THE CALCULUS

Thomas Wieting

Reed College

2006

THE CALCULUS

0 Introduction

1° The object of this brief text is to develop the Calculus. The basic players are Numbers and Functions, the basic actions are Differentiation and Integration, and the basic result is the Fundamental Theorem. From these roots spring applications of astonishing variety. For an example, we sketch the explanation of the Rainbow by René Descartes (1637a).

2° For style, we adopt a spare, relentlessly formal tone. In this way, we hope to encourage in young students that cheerfully implacable attitude of mind which is characteristic of the serious study of Mathematics.

3° In the last section, we have compiled problems which will provide practice and encourage reflection.

1 Real Numbers

4° The concept of Real Number derives from the actions of *counting* and *measuring*. By common experience with such actions, we are led to introduce numbers such as:

$$1,\ 2,\ \frac{5}{2},\ \sqrt{7},\ \pi,\ \ldots$$

These are instances of real numbers. We are also led to combine two numbers in the form of a *sum* or a *product*:

$$\frac{5}{2} + \sqrt{7},\quad \sqrt{7}\pi$$

These are instances of the Operations of Addition and Multiplication. Finally, we are led to relate two numbers:

$$\frac{5}{2} < \pi$$

This is an instance of the Relation of Order. Let us gather together all the real numbers into one set:

$$\mathcal{R}$$

and let us describe a crafted array of properties for Addition, Multiplication, and Order on the set $\mathcal{R}$. These properties are sufficient to produce all other properties by logical inference. They are the hypotheses from which we develop the Calculus.

Addition and Multiplication

5° Given any numbers x and y in $\mathcal{R}$, let us write:

$$x + y \quad \text{and} \quad x \times y$$

to stand for the *sum* and the *product* of x and y. We **assume** first that:

(•) for any numbers x and y in $\mathcal{R}$, $x + y = y + x$ and $x \times y = y \times x$

(•) for any numbers x, y, and z in $\mathcal{R}$, $(x + y) + z = x + (y + z)$ and $(x \times y) \times z = x \times (y \times z)$

(•) for any numbers x, y, and z in $\mathcal{R}$, $x \times (y + z) = (x \times y) + (x \times z)$

One refers to these properties as the Laws of Commutation, Association, and Distribution. We **assume** second that:

(•) there exist numbers 0 and 1 in $\mathcal{R}$ such that, for any number x in $\mathcal{R}$, $0 + x = x$ and $1 \times x = x$; moreover, $0 \neq 1$

One can show that the indicated numbers 0 and 1 in $\mathcal{R}$ are unique. One refers to 0 and 1 as the *neutral* numbers in $\mathcal{R}$ for addition and multiplication and one calls them *zero* and *one*. We **assume** third that:

(•) for any number x in $\mathcal{R}$, there exists a number y in $\mathcal{R}$ such that $x + y = 0$

One can show that, for any number x in $\mathcal{R}$, the indicated number y in $\mathcal{R}$ is unique. One refers to y as the *additive inverse* of x and one denotes it by $-x$. Noting the symmetric relation between x and y, one may infer that not only $-x = y$ but also $-y = x$, which is to say that $-(-x) = x$. For any numbers x and y in $\mathcal{R}$, one defines:

$$x - y := x + (-y)$$

and one refers to $x - y$ as the *difference* between x and y. At this point, one can prove the following properties:

(1) for any number x in $\mathcal{R}$, if $x + x = x$ then $x = 0$

(2) for any number x in $\mathcal{R}$, $0 \times x = 0$ and $(-1) \times x = -x$

(3) for any numbers x and y in $\mathcal{R}$, $-(x+y) = (-x)+(-y)$, $-(x-y) = y - x$, $(-x) \times y = -(x \times y)$, and $(-x) \times (-y) = x \times y$

We **assume** fourth that:

(•) for any number x in $\mathcal{R}$, if $x \neq 0$ then there exists a number y in $\mathcal{R}$ such that $x \times y = 1$

One can show that, for any number x in $\mathcal{R}$, if $x \neq 0$ then the indicated number y in $\mathcal{R}$ is unique. One refers to y as the *multiplicative inverse* of x and one denotes it by x^{-1}. By property (2), $x^{-1} \neq 0$. Noting the symmetric relation between x and y, one may infer that not only $x^{-1} = y$ but also $y^{-1} = x$, which is to say that $(x^{-1})^{-1} = x$. For any numbers x and y in $\mathcal{R}$, if $y \neq 0$ then one defines:

$$x/y \equiv \frac{x}{y} := x \times y^{-1}$$

and one refers to x/y as the *quotient* of x and y. Now one can prove the following properties:

(4) for any number x in $\mathcal{R}$, if $x \times x = x$ then $x = 0$ or $x = 1$

(5) for any numbers x and y in $\mathcal{R}$, if $x \neq 0$ and $y \neq 0$ then $x \times y \neq 0$, $(x \times y)^{-1} = x^{-1} \times y^{-1}$, and $(x/y)^{-1} = y/x$

Order

6° Given any numbers x and y in $\mathcal{R}$, let us write:

$$x < y$$

to state that x is *less than* y. We **assume** fifth that:

(•) for any number x in $\mathcal{R}$, $x \not< x$

(•) for any numbers x and y in $\mathcal{R}$, if $x < y$ then $y \not< x$

(•) for any numbers x, y, and z in $\mathcal{R}$, if $x < y$ and $y < z$ then $x < z$

(•) for any numbers x and y in $\mathcal{R}$, if $x \neq y$ then $x < y$ or $y < x$

With reference to the first of the foregoing conditions, one says that the order relation on $\mathcal{R}$ is *antireflexive*; to the second, *antisymmetric*; to the third, *transitive*; to the fourth, *linear*. We **assume** sixth that:

(•) for any numbers x, y, and z in $\mathcal{R}$, if $x < y$ then $x + z < y + z$

(•) for any numbers x, y, and z in $\mathcal{R}$, if $x < y$ and $z < 0$ then $y \times z < x \times z$

(•) for any numbers x, y, and z of X, if $x < y$ and $0 < z$ then $x \times z < y \times z$

In this context, one can readily show that:

(6) for any numbers x and y in $\mathcal{R}$, if $0 < x$ and $0 < y$ then $0 < x + y$ and $0 < x \times y$

(7) for any numbers x and y in $\mathcal{R}$, $x < y$ iff $0 < y - x$

(8) $-1 < 0 < 1$

(9) for any numbers x and y in $\mathcal{R}$, if $0 < x < y$ then $0 < y^{-1} < x^{-1}$

In particular, if $0 < 1 < y$ then $0 < y^{-1} < 1$.

7° For convenience of expression, let us write $x \leq y$ to mean that $x < y$ or $x = y$.

8° One denotes the subset of $\mathcal{R}$ consisting of all numbers x for which $x < 0$ by $\mathcal{R}^-$ and one refers to the numbers in $\mathcal{R}^-$ as *negative*. Similarly, one denotes the subset of $\mathcal{R}$ consisting of all numbers x for which $0 < x$ by $\mathcal{R}^+$ and one refers to the numbers in $\mathcal{R}^+$ as *positive*. Clearly, the sets $\mathcal{R}^-$, $\{0\}$, and $\mathcal{R}^+$ comprise a partition of $\mathcal{R}$:

$$\mathcal{R} = \mathcal{R}^- \cup \{0\} \cup \mathcal{R}^+$$

Of course, for any number x in $\mathcal{R}$, $x \in \mathcal{R}^+$ iff $-x \in \mathcal{R}^-$.

Absolute Value

9° For any number x in $\mathcal{R}$, one defines the *absolute value* of x as follows:

$$|x| := \begin{cases} -x & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ x & \text{if } 0 < x \end{cases}$$

Obviously, $0 \leq |x|$, $|x| = 0$ iff $x = 0$, and $|-x| = |x|$. By checking cases, one can show that:

(10) for any numbers x and y in $\mathcal{R}$, $|x \times y| = |x| \times |y|$

(11) for any numbers x and y in $\mathcal{R}$, $|x + y| \leq |x| + |y|$

Noting $|y| = |x + (y - x)| \leq |x| + |y - x|$ and $|x| = |y + (x - y)| \leq |y| + |x - y|$, one may infer that:

(12) for any numbers x and y in $\mathcal{R}$, $|\,|x| - |y|\,| \leq |x - y|$

The Completeness Principles

10° Now let X be any subset of $\mathcal{R}$. One says that X admits a *smallest* number iff there is a number a in X such that, for each number x in X, $a \leq x$. Such a number a in X would be unique, if it exists. It would of course be the smallest number in X. In turn, one says that X admits a *largest* number iff there is a number b in X such that, for each number x in X, $y \leq b$. Such a number b in X would be unique, if it exists. It would of course be the largest number in X. Let z be any number in $\mathcal{R}$. One says that z is a *lower bound* for X iff, for each number x in X, $z \leq x$. In turn, one says that z is an *upper bound* for X iff, for each number x in X, $x \leq z$. We shall denote by X_* and by X^* the subsets of $\mathcal{R}$ consisting of all lower bounds for X and of all upper bounds for X, respectively. Of course, either X_* or X^* may be empty. We **assume** finally that the following conditions hold:

(GLB) for any subset X of $\mathcal{R}$, if $X \neq \emptyset$ and $X_* \neq \emptyset$ then X_* admits a largest number

(LUB) for any subset X of $\mathcal{R}$, if $X \neq \emptyset$ and $X^* \neq \emptyset$ then X^* admits a smallest number

One refers to condition (GLB) as the *Greatest Lower Bound Principle* and to condition (LUB) as the *Least Upper Bound Principle* for $\mathcal{R}$. They are the *Completeness Principles*. The basic theorems of the Calculus depend upon these principles for proof.

11° Actually, the two principles are logically equivalent. That is, (GLB) is true iff (LUB) is true. To prove the equivalence, we proceed as follows. Let X be any subset of $\mathcal{R}$ and let $-X$ be the subset of $\mathcal{R}$ consisting of all numbers of the form $-x$, where x is any number in X. Of course, $-(-X) = X$. Obviously, $X \neq \emptyset$ iff $-X \neq \emptyset$. Moreover, $(-X)^* = -X_*$, so $X_* \neq \emptyset$ iff $(-X)^* \neq \emptyset$. Finally, for any number z in $\mathcal{R}$, z is the largest number in X_* iff $-z$ is the smallest number in $(-X)^*$. These observations entail that (GLB) is true iff (LUB) is true. ♮

Integers

12° Let Y be any subset of $\mathcal{R}$. Let us say that Y is *inductive* iff $1 \in Y$ and, for any number y in $\mathcal{R}$, if $y \in Y$ then $y + 1 \in Y$. Let $\mathbf{Y}$ be the family of all inductive subsets of $\mathcal{R}$. Clearly, $\mathcal{R}^+ \in \mathbf{Y}$. Hence, $\mathbf{Y}$ is not empty. Let $\mathcal{Z}^+ := \cap\mathbf{Y}$. Clearly, $\mathcal{Z}^+ \in \mathbf{Y}$. One may say that $\mathcal{Z}^+$ is the smallest among all inductive subsets of $\mathcal{R}$. In particular, $\mathcal{Z}^+ \subseteq \mathcal{R}^+$. One refers to the numbers j in $\mathcal{Z}^+$ as *positive integers*.

13° Let $\mathcal{Z}^- := -\mathcal{Z}^+$. One refers to the numbers j in $\mathcal{Z}^-$ as *negative integers*. Let:

$$\mathcal{Z} := \mathcal{Z}^- \cup \{0\} \cup \mathcal{Z}^+$$

One refers to the numbers j in $\mathcal{Z}$ simply as *integers*.

Notation

14° Let x, y, and z be any numbers in $\mathcal{R}$ and let j be any integer in $\mathcal{Z}$. From now on, we usually write:

$$x \cdot y \quad \text{or simply} \quad xy$$

instead of $x \times y$. If $2 \leq j$ then we write:

$$z^j$$

instead of $z \times z \times \cdots \times z$, where z occurs j times. We interpret z^0 to be 1 and z^1 to be z. If $z \neq 0$ and $j < 0$ then we write z^j instead of $(z^{-1})^{-j}$. Finally, if $0 \leq j$ then we write $j!$ for the *factorials*:

$$0! := 1,\ 1! := 1,\ 2! := 2,\ 3! := 6,\ 4! := 24,\ 5! := 120,\ 6! := 720,\ \ldots$$

15° Let us explain the relation between $\mathcal{Z}^+$ and the widely applied method of argument called Mathematical Induction. One begins with a statement:

$$S_k$$

depending upon the integer k in $\mathcal{Z}^+$. For example:

S_k : the sum of the squares of all integers j in $\mathcal{Z}^+$ for which $1 \leq j \leq k$ equals $k(k+1)(2k+1)/6 \quad \cdots$ briefly:

$$\sum_{j=1}^{k} j^2 = \frac{1}{6}k(k+1)(2k+1)$$

Visibly, there are infinitely many such statements:

$$S_1, S_2, S_3, S_4, S_5, \ \ldots$$

One wishes to prove them all but, of course, one cannot do so by proving them one at a time. Rather, one applies Mathematical Induction to achieve the effect, in two strokes. Specifically, one proves that the following two statements are true:

(MI.1) S_1

(MI.2) for any positive integer k, IF S_k is true THEN S_{k+1} is true

One then asserts that all the foregoing statements are true. To defend the assertion, one introduces the subset Y of $\mathcal{Z}^+$ consisting of all positive integers k such that S_k is true, one notes that statements (MI.1) and (MI.2) mean that Y is inductive, and one affirms that $Y = \mathcal{Z}^+$.

16° Let us illustrate argument by Mathematical Induction by proving the following basic fact:

(13) for any integers j and k in $\mathcal{Z}$, $j+k$ and $j \times k$ are integers in $\mathcal{Z}$

By property (3), we may restrict our attention to positive integers. Let k be any positive integer. Let A_k and M_k be the statements:

A_k : for any positive integer j, $j+k$ is a positive integer

M_k : for any positive integer j, $j \times k$ is a positive integer

Obviously, A_1 is true. Moreover, for any positive integers j and k:

$$j + (k + 1) = (j + k) + 1$$

Hence, if A_k is true then A_{k+1} is true. By Mathematical Induction, all the statements A_k are true. Obviously, M_1 is true. Moreover, for any positive integers j and k:

$$j \times (k + 1) = (j \times k) + j$$

Hence, if M_k is true then M_{k+1} is true, because A_j is true. By Mathematical Induction, all the statements M_k are true. We conclude that, for any positive integers j and k, $j + k$ and $j \times k$ are positive integers. $\natural$

The Least and Greatest Integer Principles

17° Let X be the subset of $\mathcal{R}$ consisting of all numbers x for which $1 \leq x$. Clearly, X is inductive. Hence, $\mathcal{Z}^+ \subseteq X$. We conclude that 1 is the smallest integer in $\mathcal{Z}^+$.

18° Now let us prove the following fundamental properties of $\mathcal{Z}$:

(LI) for any subset Y of $\mathcal{Z}$, if $Y \neq \emptyset$ and $Y_* \neq \emptyset$ then Y admits a smallest integer

(GI) for any subset Y of $\mathcal{Z}$, if $Y \neq \emptyset$ and $Y^* \neq \emptyset$ then Y admits a largest integer

Let Y be any subset of $\mathcal{Z}$ such that $Y \neq \emptyset$ and $Y_* \neq \emptyset$. By the Greatest Lower Bound Principle, we can introduce the largest number b in Y_*. We contend that $b \in Y$. Let us suppose, to the contrary, that $b \notin Y$. Since $b + 1 \notin Y_*$, we could introduce an integer j in Y such that $b < j < b + 1$. Since $j \notin Y_*$, we could, in turn, introduce an integer k in Y such that $b < k < j < b + 1$. We would infer that $j - k \in \mathcal{Z}^+$ and $0 < j - k < 1$, contradicting the fact that 1 is the smallest integer in $\mathcal{Z}^+$. Hence, $b \in Y$, so that b is the smallest integer in Y. We conclude that (LI) is true. In similar manner, one can prove that (GI) is true. $\natural$

19° One refers to property (LI) as the *Least Integer Principle* and to property (GI) as the *Greatest Integer Principle* for $\mathcal{Z}$.

The Principle of Archimedes

20° We contend that $\mathcal{Z}^* = \emptyset$. Let us suppose, to the contrary, that $\mathcal{Z}^* \neq \emptyset$. By the Least Upper Bound Principle, we could introduce the smallest number a in $\mathcal{Z}^*$. Since $a - 1 \notin \mathcal{Z}^*$, we could then introduce a number k in $\mathcal{Z}$ such that

$a-1 < k$. It would follow that $k+1 \in \mathcal{Z}$ and that $a < k+1$, in contradiction to the definition of a. Hence, $\mathcal{Z}^* = \emptyset$. One calls this conclusion the *Principle of Archimedes*:

(14) for any number x in $\mathcal{R}$, there is some integer j in $\mathcal{Z}$ such that $x < j$

It follows that, for any number y in $\mathcal{R}$, if $0 < y$ then there is some integer k in $\mathcal{Z}^+$ such that $0 < y^{-1} < k$, hence $0 < k^{-1} < y$. See property (9).

21° By similar argument, one can show that $\mathcal{Z}_* = \emptyset$, hence that, for any number y in $\mathcal{R}$, there is some integer k in $\mathcal{Z}$ such that $k < y$.

Rationals

22° Let $\mathcal{Q}$ be the subset of $\mathcal{R}$ consisting of all numbers of the form:

$$j/k \equiv \frac{j}{k}$$

where j and k are any integers in $\mathcal{Z}$ for which $k \neq 0$. One refers to the numbers in $\mathcal{Q}$ as *rationals*. One can easily prove that:

(15) for any rationals r and s in $\mathcal{Q}$, $r+s$ and $r \cdot s$ are rationals in $\mathcal{Q}$

In fact:

$$\frac{j}{k} + \frac{\ell}{m} = \frac{j \cdot m + k \cdot \ell}{k \cdot m}$$

and:

$$\frac{j}{k} \cdot \frac{\ell}{m} = \frac{j \cdot \ell}{k \cdot m}$$

where j, k, ℓ, and m are any integers in $\mathcal{Z}$ for which $k \neq 0$ and $m \neq 0$.

23° Of course, we have the now familiar partition of $\mathcal{Q}$:

$$\mathcal{Q} = \mathcal{Q}^- \cup \{0\} \cup \mathcal{Q}^+$$

where $\mathcal{Q}^-$ consists of the negative rationals and where $\mathcal{Q}^+$ consists of the positive rationals.

24° Let us prove the following basic property of $\mathcal{Q}$:

(16) for any numbers x and y in $\mathcal{R}$, if $x < y$ then there is a rational r in $\mathcal{Q}$ such that $x < r < y$

Of course, $0 < y - x$. By the Principle of Archimedes, we can introduce an integer k in $\mathcal{Z}^+$ such that $k^{-1} < y - x$. Clearly:

$$1 < ky - kx$$

Let Y be the subset of $\mathcal{Z}$ consisting of all integers ℓ for which $kx < \ell$. By the Principle of Archimedes, $Y \neq \emptyset$. Obviously, kx is a lower bound for Y, so $Y_* \neq \emptyset$. By the Least Integer Principle, we can introduce the smallest integer j in Y. Clearly:

$$j - 1 \leq kx < j$$

Hence, $j < ky$. Therefore:

$$x < \frac{j}{k} < y$$

Now we can take r to be j/k. ♮

25° With reference to property (16), one says that $\mathcal{Q}$ is *dense* in $\mathcal{R}$.

Irrationals

26° Let $\mathcal{I}$ be the subset of $\mathcal{R}$ complementary to $\mathcal{Q}$:

$$\mathcal{I} := \mathcal{R} \backslash \mathcal{Q}$$

One refers to the numbers in $\mathcal{I}$ as *irrationals*. Let us prove that:

(17) for any numbers x and y in $\mathcal{R}$, if $x < y$ then there is an irrational z in $\mathcal{I}$ such that $x < z < y$

That is, let us prove that $\mathcal{I}$ is dense in $\mathcal{R}$. For that purpose, we need only prove that:

$$\mathcal{I} \cap \mathcal{R}^+ \neq \emptyset$$

Let us assume for the moment that we have done so. Let $\bar{z}$ be any number in $\mathcal{I} \cap \mathcal{R}^+$. Let x and y be any numbers in $\mathcal{R}$ such that $x < y$. By property (16), we can introduce a rational number r in $\mathcal{Q}$ such that $x < r < y$. Let k be an integer in $\mathcal{Z}^+$ such that $\bar{z}/k < y - r$. Let $z := r + (\bar{z}/k)$. Clearly, $z \in \mathcal{I}$ and $x < z < y$. We conclude that $\mathcal{I}$ is dense in $\mathcal{R}$.

27° Now let us proceed to prove that $\mathcal{I} \cap \mathcal{R}^+ \neq \emptyset$. For that purpose, let us introduce the subset X of $\mathcal{Q}^+$ consisting of all positive rationals r for which $r^2 < 2$. Obviously, $X \neq \emptyset$ and $X^* \neq \emptyset$. By the Least Upper Bound Principle, we can introduce the smallest number z in X^*. Obviously, $0 < z$. Let us prove that $z^2 = 2$. To that end, let us suppose that $z^2 < 2$. Under this supposition, we could introduce the positive number $v := 2 - z^2$. In turn, we could introduce a positive number u such that $u < 1$ and $u(2z+1) < v$. We would find that $(z+u)^2 = z^2 + 2zu + u^2 = z^2 + u(2z+u) < z^2 + u(2z+1) < z^2 + v = 2$. By property (16), we could introduce a rational number r such that $z < r < z + u$. It would follow that $r^2 < 2$. Hence, $r \in X$ but $z < r$. By this contradiction, we conclude that $2 \leq z^2$. Let us suppose that $2 < z^2$. Under this supposition, we could introduce the positive number $v := z^2 - 2$. In turn, we could introduce a positive number u such that $2uz < v$. We would find that $2 = z^2 - v < z^2 - 2uz < z^2 - 2zu + u^2 = (z-u)^2$. Of course, for any r in X, $r^2 < 2 < (z-u)^2$, so $r < z - u$. Hence, $z - u \in X^*$ but $z - u < z$. By this contradiction, we conclude that $z^2 = 2$.

28° For any positive numbers z_1 and z_2 in $\mathcal{R}$, if $z_1^2 = z_2^2$ then $z_1 = z_2$, because $0 = z_1^2 - z_2^2 = (z_1 - z_2)(z_1 + z_2)$. See property (5) in article 2°. Let us adopt the conventional notation:

$$z = \sqrt{2}$$

29° Finally, let us prove the ancient and profoundly significant fact that $\sqrt{2}$ is irrational. To that end, let Y be the subset of $\mathcal{Z}^+$ consisting of all positive integers j for which there exists some positive integer k such that:

$$\sqrt{2} = \frac{j}{k}$$

Let us suppose that $Y \neq \emptyset$. By the Least Integer Principle, we could introduce the smallest integer $j_\circ$ in Y. In turn, by definition, we could introduce a positive integer $k_\circ$ such that $\sqrt{2} = j_\circ / k_\circ$. Obviously, $j_\circ^2 = 2k_\circ^2$. Hence, $j_\circ$ would be even. See the following article. We could introduce a positive integer $j_\bullet$ such that $j_\circ = 2j_\bullet$. Obviously, $k_\circ^2 = 2j_\bullet^2$. Hence, $k_\circ$ would be even. We could introduce a positive integer $k_\bullet$ such that $k_\circ = 2k_\bullet$. We would find that $j_\bullet^2 = 2k_\bullet^2$ and so:

$$\sqrt{2} = \frac{j_\bullet}{k_\bullet}$$

Hence, $j_\bullet \in Y$ but $j_\bullet < j_\circ$. By this contradiction, we infer that $Y = \emptyset$. We conclude that $\sqrt{2}$ is irrational. ♮

30° Let j be any integer in $\mathcal{Z}$. We say that j is *even* iff there is some integer ℓ in $\mathcal{Z}$ such that $j = 2\ell$. We say that j is *odd* iff there is some integer ℓ in $\mathcal{Z}$ such that $j = 2\ell + 1$. By Mathematical Induction, one can easily prove that, for any integer j in $\mathcal{Z}$, j is even or odd. Obviously, it cannot be both. Moreover, j is even iff j^2 is even and j is odd iff j^2 is odd.

Integral and Fractional Parts

31° Let x be any number in $\mathcal{R}$. Let Y be the subset of $\mathcal{Z}$ consisting of all integers j such that $j \leq x$. Clearly, $Y \neq \emptyset$ and $Y^* \neq \emptyset$. By the Greatest Integer Principle, we can introduce the largest integer k in Y. Obviously, $k \leq x < k+1$, so $0 \leq x - k < 1$. One refers to k as the *integral part* of x and one denotes it by $[x]$. One refers to $x - k$ as the *fractional part* of x and one denotes it by (x). Obviously:

$$x = [x] + (x), \quad [x] \in \mathcal{Z}, \quad 0 \leq (x) < 1$$

Representation of Numbers in Base b

32° Let b be any integer for which $2 \leq b$. We refer to b as the *base*. Let D be the subset of $\mathcal{Z}$ consisting of all integers d such that $0 \leq d < b$. We refer to the integers in D as the *digits*. In turn, let:

$$\ldots, \delta_k, \ldots, \delta_{-3}, \delta_{-2}, \delta_{-1}, \delta_0, \delta_1, \delta_2, \delta_3, \ldots, \delta_\ell, \ldots$$

be any series of digits, indexed by the integers. Let δ stand for the series just displayed. For every integer j, we refer to δ_j as the j-th *term* of δ. Let Δ be the subset of $\mathcal{Z}$ consisting of all integers k such that, for every integer ℓ, if $k < \ell$ then $\delta_\ell = 0$. It may happen that $\Delta = \emptyset$. Let us exclude that case from consideration. In effect, we accept only those series' δ which *terminate* to the right in 0. It may happen that $\Delta_* = \emptyset$. In that case, δ would be trivial, in the sense that all the terms would equal 0. Let us exclude that case from consideration, as well. By these exclusions, we are justified to introduce the smallest integer in Δ. Let us denote it by $|\delta|$. Obviously:

$$\delta_{|\delta|} \neq 0$$

and, for any integer j, if $|\delta| < j$ then $\delta_j = 0$. Finally, it may happen that there is some integer ℓ such that, for every integer k, if $k < \ell$ then $\delta_k = b - 1$. Let us exclude that case from consideration. In effect, we reject those series' which terminate to the left in $b-1$. For later reference, let us say that a series of digits is *normal* iff it survives the foregoing exclusions, which is to say that it is nontrivial, it terminates to the right in 0, and it does not terminate to the left in $b - 1$. For such a series, we will show that the expression:

$$\sum_{j=-\infty}^{|\delta|} \delta_j b^j = \cdots + \delta_{-2} b^{-2} + \delta_{-1} b^{-1} + \delta_0 + \delta_1 b^1 + \delta_2 b^2 + \cdots + \delta_{|\delta|} b^{|\delta|}$$

represents a positive number x.

