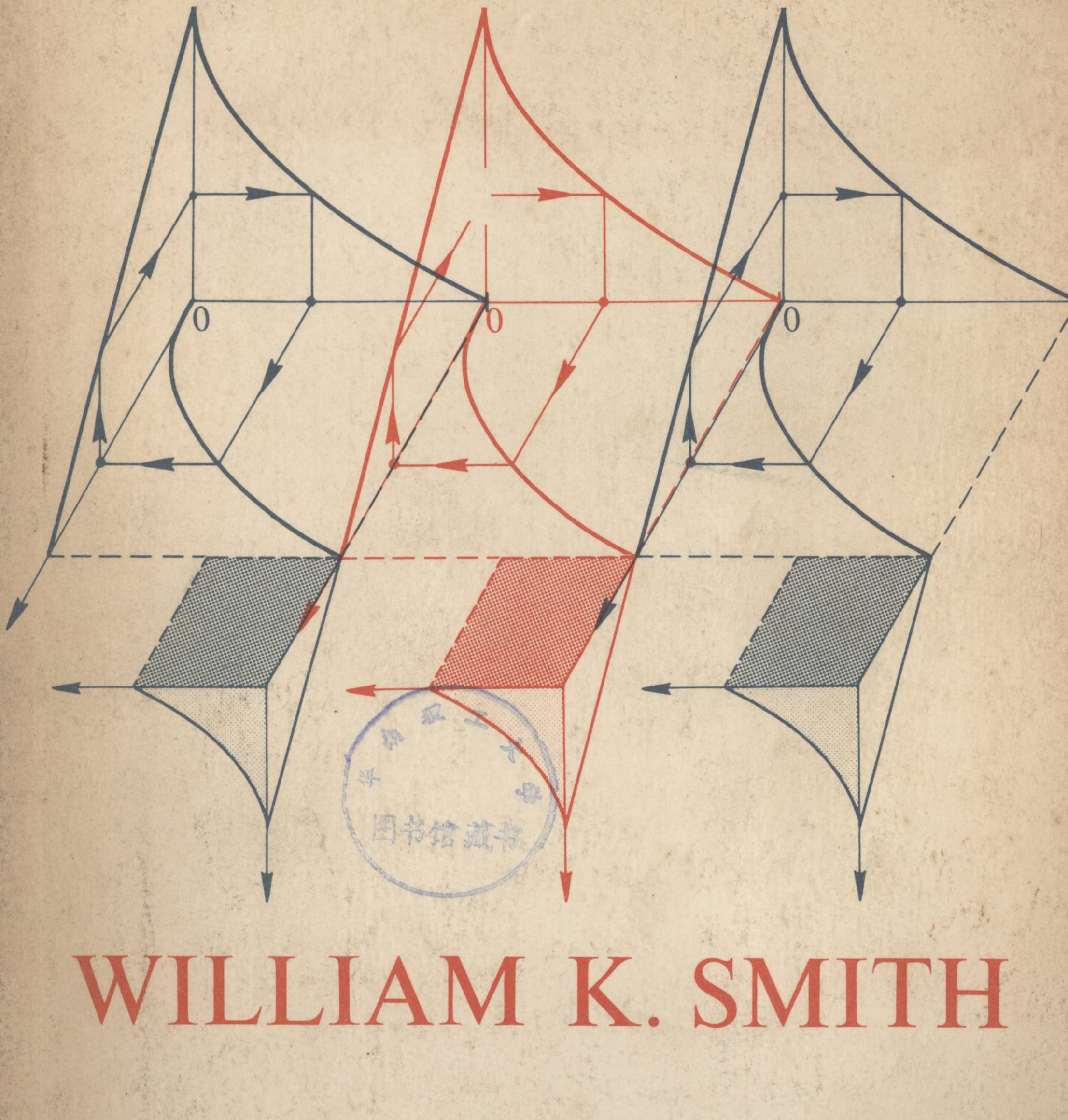
INVERSE FUNCTIONS



WILLIAM K. SMITH

CMILLAN MATHEMATICS PAPERBACKS

INVERSE FUNCTIONS

William K. Smith
NEW COLLEGE

THE MACMILLAN COMPANY · *NEW YORK*
COLLIER-MACMILLAN LIMITED · *LONDON*

First Printing

Library of Congress catalog card number: 66-11690

THE MACMILLAN COMPANY, New York
COLLIER-MACMILLAN CANADA, LTD., Toronto, Ontario

PRINTED IN THE UNITED STATES OF AMERICA

INVERSE FUNCTIONS

TO S. F. D.

