$.

Perhaps the authors intended to describe the relation between integers given by "***a*** is a divisor of ***b***" (written $a \mid b$), which is not symmetric.

The feature that makes a function a special type of relation is the fact that for each element of the domain there is only one function value in the codomain. This often does not get emphasised very well .

MAT, pg. 98, says

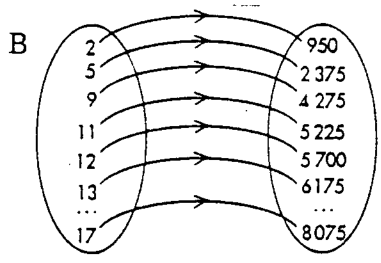
> When a value given to x determines a value of y, then y is called *a function of* x. Sometimes a value given to x may determine more than one value for y. For example,
> if $y = \sin^{-1} x$, and $x = \frac{1}{2}$, then $y = 30°$ or $150°$
>
> $$= \frac{\pi}{6} \text{ or } \frac{5\pi}{6} \text{ in radians.}$$
>
> If $y = \sqrt{x}$, and $x = 4$, then $y = +2$ or -2.

This violates the very essence of the definition of a function: the fact that there exists only one function or output value for each input value.

EM 8, pg. 97, establishes the concept of a function well, but then confuses the issue by the following reference to the "function $x^2 + y^2 = r^2$"

> **6.9 The function $x^2 + y^2 = r^2$**

MP 8, pg. 98, describes the concept of function as follows:



In B we show the table in another form; elements of one set (time) are paired off with elements of the other set (km).

This information can also be given in the form of a set of ordered number pairs.

C {(2; 950), (5; 2 375), (11; 5 225), ..., (13; 6 175), (17; 8 075), ...}

The concept of a **function** is defined in mathematics by this process of pairing off elements of one set with the elements of the other set.

This "process of pairing off elements of one set with the elements of another" suggests that all functions are bijective and does not establish the crux of the matter, which is that a function assigns to each element in a set **A** exactly one element in a set **B**.

Apart from this example stating a general characteristic of a function that is false, it is a complicated example. The listing of the set C is confusing, with an equality sign having been omitted. The pair (9;4275) has been left out of the list, and the second row of dots should precede, not follow, the element (17;8075).

MGM 8, pg. 144, contains a similar error, defining a function as follows:

A function *f* is a one-to-one mapping according to a rule, formula, graph or description of the elements of a set *X* (the domain) into the elements of set *Y* (the range).

3. Independent and dependent variables

When considering functions there has to be reference to the underlying sets. A function is a correspondence between two sets and an attempt at discussing a "function" fails without saying "A function **from a set A to a set B**." The elements of **A** are then defined as **independent** variables and the function values, which are found in **B,** are the **dependent** variables. Both approaches to the concept of functions establish clearly the distinction between independent and dependent variables.

Thus independent and dependent variables are not difficult to establish, as some texts have implied. There have been attempts to illustrate these by means of natural everyday examples. However this "natural" distinction between independent and dependent variables can be misleading. An implication that there is only a one way dependence or that only one variable can be dependent or independent may arise.

MP 8, pg. 100, has a second defining statement concerning functions in terms of independent and dependent variables. However this statement comes before any discussion of independent and dependent variables.

> A **function** is characterized by this property: for each value of the independent variable there is **only one** value of the dependent variable.

MP 8, pg. 101, handles the concept of independent and dependent variables in the following way.

> With functional relations it is sometimes unclear which variable is the dependent one and which is the independent one. For example, you pour a cup of coffee but you forget to drink it and it gradually cools down. The two variables are temperature and time. To determine which is the independent variable, one must decide which of these two arguments is correct.
>
> (a) The time that passes is dependent on the temperature of the coffee.
> (b) The temperature depends on how much time has passed.
>
> Clearly (b) is correct, and the temperature is the dependent variable in this case.

The independent and dependent variables can be changed around if we were to try to determine how long the coffee had been standing by measuring the temperature. In this case temperature is the independent variable and time the dependent variable. In forensic medicine the time of death is established by measuring the temperature of the body. This is another case where temperature is the independent variable and time the dependent variable.

4. Range and domain

Implicit in definitions of a particular function, is the domain. Examples from textbooks have shown that quite often when a function is defined the domain and codomain are left out of the definition. Two different functions arise from one rule if the domains are different. For example:

$$f:\mathbf{Z}\rightarrow\mathbf{Z} \quad \text{defined by } f(x)=2x+1$$

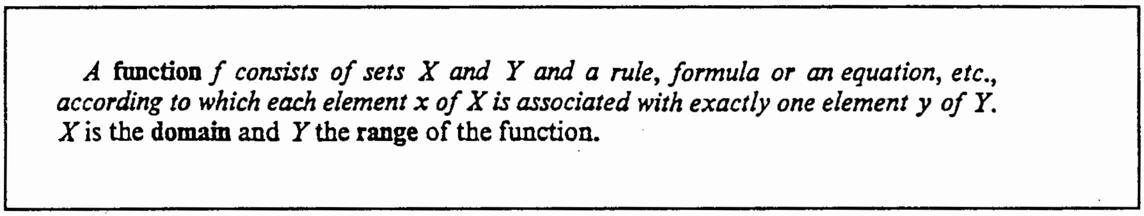
and

$$f:\mathbf{R}\rightarrow\mathbf{R} \quad \text{defined by } f(x)=2x+1$$

These are different functions because their domains are different.

Several texts have incorrect ideas when it comes to the topic of range and domain. The most common of these errors is that range and codomain are the same set. A function from **A** into **B** means that **B** is the codomain, and the range may be a strict subset of it.

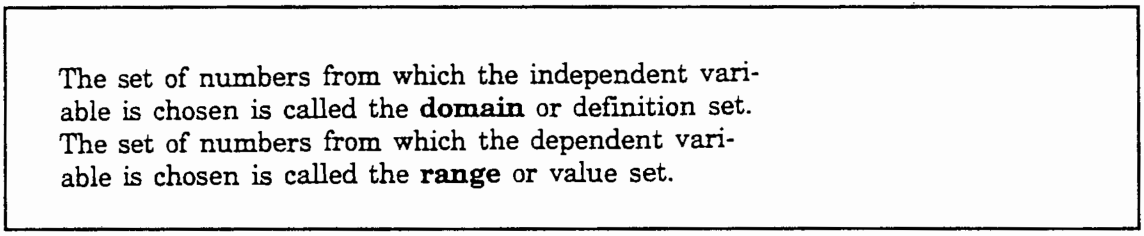
MGM 9&10, pg.28, says:



A **function** *f consists of sets X and Y and a rule, formula or an equation, etc., according to which each element x of X is associated with exactly one element y of Y.* *X* is the **domain** and *Y* the **range** of the function.

Another common misconception is that for a function from a set **A** to a set **B** the domain may be a subset of **A**.

MP 8, pg. 101,discusses the concepts of range and domain in the following way.



The set of numbers from which the independent variable is chosen is called the **domain** or definition set. The set of numbers from which the dependent variable is chosen is called the **range** or value set.

These are confusing descriptions of the concepts of range and domain. The meaning of the "set of numbers from which the independent variable is chosen" is not clear. For example, consider:

$$f:\mathbf{X}\rightarrow\mathbf{R} \qquad \text{where} \qquad \mathbf{X}=\{\,x\in\mathbf{R} \mid 0\le x\le 2\,\}$$

In this example the independent variables are "chosen" from the set of real numbers, but **X** is the domain. Domain is part of the definition of a particular function. Referring to functions of a real variable, Stewart says, "If a function is given by a formula and the domain is not stated explicitly, the convention is that the domain is the set of all numbers for which the formula makes sense and defines a real number." [4]

The use of the word "chosen" when describing the range is not correct. The elements of the range are determined by the function rule and the domain. Once those have been specified the range is automatically specified, with no "choosing" taking place. The range is the set of values the function takes on.

MP 8, pg. 105, also says:

> Suppose for example f: $y = 2x + 1$.
> The domain is {1; 2; 3; 4; 5; 6; 7; 8; 9; 10} and
> the range is {1; 2; 3; 4; 5; 6; 7; 8; 9; 10; ...; 19; 20}.
>
> Substitute all the elements of the domain into the equation one by one. The following function values are obtained.

x	$\xrightarrow{\quad 2x+1 \quad}$	y
1	$\longrightarrow$	3
2	$\longrightarrow$	5
3	$\longrightarrow$	7
4	$\longrightarrow$	9
5	$\longrightarrow$	11
6	$\longrightarrow$	13
7	$\longrightarrow$	15
8	$\longrightarrow$	17
9	$\longrightarrow$	19
10	$\longrightarrow$	21

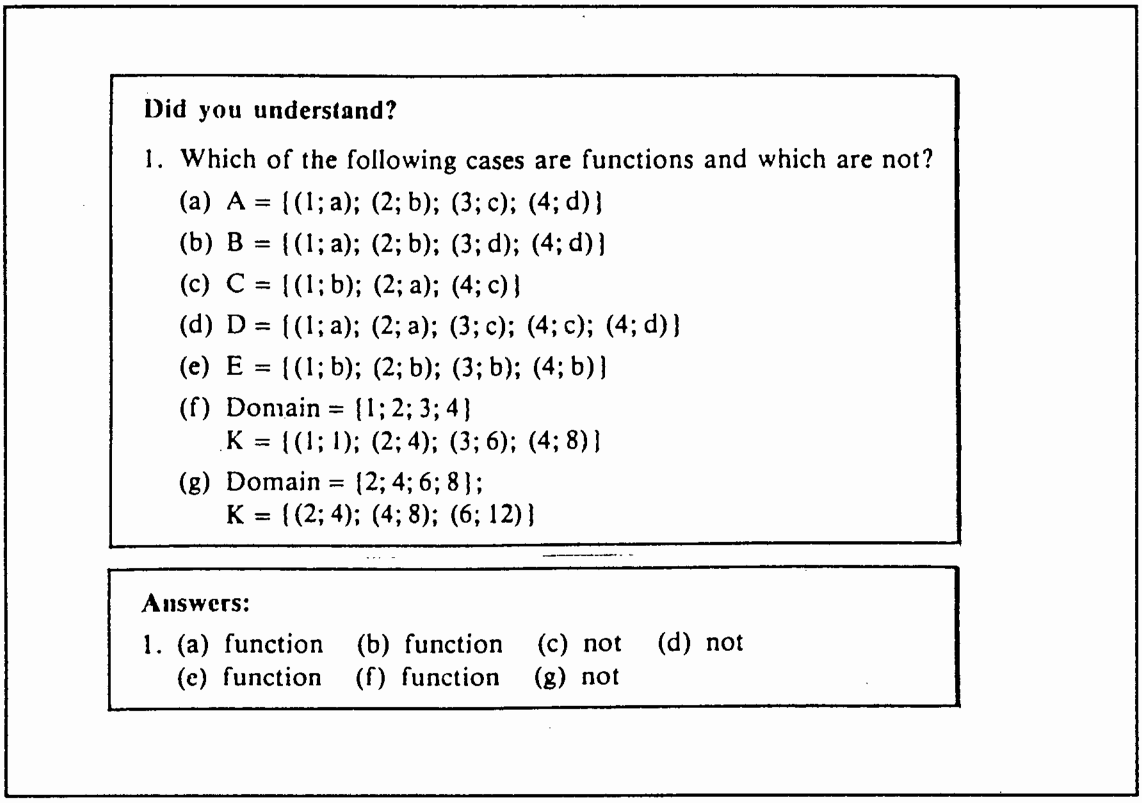
> - The **range** is that set of **possible** values from which the function values can be taken.
> - If there is an element of the domain for which **no** function value can be found in the range, this means that the function is not defined for that value of the domain. (In the example above the function is not defined for $x = 10$.)
> - The **set of function values** and the **range** of a function are not necessarily the same set.
> - In the same way the domain and the set of substituted values is not necessarily the same set

Considering the first and third points, range is being confused with codomain. Range is **exactly** the set of function values.

Point two contradicts the definition of domain. If a function value does not make sense for a particular value then that value is not in the domain. This misconception could arise from the previous statement in MP 8, pg. 101, which implied that there was a choice involved in the elements of the range. This is a lack of understanding that the function rule and domain are the determining factors of range. The example, given above the points, states an incorrect range. Once again this is a consequence of thinking that the values in the range can be "chosen". This example can be corrected by saying that the range is {3;5;7;11; ... ; 21} or that the codomain is {1;2;3; ..... ;21}.

The last point of this paragraph in MP 8 is also incorrect. The phrase "set of substituted values" needs to be reconsidered. The set of all values that are substituted into the function rule, without causing the rule to be undefined, is just the domain.

EX.M 8, pg. 143, has the following exercise:



Did you understand?

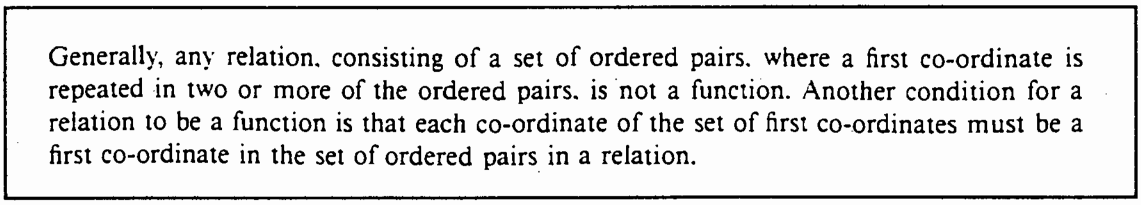
1. Which of the following cases are functions and which are not?
   (a) A = {(1; a); (2; b); (3; c); (4; d)}
   (b) B = {(1; a); (2; b); (3; d); (4; d)}
   (c) C = {(1; b); (2; a); (4; c)}
   (d) D = {(1; a); (2; a); (3; c); (4; c); (4; d)}
   (e) E = {(1; b); (2; b); (3; b); (4; b)}
   (f) Domain = {1; 2; 3; 4}
   K = {(1; 1); (2; 4); (3; 6); (4; 8)}
   (g) Domain = {2; 4; 6; 8};
   K = {(2; 4); (4; 8); (6; 12)}

Answers:

1. (a) function (b) function (c) not (d) not
   (e) function (f) function (g) not

Their claim that (c) is not a function is false, since no domain has been specified.

UM 9&10, pg. 25, says:



Generally, any relation, consisting of a set of ordered pairs, where a first co-ordinate is repeated in two or more of the ordered pairs, is not a function. Another condition for a relation to be a function is that each co-ordinate of the set of first co-ordinates must be a first co-ordinate in the set of ordered pairs in a relation.

The set of first co-ordinates, as the statement implies, consists of those co-ordinates that appear first in an ordered pair. The second sentence is confusing because it states the obvious (each first co-ordinate must be a first co-ordinate.)

UM 9&10, pg. 25, has an example to try to illustrate their previous statement:

Example

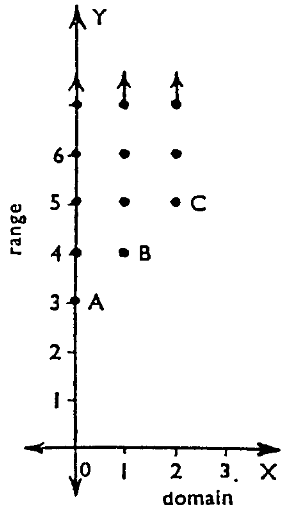
The set of first co-ordinates $X = \{1; 2; 3; 4\}$. The set of second co-ordinates $Y = \{3; 4; 5; 6\}$. The relation $\{(1; 3); (3; 3); (4; 3)\}$ is not a function because the first co-ordinate 2 does not occur in the relation.

In this example the relation is a function but its domain is not **X**.

In the following example from MGM 8, pg. 133, the first (b) is a question and the second (b) is the corresponding answer. The range of the relation is given in the answer, but the accompanying graph that is given (shown as a diagram and as a set), shows an incorrect range.

(*b*) Represent the open sentence $y > x + 3$ graphically, where the domain is $\{x \mid x \in N_0,\ 0 \leq x < 3\}$ and the range $y \in$ real numbers.

(*b*) The domain is restricted to $\{0;\ 1;\ 2\}$
The required graph of the solution set is $\{$points on Y-axis above $A\}$ $\cup$ $\{$points vertically above B$\}$ $\cup$ $\{$points vertically above $C\}$
Therefore, for the domain $\{0;\ 1;\ 2\}$ the required range is $\{y \mid y$ is real and $y > 3\}$, i.e. the required graph is the set of isolated points $(0;\ 4)$; $(0;\ 5)$; $(0;\ 6)$; ... and
$(1;\ 5)$; $(1;\ 6)$; $(1;\ 7)$; ... and
$(2;\ 6)$; $(2;\ 7)$; $(2;\ 8)$; ...

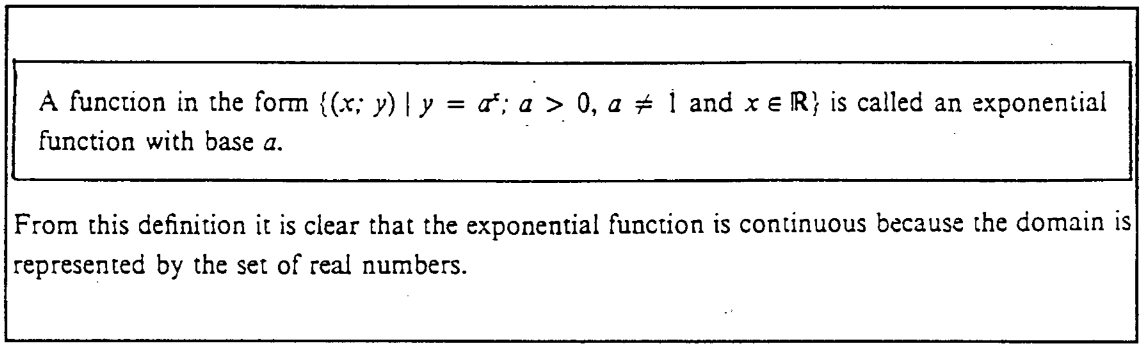


It is incorrect to refer to the "graph of the solution set", but rather "the graph of the **relation**". The graph of a relation or function is often represented in set form which is called "the solution set". The required graph, in fact, is not a set of isolated points as said in the listing of the points, and the diagram shows isolated points as well as the points (0;3), (1;4) and (2;5). The diagram should be the vertical, continuous lines and not including the points at **A**, **B** and **C**.

5. Continuity

The concept of continuity of a function is not an easy one to explain in rigorous terms. Many texts do not comment on continuity, while several others try to discuss the concept but attempt to avoid rigour and in doing so make errors concerning continuity.

EM 10, pg. 2, defines the exponential function as follows:



A function in the form $\{(x; y) \mid y = a^x; a > 0, a \neq 1 \text{ and } x \in \mathbb{R}\}$ is called an exponential function with base a.

From this definition it is clear that the exponential function is continuous because the domain is represented by the set of real numbers.

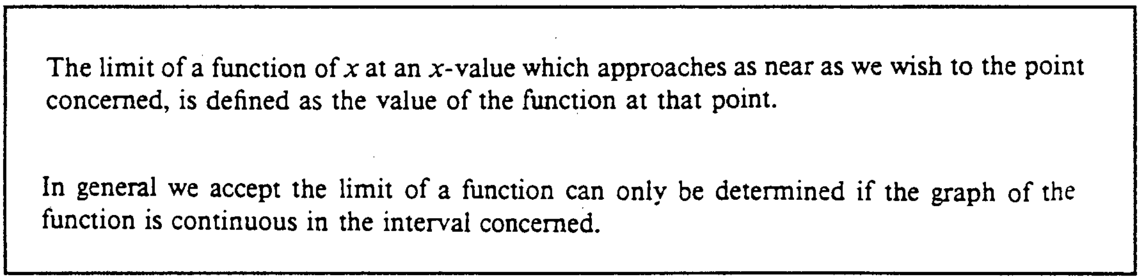
Note that the use of set notation is incorrect. The conditions $a > 0, a \neq 1$ must be written outside the set brackets. When these conditions are written inside the brackets the set is not describing a function but in fact the upper half of the real plane, i.e. $\{\ (x, y) \in \mathbf{R} \times \mathbf{R} \mid y > 0\ \}$ (as mentioned in chapter 2).

The main error that needs to be noted in this extract is that continuity is not a consequence of the fact that the domain of the exponential function is the set of real numbers. A counter-example is the Dirichlet function:

$$f: \mathbf{R} \to \mathbf{R} \text{ defined by } f(x) = \begin{cases} 1 & x \in \mathbf{Q} \\ -1 & x \notin \mathbf{Q} \end{cases}$$

This is defined on the whole of **R** but it is not continuous.

UM 9&10, pg.122, has two statements which are related to continuity.



The limit of a function of x at an x-value which approaches as near as we wish to the point concerned, is defined as the value of the function at that point.

In general we accept the limit of a function can only be determined if the graph of the function is continuous in the interval concerned.

Clearly these statements do not clarify the concepts of limits and continuity. It may be possible to determine the limit of function that is not continuous. A simple example shows that the statements in this extract do not hold. Consider the function defined as follows:

$$f(x) = \begin{cases} 2 & x = 1 \\ x & \forall x \in \mathbf{R}, x \neq 1 \end{cases}$$

Then $\lim_{x \to 1} f(x) = 1$ but $f(1) = 2$

6. Inverses

The inverse of a function is often described as the "reverse procedure" of a function, taking function or output values back to the input values or elements of the domain.

Usually texts describe the inverse of a function by the procedure which just interchanges the independent and dependent variables, with the new domain being the former range. This way of describing an inverse relation is correct, but when functions are not discussed in terms of relations confusion can result from this approach

It has been found that texts do not clearly state that an inverse of a function need not be a function, unless specifically defined (as in the case of the inverse **trigonometric functions**).

NMM 9, pg. 52, says:

> (2) The notation for the inverse of R is R^{-1}, of G is G^{-1}, etc. (If the mapping is unique, the inverse is also a function, e.g. the inverse of f is f^{-1}.)

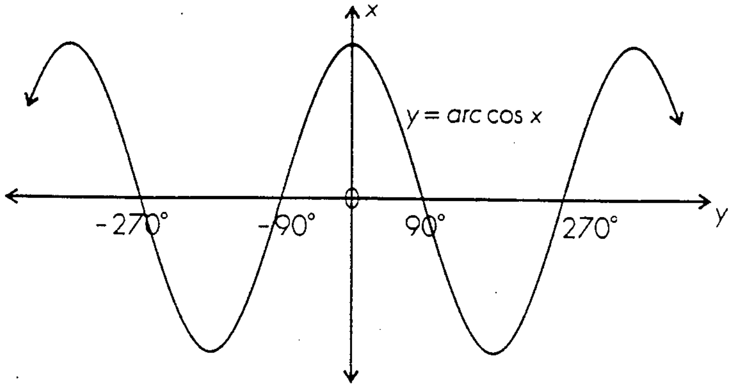
The condition for the inverse of a function to be a function is unclear in this extract. The concept of injectivity needs to be introduced. Uniqueness of a mapping as it is used in this extract does not make sense. It should be clearly stated that if the mapping is a one-to-one function then the inverse is a function.

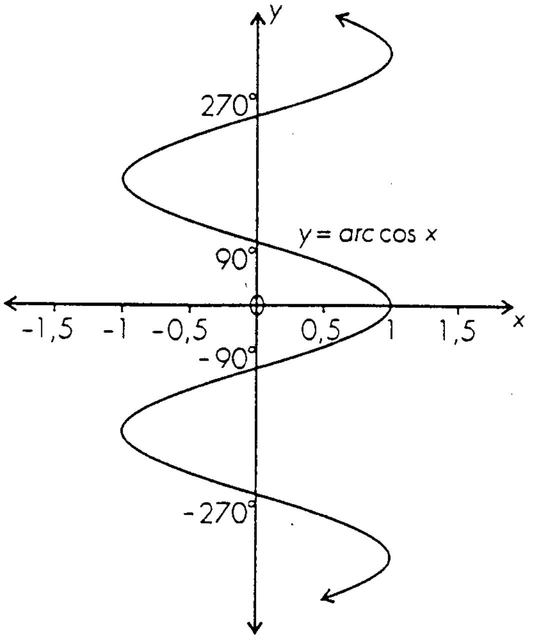
MP 9, pp. 275 and 276, refers to the inverse trigonometric functions and has graphical representation as follows:

- arc sin, arc cos and arc tan are not functions, although the calculator provides only one value for θ.

We know that any input value x produces only one output value y. So $y = \cos x$ is a function, where x represents the degrees and y represents a trigonometric ratio.

Note, however, that the same output value (for example 0,819) can be produced by many different input values. Because of this, the inverse of $y = \cos x$ is not a function. In standard form we write the inverse $x = \cos y$ as $y = \text{arc} \cos x$, where x (the input value) represents a trigonometric ratio and y (the output value) represents the degrees. Note that the x and y axes in the second sketch opposite have been swapped around, but that they have been rotated back to their standard positions.





The inverse function of the cosine function has been incorrectly defined in this text.

The standard definition of the inverse of the cosine function is given by Stewart in the book "Calculus", as follows:

$\cos^{-1} x = y \quad \Leftrightarrow \quad \cos y = x \qquad \text{and} \qquad 0 \leq y \leq \pi$

Thus part of the definition of the inverse cosine **function** is the restriction $0 \leq y \leq \pi$.

4. DIVISIBILITY AND THE REMAINDER THEOREM

1. Introduction

This chapter examines the treatment of polynomials, the division algorithm and the remainder theorem.

A number of cases have been found in text books where polynomials are incorrectly defined, the need for the division algorithm is invariably ignored and the remainder theorem is incorrectly stated and proved. The concept of divisibility in the ring of integers is confused with divisibility in the ring of polynomials.

2. Definition of a polynomial

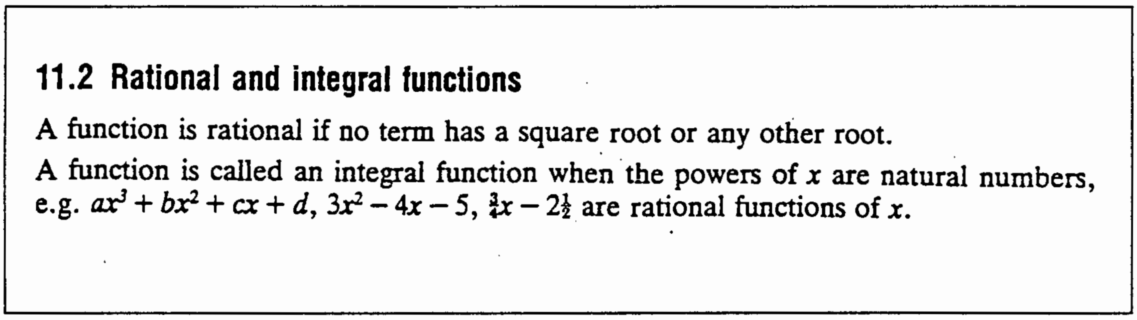
The standard definition of a polynomial is :

"A function *f* is a polynomial if $f(x) = a_n x^n + a_{n-1} x^{n-1} + \; .... \; + a_1 x + a_0$ and $a_o, a_1, \; .... \; a_n$ are real numbers and $a_n \neq 0$ and the exponents are non-negative integers."[4]

In the school syllabus, polynomials are restricted to those having rational coefficients. Several text books refer to "rational" and "integral" functions, and multinomials in their discussion of polynomials, using these terms in non-standard ways. For example, a **rational** function is usually understood to be a quotient of polynomials, an **integral** function is a function of a complex variable holomorphic over the complex numbers, and the term "multinomial" does not appear in the Concise Oxford Dictionary of Mathematics.[6]

These terms may be found in text books from the late nineteenth century but have long since fallen into disuse. For example, the text "Higher Algebra", by Hall and Knight (chapters IX and XXXV) [7] uses the term "rational integral function" for "polynomial", and "dimension" for "degree".