33° Let δ be a normal series of digits. Let (δ) and $[\delta]$ be the fractional and integral parts of δ, defined as follows:

$$(\delta): \qquad (\delta)_j = \begin{cases} \delta_j & \text{if } j < 0 \\ 0 & \text{if } 0 \leq j \end{cases}$$

and:

$$[\delta]: \qquad [\delta]_j = \begin{cases} 0 & \text{if } j < 0 \\ \delta_j & \text{if } 0 \leq j \end{cases}$$

Obviously, if not trivial then (δ) and $[\delta]$ are normal. Let us justify representation of the number (x) by (δ):

$$(x) = \cdots + \delta_{-3}b^{-3} + \delta_{-2}b^{-2} + \delta_{-1}b^{-1}$$

and of the number $[x]$ by $[\delta]$:

$$[x] = \delta_0 + \delta_1 b^1 + \delta_2 b^2 + \cdots + \delta_{|\delta|} b^{|\delta|}$$

For $[x]$, we note that the summation is finite and that $[x]$ is a nonnegative integer. It requires no explanation. For (x), we note that, for any negative integer j:

$$\begin{aligned} 0 &\leq \delta_j b^j + \cdots + \delta_{-3}b^{-3} + \delta_{-2}b^{-2} + \delta_{-1}b^{-1} \\ &\leq (b-1)b^j + \cdots + (b-1)b^{-3} + (b-1)b^{-2} + (b-1)b^{-1} \\ &= (b-1)b^j(1 + b + b^2 + \cdots + b^{-j-1}) \\ &= (b-1)b^j(b^{-j} - 1)(b-1)^{-1} \\ &= 1 - b^j \\ &< 1 \end{aligned}$$

Let X be the subset of $\mathcal{R}$ consisting of all numbers of the form:

$$\delta_j b^j + \cdots + \delta_{-3}b^{-3} + \delta_{-2}b^{-2} + \delta_{-1}b^{-1}$$

where j is any negative integer. Obviously, $X \neq \emptyset$ and $1 \in X^*$. We are led to interpret (x) as the smallest number in X^*. Since δ does not terminate to the left in $b-1$, we can introduce a negative integer k such that $\delta_k \leq b-2$. Clearly, $1 - b^k \in X^*$. We infer that:

$$0 \leq (x) \leq 1 - b^k < 1$$

Now we are justified to present x as the following sum:

$$x = [x] + (x)$$

Obviously, $0 < [x]$ or $0 < (x)$. Hence, $0 < x$.

34° In turn, let δ' and δ'' be any two normal series' of digits. Let x' and x'' be the positive numbers represented by δ' and δ'', respectively. Let Δ be the subset of $\mathcal{Z}$ consisting of all integers j such that $\delta'_j \neq \delta''_j$. Since δ' and δ'' terminate to the right in 0, $\Delta^* \neq \emptyset$. Let δ' and δ'' be distinct, so that $\Delta \neq \emptyset$. Let ℓ be the largest integer in Δ. Let us write:

$$\delta' < \delta''$$

to express the condition that $\delta'_\ell < \delta''_\ell$. Now let us assume that $\delta' < \delta''$. We will show that $x' < x''$. Since δ' does not terminate to the left in $b-1$, we can introduce an integer k such that $\delta'_k \leq b-2$. We find that:

$$\begin{aligned}
\sum_{j=-\infty}^{k-1} \delta'_j b^j \;+\; \delta'_k b^k \;+ \sum_{j=k+1}^{\ell-1} \delta'_j b^j \;+\; \delta'_\ell b^\ell \\
&< \sum_{j=-\infty}^{\ell-1} (b-1)b^j \;+\; \delta'_\ell b^\ell \\
&= \;b^\ell \;+\; \delta'_\ell b^\ell \\
&\leq \;\delta''_\ell b^\ell
\end{aligned}$$

Hence, $x' < x''$.

35° Finally, let x be any positive number. We will show that there is a normal series δ of digits which represents x. Let Δ be the subset of $\mathcal{Z}$ consisting of all integers j such that $b^j \leq x$. Clearly, $\Delta \neq \emptyset$ and $\Delta^* \neq \emptyset$. Let ℓ be the largest integer in Δ, so that $b^\ell \leq x < b^{\ell+1}$. In turn, let d be the largest digit in D such that $db^\ell \leq x$. Obviously, $1 \leq d$ and:

$$db^\ell \leq x < (d+1)b^\ell$$

so that:

$$0 \leq x - db^\ell < b^\ell$$

At this point, we specify δ in part, as follows:

$$\delta_j = \begin{cases} d & \text{if } \ell = j \\ 0 & \text{if } \ell < j \end{cases}$$

If $x = db^\ell$ then we complete the specification of δ, as follows:

$$\delta_j = 0 \quad \text{if } j < \ell$$

If not, then we form the subset Γ of $\mathcal{Z}$ consisting of all integers j such that $b^j \leq x - d\,b^\ell$. Clearly, $\Gamma \neq \emptyset$ and $\Gamma^* \neq \emptyset$. Let k be the largest integer in Γ. Obviously, $k < \ell$ and $b^k \leq x - d\,b^\ell < b^{k+1}$. In turn, let c be the largest digit in D such that $c\,b^k \leq x - d\,b^\ell$. Obviously, $1 \leq c$ and:

$$c\,b^k \leq x - d\,b^\ell < (c+1)b^k$$

so that:

$$0 \leq x - (d\,b^\ell + c\,b^k) < b^k$$

We continue the specification δ, as follows:

$$\delta_j = \begin{cases} c & \text{if } k = j \\ 0 & \text{if } k < j < \ell \end{cases}$$

Applying the foregoing pattern recursively, we may proceed to specify δ completely. So specified, δ is nontrivial and it terminates to the right in 0. We contend that it does not terminate to the left in $b-1$. Let us suppose, to the contrary, that there is an integer j such that, for any integer i, if $i < j$ then $\delta_i = b - 1$. We would find that:

$$x - \sum_{i=j}^{\ell} \delta_i b^i = \sum_{i=-\infty}^{j-1} (b-1)b^i = b^j$$

contradicting the specification of δ_j. Hence, δ is normal. By design, for any integer j, if $j \leq \ell$ then:

$$\delta_j b^j \leq x - \sum_{i=j+1}^{\ell} \delta_i b^i < (\delta_j + 1)b^j$$

so that:

$$0 \leq x - \sum_{i=j}^{\ell} \delta_i b^i < b^j$$

Therefore, δ represents x. ♮

36° By the foregoing articles, we obtain a bijective relation between normal series' δ of digits and positive numbers x:

$$\begin{aligned} x &= \sum_{j=-\infty}^{|\delta|} \delta_j b^j \\ &= \cdots + \delta_{-2}b^{-2} + \delta_{-1}b^{-1} + \delta_0 + \delta_1 b^1 + \delta_2 b^2 + \cdots + \delta_{|\delta|}b^{|\delta|} \\ &= \delta_{|\delta|}\delta_{|\delta|-1} \cdots \delta_3\delta_2\delta_1\delta_0 \bullet \delta_{-1}\delta_{-2}\delta_{-3}\delta_{-4} \cdots \end{aligned}$$

The "point" • marks the break between the integral part and the fractional part of x. Of course, one sets the base b in advance. One refers to δ as the *base b representation* of x.

The Real Line

37° One can visualize many of the properties of $\mathcal{R}$ by means of a diagram, called the Real Line:

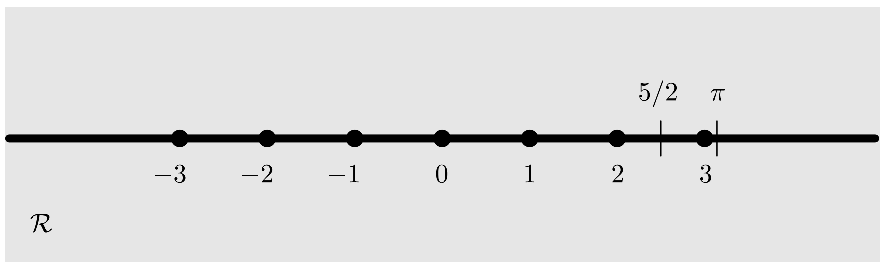


Figure 1: The Real Line

However, for our development of the Calculus, we will not substitute naive properties of the diagram for careful reasoning from the basic hypotheses.

2 Functions

1° The object of this section is to introduce the basic concepts of Interval and Function. In preparation for subsequent sections, we isolate several other related concepts.

Intervals

2° Let X be any subset of $\mathcal{R}$. Let us say that X is an *interval* in $\mathcal{R}$ iff, for any numbers x, y, and z in $\mathcal{R}$, if $x \in X$, $y \in X$, and $x < z < y$ then $z \in X$. Just as well, we may say that X satisfies the Intermediate Value Property. We contend that if X is an interval in $\mathcal{R}$ then it must fall into one of eleven types. The first two types are comprised of the *trivial* intervals:

$$\emptyset \quad \text{and} \quad [a, a] \equiv \{a\}$$

where a is any number in $\mathcal{R}$. The next four types are comprised of the *finite* intervals:

$$(a, b), \ \ [a, b], \ \ [a, b), \ \ \text{and} \ \ (a, b]$$

where a and b are any numbers in $\mathcal{R}$ for which $a < b$. By definition, for any number x in $\mathcal{R}$:

$$\begin{aligned} x \in (a, b) &\quad \text{iff} \quad a < x < b \\ x \in [a, b] &\quad \text{iff} \quad a \leq x \leq b \\ x \in [a, b) &\quad \text{iff} \quad a \leq x < b \\ x \in (a, b] &\quad \text{iff} \quad a < x \leq b \end{aligned}$$

The next four types are comprised of the *semifinite* intervals:

$$(a, \rightarrow), \ \ [a, \rightarrow), \ \ (\leftarrow, b), \ \ \text{and} \ \ (\leftarrow, b]$$

where a and b are any numbers in $\mathcal{R}$. By definition, for any number x in $\mathcal{R}$:

$$\begin{aligned} x \in (a, \rightarrow) &\quad \text{iff} \quad a < x \\ x \in [a, \rightarrow) &\quad \text{iff} \quad a \leq x \\ x \in (\leftarrow, b) &\quad \text{iff} \quad x < b \\ x \in (\leftarrow, b] &\quad \text{iff} \quad x \leq b \end{aligned}$$

The last type is comprised of the *infinite* interval:

$$\mathcal{R}$$

3° Obviously, if X falls into one of the eleven types just described then it is an interval in $\mathcal{R}$. Let us prove the converse. Let X be an interval in $\mathcal{R}$. Of course, we may assume that X contains more than one number, which is to say that X is *nontrivial*. It may happen that $X_* = \emptyset$ and $X^* = \emptyset$. In that case, $X = \mathcal{R}$. It may happen that $X_* \neq \emptyset$ and $X^* \neq \emptyset$. By the Completeness Principles, we can introduce the largest number a in X_* and the smallest number b in X^*. Obviously, $a < b$. Of course, X itself may or may not admit a smallest number and it may or may not admit a largest number. The four alternatives correspond to the four types $[\,a, b\,]$, $[\,a, b)$, $(a, b\,]$, and (a, b) of finite intervals. In turn, it may happen that $X_* \neq \emptyset$ while $X^* = \emptyset$. Of course, X itself may or may not admit a smallest number. The two alternatives correspond to the two types $[a, \rightarrow)$ and $(a, \rightarrow)$ of semifinite intervals. Finally, it may happen that $X_* = \emptyset$ while $X^* \neq \emptyset$. Of course, X itself may or may not admit a largest number. The two alternatives correspond to the two types $(\leftarrow, b]$ and $(\leftarrow, b)$ of semifinite intervals. ♮

4° One says that the intervals:

$$\emptyset,\ \mathcal{R},\ (a, b),\ (a, \rightarrow),\ (\leftarrow, b)$$

are *open* and that the intervals:

$$\emptyset,\ \mathcal{R},\ [a, a],\ [a, b],\ [a, \rightarrow),\ (\leftarrow, b]$$

are *closed*. We refer to a and b as the (left and right) *endpoints* of the relevant intervals.

Neighborhoods

5° Let a and u be any numbers for which $0 < u$. Let $N_u(a)$ stand for the finite open interval $(a - u, a + u)$:

$$N_u(a) := (a - u, a + u)$$

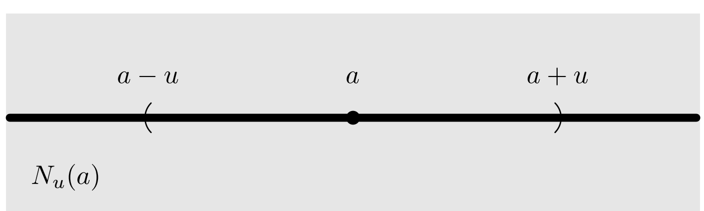


Figure 2: A Neighborhood

We refer to $N_u(a)$ as the *neighborhood* of a having *radius* u. One can easily prove that, for any number x in $\mathcal{R}$:

$$x \in N_u(a) \quad \text{iff} \quad a - u < x < a + u \quad \text{iff} \quad |x - a| < u$$

Functions

6° The concept of Function derives from common experience with relations between numbers. Very often, the relations express useful constructions or important principles. Let us describe the concept in a form suitable to our study. Let X be any subset of $\mathcal{R}$. Let F be any "rule" which assigns to each number x in X a unique *value* $F(x)$. In this context, we say that F is a *function* having *domain* X.

7° For examples, let $X_1 := \mathcal{R}$ and let F_1 be defined as follows:

$$F_1(x) := x^2$$

where x is any number in X_1. In turn, let $X_2 := \mathcal{R}^- \cup \mathcal{R}^+$ and let F_2 be defined as follows:

$$F_2(x) := \frac{1}{x}$$

where x is any number in X_2. Finally, let $X_3 := \mathcal{R}$ and let F_3 be defined as follows:

$$F_3(x) := \frac{1 - x^2}{1 + x^2}$$

where x is any number in X_3.

8° Let S be any subset of X. We define:

$$F(S)$$

to be the subset of $\mathcal{R}$ consisting of all numbers y such that there is some number x in S for which $y = F(x)$. This notation will prove convenient from time to time. In particular, we refer to:

$$F(X)$$

as the *range* of F. In the first of the foregoing examples, $F_1(X_1) = [0, \rightarrow)$. In the second, $F_2(X_2) = X_2$. In the third, $F_3(X_3) = (-1, 1]$.

Graphs of Functions

9° Let X be any subset of $\mathcal{R}$ and let F be a function for which the domain is X. By the *graph* of F, we mean the set Γ consisting of all "ordered pairs" (x, y) of numbers in $\mathcal{R}$ such that $x \in X$ and $y = F(x)$. The graph Γ provides a convenient visual form for analysis of the function F.

10° For the examples described in article 7°, the graphs have the form displayed in the following figure.

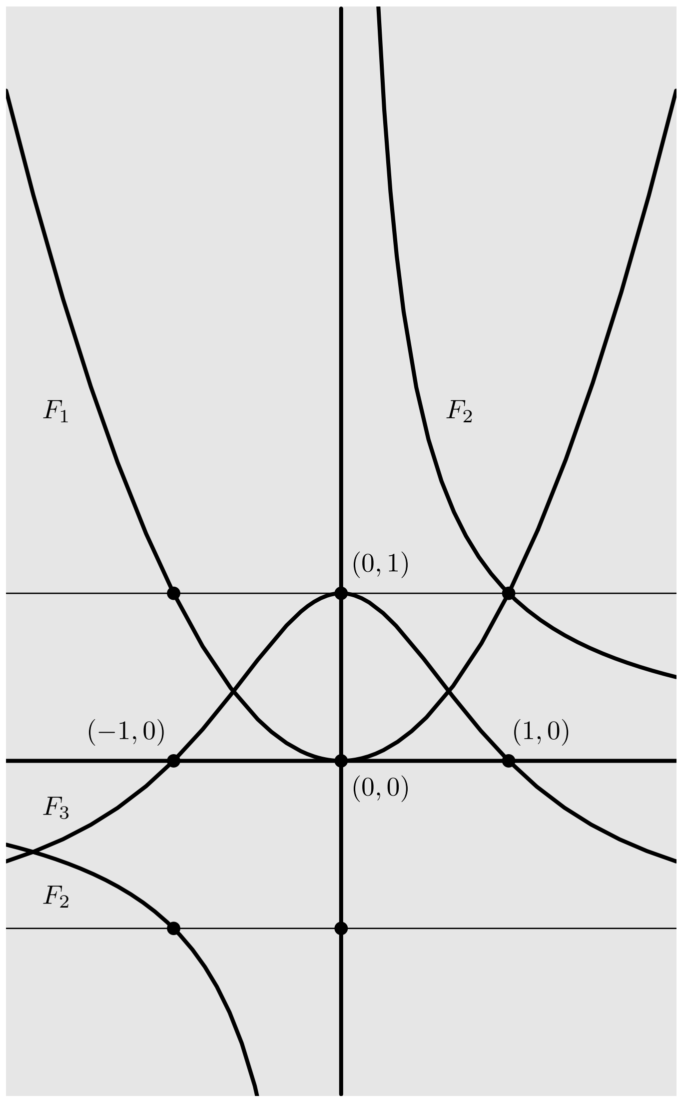


Figure 3: Graphs of Functions

Operations on Functions

11° In terms of the operations of addition and multiplication on $\mathcal{R}$, we can define corresponding operations on functions. Thus, let X be any subset of $\mathcal{R}$ and let F, G, and H be functions having common domain X. Let c be any number in $\mathcal{R}$. We define the *sum* of F and G and the *product* of F and G as follows :

$$(F + G)(x) := F(x) + G(x)$$
$$(F \times G)(x) := F(x) \times G(x)$$

and we define the *scalar product* of c and H as follows:

$$(c \times H)(x) := c \times H(x)$$

where x is any number in X. Of course, we will often write $F \cdot G$ or simply FG instead of $F \times G$ and $c \cdot H$ or simply cH instead of $c \times H$. In turn, we define the *difference* between F and G as follows:

$$(F - G)(x) := F(x) - G(x)$$

and the *quotient* of F and G as follows:

$$(F/G)(x) \equiv \frac{F}{G}(x) := \frac{F(x)}{G(x)} \equiv F(x)/G(x)$$

where x is any number in X. Naturally, for the latter definition, we presume that $G(X) \subseteq \mathcal{R}^- \cup \mathcal{R}^+$.

12° Let X be any subset of $\mathcal{R}$ and let F be a function having domain X. We define the *absolute value* of F as follows:

$$|F|(x) := |F(x)|$$

where x is any number in X.

Compositions of Functions

13° Let X and Y be any subsets of $\mathcal{R}$, let F be a function having domain X, and let G be a function having domain Y. Let us assume that $F(X) \subseteq Y$. That is, let us assume that, for each number x in X, the value $F(x)$ is in Y. In this situation, we can form a new function H having domain X, as follows:

$$H(x) := G(F(x))$$

where x is any number in X. We refer to H as the *composition* of F and G and we denote it by $G \circ F$.

14° Let $X_4 := \mathcal{R}^- \cup \mathcal{R}^+$ and let F_4 be defined as follows:

$$F_4(x) := x + \frac{1}{x}$$

where x is any number in X_4. With reference to article 7°, we have:

$$(F_3 \circ F_4)(x) = F_3(F_4(x)) = \frac{1 - (1 + 1/x)^2}{1 + (1 + 1/x)^2} = \frac{x^4 - \ \ x^2 - 1}{x^4 + 3x^2 + 1}$$

where x is any number in X_4.

15° Let us return to the context of article 13°. It may happen that not only $F(X) \subseteq Y$ but also $G(Y) \subseteq X$. If so then we can form not only $G \circ F$ but also $F \circ G$. Moreover, it may happen that, for any number x in X, $G(F(x)) = x$ and, for any number y in Y, $F(G(y)) = y$. If so then we say that F and G are *inverse* to one another.

16° One may look ahead to **Section 7** for significant examples. One will find that the logarithm function L and the exponential function E are inverse to one another and that the power functions P_a and $P_{1/a}$ $(a \neq 0)$ are inverse to one another. For now, let us simply note that F_2 is inverse to itself:

$$F_2(F_2(x)) = 1/(1/x) = x$$

where x is any number in X_2. Let us also point to the following example, inverse to itself. Let $X_5 := (-1, \rightarrow)$ and let F_5 be defined as follows:

$$F_5(x) := \frac{1 - x}{1 + x}$$

where x is any number in X_5. We find that $F_5(X_5) = X_5$ and that:

$$F_5(F_5(x)) = \frac{1 - \dfrac{1 - x}{1 + x}}{1 + \dfrac{1 - x}{1 + x}} = x$$

where x is any number in X_5.

The Order Relation on Functions

17° We can apply the order relation on $\mathcal{R}$ to define an order relation on functions. Thus, let X be any subset of $\mathcal{R}$ and let F and G be functions having common domain X. We write $F \leq G$ to express the condition that, for any number x in X, $F(x) \leq G(x)$.

Constant Functions

18° Let X be any subset of $\mathcal{R}$ and let F be a function having domain X. We say that F is *constant* iff there is a number z in $\mathcal{R}$ such that, for any number x in X, $F(x) = z$. That is, $F(X) = \{z\}$. We also say that F is *constantly* z. On occasion, we denote F by $\hat{z}$.

Bounded Functions

19° Let X be any subset of $\mathcal{R}$ and let F be a function having domain X. We say that F is *bounded* iff there are numbers c and d in $\mathcal{R}$ such that, for any number x in X, $c \leq F(x) \leq d$. Obviously, F is bounded iff $F(X)_* \neq \emptyset$ and $F(X)^* \neq \emptyset$.

Extreme Values

20° Let X be any subset of $\mathcal{R}$ and let F be a function having domain X. It may happen that there is a number c in X such that, for any number x in X, $F(c) \leq F(x)$. In such a case, we refer to c as a *minimum* number for F and to $F(c)$ as the *minimum* value for F. Similarly, it may happen that there is a number d in X such that, for any number x in X, $F(x) \leq F(d)$. In such a case, we refer to d as a *maximum* number for F and to $F(d)$ as the *maximum* value for F. For illustrations, one may inspect the foregoing examples F_1, F_2, F_3, F_4, and F_5.

21° We refer to the minimum and maximum values as *extreme* values for F and to the minimum and maximum numbers as *extreme* numbers for F. In due course, we will find criteria under which such numbers and values must exist and we will develop techniques for finding them.

Monotone Functions

22° Let X be any subset of $\mathcal{R}$ and let F be a function having domain X. We say that F is *decreasing* iff, for any numbers x and y in X, if $x < y$ then $F(y) \leq F(x)$. We say that F is *increasing* iff, for any numbers x and y in X, if $x < y$ then $F(x) \leq F(y)$. We say that F is *strictly decreasing* iff, for any numbers x and y in X, if $x < y$ then $F(y) < F(x)$. We say that F is *strictly increasing* iff, for any numbers x and y in X, if $x < y$ then $F(x) < F(y)$.

23° We say that F is *monotone* iff F is decreasing or increasing. We say that F is *strictly monotone* iff F is strictly decreasing or strictly increasing.

3 Continuous Functions

1° In this section, we study the condition of Continuity for functions defined on subsets of $\mathcal{R}$ and we prove three basic theorems: the Uniform Continuity Theorem, the Intermediate Value Theorem, and the Extreme Value Theorem. These theorems play critical roles in our subsequent study of differentiable and integrable functions.

Continuity at a

2° Let X be any subset of $\mathcal{R}$ and let F be a function having domain X. Let a be any number in X and let $b := F(a)$. We say that F is *continuous* at a iff, for any number v in $\mathcal{R}^+$, there is some number u in $\mathcal{R}^+$ such that:

$$F(X \cap N_u(a)) \subseteq N_v(b)$$

The displayed condition means that, for any number x in $\mathcal{R}$, if $x \in X \cap N_u(a)$ then $F(x) \in N_v(b)$. Just as well, it means that, for any number x in $\mathcal{R}$, if $x \in X$ and $|x - a| < u$ then $|F(x) - b| < v$.

3° Let us consider the examples described in article 7° of **Section 2**. First, we have $X_1 = \mathcal{R}$ and:

$$F_1(x) = x^2$$

where x is any number in X_1. Let a be any number in X_1. Let us prove that F_1 is continuous at a. To that end, let v be any number in $\mathcal{R}^+$. We must produce a number u in $\mathcal{R}^+$ such that, for any number x in X_1, if $|x - a| < u$ then $|x^2 - a^2| < v$. By property (12) in **Section 1**:

$$|x| - |a| \leq |x - a|$$

Hence, if $|x - a| < u$ then $|x| < u + |a|$ and therefore:

$$|x^2 - a^2| = |x - a||x + a| \leq |x - a|(|x| + |a|) < u(u + 2|a|)$$

Let u be the smaller of 1 and $v/(1 + 2|a|)$. Obviously, $u(u + 2|a|) \leq v$. We conclude that, for any number x in X_1, if $|x - a| < u$ then $|x^2 - a^2| < v$. We have proved that F_1 is continuous at a. ♮

4° Second, we have $X_2 = \mathcal{R}^- \cup \mathcal{R}^+$ and:

$$F_2(x) = \frac{1}{x}$$

where x is any number in X_2. Let a be any number in X_2. Let us prove that F_2 is continuous at a. To that end, let v be any number in $\mathcal{R}^+$. We must produce a number u in $\mathcal{R}^+$ such that, for any number x in X_2, if $|x-a|<u$ then:

$$|\frac{1}{x}-\frac{1}{a}|<v$$

Let u be the smaller of $|a|/2$ and $a^2v/2$. For any number x in X_2, if $|x-a|<u$ then $|a|-|x|<u$, so that $0<|a|/2<|x|$, hence $0<1/|x|<2/|a|$, and therefore:

$$|\frac{1}{x}-\frac{1}{a}|=\frac{1}{|x|}\frac{1}{|a|}|x-a|<\frac{2}{|a|}\frac{1}{|a|}u\leq v$$

We have proved that F_2 is continuous at a. ♮

5° Third, we have $X_3=\mathcal{R}$ and:

$$F_3(x)=\frac{1-x^2}{1+x^2}$$

where x is any number in X_3. Let a be any number in X_3. Let us prove that F_3 is continuous at a. To that end, let v be any number in $\mathcal{R}^+$. We must produce a number u in $\mathcal{R}^+$ such that, for any number x in X_3, if $|x-a|<u$ then:

$$|\frac{1-x^2}{1+x^2}-\frac{1-a^2}{1+a^2}|<v$$

Let $u:=v/2$. For any number x in X_3, if $|x-a|<u$ then:

$$\begin{aligned}|\frac{1-x^2}{1+x^2}-\frac{1-a^2}{1+a^2}| &= 2(\frac{1}{1+x^2})(\frac{1}{1+a^2})|x^2-a^2|\\ &\leq 2|x-a|(\frac{1}{1+x^2})(\frac{1}{1+a^2})(|x|+|a|)\\ &\leq 2|x-a|\\ &< v\end{aligned}$$

because $|x|+|a|\leq 2|x|$ or $|x|+|a|\leq 2|a|$ and because $2|x|/(1+x^2)\leq 1$ and $2|a|/(1+a^2)\leq 1$. We have proved that F_3 is continuous at a. ♮

Continuity and Uniform Continuity on X

6° In the first two of the foregoing examples, our design of u depended not only upon v but also upon a. It could not have been otherwise. However, in the third of the foregoing examples, our design of u depended only upon v. This subtle refinement proves important. It leads to the following distinction between Continuity and Uniform Continuity.

7° Let X be any subset of $\mathcal{R}$ and let F be a function having domain X. We say that F is *continuous* on X iff, for any number a in X, F is continuous at a. We can display the condition of Continuity schematically as follows:

(for all a in X)(for all v in $\mathcal{R}^+$)(there exists u in $\mathcal{R}^+$)(for all x in X)
[if $|x - a| < u$ then $|F(x) - F(a)| < v$]

Just as well:

(for all v in $\mathcal{R}^+$)(for all a in X)(there exists u in $\mathcal{R}^+$)(for all x in X)
[if $|x - a| < u$ then $|F(x) - F(a)| < v$]

In contrast, we can display the stronger condition of Uniform Continuity as follows:

(for all v in $\mathcal{R}^+$)(there exists u in $\mathcal{R}^+$)(for all a in X)(for all x in X)
[if $|x - a| < u$ then $|F(x) - F(a)| < v$]

Just as well:

(for all v in $\mathcal{R}^+$)(there exists u in $\mathcal{R}^+$)(for all x in X)(for all y in X)
[if $|x - y| < u$ then $|F(x) - F(y)| < v$]

We say that F is *uniformly continuous* on X iff, for any number v in $\mathcal{R}^+$, there is some number u in $\mathcal{R}^+$ such that, for any numbers x and y in X, if $|x - y| < u$ then $|F(x) - F(y)| < v$. See the following article 14°.

8° Under this terminology, we may say that F_1 and F_2 are continuous on X (but not uniformly so) while F_3 is uniformly continuous on X.