Preface

THE PURPOSE of this small work is to make available to students (primarily of calculus) a reasonably complete, self-contained treatment of inverse functions, a subject which is discussed in most calculus books with distressing brevity. The result, frequently, is an inadequate understanding of the concept, especially in the case of the inverse trigonometric functions.

When writing about functions one must choose between defining a function as a set of ordered pairs or as a mapping. Although there are very strong reasons for using the former approach, my own preference is to emphasize the dynamic rather than the static form of the definition (to quote Professor R. P. Agnew). For one thing, it is the rôle of a function in linear algebra and topology (and elsewhere) to provide a mapping from one space to another; and for another, in the present work this view enables me more easily to indulge my choice to define inverse function in terms of the operation of composition of functions.

The following two conventions are used. (1) The end of a proof is indicated by the symbol ▌. This is in accordance with the custom initiated by Professor Paul R. Halmos. (2) The exercises have been classified as A, B, or C according to the following scheme: The more or less routine drill exercises are in class A, the exercises designed to further the theoretical development of the text are in class B, and the class C exercises mostly have to do with the Cartesian product of two sets and functions defined on a subset of the plane. Answers or hints or outlines for proof are provided for well over half of the exercises.

Various drafts of the manuscript were typed by Mrs. Florence Valentine and by Mrs. Nancy Hall. In both cases the work was of high quality; I am glad it is not my function to provide a ranking.

William K. Smith

Contents

SECTION 1
Introduction

ONE OF THE PARTS of elementary calculus that is frequently—indeed, almost invariably—given short shrift is that concerning inverse functions. The subject arises when the logarithmic and exponential functions are discussed. For example, if, as is sometimes the practice, the exponential function is introduced first, then the logarithmic function may be defined by saying

$$x = \log_e y \qquad \text{means} \qquad y = e^x;$$

then, perhaps mysteriously, the letters x and y are interchanged so that everything is in order, and we can write $y = \log_e x$.

Or, it may occur that in the study of the trigonometric functions the following definition is introduced:

$$x = \sin^{-1} y \qquad \text{or} \qquad x = \text{arc}\sin y \qquad \text{means} \qquad y = \sin x.$$

This *can* present a more serious problem because, in the precise sense of the word, the sine function does *not* have an inverse. Usually, this difficulty is circumvented by an appeal to "principal value," but this device does not always produce the desired level of understanding.

With the recent increased emphasis on structure, and in particular on algebraic structure, at the elementary level, and with the agreement that the definition of a function means what in the old days was called a "single-valued function," it becomes essential that any discussion of inverse functions bring out clearly and unambiguously two points:

1. The answer to the question, inverse with respect to *what operation?*
2. It is the exception, rather than the rule, for a function to have an inverse.

Our purpose here is to present a clear, rigorous, well-motivated exposition of inverse functions, using the above two points as guiding principles. As much as possible we have tried to keep within standard conventions as regards terminology and notation. Although our primary aim is to develop the concepts needed in an elementary calculus course, that is, in terms of functions whose domain and range are subsets of the field R of real numbers, we have kept the treatment as general as possible.

In Section 2 we list briefly the essential ideas from set theory, define a function, and describe the notation for and operations on functions. In Section 3 we develop the definition of an inverse function and illustrate the concept with examples. Section 4 contains a short description of a graphing technique which is sometimes useful. In Section 5 we prove, in full generality, a few necessary theorems relating to the existence of inverse functions, and in Section 6 these results are applied to functions which have as domains and ranges subsets of the field of real numbers. The final section is devoted to a lengthy discussion of the problem arising when a function does not have an inverse; in particular, the inverse trigonometric functions are treated in detail.

SECTION 2
Notation and Prerequisites

IN THIS SECTION we list some of the necessary concepts that will be used in the subsequent sections. In particular, we shall establish agreements about notation and about the definition of a function.

2.1 Sets

We will have occasion to work with *sets*, by which we mean simply collections of objects, usually referred to as *elements*. For example, a set A may consist of the first four positive integers; we indicate this as $A = \{1, 2, 3, 4\}$. Or, a set B may consist of all the real numbers between 0 and 1, inclusive. This is denoted by

$$B = \{x \mid 0 \le x \le 1\},$$

which may be verbalized as "the set of all x satisfying the condition that x is greater than or equal to 0 and less than or equal to 1." The set B described above is the *closed* (end points included) unit interval, sometimes designated by the symbol $[0, 1]$. The *open* interval (end points excluded) between a and b is the set

$$(a, b) = \{x \mid a < x < b\}.$$

We indicate that an element x is in the set X by writing $x \in X$, read "x is an element of X." Thus, using the set B defined above, $\frac{1}{2} \in B$, but 2 is not in B; this is indicated as $2 \notin B$.