NMM 9, pg. 97, says:



11.2 Rational and integral functions

A function is rational if no term has a square root or any other root.

A function is called an integral function when the powers of x are natural numbers, e.g. $ax^3 + bx^2 + cx + d$, $3x^2 - 4x - 5$, $\frac{3}{4}x - 2\frac{1}{2}$ are rational functions of x.

UM 9&10, pg. 72, gives the following definition of a polynomial:

> **7.2 RATIONAL INTEGRAL EXPRESSION OF x:**
>
> **Definition**
>
> This is an expression where no term contains a square or other root and the powers of x are positive integers.

This is a description rather than a definition, and could be interpreted as excluding multiples of x^0 (i.e. constant terms) from polynomials. The authors of MGMSC, on pg.116, imply that a multinomial is a sum of powers of the variable with no restriction that the exponents are integral. For example:

> $2x^{2\frac{1}{2}} - 3x^{1\frac{1}{2}} - 5x^{1\frac{1}{2}} + 3$ is a **multinomial** and **not** a polynomial.

They then refer to polynomials in the following way:

> The expression $2x^3 - 3x^2 - 4x + 5$ is a **rational integral multinomial.** It is **rational** because the coefficient of the terms, viz. 2, −3, −4 and 5, are rational numbers and do not involve irrationals such as $\sqrt{3}$. It is **integral** because the exponents 3, 2 and 1 are positive integers.
>
> A rational integral multinomial is a **polynomial,** although a multinomial need not necessarily be a polynomial. The **exponents** of a polynomial are **natural numbers.**

3. The division algorithms for integers and polynomials

The division algorithm is an important result about polynomial rings. It bears resemblance to the division algorithm for integers. The remainder theorem is built upon the division algorithm and is strictly speaking a corollary. The division algorithm is not acknowledged in texts.

Since the two division algorithms are fundamental to the theory of polynomials (respectively number theory), they require careful proofs. The proofs are similar in structure and rely on arguments involving well-ordering or infinite descent (both of which are equivalent to the principle of mathematical induction).

It is appropriate to first consider the division algorithm for integers and polynomials.

Theorem (Division Algorithm for Integers)

If N and M are positive integers, then there exist unique positive integers q and r such that

$$N = qM + r \quad \text{and} \quad 0 \leq r < M$$

The integer q is called the quotient and r is the remainder.

Proof: (see [5], pp. 26 and 27)

Theorem (Division Algorithm for Polynomials)

The division algorithm for polynomials says that given two polynomials f and g, and g is not the zero polynomial, then there are polynomials m and r such that:

$$f = mg + r$$

where either i) r is the zero polynomial or
ii) r is a non-zero polynomial and $\deg r < \deg g$.

Proof: (see [5], pg. 49)

MGMSC, pg. 117, states the division algorithm, without referring to it as such, but states it for multinomials. According to their definition of a multinomial (see above), this statement is then not true because the division algorithm does not hold for non-polynomials.

> **It is always possible to write a multinomial $f(x)$ in the form**
> $$f(x) \equiv (x - q) \times p(x) + R$$
> **where R is a constant. All that is necessary is to divide $f(x)$ by $x - q$ until the remainder is independent of x.**

4. Statement and Proof of the Remainder Theorem

It has been found that discussion and proofs of the remainder theorem in school texts ignore the division algorithm. The claims and proofs found in texts actually rely heavily

on the algorithm in order to hold. Texts do not realise that the remainder theorem can be regarded as a corollary of the division algorithm.
Although most of the proofs attempted and most commonly used need the division algorithm there is an escape route, which one text [8] has used. The following is a version based on their proof:

A proof of the remainder theorem without using the division algorithm:

From the theory of geometric sequences we have:

$$1 + r + r^2 + \;\ldots.\; r^{n-1} = \frac{1-r^n}{1-r} \quad (r \neq 1)$$

$$\Leftrightarrow \quad 1 - r^n = (1-r)(1 + r + r^2 + \;\ldots.\; + r^{n-1})$$

Let $r = \dfrac{h}{x}$

Then $1 - \dfrac{h^n}{x^n} = (1 - \dfrac{h}{x})(1 + \dfrac{h}{x} + \dfrac{h^2}{x^2} + \;\ldots.\; + \dfrac{h^{n-1}}{x^{n-1}})$

$\Rightarrow \quad x^n - h^n = (x-h)(x^{n-1} + x^{n-2}h + \;\ldots.\; + h^{n-1})$

$\Rightarrow \quad x - h$ is a factor of $x^n - h^n$

Let $f(x) = a_n x^n + a_{n-1}x^{n-1} + \;\ldots.\; + a_1 x + a_0$

Then $f(h) = a_n h^n + a_{n-1}h^{n-1} + \;\ldots.\; + a_1 h + a_0$

$$\Rightarrow \quad f(x) - f(h) = a_n(x^n - h^n) + \;\ldots.\; + a_1(x-h) \quad (1)$$

Now $x - h$ is a factor of $x^n - h^n, x^{n-1} - h^{n-1}, \;\ldots.\;, x - h$.
Hence $x - h$ is a factor of the RHS of (1).

$\Rightarrow \quad f(x) - f(h) = (x-h) \times Q(x)$ where $Q(x)$ is a polynomial
$\Rightarrow \quad f(x) = (x-h) \times Q(x) + f(h)$

The term $(x-h) \times Q(x)$ is divisible by $x - h$ and $f(h)$ contains no term in x.
There fore when $f(x)$ is divided by $x - h$, a constant remainder of $f(h)$ is left.

Even though this proof does not use the division algorithm, most of the texts that have been studied use proofs that do require the division algorithm. The following are examples of such

MMC, pg. 24, states the remainder theorem as follows:

> *REMAINDER THEOREM:*
>
> If any polynomial $f(x)$ is divided by $ax + b$ until there is a remainder free of x, then that remainder is
>
> $$f\left(-\frac{b}{a}\right)$$
>
> So we may write:
>
> $$f(x) = (ax+b) \cdot Q(x) + R$$

The fact that $f(x) = (ax + b) \cdot Q(x) + R$ is not a consequence of the remainder being $f\left(-\frac{b}{a}\right)$. This would imply that the division algorithm is a consequence of the remainder theorem.

Without the division algorithm it is not clear why a polynomial can be divided by $ax + b$ until the remainder contains no x, or why the polynomial $f(x)$ can be written as follows: $f(x) = (ax + b)Q(x) + R$, where R is independent of x.

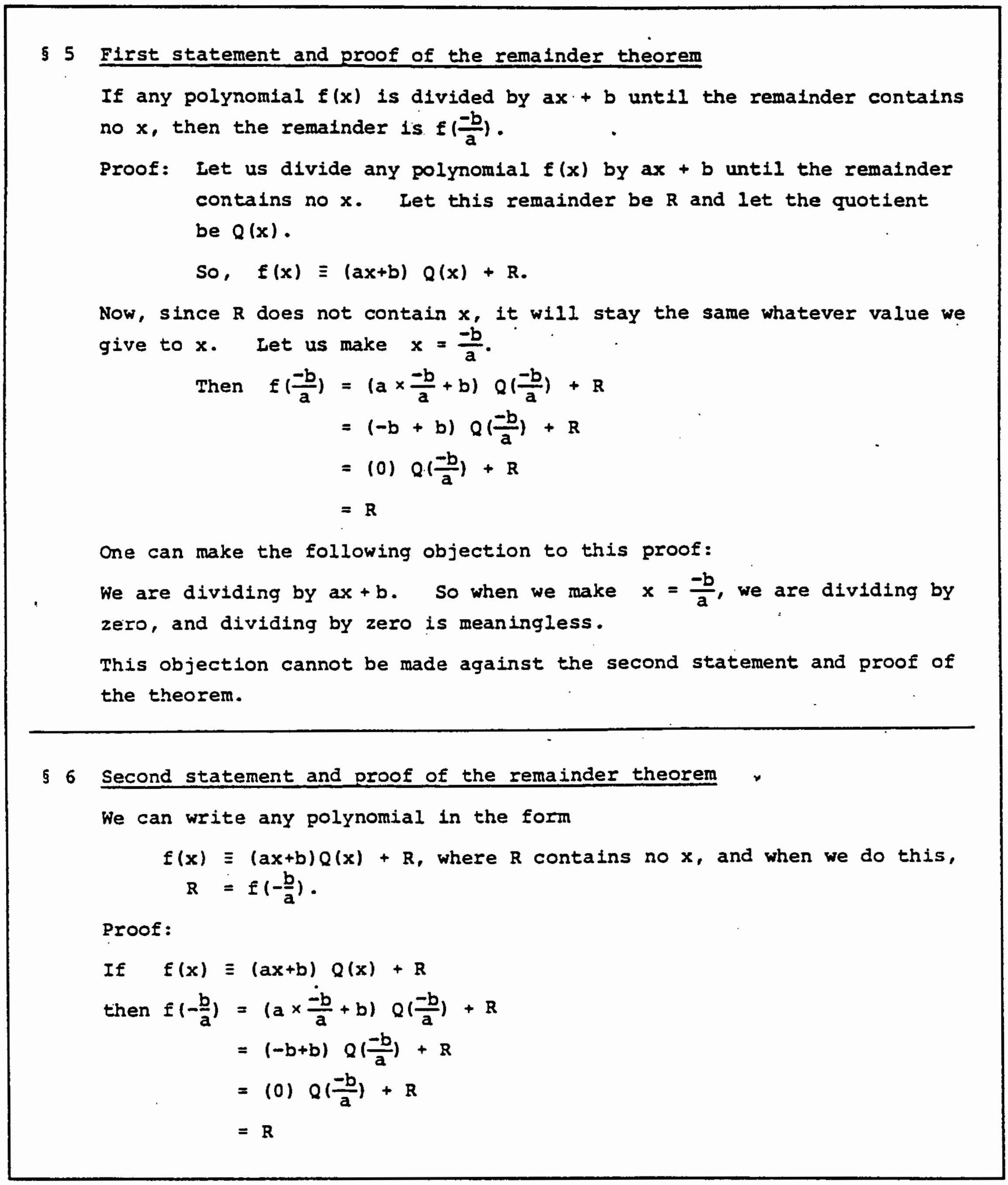
The terms **rational integral** function/expression/multinomial have been used by several texts when referring to polynomials. This outdated nomenclature has been discussed in section 1 of this chapter, but it is necessary to note that the division algorithm and remainder theorem do not hold only for polynomials with rational coefficients.

NMM 9, pg. 98, says:

> ### 11.4 The remainder theorem
>
> **When any rational, integral $f(x)$ is divided by $x - a$ until the remainder is independent of x, it is $f(a)$**

In the above statement “it” in the phrase “it is $f(a)$” is imprecise, since it could be interpreted as referring to $f(x)$ and not the remainder as required.

JM 9&10, pg.33, proves the remainder theorem as follows:



§ 5 First statement and proof of the remainder theorem

If any polynomial $f(x)$ is divided by $ax + b$ until the remainder contains no x, then the remainder is $f(\frac{-b}{a})$.

Proof: Let us divide any polynomial $f(x)$ by $ax + b$ until the remainder contains no x. Let this remainder be R and let the quotient be $Q(x)$.

So, $f(x) \equiv (ax+b)\ Q(x) + R$.

Now, since R does not contain x, it will stay the same whatever value we give to x. Let us make $x = \frac{-b}{a}$.

$$
\begin{aligned}
\text{Then}\quad f(\tfrac{-b}{a}) &= (a \times \tfrac{-b}{a} + b)\ Q(\tfrac{-b}{a}) + R \\
&= (-b + b)\ Q(\tfrac{-b}{a}) + R \\
&= (0)\ Q(\tfrac{-b}{a}) + R \\
&= R
\end{aligned}
$$

One can make the following objection to this proof:

We are dividing by $ax + b$. So when we make $x = \frac{-b}{a}$, we are dividing by zero, and dividing by zero is meaningless.

This objection cannot be made against the second statement and proof of the theorem.

§ 6 Second statement and proof of the remainder theorem

We can write any polynomial in the form

$f(x) \equiv (ax+b)Q(x) + R$, where R contains no x, and when we do this,
$R = f(-\frac{b}{a})$.

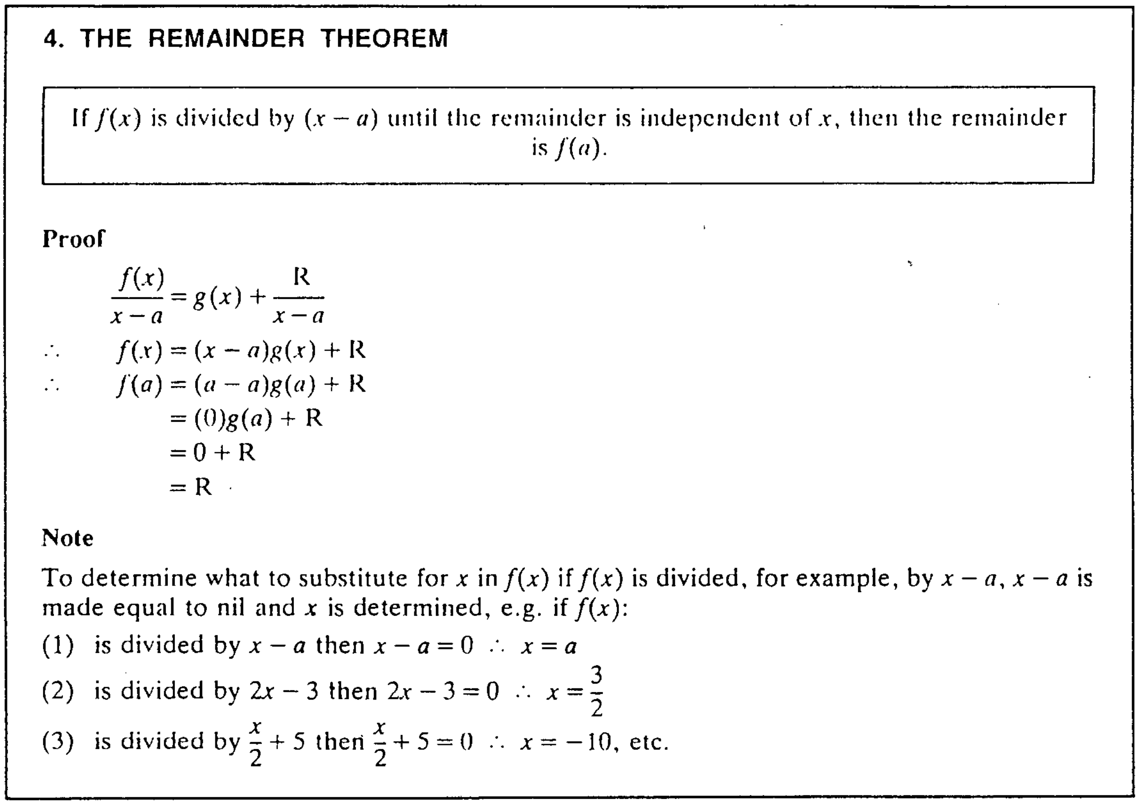
Proof:

$$
\begin{aligned}
\text{If}\quad f(x) &\equiv (ax+b)\ Q(x) + R \\
\text{then}\quad f(-\tfrac{b}{a}) &= (a \times \tfrac{-b}{a} + b)\ Q(\tfrac{-b}{a}) + R \\
&= (-b+b)\ Q(\tfrac{-b}{a}) + R \\
&= (0)\ Q(\tfrac{-b}{a}) + R \\
&= R
\end{aligned}
$$

The fact that $f(x) = (ax + b)Q(x) + R$, where R contains no x, is merely an application of the division algorithm, with $g(x) = (ax + b)$ and R is a function of degree less than $g(x)$ and therefore has degree zero, that is it contains no x.

In the first statement and proof of the remainder theorem, as given by JM 9&10, an objection is made about dividing by $ax+b$ when $x=\frac{-b}{a}$. There is a claim that division by zero occurs and therefore is undefined. This is an invalid objection and arises from confusing division in the ring of polynomials with division in the ring of integers. In the ring of polynomials division is defined for all polynomials except the zero polynomial, and hence is defined for $p(x)=ax+b$, since $a\neq 0$.

MT, pg. 112, in the following statement and proof, neglects to specify that $f(x)$ must be a polynomial. It also does not use the division algorithm and makes some incorrect implications.



4. THE REMAINDER THEOREM

If $f(x)$ is divided by $(x-a)$ until the remainder is independent of x, then the remainder is $f(a)$.

Proof

$$\frac{f(x)}{x-a}=g(x)+\frac{R}{x-a}$$
$$\therefore \quad f(x)=(x-a)g(x)+R$$
$$\therefore \quad f(a)=(a-a)g(a)+R$$
$$=(0)g(a)+R$$
$$=0+R$$
$$=R$$

Note

To determine what to substitute for x in $f(x)$ if $f(x)$ is divided, for example, by $x-a$, $x-a$ is made equal to nil and x is determined, e.g. if $f(x)$:

(1) is divided by $x-a$ then $x-a=0 \;\therefore\; x=a$

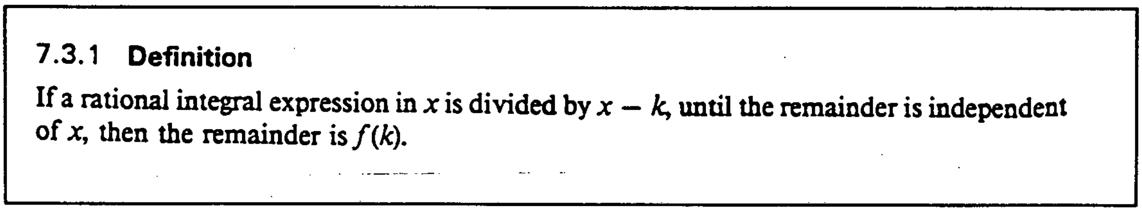
(2) is divided by $2x-3$ then $2x-3=0 \;\therefore\; x=\frac{3}{2}$

(3) is divided by $\frac{x}{2}+5$ then $\frac{x}{2}+5=0 \;\therefore\; x=-10$, etc.

For the first line of the proof to be defined it must be assumed that $x\neq a$. It is therefore not permitted to make the substitution $x=a$ in the second line. However the first line is not necessary as the second line is an application of the division algorithm.

In the "**Note**" the text makes a similar mistake by saying, "If $f(x)$ is divided by $x-a$ then $x-a=0 \;\therefore\; x=a$." (Similarly points 2 and 3 are also incorrect.) Another error concerning terminology used is the word "nil". A more appropriate word would be zero.

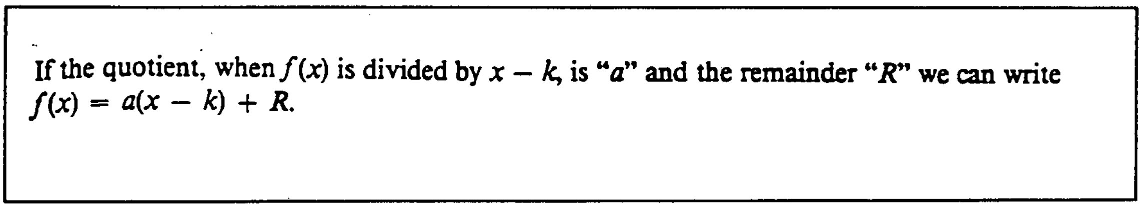
UM 9&10, pg. 72, states the remainder theorem but incorrectly calls it a "definition":



7.3.1 Definition

If a rational integral expression in x is divided by $x - k$, until the remainder is independent of x, then the remainder is $f(k)$.

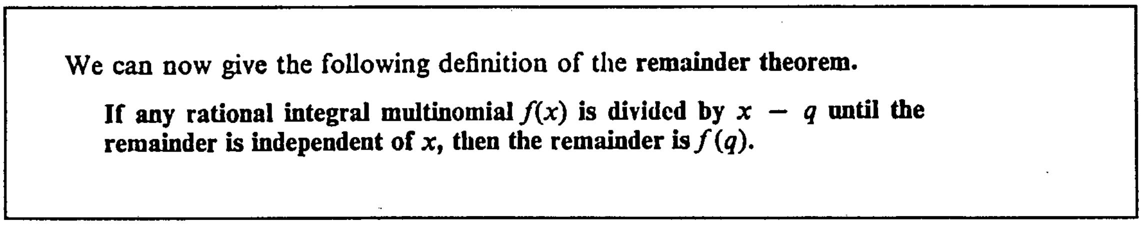
In the proof of the remainder theorem given by UM 9&10, pg. 73, the first two lines are given as follows:



If the quotient, when $f(x)$ is divided by $x - k$, is "a" and the remainder "R" we can write $f(x) = a(x - k) + R$.

Again a fact that requires the division algorithm is established without reference to the algorithm. In this proof "a" is given as the quotient which is confusing, suggesting that the quotient is a constant.

MGMSC, pg. 118, also refers to the "definition" of the remainder theorem:



We can now give the following definition of the **remainder theorem.**

If any rational integral multinomial $f(x)$ is divided by $x - q$ until the remainder is independent of x, then the remainder is $f(q)$.

A further example which illustrates the disregard for the division algorithm appears in the following extract as an "extension of the remainder theorem", while it is in fact a statement of the division algorithm when the divisor is a quadratic, a cubic or a polynomial of degree n.

MGMSC, pg. 121, says the following:

> **8.4 Extension of the remainder theorem**
>
> 1. If $f(x)$ is divided by $(x - a)(x - b)$, the remainder is $cx + d$ and we may write
> $$f(x) \equiv (x - a)(x - b)p(x) + cx + d$$
> 2. If $f(x)$ is divided by $(x - a)(x - b)(x - c)$, the remainder is $kx^2 + qx + r$ and we write:
> $$f(x) \equiv (x - a)(x - b)(x - c)p(x) + kx^2 + qx + r$$
> 3. If $f(x)$ is divided by $(x - x_1)(x - x_2) \ldots (x - x_n)$, the remainder is $a_0x^{n-1} + a_1x^{n-2} \ldots + a_{n-1}x + a_n$

5. Application of the division algorithm

The following extract is a case where there is an attempt to use the division algorithm, but there is a confusion between the algorithm for the ring of polynomials and the algorithm for the ring of integers.

The extract is taken from February 1990 issue of "Spectrum"[9]. It is an attempt to offer a solution that is simpler than one that was given in an article of an earlier issue of the same journal.

> Wiskunde-olimpiades speel 'n baie belangrike rol in die identifisering en ontwikkeling van potensiële wiskundiges.
>
> Die volgende interessante probleem het in 'n Internasionale Wiskunde-olimpiade van 1988 verskyn. Daar word beweer dat hierdie probleem, wat op die oog af eenvoudig lyk, een van die ingewikkeldste probleme is wat nog in 'n Internasionale Wiskunde-olimpiade verskyn het. Daar word gevra om te bewys dat as a en b positiewe heelgetalle is só dat $a^2 + b^2$ presies deelbaar is deur $(ab + 1)$ dan is die kwosiënt
> $$\frac{a^2 + b^2}{ab + 1}$$
> 'n volkome vierkant.
>
> Die GOK-Wiskunde en Rekenaardosente het koppe bymekaar gesit en met die volgende oplossing vorendag gekom.

$$
\begin{array}{r|l}
 & \frac{a}{b}-\frac{1}{b^2}+\frac{1}{ab^3}-\frac{1}{a^2b^4}+\dots \\
\hline
ab+1 & a^2 \qquad + b^2 \\
 & \underline{a^2+\frac{a}{b}} \\
 & \quad -\frac{a}{b}+b^2 \\
 & \quad \underline{-\frac{a}{b}-\frac{1}{b^2}} \\
 & \qquad \frac{1}{b^2}+b^2 \\
 & \qquad \underline{\frac{1}{b^2}+\frac{1}{ab^3}} \\
 & \qquad\quad -\frac{1}{ab^3}+b^2 \\
 & \qquad\quad \underline{-\frac{1}{ab^3}-\frac{1}{a^2b^4}} \\
 & \qquad\qquad\quad \frac{1}{a^2b^4}+b^2 \\
 & \qquad\qquad\qquad \vdots
\end{array}
$$

Uit die delingsalgoritme volg dat die

Res $= \frac{(-1)^k}{a^{k-2}b^k} + b^2$ met $k \geqslant 1, k \in \mathrm{N}$.

Omdat $a^2 + b^2$ presies deelbaar is deur $ab + 1$ moet

$$\frac{(-1)^k}{a^{k-2}b^k} + b^2 = 0$$

vir een of ander k.

Uit $\frac{(-1)^k}{a^{k-2}b^k} + b^2 = 0$ volg dan dat $a^{k-2}b^{k+2} = (-1)^{k+1}$

As $k = 1$, dan is $a^{-1}b^3 = 1$, waaruit volg dat $a = b^3$.

Vir $k \geqslant 2$ en k ewe sal $a^{k-2}b^{k+2} = -1$, wat nie moontlik is nie, aangesien a en b positiewe heelgetalle is.

Vir $k \geqslant 3$ en k onewe sal $a^{k-2}b^{k+2} = 1$, wat slegs moontlik is as $a = b = 1$.

Dit is dan duidelik dat die res nul sal wees vir $k = 1$ in welke geval $a = b^3$.
Gevolglik sal

$$\frac{a^2+b^2}{ab+1} = \frac{b^6+b^2}{b^4+1} = \frac{b^2(b^4+1)}{b^4+1} = b^2$$

If we are to regard $a^2 + b^2$ and $ab + 1$ as polynomials in a and b, then the long division that is executed yields a quotient which is not a polynomial. So the division algorithm for polynomials has not been applied properly. In terms of division of polynomials $a^2 + b^2$ is not divisible by $ab + 1$.

In this extract there appears to be great confusion between the two versions of the division algorithm as it is not clear which one is being used. The algorithm for polynomials does not apply, as established, so it would seem that the one for integers should be used. Certainly for $a^2 + b^2$ and $ab + 1$ as integers one can apply the division algorithm, but the quotient and remainder will be integers. Considering the proof that has been carried out, one can see that the quotient is not an integer. The algorithm for integers is also not used correctly.

5. REAL NUMBERS

1. Introduction

The discussions found in textbooks of rational and irrational numbers bring to light four main items of concern:

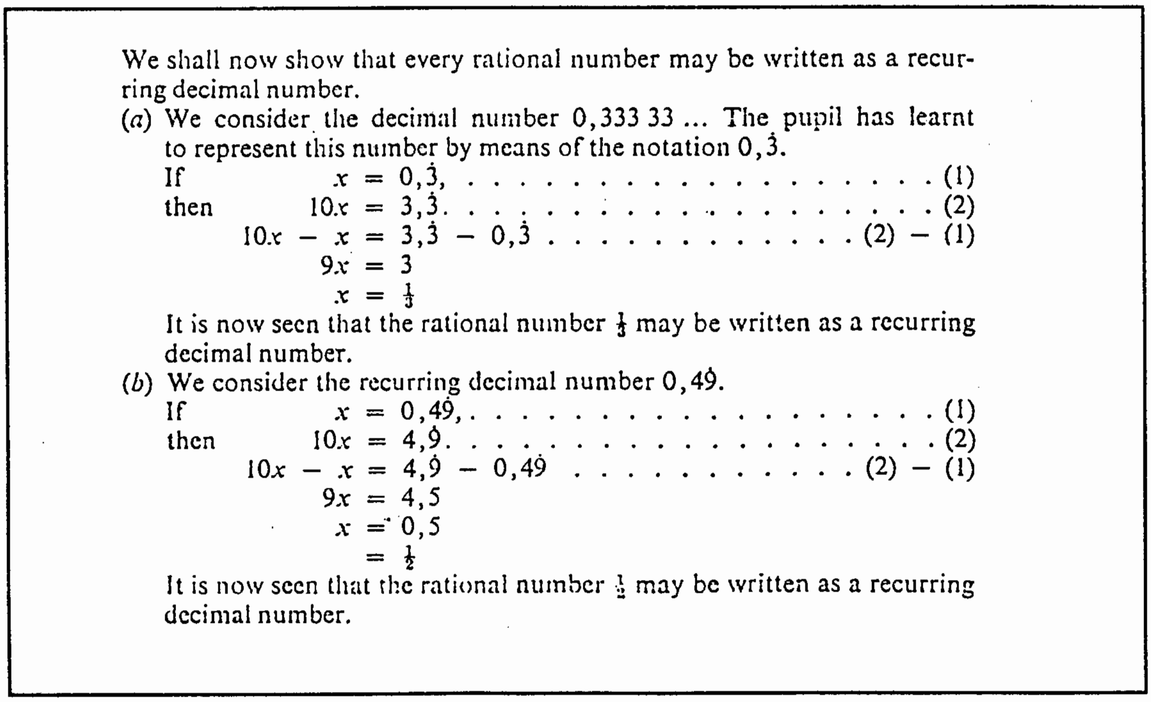
- the explanations of why rational numbers are precisely those numbers whose decimal expansions are recurring or terminating,
- incorrect proofs of irrationality,
- errors in discussion of "density" of the rational numbers in the real numbers,
- careless statements about the representation of rational numbers on the number line.