Basic Properties

9° Let X and Y be any subsets of $\mathcal{R}$, let F be a function having domain X, and let G be a function having domain Y. Let us assume that $F(X) \subseteq Y$. Let a be any number in X, let $b := F(a)$, and let $c := G(b)$, so that:

$$(G \circ F)(a) = G(F(a)) = G(b) = c$$

Let us prove that:

(1) if F is continuous at a and if G is continuous at b then $G \circ F$ is continuous at a

To that end, let w be any number in $\mathcal{R}^+$. Since G is continuous at b, we can introduce a number v in $\mathcal{R}^+$ such that, for any number y in Y, if $|y - b| < v$ then $|G(y) - c| < w$. Since F is continuous at a, we can introduce a number

u in $\mathcal{R}^+$ such that, for any number x in X, if $|x-a|<u$ then $|F(x)-b|<v$. Hence, for any number x in X, if $|x-a|<u$ then:

$$|(G\circ F)(x)-c| = |G(F(x))-c| = |G(y)-c| < w$$

where $y := F(x)$. We conclude that $G\circ F$ is continuous at a. ♮

10° Let X be any subset of $\mathcal{R}$, let F be a function having domain X, let a be a number in X, and let $b := F(a)$. Let us prove that:

(2) if F is continuous at a then $|F|$ is continuous at a

Of course, $|F|(a) = |b|$ and, for any number x in X:

$$|\,|F|(x)-|b|\,| = |\,|F(x)|-|b|\,| \leq |F(x)-b|$$

Let v be any number in $\mathcal{R}^+$. Since F is continuous at a, we can introduce a number u in $\mathcal{R}^+$ such that, for any number x in X, if $|x-a|<u$ then $|F(x)-b|<v$, hence $|\,|F|(x)-|b|\,|<v$. We conclude that $|F|$ is continuous at a. ♮

11° Let X be any subset of $\mathcal{R}$ and let F, G, and H be functions having common domain X. Let c be any number in $\mathcal{R}$. Let a be any number in X. Let us prove that:

(3) if F and G are continuous at a then $F+G$ and $F\cdot G$ are continuous at a

Let $b := F(a)$ and $c := G(a)$, so that $(F+G)(a) = b+c$ and $(F\cdot G)(a) = b\cdot c$. Let v be any number in $\mathcal{R}^+$. Since F is continuous at a, we can introduce a number u' in $\mathcal{R}^+$ such that, for any number x in X, if $|x-a|<u'$ then $|F(x)-b|<v/2$. Since G is continuous at a, we can introduce a number u'' in $\mathcal{R}^+$ such that, for any number x in X, if $|x-a|<u''$ then $|G(x)-c|<v/2$. Let u be the smaller of u' and u''. For any number x in X, if $|x-a|<u$ then:

$$\begin{aligned}|(F+G)(x)-(b+c)| &= |(F(x)-b)+(G(x)-c)|\\ &\leq |F(x)-b|+|G(x)-c|\\ &< (v/2)+(v/2)\\ &= v\end{aligned}$$

We conclude that $F+G$ is continuous at a. Again, let v be any number in $\mathcal{R}^+$. Let r', s, and r'' be any numbers in $\mathcal{R}^+$ for which:

$$r'|c|\leq 1,\quad s := |b|+r'\cdot(v/2),\quad r''s\leq 1$$

Since F is continuous at a, we can introduce a number u' in $\mathcal{R}^+$ such that, for any number x in X, if $|x-a|<u'$ then $|F(x)-b|<r'\cdot(v/2)$, so that:

$$|F(x)|<|b|+r'\cdot(v/2)=s$$

Since G is continuous at a, we can introduce a number u'' in $\mathcal{R}^+$ such that, for any number x in X, if $|x-a|<u''$ then $|G(x)-c|<r''\cdot(v/2)$. Let u be the smaller of u' and u''. For any number x in X, if $|x-a|<u$ then:

$$\begin{aligned}
|(F\cdot G)(x)-(b\cdot c)| &= |F(x)\cdot G(x)-b\cdot c|\\
&= |F(x)\cdot G(x)-F(x)\cdot c+F(x)\cdot c-b\cdot c|\\
&= |F(x)\cdot(G(x)-c)+(F(x)-b)\cdot c|\\
&\leq |F(x)|\cdot|G(x)-c|+|F(x)-b|\cdot|c|\\
&< s\cdot r''\cdot(v/2)\ +\ r'\cdot(v/2)\cdot|c|\\
&\leq (v/2)+(v/2)\\
&= v
\end{aligned}$$

We conclude that $F\cdot G$ is continuous at a. ♮

12° If F is constant with constant value c then of course F is continuous at a. If $G=H$ as well then $F\cdot G=c\cdot H$. We conclude that:

(4) if H is continuous at a then $c\cdot H$ is continuous at a

13° If G is continuous at a and if $G(X)\subseteq X_2$ then $1/G$ is continuous at a, because $1/G=F_2\circ G$. We conclude that:

(5) if F and G are continuous at a and if $G(X)\subseteq X_2$ then F/G is continuous at a

because $F/G=F\cdot(1/G)=F\cdot(F_2\circ G)$. See articles 4° and 9°.

The Uniform Continuity Theorem

14° In general, the condition of uniform continuity is stronger than the condition of continuity unimproved. However, in certain contexts, the two conditions are equivalent. This equivalence yields important consequences for the study of Integration.

15° Let X be any subset of $\mathcal{R}$ and let F be a function having domain X. Let us prove that:

(6) if X is a closed finite interval in $\mathcal{R}$ and if F is continuous on X then F is uniformly continuous on X

We can introduce numbers a and b in $\mathcal{R}$ such that $a < b$ and $X = [a, b]$. Let v be any number in $\mathcal{R}^+$. We must produce a number u in $\mathcal{R}^+$ such that, for any numbers x and y in $[\,a, b\,]$, if $|x - y| < u$ then $|F(x) - F(y)| < v$. To that end, let C be the subset of $[a, b]$ consisting of all numbers c in $[a, b]$ for which:

($\circ$) there exists a number u_c in $\mathcal{R}^+$ such that, for any numbers x and y in $[a, c]$, if $|x - y| < u_c$ then $|F(x) - F(y)| < v$

Obviously, $C \neq \emptyset$ because $a \in C$. Obviously, $C^* \neq \emptyset$ because $b \in C^*$. By the Least Upper Bound Principle, we can introduce the smallest number d in C^*. Obviously, $a \leq d \leq b$. We plan to prove that $d \in C$ and $d = b$. That done, we may introduce a number $u := u_b$ in $\mathcal{R}^+$ such that, for any numbers x and y in $[\,a, b\,]$, if $|x - y| < u$ then $|F(x) - F(y)| < v$. In turn, we may conclude that F is uniformly continuous on X. Let us prove that $d \in C$ and $d = b$. Since F is continuous at d, we can introduce a number t in $\mathcal{R}^+$ such that, for any number z in $[a, b]$, if $|z - d| < t$ then $|F(z) - F(d)| < v/2$. Since $d - (t/2) \notin C^*$, we can introduce a number c in C for which $d - (t/2) < c \leq d$. By definition, we can introduce a number u_c in $\mathcal{R}^+$ such that, for any numbers x and y in $[a, c]$, if $|x - y| < u_c$ then $|F(x) - F(y)| < v$. Let e be any number in $[\,d, d + (t/2)) \cap [\,a, b]$. To illustrate the current argument, let us introduce the following *counterfactual* figure.

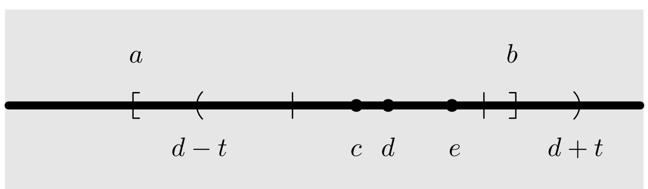


Figure 4: Contrary to Fact

Obviously, $d \leq e \leq b$. We plan to prove that $e \in C$. That done, we may infer that $e \leq d$. In turn, we may conclude that $d = e$, so that $d \in C$, and we may conclude that $(d, d + (t/2)) \cap [\,a, b\,] = \emptyset$, so that $d = b$. Let us prove that $e \in C$. Let u_e be the smaller of $t/2$ and u_c. For any numbers x and y in $[\,a, e\,]$, if $|x - y| < u_e$ and if both $x \in [\,a, c\,]$ and $y \in [\,a, c\,]$ then of course $|F(x) - F(y)| < v$. If either $x \in (c, e]$ or $y \in (c, e]$ then both $x \in (d - t, d + t)$ and $y \in (d - t, d + t)$, so that $|F(x) - F(y)| = |(F(x) - d) - (F(y) - d)| \leq |F(x) - d| + |F(y) - d| < (v/2) + (v/2) = v$. We conclude that $e \in C$. $\natural$

16° We refer to property (6) as the Uniform Continuity Theorem.

17° Let X be any subset of $\mathcal{R}$ and let F be a function having domain X. Let Y be the range of F. Let us prove that:

(7) if X is an interval in $\mathcal{R}$ and if F is continuous on X then Y is an interval in $\mathcal{R}$

To that end, let c, d, and y be any numbers in $\mathcal{R}$ such that $c \in Y$, $d \in Y$, and $c < y < d$. We must prove that $y \in Y$. Of course, we can introduce numbers a and b in X such that $F(a) = c$ and $F(b) = d$. Obviously, $a \neq b$. However, it may happen that $a < b$ or it may happen that $b < a$. Let us assume that $a < b$. Let R be the subset of $[\,a, b\,]$ consisting of all numbers r in $[\,a, b\,]$ for which $F(r) < y$. Obviously, $R \neq \emptyset$ because $a \in R$. Obviously, $R^* \neq \emptyset$ because $b \in R^*$. By the Least Upper Bound Principle, we can introduce the smallest number s in R^*. Obviously, $a \leq s \leq b$. We contend that $F(s) = y$. Let us suppose that $F(s) < y$. Under this supposition, we could introduce the number $v := y - F(s)$ in $\mathcal{R}^+$. Since F is continuous at s, we could introduce a number u in $\mathcal{R}^+$ such that, for any number x in X, if $|x - s| < u$ then $|F(x) - F(s)| < v$, so that $F(x) < y$. We could then introduce numbers r in $(s, s + u) \cap [\,a, b\,]$ for which the following contradictory properties hold: $s < r$ and $F(r) < y$. We infer that $y \leq F(s)$. Let us suppose that $y < F(s)$. Under this supposition, we could introduce the number $v := F(s) - y$ in $\mathcal{R}^+$. Since F is continuous at s, we could introduce a number u in $\mathcal{R}^+$ such that, for any number x in X, if $|x - s| < u$ then $|F(x) - F(s)| < v$, so that $y < F(x)$. We could then introduce numbers r in $(s - u, s) \cap [\,a, b\,]$ for which the following contradictory properties hold: $r \in R$ and $y < F(r)$. We conclude that $y = F(s)$. We are half way done. Let us assume that $b < a$. Let R be the subset of $[\,b, a\,]$ consisting of all numbers r in $[\,b, a\,]$ for which $y < F(r)$. Obviously, $R \neq \emptyset$ because $b \in R$. Obviously, $R^* \neq \emptyset$ because $a \in R^*$. By the Least Upper Bound Principle, we can introduce the smallest number s in R^*. Obviously, $b \leq s \leq a$. We contend that $F(s) = y$. Let us suppose that $F(s) < y$. Under this supposition, we could introduce the number $v := y - F(s)$ in $\mathcal{R}^+$. Since F is continuous at s, we could introduce a number u in $\mathcal{R}^+$ such that, for any number x in X, if $|x - s| < u$ then $|F(x) - F(s)| < v$, so that $F(x) < y$. We could then introduce numbers r in $(s - u, s) \cap [\,b, a\,]$ for which the following contradictory properties hold: $r \in R$ and $F(r) < y$. We infer that $y \leq F(s)$. Let us suppose that $y < F(s)$. Under this supposition, we could introduce the number $v := F(s) - y$ in $\mathcal{R}^+$. Since F is continuous at s, we could introduce a number u in $\mathcal{R}^+$ such that, for any number x in X, if $|x - s| < u$ then $|F(x) - F(s)| < v$, so that $y < F(x)$. We could then introduce numbers r in $(s, s + u) \cap [\,b, a\,]$ for which the following contradictory properties hold: $s < r$ and $y < F(r)$. We conclude that $y = F(s)$. $\natural$

18° We refer to property (7) as the Intermediate Value Theorem.

The Extreme Value Theorem

19° Let X be any subset of $\mathcal{R}$ and let F be a function having domain X. Let Y be the range of F. Let us prove that:

(8) if X is a closed finite interval in $\mathcal{R}$ and if F is continuous on X then Y is a closed finite interval in $\mathcal{R}$, or else trivial

Obviously, Y is a nonempty trivial interval in $\mathcal{R}$ iff F is constant. Of course, we can ignore this distracting case. Let us introduce numbers a and b in $\mathcal{R}$ such that $a < b$ and $X = [a, b]$. We must produce numbers c and d in $\mathcal{R}$ such that $c < d$ and $Y = F([a, b]) = [c, d]$. By the Intermediate Value Theorem, Y is an interval in $\mathcal{R}$. By the Uniform Continuity Theorem, we can introduce a number u in $\mathcal{R}^+$ such that, for any numbers x and y in $[a, b]$, if $|x - y| < u$ then $|F(x) - F(y)| < 1$, so that:

$$F(y) - 1 < F(x) < F(y) + 1$$

By the Principle of Archimedes, we can introduce an integer k in $\mathcal{Z}^+$ for which $(b - a)/k < u$. We are led to define the finite chain:

$$a = y_0 < y_1 < y_2 < \cdots < y_{k-1} < y_k = b$$

of numbers in $[a, b]$, where:

$$y_j := a + \frac{j}{k}(b - a)$$

Clearly, for any number x in $[a, b]$, there is some integer j $(1 \leq j \leq k)$ such that $x \in [y_{j-1}, y_j]$. Hence, $|x - y_j| < u$ and $|F(x) - F(y_j)| < 1$. We infer that:

$$F(X) \subseteq \bigcup_{j=1}^{k} (F(y_j) - 1, F(y_j) + 1)$$

We conclude that F is bounded, so that Y is a finite interval in $\mathcal{R}$. Hence, we can introduce numbers c and d in $\mathcal{R}$ such that $c < d$ and Y equals:

$$(c, d),\ [c, d],\ [c, d),\ \text{or } (c, d]$$

Let us suppose that $c \notin Y$. Under this supposition, we could introduce the function G having domain $[a, b]$, defined as follows:

$$G(x) := \frac{1}{F(x) - c}$$

where x is any number in $[a, b]$. Obviously, G would be continuous on $[a, b]$. However, G would be unbounded, contradicting the conclusion just drawn for F. We infer that $c \in Y$. Let us suppose that $d \notin Y$. Under this supposition, we could introduce the function H having domain $[a, b]$, defined as follows:

$$H(x) := \frac{1}{d - F(x)}$$

where x is any number in $[a, b]$. Obviously, H would be continuous on $[a, b]$. However, H would be unbounded, contradicting the conclusion just drawn for F. We infer that $d \in Y$. We conclude that $Y = [c, d]$. ♮

20° We refer to property (8) as the Extreme Value Theorem. Obviously, c is the smallest value of F and d is the largest value of F.

4 Differentiation

1° In this section, we investigate the condition of Differentiablity for functions defined on nontrivial intervals in $\mathcal{R}$. Subject to this condition, one can measure the *rate of change* of the values of a function. After establishing the computational properties of differentiable functions, we prove the basic Mean Value Theorem. We then show that we can obtain important information about a function from specific properties of its derivative. In particular, we study the (related) problems of sketching the graph of a function and of determining the extreme values of a function. Finally, we prove the grand Theorem of Taylor.

Differentiability at a

2° Let X be any nontrivial interval in $\mathcal{R}$ and let F be a function having domain X. Let a be any number in X. Let:

$$X_a := X \backslash \{a\}$$

be the subset of X defined by excising a from X. Thus, for any number x in $\mathcal{R}$, $x \in X_a$ iff $x \in X$ and $x \neq a$. Let F_a be the function having domain X_a, defined as follows:

$$F_a(x) := \frac{1}{x-a}(F(x) - F(a))$$

where x is any number in X_a. We refer to F_a as the *quotient function* for F at a.

3° We say that F is *differentiable* at a iff there is some number q in $\mathcal{R}$ such that, for any number v in $\mathcal{R}^+$, there is some number u in $\mathcal{R}^+$ such that:

$$F_a(X_a \cap N_u(a)) \subseteq N_v(q)$$

The displayed condition means that, for any number x in $\mathcal{R}$, if $x \in X_a \cap N_u(a)$ then $F_a(x) \in N_v(q)$. Just as well, it means that, for any number x in $\mathcal{R}$, if $x \in X$ and $0 < |x - a| < u$ then $|F_a(x) - q| < v$. That is:

$$\left| \frac{1}{x-a}(F(x) - F(a)) - q \right| < v$$

See Figure 5.

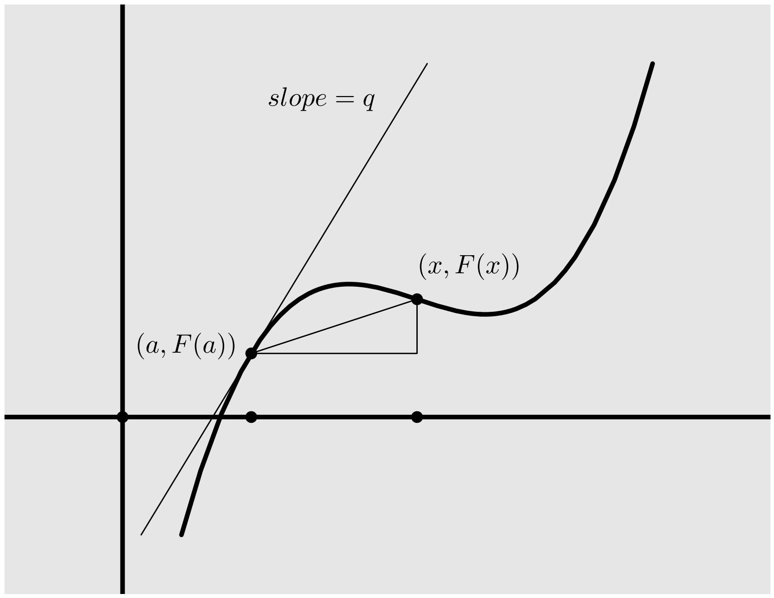


Figure 5: A Derivative

4° Let us show that there may be no such number q. For example, let $X := \mathcal{R}$, $a := 0$, and:

$$F(x) := |x|$$

where x is any number in X. Clearly:

$$F_0(x) = \frac{1}{x}|x| = \begin{cases} -1 & \text{if } x < 0 \\ 1 & \text{if } 0 < x \end{cases}$$

where x is any number in X_0. Obviously, for this example, there is no such number q.

5° Let us prove that if there is such a number q then it is unique. To that end, let r and s be numbers in $\mathcal{R}$ which satisfy the stated condition. Let v be any number in $\mathcal{R}^+$. We can introduce a number u' in $\mathcal{R}^+$ such that, for any number x in X, if $0 < |x - a| < u'$ then $|F_a(x) - r| < (v/2)$. In turn, we can introduce a number u'' in $\mathcal{R}^+$ such that, for any number x in X, if

$0 < |x - a| < u''$ then $|F_a(x) - s| < (v/2)$. Let u be the smaller of u' and u''. Let x be any number in $X_a \cap N_u(a)$. Clearly:

$$\begin{aligned} |r - s| &= |(r - F_a(x)) + (F_a(x) - s)| \\ &\leq |r - F_a(x)| + |F_a(x) - s| \\ &< (v/2) + (v/2) \\ &= v \end{aligned}$$

We conclude that $r = s$. ♮

6° Let F be differentiable at a. We refer to the uniquely determined number q as the *derivative* of F at a and we denote it by:

$$F'(a)$$

7° For illustration, let us consider the first among our standing examples. We have $X_1 = \mathcal{R}$ and:

$$F_1(x) = x^2$$

where x is any number in X_1. Let a be any number in X_1. Let us prove that F_1 is differentiable at a and that:

$$F_1'(a) = 2a$$

To that end, let v be any number in $\mathcal{R}^+$. We must produce a number u in $\mathcal{R}^+$ such that, for any number x in X_1, if $0 < |x - a| < u$ then:

$$\left| \frac{1}{x - a}(x^2 - a^2) - 2a \right| < v$$

Let $u := v$. Obviously, for any number x in X_1, if $0 < |x - a| < u$ then:

$$\left| \frac{1}{x - a}(x^2 - a^2) - 2a \right| = \left| \frac{1}{x - a}(x - a)(x + a) - 2a \right| = \left| x - a \right| < v$$

We conclude that F_1 is differentiable at a and that $F_1'(a) = 2a$. ♮

8° Now let us consider the second among our standing examples. We have $X_2 = \mathcal{R}^- \cup \mathcal{R}^+$ and:

$$F_2(x) = \frac{1}{x}$$

where x is any number in X_2. Technically, we should restrict attention either to the interval $\mathcal{R}^-$ or to the interval $\mathcal{R}^+$. In effect, we will argue both cases at once. Let a be any number in X_2. Let us prove that F_2 is differentiable at a and that:

$$F_2'(a) = -\frac{1}{a^2}$$

To that end, let v be any number in $\mathcal{R}^+$. We must produce a number u in $\mathcal{R}^+$ such that, for any number x in X_1, if $0 < |x-a| < u$ then:

$$\Big| \frac{1}{x-a}(\frac{1}{x} - \frac{1}{a}) + \frac{1}{a^2} \Big| < v$$

Let us look back to article 4° in **Section 3**. Let u be the smaller of $|a|/2$ and $|a|^3 v/2$. For any number x in X_2, if $0 < |x-a| < u$ then $0 < 1/|x| < 2/|a|$ and:

$$\begin{aligned}
\Big| \frac{1}{x-a}(\frac{1}{x} - \frac{1}{a}) + \frac{1}{a^2} \Big| &= \Big| \frac{1}{x-a}(\frac{a-x}{xa}) + \frac{1}{a^2} \Big| \\
&= \Big| \frac{1}{xa} - \frac{1}{a^2} \Big| \\
&= \frac{1}{|a|} \Big| \frac{1}{x} - \frac{1}{a} \Big| \\
&< \frac{1}{|a|} \frac{2}{|a|} \frac{1}{|a|} |x-a| \\
&< v
\end{aligned}$$

We conclude that F_2 is differentiable at a and that $F_2'(a) = -1/a^2$. ♮

Differentiability on X

9° Let X be any nontrivial interval in $\mathcal{R}$ and let F be a function having domain X. We say that F is *differentiable* on X iff, for any number a in X, F is differentiable at a. Given that F is differentiable on X, we obtain a new function having domain X, which, by definition, assigns to each number a in X the value $F'(a)$. Of course, we denote the new function by F' and we refer to it as the *derivative* of F.

10° For the examples F_1 and F_2, we have:

$$F_1'(x) = 2x$$

where x is any number in X_1, and:

$$F_2'(x) = -\frac{1}{x^2}$$

where x is any number in X_2.

Continuous Extensions at a

11° Let us reformulate the condition of differentiability for F at a in terms of the *question* of continuity for F_a at a. By this maneuver, we will be able to design efficient proofs of the properties of differentiable functions, soon to follow.

12° Let X be a nontrivial interval in $\mathcal{R}$, let a be any number in X, and let F_a be a function having domain X_a. Let G be a function having domain X. We say that G is a *continuous extension* of F_a at a iff:

(∘) for each number x in X_a, $G(x) = F_a(x)$
(∘) G is continuous at a

In effect, G supplies the "missing value" of F_a at a. Moreover, the supplied value is coherent with the given values of F_a.

13° With reference to articles 4° and 5°, let us note that a continuous extension of F_a at a may not exist. However, if it does exist then it is unique; we denote it by $\bar{F}_a$.

14° Let X be any nontrivial interval in $\mathcal{R}$ and let F be a function having domain X. Let a be any number in X. By comparison of definitions, we find that F is differentiable at a iff F_a admits a continuous extension at a. Obviously, the supplied value for F_a at a would be $F'(a)$:

$$\bar{F}_a(a) = F'(a)$$

15° Let us consider the examples F_1 and F_2. We find that $(F_1)_a(x) = x + a$, where x is any number in $(X_1)_a$, and that $(F_2)_a(x) = -1/ax$, where x is any number in $(X_2)_a$. In both cases, these quotient functions admit continuous extensions at a:

$$(\overline{F_1})_a(x) = x + a \quad \text{and} \quad (\overline{F_2})_a(x) = -1/ax$$

Hence, $F_1'(a) = (\overline{F_1})_a(a) = 2a$ and $F_2'(a) = (\overline{F_2})_a(x) = -1/a^2$. See articles 7° and 8°.

Differentiable Functions are Continuous

16° Let X be any nontrivial interval in $\mathcal{R}$ and let F be a function having domain X. Let a be any number in X. Let us prove that:

(1) if F is differentiable at a then F is continuous at a

To that end, let us introduce the continuous extension $\bar{F}_a$ of F_a at a. We have:

$$F(x) = F(a) + \bar{F}_a(x) \cdot (x - a)$$

where x is any number in X. Since $\bar{F}_a$ is continuous at a, F is continuous at a. ♮

17° Obviously, if F is differentiable on X then F is continuous on X.

The Composition Rule

18° Let X and Y be any subsets of $\mathcal{R}$, let F be a function having domain X, and let G be a function having domain Y. Let us assume that $F(X) \subseteq Y$. Let a be any number in X and let $b := F(a)$. Let us prove that:

(2) if F is differentiable at a and if G is differentiable at b then $G \circ F$ is differentiable at a and:

$$(G \circ F)'(a) = G'(b) \cdot F'(a)$$

To that end, let us introduce the continuous extension $\bar{F}_a$ of F_a at a and the continuous extension $\bar{G}_b$ of G_b at b. We have:

$$\begin{aligned} (G \circ F)_a(x) &= \frac{1}{x-a}\Big((G \circ F)(x) - (G \circ F)(a)\Big) \\ &= \frac{1}{x-a}\Big(G(F(x)) - G(F(a))\Big) \\ &= \frac{1}{x-a}\,\bar{G}_b(F(x)) \cdot \Big(F(x) - F(a)\Big) \\ &= \bar{G}_b(F(x)) \cdot F_a(x) \end{aligned}$$

where x is any number in X_a. Clearly, $(\bar{G}_b \circ F) \cdot \bar{F}_a$ is continuous at a. Hence, $(G \circ F)_a$ admits a continuous extension at a. In fact:

$$(\overline{G \circ F})_a = (\bar{G}_b \circ F) \cdot \bar{F}_a$$

We conclude that:

$$(G \circ F)'(a) = (\overline{G \circ F})_a(a) = \bar{G}_b(b) \cdot \bar{F}_a(a) = G'(b) \cdot F'(a)$$

♮

19° We refer to the rule expressed in property (2) as the Composition Rule.

