A. The "Foundational Theorems"

0. Introduction.
1. The Intermediate Value Theorem.
2. The Maximum Value Theorem.
3. The Integration Theorem.
4. Other Theorems.

A.0. Introduction

The proofs of the foundational theorems, **A.1–3** below, are based on a couple of definitions and facts.

Definition. *A digit is an element of the set* $\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$.

Fact 1. *Corresponding to any sequence of digits* $c_1, c_2, c_3, \ldots$, *there is a unique real number*

$$.c_1c_2 \cdots := c_1 \cdot 10^{-1} \; + \; c_2 \cdot 10^{-2} \; + \cdots$$

in the interval $[0, 1]$.

We will not prove Fact 1, but accept it as true. It could be part of a definition of the real numbers—a full definition would need to point out that ocasionally two different sequences need to considered as the same real—or it could be derived from other definitions. We also need the definition of continuous function and a simple fact about continuous functions that is immediate from the definition.

Definition. Let f be a function from $[0, 1]$ to $\mathbf{R}$ and let c be any point of $[0, 1]$. We say f is *continuous at* c if, given any $\epsilon > 0$, there is an open interval (a, b) containing c such that $|f(x) - f(c)| < \epsilon$ for all $x \in (a, b) \cap [0, 1]$.

Fact 2. *Suppose* f *is continuous at* c. *Let* k *be any real number. If* $f(c) < k$ *(respectively,* $f(c) > k$*) then there is an open interval* (a, b) *containing* c *such that* $f(x) < k$ *(respectively,* $f(x) > k$*) for all* $x \in (a, b) \cap [0, 1]$.

Proof. Let $\epsilon := \frac{1}{2}|f(c) - k|$, and apply the definition. /////

Apart from these definitions and facts and some basic logic and arithmetic, the proofs we present are entirely self-contained. In particular, we make no reference to the least upper bound principle (though of course it is implicit in Fact 1). This does not mean that this section will be easy for an undergraduate to understand. To the contrary, this section is presented in the style that a typical mathematician would learn in graduate school and employ in professional writing. This style is very efficient and accurate in conveying arguments, but only a reader who is thoroughly acquainted with it and understands its conventions is likely to be comfortable reading it. These notes do not include any commentary on the place the foundational theorems have in calculus. For that, you need to look elsewhere.

Note that the theorems are all stated for the closed unit interval. In this we suffer no essential loss of generality. The function $g(x) := a + (b - a)x$ is a continuous function with continuous inverse of interval $[0, 1]$ onto the interval $[a, b]$. Together with the fact that a composition of continuous functions is continuous, this enables one to generalize the theorems to arbitrary intervals.

A.1. The Intermediate Value Theorem.

Theorem. *Suppose* $f : [0, 1] \to \mathbf{R}$ *is continuous and*

$$f(0) < 0 \leq f(1).$$

Then there is $c \in [0, 1]$ *such that* $f(c) = 0$.

Proof. Consider the numbers $f(\frac{d+1}{10})$, where d is a digit. Note that $0 \leq f(1) = f(\frac{9+1}{10})$. Let c_1 be the smallest digit such that $0 \leq f(\frac{c_1+1}{10})$. Then

$$f(\frac{c_1}{10}) < 0 \leq f(\frac{c_1 + 1}{10}).$$

Let c_2 be the smallest digit such that $0 \leq f(\frac{c_1}{10} + \frac{c_2+1}{100})$. Then

$$f(\frac{c_1}{10} + \frac{c_2}{100}) < 0 \leq f(\frac{c_1}{10} + \frac{c_2 + 1}{100}).$$

We continue in this fashion, choosing for each positive integer k the least digit c_k such that

$$f(0.c_1c_2 \cdots c_k) < 0 \leq f(0.c_1c_2 \cdots c_k + 10^{-k}).$$

Let $c := 0.c_1c_2 \cdots$, the full, unending decimal, *i.e.*, the real number that we approach as we continue the process of choosing the c_i indefinitely. Every open interval containing c contains a point at which the value of f is strictly less than 0 and a point at which the value of f is greater than or equal to 0. By Fact 2, therefore, $f(c) = 0$. /////

A.2. The Maximum Value Theorem

In proving the Maximum Value Theorem, we shall use the following notion. Suppose A and B are sets of real numbers. We shall say that A *dominates* B if:

$$\text{for every } b \in B \text{ there is } a \in A \text{ such that } a \geq b.$$

Note the following:

i) The relation of dominance is reflexive and transitive. In addition, if A dominates every set in some family of sets, then it dominates the union of the family. (These assertions are immediate from the definition.)

ii) If A does *not* dominate B, then B dominates A. (If A does *not* dominate B then there is some element—call it b_0—in B that is not less than or equal to *any* element of A. Then $b_0 > a$ for all $a \in A$, and this means that B dominates A.)

iii) Given any finite family of sets of real numbers, there is one of them that dominates the union of all of them. (This follows from *i*) and *ii*).)