Equality for sets is defined as *strict identity*; thus, $A = B$ means A and B contain precisely the same elements. One way of indicating this is by saying $x \in A$ implies $x \in B$ and $y \in B$ implies $y \in A$. A simpler way of writing this is by using the symbol $\Rightarrow$ for "implies"; also, the symbol $\Leftrightarrow$ may be used for "implies and is implied by" or, equivalently, for "if and only if" or "is logically equivalent to." Thus, the above description of equality can be expressed symbolically as

$$A = B \Leftrightarrow \begin{cases} \text{(i)} \quad x \in A \Rightarrow x \in B, \\ \qquad\qquad \text{and} \\ \text{(ii)} \quad y \in B \Rightarrow y \in A. \end{cases}$$

If it should happen that only condition (i) of the above description holds, i.e., if it should happen that $x \in A \Rightarrow x \in B$, then we say that A is a *subset* of B, and write $A \subset B$. Notice that we always have $A \subset A$. If $A \subset B$ but $A \neq B$, i.e., if there exists an element in B which is not in A, then we say that A is a *proper* subset of B. For example, the open interval (0, 1) is a proper subset of the closed interval [0, 1], since $x \in (0, 1) \Rightarrow x \in [0, 1]$, and additionally, $0 \in [0, 1]$ but $0 \notin (0, 1)$; similarly, $1 \in [0, 1]$, but $1 \notin (0, 1)$.

In addition to the two *relations* (= and $\subset$) between sets, there are several *operations* for which we will have some use: *union* and *intersection.* The union of the sets A and B, $A \cup B$, is the set of all elements in either A *or* B:

$$A \cup B = \{x \mid x \in A \text{ or } x \in B\}.$$

The intersection of the sets A and B, $A \cap B$, is the set of all elements in both A *and* B:

$$A \cap B = \{x \mid x \in A \text{ and } x \in B\}.$$

In order that no exceptions need be made we also define the *empty set* $\varnothing$ to be the set with no elements. When two sets have no elements in common (their intersection $= \varnothing$) they are said to be *disjoint.*

Thus, if $A = \{1, 2, 3, 4\}$, $B = \{3, 4, 5, 6\}$, $C = \{5, 6, 7\}$, then

$$\begin{aligned} A \cup B &= \{1, 2, 3, 4, 5, 6\} \\ A \cap B &= \{3, 4\} \\ A \cup C &= \{1, 2, 3, 4, 5, 6, 7\} \\ A \cap C &= \varnothing. \end{aligned}$$

EXERCISES

A

1. Let $A = \{1, 2, 3, 4\}$, $B = \{3, 4, 5, 6, 7\}$, $C = \{1, 3, 5, 7, 8\}$.
Find:

(a) $A \cup B$.
(b) $A \cup C$.
(c) $B \cup C$.
(d) $A \cap B$.
(e) $A \cap C$.
(f) $B \cap C$.
(g) $(A \cup B) \cup C$.
(h) $A \cap (B \cap C)$.
(i) $A \cap (B \cup C)$.
(j) $(A \cap B) \cup (A \cap C)$.
(k) $A \cup (B \cap C)$.
(l) $(A \cup B) \cap (A \cup C)$.

2. Let $A = \{x \mid 0 < x < 2\}$, $B = \{x \mid 1 < x < 5\}$, $C = \{x \mid 4 \le x \le 10\}$, where x stands for a real number.
Find:

(a) $A \cup B$.
(b) $A \cup C$.
(c) $B \cup C$.
(d) $A \cap B$.
(e) $A \cap C$.
(f) $B \cap C$.
(g) $(A \cup B) \cup C$.
(h) $(A \cap B) \cap C$.
(i) $A \cap (B \cup C)$.
(j) $(A \cap B) \cup (A \cap C)$.
(k) $A \cup (B \cap C)$.
(l) $(A \cup B) \cap (A \cup C)$.
(m) $A \cup A$.
(n) $B \cap B$.
(o) $(A \cap B) \cap B$.
(p) $(A \cap B) \cup A$.