This chapter deals with the various ways texts handle these ideas.

2. Rational numbers

A rational number is of the form $\frac{a}{b}$, where a and b are integers and $b \neq 0$. Every rational number can be expressed as a recurring decimal. This result is easy to prove by a simple division process, but none of the text books referred to in this section did so.

NAM 8 , pg. 3, states the result, but then deals with it as follows:



We shall now show that every rational number may be written as a recurring decimal number.

(*a*) We consider the decimal number 0,333 33 ... The pupil has learnt to represent this number by means of the notation $0{,}\dot{3}$.

$$\begin{array}{llrcll}
\text{If} & & x & = & 0{,}\dot{3}, \ldots\ldots\ldots & (1)\\
\text{then} & & 10x & = & 3{,}\dot{3}. \ldots\ldots\ldots & (2)\\
 & 10x - & x & = & 3{,}\dot{3} - 0{,}\dot{3} \ldots\ldots\ldots & (2)-(1)\\
 & & 9x & = & 3 & \\
 & & x & = & \frac{1}{3} &
\end{array}$$

It is now seen that the rational number $\frac{1}{3}$ may be written as a recurring decimal number.

(*b*) We consider the recurring decimal number $0{,}4\dot{9}$.

$$\begin{array}{llrcll}
\text{If} & & x & = & 0{,}4\dot{9}, \ldots\ldots\ldots & (1)\\
\text{then} & & 10x & = & 4{,}\dot{9}. \ldots\ldots\ldots & (2)\\
 & 10x - & x & = & 4{,}\dot{9} - 0{,}4\dot{9} \ldots\ldots\ldots & (2)-(1)\\
 & & 9x & = & 4{,}5 & \\
 & & x & = & 0{,}5 & \\
 & & & = & \frac{1}{2} &
\end{array}$$

It is now seen that the rational number $\frac{1}{2}$ may be written as a recurring decimal number.

(*c*) In like manner it can be shown that $1 = 0{,}\dot{9}$; $2 = 1{,}\dot{9}$; $\frac{3}{4} = 0{,}75 = 0{,}74\dot{9}$; $\frac{84}{99} = 0{,}\dot{8}\dot{4}$; $\frac{35}{333} = 0{,}\dot{1}0\dot{5}$, etc.

(*d*) From the examples given, and from many similar examples, we conclude:

Every rational number can be written as a recurring decimal number.

In this extract there is a claim that every rational may be written as a recurring decimal number. Although the claim is true, it is not what is shown in this extract. Instead the approach suggests why in general a recurring decimal represents a rational number.

NMM 9, pg. 8, says the following:

Every rational number has a decimal equivalent, which is either a terminating decimal or a non-terminating recurring decimal.

A proof of this statement is not given, but examples show that every non-terminating recurring decimal is a rational number.

(*b*) Every non-terminating recurring decimal may be expressed as a fraction by the method shown below:

Example

Express in fraction form:

(1) $0{,}\dot{7}$ (2) $0{,}\dot{3}\dot{6}$ (3) $0{,}3\dot{7}\dot{2}$ (4) $0{,}\dot{4}6\dot{8}$

(1) Let
$$\begin{array}{lrl} & x &= 0{,}\dot{7} \\ \text{Then} & 10x &= 7{,}777\ldots \\ & x &= \underline{0{,}777\ldots} \\ \text{Subtract: } \therefore & 9x &= 7 \\ \therefore & x &= \frac{7}{9} \\ \therefore & 0{,}\dot{7} &= \frac{7}{9} \end{array}$$

(2) Let
$$\begin{array}{lrl} & x &= 0{,}\dot{3}\dot{6} \\ \text{Then} & 100x &= 36{,}3636\ldots \\ & x &= \underline{\ \ 0{,}3636\ldots} \\ \text{Subtract: } \therefore & 99x &= 36 \\ \therefore & x &= \frac{36}{99} = \frac{4}{11} \\ \therefore & 0{,}\dot{3}\dot{6} &= \frac{4}{11} \end{array}$$

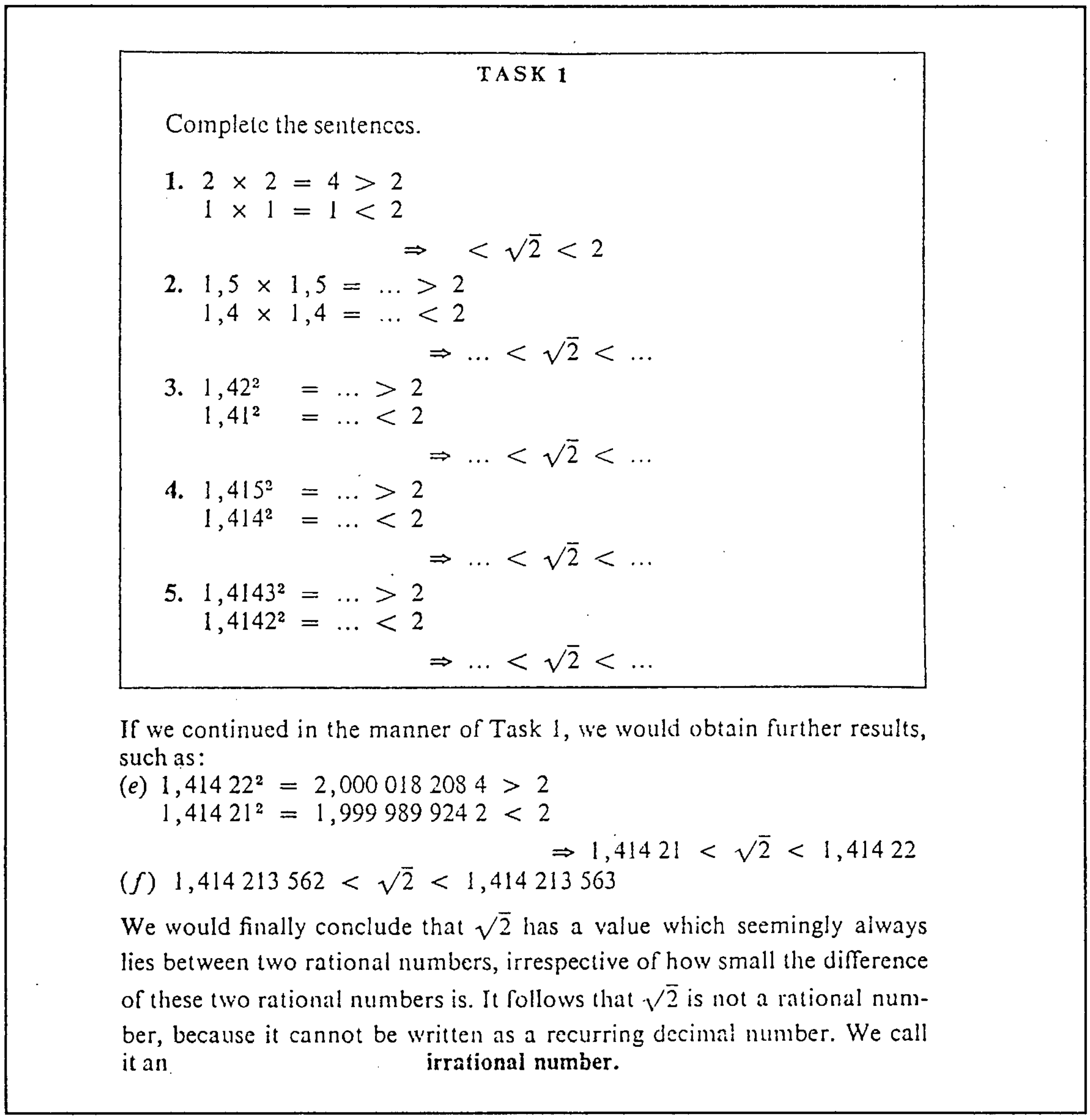
(3) Let
$$\begin{array}{lrl} & x &= 0{,}3\dot{7}\dot{2} \\ \text{Then} & 100x &= 37{,}27272\ldots \\ & x &= \underline{\ \ 0{,}37272\ldots} \\ \text{Subtract: } \therefore & 99x &= 36{,}9 \\ \therefore & x &= \frac{36{,}9}{99} = \frac{369}{990} \\ & &= \frac{41}{110} \\ \therefore & 0{,}3\dot{7}\dot{2} &= \frac{41}{110} \end{array}$$

(4) Let
$$\begin{array}{lrl} & x &= 0{,}\dot{4}6\dot{8} \\ \text{Then} & 1000x &= 468{,}468468\ldots \\ & x &= \underline{\ \ \ 0{,}468468\ldots} \\ \text{Subtract: } \therefore & 999x &= 468 \\ \therefore & x &= \frac{468}{999} = \frac{52}{111} \\ \therefore & 0{,}\dot{4}6\dot{8} &= \frac{52}{111} \end{array}$$

3. Irrational numbers

Several text books attempt to prove irrationality (usually of $\sqrt{2}$) by working out a few terms of the decimal expansion of the number and then asserting that no recurrence takes place.

NAM 8, pg. 4, proceeds as follows:



TASK 1

Complete the sentences.

1. $2 \times 2 = 4 > 2$
   $1 \times 1 = 1 < 2$
   $\Rightarrow \quad < \sqrt{2} < 2$
2. $1{,}5 \times 1{,}5 = \ldots > 2$
   $1{,}4 \times 1{,}4 = \ldots < 2$
   $\Rightarrow \ldots < \sqrt{2} < \ldots$
3. $1{,}42^2 = \ldots > 2$
   $1{,}41^2 = \ldots < 2$
   $\Rightarrow \ldots < \sqrt{2} < \ldots$
4. $1{,}415^2 = \ldots > 2$
   $1{,}414^2 = \ldots < 2$
   $\Rightarrow \ldots < \sqrt{2} < \ldots$
5. $1{,}4143^2 = \ldots > 2$
   $1{,}4142^2 = \ldots < 2$
   $\Rightarrow \ldots < \sqrt{2} < \ldots$

If we continued in the manner of Task 1, we would obtain further results, such as:

(*e*) $1{,}414\,22^2 = 2{,}000\,018\,208\,4 > 2$
$1{,}414\,21^2 = 1{,}999\,989\,924\,2 < 2$
$\Rightarrow 1{,}414\,21 < \sqrt{2} < 1{,}414\,22$

(*f*) $1{,}414\,213\,562 < \sqrt{2} < 1{,}414\,213\,563$

We would finally conclude that $\sqrt{2}$ has a value which seemingly always lies between two rational numbers, irrespective of how small the difference of these two rational numbers is. It follows that $\sqrt{2}$ is not a rational number, because it cannot be written as a recurring decimal number. We call it an **irrational number.**

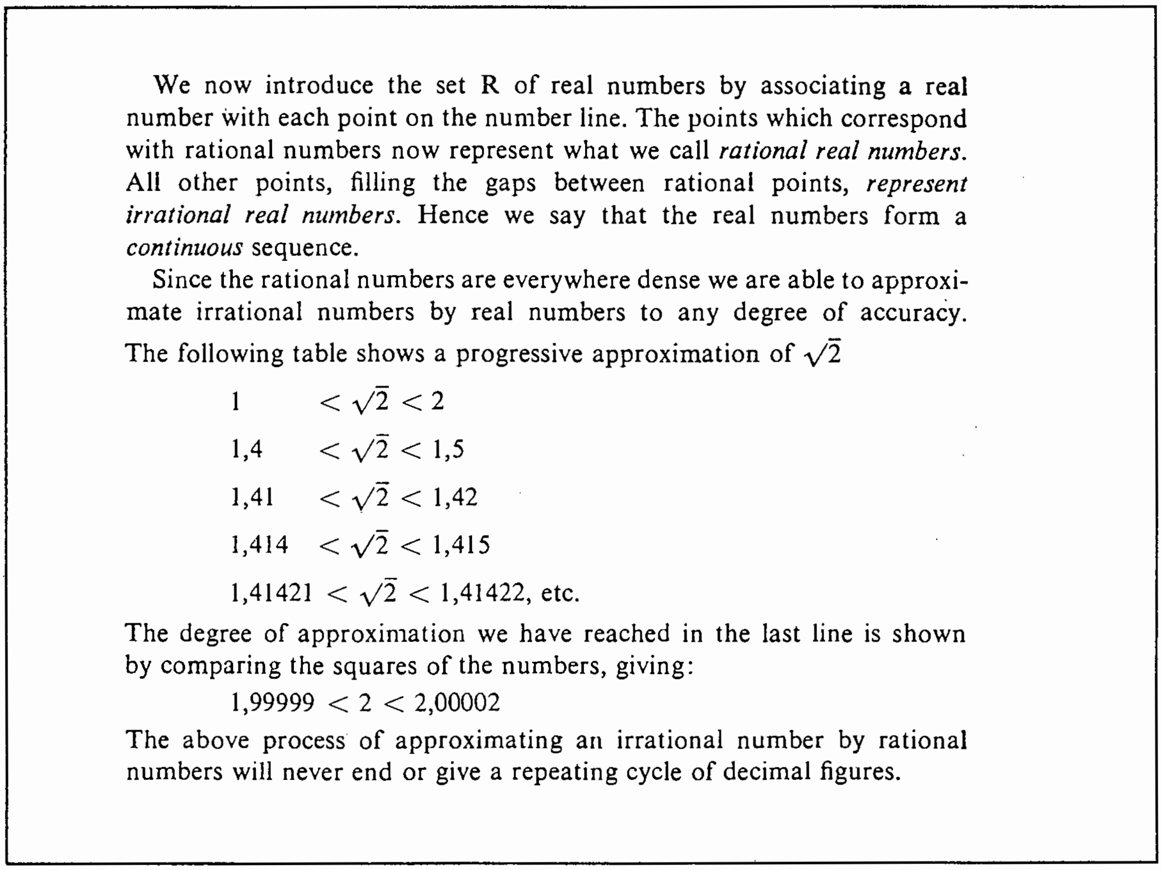
In fact, this extract does not even assert that no recurrence has been found. It seems to imply that any number which lies between two rational numbers, irrespective of how small the difference of these two numbers is, is irrational. Thus one could apparently prove that $\frac{1}{7}$ is irrational as follows:

$$0{,}7 < 1 < 1{,}05 \quad \Rightarrow \quad 0{,}1 < \frac{1}{7} < 0{,}15$$

$$0{,}98 < 1 < 1{,}001 \quad \Rightarrow \quad 0{,}14 < \frac{1}{7} < 0{,}143$$

$$0{,}994 < 1 < 1{,}0003 \quad \Rightarrow \quad 0{,}142 < \frac{1}{7} < 0{,}1429$$

MM 9&10, pg. 17, discusses irrationality of $\sqrt{2}$ in a similar vein:



We now introduce the set R of real numbers by associating a real number with each point on the number line. The points which correspond with rational numbers now represent what we call *rational real numbers.* All other points, filling the gaps between rational points, *represent irrational real numbers.* Hence we say that the real numbers form a *continuous* sequence.

Since the rational numbers are everywhere dense we are able to approximate irrational numbers by real numbers to any degree of accuracy. The following table shows a progressive approximation of $\sqrt{2}$

$$1 < \sqrt{2} < 2$$
$$1{,}4 < \sqrt{2} < 1{,}5$$
$$1{,}41 < \sqrt{2} < 1{,}42$$
$$1{,}414 < \sqrt{2} < 1{,}415$$
$$1{,}41421 < \sqrt{2} < 1{,}41422, \text{etc.}$$

The degree of approximation we have reached in the last line is shown by comparing the squares of the numbers, giving:

$$1{,}99999 < 2 < 2{,}00002$$

The above process of approximating an irrational number by rational numbers will never end or give a repeating cycle of decimal figures.

Again this does not establish soundly that the decimal expansion of $\sqrt{2}$ will never end in a repeating cycle of decimal figures. In addition to this proof being incomplete there is an error at the end of the first paragraph in the extract. Here the text says that real numbers form a continuous **sequence**. A real valued sequence is the range of a function from the natural numbers to the real numbers, and therefore is a countable collection of numbers. The real numbers, as a whole, are **uncountable** and thus cannot be referred to as a sequence. A second error occurs in the second line of the second paragraph: the word "real" should be replaced by the word "rational".

MGM 8, pg. 33, says that irrational numbers cannot be written precisely in the form $\frac{a}{b}$ where a and b are integers, $b \neq 0$, and classifies the irrational numbers into **surds** and **non-surds**.

> (a) Surds
>
> A surd is a radical number like $\sqrt{5}$ or $\sqrt[3]{7}$ which cannot be determined exactly but to a degree of accuracy to one, two, three, etc. decimal places.
>
> Using either the method of successive approximations or division discussed in Standard 7 or by tables we find that
>
> $\sqrt{2}$ = 1,4 correct to 1 decimal place
> = 1,41 correct to 2 decimal places
> = 1,414 correct to 3 decimal places.
> = 1,414 2 correct to 4 decimal places. Thus we find that
>
> $1{,}414 < \sqrt{2} < 1{,}415$. If we continue this process we have $1{,}414\,21 < \sqrt{2} < 1{,}414\,22$. Although it is impossible for us to distinguish between points representing 1,414 21 and 1,414 22 on the number line $\sqrt{2}$ still lies between them. Between points representing values like $\sqrt{2}$ there are infinitesimally small openings on the number line for rational numbers. A number line for irrational numbers is also dense, e.g. $\sqrt{2{,}066\,01}$; $\sqrt{2{,}000\,02}$; $\sqrt{2{,}000\,03}$ etc. are all irrational numbers, very close to one another on the number line and there is an infinite number of them. $\sqrt{5}$; $\sqrt[3]{7}$; $\sqrt[7]{9}$; $2 - \sqrt[5]{8}$ are other examples of radicals (surds) which can approximately only be obtained in the form $\frac{a}{b}$.

In this example it appears that a similar method, as used in the extract prior to this one, is being attempted. The paragraph offers no clarity on why the $\sqrt{2}$ is irrational. It may be difficult, but it is **not** impossible to distinguish between 1,41421 and 1,41422 on the number line, given a large enough scale.

The text claims that irrational numbers are dense and tries to substantiate this with an example of a list of irrational numbers that are very close together. One would presume that there is a printing error and that the first number listed should be $\sqrt{2{,}00001}$. The use of the "etc." would imply that this list of "irrational numbers" can be continued by increments of 0,00001. By doing this we would reach $\sqrt{2{,}01640}$ which is a rational number (equal to 1,42). The numbers $\sqrt{2{,}00001}$, $\sqrt{2{,}00002}$,etc. do in fact get closer together (since $\sqrt{a+nh} - \sqrt{a+(n-1)h} = \dfrac{h}{\sqrt{a+nh} + \sqrt{a+(n-1)h}} \to 0$ as $n \to \infty$), but the sequence nevertheless tends to infinity. This sequence therefore has no relevance to the subject of density of the irrational numbers.

This text also notes five points after the above paragraph, of which the second and third are as follows:

> 2. If $a < b$ then $\sqrt{a} < \sqrt{b}$, $\sqrt[3]{a} < \sqrt[3]{b}$ etc.
> 3. If $a < x^2 < b$ then $\sqrt{a} < x < \sqrt{b}$ and conversely. Also if $a > x^3 > b$

It must be specified that a, b and x are positive. In point 3 if $x = -3$, $a = 8$ and $b = 10$ then the statement is false.

MGM 8, pg. 33, gives a short paragraph on "non-surds" as follows:

> (*b*) **Non-surds**
>
> Irrational numbers need not be surds. All non-terminating but non-recurring decimal fractions are irrational. Thus $\pi = \dfrac{\text{circumference of any circle}}{\text{diameter of the same circle}}$ is an irrational number. Remember that the values assigned to π such as $\frac{22}{7}$; 3,14; 3,1416 are all rational numbers but they are only approximate values of π. π can be obtained from formulae such as $\dfrac{\pi}{8} = \dfrac{1}{1 \times 3} + \dfrac{1}{5 \times 7} + \dfrac{1}{9 \times 11} + \dfrac{1}{13 \times 15} + \cdots$
>
> Using computers π has already been calculated to more than a thousand digits after the decimal sign and is not a recurring or a terminating decimal.

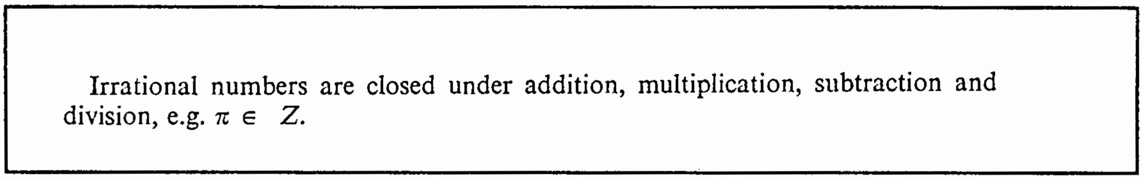
The irrationality of π is not established by the fact that computers have calculated its value to more than a thousand decimal places and have not yet found recurrence. The "formula" that is given for π does not show irrationality either.

Because this study has been restricted to South African texts, the following extract is strictly not a part of this study. It is taken from an American journal and it is included as a matter of interest because it resembles extracts that have already been discussed.

> **FFF #120. A Quick (?) Proof of Irrationality**
>
> **Problem.** Explain why $4^{1/2}$ is rational while $5^{1/2}$ is irrational.
>
> *Solution.* $4^{1/2} = 2$ which is rational. $5^{1/2}$, in its decimal form, does not terminate or repeat and therefore cannot be written as an integer over an integer.

This extract was submitted to the journal by a professor at an American university. It was taken from an unnamed American text book and the solution from the teacher's edition. The question raised is: how does one prove that the decimal expansion of $5^{\frac{1}{2}}$ has non-repeating decimal expansion without first proving irrationality?[10]

MGM 8, pg. 34, makes an incorrect statement concerning properties of operations on the irrational numbers, and then gives an unrelated example.

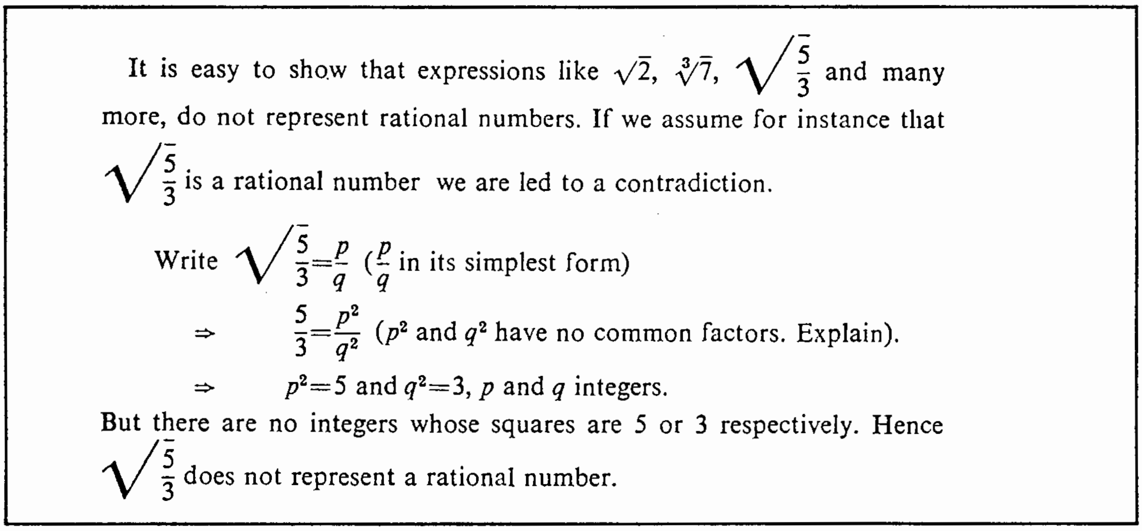


Irrational numbers are closed under addition, multiplication, subtraction and division, e.g. $\pi \in Z$.

The irrational numbers are not closed under any of the operations mentioned. The example of $\pi \in \mathbf{Z}$ has nothing to do with the question of closure under any of the operations. Another point to note is that the symbol **Z** is usually used for the set of integers and not for irrational numbers.

The irrationality of numbers is often proved using "proof by contradiction".

MM 9&10, pg. 16, has the following proof:



It is easy to show that expressions like $\sqrt{2}$, $\sqrt[3]{7}$, $\sqrt{\frac{5}{3}}$ and many more, do not represent rational numbers. If we assume for instance that $\sqrt{\frac{5}{3}}$ is a rational number we are led to a contradiction.

Write $\sqrt{\frac{5}{3}}=\frac{p}{q}$ ($\frac{p}{q}$ in its simplest form)

$\Rightarrow \quad \frac{5}{3}=\frac{p^2}{q^2}$ (p^2 and q^2 have no common factors. Explain).

$\Rightarrow \quad p^2=5$ and $q^2=3$, p and q integers.

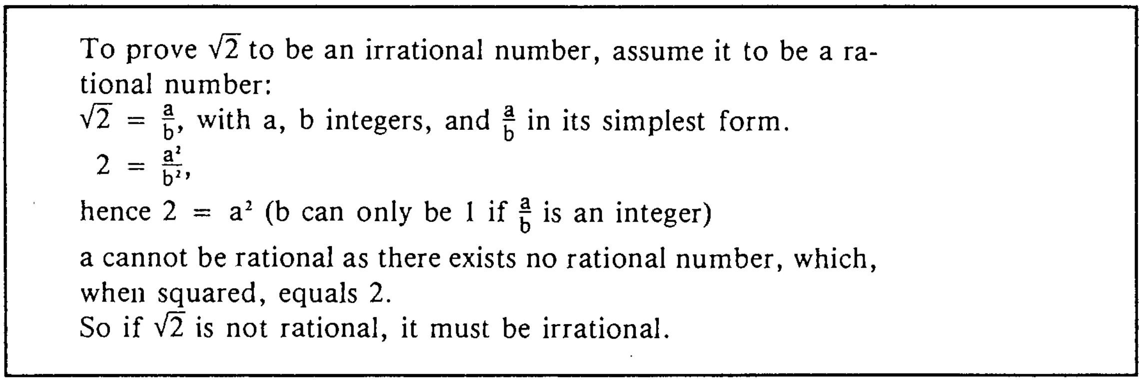
But there are no integers whose squares are 5 or 3 respectively. Hence $\sqrt{\frac{5}{3}}$ does not represent a rational number.

The second implication in the above extract is not valid. This can be shown to break down if the above proof is imitated for $\sqrt{\frac{8}{2}}$. If this is done, i.e. the same argument is carried out for $\sqrt{\frac{8}{2}}$, then the conclusion that $p^2 = 8$ and $q^2 = 2$, where p and q are integers, must be made. Since there are no such integers p and q that make $p^2 = 8$ and $q^2 = 2$, the logical conclusion in this proof would be that $\sqrt{\frac{8}{2}}$ is not rational. However it is a rational number. Thus in the above extract it cannot be said that $\frac{5}{3} = \frac{p^2}{q^2} \Rightarrow p^2 = 5$ and $q^2 = 3$, p and q integers, without further justification. Two facts that need verification and are required in this proof are:

1. If $\frac{a}{b} = \frac{c}{d}$ and $(a,b) = (c,d) = 1$ then $a = c$ and $b = d$
2. If $(p,q) = 1$ then $(p^2,q^2) = 1$.