20° Obviously, if F is differentiable on X and G is differentiable on Y then $G \circ F$ is differentiable on X and:

$$(G \circ F)' = (G' \circ F) \cdot F'$$

Computational Properties

21° Let X be any subset of $\mathcal{R}$ and let F, G, and H be functions having common domain X. Let c be any number in $\mathcal{R}$. Let a be any number in X. Let us prove that:

(3) if F and G are differentiable at a then $F+G$ and $F \cdot G$ are differentiable at a and:

$$(F+G)'(a) = F'(a) + G'(a)$$

and:

$$(F \cdot G)'(a) = F(a) \cdot G'(a) + F'(a) \cdot G(a)$$

We have:

$$\begin{aligned}
(F+G)_a(x) &= \frac{1}{x-a}\Big((F+G)(x) - (F+G)(a)\Big) \\
&= \frac{1}{x-a}\Big((F(x) - F(a)) + (G(x) - G(a))\Big) \\
&= \frac{1}{x-a}\Big(F(x) - F(a)\Big) + \frac{1}{x-a}\Big(G(x) - G(a)\Big) \\
&= F_a(x) + G_a(x)
\end{aligned}$$

where x is any number in X_a. Clearly, $(F+G)_a$ admits a continuous extension at a. In fact:

$$(\overline{F+G})_a = \bar{F}_a + \bar{G}_a$$

Hence:

$$(F+G)'(a) = (\overline{F+G})_a(a) = \bar{F}_a(a) + \bar{G}_a(a) = F'(a) + G'(a)$$

Moreover:

$$\begin{aligned}
(F \cdot G)_a(x) &= \frac{1}{x-a}\Big((F \cdot G)(x) - (F \cdot G)(a)\Big) \\
&= \frac{1}{x-a}\Big(F(x) \cdot (G(x) - G(a)) + (F(x) - F(a)) \cdot G(a)\Big) \\
&= F(x) \cdot \frac{1}{x-a}(G(x) - G(a)) + \frac{1}{x-a}(F(x) - F(a)) \cdot G(a) \\
&= F(x) \cdot G_a(x) + F_a(x) \cdot G(a)
\end{aligned}$$

where x is any number in X_a. Clearly, $(F \cdot G)_a$ admits a continuous extension at a. In fact:

$$(\overline{F \cdot G})_a = F \cdot \bar{G}_a + \bar{F}_a \cdot G(a)$$

Hence:

$$(F \cdot G)'(a) = (\overline{F \cdot G})_a(a) = F(a) \cdot \bar{G}_a(a) + \bar{F}_a(a) \cdot G(a) = F(a) \cdot G'(a) + F'(a) \cdot G(a)$$

♮

22° We refer to the rules expressed in property (3) as the Sum Rule and the Product Rule.

23° Obviously, if F and G are differentiable on X then $F + G$ and $F \cdot G$ are differentiable on X and:

$$(F + G)' = F' + G'$$

and:

$$(F \cdot G)' = F \cdot G' + F' \cdot G$$

24° If F is constant with constant value c then of course F is differentiable at a and $F'(a) = 0$. If $G = H$ as well then $F \cdot G = c \cdot H$. By the Product Rule, we conclude that:

(4) if H is differentiable at a then $c \cdot H$ is differentiable at a and:

$$(c \cdot H)'(a) = c \cdot H'(a)$$

25° If G is differentiable at a and if $G(X) \subseteq X_2$ then $1/G = F_2 \circ G$. By the Composition Rule and by article 8°, $1/G$ is differentiable at a and:

$$(\frac{1}{G})'(a) = F_2'(G(a)) \cdot G'(a) = -\frac{G'(a)}{G^2(a)}$$

We conclude that:

(5) if F and G are differentiable at a and if $G(X) \subseteq X_2$ then F/G is differentiable at a and:

$$(\frac{F}{G})'(a) = \frac{G(a) \cdot F'(a) - F(a) \cdot G'(a)}{G^2(a)}$$

because $F/G = F \cdot (1/G)$ and, by the Product Rule:

$$(F \cdot \frac{1}{G})'(a) = -F(a)\frac{G'(a)}{G^2(a)} + F'(a) \cdot \frac{1}{G(a)}$$

26° We refer to the rule expressed in property (5) as the Quotient Rule.

27° Obviously, if F and G are differentiable on X and $G(X) \subseteq X_2$ then F/G is differentiable on X and:

$$(\frac{F}{G})' = \frac{G \cdot F' - F \cdot G'}{G^2}$$

Notation

28° On occasion, we denote the derivative of F at a not by $F'(a)$ but by:

$$\frac{d}{dx}F(x)\Big|_{x=a}$$

and the derivative of F not by F' but by:

$$\frac{d}{dx}F(x)$$

Of course, we might choose some other symbol, such as y, in place of x. By this notation, we gain a certain flexibility. We can compute derivatives without engaging in the sometimes circuitous process of "naming" the function. For instance, the following expressions are now clear and meaningful:

$$\frac{d}{dx}x^2 = 2x$$

and:

$$\frac{d}{dx}x^2\Big|_{x=3} = 6$$

29° In the foregoing expressions, the unnamed function is F_1. Let us apply the flexible notation to the functions F_2, F_3, F_4, and F_5 as well:

$$\frac{d}{dx}x^{-1} = -x^{-2}$$

$$\frac{d}{dx}\frac{1-x^2}{1+x^2} = \frac{(1+x^2)(-2x) - (1-x^2)(2x)}{(1+x^2)^2} = -\frac{4x}{(1+x^2)^2}$$

$$\frac{d}{dy}(y + y^{-1}) = 1 - y^{-2}$$

$$\frac{d}{dx}\frac{1-x}{1+x} = \frac{(1+x)(-1) - (1-x)(1)}{(1+x)^2} = -\frac{2}{(1+x)^2}$$

Positive Powers

30° By the Product Rule and by Mathematical Induction, one can easily prove that:

(6) for any integer k in $\mathcal{Z}^+$:

$$\frac{d}{dx}x^k = k \cdot x^{k-1}$$

To that end, one would apply the following inductive pattern of calculation:

$$\frac{d}{dx}x = 1$$

and:

$$\frac{d}{dx}x^{k+1} = \frac{d}{dx}(x \cdot x^k) = (\frac{d}{dx}x) \cdot x^k + x \cdot \frac{d}{dx}x^k = 1 \cdot x^k + x \cdot k \cdot x^{k-1} = (k+1) \cdot x^k$$

$F'(c) = 0$ at Extreme Numbers

31° Let X be any nontrivial open interval in $\mathcal{R}$ and let F be a function having domain X. Let F be differentiable on X. Let us prove that:

(7) for any number ϵ in X, if ϵ is an extreme number for F then $F'(\epsilon) = 0$

Let us suppose that $F'(\epsilon) < 0$. Under this supposition, we could introduce a number v in $\mathcal{R}^+$ for which $F'(\epsilon) + v < 0$. Since F is differentiable at ϵ, we could introduce a number u in $\mathcal{R}^+$ such that, for any number x in X, if $0 < |x - \epsilon| < u$ then:

$$\frac{1}{x-\epsilon}(F(x) - F(\epsilon)) - F'(\epsilon) \leq \left| \frac{1}{x-\epsilon}(F(x) - F(\epsilon)) - F'(\epsilon) \right| < v$$

so that:

$$\frac{1}{x-\epsilon}(F(x) - F(\epsilon)) < 0$$

Finally, we could introduce numbers x' and x'' in X for which $\epsilon - u < x' < \epsilon < x'' < \epsilon + u$. We would find that $F(x'') < F(\epsilon) < F(x')$, contrary to the assumption that ϵ is an extreme number for F. In turn, let us suppose that $0 < F'(\epsilon)$. Under this supposition, we could introduce a number v in $\mathcal{R}^+$

for which $0 < F'(\epsilon) - v$. Since F is differentiable at ϵ, we could introduce a number u in $\mathcal{R}^+$ such that, for any number x in X, if $0 < |x - \epsilon| < u$ then:

$$F'(\epsilon) - \frac{1}{x-\epsilon}(F(x) - F(\epsilon)) \leq \left| \frac{1}{x-\epsilon}(F(x) - F(\epsilon)) - F'(\epsilon) \right| < v$$

so that:

$$0 < \frac{1}{x-\epsilon}(F(x) - F(\epsilon))$$

Finally, we could introduce numbers x' and x'' in X for which $\epsilon - u < x' < \epsilon < x'' < \epsilon + u$. We would find that $F(x') < F(\epsilon) < F(x'')$, contrary to the assumption that ϵ is an extreme number for F. We conclude that $F'(\epsilon) = 0$. ♮

32° One should note that, in the foregoing argument, our introduction of x' and x'' depended upon the assumption that X is an open interval.

The Mean Value Theorem

33° Let X be any nontrivial interval in $\mathcal{R}$ and let F be a function having domain X. Let F be differentiable on X. Let a and b be any numbers in X such that $a \neq b$. For any number c in X, we say that c lies *between* a and b iff $a < c < b$ if $a < b$ and $b < c < a$ if $b < a$. Let us prove that:

(8) there is a number c lying between a and b such that:

$$F'(c) = \frac{1}{b-a}(F(b) - F(a))$$

Obviously, property (8) is symmetric in a and b. Hence, without loss, we may assume that $a < b$. Let us introduce the auxiliary function G having domain $[a, b]$, defined as follows:

$$G(x) = \big(F(x) - F(a)\big) - \frac{1}{b-a}\big(F(b) - F(a)\big) \cdot \big(x - a\big)$$

where x is any number in $[\,a, b\,]$. Clearly, $G(a) = 0 = G(b)$. Moreover, G is differentiable on X and:

$$G'(x) = F'(x) - \frac{1}{b-a}\big(F(b) - F(a)\big)$$

where x is any number in $[\,a, b\,]$. Let us prove that there is a number c in (a, b) such that $G'(c) = 0$. That done, we may conclude that:

$$F'(c) = \frac{1}{b-a}\big(F(b) - F(a)\big)$$

By the Extreme Value Theorem, we can introduce numbers r and s in $\mathcal{R}$ such that:

$$G([a,b]) = [r,s]$$

Let p and q be numbers in $[\,a,b\,]$ for which $G(p) = r$ and $G(q) = s$. By definition, p and q are extreme numbers for G. It may happen that both p and q are endpoints of $[a,b]$ or it may not. If so then $r = 0 = s$. In that case, G is constantly 0 and we may take c to be any number in (a,b). If not then $p \in (a,b)$ or $q \in (a,b)$. In that case, by property (7), we may take c to be p if $p \in (a,b)$ and we may take c to be q if $q \in (a,b)$. $\natural$

34° We refer to property (8) as the Mean Value Theorem. From this theorem, we can derive many important consequences. For instance, let us assume that F is differentiable on X and that F' is constantly 0. By the Mean Value Theorem, we infer that F itself must be constant. This simple property figures at several points in **Section 7**.

Sketching Graphs of Functions

35° Let X be any nontrivial open interval in $\mathcal{R}$ and let F be a function having domain X. Let F be differentiable on X. Let us prove that:

(9) if $F'(X) \subseteq \mathcal{R}^-$ then F is strictly decreasing and if $F'(X) \subseteq \mathcal{R}^+$ then F is strictly increasing

Let us assume that $F'(X) \subseteq \mathcal{R}^-$. Let a and b be any numbers in X such that $a < b$. By the Mean Value Theorem, we can introduce a number c in (a,b) such that:

$$F(b) - F(a) = F'(c)(b-a)$$

Clearly, $F(b) < F(a)$, because $F'(c) < 0$. In turn, let us assume that $F'(X) \subseteq \mathcal{R}^+$. Let a and b be any numbers in X such that $a < b$. By the Mean Value Theorem, we can introduce a number c in (a,b) such that:

$$F(b) - F(a) = F'(c)(b-a)$$

Clearly, $F(a) < F(b)$, because $0 < F'(c)$. $\natural$

36° Property (9) yields a method for sketching the graphs of differentiable functions. Given a differentiable function F having domain X, one proceeds to determine the subintervals of X on which the values of F' are negative and the subintervals of X on which the values of F' are positive. On such a subinterval, the restriction of F would be, correspondingly, strictly decreasing and strictly increasing. The found intervals meet at *critical* numbers ϵ, for which $F'(\epsilon) = 0$. See article 41°. The ordered pairs $(\epsilon, F(\epsilon))$ serve as reference positions from which one can sketch the graph of F.

37° For illustration, let us recall the fourth function in our stock of examples. We have $X_4 = \mathcal{R}^- \cup \mathcal{R}^+$ and:

$$F_4(x) = x + \frac{1}{x}$$

where x is any number in X_4. Technically, we should restrict attention either to the interval $\mathcal{R}^-$ or to the interval $\mathcal{R}^+$. In effect, we will argue both cases at once. Clearly:

$$F_4'(x) = 1 - \frac{1}{x^2}$$

where x is any number in X_4. Hence, $F_4'(x) < 0$ iff $0 < x^2 < 1$, $F_4'(x) = 0$ iff $x^2 = 1$, and $0 < F_4'(x)$ iff $1 < x^2$. The critical numbers are -1 and 1. Obviously, $F(-1) = -2$ and $F(1) = 2$. The graph of F has the form sketched in Figure 6.

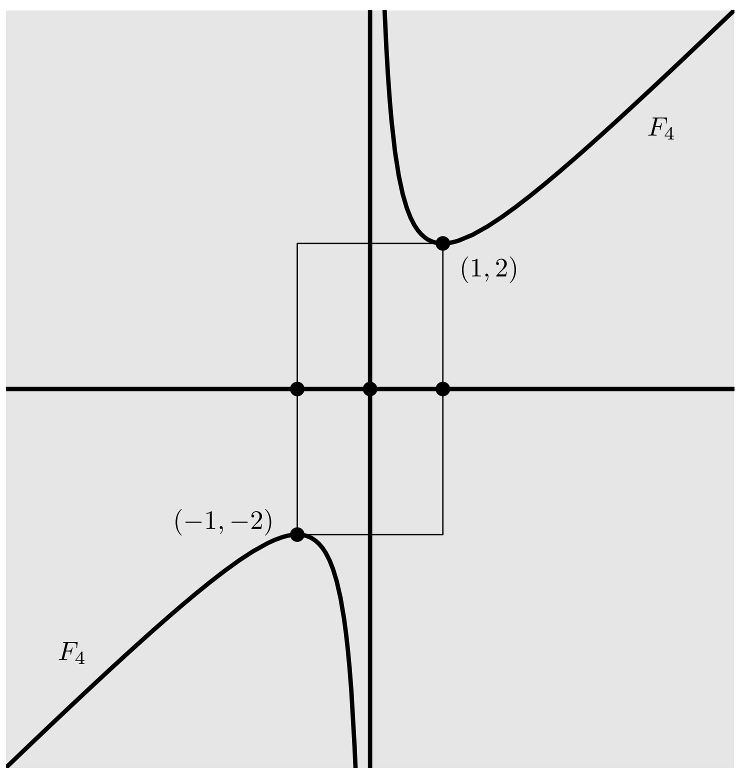


Figure 6: Sketch of a Graph

Extreme Value Problems

38° Let X be any nontrivial interval in $\mathcal{R}$ and let F be a function having domain X. Very often, one wants to find the extreme numbers ϵ for F and the corresponding extreme values $F(\epsilon)$. For that purpose, one asks two questions:

(∘) Do such numbers exist?
(∘) If so, where are they?

The Extreme Value Theorem provides a useful response to the first question. One should note that it requires X to be closed and finite. Property (7) provides a useful response to the second question. One should note that it requires X to be open. Moreover, one should take care to distinguish between critical numbers and extreme numbers for F. In general, we must say that the most reliable responses to the two questions descend from a clear sketch of the graph of F.

39° Let us consider a simple example. One will find a substantial, historically interesting example in **Section 8**. Let a, b, and c be any numbers in $\mathcal{R}$. Let F be the function having domain $\mathcal{R}$, defined as follows:

$$F(x) := (x-a)^2 + (x-b)^2 + (x-c)^2$$

where x is any number in $\mathcal{R}$. Clearly:

$$F'(x) = 2(x-a) + 2(x-b) + 2(x-c) = 2(3x - (a+b+c))$$

where x is any number in $\mathcal{R}$. Obviously, $F'(x) < 0$ iff $3x < a+b+c$, $F'(x) = 0$ iff $3x = a+b+c$, and $0 < F'(x)$ iff $a+b+c < 3x$. We infer that the restriction of F to $(\leftarrow, (a+b+c)/3)$ is strictly decreasing and the restriction of F to $((a+b+c)/3, \rightarrow)$ is strictly increasing. We conclude that $\epsilon := (a+b+c)/3$ is a minimum number for F and there are no others. There is no maximum number for F. For the record:

$$F(\epsilon) = \frac{2}{3}(a^2 + b^2 + c^2 - ab - ac - bc)$$

The Inversion Theorem

40° Let X be any nontrivial open interval in $\mathcal{R}$ and let F be a function having domain X. Let F be differentiable on X. By the Intermediate Value Theorem, $Y := F(X)$ is an interval in $\mathcal{R}$. Let us assume that $F'(X) \subseteq \mathcal{R}^-$ or $F'(X) \subseteq \mathcal{R}^+$. In the first case, F is strictly decreasing and in the second case F is strictly increasing. In either case, we can introduce the function G having domain Y such that F and G are inverse to one another. Of course, G would be, correspondingly, strictly decreasing or strictly increasing. Obviously, Y is a nontrivial open interval in $\mathcal{R}$. Let us prove that:

(10) G is differentiable on Y, and, for any number y in Y:

$$G'(y) = \frac{1}{F'(x)}$$

where $x := G(y)$.

Let us focus upon the second of the foregoing cases, in which F is strictly increasing. One can derive the first case from the second case by replacing F by $-F$. Let b be any number in Y and let $a := G(b)$. Let us prove that G is continuous at b. To that end, let u be any number in $\mathcal{R}^+$. We must produce a number v in $\mathcal{R}^+$ such that $G(Y \cap N_v(b)) \subseteq N_u(a)$. Let p and q be the numbers in X such that:

$$X \cap N_u(a) = (p, q)$$

Let $r := F(p)$ and $s := F(q)$. Of course, $r < b < s$. Let v be a number in $\mathcal{R}^+$ such that:

$$N_v(b) \subseteq (r, s)$$

Clearly:

$$G(Y \cap N_v(b)) \subseteq N_u(a)$$

We conclude that G is continuous at b. In turn, let us prove that G is differentiable at b and that $G'(b) = 1/F'(a)$. To that end, let us introduce the continuous extension $\bar{F}_a$ of F_a at a. Let y be any number in Y_b and let $x := G(y)$. We have:

$$G_b(y) = \frac{1}{y-b}(G(y) - G(b)) = \frac{1}{F(x) - F(a)}(x - a) = \frac{1}{F_a(x)}$$

We infer that:

$$G_b = \frac{1}{F_a \circ G}$$

Clearly, G_b admits a continuous extension at b, because G is continuous at b. In fact:

$$\bar{G}_b = \frac{1}{\bar{F}_a \circ G}$$

We conclude that G is differentiable at b and:

$$G'(b) = \bar{G}_b(b) = \frac{1}{\bar{F}_a(G(b))} = \frac{1}{\bar{F}_a(a)} = \frac{1}{F'(a)}$$

The Range of F' is an Interval

41° Let X be any nontrivial interval in $\mathcal{R}$ and let F be a function having domain X. Let F be differentiable on X. Of course, the derivative F' of F need not itself be differentiable on X. In fact, F' need not even be continuous on X. Nevertheless, we can prove that:

(11) if F is differentiable on X then the range $Y' := F'(X)$ of F' is an interval in $\mathcal{R}$

To prove this remarkable property, let a and b be any numbers in X for which $a < b$. Under the stated assumption, $\bar{F}_a$ and $\bar{F}_b$ are continuous on X. Moreover, by definition:

$$\bar{F}_a(b) = F_a(b) = F_b(a) = \bar{F}_b(a)$$

Let τ be the common value of these numbers. Let $r := F'(a)$ and $s := F'(b)$. Let us assume that $r \neq s$. Let t be any number in $\mathcal{R}$ lying between r and s. We must prove that $t \in F'(X)$. To that end, let us simply note that one of the following three conditions must hold:

(∘) t lies between $r = \bar{F}_a(a)$ and $\tau = \bar{F}_a(b)$
(∘) $t = \tau$
(∘) t lies between $s = \bar{F}_b(b)$ and $\tau = \bar{F}_b(a)$

In the first case, the Intermediate Value Theorem implies that t lies in the interval $F_a((a,b))$. In the third case, the Intermediate Value Theorem implies that t lies in the interval $F_b((a,b))$. Hence, in all three cases, the Mean Value Theorem implies that:

$$t \in F'((a,b)) \subseteq F'(X)$$

♮

The Theorem of Taylor

42° Let X be any nontrivial interval in $\mathcal{R}$ and let F be a function having domain X. It may happen that F is differentiable on X. If so then we can form the (first) derivative F' of F. It may happen that F' is differentiable on X. If so then we can form the second derivative $F'' \equiv (F')'$ of F. It may happen that F'' is differentiable on X. If so then we can form the third derivative $F''' \equiv ((F')')'$ of F. Of course, we can continue these computations for as long as the relevant functions are differentiable on X:

$$F^{(0)} := F,\ \ F^{(1)} := F',\ \ F^{(2)} := F'',\ \ F^{(3)} := F''',\ \ \dots\ \ ,\ \ F^{(k)},\ \ \dots$$

For convenience, we interpret $F^{(0)}$ to be F itself. For any integer k in $\mathcal{Z}^+ \cup \{0\}$, we refer to $F^{(k)}$ as the *k-th derivative* of F. We write:

$$F \in \mathcal{D}^k$$

to express the condition that the k-th derivative of F exists. Of course, $F \in \mathcal{D}^{k+1}$ implies that $F \in \mathcal{D}^k$.

43° Very often, one denotes $F^{(k)}$ by:

$$\frac{d^k}{dx}F(x)$$

44° Let X be any nontrivial interval in $\mathcal{R}$ and let F be a function having domain X. Let k be any integer in $\mathcal{Z}^+ \cup \{0\}$. Let a be any number in X. Let us assume that $F \in \mathcal{D}^k$. Under this assumption, we define the k-th *Taylor Polynomial* for F at a as follows:

$$(T_a^k F)(x) := \sum_{j=0}^{k} \frac{1}{j!} F^{(j)}(a)(x-a)^j$$

where x is any number in $\mathcal{R}$. Let us display the cases in which $k = 0$, $k = 1$, $k = 2$, and $k = 3$:

$$\begin{aligned}
&(T_a^0 F)(x) = F(a)\\
&(T_a^1 F)(x) = F(a) + F'(a)(x-a)\\
&(T_a^2 F)(x) = F(a) + F'(a)(x-a) + \frac{1}{2}F''(a)(x-a)^2\\
&(T_a^3 F)(x) = F(a) + F'(a)(x-a) + \frac{1}{2}F''(a)(x-a)^2 + \frac{1}{6}F'''(a)(x-a)^3
\end{aligned}$$

These cases show the sense and the merit of the summation sign.

45° In turn, we define the k-th *Remainder Function* for F at a as follows:

$$(R_a^k F)(x) := F(x) - (T_a^k F)(x)$$

so that:

$$F(x) = (T_a^k F)(x) + (R_a^k F)(x)$$

where x is any number in X.

46° Let us prove that, for any integer k in $\mathcal{Z}^+ \cup \{0\}$ and:

(12) for any function F having domain X, if $F \in \mathcal{D}^{k+1}$ then, for any numbers a and x in X, if $x \neq a$ then there is a number c lying between a and x such that:

$$(R_a^k F)(x) := \frac{1}{(k+1)!} F^{(k+1)}(c)(x-a)^{k+1}$$

To that end, we apply the method of Mathematical Induction. For each integer k in $\mathcal{Z}^+ \cup \{0\}$, let T_k stand for property (12). Let us note that T_0 coincides with the Mean Value Theorem. We have already proved that it is true. Let k be any integer in $\mathcal{Z}^+ \cup \{0\}$. Let us assume that T_k is true. We must prove that T_{k+1} is true. Let F be any function having domain X such that $F \in \mathcal{D}^{k+2}$. Let a and x be any numbers in X for which $a \neq x$. We must produce a number c lying between a and x such that:

$$(R_a^{k+1} F)(x) = \frac{1}{(k+2)!} F^{(k+2)}(c)(x-a)^{k+2}$$

For that purpose, let us introduce the function G having domain X, defined as follows:

$$G(y) := (R_a^{k+1} F)(y)(x-a)^{k+2} - (y-a)^{k+2}(R_a^{k+1} F)(x)$$

where y is any number in X. Clearly, $G(x) = 0$ and $G(a) = 0$ as well, because $(R_a^{k+1} F)(a) = 0$. Moreover, G is differentiable on X and:

$$G'(y) = (R_a^{k+1} F)'(y)(x-a)^{k+2} - (k+2)(y-a)^{k+1}(R_a^{k+1} F)(x)$$

where y is any number in X. By the Mean Value Theorem, we can introduce a number d lying between a and x such that $G'(d) = 0$. That is:

$$(R_a^{k+1} F)'(d)(x-a)^{k+2} = (k+2)(d-a)^{k+1}(R_a^{k+1} F)(x)$$

However:

$$F = T_a^{k+1} F + R_a^{k+1} F, \quad F' = T_a^k F' + R_a^k F', \quad \text{and} \quad (T_a^{k+1} F)' = T_a^k F'$$

Hence:

$$(R_a^{k+1} F)' = R_a^k F'$$

We infer that:

$$(R_a^k F')(d)(x-a)^{k+2} = (k+2)(d-a)^{k+1}(R_a^{k+1} F)(x)$$

Of course, $F' \in \mathcal{D}^{k+1}$. By T_k, we can introduce a number c lying between a and d such that:

$$(R_a^k F')(d) = \frac{1}{(k+1)!}(F')^{(k+1)}(c)(d-a)^{k+1}$$

We conclude that:

$$(R_a^{k+1} F)(x) = \frac{1}{(k+2)!}F^{(k+2)}(c)(x-a)^{k+2}$$

♮

47° We refer to property (12) as the Theorem of Taylor. With reference to this theorem, we can take:

$$(T_a^k F)(x) = \sum_{j=0}^{k} \frac{1}{j!}F^{(j)}(a)(x-a)^j$$

to be an estimate of $F(x)$ and we can measure the error in the estimate by finding an upper bound for:

$$\big|(R_a^k F)(x)\big| = \frac{1}{(k+1)!}\big|F^{(k+1)}(c)\big|\big|x-a\big|^{k+1}$$

Of course, we would choose a to be a number in X at which we can readily compute the derivatives $F^{(j)}(a)$. However, in practice, it is not easy to bound the error. For examples, see articles 37° and 38° in **Section 7**.

5 Integration

1° We devote this section to a study of the condition of Integrability for bounded functions on closed finite intervals in $\mathcal{R}$. Subject to this condition, one can measure the *area* of the ordinate set of a function. We prove the computational properties of integrable functions and we isolate two conditions which are sufficient for integrability.

Integrable Functions

2° Let a and b be any numbers in $\mathcal{R}$ for which $a < b$. Let F be a function having domain $[a, b]$. We assume that F is bounded. Let:

$$\Omega^-(F) \quad \text{and} \quad \Omega^+(F)$$

be the negative and positive parts of the *ordinate set* $\Omega(F)$ for F. By definition, for any ordered pair (x, z) of numbers in $\mathcal{R}$:

$$(x, z) \in \Omega^-(F) \quad \textit{iff} \quad a \leq x \leq b \ \text{ and } \ F(x) \leq z \leq 0$$

and:

$$(x, z) \in \Omega^+(F) \quad \textit{iff} \quad a \leq x \leq b \ \text{ and } \ 0 \leq z \leq F(x)$$

while:

$$\Omega(F) = \Omega^-(F) \cup \Omega^+(F)$$

In due course, we will find that the condition of integrability for F coincides with the condition that $\Omega^-(F)$ and $\Omega^+(F)$ have well defined areas and we will find that the integral of F coincides with the "signed" area of $\Omega(F)$, where the area of $\Omega^-(F)$ counts as negative and the area of $\Omega^+(F)$ counts as positive.

3° Let us develop the condition of integrability for F. For that purpose, we require several technical definitions, by which we can reduce the measure of area for ordinate sets in general to the measure of area for rectangles in particular.

4° By a *partition* of $[a, b]$, we mean any finite subset P of $[a, b]$ such that $a \in P$ and $b \in P$. We can display P as an indexed chain of numbers in $[a, b]$ starting at a and ending at b:

$$a = p_0 < p_1 < p_2 < \cdots < p_{k-1} < p_k = b$$

We refer to the integers j $(0 \leq j \leq k)$ as *indices*.