3. Let $A = \{x \in R \mid 3 < x < 5\}$, $B = \{x \in R \mid 3.5 < x < 4.5\}$,
$C = \{x \in R \mid 3.9 < x < 4.1\}$, $D = \{x \in R \mid 3.99 < x < 4.01\}$,
$E = \{x \in R \mid 3.999 < x < 4.001\}$. R means the set of real numbers.
Find:

(a) $A \cap B$.
(b) $C \cap E$.
(c) $(B \cap C) \cap D$.
(d) $(A \cup C) \cup E$.
(e) $A \cap B \cap C \cap D \cap E$.
(f) $A \cup B \cup C \cup D \cup E$.
(g) $A \cap E$.
(h) $A \cup E$.

4. Identify differently each of the following sets. In every case $x \in R$, i.e., x stands for a real number.

(a) $A = \{x \mid x^2 = 4\}$.
(b) $B = \{x \mid x = \sqrt{4}\}$. (Answer: $B = \{2\}$.)
(c) $C = \{x \mid x = x\}$.
(d) $D = \{x \mid (x - 3)(x + 7)(x - 11) = 0\}$.
(e) $E = \{x \mid x = \sqrt{-7}\}$. (Answer: $E = \varnothing$.)

(f) $F = \{x \mid x = x^2\}$.
(g) $G = \{x \mid 2x + 3 = 15\}$.
(h) $H = \{x \mid x \neq x\}$.

5. Refer to the sets in Exercise A1. Verify that:
(a) $(A \cap B) \subset A$.
(b) $(A \cap B) \subset B$.
(c) $A \subset (A \cup B)$.
(d) $B \subset (A \cup B)$.
(e) $A \cap B = B \cap A$.

6. Same as Exercise A5, only use the sets in Exercise A2.

7. Let $A = (1, 3) = \{x \mid 1 < x < 3\}$, $B = [3, 5] = \{x \mid 3 \leq x \leq 5\}$, $C = (5, 6) = \{x \mid 5 < x < 6\}$.
Find:
(a) $A \cup B$. What symbol might be used for this kind of interval?
(b) $B \cup C$. Suggest a symbol. (Answer: [3,6).)
(c) $A \cap B$.

B

1. (a) If $A \subset B$, what can be said about $A \cap B$? About $A \cup B$?
(b) If $A \cap B = A$, what can be said about A and B?
(c) If $A \cup B = B$, what can be said about A and B?

2. If $A \cup B = A \cup C$, does $B = C$? Find an example to support your answer.

3. If $A \cap B = A \cap C$, does $B = C$? Find an example to support your answer.

4. (a) What can be said about $A \cup (B \cup C)$ and $(A \cup B) \cup C$?
(b) Is $A \cup B \cup C$ meaningful?

5. We consider $\{2\}$ and 2. (A set with one element is called a *singleton*.)
(a) Is $2 \in \{2\}$?
(b) Is $2 \subset \{2\}$?
(c) Is $\{2\} \in 2$?
(d) Is $2 = \{2\}$?
(e) Is $\{2\} \subset 2$?
(f) Is $\{2\} \subset \{2\}$?

6. Prove from the definition that $A \cap B = B \cap A$ and $A \cup B = B \cup A$.

7. If A and B are intervals in R, is $A \cap B$ an interval? What about $A \cup B$? (One way to handle this question is to consider the various possible cases. Also, examine Exercises A3 and A7.)

8. Let B be any set. Prove that $\varnothing \subset B$.
Hint: Suppose $\varnothing$ is not a subset of B. Use the definition of the relation $A \subset B$ and find a contradiction.

C

Let A and B be sets. We define the *cartesian product* of A and B, $A \times B$, by the equation

$$A \times B = \{(a, b) \mid a \in A, b \in B\},$$

where (a, b)* stands for the *ordered pair* consisting of first element a and second element b. For example, if $A = \{1, 2\}$, $B = \{x, y, z\}$, then

$$A \times B = \{(1, x), (1, y), (1, z), (2, x), (2, y), (2, z)\}.$$

1. For the above example, find $B \times A$. Does $A \times B = B \times A$?

2. For the above example, find $A \times A$ and $B \times B$.

3. Suppose A has 5 elements and B has 7 elements. How many elements are there in $A \times B$? In $B \times A$?

4. Let A be as above. Find $A \times \varnothing$.

5. Suppose $A \times B = \varnothing$. What can be said about A and B?

6. Let R = the set of real numbers. What is the geometric interpretation of $R \times R$? Notice that

$$R \times R = \{(x, y) \mid x \in R, y \in R\}.$$

7. See Exercise C6 above. In this exercise we introduce a concept from analytic geometry which will be of use in Section 6. First a definition.