(Here (m,n) is defined as the greatest common divisor of the integers m and n.)

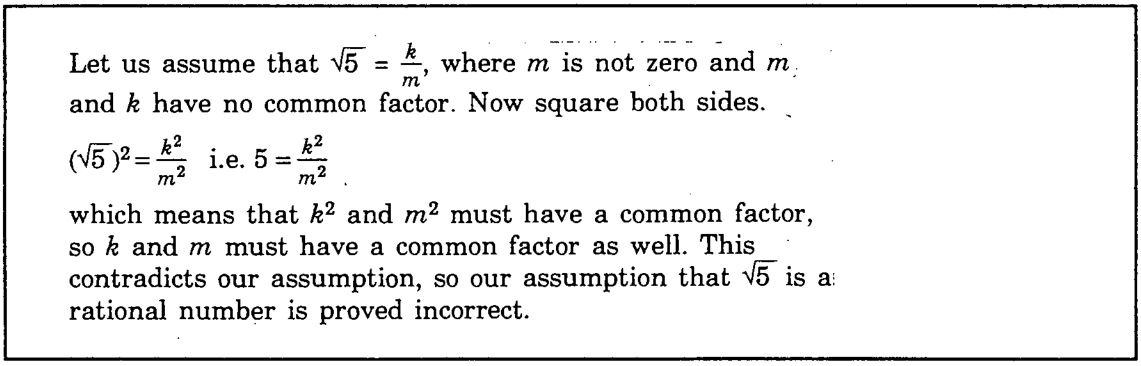
PM 7, pg. 3, says the following:



To prove $\sqrt{2}$ to be an irrational number, assume it to be a rational number:
$\sqrt{2} = \frac{a}{b}$, with a, b integers, and $\frac{a}{b}$ in its simplest form.
$2 = \frac{a^2}{b^2}$,
hence $2 = a^2$ (b can only be 1 if $\frac{a}{b}$ is an integer)
a cannot be rational as there exists no rational number, which, when squared, equals 2.
So if $\sqrt{2}$ is not rational, it must be irrational.

This argument is difficult to understand. The deduction "hence $2 = a^2$" is not justified by the statement in brackets. It is not true that $b = 1$, so $2 = a^2$ is untrue. (It is not even true that $\frac{a}{b}$ is an integer.) The conclusion "*a* cannot be rational as there exists no rational number, which, when squared, equals 2" is in no way a deduction from "$2 = a^2$". (It is in fact a restatement of the theorem that requires proof). The word "rational" needs to be replaced by "integer" to make the logical inference valid, but it does not correct the proof.

MP 9, pg. 10 , contains a similar error:



Let us assume that $\sqrt{5} = \frac{k}{m}$, where m is not zero and m and k have no common factor. Now square both sides.

$(\sqrt{5})^2 = \frac{k^2}{m^2}$ i.e. $5 = \frac{k^2}{m^2}$.

which means that k^2 and m^2 must have a common factor, so k and m must have a common factor as well. This contradicts our assumption, so our assumption that $\sqrt{5}$ is a rational number is proved incorrect.

The argument is clearly suspect, since if "5" is replaced by "4", we apparently are able to prove that $\sqrt{4}$ is irrational. The proof again requires verification of the fact: "If $(p,q)=1$ then $(p^2,q^2)=1$".

4. Density

The correct definition of density is as follows: A set $\mathbf{A} \subset \mathbf{R}$ is **dense in R** if for all $x<y$ in $\mathbf{R}$ there exists an $a \in \mathbf{A}: x<a<y$. Note that there is reference to a superset which is part of the definition of the concept of density.

With regard to the topic of real numbers, one of the most common issues of concern in text books is that of density. Texts have most frequently used the following defining idea for the concept of density:

For want of a name we shall refer to the following principle as **order density.** A set $\mathbf{A} \subset \mathbf{R}$ is **"order dense"** if for all $x<y \in \mathbf{A}$ there exists an $a \in \mathbf{A}: x<a<y$.

A number of texts use the definition of density incorrectly. This is a small but significant change in the definition of density and as used in all these cases concerning rational numbers: the points x and y are restricted to being **rational** numbers. This becomes a case of **order** density.

Very often in texts density is explained using the criterion for order density. The Afrikaans text WMI 8, pg. 41, writes about order density but refers to it as density. The English translation is as follows:

> So, another rational number can always be placed halfway between any two rational numbers, no matter how close they are together: the points that represent the rational numbers are practically next to each other and hence mathematicians say that the collection of points, that represent the rational numbers on the number line, is dense.
>
> **The collection of rational numbers is dense because between any two rational numbers another rational number can be found.**

This is the first of several examples that write about density, but are strictly speaking referring to "order density".

Order density does not imply density. Before looking at more examples of the confusion between the two types of density consider the following simple example:

The interval $(0,1)$ is "order dense", but not dense in **R.**

Textbooks claim that because there are infinitely many rational numbers between any two rational numbers, the rational numbers are dense in the real numbers. This is not a criterion for density. The following is an example of a set that exhibits the property that between any two elements of the set there are infinitely many other elements of the set, yet is not dense. Indeed, we shall see that it is nowhere dense (the interior of its closure is empty). The example is a modification of the Cantor set.

The Cantor set is constructed as follows:

Let $I_0 = [0,1]$

$I_1 = [0,\frac{1}{3}] \cup [\frac{2}{3},1]$

$I_2 = [0,\frac{1}{9}] \cup [\frac{2}{9},\frac{3}{9}] \cup [\frac{6}{9},\frac{7}{9}] \cup [\frac{8}{9},1]$

I_n is defined inductively with I_{n+1} obtained by omitting open middle thirds of each interval that makes up I_n.

This yields a contracting family of closed sets .

Let $C = \bigcap\{ I_n \mid n \in N \}$ where $N = \{0,1,2,3.....\}$

C is called the Cantor set, and is non-empty because $[0,1]$ is compact and the I_n's form a family of closed sets satisfying the finite intersection property (F.I.P.)

C is uncountable and every point in C has a unique ternary expansion of the form

$$\sum_{n=1}^{\infty} \frac{a_n}{3^n} \qquad \text{where } a_n \text{ is 0 or 2.}$$

The proofs of these statements can be found in any standard text on real analysis.

Consider now a similar construction, except that now closed middle thirds are deleted from $(0,1)$ etc. In other words the following sets are constructed.

Let $I_0 = (0,1)$

$I_1 = (0,\frac{1}{3}) \cup (\frac{2}{3},1)$

$I_2 = (0,\frac{1}{9}) \cup (\frac{2}{9},\frac{3}{9}) \cup (\frac{6}{9},\frac{7}{9}) \cup (\frac{8}{9},1)$

Let $A = \bigcap\{ I_n \mid n \in N \}$

In comparing A with C it is clear that a countable collection of numbers have been omitted from the Cantor set to construct A (i.e. the end points of each interval in I_n $\forall\, n \in N$.)

Since C is uncountable, what remains after omitting this countable number of end points is an uncountable set. Hence A is non-empty.

The elements that have been deleted from **C**, in order to form **A**, each have ternary expansions that end in a recurring string of 0's or 2's. Thus the elements that remain (i.e. elements of **A**) end with non-recurring strings of 0's or 2's.

CLAIM:

For all $a, b \in \mathbf{A}$ such that $a > b$ there exists a $c \in \mathbf{A}$ such that $b < c < a$

Proof:

Take any $a, b \in \mathbf{A}$ with $a > b$

$$a = \sum_{n=1}^{\infty} \frac{a_n}{3^n} \quad \text{and} \quad b = \sum_{n=1}^{\infty} \frac{b_n}{3^n}$$

where each a_n and b_n is a 0 or 2 and there are no tails of 0's or 2's.

For the sake of notation $a = 0{,}a_1a_2a_3\ldots.$
$b = 0{,}b_1b_2b_3\ldots.$

Now there is a first m such that $a_m > b_m$ (since $a > b$)
i.e. $a_m = 2$ and $b_m = 0$ and $a_i = b_i$ $(1 \le i < m)$.

Then there is a $k > m$ such that $b_k = 0$ (since b does not have tail of 2's)

Let $\mathbf{c} = 0{,}c_1c_2c_3\ldots.$ where $c_i = b_i$ if $i \neq k$ and $c_k = 2$.

Clearly $\mathbf{c} \in \mathbf{A}$ and $b < c < a$.

Since a and b were arbitrarily chosen, there are infinitely many members of **A** between any two members of **A**.

In TM 7, pg. 10, amongst other misconceptions, the confusion between order density and density arises:

- Any rational number has a fixed position on the number line, i.e. Q is an **ordered** set.
- Any irrational number also has a fixed position on the number line, i.e. Q′ is an ordered set.

- Since $\mathbb{R}$ is made up of $\mathbb{Q}$ and $\mathbb{Q}'$, which are both ordered sets, it follows that $\mathbb{R}$ is an ordered set, i.e. any real number has a fixed position on the number line.
- There is an infinite number of rational numbers and an infinite number of irrational numbers.
- The rational numbers and the irrational numbers are infinitely closely packed on the number line.
- We say that $\mathbb{Q}$ and $\mathbb{Q}'$ are both **dense**.
- It follows, therefore, that $\mathbb{R}$ is also **dense**, i.e. the real numbers are infinitely closely packed on the number line.
- The real numbers are ordered and each number has a fixed position on the number line.
- For this reason, any given real numbers can be written in ascending or descending order.

Going in order of the points above, the first point implies that **Q** is an ordered set because every element (every rational number) has a fixed position on the number line. This does not give a clear indication of what an ordered set is. The real numbers are ordered by the relation < and the rational numbers are ordered because they are a subset of an ordered set.

The fourth, fifth, sixth and seventh points do not give a clear explanation of the term **dense.** This is also the first time it is mentioned in this text book, so nowhere prior to this has an explanation been given. From these points it is understood that "infinitely closely packed" is supposed to mean "density", but from these points the actual concept of "infinitely closely packed" is vague.

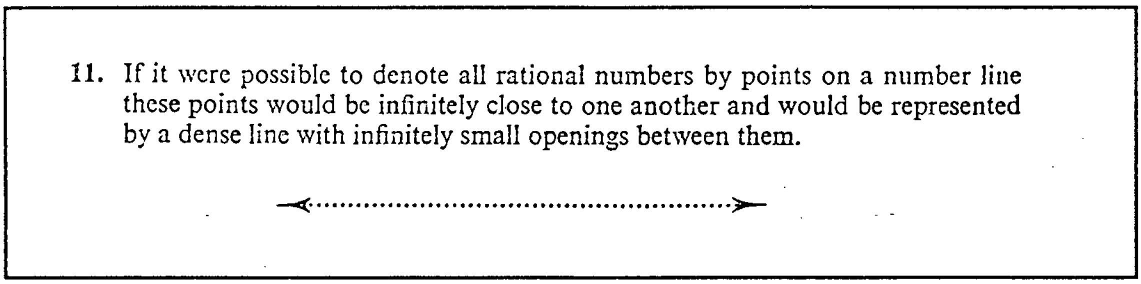
The seventh point implies that **R** is "dense" because it is the union of two "dense" sets (where "dense" is infinitely closely packed). If this is to be considered in terms of the correct meaning of density then it is true, but in terms of "infinitely closely packed", then it is not always true. For example, take two intervals, [0,1] and [2,3]. Each of these is "infinitely closely packed", but their union is not.

According to what is written in this text the last point is a result of the eighth point. The fact that the real numbers are ordered does mean any two real numbers are comparable. In fact any finite collection of real numbers are comparable in that they can be written in ascending or descending order, but not "any given real numbers". If the "any given real numbers" were the whole set of real numbers, the ability to write them in ascending or descending order would imply that a one-to-one correspondence with the natural numbers exists. This is clearly not true, and so there can't, in general, be a listing of "any given real numbers" in descending or ascending order.

MM 9&10, pg. 15, refers to density, when it is really describing order density:

> There is therefore an infinity of rational numbers between every two different rational numbers however close together these two rational numbers may be! Hence we say that the set of rational numbers is *dense*.

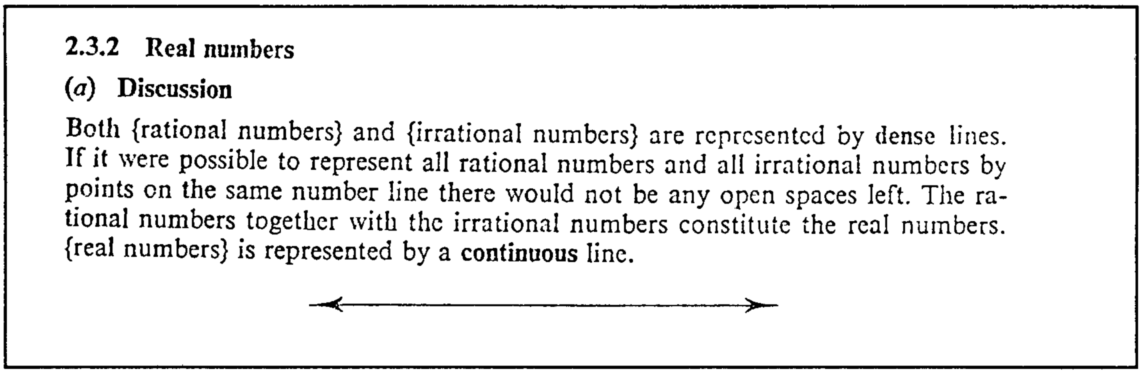
In MGM 8, pg. 31, the following is stated:



11. If it were possible to denote all rational numbers by points on a number line these points would be infinitely close to one another and would be represented by a dense line with infinitely small openings between them.

The idea that there are "infinitely small openings" together with the illustration of the dotted line does not offer any clarity on the subject of density.

In the same text on pg. 34, the following is written:



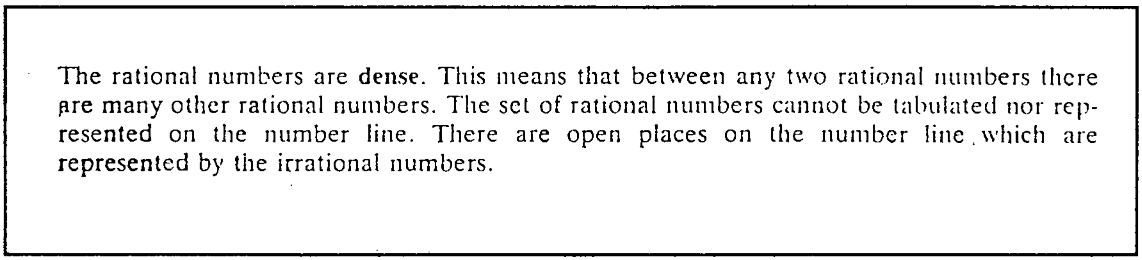
2.3.2 Real numbers

(*a*) **Discussion**

Both {rational numbers} and {irrational numbers} are represented by dense lines. If it were possible to represent all rational numbers and all irrational numbers by points on the same number line there would not be any open spaces left. The rational numbers together with the irrational numbers constitute the real numbers. {real numbers} is represented by a **continuous** line.

In this extract density is again not clearly defined. This representation of the real numbers by means of a solid continuous line together with the previous example makes the concept of density confusing.

MT 9, pg. 3, has the following paragraph on density:



The rational numbers are dense. This means that between any two rational numbers there are many other rational numbers. The set of rational numbers cannot be tabulated nor represented on the number line. There are open places on the number line which are represented by the irrational numbers.

This example again indicates the misuse of the term "dense". When referring to a **dense** set there has to be an overlying or superset. We talk of the rational numbers being dense in the set of real numbers. When considering order density one can do so without referring to a superset.

According to the extract "the set of rational numbers cannot be tabulated or represented on the number line". This would question the fact that the rational numbers are constructible and hence can be represented on the number line. The question of tabulating the rational numbers is one of whether it is possible to list or count the rational numbers. Cantor's construction of a one-to-one correspondence between the rational numbers and the natural numbers will give us a means of tabulating.

MT 9, pg. 5 , then says the following:

> Any point on this number line represents a real number and conversely: every real number can be linked to a point on the number line.

The text is now saying that every real number can be linked to a point on the number line. In other words any real number can be represented by a point on the number line, whereas previously (pg. 3: see above) it claimed that the rational numbers cannot be represented on the number line.

NAM 8 , pg. 3, attempts to do such a tabulation, but it is not a tabulation in the sense of a one-to-one correspondence with **N**.

> The s.o.a. rational numbers may be tabulated as follows:
>
> $$\begin{array}{ll} Q = & \left\{\frac{1}{1};\ \frac{2}{1};\ \frac{3}{1};\ \cdots \right. \\ & \quad \frac{1}{2};\ \frac{2}{2};\ \frac{3}{2};\ \cdots \\ & \quad \frac{1}{3};\ \frac{3}{3};\ \frac{3}{3};\ \cdots \\ & \quad \left. \ldots\ldots\ldots\ldots\ldots \right\} \\ & \cup\{0\} \\ & \cup\left\{\frac{-1}{1};\ \frac{-2}{1};\ \frac{-3}{1};\ \cdots \right. \\ & \quad \frac{-1}{2};\ \frac{-2}{2};\ \frac{-3}{2};\ \cdots \\ & \quad \frac{-1}{3};\ \frac{-2}{3};\ \frac{-3}{3};\ \cdots \\ & \quad \left. \ldots\ldots\ldots\ldots\ldots\ldots \right\} \end{array}$$

6. EXPONENTS

1. Introduction

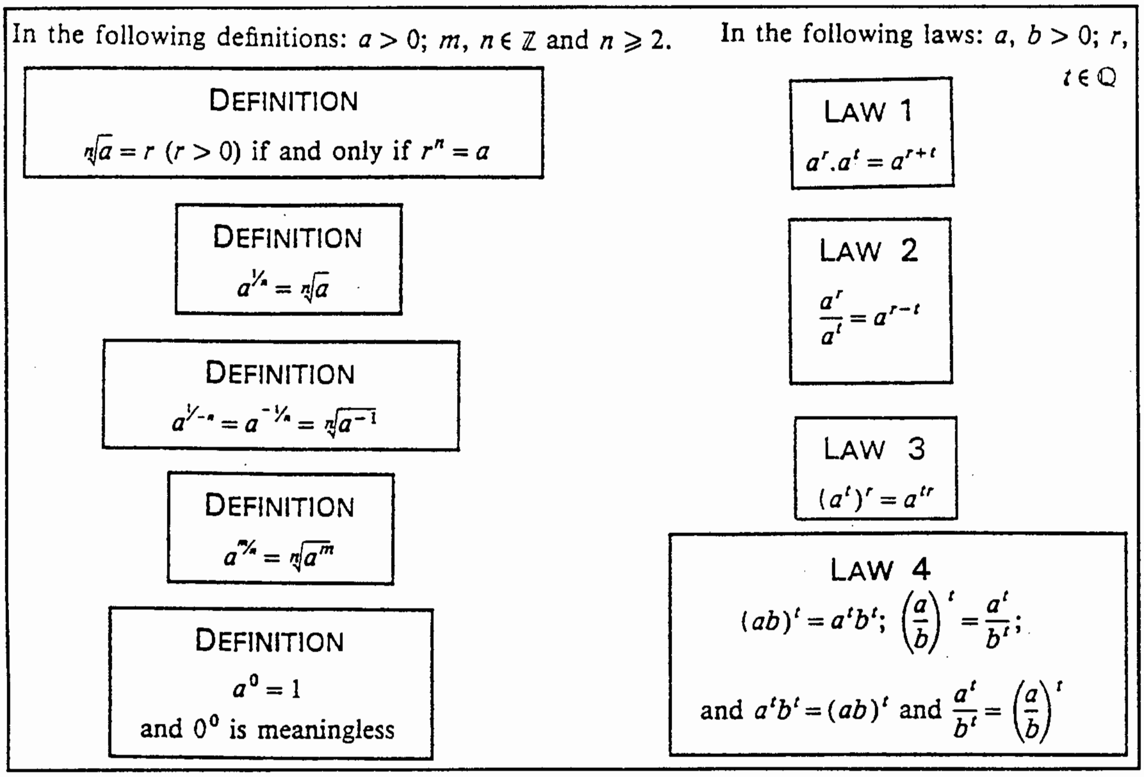
There are two different approaches that texts adopt concerning the topic of exponents. Some establish definitions and then lay down laws concerning exponents. Others establish one definition and then prove relationships which other texts regard as definitions.

In this section these different approaches will be discussed. The meaning of $a^{m/n}$ will be discussed as well as the concepts of square roots and absolute values, which follow from this definition.

2. Definitions and Laws

It is easy to establish the meaning $a^m, m \in \mathbf{Z}^+$, after which the laws pertaining to this definition are proved. However, to extend to negative integer and rational exponents poses a possible debate. Do we give artificial meaning and then show that they obey the same laws that hold for positive integer exponents? Do we give artificial meaning and assume, as part of the definition, that the laws hold? Can meaning be established by what is known for positive integers?

CM 9, pp. 197-199, gives the following as definitions and laws:



In the following definitions: $a > 0$; $m, n \in \mathbb{Z}$ and $n \geqslant 2$.

DEFINITION

$\sqrt[n]{a} = r$ $(r > 0)$ if and only if $r^n = a$

DEFINITION

$a^{1/n} = \sqrt[n]{a}$

DEFINITION

$a^{1/-n} = a^{-1/n} = \sqrt[n]{a^{-1}}$

DEFINITION

$a^{m/n} = \sqrt[n]{a^m}$

DEFINITION

$a^0 = 1$

and 0^0 is meaningless

In the following laws: $a, b > 0$; $r, t \in \mathbb{Q}$

LAW 1

$a^r . a^t = a^{r+t}$

LAW 2

$\frac{a^r}{a^t} = a^{r-t}$

LAW 3

$(a^t)^r = a^{tr}$

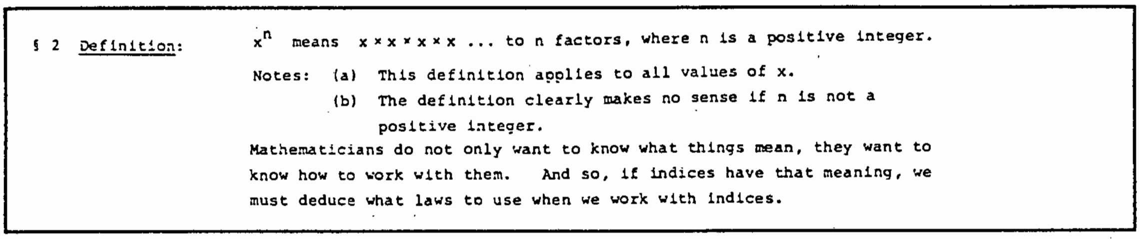
LAW 4

$(ab)^t = a^t b^t$; $\left(\frac{a}{b}\right)^t = \frac{a^t}{b^t}$;

and $a^t b^t = (ab)^t$ and $\frac{a^t}{b^t} = \left(\frac{a}{b}\right)^t$

In the above extract it can be seen as a matter of "laying down the law", whereas in the next extract the laws are proved using the definition of a^m and algebraic laws.

JM 9&10, pg. 65, gives the definition of x^m and then uses this to prove the laws of exponents.



§ 2 Definition: x^n means $x \times x \times x \times x$ ... to n factors, where n is a positive integer.

Notes: (a) This definition applies to all values of x.
(b) The definition clearly makes no sense if n is not a positive integer.

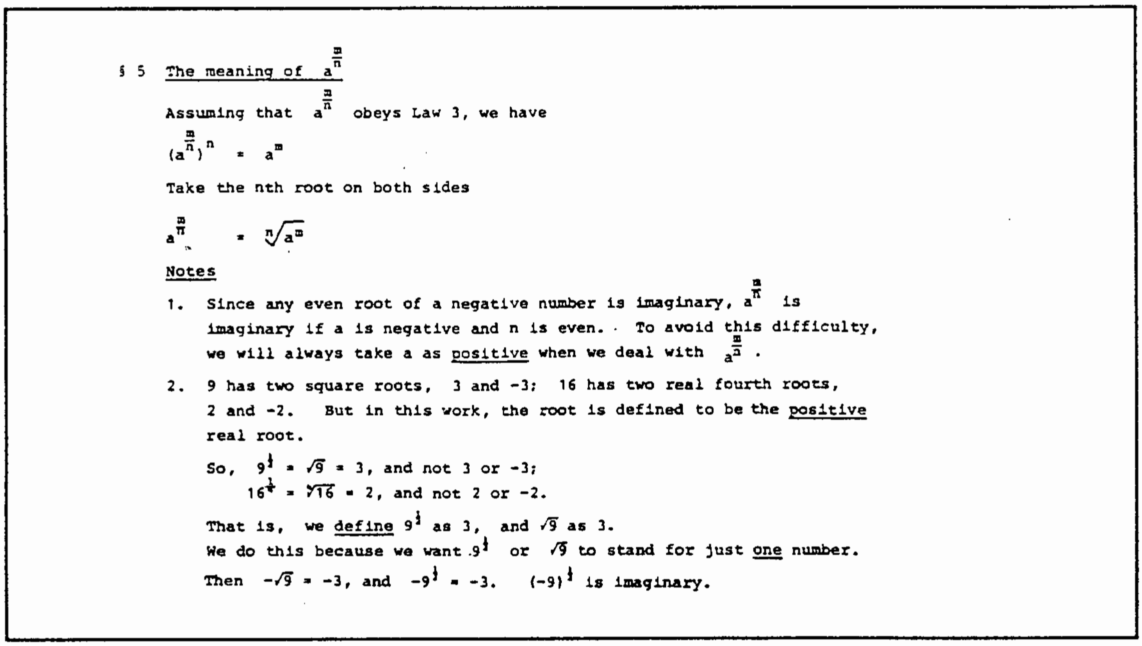
Mathematicians do not only want to know what things mean, they want to know how to work with them. And so, if indices have that meaning, we must deduce what laws to use when we work with indices.

3. Rational exponents

The main area of confusion is that of defining or explaining the meaning of powers which have rational exponents.

JM 9&10 gives the definition of x^m and then develops the theory of indices from that. The definition is given where the exponents are positive integer values and then laws are proved for this case. The extension to rational valued exponents is based on the assumption that the laws established for the integral exponents will also hold for rational exponents.

The following extract is taken from the above mentioned text, pg. 66, concerning rational exponents.



§ 5 The meaning of $a^{\frac{m}{n}}$

Assuming that $a^{\frac{m}{n}}$ obeys Law 3, we have

$(a^{\frac{m}{n}})^n = a^m$

Take the nth root on both sides

$a^{\frac{m}{n}} = \sqrt[n]{a^m}$

Notes

1. Since any even root of a negative number is imaginary, $a^{\frac{m}{n}}$ is imaginary if a is negative and n is even. To avoid this difficulty, we will always take a as positive when we deal with $a^{\frac{m}{n}}$.
2. 9 has two square roots, 3 and -3; 16 has two real fourth roots, 2 and -2. But in this work, the root is defined to be the positive real root.

So, $9^{\frac{1}{2}} = \sqrt{9} = 3$, and not 3 or -3;
$16^{\frac{1}{4}} = \sqrt[4]{16} = 2$, and not 2 or -2.

That is, we define $9^{\frac{1}{2}}$ as 3, and $\sqrt{9}$ as 3.
We do this because we want $9^{\frac{1}{2}}$ or $\sqrt{9}$ to stand for just one number.
Then $-\sqrt{9} = -3$, and $-9^{\frac{1}{2}} = -3$. $(-9)^{\frac{1}{2}}$ is imaginary.