5° Let P be any partition of $[a,b]$. For each index j $(1 \leq j \leq k)$, let:

$$\underline{n}_j(F,P)$$

be the largest number in $F([p_{j-1},p_j])_*$ and let:

$$\overline{n}_j(F,P)$$

be the smallest number in $F([p_{j-1},p_j])^*$. Without serious loss, one can regard $\underline{n}_j(F,P)$ as the smallest value of F and $\overline{n}_j(F,P)$ as the largest value of F on the subinterval $[p_{j-1},p_j]$ of $[a,b]$. In turn, let:

$$\underline{\Sigma}(F,P) = \sum_{j=1}^{k} \underline{n}_j(F,P)\,(p_j - p_{j-1})$$

and let:

$$\overline{\Sigma}(F,P) = \sum_{j=1}^{k} \overline{n}_j(F,P)\,(p_j - p_{j-1})$$

We refer to:

$$\underline{\Sigma}(F,P) \quad \text{and} \quad \overline{\Sigma}(F,P)$$

as the *lower estimate* and the *upper estimate*, relative to P, of the (signed) area of $\Omega(F)$. Obviously:

$$\underline{\Sigma}(F,P) \leq \overline{\Sigma}(F,P)$$

See Figure 7.

6° Let us prove that:

(1) for any partitions P and Q of $[a,b]$, $\underline{\Sigma}(F,P) \leq \overline{\Sigma}(F,Q)$

To that end, let r be any number in $[a,b]$ such that $r \notin P$. Let $R := P \cup \{r\}$ be the partition of $[a,b]$ produced by adding the number r to P. Isolating the relevant index j, we can display the partitions P and R as follows:

$$a = p_0 < p_1 < p_2 < \cdots p_{j-1} < \quad p_j < \cdots < p_{k-1} < p_k = b$$

$$a = p_0 < p_1 < p_2 < \cdots p_{j-1} < r < p_j < \cdots < p_{k-1} < p_k = b$$

Let $\underline{n}'_j(F,R)$ be the largest number in $F([p_{j-1},r])_*$ and let $\underline{n}''_j(F,R)$ be the largest number in $F([r,p_j])_*$. Clearly:

$$\underline{n}_j(F,P) \leq \underline{n}'_j(F,R) \quad \text{and} \quad \underline{n}_j(F,P) \leq \underline{n}''_j(F,R)$$

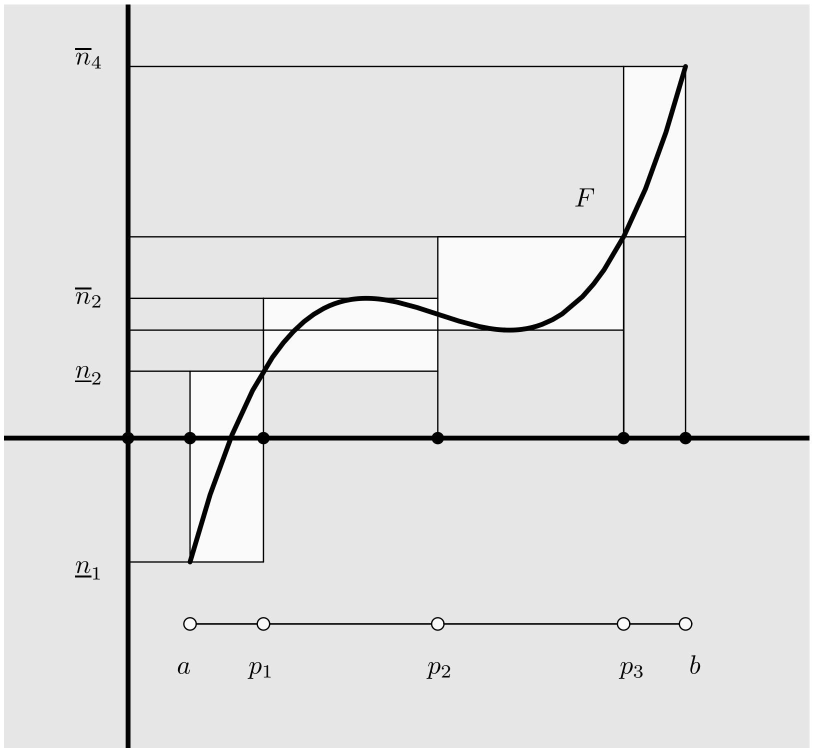


Figure 7; Lower and Upper Sums

Hence:

$$\underline{n}_j(F,P)(p_j - p_{j-1}) \leq \underline{n}'_j(F,R)(r - p_{j-1}) + \underline{n}''_j(F,R)(p_j - r)$$

We infer that:

$$\underline{\Sigma}(F,P) \leq \underline{\Sigma}(F,R)$$

By applying the foregoing observations "one new r at a time," we may infer that, for any partition R of $[\,a,b\,]$, if $P \subseteq R$ then $\underline{\Sigma}(F,P) \leq \underline{\Sigma}(F,R)$. In turn, let r be any number in $[\,a,b\,]$ such that $r \notin Q$. Let $R := Q \cup \{r\}$ be the partition of $[\,a,b\,]$ produced by adding the number r to Q. Isolating the relevant index j, we can display the partitions Q and R as follows:

$$a = q_0 < q_1 < q_2 < \cdots q_{j-1} < \quad q_j < \cdots < q_{\ell-1} < q_\ell = b$$

$$a = q_0 < q_1 < q_2 < \cdots q_{j-1} < r < q_j < \cdots < q_{\ell-1} < q_\ell = b$$

Let $\overline{n}'_j(F,R)$ be the smallest number in $F([\,p_{j-1}, r\,])^*$ and let $\overline{n}''_j(F,R)$ be the smallest number in $F([\,r, p_j\,])^*$. Clearly:

$$\overline{n}'_j(F,R) \leq \underline{n}_j(F,Q) \quad \text{and} \quad \overline{n}''_j(F,R) \leq \overline{n}_j(F,Q)$$

Hence:

$$\overline{n}'_j(F,R)(r - p_{j-1}) + \overline{n}''_j(F,R)(p_j - r) \leq \overline{n}_j(F,Q)(p_j - p_{j-1})$$

We infer that:

$$\overline{\Sigma}(F,R) \leq \overline{\Sigma}(F,Q)$$

By applying the foregoing observations "one new r at a time," we may infer that, for any partition R of $[\,a,b\,]$, if $Q \subseteq R$ then $\overline{\Sigma}(F,R) \leq \overline{\Sigma}(F,Q)$. Finally, let R be any partition of $[\,a,b\,]$ such that $P \subseteq R$ and $Q \subseteq R$. For example, let $R := P \cup Q$. We have:

$$\underline{\Sigma}(F,P) \leq \underline{\Sigma}(F,R) \leq \overline{\Sigma}(F,R) \leq \overline{\Sigma}(F,Q)$$

$\natural$

7° Now let $\underline{\boldsymbol{\Sigma}}(F)$ be the family of all numbers in $\mathcal{R}$ of the form:

$$\underline{\Sigma}(F,P)$$

where P runs through all partitions of $[\,a,b\,]$, and let $\overline{\boldsymbol{\Sigma}}(F)$ be the family of all numbers in $\mathcal{R}$ of the form:

$$\overline{\Sigma}(F,Q)$$

where Q runs through all partitions of $[\,a,b\,]$. By property (1):

$$\underline{\boldsymbol{\Sigma}}(F) \subseteq \overline{\boldsymbol{\Sigma}}(F)_* \quad \text{and} \quad \overline{\boldsymbol{\Sigma}}(F) \subseteq \underline{\boldsymbol{\Sigma}}(F)^*$$

Let:

$$\underline{\Sigma}(F)$$

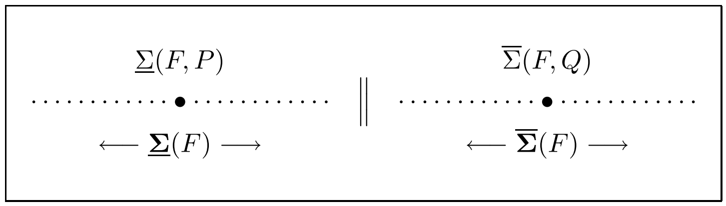
be the smallest number in $\underline{\boldsymbol{\Sigma}}(F)^*$ and let:

$$\overline{\Sigma}(F)$$

be the largest number in $\overline{\boldsymbol{\Sigma}}(F)_*$. One may regard $\underline{\Sigma}(F)$ as the "best" lower estimate and $\overline{\Sigma}(F)$ as the "best" upper estimate by means of rectangles of the area of $\Omega(F)$. Clearly:

$$\underline{\Sigma}(F) \leq \overline{\Sigma}(F)$$

Schematically:



8° Let a and b be any numbers in $\mathcal{R}$ for which $a < b$. Let F be a function having domain $[\,a,b\,]$. We assume that F is bounded. We say that F is *integrable* iff:

$$\underline{\Sigma}(F) = \overline{\Sigma}(F)$$

In such a case, we define the *integral* of F to be the common value of the foregoing numbers and we denote it by:

$$\int_a^b F$$

Hence:

$$\underline{\Sigma}(F) = \int_a^b F = \overline{\Sigma}(F)$$

9° Clearly, for any partition R of $[\,a,b\,]$:

$$\overline{\Sigma}(F) - \underline{\Sigma}(F) \leq \overline{\Sigma}(F,R) - \underline{\Sigma}(F,R)$$

We infer that F is integrable iff, for any number v in $\mathcal{R}^+$, there is some partition R of $[\,a,b\,]$ such that:

$$\overline{\Sigma}(F,R) - \underline{\Sigma}(F,R) < v$$

10° For illustration, let us consider the first among our standing examples. Let $a := 0$ and $b := 1$. Let F be the function having domain $[\,0,1\,]$, defined as follows:

$$F(x) := x^2$$

where x is any number in $[\,0,1\,]$. Let us prove that F is integrable and that:

$$\int_0^1 F = \frac{1}{3}$$

To that end, let k be any integer in $\mathcal{Z}^+$ and let:

$$p_j := \frac{j}{k}$$

where j is any integer for which $0 \le j \le k$. Let P be the partition of $[\,0,1\,]$ composed of the numbers just defined. For each index j $(1 \le j \le k)$:

$$\underline{n}_j(F,P) = \frac{1}{k^2}(j-1)^2 \quad \text{and} \quad \overline{n}_j(F,P) = \frac{1}{k^2}j^2$$

Hence:

$$\underline{\Sigma}(F,P) = \sum_{j=0}^{k-1} \frac{1}{k^2}j^2\frac{1}{k} = \frac{1}{k^3}\frac{1}{6}(k-1)k(2k-1) = \frac{1}{6}(1-\frac{1}{k})(2-\frac{1}{k})$$

$$\overline{\Sigma}(F,P) = \sum_{j=1}^{k} \frac{1}{k^2}j^2\frac{1}{k} = \frac{1}{k^3}\frac{1}{6}k(k+1)(2k+1) = \frac{1}{6}(1+\frac{1}{k})(2+\frac{1}{k})$$

See article 12° in **Section 1**. Obviously:

$$\underline{\Sigma}(F,P) < \frac{1}{3} < \overline{\Sigma}(F,P) = \underline{\Sigma}(F,P) + \frac{1}{k}$$

We conclude that $\underline{\Sigma}(F) = 1/3 = \overline{\Sigma}(F)$. Hence, F is integrable and:

$$\int_0^1 F = \frac{1}{3}$$

♮

11° Now let us consider a peculiar example. Let $a := 0$ and $b := 1$. Let F be the function having domain $[0,1]$, defined as follows:

$$F(x) := \begin{cases} 0 & \text{if } x \in \mathcal{Q} \\ 1 & \text{if } x \in \mathcal{I} := \mathcal{R}\backslash\mathcal{Q} \end{cases}$$

where x is any number in $[\,0,1\,]$. Let P be any partition of $[\,0,1\,]$. For each index j $(1 \le j \le k)$, we have:

$$\underline{n}_j(F,P) = 0 \qquad \text{and} \qquad \overline{n}_j(F,P) = 1$$

because both $\mathcal{Q}$ and $\mathcal{I}$ are dense in $\mathcal{R}$. Hence, $\underline{\Sigma}(F,P) = 0$ and $\overline{\Sigma}(F,P) = 1$. In turn, $\underline{\Sigma}(F) = 0 < 1 = \overline{\Sigma}(F)$. We conclude that F is not integrable. ♮

12° In the following articles 19° and 20°, we will acquire two simple conditions which faithfully imply integrability. In **Section 6**, we will acquire, by the Fundamental Theorem, an efficient technique for computing the values of the integrals.

Computational Properties

13° Let a and b be any numbers in $\mathcal{R}$ for which $a < b$. Let F, G, and H be functions having common domain $[a, b]$. Let c be a number in $\mathcal{R}$. We assume that F, G, and H are bounded. Let us prove that:

(2) if F and G are integrable then $F + G$ is integrable and:

$$\int_a^b (F+G) = \int_a^b F + \int_a^b G$$

To that end, let v be any number in $\mathcal{R}^+$. Since F is integrable, we can introduce a partition P' of $[a, b]$ such that:

$$\int_a^b F - (v/2) < \underline{\Sigma}(F, P') \leq \overline{\Sigma}(F, P') < \int_a^b F + (v/2)$$

Since G is integrable, we can introduce a partition P'' of $[a, b]$ such that:

$$\int_a^b G - (v/2) < \underline{\Sigma}(G, P'') \leq \overline{\Sigma}(G, P'') < \int_a^b G + (v/2)$$

Let $P := P' \cup P''$:

$$a = p_0 < p_1 < p_2 < \cdots < p_{k-1} < p_k = b$$

For each index j $(1 \leq j \leq k)$ and for any number x in $[p_{j-1}, p_j]$:

$$\underline{n}_j(F, P) \leq F(x) \leq \overline{n}_j(F, P)$$

$$\underline{n}_j(G, P) \leq G(x) \leq \overline{n}_j(G, P)$$

so:

$$\underline{n}_j(F, P) + \underline{n}_j(G, P) \leq F(x) + G(x) \leq \overline{n}_j(F, P) + \overline{n}_j(G, P)$$

Hence:

$$\underline{n}_j(F, P) + \underline{n}_j(G, P) \leq \underline{n}_j(F + G, P)$$

$$\overline{n}_j(F + G, P) \leq \overline{n}_j(F, P) + \overline{n}_j(G, P)$$

We infer that:

$$
\begin{aligned}
(\int_a^b F + \int_a^b G) - v &< \underline{\Sigma}(F, P') + \underline{\Sigma}(G, P'') \\
&\leq \underline{\Sigma}(F, P) + \underline{\Sigma}(G, P) \\
&\leq \underline{\Sigma}(F + G, P) \\
&\leq \overline{\Sigma}(F + G, P) \\
&\leq \overline{\Sigma}(F, P) + \overline{\Sigma}(G, P) \\
&\leq \underline{\Sigma}(F, P') + \underline{\Sigma}(G, P'') \\
&< (\int_a^b F + \int_a^b G) + v
\end{aligned}
$$

We conclude that $F + G$ is integrable and:

$$\int_a^b (F + G) = \int_a^b F + \int_a^b G$$

♮

14° In turn, let us prove that:

(3) if H is integrable then $c \cdot H$ is integrable and:

$$\int_a^b c \cdot H = c \cdot \int_a^b H$$

To that end, we will focus upon the cases in which $c = -1$ and $0 < c$. That will be sufficient, because the case in which $c = 0$ is obviously true and the case in which $c < 0$ can be reduced to the cases just described:

$$\int_a^b c \cdot H = \int_a^b (-|c|) \cdot H = (-|c|) \cdot \int_a^b H = c \cdot \int_a^b H$$

Let us assume that $c = -1$. Clearly, for any partition P of $[a, b]$ and for any index j $(1 \leq j \leq k)$:

$$\underline{n}_j(-H, P) = -\overline{n}_j(H, P) \quad \text{and} \quad \overline{n}_j(-H, P) = -\underline{n}_j(H, P)$$

Hence:

$$\underline{\Sigma}(-H, P) = -\overline{\Sigma}(H, P) \quad \text{and} \quad \overline{\Sigma}(-H, P) = -\underline{\Sigma}(H, P)$$

We infer that:

$$\underline{\Sigma}(-H) = -\overline{\Sigma}(H) = -\underline{\Sigma}(H) = \underline{\Sigma}(-H)$$

We conclude that $-H$ is integrable and:

$$\int_a^b (-H) = -\int_a^b H$$

Let us assume that $0 < c$. Clearly, for any partition P of $[\,a, b\,]$ and for any index j $(1 \leq j \leq k)$:

$$\underline{n}_j(c \cdot H, P) = c \cdot \underline{n}_j(H, P) \quad \text{and} \quad \overline{n}_j(c \cdot H, P) = c \cdot \overline{n}_j(H, P)$$

Hence:

$$\underline{\Sigma}(c \cdot H, P) = c \cdot \underline{\Sigma}(H, P) \quad \text{and} \quad \overline{\Sigma}(c \cdot H, P) = c \cdot \overline{\Sigma}(H, P)$$

We infer that:

$$\underline{\Sigma}(c \cdot H) = c \cdot \underline{\Sigma}(H) = c \cdot \overline{\Sigma}(H) = \underline{\Sigma}(c \cdot H)$$

We conclude that $c \cdot H$ is integrable and:

$$\int_a^b c \cdot H = c \cdot \int_a^b H$$

♮

The Order Property

15° Let a and b be any numbers in $\mathcal{R}$ for which $a < b$. Let F and G be functions having common domain $[\,a, b\,]$. We assume that F and G are bounded. Let us prove that:

(4) if F and G are integrable and if $F \leq G$ then:

$$\int_a^b F \leq \int_a^b G$$

To that end, let us introduce the functions $\hat{0}$ and $H := G - F$ having common domain $[\,a, b\,]$. Of course, the former is constantly 0 on $[\,a, b\,]$. By the preceding articles 13° and 14°, H is integrable. Clearly, $\hat{0} \leq H$. Hence, for any partition P of $[\,a, b\,]$:

$$0 \leq \underline{\Sigma}(H, P)$$

We conclude that:

$$0 \leq \int_a^b H = \int_a^b G - \int_a^b F$$

That is:

$$\int_a^b F \leq \int_a^b G$$

♮

16° Let a and b be numbers in X for which $a < b$. Let F be a function having domain $[a, b]$. We assume that F is bounded. Let us prove that:

(5) if F is integrable then $|F|$ is integrable and:

$$\Big| \int_a^b F \Big| \leq \int_a^b |F|$$

To that end, let P be any partition of $[a, b]$. We find that:

$$\begin{aligned} \overline{\Sigma}(|F|, P) - \underline{\Sigma}(|F|, P) &= \sum_{j=1}^{k} (\overline{n}_j(|F|, P) - \underline{n}_j(|F|, P))(p_j - p_{j-1}) \\ (!) \qquad &\leq \sum_{j=1}^{k} (\overline{n}_j(F, P) - \underline{n}_j(F, P))(p_j - p_{j-1}) \\ &= \overline{\Sigma}(F, P) - \underline{\Sigma}(F, P) \end{aligned}$$

See the following article. By article 9°, we infer that $|F|$ is integrable. Since:

$$-|F| \leq F \leq |F|$$

we infer that:

$$-\int_a^b |F| \leq \int_a^b F \leq \int_a^b |F|$$

We conclude that:

$$\Big| \int_a^b F \Big| \leq \int_a^b |F|$$

17° Let us defend the inequality:

$$\overline{n}_j(|F|, P) - \underline{n}_j(|F|, P) \leq \overline{n}_j(F, P) - \underline{n}_j(F, P)$$

which figured as the critical move in the foregoing argument. For that purpose, we note that, for any numbers x and y in $[p_{j-1}, p_j]$:

$$\underline{n}_j(F,P) \leq F(x) \leq \overline{n}_j(F,P) \quad \text{and} \quad \underline{n}_j(F,P) \leq F(y) \leq \overline{n}_j(F,P)$$

Hence:

$$-(\overline{n}_j(F,P) - \underline{n}_j(F,P)) \leq F(x) - F(y) \leq (\overline{n}_j(F,P) - \underline{n}_j(F,P))$$

and so:

$$|F(x) - F(y)| \leq \overline{n}_j(F,P) - \underline{n}_j(F,P)$$

Let z be any number in $\mathcal{R}$ such that:

$$z < \overline{n}_j(|F|,P) - \underline{n}_j(|F|,P)$$

Clearly, we can introduce a number x in $[p_{j-1}, p_j]$ such that:

$$z + \underline{n}_j(|F|,P) < |F|(x) < \overline{n}_j(|F|,P)$$

In turn, we can introduce a number y in $[p_{j-1}, p_j]$ such that:

$$\underline{n}_j(|F|,P) < |F|(y) < |F|(x) - z < \overline{n}_j(|F|,P) - z$$

Hence:

$$z < |F(x)| - |F(y)| \leq |F(x) - F(y)| \leq \overline{n}_j(F,P) - \underline{n}_j(F,P)$$

We conclude that:

$$\overline{n}_j(|F|,P) - \underline{n}_j(|F|,P) \leq \overline{n}_j(F,P) - \underline{n}_j(F,P)$$

♮

Notation

18° On occasion, we denote the integral of F not by $\int_a^b F$ but by:

$$\int_a^b F(x)dx$$

Of course, we might choose some other symbol, such as y, in place of x. By this notation, we gain a certain flexibility. We can compute integrals without engaging in the sometimes circuitous process of "naming" the function. For instance, the following expression is now clear and meaningful:

$$\int_0^1 x^2 dx = \frac{1}{3}$$

Monotonic Functions are Integrable

19° Let a and b be any numbers in $\mathcal{R}$ for which $a < b$. Let F be a function having domain $[a, b]$. Let us prove that:

(6) if F is monotone then F is integrable

By definition, F is decreasing or increasing. Obviously, in either case, it is bounded. Moreover, $-F$ is decreasing iff F is increasing. Hence, without loss, we can focus upon the second case. Let v be any number in $\mathcal{R}^+$. Let k be any integer in $\mathcal{Z}^+$ such that $(F(b) - F(a))(b - a)/k < v$ and let:

$$p_j := a + \frac{j}{k}(b - a)$$

where j is any integer for which $0 \leq j \leq k$. Let P be the partition of $[a, b]$ composed of the numbers just defined. For each index j ($1 \leq j \leq k$):

$$\underline{n}_j(F, P) = F(p_{j-1}) \quad \text{and} \quad \overline{n}_j(F, P) = F(p_j)$$

Hence:

$$\overline{\Sigma}(F, P) - \underline{\Sigma}(F, P) = F(b)\frac{1}{k}(b - a) - F(a)\frac{1}{k}(b - a) < v$$

and:

$$\underline{\Sigma}(F, P) \leq \underline{\Sigma}(F) \leq \overline{\Sigma}(F) \leq \overline{\Sigma}(F, P) < \underline{\Sigma}(F, P) + v$$

We conclude that:

$$\underline{\Sigma}(F) = \overline{\Sigma}(F)$$

hence that F is integrable. $\natural$

Continuous Functions are Integrable

20° Let a and b be any numbers in $\mathcal{R}$ for which $a < b$. Let F be a function having domain $[a, b]$. Let us prove that:

(7) if F is continuous on $[a, b]$ then F integrable

By the Extreme Value Theorem, F is bounded. Let v be any number in $\mathcal{R}^+$. By the Uniform Continuity Theorem, we can introduce a number u in $\mathcal{R}^+$ such that, for any numbers x and y in $[a, b]$, if $|x - y| < u$ then

$|F(x)-F(y)| < v/(b-a)$. Let k be any integer in $\mathcal{Z}^+$ such that $(b-a)/k < u$ and let:

$$p_j := a + \frac{j}{k}(b-a)$$

where j is any integer for which $0 \leq j \leq k$. Let P be the partition of $[\,a, b\,]$ composed of the numbers just defined. For each index j $(1 \leq j \leq k)$ and for any numbers x and y in $[p_{j-1}, p_j\,]$:

$$\overline{n}_j(F, P) - \underline{n}_j(F, P) < v/(b-a)$$

because $|x-y| < u$ and $|F(x) - F(y)| < v/(b-a)$. Hence:

$$\overline{\Sigma}(F, P) - \underline{\Sigma}(F, P) = \sum_{j=1}^{k} (\overline{n}_j(F, P) - \underline{n}_j(F, P))\frac{1}{k}(b-a) < v$$

and:

$$\underline{\Sigma}(F, P) \leq \underline{\Sigma}(F) \leq \overline{\Sigma}(F) \leq \overline{\Sigma}(F, P) < \underline{\Sigma}(F, P) + v$$

We conclude that:

$$\underline{\Sigma}(F) = \overline{\Sigma}(F)$$

hence that F is integrable. ♮

The Property of Additivity

21° Let a, b, and c be any numbers in $\mathcal{R}$ for which $a < b < c$. Let F be a function having domain $[\,a, c\,]$ and let G_1 and G_2 be the restrictions of F to $[\,a, b\,]$ and $[\,b, c\,]$. Let us prove that:

(8) F is integrable iff G_1 and G_2 are integrable and:

$$\int_a^c F = \int_a^b G_1 + \int_b^c G_2$$

Let us assume that F is integrable. Let R be any partition of $[\,a, c\,]$. Let $R^* = R \cup \{b\}$, $R_1 := R^* \cap [\,a, b\,]$, and $R_2 := R^* \cap [\,b, c\,]$. Obviously:

$$\underline{\Sigma}(G_1, R_1) - \overline{\Sigma}(G_1, R_1) \leq \underline{\Sigma}(F, R^*) - \overline{\Sigma}(F, R^*) \leq \underline{\Sigma}(F, R) - \overline{\Sigma}(F, R)$$

and:

$$\underline{\Sigma}(G_2, R_2) - \overline{\Sigma}(G_2, R_2) \leq \underline{\Sigma}(F, R^*) - \overline{\Sigma}(F, R^*) \leq \underline{\Sigma}(F, R) - \overline{\Sigma}(F, R)$$

We infer that G_1 and G_2 are integrable. Now let us assume that G_1 and G_2 are integrable. Let v be any number in $\mathcal{R}^+$. Since G_1 is integrable, we can introduce a partition R_1 of $[a, b]$ such that:

$$\int_a^b G_1 - (v/2) < \underline{\Sigma}(G_1, R_1) \leq \overline{\Sigma}(G_1, R_1) < \int_a^b G_1 + (v/2)$$

Since G_2 is integrable, we can introduce a partition R_2 of $[b, c]$ such that:

$$\int_b^c G_2 - (v/2) < \underline{\Sigma}(G_2, R_2) \leq \overline{\Sigma}(G_2, R_2) < \int_b^c G_2 + (v/2)$$

Let R be the partition of $[a, c]$ formed by listing first the terms of R_1, then the terms of R_2: $R := R_1 \cup R_2$. Clearly:

$$\begin{aligned}
(\int_a^b G_1 + \int_b^c G_2) - v &< \underline{\Sigma}(G_1, R_1) + \underline{\Sigma}(G_2, R_2) \\
&= \underline{\Sigma}(F, R) \\
&\leq \overline{\Sigma}(F, R) \\
&= \overline{\Sigma}(G_1, R_1) + \overline{\Sigma}(G_2, R_2) \\
&< (\int_a^b G_1 + \int_a^c G_2) + v
\end{aligned}$$

We conclude that F is integrable and:

$$\int_a^c F = \int_a^b G_1 + \int_b^c G_2$$

♮

The Oriented Integral

22° Let X be any nontrivial interval in $\mathcal{R}$ and let F be a function having domain X. Let us assume that, for any numbers a and b in X, if $a < b$ then the restriction of F to $[a, b]$ is bounded and integrable. In turn, let a and b be any numbers in X, unconstrained. Of course, we may have $a < b$, $a = b$, or $b < a$. In the first case, the following expresson is clear and meaningful:

$$\int_a^b F$$

Let us extend the sense of the expression to cover all three cases, as follows:

$$\int_a^b F := \begin{cases} \displaystyle\int_a^b F & \text{if } a < b \\ 0 & \text{if } a = b \\ \displaystyle -\int_b^a F & \text{if } b < a \end{cases}$$

Let us prove that:

(9) for any numbers a, b, and c in X:

$$\int_a^c F = \int_a^b F + \int_b^c F$$

By simple inspection, we find that if any two of the three numbers a, b, and c are equal then the displayed relation is true. Hence, we need only consider the six cases in which the numbers a, b, and c are mutually distinct. We do so by the following unconventional method. We select one of the six cases at random, prove the displayed relation for that case, then declare the remaining five cases to be true because, under random selection, any one of them *could* have been selected. Now let us select, at random, the case: $c < a < b$. By property (8), we have:

$$\int_c^b F = \int_c^a F + \int_a^b F$$

Hence:

$$\int_a^c F = -\int_c^a F = \int_a^b F - \int_c^b F = \int_a^b F + \int_b^c F$$

We rest our cases. They who are not convinced by this method should check the other cases, one by one. ☺

6 THE FUNDAMENTAL THEOREM

1° In this section, we develop the Fundamental Theorem of our subject, which relates the actions of differentiation and integration. To that end, we introduce the concepts of Antiderivative and Indefinite Integral. We also develop two important methods for the calculation of integrals: Integration by Parts and Integration by Substitution.