Definition. *The two points* $P_1(x_1, y_1)$ and $P_2(x_2, y_2)$ are symmetric with respect to the line $y = x$ *means exactly that the line* $y = x$ *is the perpendicular bisector of the line segment joining* P_1 *and* P_2.

---

*Note that the (a, b) here does *not* mean an open interval. Usually the context makes it clear which meaning should be attached to (a, b).

We now ask you to prove the following assertion:

Theorem.

$$\left\{\begin{matrix} P_1(x_1, y_1) \text{ and } P_2(x_2, y_2) \text{ are symmetric} \\ \text{with respect to the line } y = x. \end{matrix}\right\} \Leftrightarrow \left\{\begin{matrix} x_2 = y_1 \\ y_2 = x_1 \end{matrix}\right\}$$

Hints:

To prove the $\Rightarrow$ part of the assertion either use some congruent triangles or, if you prefer analytic methods:

(1) Use the normal form of the equation of the line $y = x$ and the expression for the distance from a line to a point to obtain

$$x_2 - y_2 = -x_1 + y_1. \qquad (*)$$

(2) Use the slope formula to obtain

$$x_2 + y_2 = x_1 + y_1. \qquad (**)$$

(3) Solve (*) and (**) simultaneously for x_2 and y_2 as unknowns.

2.2 Functions

Of the various ways of defining functions, the most satisfactory for our present purposes is that which treats a function as a mapping from one set to another. Although we shall be concerned largely with functions which map a subset (perhaps all) of the real numbers into a subset of the real numbers, we give the definition in its general form.

Definition 1. *A* function f *consists of a set* X, *a set* Y *(which may equal* X*), and a rule which makes correspond to each element* $x \in X$ *exactly one element* $f(x) \in Y$. *The set* X *is called the* domain of f *and the subset of* Y *defined as*

$$f(X) = \{y \in Y \mid y = f(x), \text{ some } x \in X\}$$

is called the range of f.

The range of f is, in other words, the set of all *images* under the mapping (see examples below).

Functions may be described in various ways. For example, if $X = \{1, 2, 3, 4\}$, $Y = \{a, b, c, d, e\}$, the rule might be given by Table 1, the element appearing in the $f(x)$ column being the image of the element opposite it in the x column.

TABLE 1

x	$f(x)$
1	a
2	b
3	c
4	a

Thus, $f(1) = a$, etc. Note that $a = f(4)$ also, i.e., a is the image of both 1 and 4, an allowable occurrence. For this example the range $f(X) = \{a, b, c\}$, a proper subset of Y.

In the sequel X and Y will often be subsets of the set R of real numbers, and in these cases the rule will usually be given by some formula or "recipe." Then the domain will usually be taken as the largest admissible subset of R.

Thus, if f has recipe $f(x) = x^2$, all real numbers are admissible, so the domain $X = R$; but the range $f(X) = \{y \mid y \geq 0\}$ is the set of all non-negative real numbers. Our earlier remark about looking on a function as providing a mapping is illustrated in Figure 1. In this sense this function maps all of R into the non-negative part of R—in geometric terms, the entire x axis is mapped into the upper half of the y axis.

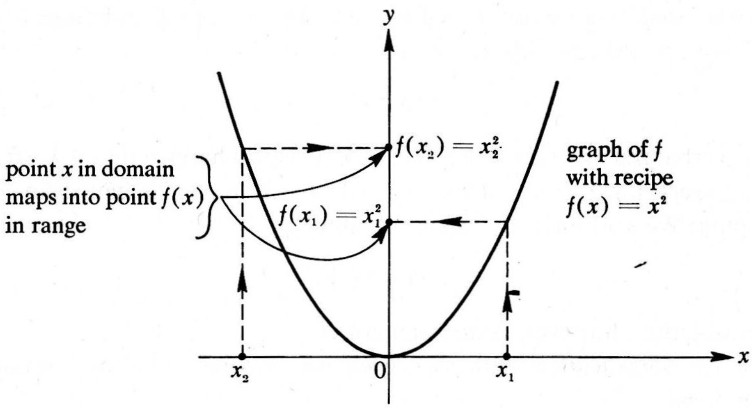


Figure 1