In this paragraph, $a^{\frac{m}{n}} = \sqrt[n]{a^m}$ is shown to hold assuming that law 3 holds. In the "notes" $a^{\frac{m}{n}}$ is said to be imaginary if a is negative and n is even. This is true if m is not even. If m and n were both 2, $a^{\frac{m}{n}}$ would not be imaginary ($a^{\frac{2}{2}} = \sqrt{a^2}$ is an alternative way of defining the absolute value of a number).

Great care is required when dealing with extending the definition of x^n where $n \in \mathbf{N}$ to x^p where $p \in \mathbf{Q}^+$. The definition that $x^{\frac{m}{n}} = \sqrt[n]{x^m}$, if adhered to, allows some cases where x is negative to be considered (e.g. absolute values, see above).

This definition also rules out any apparent paradoxes such as:

$$\sqrt{(-2)^2} = (-2)^{\frac{2}{2}} = (-2)^1 = -2$$

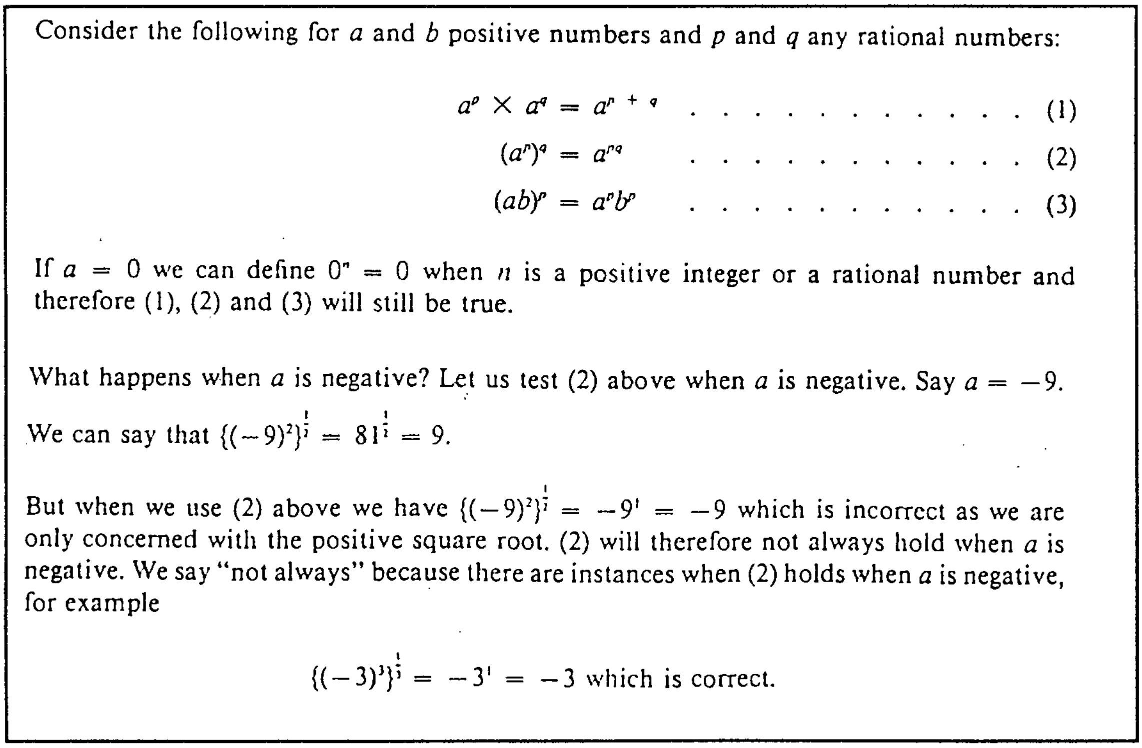
OR

$\sqrt{(-2)^2} = ((-2)^{\frac{1}{2}})^2$ which is undefined since $(-2)^{\frac{1}{2}}$ is undefined

OR

$\sqrt{(-2)^2} = \sqrt{4} = 2$ (which is the correct answer).

UM 9&10, pg. 87 and 88, says the following concerning rational exponents:



> Consider the following for a and b positive numbers and p and q any rational numbers:
>
> $$a^p \times a^q = a^{p+q} \quad \ldots\ldots\ldots\ldots \quad (1)$$
> $$(a^p)^q = a^{pq} \quad \ldots\ldots\ldots\ldots \quad (2)$$
> $$(ab)^p = a^p b^p \quad \ldots\ldots\ldots\ldots \quad (3)$$
>
> If $a = 0$ we can define $0^n = 0$ when n is a positive integer or a rational number and therefore (1), (2) and (3) will still be true.
>
> What happens when a is negative? Let us test (2) above when a is negative. Say $a = -9$.
>
> We can say that $\{(-9)^2\}^{\frac{1}{2}} = 81^{\frac{1}{2}} = 9$.
>
> But when we use (2) above we have $\{(-9)^2\}^{\frac{1}{2}} = -9^1 = -9$ which is incorrect as we are only concerned with the positive square root. (2) will therefore not always hold when a is negative. We say "not always" because there are instances when (2) holds when a is negative, for example
>
> $\{(-3)^3\}^{\frac{1}{3}} = -3^1 = -3$ which is correct.

This text confuses the issue of the law $(a^p)^q = a^{pq}$ when a is a negative number. A contradiction is obtained because it is not clear which operation must be performed first: squaring -9 first, or multiplying the exponents first. The standard would be to first square -9 and then take it to the power $\frac{1}{2}$. The reason why the second calculation is incorrect is because of this and not because " we are only concerned with the positive square root". This order of performing the operations would be in keeping with the definition $a^{\frac{m}{n}} = \sqrt[n]{a^m}$ where the operation "raising to the power m" is done first. (As is the case with the absolute value of a number given by $|a| = \sqrt{a^2}$.)

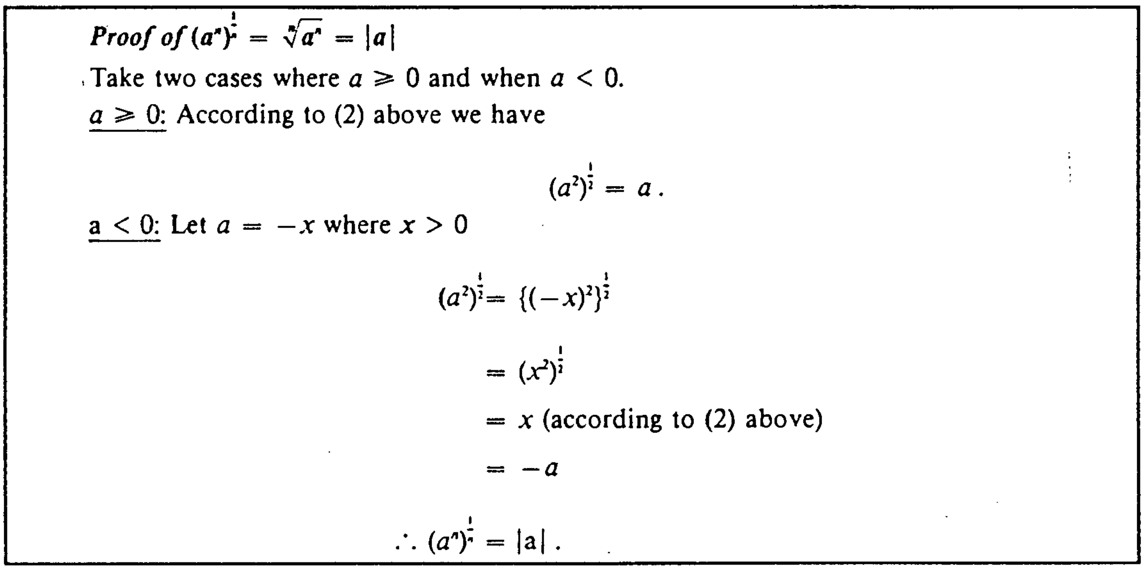
UM 9&10, pg. 88, continues as follows:

> *General rule*
>
> We can now say that for any real numbers a and n and $n \neq 0$:
>
> $$(a^n)^{\frac{1}{n}} = \left| \begin{array}{ll} a \text{ if } a \geqslant 0 & \ldots\ldots\ldots\ldots (1) \\ -a \text{ if } a < 0 & \ldots\ldots\ldots\ldots (2) \end{array} \right.$$
>
> *or*
>
> $$(a^n)^{\frac{1}{n}} = \sqrt[n]{a^n} = |a|$$
>
> Take a numerical example. Let $a = 2$ and $n = 3$.
>
> $(a^n)^{\frac{1}{n}} = (2^3)^{\frac{1}{3}} = 2^1 = 2$. This proves (1) above.
>
> Now let $a = -2$.
>
> $(a^n)^{\frac{1}{n}} = \{(-2)^3\}^{\frac{1}{3}} = -8^{\frac{1}{3}} = -2$ which proves (2) above.
>
> Now take the example when $a = -3$ and $n = 2$.
>
> $\{(-3)^2\}^{\frac{1}{2}} = -3^1 = -3$ when we apply (2) above.

The general rule as stated in this extract is incorrect. The example that $\{(-3)^{\frac{1}{3}}\}^3 = -3$ is an example that shows that the general rule is incorrect, since $|-3| = 3$. In the example where $a = -2$ the text claims that "(2) above" is proved. This is clearly not true as $|-2| = 2$. Apart from the general rule being incorrect there are contradictions within the extract. The general rule as stated is used incorrectly in the paragraph. According to (2) of the general rule, $\{(-2)^{\frac{1}{3}}\}^3 = -(-2) = 2$, which is not what the text establishes. Similarly the text says "when $a = -3$ and $n = 2$, $\{(-3)^{\frac{1}{2}}\}^2 = (-3)^1 = -3$

when we apply (2) above". This is not an application of (2), and is also incorrect as the text had established before, so there is a contradiction within the text.

UM 9&10, pg. 88, says the following:



Proof of $(a^n)^{\frac{1}{n}} = \sqrt[n]{a^n} = |a|$

Take two cases where $a \geqslant 0$ and when $a < 0$.

$\underline{a \geqslant 0}$: According to (2) above we have

$$(a^2)^{\frac{1}{2}} = a\,.$$

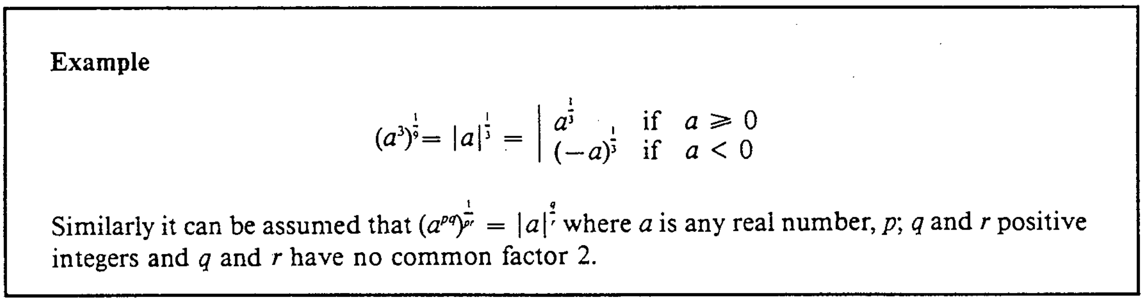
$\underline{a < 0}$: Let $a = -x$ where $x > 0$

$$\begin{aligned}(a^2)^{\frac{1}{2}} &= \{(-x)^2\}^{\frac{1}{2}}\\ &= (x^2)^{\frac{1}{2}}\\ &= x \text{ (according to (2) above)}\\ &= -a\end{aligned}$$

$$\therefore (a^n)^{\frac{1}{n}} = |a|\,.$$

The rule stated is true only if n is even. The proof given for the general rule uses $n = 2$, but doesn't establish generality.

UM 9&10, pg. 89, continues with an example:



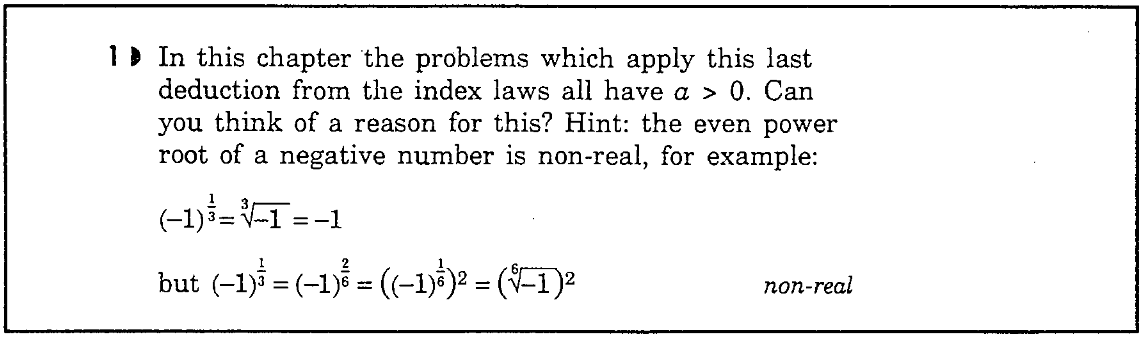
Example

$$(a^3)^{\frac{1}{9}} = |a|^{\frac{1}{3}} = \left|\begin{array}{ll} a^{\frac{1}{3}} & \text{if } a \geqslant 0 \\ (-a)^{\frac{1}{3}} & \text{if } a < 0 \end{array}\right.$$

Similarly it can be assumed that $(a^{pq})^{\frac{1}{pr}} = |a|^{\frac{q}{r}}$ where a is any real number, p; q and r positive integers and q and r have no common factor 2.

This is clearly not true. A simple counterexample to show that it is not true is as follows: if $a = -8$ then $((-8)^3)^{\frac{1}{9}} = (-512)^{\frac{1}{9}} = -2$ whereas $|-8|^{\frac{1}{3}} = 8^{\frac{1}{3}} = 2$. There is a great deal of confusion in this text concerning the meaning of $a^{\frac{m}{n}}$.

MP 9, pg. 156, writes the following:



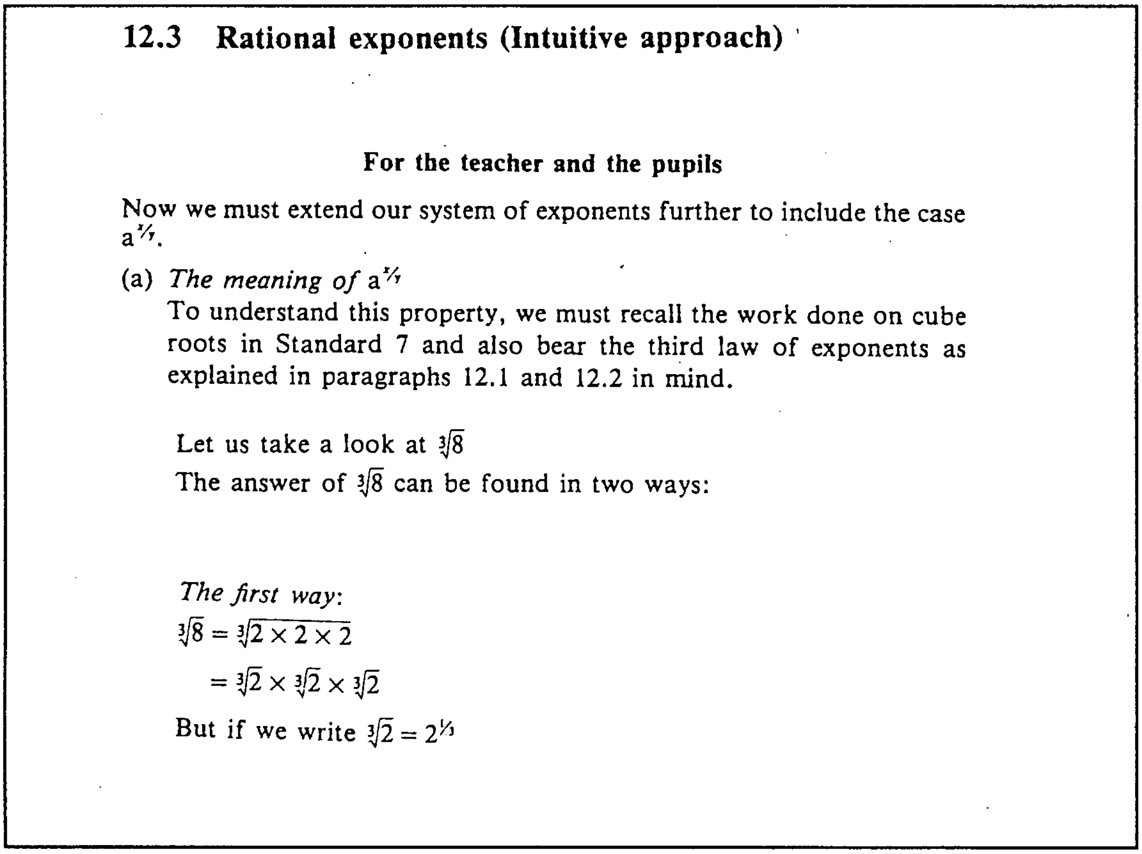
1 ▶ In this chapter the problems which apply this last deduction from the index laws all have $a > 0$. Can you think of a reason for this? Hint: the even power root of a negative number is non-real, for example:

$(-1)^{\frac{1}{3}} = \sqrt[3]{-1} = -1$

but $(-1)^{\frac{1}{3}} = (-1)^{\frac{2}{6}} = ((-1)^{\frac{1}{6}})^2 = (\sqrt[6]{-1})^2$ *non-real*

There is a difference between $(-1)^{\frac{1}{3}}$ and $(\sqrt[6]{-1})^2$, so they cannot be equated as has been done in this extract. From the definition $a^{p/q} = \sqrt[q]{a^p}$, $(-1)^{\frac{2}{6}} = \sqrt[6]{(-1)^2}$ and is therefore equal to 1. The authors have recognised that for the index laws to hold the bases have to be positive, but the reason given as shown in this extract is not correct. The argument shown here should rather lead to the contradiction that $-1 = 1$.

EX.M 8, pg. 291, writes about an "intuitive" approach towards rational exponents as follows:



12.3 Rational exponents (Intuitive approach)

For the teacher and the pupils

Now we must extend our system of exponents further to include the case $a^{1/3}$.

(a) *The meaning of* $a^{1/3}$
To understand this property, we must recall the work done on cube roots in Standard 7 and also bear the third law of exponents as explained in paragraphs 12.1 and 12.2 in mind.

Let us take a look at $\sqrt[3]{8}$
The answer of $\sqrt[3]{8}$ can be found in two ways:

The first way:
$\sqrt[3]{8} = \sqrt[3]{2 \times 2 \times 2}$
$= \sqrt[3]{2} \times \sqrt[3]{2} \times \sqrt[3]{2}$
But if we write $\sqrt[3]{2} = 2^{1/3}$

then $\sqrt[3]{8} = 2^{1/3} \times 2^{1/3} \times 2^{1/3}$

$= 2^{1/3+1/3+1/3}$

$= 2^1$

$\therefore \quad \sqrt[3]{8} = 2$

The second way:

$\sqrt[3]{8} = \sqrt[3]{2 \times 2 \times 2} = \sqrt[3]{2^3}$

But if we write $\sqrt[3]{2^3} = (2^3)^{1/3}$

then $\sqrt[3]{8} = (2^3)^{1/3}$

$= 2^{3 \times 1/3}$

$= 2^1$

$= 2$

The answer is the same in both cases:

$\therefore \quad \sqrt[3]{8} = \sqrt[3]{2^3} = (2^3)^{1/3}$

The extract refers to $a^{\frac{x}{y}}$ as a **property** (which it is not) and then asks for "the answer of $\sqrt[3]{8}$" when there is no question. These sentences need to be rephrased. The methods used to show that $\sqrt[3]{8}$ has value equal to 2, is circular. The value of $\sqrt[3]{8}$ is established, by definition of cube roots, i.e. $\sqrt[3]{8}$ is defined to be the number which is cubed to give 8. The first line in the "proof" already establishes this fact, i.e. that $\sqrt[3]{8} = 2$ since it uses the fact that $8 = 2 \times 2 \times 2$. This whole section is to find out what the meaning of $a^{\frac{x}{y}}$ is, but it is using $2^{\frac{1}{3}}$ to show the meaning of $\sqrt[3]{8}$.

The text continues with the following:

Remember: 2^3 means that 2 must be taken as a factor three times. Likewise, in $8^{1/3}$ the exponent $\frac{1}{3}$ means that one-third of the number of factors of 8 must be calculated.

$\therefore \quad 8^{1/3} = (2 \times 2 \times 2)^{1/3}$

$\frac{1}{3}$ of the number of factors is 2

$\therefore \quad 8^{1/3} = 2$

The text writes that "in $8^{\frac{1}{3}}$ the exponent $\frac{1}{3}$ means that one-third of the number of factors of 8 must be calculated", which is confusing. There are four factors of 8, namely 1,2,4 and 8, so according to this statement the value that is sought after is $\frac{4}{3}$ .The definition of $2^{\frac{1}{3}}$ would be confusing in this light.

> Similarly,
>
> $$\sqrt[3]{125} = 125^{1/3} = (5^3)^{1/3} = 5$$
> $$\sqrt[5]{32} = 32^{1/5} = (2^5)^{1/5} = 2$$
> $$\sqrt[3]{16} = 16^{1/3} = (2^4)^{1/3} = 2^{4/3}$$
> $$\sqrt[3]{x^6} = x^{6/3} = x^2$$
> $$\sqrt[4]{a^7} = a^{7/4}$$
>
> Thus, in order to find the actual exponent of "a" in $\sqrt[4]{a^7}$, a quarter of of the exponents of "a", i.e. $\frac{1}{4}$ of 7, must be calculated.
>
> Likewise, $a^{7/4} = \sqrt[4]{a^7}$
>
> or $x^{2/5} = \sqrt[5]{x^2}$

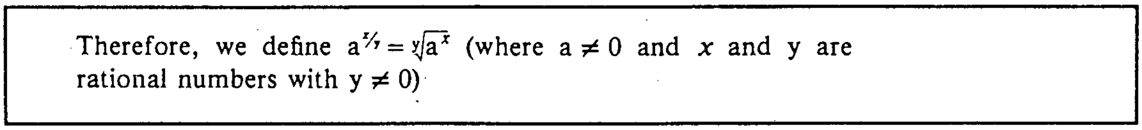
The fourth last and third last lines of this extract are confusing when it is written that "a quarter of the exponents of 'a' must be calculated". This seems to imply that a has more than one exponent.

This extract gives no clarification on the meaning of rational exponents. The fact used in the argument that $2^{\frac{1}{3}} \times 2^{\frac{1}{3}} \times 2^{\frac{1}{3}} = 2$ or $(2^{\frac{1}{3}})^3 = 2$ stems from the definition of the cube root of 8. The concept of $2^{\frac{1}{3}}$ is less fundamental than $8^{\frac{1}{3}}$ and therefore the definition of $2^{\frac{1}{3}}$ comes after defining the cube root of 8, and not the other way around as shown in this text (which was to define $\sqrt[3]{8}$ using $2^{\frac{1}{3}}$).

Tickey de Jager , in his book "More Than Just Maths"[11], says "$a^o, a^{-n}, a^{\frac{3}{4}}$ have no meaning; we are going to give them artificial meanings, but these artificial meanings must be such that they agree with the rest of algebra, that is, so that they obey the laws which we established for positive integral indices."

From this it is clear that a set of definitions should be constructed which are then shown to obey the laws of algebra. This would alleviate any difficulties in trying to describe what is meant by $8^{\frac{1}{3}}$ in similar terms that describe the meaning of 2^3 .

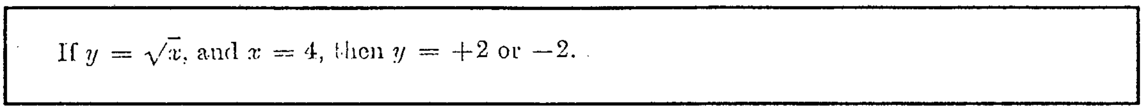
After the paragraph, as shown in the extract above, EX.M 8, pg. 293, gives a definition as follows:



Therefore, we define $a^{x/y} = \sqrt[y]{a^x}$ (where $a \neq 0$ and x and y are rational numbers with $y \neq 0$)

Here $a \geq 0$ is required, and x and y should be integers.

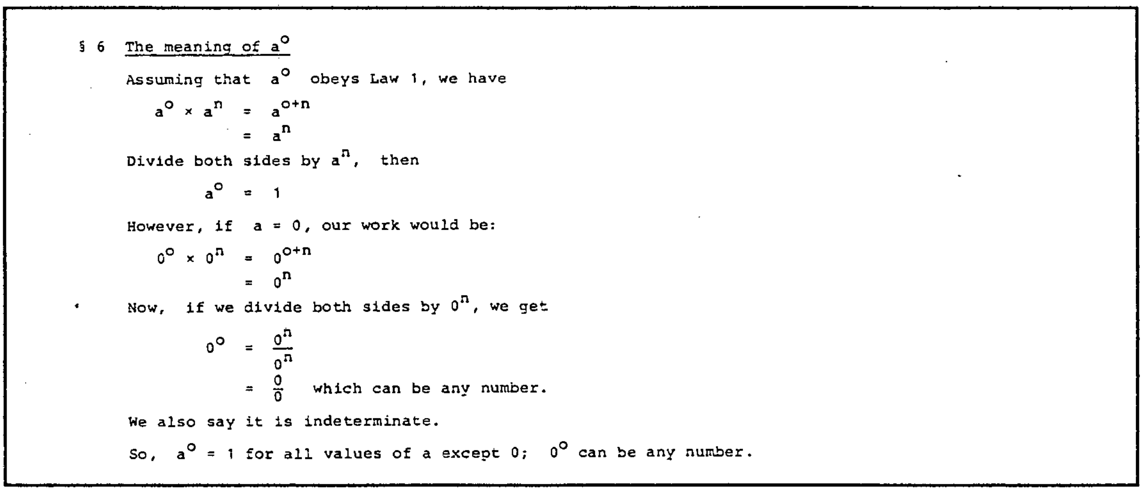
MAT , pg. 98, writes the following:



If $y = \sqrt{x}$, and $x = 4$, then $y = +2$ or -2.

The meaning of $\sqrt{a}$ is taken to be the positive square root . Any positive number has two square roots, i.e. two solutions to the equation $x^2 = a$ ($a \geq 0$) but the symbol $\sqrt{a}$ represents the positive square root.

As shown at the beginning of this chapter, CM 9 gives meaning to a^0, whereas JM 9&10, pg.67, assumes that the laws established for positive integer exponents hold and then proves that $a^0 = 1$ as follows:



§ 6 The meaning of a^0

Assuming that a^0 obeys Law 1, we have

$$a^0 \times a^n = a^{0+n}$$
$$= a^n$$

Divide both sides by a^n, then

$$a^0 = 1$$

However, if $a = 0$, our work would be:

$$0^0 \times 0^n = 0^{0+n}$$
$$= 0^n$$

Now, if we divide both sides by 0^n, we get

$$0^0 = \frac{0^n}{0^n}$$
$$= \frac{0}{0} \quad \text{which can be any number.}$$

We also say it is indeterminate.

So, $a^0 = 1$ for all values of a except 0; 0^0 can be any number.

The text tries to show what is meant by 0^0, by following the same steps used to show that $a^0 = 1$. This is invalid because the steps used in the "proof " are only valid for non-zero bases, otherwise division by zero would occur and that is undefined. Division by zero is undefined (this is discussed later in chapter 8).

7. LOGARITHMS

1. Introduction

Historically logarithms played a role in making calculations involving difficult multiplication and division easier to solve. Nowadays calculators make multiplication and division quick and easy and have replaced the need for logs.

The approach to the theory of logarithms changed a great deal after calculators were available and made part of the school syllabus. After this, the computations using logarithms were easier and more accurate than those that relied on logarithmic tables. This can be seen in texts that used log tables and where inaccurate answers were produced.