Antiderivatives

2° Let X be any nontrivial interval in $\mathcal{R}$ and let F be a function having domain X. We inquire whether or not there exists a function G having domain X such that G is differentiable on X and $G' = F$. That is:

$$G'(x) = F(x)$$

where x is any number in X. We refer to such a function G (if it exists) as an *antiderivative* of F.

Essential Uniqueness

3° If such a function G does exist then it is essentially unique. Thus, for any functions G_1 and G_2 having common domain X, if G_1 and G_2 are antiderivatives of F then $(G_1 - G_2)' = F - F = 0$, hence $G_1 - G_2$ is constant. That is, there is a number z in $\mathcal{R}$ such that:

$$G_1(x) = G_2(x) + z$$

where x is any number in X.

4° However, such a function G may not exist. For instance, that would be so if the range of F is not an interval. See article 41° in **Section 4**.

Existence: Indefinite Integrals

5° Now let us prove that:

(1) if F is continuous then, indeed, there exists a function G such that G is an antiderivative of F

To that end, let a be any number in X. Let G_a be the function having domain X, defined as follows:

$$G_a(x) := \int_a^x F$$

where x is any number in X. We refer to G_a as the *indefinite integral* of F at a. Let us prove that G_a is an antiderivative of F. Thus, let b be any number in X. We must prove that G_a is differentiable at b and that $G'_a(b) = F(b)$. For convenience, let $\hat{F(b)}$ denote the constant function having domain X and constant value $F(b)$. Let v be any number in $\mathcal{R}^+$. Since F is continuous at b, we can introduce a number u in $\mathcal{R}^+$ such that, for any number y in X, if $|y-b| < u$ then $|F(y) - F(b)| < v$. Hence, for any number y in X, if $0 < |y-b| < u$ then:

$$\begin{aligned}
\Big|\frac{1}{y-b}(G_a(y) - G_a(b)) \ - \ F(b)\Big| &= \Big|\frac{1}{y-b}(\int_a^y F - \int_a^b F) \ - \ F(b)\Big| \\
&= \Big|\frac{1}{y-b}\int_b^y F \ - \ \frac{1}{y-b}\int_b^y \hat{F(b)}\Big| \\
&= \Big|\frac{1}{y-b}\int_b^y (F - \hat{F(b)})\Big| \\
&\leq \frac{1}{|y-b|}\Big|\int_b^y |F - \hat{F(b)}|\Big| \\
&< \frac{1}{|y-b|}\cdot v\cdot |y-b| \\
&= v
\end{aligned}$$

We conclude that $G'_a(b) = F(b)$. ♮

Computation

6° Let X be any nontrivial interval in $\mathcal{R}$ and let F be a function having domain X. Let F be continuous. Let G be an antiderivative of F. Let us prove that:

(2) for any numbers a and b in X:

$$\int_a^b F = G(b) - G(a)$$

Since $G - G_a$ is constant, we can introduce a number z in $\mathcal{R}$ such that:

$$G(x) = G_a(x) + z$$

where x is any number in X. Hence:

$$G(b) - G(a) = (G_a(b) + z) - (G_a(a) + z) = G_a(b) = \int_a^b F$$

♮

7° In summary, we may say that Indefinite Integrals yield Antiderivatives and that Antiderivatives yield Definite Integrals. These assertions, explained by relations (1) and (2), comprise the Fundamental Theorem.

8° By the notation introduced in article 28° of **Section 4**, we can express relation (1) as follows:

$$\frac{d}{dx}\int_a^x F = F(x)$$

In turn, by the following convenient notation:

$$G(x)\Big|_a^b = G(b) - G(a)$$

and by the notation introduced in article 19° of **Section 5**, we can express relation (2) as follows:

$$\int_a^b F(x)dx = G(x)\Big|_a^b$$

For illustration, we offer the following expressions:

$$\frac{d}{dx}\int_a^x y^2 dy = x^2$$

and:

$$-\int_a^b \frac{2}{(1+x)^2}dx = \frac{1-x}{1+x}\Big|_a^b$$

See article 29° in **Section 4**.

Integration by Parts

9° Let X be any nontrivial interval in $\mathcal{R}$ and let F and G be functions having common domain X. Let F and G be differentiable. Let a and b be any numbers in X. By the Product Rule, $F \cdot G$ is an antiderivative for $F \cdot G' + F' \cdot G$. By the Fundamental Theorem:

$$(\circ) \qquad \int_a^b F \cdot G' \;+\; \int_a^b F' \cdot G \;=\; F(b)G(b) - F(a)G(a)$$

Just as well:

$$\int_a^b F(x)G'(x)dx \;+\; \int_a^b F'(x)G(x)dx \;=\; F(b)G(b) - F(a)G(a)$$

10° Given F and G, we can compute the integral:

$$\int_a^b F' \cdot G$$

by computing the integral:

$$\int_a^b F \cdot G'$$

and by applying the foregoing relation. By this procedure, we may gain an advantage, because the latter integral may be much easier to compute than the former. We refer to the procedure as Integration by Parts.

11° Let us work out an example of Integration by Parts. To raise interest, we borrow the logarithm function L from the following section. We have:

$$\int_1^2 x\frac{d}{dx}log(x)dx + \int_1^2 \frac{d}{dx}(\frac{1}{2}x^2)log(x)dx = \int_1^2 \hat{1}\,dx + \int_1^2 x\,log(x)dx$$

Hence:

$$\int_1^2 x\,log(x)dx = -\int_1^2 \hat{1}\,dx \ + \ \frac{1}{2}x^2 log(x)\Big|_1^2 = 2\,log(2) - 1$$

Integration by Substitution

12° Let X and Y be any nontrivial intervals in $\mathcal{R}$, let F be a function having domain X, and let G be a function having domain Y. Let us assume that $F(X) \subseteq Y$. Let F be differentiable, let F' be continuous, and let G be continuous. Let a and b be any numbers in X and let:

$$c := F(a) \ \text{ and } \ d := F(b)$$

Let H be an antiderivative for G. By the Composition Rule, $H \circ F$ is an antiderivative for $(G \circ F) \cdot F'$. Of course:

$$(H \circ F)(a) = H(c) \ \text{ and } \ (H \circ F)(b) = H(d)$$

By the Fundamental Theorem:

$$(\circ) \qquad \int_a^b (G \circ F) \cdot F' \ = \ \int_c^d G$$

Just as well:

$$\int_a^b G(F(x))F'(x)dx \ = \ \int_c^d G(y)dy$$

13° Given G, we can try to design F so that the integral:

$$\int_a^b (G \circ F) \cdot F'$$

is much easier to compute than the integral:

$$\int_c^d G$$

If successful, we gain an advantage, though the design of F may be troublesome. We refer to this procedure as Integration by Substitution.

14° Let us work out an example of Integration by Substitution. To raise interest, we borrow the exponential function E and the power function $P_{1/2}$ from the following section. By the substitution:

$$y := log(x^2 - 1), \qquad 1 < x$$

we have:

$$\frac{dy}{dx} = \frac{2x}{x^2 - 1}$$

and

$$\frac{exp(2y)}{\sqrt{exp(y) + 1}} = \frac{(x^2 - 1)^2}{x}$$

Hence:

$$\begin{aligned}
\int_{log(3)}^{log(8)} \frac{exp(2y)}{\sqrt{exp(y) + 1}} dy &= \int_2^3 \frac{(x^2 - 1)^2}{x} \frac{2x}{x^2 - 1} dx \\
&= \int_2^3 2(x^2 - 1) dx \\
&= 2(\frac{1}{3}x^3 - x)\Big|_2^3 \\
&= \frac{32}{3}
\end{aligned}$$

In retrospect, we see that the good effect of the substitution is to turn the problematic expression $\sqrt{exp(y) + 1}$ into x.

7 Classical Functions

1° Now we turn to the definition and analysis of the *classical functions*, which figure most frequently in the applications of differential and integral calculus.

The Logarithm Function

2° Let us prove that there is exactly one function L having domain $\mathcal{R}^+$ and meeting the following conditions:

(1) $L(1) = 0$
(2) L is differentiable on $\mathcal{R}^+$ and, for each number x in $\mathcal{R}^+$, $L'(x) = 1/x$

One refers to L as the *logarithm* function. One usually writes $log(x)$ instead of $L(x)$.

3° To prove that L exists, we argue as follows. For each x in $\mathcal{R}^+$, let:

$$L(x) = \int_1^x \frac{1}{z} dz$$

Obviously, $L(1) = 0$. By the Fundamental Theorem, L is differentiable on $\mathcal{R}^+$ and, for each x in $\mathcal{R}^+$, $L'(x) = 1/x$.

4° To prove that L is unique, we argue as follows. Let L_1 and L_2 be functions having common domain $\mathcal{R}^+$ and meeting the conditions (1) and (2). Obviously, $L_1' = L_2'$. Hence, $L_1 - L_2$ is constant. Moreover, $L_1(1) - L_2(1) = 0$. Hence, $L_1 - L_2$ is constantly 0. Therefore, $L_1 = L_2$. ♮

5° Let us prove that, for any numbers x, y, and z in $\mathcal{R}^+$:

$$(\circ) \qquad L(xy) = L(x) + L(y) \quad \text{and} \quad L(1/z) = -L(z)$$

To that end, let y be any number in $\mathcal{R}^+$. Let $\bar{L}$ be the function having domain $\mathcal{R}^+$, defined as follows:

$$\bar{L}(x) := L(xy) - L(y)$$

where x is any number in $\mathcal{R}^+$. Clearly, for any number x in $\mathcal{R}^+$:

$$\bar{L}'(x) = (1/xy) \cdot y = 1/x$$

Moreover, $\bar{L}(1) = 0$. Hence, $\bar{L} = L$. We conclude that, for any numbers x and y in $\mathcal{R}^+$, $L(xy) = L(x) + L(y)$. In turn, let z be any number in $\mathcal{R}^+$. We find that:

$$0 = L(1) = L(z \cdot (1/z)) = L(z) + L(1/z)$$

We conclude that, for any number z in $\mathcal{R}^+$, $L(1/z) = -L(z)$. ♮

6° Let us note that the values of L' are positive numbers. Hence, L is strictly increasing. By the Intermediate Value Theorem, $L(\mathcal{R}^+)$ is an interval. Let us prove that, in fact, $L(\mathcal{R}^+) = \mathcal{R}$. Obviously, $0 = L(1) < L(2)$. By article 5° and by Mathematical Induction, one can easily prove that, for any integer j in $\mathcal{Z}^+$, $L(2^j) = jL(2)$. By the Principle of Archimedes, we infer that $L(\mathcal{R}^+)^* = \emptyset$. By article 5°, $-L(\mathcal{R}^+) = L(\mathcal{R}^+)$. We infer that $L(\mathcal{R}^+)_* = \emptyset$, as well. We conclude that $L(\mathcal{R}^+) = \mathcal{R}$. ♮

7° In particular, we can introduce the number e in $\mathcal{R}^+$ such that:

$$(\circ) \qquad L(e) = \int_1^e \frac{1}{z} dz = 1$$

The Exponential Function

8° Let us prove that there is exactly one function E having domain $\mathcal{R}$ and meeting the following conditions:

(3) $E(0) = 1$
(4) E is differentiable on $\mathcal{R}$ and, for each number y in $\mathcal{R}$, $E'(y) = E(y)$

One refers to E as the *exponential* function. One usually writes $exp(y)$ instead of $E(y)$.

9° To prove that E exists, we argue as follows. By article 6°, we can introduce the function E having domain $\mathcal{R}$ such that L and E are inverse to one another. Of course, $E(\mathcal{R}) = \mathcal{R}^+$. Obviously, $E(0) = 1$ because $L(1) = 0$. By the Inversion Theorem, E is differentiable on $\mathcal{R}$ and, for each number y in $\mathcal{R}$:

$$E'(y) = 1/L'(x) = 1/(1/x) = x$$

where $x := E(y)$. See article 40° in **Section 4**.

10° To prove that E is unique, we argue as follows. Let E_1 and E_2 be functions having common domain $\mathcal{R}$ and meeting the conditions (3) and (4). We may assume that the values of E_2 are positive, since that is true for the particular case of E itself. Obviously:

$$(E_1/E_2)' = (E_2 \cdot E_1' - E_1 \cdot E_2')/E_2^2 = (E_2 \cdot E_1 - E_1 \cdot E_2)/E_1^2 = 0$$

Hence, E_1/E_2 is constant. Moreover, $E_1(0)/E_2(0) = 1$. Hence, E_1/E_2 is constantly 1. Therefore, $E_1 = E_2$. ♮

11° Clearly, for any numbers x, y, and z in $\mathcal{R}$:

(◦) $$E(x+y) = E(x) \cdot E(y) \quad \text{and} \quad E(-z) = 1/E(z)$$

because:

$$L(E(x+y)) = x + y = L(E(x)) + L(E(y)) = L(E(x) \cdot E(y))$$

and:

$$L(E(-z)) = -z = -L(E(z)) = L(1/E(z))$$

12° Obviously:

(◦) $$E(1) = e$$

13° The following graphs display the relation between L and E.

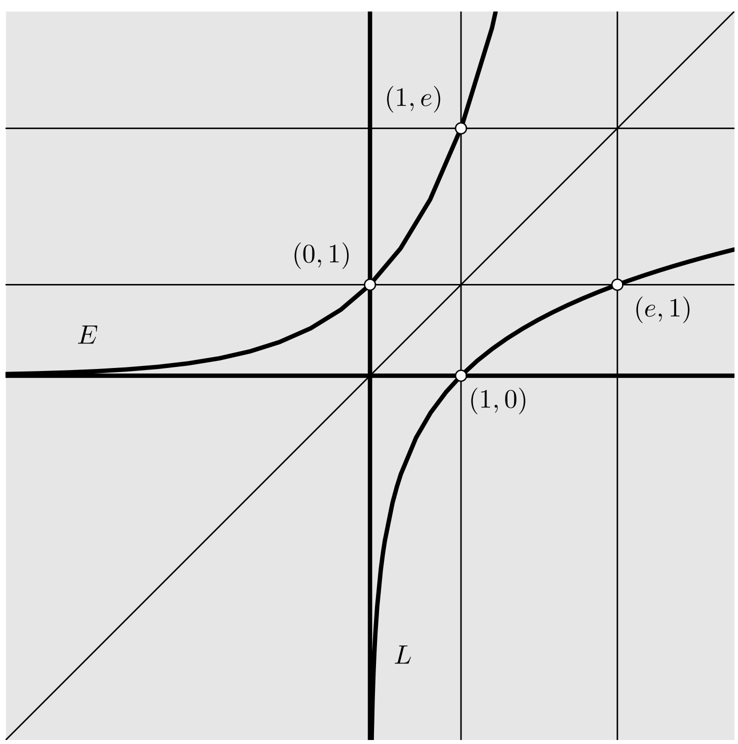


Figure 8: Logarithm and Exponential Functions

The Power Functions

14° Let a be any number in $\mathcal{R}$. By the *power function* with *exponent* a, we mean the function P_a having domain $\mathcal{R}^+$, defined in terms of L and E as follows:

$$(\circ) \qquad P_a(x) := E(aL(x))$$

where x is any number in $\mathcal{R}^+$. One usually writes x^a instead of $P_a(x)$, so that:

$$x^a = exp(a\,log(x))$$

In particular:

$$\begin{aligned} P_{-1}(x) &= E(-L(x)) = E(L(1/x)) = 1/x \\ P_0(x) &= E(0 \cdot L(x)) = E(0) = 1 \\ P_1(x) &= E(L(x)) = x \\ P_2(x) &= E(2 \cdot L(x)) = E(L(x) + L(x)) = E(L(x)) \cdot E(L(x)) = x \cdot x \end{aligned}$$

and so forth. Moreover:

$$\begin{aligned} P_{1/2}(x) \cdot P_{1/2}(x) &= E((1/2)L(x)) \cdot E((1/2)L(x)) \\ &= E((1/2)L(x) + (1/2)L(x)) \\ &= E(L(x)) \\ &= x \end{aligned}$$

so that, in familiar notation:

$$(\circ) \qquad P_{1/2}(x) = \sqrt{x}$$

15° Let us prove that, for any numbers a and b in $\mathcal{R}$:

$$(\circ) \qquad P_a \cdot P_b = P_{a+b} \quad \text{and} \quad P_a \circ P_b = P_{ab}$$

Thus, for any number x in $\mathcal{R}^+$:

$$\begin{aligned} (P_a \cdot P_b)(x) &= P_a(x) \cdot P_b(x) \\ &= E(aL(x)) \cdot E(bL(x)) \\ &= E(aL(x) + bL(x)) \\ &= E((a+b)L(x)) \\ &= P_{a+b}(x) \end{aligned}$$

and:

$$\begin{aligned}(P_a \circ P_b)(x) &= P_a(P_b(x)) \\ &= E(aL(P_b(x))) \\ &= E(aL(E(bL(x)))) \\ &= E(abL(x)) \\ &= P_{ab}(x)\end{aligned}$$

16° Let us prove that, for any number a in $\mathcal{R}$:

$$(\circ) \qquad P'_a = a \cdot P_{a-1}$$

Thus, for any number x in $\mathcal{R}^+$:

$$\begin{aligned}P'_a(x) &= E'(aL(x)) \cdot a \cdot L'(x) \\ &= E(aL(x)) \cdot a \cdot (1/x) \\ &= a \cdot P_a(x) \cdot P_{-1}(x) \\ &= a \cdot P_{a-1}(x) \\ &= (a \cdot P_{a-1})(x)\end{aligned}$$

17° One can rewrite the foregoing relations as follows:

$$x^{a+b} = x^a \cdot x^b, \quad x^{ab} = (x^b)^a, \quad \frac{d}{dx}x^a = a \cdot x^{a-1}$$

18° The graphs in Figure 9 show the pattern of the power functions:

$$P_a \qquad (a \in \mathcal{R})$$

The Trigonometric Functions

19° We contend that there is exactly one pair of functions C and S having common domain $\mathcal{R}$ and meeting the following conditions:

(5) $C(0) = 1$ and $S(0) = 0$

(6) C and S are differentiable on $\mathcal{R}$ and, for each number z in $\mathcal{R}$, $C'(z) = -S(z)$ and $S'(z) = C(z)$

One refers to C and S as the *trigonometric* functions, in particular, to C as the *cosine* function and to S as the *sine* function. One usually writes $cos(z)$ instead of $C(z)$ and $sin(z)$ instead of $S(z)$.

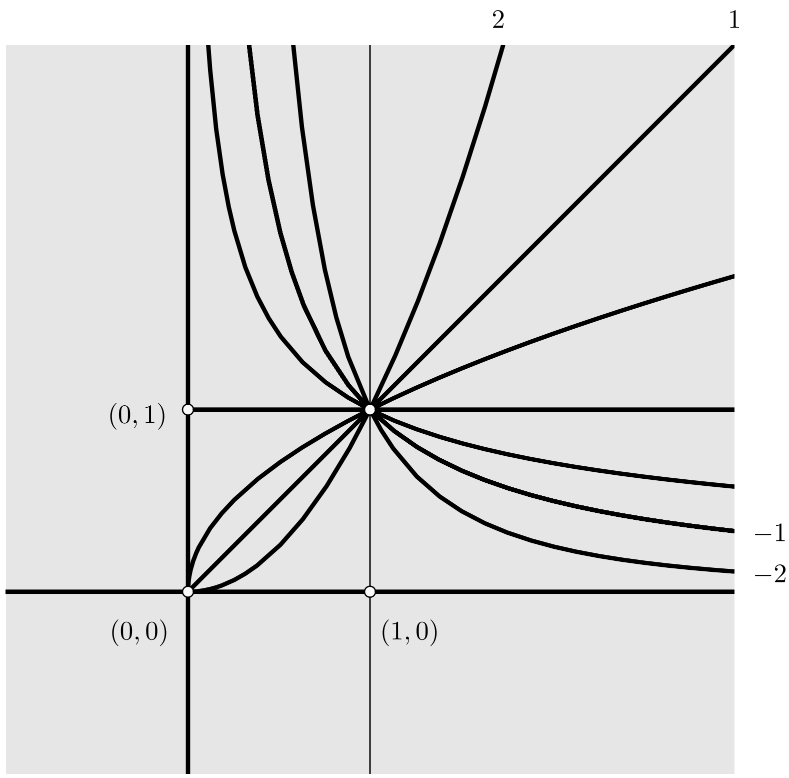


Figure 9: Power Functions

20° To prove that C and S exist, we require the theory of Power Series. We defer that story to a later day. For now, we simply assume that C and S exist. However, for suggestions of arguments, one may look ahead to article 29°.

21° To prove that C and S are unique, we argue as follows. Let C_1 and S_1 and C_2 and S_2 be pairs of functions having common domain $\mathcal{R}$ and meeting the conditions (5) and (6). Let F be the function having domain $\mathcal{R}$, defined as follows:

$$F = (C_1 - C_2)^2 + (S_1 - S_2)^2$$

Clearly:

$$F' = 2(C_1 - C_2)(S_2 - S_1) + 2(S_1 - S_2)(C_1 - C_2) = 0$$

Moreover, $F(0) = (1-1)^2 + (0-0)^2 = 0$. Hence, F is constantly 0. Therefore, $C_1 = C_2$ and $S_1 = S_2$. ♮

22° Let us prove that, for any number z in $\mathcal{R}$:

$$C(z)^2 + S(z)^2 = 1 \tag{$\circ$}$$

To that end, let F be the function having domain $\mathcal{R}$, defined as follows:

$$F = C^2 + S^2$$

Clearly:

$$F' = 2CC' + 2SS' = -2CS + 2CS = 0$$

Moreover, $F(0) = 1^2 + 0^2 = 1$. Hence, F is constantly 1. ♮

23° Let us prove that, for any number z in $\mathcal{R}$:

$$C(-z) = C(z) \quad \text{and} \quad S(-z) = -S(z) \tag{$\circ$}$$

To that end, let $\bar{C}$ and $\bar{S}$ be the functions having common domain $\mathcal{R}$, defined as follows:

$$\begin{aligned} \bar{C}(z) &:= C(-z) \\ \bar{S}(z) &:= -S(-z) \end{aligned}$$

where z is any number in $\mathcal{R}$. Clearly, $\bar{C}' = -\bar{S}$ and $\bar{S}' = \bar{C}$. Moreover, $\bar{C}(0) = 1$ and $\bar{S}(0) = 0$. Hence, $\bar{C} = C$ and $\bar{S} = S$. ♮

24° Let us prove that, for any numbers x and y in $\mathcal{R}$:

$$C(x + y) = C(x)C(y) - S(x)S(y) \tag{$\circ$}$$

and:

$$S(x + y) = S(x)C(y) + C(x)S(y) \tag{$\circ$}$$

To that end, let y be any number in $\mathcal{R}$ and let $\bar{C}$ and $\bar{S}$ be functions having common domain $\mathcal{R}$, defined as follows:

$$\begin{aligned} \bar{C}(x) &:= C(y)C(x + y) + S(y)S(x + y) \\ \bar{S}(x) &:= -S(y)C(x + y) + C(y)S(x + y) \end{aligned}$$

where x is any number in $\mathcal{R}$. Clearly:

$$\bar{C}(x)C(y) - \bar{S}(x)S(y) = C(x + y)$$

and:

$$\bar{C}(x)S(y) + \bar{S}(x)C(y) = S(x + y)$$

One can easily check that $\bar{C}' = -\bar{S}$ and $\bar{S}' = \bar{C}$. Moreover, $\bar{C}(0) = 1$ and $\bar{S}(0) = 0$. We infer that $\bar{C} = C$ and $\bar{S} = S$. ♮

25° Let Z be the subset of $\mathcal{R}$ consisting of all numbers z in $\mathcal{R}$ such that:

$$C(z) = 0$$

We refer to the numbers z in Z as *zeroes* of C. Let us suppose that $Z = \emptyset$. Of course, $C(0) = 1$. By the Intermediate Value Theorem, we would infer that, for any number z in $\mathcal{R}$, $0 < C(z)$. Since $S' = C$, S would be strictly increasing. Of course, $S(0) = 0$. Hence, $0 < S(1)$. Let b be any number in $\mathcal{R}$ such that $1 < b$ and $C(1) < S(1)(b-1)$. By the Mean Value Theorem, we could introduce a number c such that $1 < c < b$ and such that:

$$C(b) = C(1) - S(c)(b-1) < C(1) - S(1)(b-1) < 0$$

contrary to the foregoing inference. We conclude that $Z \neq \emptyset$.

26° Let $Z^- := Z \cap \mathcal{R}^-$ and $Z^+ := Z \cap \mathcal{R}^+$, so that $Z = Z^- \cup Z^+$. By article 23°, $Z^- = -Z^+$. Hence, $Z^+ \neq \emptyset$. Let $\bar{z}$ be the largest number in $(Z^+)_*$. By routine argument, one can prove that $0 < \bar{z}$ and $C(\bar{z}) = 0$. We may say that $\bar{z}$ is the smallest positive zero of C. By convention, one denotes $2\bar{z}$ by π, so that $\bar{z} = \pi/2$.

27° Clearly, the values of C and S on $(0, \pi/2)$ are positive. Moreover, C is strictly decreasing on $(0, \pi/2)$ and S is strictly increasing on $(0, \pi/2)$. Finally:

$$(\circ) \qquad C(0) = 1,\ C(\frac{\pi}{2}) = 0,\ S(0) = 0,\ S(\frac{\pi}{2}) = 1$$

28° Clearly, for any number z in $\mathcal{R}$:

$$(\circ) \qquad C(z + \frac{\pi}{2}) = C(z)C(\frac{\pi}{2}) - S(z)S(\frac{\pi}{2}) = -S(z)$$

and:

$$(\circ) \qquad S(z + \frac{\pi}{2}) = S(z)C(\frac{\pi}{2}) + C(z)S(\frac{\pi}{2}) = \quad C(z)$$

By repeated application of these relations, we find that, for any number z in $\mathcal{R}$:

$$(\circ) \qquad C(z + 2\pi) = C(z) \qquad \text{and} \qquad S(z + 2\pi) = S(z)$$

We say that C and S are *periodic* with period 2π.

29° In Figures 10 and 11, we organize many of the properties of C and S. In particular, in the Unit Circle Diagram (Figure 11), we interpret z to be the *radian measure* of the *oriented angle* comprised of the *initial ray*, in standard position, issuing from $(0,0)$ and passing through:

$$(1,0) = (C(0), S(0))$$

and the *terminal ray*, in variable position defined by z, issuing from $(0,0)$ and passing through:

$$(x,y) = (C(z), S(z))$$

The oriented angle opens clockwise if $z < 0$ and counterclockwise if $0 < z$. In effect, z is the (signed) length of the circular arc joining $(C(0), S(0))$ to $(C(z), S(z))$ and $C(z)$ and $S(z)$ are the first and second coordinates of the position on the unit circle defined by z. However, we must acknowledge that the interpretation just stated lacks precision, because we have not formally defined the concept of *arc length*. In fact, such a definition requires great care. It proves most efficient to define C and S independently, in terms of Power Series, then to define the length of the arc joining $(C(0), S(0))$ to $(C(z), S(z))$ to be z.

30° By the way, one converts radian measure z to *degree* measure ζ as follows:

$$\zeta^\circ = \frac{180^\circ}{\pi} z$$

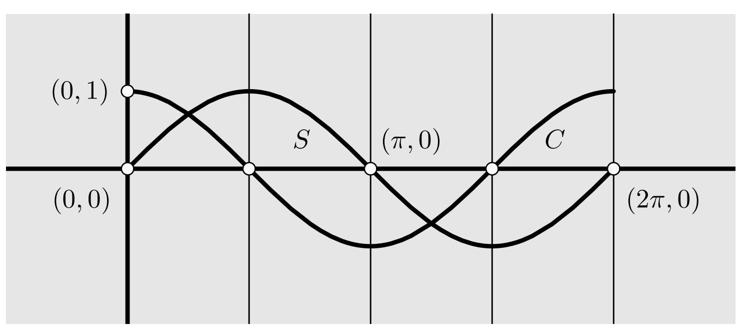


Figure 10: Trigonometric Functions

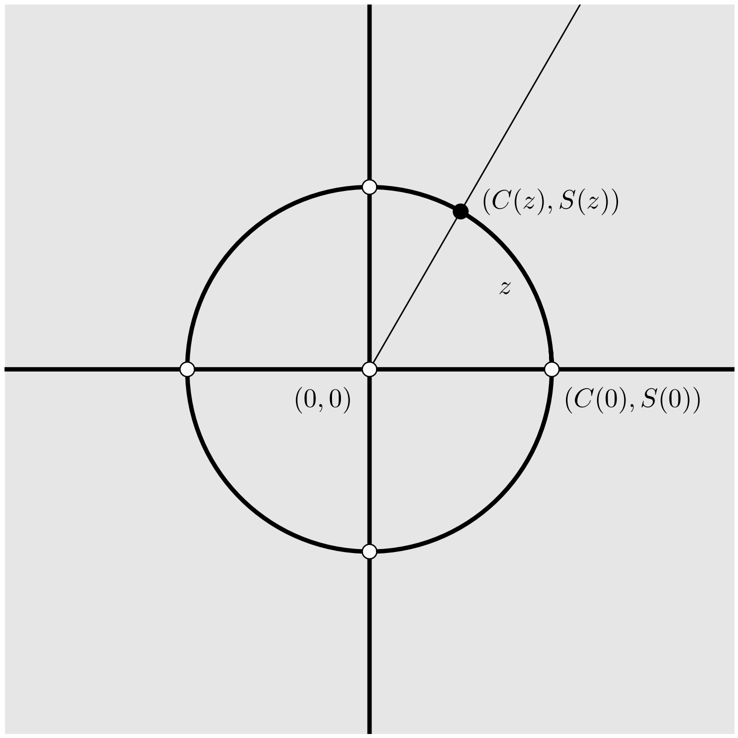


Figure 11: The Unit Circle

The Inverse Trigonometric Functions

31° Obviously, the restriction of C to $(0, \pi)$ is strictly decreasing and the restriction of S to $(-\pi/2, \pi/2)$ is strictly increasing. Moreover:

$$C((0,\pi)) = (-1,1) = S((-\pi/2,\pi/2))$$

Hence, we can introduce the functions $\hat{C}$ and $\hat{S}$ having common domain $(-1, 1)$ such that C and $\hat{C}$ are inverse to one another and S and $\hat{S}$ are inverse to one another. For any y in $(-1, 1)$, we have:

$$\begin{aligned} \hat{C}'(y) &= \frac{1}{C'(x)} = -\frac{1}{S(x)} = -\frac{1}{\sqrt{1-C^2(x)}} = -\frac{1}{\sqrt{1-y^2}} \\ \hat{S}'(y) &= \frac{1}{S'(x)} = \frac{1}{C(x)} = \frac{1}{\sqrt{1-S^2(x)}} = \frac{1}{\sqrt{1-y^2}} \end{aligned}$$

where $x := \hat{C}(y)$ or $x := \hat{S}(y)$, as needed.

32° One usually writes $arccos(y)$ instead of $\hat{C}(y)$ and $arcsin(y)$ instead of $\hat{S}(y)$. Hence:

$$\frac{d}{dy}arccos(y) = -\frac{1}{\sqrt{1-y^2}}$$

and:

$$\frac{d}{dy}arcsin(y) = \quad \frac{1}{\sqrt{1-y^2}}$$

The Tangent Function and its Inverse

33° Let T be the function having domain $(-\pi/2, \pi/2)$, defined in terms of C and S as follows:

$$T(z) = \frac{S(z)}{C(z)}$$

where z is any number in $(-\pi/2, \pi/2)$. One refers to T as the *tangent* function. One usually writes $tan(z)$ instead of $T(z)$, so that:

$$tan(z) = \frac{sin(z)}{cos(z)}$$

By the Quotient Rule, we find that:

$$T'(z) = \frac{C(z)S'(z) - C'(z)S(z)}{C^2(z)} = \frac{1}{C^2(z)}$$

where z is any number in $(-\pi/2, \pi/2)$.

34° Obviously, $T'((-\pi/2, \pi/2)) \subseteq \mathcal{R}^+$, so T is strictly increasing. Since:

$$S(-\pi/2) = -1, \ \ C(-\pi/2) = 0 = C(\pi/2), \ \ S(\pi/2) = 1$$

we find that:

$$T((-\pi/2, \pi/2))_* = \emptyset = T((-\pi/2, \pi/2))^*$$

By the Intermediate Value Theorem, $T(X) = \mathcal{R}$. At this point, we can introduce the function A having domain $\mathcal{R}$ such that T and A are inverse to one another. For any number y in $\mathcal{R}$, we have:

$$A'(y) = \frac{1}{T'(x)} = C^2(x) = \frac{1}{1+T^2(x)} = \frac{1}{1+y^2}$$

where $x := A(y)$.

35° One usually writes $arctan(y)$ instead of $A(y)$. Hence:

$$\frac{d}{dy}arctan(y) = \frac{1}{1+y^2}$$

Estimates of e and π

36° Let us make rough estimates of e and π. For e, we use a Taylor Polynomial for E and the corresponding Remainder Function. For π, we use an indirect computation involving T and A.

37° By the Mean Value Theorem, there is some number c in $\mathcal{R}$ such that $1 < c < 2$ and:

$$L(2) - L(1) = L'(c)(2-1) = \frac{1}{c}(2-1)$$

Hence, $1/2 < L(2)$ and $1 < L(4)$. We infer that $e = E(1) < 4$. By the Theorem of Taylor, there is a number d in $\mathcal{R}$ such that $0 < d < 1$ and such that:

$$(R_0^6 E)(1) = \frac{1}{7!}E(d)(1-0)^7$$

Hence:

$$\begin{aligned} e &= E(1) \\ &= (T_0^6 E)(1) + (R_0^6 E)(1) \\ &= \Big(\frac{1}{1} + \frac{1}{1} + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \frac{1}{120} + \frac{1}{720}\Big) + \Big(\frac{1}{5040}E(d)\Big) \end{aligned}$$

We infer that:

$$\frac{13699}{5040} < e < \frac{13699}{5040} + \frac{4}{5040} = \frac{13703}{5040}$$

By long division:

$$2.718 < e < 2.719$$

38° One can easily verify that:

$$T(x+y) = \frac{T(x) + T(y)}{1 - T(x)T(y)}$$

where x and y are any numbers in $(-\pi/2, \pi/2)$ for which $x + y$ is also in $(-\pi/2, \pi/2)$. Let:

$$r := A(\frac{1}{5}) \quad \text{and} \quad s := 4r - \frac{1}{4}\pi$$

We find that:

$$T(2r) = \frac{T(r)+T(r)}{1-T(r)T(r)} = \frac{(1/5)+(1/5)}{1-(1/5)(1/5)} = \frac{5}{12}$$

$$T(4r) = \frac{T(2r)+T(2r)}{1-T(2r)T(2r)} = \frac{(5/12)+(5/12)}{1-(5/12)(5/12)} = \frac{120}{119}$$

and:

$$T(s) = \frac{T(4r)-T(\pi/4)}{1+T(4r)T(\pi/4)} = \frac{(120/119)-1}{1+(120/119)} = \frac{1}{239}$$

Hence:

$$\pi = 16A(\frac{1}{5}) - 4A(\frac{1}{239})$$

By the Fundamental Theorem:

$$A(y) = \int_0^y \frac{1}{1+z^2}dz$$

where y is any number in $\mathcal{R}$. Clearly:

$$1 - z^2 < \frac{1}{1+z^2} < 1$$

and:

$$1 - z^2 + z^4 - z^6 < \frac{1}{1+z^2} < 1 - z^2 + z^4$$

where z is any number in $\mathcal{R}$. By integration:

$$y - \frac{1}{3}y^3 < A(y) < y$$

and:

$$y - \frac{1}{3}y^3 + \frac{1}{5}y^5 - \frac{1}{7}y^7 < A(y) < y - \frac{1}{3}y^3 + \frac{1}{5}y^5$$

where y is any number in $\mathcal{R}$. Hence:

$$(-4)(\frac{1}{239}) < (-4)A(\frac{1}{239}) < (-4)\big(\frac{1}{239} - \frac{1}{3}(\frac{1}{239})^3\big)$$

and:

$$16\big(\frac{1}{5} - \frac{1}{3}(\frac{1}{5})^3 + \frac{1}{5}(\frac{1}{5})^5 - \frac{1}{7}(\frac{1}{5})^7\big) < 16A(\frac{1}{5}) < 16\big(\frac{1}{5} - \frac{1}{3}(\frac{1}{5})^3 + \frac{1}{5}(\frac{1}{5})^5\big)$$

We infer that:

$$\frac{1231847548}{392109375} < \pi < \frac{670143059704}{213311234375}$$

By long division:

$$3.141 < \pi < 3.142$$

8 The Cartesian Rainbow

1° In 1637a, René Descartes published one of the celebrated works in the history of thought: **Discours de la Méthod**. In the preface to this work, Descartes set forth the central precepts of his scientific method, by which one would "acquire knowledge and avoid error." In the work itself, he presented three substantial discourses, which served as grand instances of successful application of his method. However, in course of time, the preface came to be viewed (and published) in its own right as a fundamental exposition of the scientific method.

The three discourses by Descartes were devoted to Geometry, Dioptrics, and Meteorology. In the first, Descartes initiated the study of geometry by arithmetic methods, that is, by means of coordinate systems. In the second, he described a quantitatively precise expression for the relation between the incident and refracted rays in context of refraction of light, the relation now known as the Law of Snell. Finally, in the third, he applied the Law of Snell to develop a compelling explanation of the Rainbow.

The Problem

2° In his **Meteorologica** (c0340b), Aristotle presented the rainbow as a problem to be solved. He required a description of:

- (•) the agents of formation of the rainbow

and he required explanations of:

- (•) its shape
- (•) its size
- (•) and its colors

Ab initio, Aristotle identified the agents of formation of the rainbow as the Sun, a rain shower, and the eye of an observer. He declared its shape to be a circular arc. These contributions have proved durable. The rest of his ideas, however, have proved misleading.

3° In his **Magnum Opus**, delivered to Pope Clement IV in 1268a, Roger Bacon reported his measurement of the angle of elevation of the peak of the rainbow at sunset: roughly 42°. This number serves as a measure of the size of the bow.

4° The concentric arcs of color in the rainbow, descending subtly through the visual spectrum:

$$red, orange, yellow, green, blue/indigo/violet$$

comprise the primary mystery. For more than two thousand years, efforts to explain the colors have developed, step for step, with efforts to explain the nature of Light itself.

Descartes' Diagrams

5° Let us summarize Descartes' explanation of the shape and size of the rainbow. In the first of the following two diagrams (Figure 12), one finds the Sun setting in the west, rain falling in the east, and an (astonished) observer taking note of the rainbow formed in the sky by the interaction of rays of light from the Sun and droplets of water in the shower.

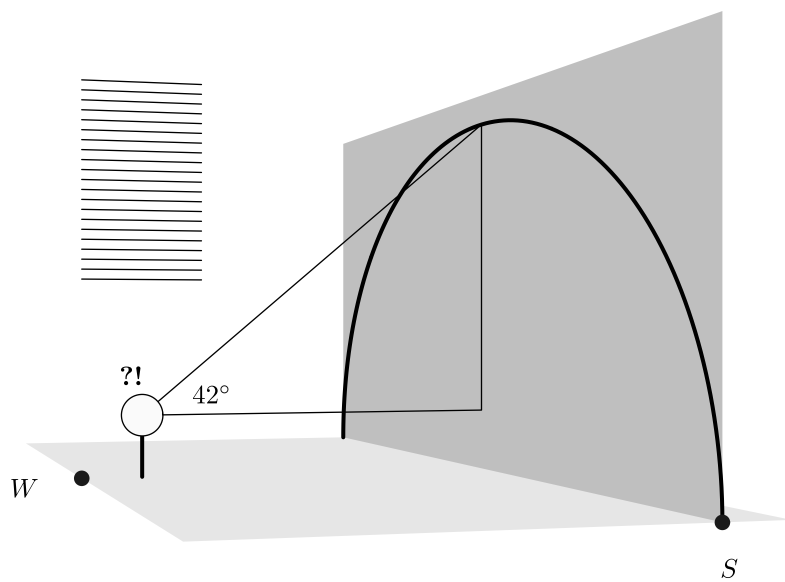


Figure 12: Observation

6° In the second diagram (Figure 13), one finds a particular ray of light and a particular raindrop magnified for inspection. The parameter y measures the elevation of the particular ray above the indicated axis of the raindrop. We have set the radius of the raindrop at one unit. In reality, the radius is roughly one millimeter. The stream of Particles composing the ray will meet the raindrop at point A, some being reflected but some being refracted into the body of the drop. Those particles which enter the drop at point A will meet the opposite surface at point B, some being refracted into the exterior but some being reflected. The particles which are reflected at point B will again meet the surface of the raindrop at point C, some being again reflected but some being refracted into the exterior. The particles which leave the raindrop at point C will have followed the path drawn in the diagram. Employing the Law of Snell, Descartes calculated the angle δ of deviation of the incident ray as a function of the parameter y:

$$(\star) \qquad \delta = \pi + 2\iota - 4\rho$$

where ι and ρ are the angles of Incidence and Refraction, respectively, as indicated in the diagram. Of course, ι and ρ are determined by y. See article 8°.

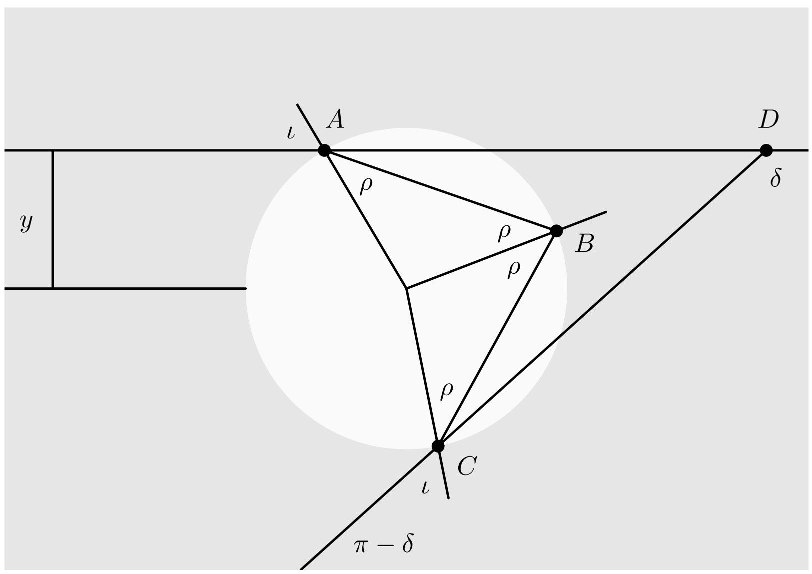


Figure 13: Deviation

Derivation of the Deviation Angle δ

7° The incident ray of light marked by the parameter y $(0 < y < 1)$ changes direction three times: at point A, at point B, and at point C. At point A, it turns clockwise through an angle of $\iota - \rho$; at point B, clockwise through an angle of $\pi - 2\rho$; and at point C, clockwise through an angle of $\iota - \rho$. Hence, the total angle δ of deviation of the incident ray is $\pi + 2\iota - 4\rho$.

The Calculations

8° In his famous tables of trigonometric functions (1612a), Bartholomeus Pitiscus recorded the values of the sine, tangent, and reciprocal cosine functions accurate to seven significant figures in steps of one sixth of one sixtieth of a degree. With immense patience, Descartes applied the tables to calculate approximate values of δ corresponding to the following values of y:

$$0.0,\ 0.1,\ 0.2,\ 0.3,\ 0.4,\ 0.5,\ 0.6,\ 0.7,\ 0.8,\ 0.9,\ 1.0$$

Intrigued by the emerging pattern, he then calculated approximate values of δ corresponding to the following values of y:

$$0.80,\ 0.81,\ 0.82,\ 0.83,\ 0.84,\ 0.85,\ 0.86,\ 0.87,\ 0.88,\ 0.89,\ 0.90$$

Assembling the numbers in a graph, he found compelling evidence that one special value of y, roughly 0.86, yielded a corresponding deviation angle δ of minimum value:

$$\frac{180^\circ}{\pi}\delta \approx 138^\circ$$

Let us denote those values of y and δ by $\bar{y}$ and $\bar{\delta}$ and let us refer to the ray with parameter $\bar{y}$ as the Cartesian Ray. Clearly, the cartesian ray will reach the eye of the observer at an angle of elevation:

$$180^\circ - \frac{180^\circ}{\pi}\bar{\delta} \approx 42^\circ$$

the angle of Bacon.

The Basic Graph

9° Let us apply the Calculus to analyze relation ($\star$). Let Δ be the *deviation* function having domain $(0, 1)$, defined as follows:

$$\delta \equiv \Delta(y) := \pi + 2\iota - 4\rho$$

where y is any number in $(0, 1)$. Of course, ι and ρ are determined by y. In fact, by Figure 13:

$$y = sin(\iota), \quad \iota = arcsin(y)$$

By the Law of Snell:

$$\boxed{sin(\iota) = \nu\, sin(\rho)}$$

Hence:

$$y = \nu sin(\rho), \quad \rho = arcsin(\frac{1}{\nu}y)$$

Now we can present Δ explicitly as a function of y:

$$\delta \equiv \Delta(y) := \pi + 2\,arcsin(y) - 4\,arcsin(\frac{1}{\nu}y)$$

where y is any number in $(0,1)$.

10° Descartes adopted the following value for the air/water Index of Refraction ν:

$$\nu = \frac{4}{3}$$

11° By differentiation, we find that:

$$\frac{d\delta}{dy} = 0 + 2\frac{d\iota}{dy} - 4\frac{d\rho}{dy} = \frac{2}{\sqrt{1-y^2}} - \frac{4}{\sqrt{\nu^2 - y^2}}$$

See article 32° in **Section 7**. By simple computation, we find that:

$$\begin{aligned} \frac{d\delta}{dy} &< 0 \quad \text{iff} \quad y < \sqrt{\frac{4-\nu^2}{3}} \\ \frac{d\delta}{dy} &= 0 \quad \text{iff} \quad y = \sqrt{\frac{4-\nu^2}{3}} \\ \frac{d\delta}{dy} &> 0 \quad \text{iff} \quad y > \sqrt{\frac{4-\nu^2}{3}} \end{aligned}$$

Clearly, the graph of Δ must take the form displayed in Figure 14. Moreover:

$$\bar{y} = \sqrt{\frac{4-\nu^2}{3}}$$

and:

$$\bar{\delta} = \pi + 2\,arcsin(\bar{y}) - 4\,arcsin(\frac{1}{\nu}\bar{y})$$

For the value $\nu = 4/3$, we find that:

$$\bar{y} = 0.8607$$

and:

$$180^\circ - \frac{180^\circ}{\pi}\bar{\delta} = 180^\circ - \frac{180^\circ}{\pi}2.4080 = 180^\circ - 138.0^\circ = 42.0^\circ$$

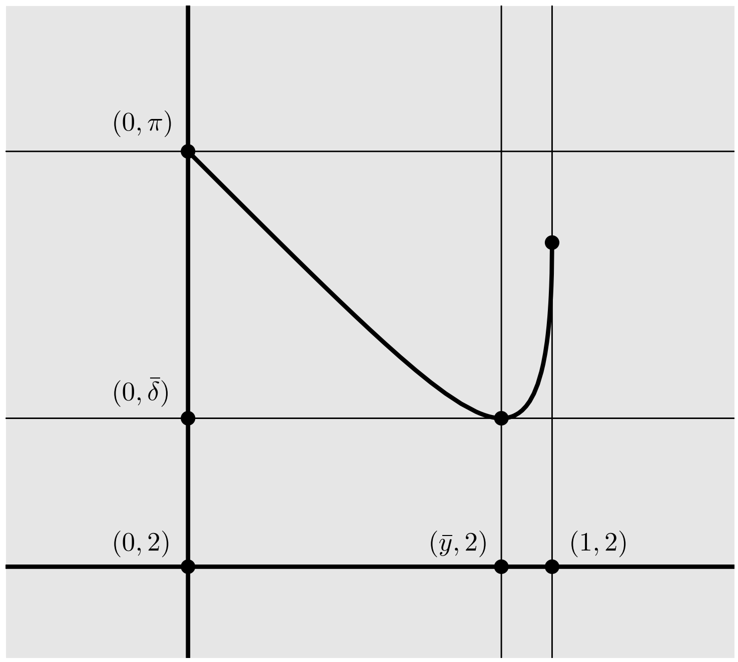


Figure 14: The Cartesian Graph

12° The foregoing analysis makes sense only if $1 < \nu < 2$.

13° Let us introduce the *elevation* function H, having domain $(1, 2)$:

$$\epsilon \equiv H(\nu) := \pi - \bar{\delta} = 4\,arcsin(\frac{1}{\nu}\sqrt{\frac{4-\nu^2}{3}}) - 2\,arcsin(\sqrt{\frac{4-\nu^2}{3}})$$

where ν is any number in $(1, 2)$. With diligence, one can show that:

$$\frac{d\epsilon}{d\nu} = \cdots\cdots\cdots = -\frac{2}{\nu}\sqrt{\frac{4-\nu^2}{\nu^2-1}}$$

Hence, H is strictly decreasing. See the following article 20°.

Interpretation: Its Size

14° The coincidence between Descartes' calculation and Bacon's measurement is, of course, striking. However, it does not by itself constitute an **explanation** of the size of the rainbow. Scientific explanation requires more than a coincidence between construction and measurement. It requires that the coincidence itself be subject to rational interpretation. The process of explaining natural phenomena is inherently regressive, terminating only when it reaches a primary layer of uncontested assent.

15° But Descartes pressed his discovery to a deeper level. He called attention to the **significance** of the minimum value of a function. Since the cartesian ray yields a deviation angle of minimum value, the light rays nearby to that ray will emerge from the raindrop *closely packed*. They will create for the eye of the observer the impression of a bright spot in the sky at an angular elevation of 42°. In contrast, the light rays far from the cartesian ray will emerge more or less evenly spaced and, in comparison with the Cartesian Pack, will create for the eye of the observer impressions substantially less bright. The following Ray Diagram (Figure 15) makes everything clear.

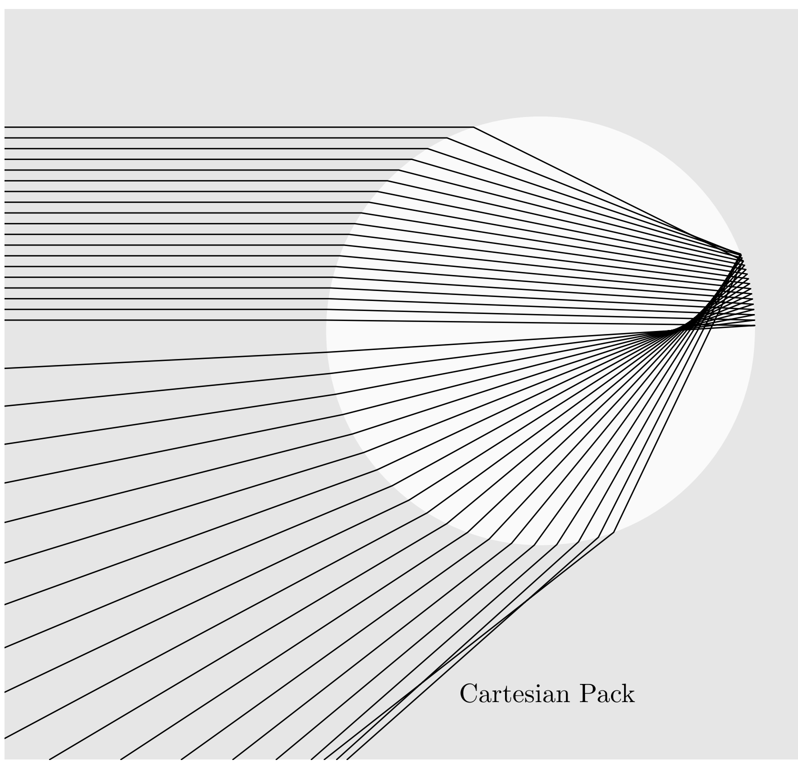


Figure 15: Ray Diagram

The Cinematic Effect

16° Of course, raindrops fall. While the bright spot seems to hang in the sky at an angular elevation of 42°, the raindrops creating it give way, moment by moment, to those above. Moreover, raindrops fall rapidly. The bright spot seems to hang continuously. It does not flicker.

Interpretation: Its Shape

17° Descartes' construction is symmetric about the line issuing from the eye of the observer, parallel to the line of the horizon. Accordingly, any raindrop for which the angle between the eye-raindrop line and the eye-horizon line is 42° will contribute to the impression of the rainbow for the observer. Technically, then, the rainbow consists of a circular cone of *directions*, with vertex at the eye of the observer and with angle of aperture equal to 42°. Hence, Descartes' construction explains not only the size but also the circular shape of the rainbow.

Jubilation

18° Descartes attempted but failed to explain the distribution of colors in the rainbow.

19° While the theory of Descartes would in due course prove to be only the first step in a complex sequence of refinements, continuing to the present day, one can hardly help but share in his jubilation:

> *"Those who have understood all which has been said in the treatise will no longer see anything in the clouds in the future for which they will not easily understand the cause."* (**Les Météors**)

Its Colors

20° In 1704a, Isaac Newton published his treatise: **Optiks**. In this work, Newton presented his theory of color and his application of that theory to numerous observations of natural bodies. In particular, in Proposition 9, Problem 4 of Book 1, Part 2, he set the following problem:

> *"By the discovered properties of light, to explain the colours of the rainbow."*

To solve the problem, Newton applied the theory of Descartes but he introduced a new feature: the parameter ν (the index of refraction for air/water) varied for various colors of visible light, being smallest for red light and largest for blue:

$$\nu_r = 1.331$$
$$\nu_b = 1.343$$

As a result, the angular elevation of the cartesian pack varied for various colors:

$$\frac{180.00^\circ}{\pi}\epsilon_r = 42.37^\circ$$
$$\frac{180.00^\circ}{\pi}\epsilon_b = 40.65^\circ$$

where $\epsilon_r := H(\nu_r)$ and $\epsilon_b := H(\nu_b)$. Newton's theory entailed that the vertical span of the rainbow should be:

$$42.37^\circ - 40.65^\circ = 1.72^\circ$$

which proved to be in rough agreement with observation.

21° One applies the term Dispersion to refer to optical phenomena which depend specifically upon the color (that is, the Frequency) of light. Thus, one may say that the distribution of colors in the rainbow is an effect of dispersion. However, one may rightly ask whether such a statement explains anything at all. The fact of dispersion appears as an empirical irreducible. Even under the sophisticated theory of Electricity and Magnetism perfected by James C. Maxwell in the Nineteenth Century, the effects of dispersion are traceable to the empirically determined parameters of Electric Permittivity and Magnetic Permeability of the medium under study. In any case, Newton did not explain, in terms of more fundamental concepts and constructions, the dependence of the index of refraction for air/water upon color.

Informed Seeing

22° Under certain conditions, streaks of green and purple appear at the lower edge of the peak of the rainbow. One refers to the streaks as Supernumerary Arcs. To the naive observer, these arcs are simply a part of the sweep of color in the rainbow. To the informed observer, however, they pose a new problem. The arcs have no "place" in the Cartesian/Newtonian theory. To explain the supernumerary arcs, one must invoke not the Particle Model but the Wave Model of Light, one must investigate the optical phenomenon of Interference, and one must analyze the Perception of Color in the Eye/Mind of the observer.