In this section a brief look will be taken of the uses of logarithms and some other errors concerning logarithms.

2. Logarithms

The theory of logarithms enables one to work with exponents that are not "nice", as in whole numbers (even rational number exponents are still considered "nice").

When using logarithms the base 10 is frequently used. MGM 8, pg. 180, says the following about using base 10:

> **13.3.2 Expressing numbers as powers of 10**
> Any number can be expressed as a power of 10, i.e. any number $= 10^x$ where x can be any rational number.

Given any real number, it cannot be expressed as 10^x, where x is a **rational** number. This is claiming that all logarithms are rational, which is not true. For example if $x = \log_{10} 5$ i.e. $5 = 10^x$ then x is irrational. This is not trivial and requires a small proof:

Suppose $5 = 10^x$ where x is rational

i.e. $x = \dfrac{p}{q}$ where p and q are natural numbers ($p > 0$ and $q > 0$ since $5 > 1$)

This implies that $5 = 10^{\frac{p}{q}}$

and therefore $5^q = 10^p$

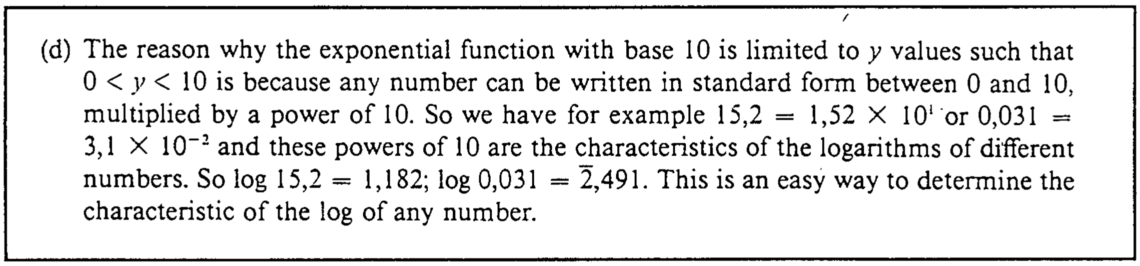
The expansion of 5^q ends with the digit 5 and the expansion of 10^p ends with the digit zero. Therefore $5^q \neq 10^p$. Therefore we get a contradiction and x cannot be rational.

The claim that "any number can be expressed as a power of 10" is incorrect because negative numbers cannot be expressed in this way. To say any number = 10^x where x can be **any** number is also incorrect. This implies that a range of numbers will satisfy this equation. It should say "where x is **some** real number".

The idea that logarithms are needed in order to make multiplication and division easier is obsolete. However modern texts kept this idea in mind. The following is a text book published in 1985, covering the 1986 syllabus. Calculators were available and were to be used in the new syllabus, replacing the need for log tables.

The influence of the old syllabus crept into this text. The "characteristic" of a logarithm is referred to and related notation is used, without any definitions or explanations.

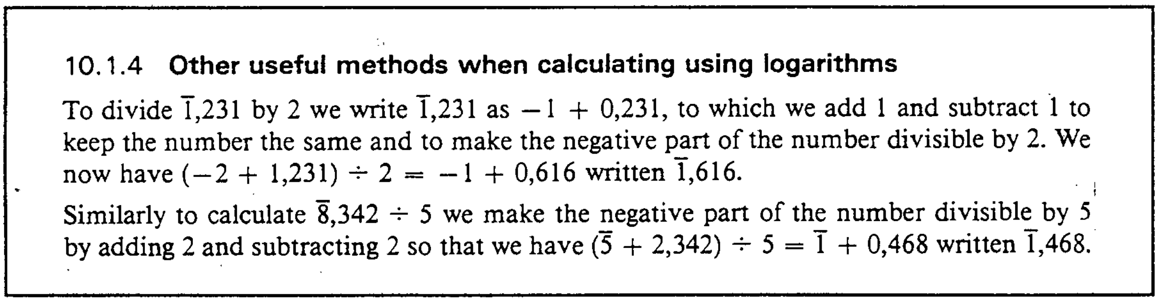
UM 9&10, pg. 91 says the following:



(d) The reason why the exponential function with base 10 is limited to y values such that $0 < y < 10$ is because any number can be written in standard form between 0 and 10, multiplied by a power of 10. So we have for example $15,2 = 1,52 \times 10^1$ or $0,031 = 3,1 \times 10^{-2}$ and these powers of 10 are the characteristics of the logarithms of different numbers. So $\log 15,2 = 1,182$; $\log 0,031 = \bar{2},491$. This is an easy way to determine the characteristic of the log of any number.

Note also that in the second line it should read, "any number can be written in the standard form of a number between 0 and 10 multiplied by a power of 10". This idea of characteristic and the "bar one" notation appears again on pg. 94, but has little use in the modern syllabus.

UM 9&10, pg. 94, says the following:

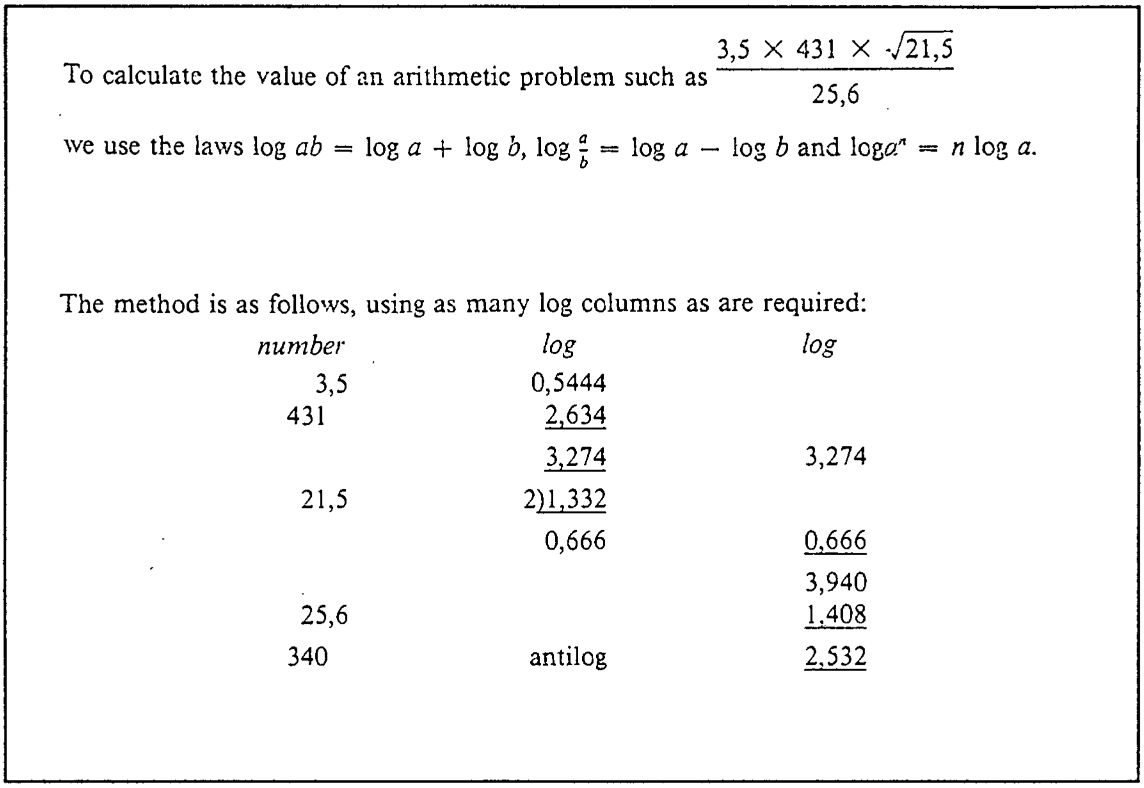


10.1.4 Other useful methods when calculating using logarithms

To divide $\bar{1},231$ by 2 we write $\bar{1},231$ as $-1 + 0,231$, to which we add 1 and subtract 1 to keep the number the same and to make the negative part of the number divisible by 2. We now have $(-2 + 1,231) \div 2 = -1 + 0,616$ written $\bar{1},616$.

Similarly to calculate $\bar{8},342 \div 5$ we make the negative part of the number divisible by 5 by adding 2 and subtracting 2 so that we have $(\bar{5} + 2,342) \div 5 = \bar{1} + 0,468$ written $\bar{1},468$.

UM 9&10, pg. 93, says:



To calculate the value of an arithmetic problem such as $\dfrac{3{,}5 \times 431 \times \sqrt{21{,}5}}{25{,}6}$

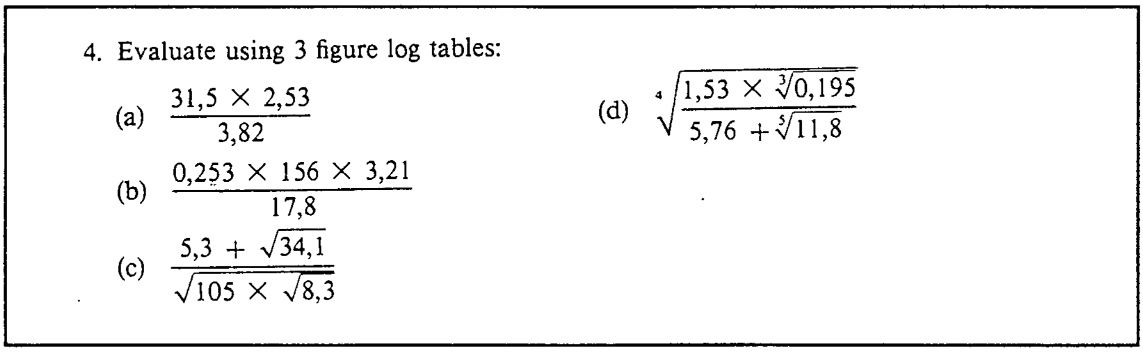
we use the laws $\log ab = \log a + \log b$, $\log \frac{a}{b} = \log a - \log b$ and $\log a^n = n \log a$.

The method is as follows, using as many log columns as are required:

number	*log*	*log*
3,5	0,5444	
431	2,634	
	3,274	3,274
21,5	2)1,332	
	0,666	0,666
		3,940
25,6		1,408
340	antilog	2,532

Logarithms are being used in an attempt to simplify multiplication and division. The operations that have been carried out are not clearly shown and it is confusing to follow even though the method is written out step by step. There are also calculation errors. The figure 0.5444 is probably a misprint for 0.544, and the sum of 0.544 and 2.634 is 3.178. The correct final answer is 273.2

The inaccuracy resulting from using 3-figure log tables is well illustrated in UM 9&10, pg. 96, Ex 4.



4. Evaluate using 3 figure log tables:

(a) $\dfrac{31{,}5 \times 2{,}53}{3{,}82}$

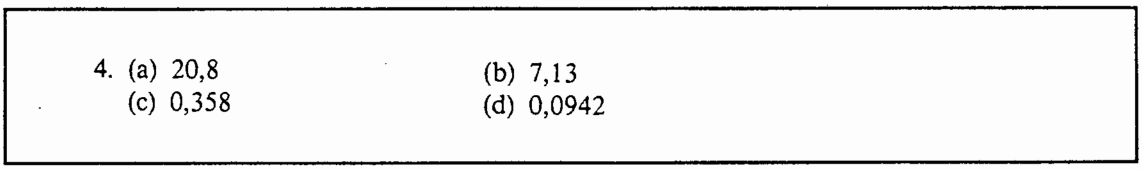
(b) $\dfrac{0{,}253 \times 156 \times 3{,}21}{17{,}8}$

(c) $\dfrac{5{,}3 + \sqrt{34{,}1}}{\sqrt{105} \times \sqrt{8{,}3}}$

(d) $\sqrt[4]{\dfrac{1{,}53 \times \sqrt[3]{0{,}195}}{5{,}76 + \sqrt[5]{11{,}8}}}$

The answers, using a calculator and rounding to three significant figures, are (a) 20,9; (b) 7,12; (c) 0,640 and (d) 0,588.

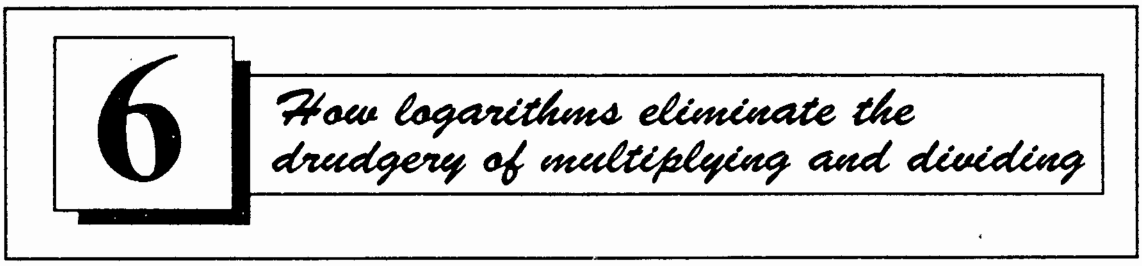
The answers in the back of the book, pg.330, are:



4. (a) 20,8 (b) 7,13
(c) 0,358 (d) 0,0942

N1 is a text published in 1994. A whole chapter is devoted to showing how logarithms could be used in calculations.

N1, pg. 30, starts the chapter with the heading:



6
How logarithms eliminate the drudgery of multiplying and dividing

A great deal of drudgery is devoted to trying to show that logarithms make calculations involving multiplication, division, powers of numbers and root extraction easier. There are no log tables in the book. The author advises readers to use a calculator to determine the logarithm of a number.

8. ALGEBRA

1. Introduction

It is interesting to note that in high school texts there are considerably more errors in the chapters on abstract algebra than in those on geometry. Apart from errors such as confusing the concept of converse, as seen earlier, geometry is fairly free of flaws. This may be attributed to the fact that at school level, geometry has been studied for much longer than topics in abstract algebra. Over the years sections on geometry have been set and unchanging, while algebra has varied with syllabus changes.

Texts of the early "New Maths" era included the topic of abstract algebra. With this inclusion came discussion of group theory, rings and fields. These topics have since been abandoned, but traces remain. Terminology from the theory of groups and fields has filtered through to modern texts. This is evident in sections which discuss the operations of multiplication, division, addition and subtraction.

Other topics under the heading of algebra, such as mathematical induction, were also introduced during the "New Maths" era. Mathematical induction is no longer in the syllabus. Extracts show that mathematical induction was not fully understood.

In this section the following algebraic topics will be discussed:

- group theory, rings and fields
- mathematical induction
- sequences
- algebraic variables and absolute values

2. Groups, rings and fields

In order to discuss the various errors that texts have concerning these concepts definitions need to be given.

A **group**, $\langle G,*\rangle$, is a non-empty set G with a binary operation $*$, such that:

1. $\forall\, a,b \in G \quad a*b \in G$ (closure)
2. $\forall\, a,b,c \quad a*(b*c) = (a*b)*c$ (associative law)
3. $\exists\, e \in G$ such that $\forall\, a \in G \quad e*a = a*e = a$ (e is the identity of G)
4. $\forall\, a \in G$ there exists an element in G (denoted by a^{-1}) such that $a^{-1}*a = a*a^{-1} = e$ (inverses)

An **abelian** group has the added criterion of commutativity: $a*b = b*a$ for all a, b in G.

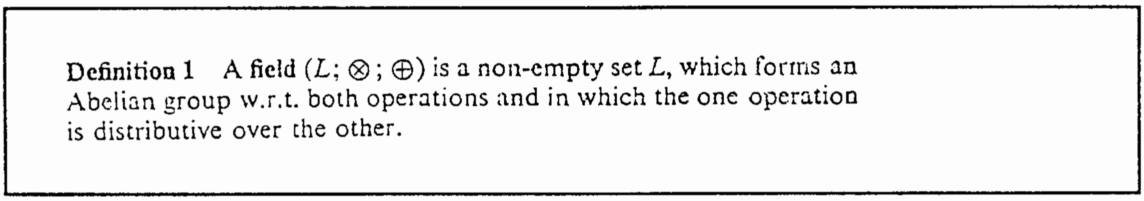
A **ring**, $\langle R,\oplus,\otimes\rangle$, is a non-empty set R together with two binary operations $\oplus$ and $\otimes$, such that:

1. $\langle R,\oplus\rangle$ is an abelian group, with the identity denoted by z
2. $\forall\, a,b,c \in R .\; a\otimes(b\oplus c) = (a\otimes b) \quad \oplus \quad (a\otimes c)$
   (distributivity of $\otimes$ over $\oplus$)

A **field** F is a non-empty set such that:

1. $\langle F,\oplus,\otimes\rangle$ is a ring
2. $\langle F\setminus\{z\},\otimes\rangle$ is an abelian group (where z is the identity of $\oplus$).

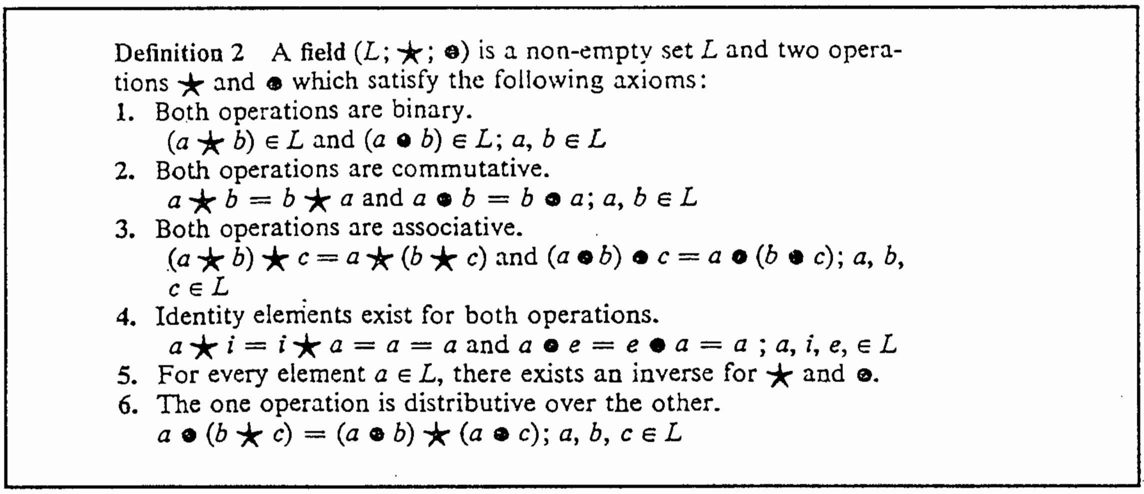
SSM 10, pg. 64, says the following about fields:



Definition 1 A field $(L;\otimes;\oplus)$ is a non-empty set L, which forms an Abelian group w.r.t. both operations and in which the one operation is distributive over the other.

The above statement is incorrect because it is essential to specify that with respect to the one operation, usually represented by $\otimes$, $\langle L,\otimes\rangle$ does not form a group, but $\langle L\setminus\{z\},\otimes\rangle$ (where z is the identity element with respect to $\oplus$) forms a group. This is because z does not have an inverse with respect to $\otimes$. The statement about distributivity is vague and should specify that $\otimes$ distributes over $\oplus$. (Note: $\oplus$ does not distribute over $\otimes$.)

SSM 10, pg. 64, has a second definition:



Definition 2 A field $(L;\bigstar;\bullet)$ is a non-empty set L and two operations $\bigstar$ and $\bullet$ which satisfy the following axioms:

1. Both operations are binary.
   $(a\bigstar b)\in L$ and $(a\bullet b)\in L$; $a, b\in L$
2. Both operations are commutative.
   $a\bigstar b = b\bigstar a$ and $a\bullet b = b\bullet a$; $a, b\in L$
3. Both operations are associative.
   $(a\bigstar b)\bigstar c = a\bigstar(b\bigstar c)$ and $(a\bullet b)\bullet c = a\bullet(b\bullet c)$; $a, b, c\in L$
4. Identity elements exist for both operations.
   $a\bigstar i = i\bigstar a = a = a$ and $a\bullet e = e\bullet a = a$; $a, i, e, \in L$
5. For every element $a\in L$, there exists an inverse for $\bigstar$ and $\bullet$.
6. The one operation is distributive over the other.
   $a\bullet(b\bigstar c) = (a\bullet b)\bigstar(a\bullet c)$; $a, b, c\in L$

In this extract there are logical quantifiers that have been omitted. In points 1,2 and 3 a "for all" statement or quantifier is required. The statement in point 4 should be replaced by "$\exists\, i, e$ such that $\forall\, a \in L \quad a*i = i*a = a$ and $a \bullet e = e \bullet a = a$". Point 5 repeats the error that $\langle L, \bullet \rangle$ is a group, since i does not have an inverse with respect to "$\bullet$". Point 6, the distributivity of "one operation over the other", is again vague. Only one of the operations is distributive over the other. In this particular case the second operation is distributive over the first.

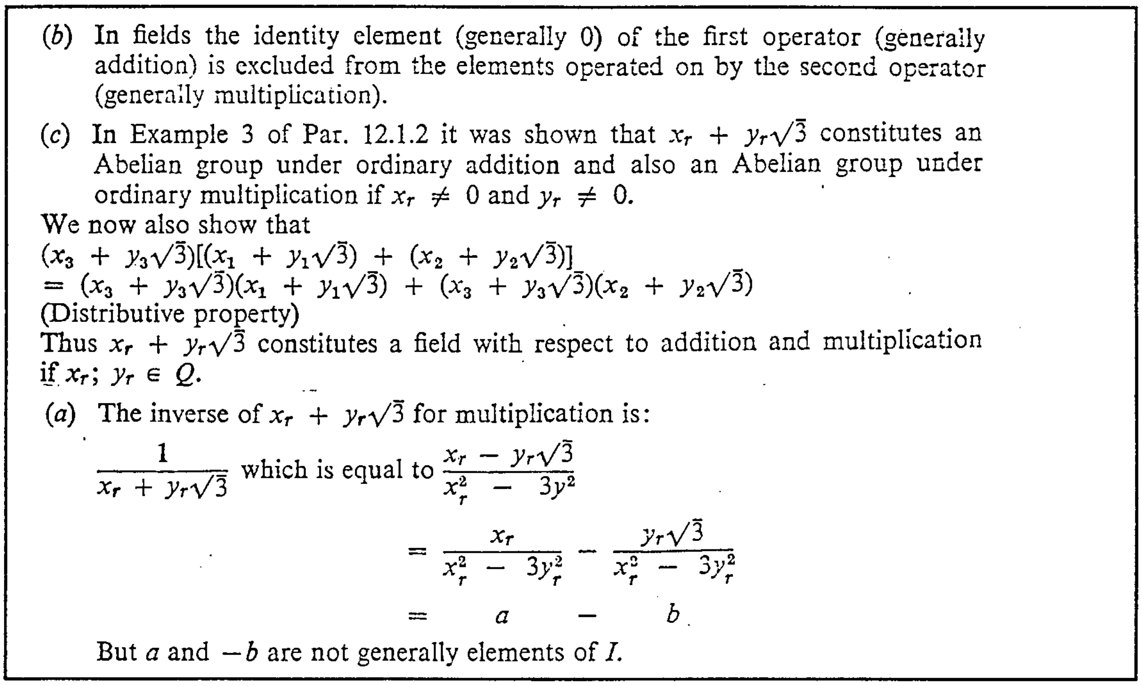
MGM 9&10, pg. 205, says the following:

> $F = \{a;\ b;\ c;\ \ldots\}$ **constitutes a field if:**
>
> **1.** The *closure property* holds for *addition* and *multiplication*,
> i.e. if $(a \oplus b) \in F$ and $a \otimes b \in F$
>
> **2.** *The operators are associative in F,*
> i.e. if $(a \oplus b) \oplus c = a \oplus (b \oplus c)$
> and $a \otimes (b \otimes c) = (a \otimes b) \otimes c.$
>
> **3.** *There exist identity elements i and e* such that
> $a \oplus i = a$
> and $a \otimes e = a$
> Very often $i = 0$ and $e = 1.$
>
> **4.** *Inverse elements exist for each $a \in F$*
> i.e. $a \oplus \tilde{a} = i$ and very often $\tilde{a} = -a$
> and $a \otimes \acute{a} = e$ and very often $a' = \dfrac{1}{a}$
> The notations $\tilde{a}$ and a' denote inverses of a.
>
> **5.** *The commutative property holds for both operations,*
> i.e. $a \oplus b = b \oplus a$
> and $a \otimes b = b \otimes a$
>
> We note that so far a field constitutes Abelian groups for both $\oplus$ and $\otimes$.
>
> **6.** *The distributive property of the second operator over the first is satisfied,*
> i.e. $a \otimes (b \oplus c) = (a \otimes b) \oplus (a \otimes c).$

Apart from errors in the use of language and logical quantifiers, (discussed in chapter 2) this extract has the same error as SSM 10, above. Part 4 says that for every element of the set there are inverses with respect to **both** operations. The identity (with respect to the first operation) does not have an inverse with respect to the second operation.

The following example is taken from MGM 9&10, pg. 206:

> **Example 3**
>
> (*a*) Why does $I = \{\text{integers}\}$ not constitute a field with respect to ordinary addition and multiplication?
> (*b*) Why must a field have at least two elements?
> (*c*) Show that $x_r + y_r\sqrt{3}$ constitutes a field under ordinary addition and multiplication if $x_r;\ y_r \in Q$ where $Q = \{\text{rational numbers}\}$
> (*d*) Why does $x_r + y_r\sqrt{3}$ not constitute a field with respect to ordinary addition and multiplication if $x_r;\ y_r \in I$ where $I = \{\text{integers}\}$?
>
> **Discussion**
>
> (*a*) -1 and $+1$ are the only integers which have inverses under multiplication. Thus $I = \{\text{integers}\}$ is not a field.



(*b*) In fields the identity element (generally 0) of the first operator (generally addition) is excluded from the elements operated on by the second operator (generally multiplication).

(*c*) In Example 3 of Par. 12.1.2 it was shown that $x_r + y_r\sqrt{3}$ constitutes an Abelian group under ordinary addition and also an Abelian group under ordinary multiplication if $x_r \neq 0$ and $y_r \neq 0$.

We now also show that

$(x_3 + y_3\sqrt{3})[(x_1 + y_1\sqrt{3}) + (x_2 + y_2\sqrt{3})]$
$= (x_3 + y_3\sqrt{3})(x_1 + y_1\sqrt{3}) + (x_3 + y_3\sqrt{3})(x_2 + y_2\sqrt{3})$

(Distributive property)

Thus $x_r + y_r\sqrt{3}$ constitutes a field with respect to addition and multiplication if x_r; $y_r \in Q$.

(*a*) The inverse of $x_r + y_r\sqrt{3}$ for multiplication is:

$$\frac{1}{x_r + y_r\sqrt{3}} \text{ which is equal to } \frac{x_r - y_r\sqrt{3}}{x_r^2 - 3y^2}$$

$$= \frac{x_r}{x_r^2 - 3y_r^2} - \frac{y_r\sqrt{3}}{x_r^2 - 3y_r^2}$$

$$= \quad a \quad - \quad b.$$

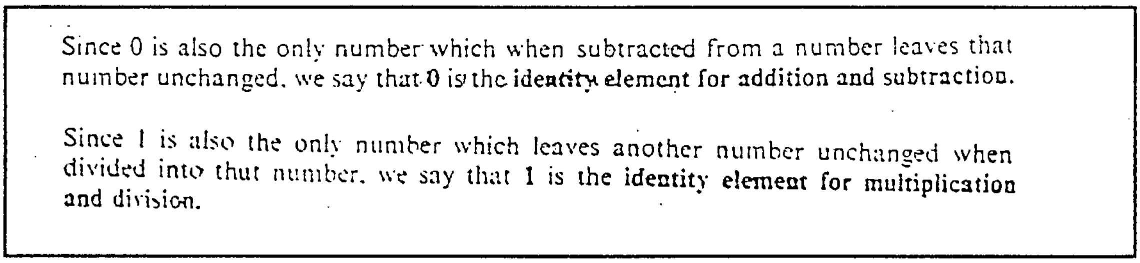
But a and $-b$ are not generally elements of I.