Existence/Uniqueness

23° One may rightly ask whether the rainbow "exists," and, if so, whether it is "unique." For a given observer, the observed rainbow is not an object but a conical assembly of directions. For distinct observers, the observed rainbows are distinct. Unlike the circumstance in which such sensory impressions as tree-like legs and a snake-like trunk could be explained to a group of blind men as aspects of the same underlying Elephant, for the aggregate of sensory impressions to which we refer as the rainbow, there is no underlying common object, unless one is content to declare it to be a State of the Atmosphere. The rainbow shares in the subtlety of distinctions between Matter and Light, between Thing and Process.

References

24° Very often, a secondary rainbow appears in the sky, above the primary bow. Can one adapt the cartesian explanation to the secondary bow? This and many other questions are treated in the following books:

The Rainbow: From Myth to Mathematics, Carl B. Boyer, 1987a

Geometry Civilized, J. L. Heilbron, 1998a

Light and Color in the Outdoors, M. G. J. Minnaert, 1993a

Introduction to Meteorological Optics, R. A. R. Tricker, 1970a

By study of these books, one will be able to form answers to such questions as the following:

(◦) Why does one see just two rainbows?

(◦) Does the size of the raindrops effect the appearance of the rainbow?

(◦) Should one expect to see a rainbow in a shower of sulphuric acid on Venus?

(◦) ...... in a shower of lead sulphate on Earth?

(◦) Would an Orca see a rainbow in a quiet sea, formed in a rising shower of air bubbles?

9 Problems

$1^{\bullet}$ Memorize the Greek alphabet:

α	alpha	A
β	beta	B
γ	gamma	Γ
δ	delta	Δ
ϵ	epsilon	E
ζ	zeta	Z
η	eta	H
θ	theta	Θ
ι	iota	I
κ	kappa	K
λ	lambda	Λ
μ	mu	M
ν	nu	N
ξ	xi	Ξ
o	omicron	O
π	pi	Π
ρ	rho	P
σ	sigma	Σ
τ	tau	T
υ	upsilon	Υ
ϕ	phi	Φ
χ	chi	X
ψ	psi	Ψ
ω	omega	Ω

$2^{\bullet}$ Show that, for any number x in $\mathcal{R}$, if $x + x = 0$ then $x = 0$. Note that, for any number y in $\mathcal{R}$, $0 \cdot y + 0 \cdot y = (0+0) \cdot y = 0 \cdot y$. Conclude that $0 \cdot y = 0$.

$3^{\bullet}$ Show that, for any numbers x and y in $\mathcal{R}$, if $xy = y$ then $x = 1$ or $y = 0$.

$4^{\bullet}$ In article 2° of **Section 1**, one finds the definition of the multiplicative inverse of a number x in $\mathcal{R}$. In that context, one finds the constraint that $x \neq 0$. Why is that constraint imposed?

$5^{\bullet}$ Show that, for any numbers x and y in $\mathcal{R}$, if $0 < x$ and $0 < y$ then $0 < x + y$ and $0 < x \cdot y$.

$6^{\bullet}$ Show that $-1 < 0 < 1$.

$7^\bullet$ Show that, for any numbers x and y in $\mathcal{R}$, if $0 < x$ and $0 < y$ then:

$$\sqrt{xy} \leq \frac{1}{2}(x + y)$$

To that end, note that $0 \leq (\sqrt{x} - \sqrt{y}\,)^2$. Conclude that if $xy = 1$ then:

$$2 \leq x + y$$

$8^\bullet$ Show that, for any number x in $\mathcal{R}$ and for any integer k in $\mathcal{Z}$, if $x \neq 1$ and $0 < k$ then:

$$1 + x + x^2 + x^3 \ \cdots \ + x^{k-1} = \frac{1 - x^k}{1 - x}$$

$9^\bullet$ Apply Mathematical Induction to show that, for any positive integer k, the sum of the first k odd integers equals k^2:

$$1 + 3 + 5 + \ \cdots \ + (2k - 1) = k^2$$

$10^\bullet$ Note that:

$$\begin{aligned} (1 - \tfrac{1}{2}) &= \tfrac{1}{2} \\ (1 - \tfrac{1}{2})(1 - \tfrac{1}{3}) &= \tfrac{1}{3} \\ (1 - \tfrac{1}{2})(1 - \tfrac{1}{3})(1 - \tfrac{1}{4}) &= \tfrac{1}{4} \end{aligned}$$

Guess the general "law." Prove it by Mathematical Induction.

$11^\bullet$ Apply Mathematical Induction to show that, for any integer k in $\mathcal{Z}$, if $3 < k$ then $2^k < k!$.

$12^\bullet$ With reference to article 29° in **Section 1**, let b be any integer for which $2 \leq b$. Let δ be any series of digits subject to the stated conditions. One says that δ is *preperiodic* iff there are integers k and ℓ such that $0 < k$ and, for any integer j, if $j + k < \ell$ then $\delta_j = \delta_{j+k}$. Let x be the positive number represented by δ. Show that δ is preperiodic iff x is rational. Find the base 60 representation of $1/7$.

$13^\bullet$ Let C be a circle in the Euclidean plane for which the radius is 1. Let P_1 be an equilateral triangle in the plane circumscribed about C and let C_1 be the circle in the plane circumscribed about P_1. Let P_2 be a square in the plane circumscribed about C_1 and let C_2 be the circle in the plane circumscribed about P_2. Let P_3 be a regular pentagon in the plane circumscribed about C_2 and let C_3 be the circle in the plane circumscribed about P_3. In general, for each positive integer j, let P_{j+1} be a regular $(j+2)$-gon in the plane circumscribed about C_j and let C_{j+1} be the circle in the plane circumscribed about P_{j+1}. Let X be the subset of $\mathcal{R}$ comprised of the radii of the various circles. Show that X^* is not empty. Find the smallest number r in X^*. See Figure 16.

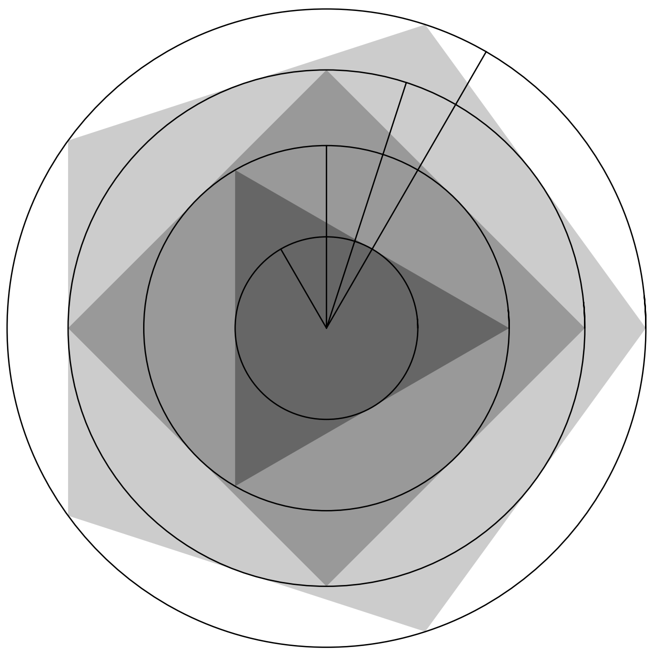


Figure 16: Radii

14• With reference to the examples of functions described in article 7° of **Section 2**, verify that $F_1(X_1) = [0, \rightarrow)$, $F_2(X_2) = X_2$, and $F_3(X_3) = (-1, 1]$.

15• Let z be any number in $\mathcal{R}$. Let $[z]$ stand for the integer k in $\mathcal{Z}$ such that:

$$k \leq z < k + 1$$

Let F be the function having domain $[-2, 2]$, defined as follows:

$$F(x) = [x^2 - 1]$$

Draw the graph of F.

16• Let r be any number in $\mathcal{R}$ for which $r < -1$ or $3 < r$ and let F_r be the function having domain $\mathcal{R}$, defined as follows:

$$F_r(x) := r\,x\,(1-x)$$

Find all numbers y in $\mathcal{R}$ such that:

$$F_r(F_r(y)) = y$$

To that end, let G be the function having domain $\mathcal{R}$, defined as follows:

$$\begin{aligned} G(x) &:= x - F_r(F_r(x)) \\ &= (1-r^2)x + r^2(1+r)x^2 - 2r^3x^3 + r^3x^4 \end{aligned}$$

Note that:

$$G(x) = 0 \quad \text{iff} \quad F_r(F_r(x)) = x$$

where x is any number in $\mathcal{R}$. Note also that G is a fourth degree polynomial and that:

$$G(0) = 0,\ G(\frac{r-1}{r}) = 0 \ \text{ because } \ F(0) = 0,\ F(\frac{r-1}{r}) = \frac{r-1}{r}$$

As a result, there must exist numbers a, b, and c in $\mathcal{R}$ such that:

$$G(x) = (x-0)(x - \frac{r-1}{r})(ax^2 + bx + c)$$

where x is any number in $\mathcal{R}$. Find a, b, and c. Of course, these numbers will depend upon r. In fact, $a = r^3$, $b = -r^2(1+r)r$, and $c = r(1+r)$. Now finish the problem.

17• Let H_0 be the function having domain $X := \mathcal{R}^- \cup \mathcal{R}^+$, defined as follows:

$$H_0(x) := \begin{cases} -1 & \text{if } x < 0 \\ \ \ 1 & \text{if } 0 < x \end{cases}$$

Let h be any number in $\mathcal{R}$. Let H be the function having domain $\mathcal{R} = X \cup \{0\}$, defined by *extending* H_0 as follows:

$$H(x) := \begin{cases} H_0(x) & \text{if } x \neq 0 \\ h & \text{if } x = 0 \end{cases}$$

Find a number h in $\mathcal{R}$ such that H is continuous at 0 or show that it cannot be done.

18• Let F be the function having domain $\mathcal{R}$, defined as follows:

$$F(x) = \frac{x-6}{x^2+1}$$

Find a specific number u in $\mathcal{R}^+$ such that, for any number x in $\mathcal{R}$, if $|x-1| < u$ then:

$$|F(x) - F(1)| < \frac{1}{100}$$

To that end, note that:

$$\begin{aligned} |F(x) - F(1)| &= |\frac{x-6}{x^2+1} - \frac{1-6}{1^2+1}| \\ &= |\frac{5x^2+2x-7}{2(x^2+1)}| \\ &\leq |5x^2+2x-7| \\ &= |(x-1)(5x+7)| \\ &\leq |x-1|(5|x|+7) \end{aligned}$$

Note also that, for any number u in $\mathcal{R}^+$, if $|x-1| < u$ then $|x| < 1+u$, so that $5|x|+7 < 5u+12$. Now find an appropriate u.

19• Let F be the function having domain $(0, 1]$, defined as follows:

$$F(x) := \frac{1}{x}$$

Note that F is continuous. Show that F is not bounded.

20• Let X be any interval in $\mathcal{R}$, let F be a function having domain X, and let a be any number in X. Let F be continuous at a and let $0 < F(a)$. Show that there is a number u in $\mathcal{R}^+$ such that, for any number x in $\mathcal{R}$, if $x \in X$ and $|x-a| < u$ then $0 < F(x)$.

21• Let F be the function having domain $(-1, \rightarrow)$, defined as follows:

$$F(x) := \frac{1-x}{1+x}$$

Apply the Basic Properties of continuous functions to show that F is continuous. Find the range of F.

22$^\bullet$ Let F be the function having domain $[0, \rightarrow)$, defined as follows:

$$F(x) := x^2$$

Note that F is continuous. Let y be any number in $[0, \rightarrow)$. Show that there is exactly one number x in $[0, \rightarrow)$ such that $F(x) = y$. That is:

$$x = \sqrt{y}$$

To that end, apply the Intermediate Value Theorem.

23$^\bullet$ Let F be a function having domain $[0, 1]$. Let F be continuous and let the range of F be a subset of $[0, 1]$. Show that there must be a number z in $[0, 1]$ such that $F(z) = z$. To that end, introduce the function G having domain $[0, 1]$, defined as follows:

$$G(x) := x - F(x)$$

Apply the Intermediate Value Theorem. Draw a diagram to illustrate the foregoing result.

24$^\bullet$ Let F be the function having domain $\mathcal{R}^+$, defined as follows:

$$F(x) = \sqrt{x}$$

Let a and x be any numbers in $\mathcal{R}^+$. Note that:

$$(\sqrt{x} - \sqrt{a})(\sqrt{x} + \sqrt{a}) = x - a$$

and that:

$$|\sqrt{x} - \sqrt{a}\,| = \frac{|x - a|}{\sqrt{x} + \sqrt{a}} \leq \frac{1}{\sqrt{a}}|x - a|$$

Apply the foregoing observation to prove that F is continuous.

25$^\bullet$ Let F be the function having domain $\mathcal{R}^+$, defined as follows:

$$F(x) = \sqrt{x}$$

Show that F is differentiable at 1 and that:

$$F'(1) = \frac{1}{2}$$

26$^\bullet$ Calculate:

$$\frac{d}{dx} log(\frac{1}{cos(x)} + tan(x))$$

27• Compute:

(01) $$\frac{d}{dx}(6 - x + 3x^2)\big|_{x=2}$$

(02) $$\frac{d}{dy}\, 7y^9 \,\big|_{y=b}$$

(03) $$\frac{d}{dz}\,\frac{1 - z^3}{1 + z^3}\,\big|_{z=0}$$

(04) $$\frac{d}{dz}\sqrt{\frac{1 - z^3}{1 + z^3}}\,\Big|_{z=c}$$

(05) $$\frac{d}{dx}(x^2\, log(x))\big|_{x=1}$$

(06) $$\frac{d}{dy}\sqrt{1 + \sqrt{y}}\,\Big|_{y=4}$$

(07) $$\frac{d}{dx}\Big((\frac{1 + x^2}{x^{1/2}})(1 + x^{1/2})^{1/2}\Big)\big|_{x=1}$$

(08) $$\frac{d}{dx}(x^3 + \frac{x^2 log(x)}{sin(x)})\big|_{x=\pi/2}$$

(09) $$\frac{d}{dx}\,\frac{log(\frac{1}{2}x)tan(\frac{\pi}{4}x)}{1 + x^2}\,|_{x=1}$$

(10) $$\frac{d}{dx} log(\frac{1 - \sqrt{1 + x}}{1 + \sqrt{1 + x}})$$

28• Describe the function F having domain $\mathcal{R}$ such that $F(0) = 3$ and, for any number x in $\mathcal{R}$, $F'(x) = -2F(x)$.

$29^\bullet$ Let F be the function having domain $\mathcal{R}^+$, defined as follows:

$$F(x) = \frac{1}{2}(1+x) - x^{1/2}$$

Find the values of x such that $F'(x) < 0$. In turn, find the values of x such that $F'(x) = 0$ and the values of x such that $0 < F'(x)$. Use the information to sketch the graph of F.

$30^\bullet$ Sketch the graph of the following function:

$$F(x) = arctan(x + \frac{1}{x})$$

where x is any number in $\mathcal{R}^+$. Portray the endpoint behaviour carefully.

$31^\bullet$ Let a, b, and c be any numbers in $\mathcal{R}$. Let F be the function defined as follows:

$$F(x) = \begin{cases} x^2 & \text{if } x < c \\ ax + b & \text{if } c \leq x \end{cases}$$

where x is any number in $\mathcal{R}$. Show that F is differentiable at c iff a, b, and c satisfy the following relations: • • • • • ?

$32^\bullet$ Let h, c, k, and T be certain numbers in $\mathcal{R}^+$. Let F be the function having domain $\mathcal{R}^+$, defined as follows:

$$F(x) = hc^2x^{-5}(exp(\frac{hc}{kTx}) - 1)^{-1}$$

Sketch the graph of F. To do so, show that there is exactly one number y in $\mathcal{R}^+$ such that:

$$(\circ) \qquad y\, exp(y)(exp(y) - 1)^{-1} = 5$$

In fact, the value of y is approximately 4.965. In turn, show that, for any number x in $\mathcal{R}^+$, if $0 < x < hc/kTy$ then $0 < F'(x)$; if $x = hc/kTy$ then $F'(x) = 0$; and if $hc/kTy < x$ then $F'(x) < 0$. Conclude that:

$$\bar{x} = \frac{hc}{kTy}$$

is the maximum number for F. Obviously:

$$\bar{x}T = \frac{hc}{ky} \approx \frac{hc}{4.965k}$$

Note that, for fixed values of h, c, and k, $\bar{x}$ and T are inversely related.

$33^\bullet$ Let F be the function defined as follows:

$$F(x) = \frac{1}{1+x^2}$$

where x is any number for which $0 \leq x \leq 1$. Let P be the partition of the interval $[0,1]$ comprised of the numbers:

$$0 = \frac{0}{4}, \frac{1}{4}, \frac{2}{4}, \frac{3}{4}, \frac{4}{4} = 1$$

Calculate the corresponding lower and upper sums:

$$L(F,P),\ U(F,P)$$

With reference to article 35° in **Section 7**, show that:

$$\int_0^1 \frac{1}{1+x^2}dx = \frac{\pi}{4}$$

Is that result consistent with your calculations?

$34^\bullet$ Let a, b, and c be any numbers in $\mathcal{R}$ but let $a \neq 0$. Let F be the *quadratic polynomial* defined as follows:

$$F(x) = ax^2 + bx + c$$

where x any number in $\mathcal{R}$. Find specific values for a, b, and c such that:

$$F(0) = 0, \ \ F(1) = 0, \ \ \int_0^1 F(x)dx = 1$$

$35^\bullet$ Show that:

$$\int_{1/2}^1 \frac{1}{x}dx = \int_1^2 \frac{1}{x}dx$$

$36^\bullet$ Find all antiderivatives for the function F having domain $(-\pi/2, \pi/2)$, defined as follows:

$$F(x) = log(cos(x))tan(x)$$

$37^\bullet$ Find all continuous functions F for which:

$$\int_0^x F(y)dy = (F(x))^2 + c$$

where c is a constant.

38• Compute:

(01) $$2\int_0^1 \frac{x}{1+x^2}dx$$

(02) $$\int_1^2 \frac{1}{x}log(x)dx$$

(03) $$\int_1^4 \sqrt{x}\,log(x)dx$$

(04) $$\int_1^e w^3log(w)dw$$

(05) $$\int_0^1 \frac{1}{3}w^2\frac{1}{1+w^3}dw$$

(06) $$\int_0^\pi sin^3(w)dw$$

39• Compute:

(01) $$\frac{d}{dx}\int_0^{log(x)} exp(y^2)dy$$

(02) $$\frac{d}{dx}\int_0^{tan(x)} arctan(w)dw\,\Big|_{x=\pi/4}$$

(03) $$\frac{d}{dx}\int_0^{x(1+x)} \frac{1-w}{1+w}dw\,\Big|_{x=1}$$

(04) $$\frac{d}{dx}\int_0^{tan(x)} arctan(w)dw\,\Big|_{x=\pi/4}$$

$40^\bullet$ Let ϵ be any number for which $0 < \epsilon < 1$. Show that, for any positive number y, there is exactly one positive number x such that:

$$y = x - \epsilon\, sin(x)$$

Let G be the function so defined:

$$y = G(y) - \epsilon\, sin(G(y))$$

where y is any positive number. Calculate:

$$G'(x - \epsilon\, sin(x))$$

where x is any positive number.

$41^\bullet$ Let F be the function defined as follows:

$$F(x) := 1 - 2(1-x)^{1/2}$$

where x is any number for which $-1 < x < 1$. Let k be any nonnegative integer. Show that:

$$(T_0^k F)(x) < F(x)$$

where x is any number for which $0 < x < 1$. Show that if k is even then:

$$F(x) < (T_0^k F)(x)$$

while if k is odd then:

$$(T_0^k F)(x) < F(x)$$

where x is any number for which $-1 < x < 0$. Sketch the graphs of F and of $T_0^2 F$.

$42^\bullet$ Let a be any number in $\mathcal{R}$. Let F be the function defined as follows:

$$F(x) = (1+x)^a = exp(a\, log(1+x))$$

where x is any number in $\mathcal{R}$ for which $-1 < x < 1$. Let k be any nonnegative integer. Calculate:

$$(T_0^k F)(x)$$

where x is any number in $\mathcal{R}$.

$43^\bullet$ With reference to article 13° in **Section 8**, show that:

$$\frac{d\epsilon}{d\nu} = -\frac{2}{\nu}\sqrt{\frac{4-\nu^2}{\nu^2-1}}$$

$44^\bullet$ Let p, q, r, s, v, and w be any numbers for which $p < 0 < r$, $s < 0 < q$, $0 < v$, and $0 < w$. Let F be the function defined as follows:

$$F(x) = \frac{1}{v}\sqrt{(x-p)^2 + q^2} \; + \; \frac{1}{w}\sqrt{(x-r)^2 + s^2}$$

where x is any number. Show that there is precisely one number $\bar{x}$ such that, for any number x, if $x < \bar{x}$ then $F'(x) < 0$ and if $\bar{x} < x$ the $0 < F'(x)$. Note that $F(\bar{x})$ is the minimum value of F. With reference to Figure 17, verify that:

$$\frac{\bar{x} - p}{\sqrt{(\bar{x}-p)^2 + q^2}} = \nu \, \frac{r - \bar{x}}{\sqrt{(\bar{x}-r)^2 + s^2}}$$

where $\nu = v/w$. That is:

(S) $$sin(\iota) = \nu \, sin(\rho)$$

Suitably interpreted, the foregoing relation comprises *Snell's Law*.

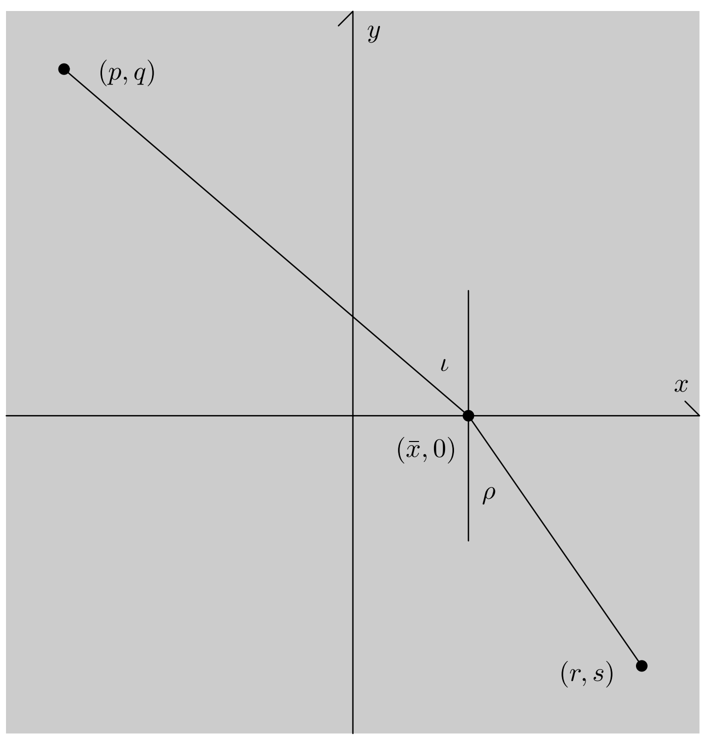


Figure 17: Snell's Law

45• Let ϵ and d be numbers for which $0 < \epsilon < 1$ and $0 < d$. Let r be the function of ϕ defined as follows:

$$\frac{r}{d - r\,cos(\phi)} = \epsilon$$

That is:

$$r = \frac{d\epsilon}{1 + \epsilon\,cos(\phi)}$$

See Figure 18. Let x and y be the functions of ϕ defined as follows:

$$x = r\,cos(\phi)$$
$$y = r\,sin(\phi)$$

Find the minimum and maximum values of x. Let a be half the difference between them. Find the minimum and maximum values of y. Let b be half the difference between them. Naturally, the values of a and b will depend upon ϵ and d. Note that $0 < b < a$. Show that:

$$a^2 - b^2 = \epsilon^2 a^2$$

With reference to Figure 18, find the relation between the angles η and ϕ.

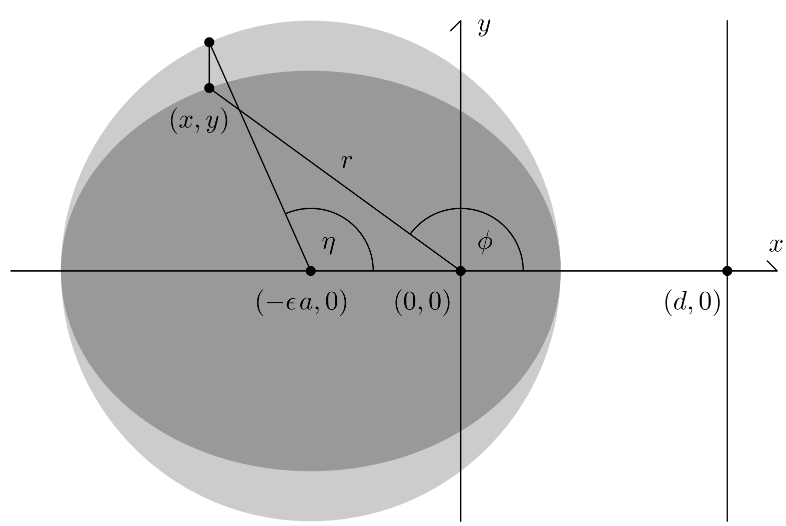


Figure 18: An Ellipse

46• Consider the standard parabola in the plane, described by the following relation between the coordinates x and y:

$$y = x^2$$

The *focus* of this parabola is the point $(0, 1/4)$. Imagine a ray of light which issues from the focus, meets the parabola at the point (x, x^2), then reflects in such a way to make *equal angles* with the perpendicular. See Figure 19. Show that, no matter what ray of light be imagined, the reflected ray is vertical. This is the principle of the *parabolic mirror*. All rays of light issuing from the focus emerge from the mirror parallel to its axis.

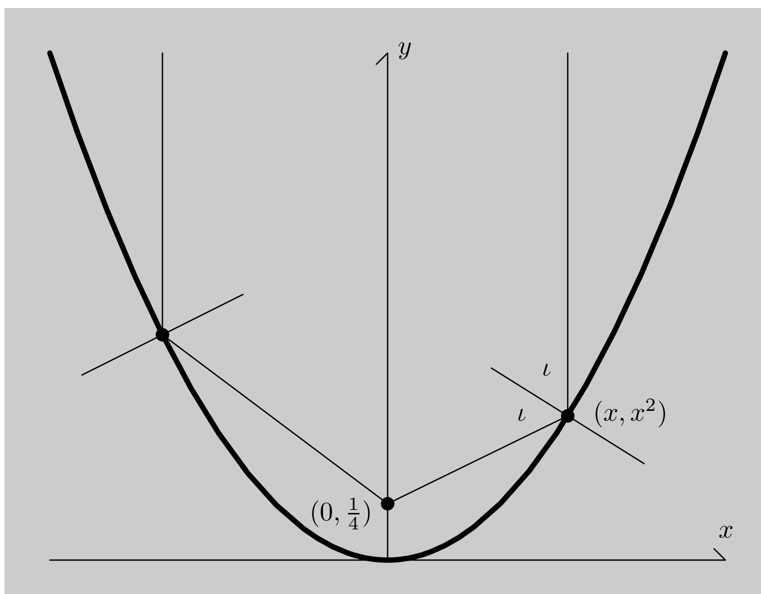


Figure 19: A Parabolic Mirror

Bibliography

History

47• For a lucid account of the history of our subject, we recommend the following book:

The History of the Calculus, Carl B. Boyer, 1959a

Questions

48• What is a Real Number? What does it mean for numbers, such as 0 and 1, to exist? Are the assumptions about numbers true? Do the assumptions lead to contradictions? These questions are the province of Mathematical Logic: the formal study of the Foundations of Mathematics. The answers to these questions prove to be subtle, complex, and fascinating. For an introduction to such matters, one might consult the essay:

Math000.pdf

posted on my website:

http://www.reed.edu/~wieting/essays.html