Though it is true that $F = \{0\}$ with the usual rules of addition and multiplication is a field it is a degenerate and uninteresting case. Therefore, while fields that are studied have two or more elements it is not a necessity.

In part c of the "discussion" it is not necessary to have $x_r \neq 0$ **and** $y_r \neq 0$, but rather $x_r \neq 0$ **or** $y_r \neq 0$.

In the last part of the discussion (labelled part "a" where it should be part "d") the second last line should read as $a - \sqrt{3}b$. This particular discussion would be better illustrated with a specific example to show that $\{x_r + y_r\sqrt{3} \mid x_r, y_r \in I\}$ with the usual multiplication, does not form a group. For example $0 + 1\sqrt{3}$ does not have a multiplicative inverse.

MGM 6, pg. 11, writes the following about the identities 0 and 1 for real numbers with respect to addition and multiplication:

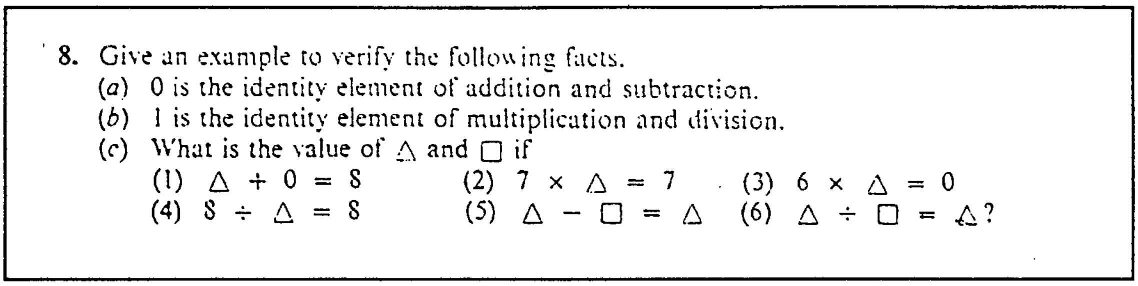


Since 0 is also the only number which when subtracted from a number leaves that number unchanged, we say that 0 is the identity element for addition and subtraction.

Since 1 is also the only number which leaves another number unchanged when divided into that number, we say that 1 is the identity element for multiplication and division.

For an element of a set *X*, say *e*, to be the **identity** of an operation $*$ the following must hold: $e*x = x*e = x$ for all elements *x* of *X*. For the operation of subtraction this is not true, i.e. $b - 0 \neq 0 - b \neq b$, so zero is not the identity of subtraction. Similarly the element 1 is not the identity of division since $x \div 1 \neq 1 \div x \neq x$.

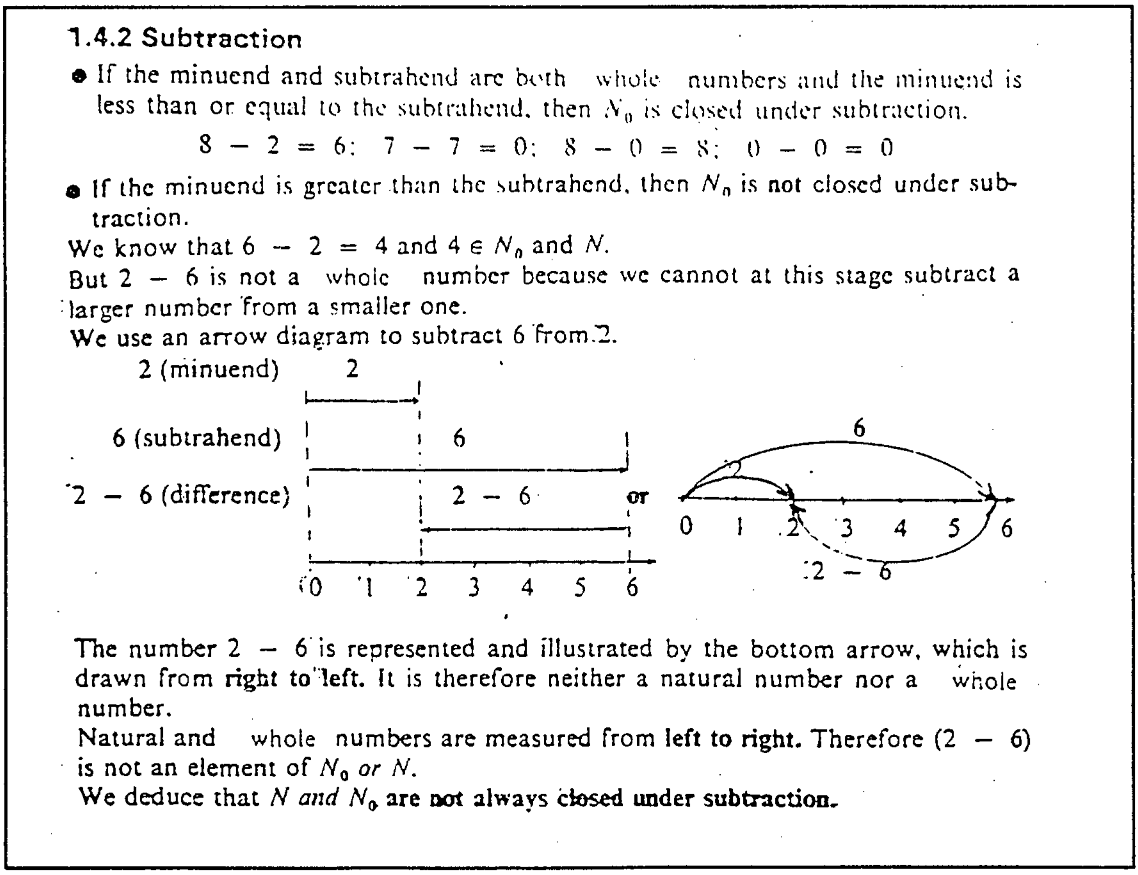
MGM 6, pg. 12, gives the following example that illustrates the above confusion again:



8. Give an example to verify the following facts.
   (*a*) 0 is the identity element of addition and subtraction.
   (*b*) 1 is the identity element of multiplication and division.
   (*c*) What is the value of △ and □ if
      (1) △ + 0 = 8 (2) 7 × △ = 7 (3) 6 × △ = 0
      (4) 8 ÷ △ = 8 (5) △ − □ = △ (6) △ ÷ □ = △?

One could say that 0 and 1 are right identities for subtraction and division respectively, but this is not particularly useful.

MGM 6, pg. 13, says the following about closure of sets under operations:



1.4.2 Subtraction

- If the minuend and subtrahend are both whole numbers and the minuend is less than or equal to the subtrahend, then N_0 is closed under subtraction.

  8 − 2 = 6; 7 − 7 = 0; 8 − 0 = 8; 0 − 0 = 0

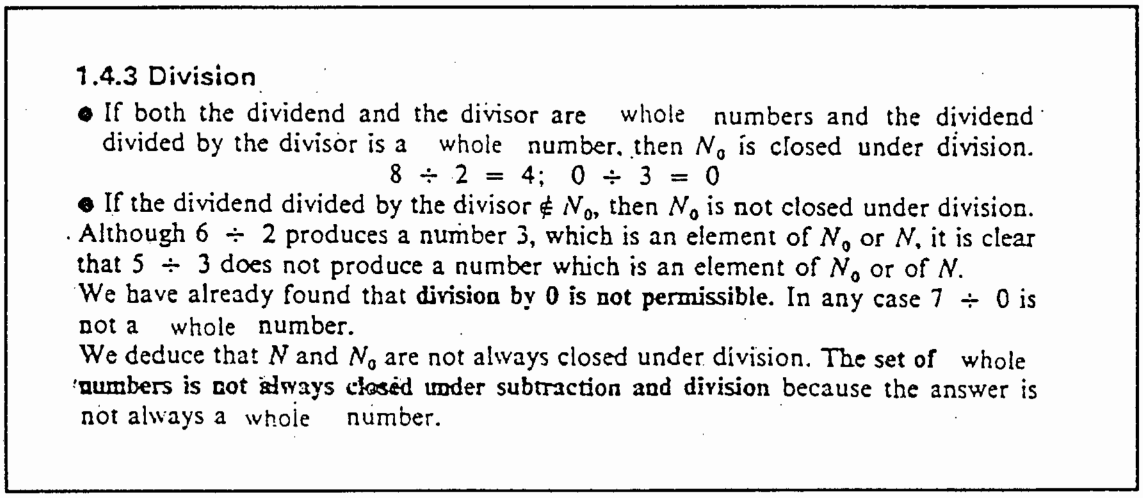
- If the minuend is greater than the subtrahend, then N_0 is not closed under subtraction.

We know that 6 − 2 = 4 and 4 ∈ N_0 and *N*.
But 2 − 6 is not a whole number because we cannot at this stage subtract a larger number from a smaller one.
We use an arrow diagram to subtract 6 from 2.

The number 2 − 6 is represented and illustrated by the bottom arrow, which is drawn from right to left. It is therefore neither a natural number nor a whole number.
Natural and whole numbers are measured from **left to right.** Therefore (2 − 6) is not an element of N_0 *or N*.
We deduce that *N and* N_0 **are not always closed under subtraction.**

For a set X to be closed under an operation $*$ it must be true that $\forall\, a,b \in X \quad a*b \in X$. There is no question of **sometimes** being closed under the operation.

Again on pg. 13 the text makes the same error concerning division and closure.



1.4.3 Division

- If both the dividend and the divisor are whole numbers and the dividend divided by the divisor is a whole number, then N_0 is closed under division.

$$8 \div 2 = 4; \quad 0 \div 3 = 0$$

- If the dividend divided by the divisor $\notin N_0$, then N_0 is not closed under division.

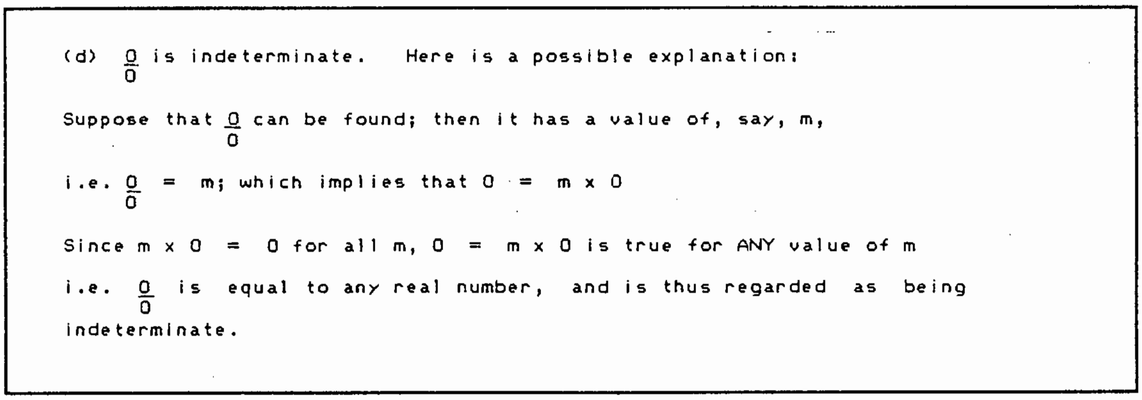
Although $6 \div 2$ produces a number 3, which is an element of N_0 or N, it is clear that $5 \div 3$ does not produce a number which is an element of N_0 or of N.

We have already found that **division by 0 is not permissible.** In any case $7 \div 0$ is not a whole number.

We deduce that N and N_0 are not always closed under division. **The set of** whole **numbers is not always closed under subtraction and division** because the answer is not always a whole number.

The reason why division by zero is undefined is an algebraic topic and several texts consider $\frac{0}{0}$ to be **any** real number.

MRS, pg. 7, says the following:



(d) $\frac{0}{0}$ is indeterminate. Here is a possible explanation:

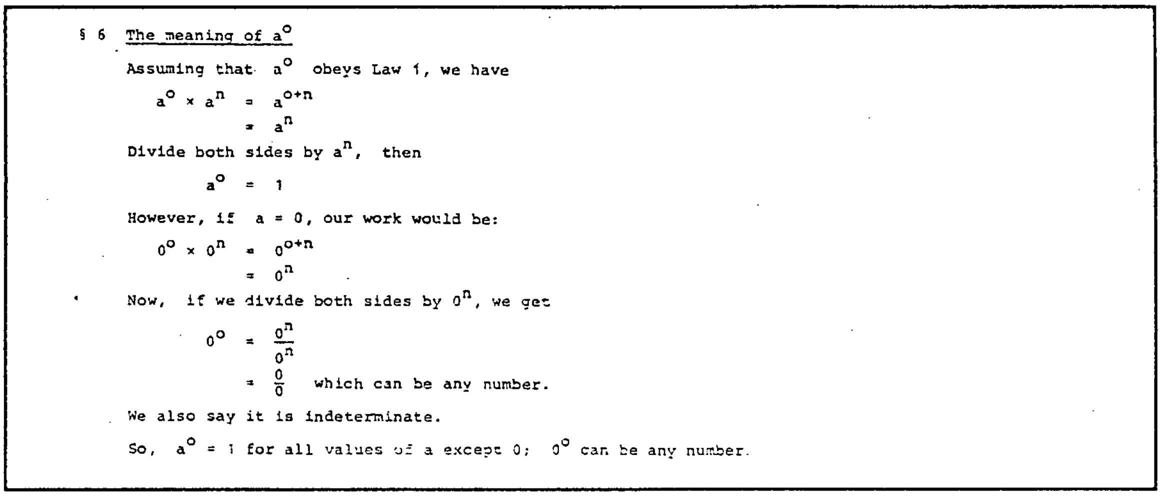
Suppose that $\frac{0}{0}$ can be found; then it has a value of, say, m,

i.e. $\frac{0}{0} = m$; which implies that $0 = m \times 0$

Since $m \times 0 = 0$ for all m, $0 = m \times 0$ is true for ANY value of m

i.e. $\frac{0}{0}$ is equal to any real number, and is thus regarded as being indeterminate.

JM 9&10, pg. 67, says:



§ 6 The meaning of a^o

Assuming that a^o obeys Law 1, we have

$$a^o \times a^n = a^{o+n}$$
$$= a^n$$

Divide both sides by a^n, then

$$a^o = 1$$

However, if $a = 0$, our work would be:

$$0^o \times 0^n = 0^{o+n}$$
$$= 0^n$$

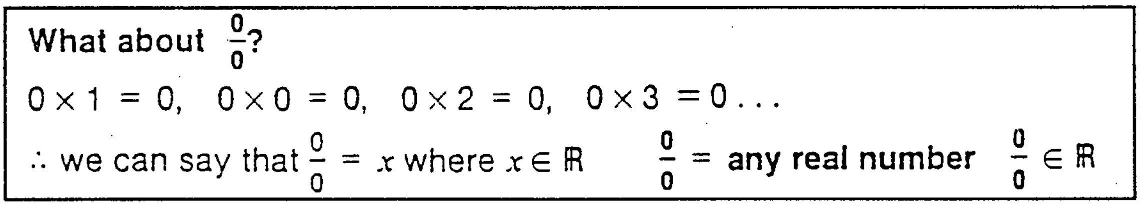
Now, if we divide both sides by 0^n, we get

$$0^o = \frac{0^n}{0^n}$$
$$= \frac{0}{0} \text{ which can be any number.}$$

We also say it is indeterminate.

So, $a^o = 1$ for all values of a except 0; 0^o can be any number.

SM, pg. 36, says:



What about $\frac{0}{0}$?

$0 \times 1 = 0, \quad 0 \times 0 = 0, \quad 0 \times 2 = 0, \quad 0 \times 3 = 0 \ldots$

∴ we can say that $\frac{0}{0} = x$ where $x \in \mathbb{R}$ $\quad$ $\frac{0}{0}$ = **any real number** $\quad$ $\frac{0}{0} \in \mathbb{R}$

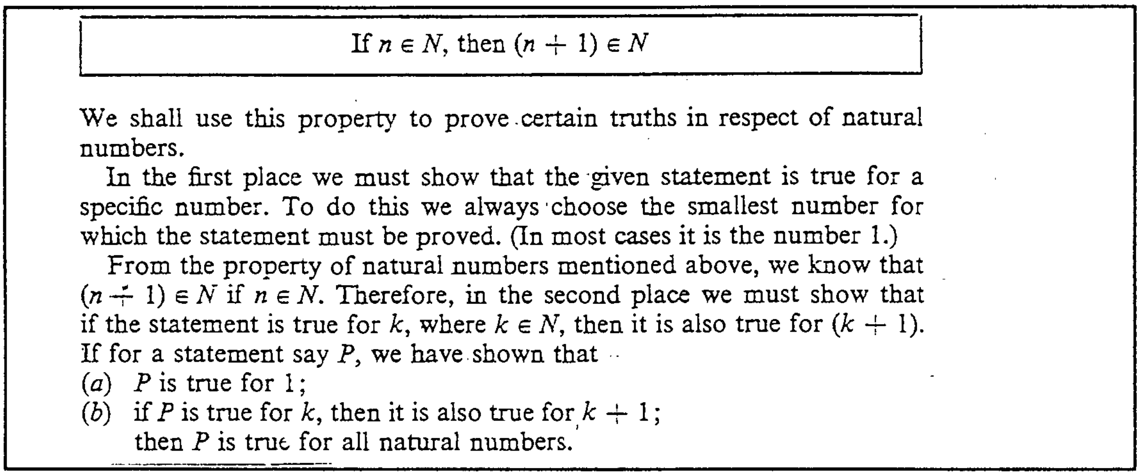
Division by zero is undefined, even if the numerator is also zero. The two cases have different reasons, but one arrives at the same conclusions. The division of two real numbers, say $A \div B$, is defined if there is a unique solution to the equation $A \cdot x = B$. For $A \neq 0$, $\frac{A}{0}$ is undefined because the solution to the equation $0 \cdot x = A$ does not exist. On the other hand $\frac{0}{0}$ is not defined because the equation $0 \cdot x = 0$ has infinitely many solutions. It is not true that $\frac{0}{0}$ is equal to any real number, as some texts have written. The equation $0 \cdot x = 0$ is indeterminate because it has infinitely many solutions. Therefore $\frac{0}{0}$ is often referred to as an indeterminate form.

3. Mathematical Induction

The principle of mathematical induction says:

Let S_n be a statement about the positive integer n. Suppose that
1. S_1 is true
2. S_{k+1} is true whenever S_k is true.
Then S_n is true for all integers n. [4]

SSM 10, pg. 25 says the following on mathematical induction:



If $n \in N$, then $(n + 1) \in N$

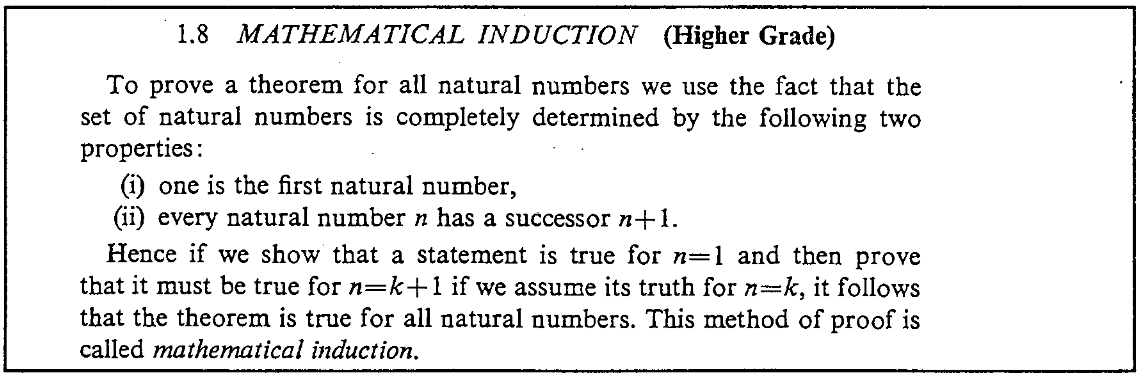
We shall use this property to prove certain truths in respect of natural numbers.

In the first place we must show that the given statement is true for a specific number. To do this we always choose the smallest number for which the statement must be proved. (In most cases it is the number 1.)

From the property of natural numbers mentioned above, we know that $(n + 1) \in N$ if $n \in N$. Therefore, in the second place we must show that if the statement is true for k, where $k \in N$, then it is also true for $(k + 1)$. If for a statement say P, we have shown that
(*a*) P is true for 1;
(*b*) if P is true for k, then it is also true for $k + 1$;
then P is true for all natural numbers.

The implication here is that the principle of mathematical induction is a consequence of the statement " If $n \in \mathbf{N}$ then $n + 1 \in \mathbf{N}$ ". This is the "successor axiom". The axiom of mathematical induction is independent of this successor axiom. An easy example to validate this independence are the real numbers. The real numbers satisfy the property "if $x \in \mathbf{R}$ then $x + 1 \in \mathbf{R}$ " but do not satisfy the axiom of mathematical induction.

MM 9&10, pg. 26, deals with mathematical induction in a similar way, as follows:



1.8 *MATHEMATICAL INDUCTION* **(Higher Grade)**

To prove a theorem for all natural numbers we use the fact that the set of natural numbers is completely determined by the following two properties:
(i) one is the first natural number,
(ii) every natural number n has a successor $n+1$.

Hence if we show that a statement is true for $n=1$ and then prove that it must be true for $n=k+1$ if we assume its truth for $n=k$, it follows that the theorem is true for all natural numbers. This method of proof is called *mathematical induction.*

4. Sequences

A sequence is a function from the natural numbers to another set. In particular a real-valued sequence is a function from the natural numbers to the real numbers. The value of a term in the sequence depends on the position of the term in the sequence and is not calculated according to predecessors using a fixed numerical process. Specific types of sequences do this, but in general it is not the case. For example the sequence 1,4,9,16,25... is such that each term is not calculated from its predecessor term, but according to its position in the sequence, i.e. $T_n = n^2$.

UM 9&10, pg. 97, says

> **11.1.1 Definitions**
>
> This is a set of numbers which follow each other in such a way that each number of the sequence can be determined from the preceding number according to a fixed numerical process. A simple sequence of numbers is the set of natural numbers. If one number is 3 then the next number is one more than 3. Therefore if a certain number of this sequence is n then the following number will be $n + 1$.

MRS, pg. 175, has the following similar statement:

> A Sequence or a Progression is an ORDERED set of numbers, where each term may be determined from its predecessor according to some rule or pattern.

MGM 9&10, pg. 323, says

> **17.1.1 Sequences**
>
> A sequence is a set of numbers which follow on one another in such a manner that each number of the sequence may be found from the preceding one by some fixed rule or law. The different numbers are known as the terms of the sequence.

These three extracts all make the same error in saying that a sequence is necessarily described by the fact that a term in a sequence is determined by its predecessor. This is true for arithmetic and geometric series. This error could perhaps have arisen in the texts because of generalising these two particular series.

MGM 9&10, pg. 323, continues by saying:

> **We may now deduce**
>
> 1. a finite (infinite) sequence consists of a finite (infinite) set of numbers,
> 2. this set is a one-to-one mapping according to a rule of the elements of the finite (infinite) set $\{1;\ 2;\ 3;\ \ldots;\ n\}$,
> 3. a sequence may be represented graphically.

In point "2", the authors seem to be attempting to define a sequence as a mapping. However it is not true that the mapping need be one-to-one. There are sequences that are constant or alternate between 1 and -1, for example, that are clearly not one-to-one mappings of the natural numbers. Point 2 is difficult to understand.

SSM 10, pg. 148, says the following:

> **1. A few concepts**
>
> If $A = \{12; 3; 21; 15; 6; 18; 9\}$ and $B = \{3; 6; 9; 12; 15; 18; 21\}$, then $A = B$. This means that A has exactly the same elements as B. However, there is still a difference between A and B. We say that the elements of B are **ordered**, because the set has a specific first element, second element, etc.
>
> A set with ordered elements is called a **sequence.**

This does not describe a sequence. There is a rule associated with a sequence that requires a definition in order to be able to calculate the value of a term at a particular position in the sequence. A mere ordering of a collection of elements does not constitute a sequence.

MRS, pg. 193, makes the following error concerning sequential notation.

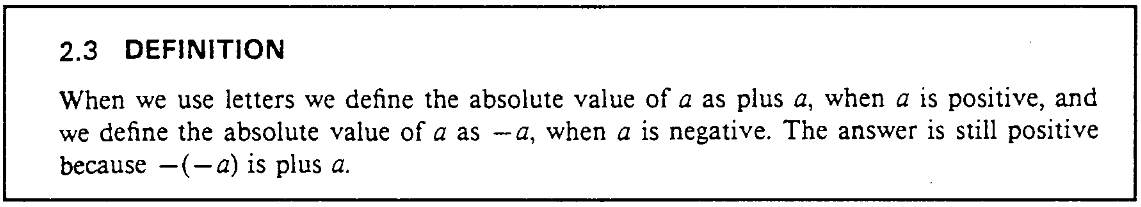
> However if $n \to \infty$, then
>
> $(0,1)^{\infty} \to$ 0,00000000000000000..........an infinite number of zeroes before we can place the one. We never reach this position!
>
> i.e. $(0,1)^{\infty} \to 0$
>
> i.e. If $-1 < r < 1$, then $r^{\infty} \to 0$

The correct use of notation is: $(0{,}1)^n \to 0$ as $n \to \infty$.

5. Algebraic Variables and Absolute Values

There appear to be difficulties concerning when a variable is positive or negative. Most high school pupils have the misconception that the variable $-x$ is negative. This has also come up in texts, when dealing with absolute values.

UM 9&10, pg. 9, says the following:

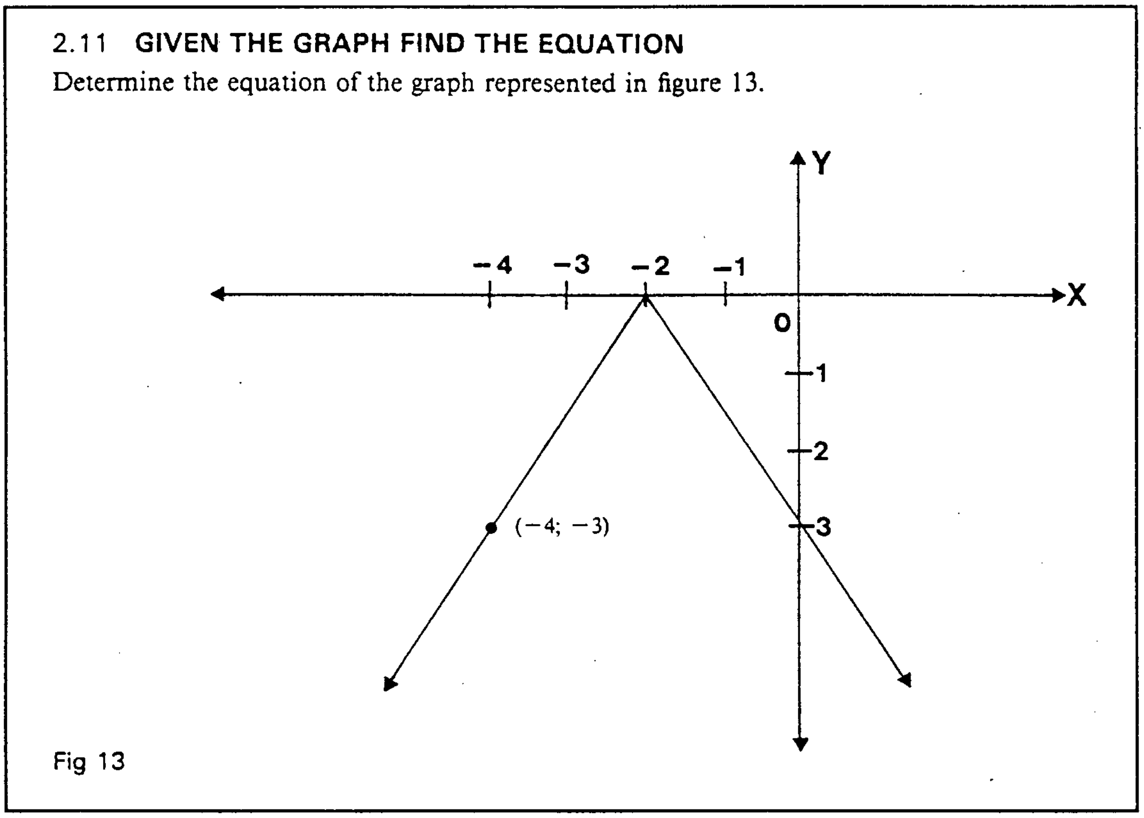


2.3 DEFINITION

When we use letters we define the absolute value of a as plus a, when a is positive, and we define the absolute value of a as $-a$, when a is negative. The answer is still positive because $-(-a)$ is plus a.

This extract clearly shows that there is no understanding that $-a$ is positive if a is itself negative. The "answer" is positive for this reason and not because "$-(-a)$ is plus a".

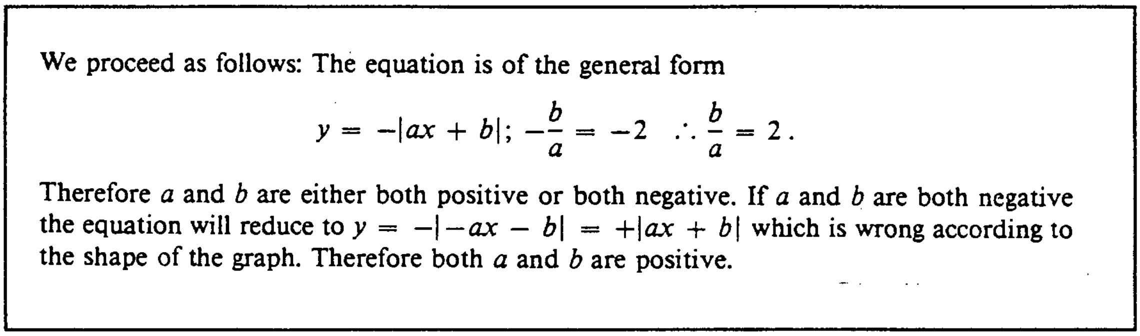
UM 9&10, pg. 20, discusses an absolute value graph as follows:



2.11 GIVEN THE GRAPH FIND THE EQUATION

Determine the equation of the graph represented in figure 13.

Fig 13



We proceed as follows: The equation is of the general form

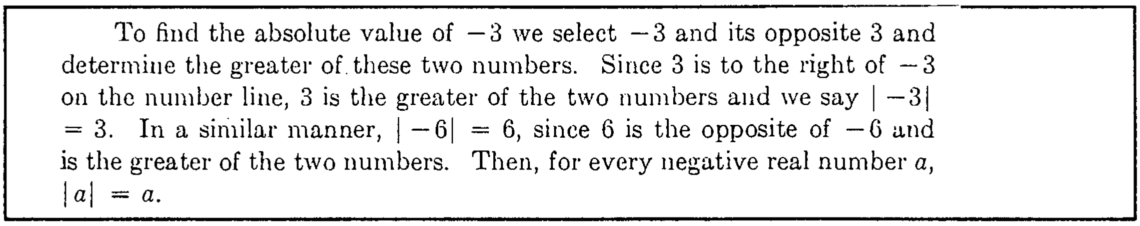
$$y = -|ax + b|; \; -\frac{b}{a} = -2 \;\; \therefore \frac{b}{a} = 2.$$

Therefore a and b are either both positive or both negative. If a and b are both negative the equation will reduce to $y = -|-ax - b| = +|ax + b|$ which is wrong according to the shape of the graph. Therefore both a and b are positive.

The first error is in deducing that $y = -|ax + b|$ becomes $y = -|-ax - b|$ **if** a and b are negative. The fact that these are the same does not depend on a and b being negative, since $|A| = |-A|$ for any A. Specifying that negative values of a and b results in $y = -|ax + b|$ becoming $y = -|-ax - b|$, could stem from the misconception that $-a$ and $-b$ are negative entities and a and b are positive entities. From this, the conclusion that "a and b are positive" is not true. The second, more serious, error is based on the misconception that $|-A| = -|A|$.

It is not clear how $-\frac{b}{a} = -2$ is obtained since $x = -2, y = 0$ implies that $0 = -|-2a + b|$ which in turn implies that $2a = b$ or $\frac{b}{a} = 2$.

IA, pg. 9, says the following on absolute values:



To find the absolute value of -3 we select -3 and its opposite 3 and determine the greater of these two numbers. Since 3 is to the right of -3 on the number line, 3 is the greater of the two numbers and we say $|-3| = 3$. In a similar manner, $|-6| = 6$, since 6 is the opposite of -6 and is the greater of the two numbers. Then, for every negative real number a, $|a| = a$.

This could very well be a misprint, but is serious enough for discussion. The fact that if $A < 0$ then $|A| = -A$ leads to the conclusion that the absolute value of a number is always positive. Many students have difficulties with understanding this and the error in the extract could aggravate the situation.

9. VECTORS

1. Introduction

During the "New Maths" era vectors were introduced into the texts. Vectors stayed in the syllabus briefly and the topic has since been replaced by co-ordinate geometry. Unlike set theory, there is little that has spilled over into texts that have since been published.

The main difficulty found concerned the approach to the dot product, and its relation to geometric concepts such as length and angle.

2. Vector concept

A vector is a quantity that has both magnitude and direction. A co-ordinate system is used so that both size and direction may be represented, e.g. two- and three-dimensional real space.

SSM 10, pg. 189, says the following:

> A second group of physical quantities consists of those quantities which are associated with magnitude as well as direction. Such quantities are called **vector quantities** or **vectors.** Speed, displacement, energy, etc. are examples of vector quantities.

The error in this extract is saying that "speed" and "energy" are vectors. This is not true, since neither have an associated direction, only magnitude.

3. The dot product

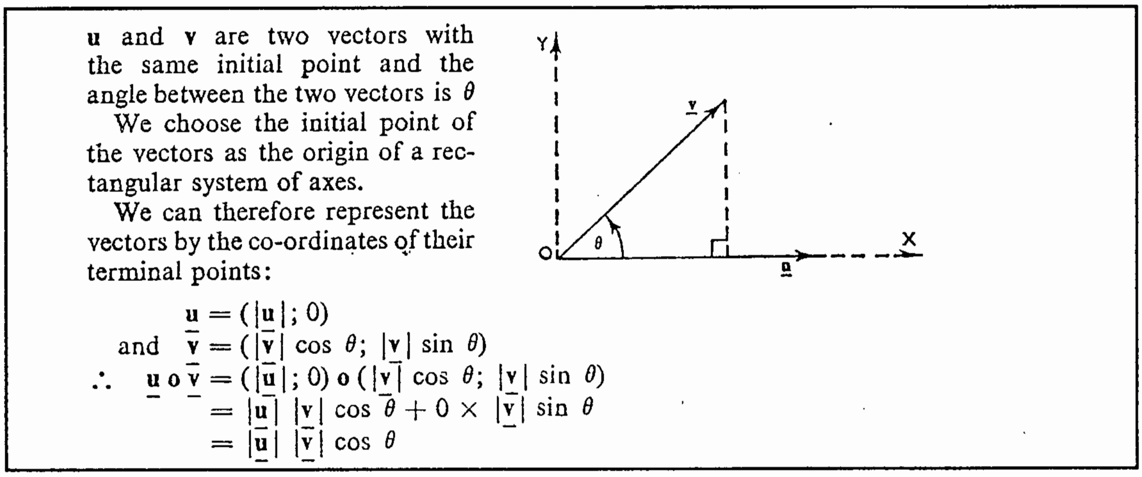
If $\underline{\mathbf{u}} = (u_1, u_2)$ and $\underline{\mathbf{v}} = (v_1, v_2)$ then the dot product of $\underline{\mathbf{u}}$ and $\underline{\mathbf{v}}$ is given by

$$\underline{\mathbf{u}} \bullet \underline{\mathbf{v}} = u_1 v_1 + u_2 v_2$$

The *length* of a vector $\underline{\mathbf{u}}$ is then defined to be $\sqrt{\underline{\mathbf{u}} \bullet \underline{\mathbf{u}}}$, and is denoted by $\left| \underline{\mathbf{u}} \right|$. The Theorem of Pythagoras shows that this definition coincides with the geometrical concept of length.

Using the cosine rule, it is readily shown ([4] pg. 652) that $\underline{u} \bullet \underline{v} = |\underline{u}| \cdot |\underline{v}| \cos\theta$ where θ is the angle between $\underline{u}$ and $\underline{v}$.

However, SSM 10, pg. 226, proves the above result as follows:



$\underline{u}$ and $\underline{v}$ are two vectors with the same initial point and the angle between the two vectors is θ

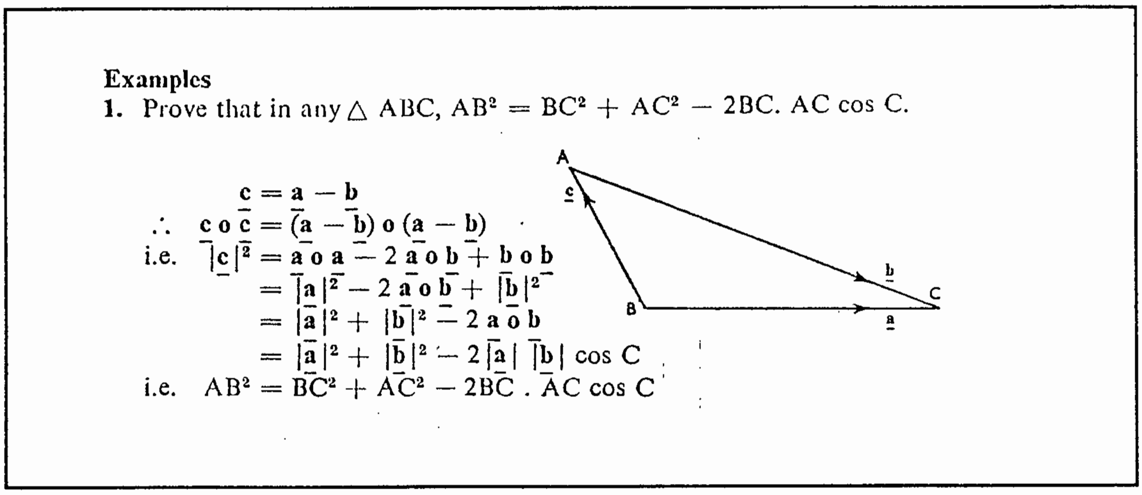
We choose the initial point of the vectors as the origin of a rectangular system of axes.

We can therefore represent the vectors by the co-ordinates of their terminal points:

$$\begin{aligned} \underline{u} &= (|\underline{u}|;\, 0) \\ \text{and}\quad \underline{v} &= (|\underline{v}| \cos \theta;\ |\underline{v}| \sin \theta) \\ \therefore\quad \underline{u} \circ \underline{v} &= (|\underline{u}|;\, 0) \circ (|\underline{v}| \cos \theta;\ |\underline{v}| \sin \theta) \\ &= |\underline{u}|\ |\underline{v}| \cos \theta + 0 \times |\underline{v}| \sin \theta \\ &= |\underline{u}|\ |\underline{v}| \cos \theta \end{aligned}$$

In this proof it is assumed that if $\underline{u}$ and $\underline{v}$ are rotated through the same angle, so that $\underline{u}$ lies on the *x*-axis, their dot product remains unchanged. This is not at all obvious from the component definition of the dot product and must be proved.

If rotation invariance is proved by means of the cosine rule, the following argument, found in SSM 10, pg. 229, is circular:



Examples

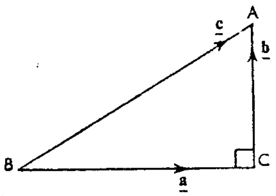
1. Prove that in any △ ABC, $AB^2 = BC^2 + AC^2 - 2BC.\,AC \cos C$.

$$\begin{aligned} \underline{c} &= \underline{a} - \underline{b} \\ \therefore\quad \underline{c} \circ \underline{c} &= (\underline{a} - \underline{b}) \circ (\underline{a} - \underline{b}) \\ \text{i.e.}\quad |\underline{c}|^2 &= \underline{a} \circ \underline{a} - 2\,\underline{a} \circ \underline{b} + \underline{b} \circ \underline{b} \\ &= |\underline{a}|^2 - 2\,\underline{a} \circ \underline{b} + |\underline{b}|^2 \\ &= |\underline{a}|^2 + |\underline{b}|^2 - 2\,\underline{a} \circ \underline{b} \\ &= |\underline{a}|^2 + |\underline{b}|^2 - 2\,|\underline{a}|\,|\underline{b}| \cos C \\ \text{i.e.}\quad AB^2 &= BC^2 + AC^2 - 2BC\,.\,AC \cos C \end{aligned}$$

On pg. 196 of SSM 10, and in general, the Theorem of Pythagoras is used to show that the vector length is the same as the geometrical length in two dimensions. The following proof in SSM 10, pg. 229, is circular because it relies on the Theorem of Pythagoras in the last step.

2. Prove the theorem of Pythagoras.

$$\underline{\mathbf{c}} = \underline{\mathbf{a}} + \underline{\mathbf{b}}$$
$$\therefore\ \underline{\mathbf{c}} \circ \underline{\mathbf{c}} = (\underline{\mathbf{a}} + \underline{\mathbf{b}}) \circ (\underline{\mathbf{a}} + \underline{\mathbf{b}})$$
$$\therefore\ |\underline{\mathbf{c}}|^2 = \underline{\mathbf{a}} \circ \underline{\mathbf{a}} + 2\,\underline{\mathbf{a}} \circ \underline{\mathbf{b}} + \underline{\mathbf{b}} \circ \underline{\mathbf{b}}$$
$$= |\underline{\mathbf{a}}|^2 + 2\,\underline{\mathbf{a}} \circ \underline{\mathbf{b}} + |\underline{\mathbf{b}}|^2$$
$$\text{But}\quad \underline{\mathbf{a}} \perp \underline{\mathbf{b}}$$
$$\therefore\ \underline{\mathbf{a}} \circ \underline{\mathbf{b}} = \underline{0}$$
$$\therefore\ |\underline{\mathbf{c}}|^2 = |\underline{\mathbf{a}}|^2 + |\underline{\mathbf{b}}|^2$$
$$\text{i.e.}\quad AB^2 = BC^2 + AC^2$$



MM 9&10, pg. 284, says the following:

Definition 2. The cosine of the angle α between two vectors $\underline{a}$ and $\underline{b}$ is defined to be

$$\cos\alpha = \frac{\underline{a}\cdot\underline{b}}{|\underline{a}|\,|\underline{b}|} = \frac{\underline{a}}{|\underline{a}|}\cdot\frac{\underline{b}}{|\underline{b}|},\ \text{provided } |\underline{a}| \neq 0 \text{ and } |\underline{b}| \neq 0.$$

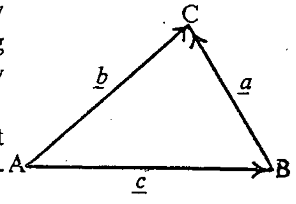
The cosine of an angle between two vectors is given as a definition in this extract. It is not obvious that this definition is equivalent to the definition in trigonometry and this needs verification.

MM 9&10, pg. 290, says:

21.2 *VECTORS IN TRIGONOMETRY* (Higher Grade)

In every mathematical system the conclusions and theorems are fundamentally based on the initial definitions and axioms. In fact the theorems are implicitly contained in the axioms (like the oak tree in the kernel of the acorn). Some theorems are almost self evident while others only follow after a long chain of previous theorems have been proved. This sequence of theorems may be altered by changing the system of axioms and thereby simplifying some theorems and complicating others. For this reason vector methods simplify the proofs of some theorems.

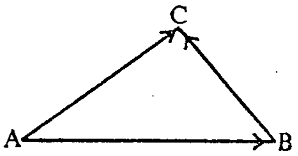
The cosine rule in trigonometry is a direct consequence of the definition of the angle between two vectors.



$$\underline{BC} = \underline{AC} - \underline{AB}$$
$$\Rightarrow\ \underline{BC}\cdot\underline{BC} = (\underline{AC} - \underline{AB})\cdot(\underline{AC} - \underline{AB})$$
$$= \underline{AC}\cdot\underline{AC} + \underline{AB}\cdot\underline{AB} - 2\underline{AC}\cdot\underline{AB}$$

(Justify this result)

$$\Rightarrow\ a^2 = b^2 + c^2 - 2bc\cos A$$



Conclusion: As a special case we have for ∠A=90° the theorem of Pythagoras:

$$a^2=b^2+c^2-2bc\cos 90°=b^2+c^2$$

The argument is circular because it has not established that the cosine of an angle in terms of the inner product is indeed equivalent to the cosine of an angle in terms of trigonometry. The fact that the inner product of a vector with itself is the square of the length of the vector needs to be verified, and requires the theorem of Pythagoras. Hence the inner product definition cannot be used to prove the cosine rule of trigonometry nor the theorem of Pythagoras.

CONCLUSION

A report in The Cape Times, 15 July 1997, brings to light the fact that there are maths and science teachers emerging from colleges without proper training. The newspaper article says that a report initiated by the National Departments of Education and Arts, Culture, Science and Technology stated that newly qualified maths and science teachers are illiterate in these fields. “Few student teachers emerging from the system could be regarded as either mathematically or scientifically literate”

The article continues by saying, “The Education Foundation, in a recent government report, said that 85% of maths teachers were qualified as teachers but only 50% had specialised in mathematics during their training.”

Thus it can be assumed that the majority of South African teachers of mathematics lack a good background in their subject. Teachers of mathematics who are poorly qualified rely heavily on textbooks in their teaching. As a result textbooks carry a great deal of authority. During this time of educational transition in South Africa, conceptual errors that arise in texts should be a great source of concern.

Both students and teachers alike tend to find it difficult to come to terms with errors in texts. If texts are not well written and are conceptually wrong this overflows into teaching and thus feeds the confusion of the students.

The more able pupil who detects textbook errors will be particularly confused. If the teacher is unable to explain such an error to the student, a conflict may occur. There is potential here for the teacher’s authority in the subject to be undermined as well as a loss of respect in terms of discipline.

This study highlights problem areas in teachers’ understanding of fundamental ideas in mathematics.

Authors of textbooks are generally experienced teachers and it can reasonably be assumed that they represent the upper end of the range of teaching ability and expertise. The fact that these authors show a lack of understanding of basic mathematical issues indicates that widespread misconceptions could exist among mathematics teachers in general.

This study highlighted the fact that conceptual errors in textbooks often arise as a result of trying to achieve less formal explanations. Mathematicians know the deeper underlying structure of the mathematics that is taught and can see conceptual errors that arise from less rigorous explanations

The benefit of a study such as this one is that a deeper knowledge and understanding of concepts for teachers of high school mathematics is provided. Although some of the ideas are more rigorous than is necessarily required for school pupils it is of benefit for the teacher to be aware of misconceptions and for teachers to be equipped with an understanding of mathematical concepts that are deeper than that which they have to teach.

Hopefully this will enable teachers to be properly equipped when faced with misconceptions in the texts and in the classroom.

An error in a mathematics textbook is the responsibility of the author and publisher. The usual procedure is for a publisher to submit the manuscript of a proposed textbook to one or more referees for an opinion on whether it is worthy of publication. Experienced editors are required to establish the accuracy of any such work.

A study such as this one may be a positive source of strength for mathematics education in this country. This could be a direct influence or at least a catalyst for creating the necessary action to address these problems.

Most of the errors highlighted by this study occurred in the treatment of new topics introduced by curriculum changes. As South Africa embarks upon "Curriculum 2005", substantial changes in terms of both content and approach are envisaged, and new textbooks are being planned and written. It is hoped that authors, publishers and education authorities act together to ensure that the new generation of textbooks will be free of the type of errors that this study has highlighted.

BIBLIOGRAPHY (Part I)

ASC	*Algebra for Senior Certificate and Matriculation* (Standards IX and X). By J. J. Dreyer and B.F. Schmidt. Juta & Company, Limited (1940). Note: This reference is an undated reprint of the 1940 edition with four chapters (A. Sets, B. Types of Numbers C. Functions, D. Series) inserted between chapters 11 and 12.
CM 9	*Classroom Mathematics, Std 9.* By P. E. J. M. Laridon, M. E. Brink, A. G. Burgess, A. M. Jawurek, A. L. Kitto, M. J. Myburgh, M. R. Pike, R. H. Rhodes-Houghton, J. Strauss, L. Strimling, R. P. van Rooyen and H. Wilson. Lexicon Publishers Johannesburg 1988.
EM 10	*Exploring Mathematics 10.* By S. W. B. Engelbrecht and A. N. P. Lubbe. Via Afrika Limited.
EM 8	*Exploring Mathematics 8.* By S. W. B. Engelbrecht and A. N. P. Lubbe. Via Afrika Limited 1988.
ES.M 7	*Essential Mathematics 7.* By B. Shapiro and D. S. A. Schell. Preskor Publishers 1974.
EX.M 8	*Exciting Maths Std 8.* By H. J. Dreyer, D. C. Badenhorst, D.V. Luthuli and D. C. J. Wessels. Lexicon Publishers Johannesburg 1992.
EX.M F3	*Exciting Maths Form III (Teacher's Edition).* By D. C. Badenhorst and H. J. Dreyer. McGraw Hill 1974.
FM 3	*Functional Mathematics for Form 3 (Std 8).* By C and N. G. Fletcher. Juta & Company Limited.
FMF III	*Functional Mathematics Form III.* By A. N. P. Lubbe and A. W. Dill Via Afrika Limited.

IA	*Intermediate Algebra.* By Donald S. Russell. Allyn and Bacon, Inc. 1969
IM Vol 1	*Integrated Mathematics Volume 1.* By G. Lambert. Shuter and Shooter 1971, Pietermaritzburg.
JM 9 & 10	*Just Mathematics 9/10 HG (1986 Syllabus).* By C. de Jager, S. Fitton and P. Blake. Maskew Miller Longman (Pty) Ltd 1985.
JMAF 1	*Juta's Modern Arithmetic Form 1* By C. Fletcher. Juta & Company, Limited 1971.
M 3	*Mathematics 3.* By J. F. A. Swart, P. J. Strydom, P. I. E. Barnard, W. A. P. von Ludwig and J. Williamson. Nasou Ltd 1980.
MASAS	*Juta's Modern Arithmetic for South African Schools.* By C. Fletcher. Juta & Company, 1969.
MAT	*Mathematics for Advanced Technical Students.* By A. R. Williams and H. Burggraaf. Nasou Limited.
MBM 9 & 10	*Modern Basic Mathematics Std 9 & 10.* By A. R. Matz, D. J. Hechter, J. A. Van Zyl and C. R. Venter. Maskew Miller (Ltd) 1975.
MGM 6	*Modern Graded Mathematics for Std 6.* By A. A. Gonin, I. J. M. Archer, G. P. L. Slabber and G. de la Rey Nel. Nasou Limited.
MGM 8	*Modern Graded Mathematics for Standard 8.* New Syllabus. By A. A. Gonin, I. J. M. Archer, G. P. L. Slabber and G. de la Rey Nel. Nasou Limited.
MGM 9 & 10	*Modern Graded Mathematics for Standards 9 & 10.* New Syllabus. By A. A. Gonin, I. J. M. Archer and G. P. L. Slabber and G. De La Rey Nel. Nasou Limited.

MGMSC	*Modern Graded Mathematics for Senior Certificate.* By A. A. Gonin, I. J. M. Archer and G. P. L. Slabber. Nasou Limited.
MM 9 & 10	*Modern Mathematics for Standards 9 & 10.* By A. P. Malan, R. L. Nero and I. P. Wait. McGraw-Hill Book Company Johannesburg 1974.
MM	*Modern Mathematics for South African Schools.* *Standard VII.* By J. J. Dreyer. Juta & Company, Limited.
MMC	*My Maths Companion (A dynamic guide for self study).* By John Small Maskew Miller Longman (Pty) Ltd 1993.
MP 8	*Mathematics Plus 8.* By W. Ladewig, R. Potgieter and J. Pretorius. Oxford University Press, Cape Town 1993.
MP 9	*Mathematics Plus 9.* By W. Ladewig, R. Potgieter and J. Pretorius. Oxford University Press, Cape Town 1995.
MRS	*Matric Revision Series: Mathematics New Syllabus.* By Sheryl Benjamin. Southern Book Publishers 1987.
MT 9	*Mathematics for Today Std 9. (Higher Grade and* *Standard Grade).* By A. J. Dekker, D. P. Visser, J. N. J. Van Rensburg. Juta & Company, Ltd 1992.
N1	*N1 Mathematics.* By Peter Shaw. Maskew Miller Longman (Pty) Ltd 1993.
NAM 8	*New Approach Mathematics for S. A. Schools, Std 8.* By A. P. Prins, J. D. Visser and W. J. A. Bester. Nasou Limited 1972.
NMM 8	*New Modern Mathematics 8.* By J. J. Dreyer and B. Dreyer. Juta & Company, Ltd.
NMM 9	*New Modern Mathematics Std 9* By J. J. Dreyer and B. Dreyer. Juta & Company 1988.

PM 5	*Primary Mathematics Std 5.* By Botes and Visser.
PM 7	*Positive Mathematics Std 7, New Syllabus.* By D. J. Visser, J. A. Visser and A. J. Schutz. Hodder and Stoughton Educational Southern Africa 1991.
SM	*Successful Mathematics Std 8.* By B. Commins, D. Duxbury and S. Makgamathe. Oxford University Press, Cape Town, 1986.
SSM 10	*Senior Secondary Mathematics for Std 10.* By H. J. Ahlers, P. J. van der Schyf. Nasou (Ltd).
TM 7	*Total Mathematics 7.* By A. J. Dekker, A. J. Wilters, A. J. Coetzee and L. J. Coetzee. Juta & Company, Ltd.
UM 9 & 10	*Understanding Mathematics 9/10. Higher and Standard Grade.* By B. van der Merwe and P. du Toit. Maskew Miller Longman (Pty) Ltd 1985.
WMI 8	*Wiskunde Met Insig.* By T. S. Welman and G. H. A. Steyn. Voortrekkerpers, Johannesburg 1969.

Note: Many South African textbooks do not reflect the date of publication, but wherever possible, dates of publication have been obtained and inserted.

BIBLIOGRAPHY (Part II)

[1] *Language in Mathematical Education.*
By David Pimm.

[2] *The South African Oxford Dictionary.*
Edited by William Branford.
Cape Town Oxford University Press 1991.

[3] *Psychology of Learning Mathematics.*
By Richard D. Skemp.
Penguin Books Ltd, Harmondsworth, Middlesex, England 1986.

[4] *Calculus.*
By James Stewart.
Brooks / Cole Publishing Company 1991.

[5] *Rings, Fields and Groups.*
By R.B.J.T. Allenby.
Edward Arnold - A division of Hodder & Stoughton 1991.

[6] *Concise Oxford Dictionary of Mathematics.*
By Christopher Clapham.
Oxford University Press 1996.

[7] *Higher Algebra.*
By H.S. Hall and S.R. Knight.
London MacMillan & Co (Ltd), New York, St Martin's Press 1957.

[8] *Modern Mathematics for Schools (2nd Edition).*
Scottish Mathematics Group.
Blackie & Son Ltd. W & R Chambers Limited 1975.

[9] *Spectrum.*
February 1990.
Foundation for Education, Science and Technology.

[10] *The College Mathematics Journal.*
Volume 28 , Number 4, September 1997.
Mathematical Association of America.

[11] *More Than Just Mathematics.*
By Tickey de Jager.
Shuter & Shooter (Pty) Ltd 1996.