Theorem. *Suppose $f : [0,1] \to \mathbf{R}$ is continuous. Then there is $c \in [0,1]$ such that for all $x \in [0,1]$, $f(c) \geq f(x)$.*

Proof. From the facts immediately preceding the statement of the theorem, we see that if $A = A_1 \cup \cdots \cup A_k$ are any subsets of $[0,1]$, then for some $i \in \{1, \ldots, k\}$, $f(A_i)$ dominates $f(A)$. Thus there exists a sequence of digits $c_1, c_2, c_3, \ldots$ such that $f\big([\frac{c_1}{10}, \frac{c_1+1}{10}]\big)$ dominates $f\big([0,1]\big)$, $f\big([\frac{c_1}{10} + \frac{c_2}{100}, \frac{c_1}{10} + \frac{c_2+1}{100}]\big)$ dominates $f\big([\frac{c_1}{10}, \frac{c_1+1}{10}]\big)$, $\ldots$*etc.* Letting J_k denote the interval $[.c_1c_2\cdots c_k, .c_1c_2\cdots c_k + 10^{-k}]$, we have in general that

$$\text{for each } k,\ f(J_k) \text{ dominates } f\big([0,1]\big).$$

Let $c := .c_1c_2\cdots$. If $I := (a,b)$ is any open interval containing c then $f(I \cap [0,1])$ dominates $f\big([0,1]\big)$, since $J_k \subseteq (a,b)$ whenever k is large enough that $10^{-k} < \min\{b-c, c-a\}$. Now, I claim that $f(c) \geq f(x)$ for all $x \in [0,1]$. If not, there is $x_0 \in [0,1]$ such that $f(x_0) > f(c)$. Then, then by Fact 2, there is an open interval I about c such that $f(x) < f(x_0)$ for all $x \in I \cap [0,1]$. But this is impossible since as we have seen $f(I \cap [0,1])$ dominates $f\big([0,1]\big)$. /////

A.3. The Integration Theorem

In dealing with integrals, we use the following notation. Let f be a bounded function on $J = [a, b]$. Let $\mathcal{P} = \{ a = x_0 < x_1 < \cdots < x_n = b \}$ be a partition of J. For each interval $[x_{i-1}, x_i]$, let λ_i be any number such that $\lambda_i \leq f(x)$ for all $x \in [x_{i-1}, x_i]$. Similarly, let μ_i be any number such that $\mu_i \geq f(x)$ for all $x \in [x_{i-1}, x_i]$. The *lower sum* determined by this data is:

$$L(f, J, \mathcal{P}, \lambda) := \sum_{i=1}^{n} \lambda_i (x_i - x_{i-1}),$$

and the upper sum determined by this data is:

$$U(f, J, \mathcal{P}, \mu) := \sum_{i=1}^{n} \mu_i (x_i - x_{i-1}).$$

Lemma. *Suppose that ϵ is some positive number such that*

$$\textit{for all data } \mathcal{P}, \lambda \textit{ and } \mu \textit{ on } [a, b],\ U(f, [a, b], \mathcal{P}, \mu) - L(f, [a, b], \mathcal{P}, \lambda) > \epsilon.$$

Then there are $w_1, w_2 \in [a, b]$ such that $|f(w_1) - f(w_2)| > \frac{1}{2} \frac{\epsilon}{b-a}$.

Proof. We prove the contrapositive. Suppose that $|f(w_1) - f(w_2)| \leq \frac{1}{2} \frac{\epsilon}{b-a}$ for all $w_1, w_2 \in [a, b]$. Then $f(a) - \frac{1}{2} \frac{\epsilon}{b-a} \leq f(x) \leq f(a) + \frac{1}{2} \frac{\epsilon}{b-a}$ for all $x \in [a, b]$. Letting $\mathcal{P} := \{ a = x_0 < x_1 = b \}$, $\lambda := f(a) - \frac{1}{2} \frac{\epsilon}{b-a}$ and $\mu := f(a) + \frac{1}{2} \frac{\epsilon}{b-a}$, we get

$$U(f, [a, b], \mathcal{P}, \mu) - L(f, [a, b], \mathcal{P}, \lambda) = (f(a)(b - a) + \frac{\epsilon}{2}) - (f(a)(b - a) - \frac{\epsilon}{2}) = \epsilon.$$

/////

Theorem. *Suppose that f is a continuous on $[0, 1]$. For any $\epsilon > 0$, there are data $\mathcal{P}$, λ and μ on $[0, 1]$ such that*

$$U(f, [0, 1], \mathcal{P}, \mu) - L(f, J, \mathcal{P}, \lambda) \leq \epsilon.$$

Proof. We will prove the contrapositive. Assume we have a bounded function f for which the conclusion fails. We shall show that there is a point at which f is not continuous. By assumption, there is a number $\epsilon > 0$ such that

$$\text{for all data } \mathcal{P}, \lambda \text{ and } \mu \text{ on } [0, 1],\ U(f, [0, 1], \mathcal{P}, \mu) - L(f, [0, 1], \mathcal{P}, \lambda) > \epsilon. \tag{1}$$

Now for at least one of the subintervals $I = [0, \frac{1}{10}], [\frac{1}{10}, \frac{2}{10}], \cdots, [\frac{9}{10}, 1]$ we must have that for all data $\mathcal{P}$, λ and μ on I, $U(f, I, \mathcal{P}, \mu) - L(f, I, \mathcal{P}, \lambda) > \frac{\epsilon}{10}$, since otherwise we could construct data $\mathcal{P}$, λ and μ on $[0, 1]$ violating (1). Let $J_1 := [\frac{c_1}{10}, \frac{c_1+1}{10}]$ be one such interval. Now repeat the argument inside J_1 to find an interval $J_2 := [.c_1c_2, .c_1c_2 + 10^{-2}]$ such that for all data $\mathcal{P}$, λ and μ on J_2, $U(f, J_2, \mathcal{P}, \mu) - L(f, J_2, \mathcal{P}, \lambda) > \frac{\epsilon}{100}$. Continuing in this fashion, we get a sequence of digits $c_1, c_2, \cdots$ and a corresponding sequence of intervals J_k such that

$$\text{for all data } \mathcal{P},\ \lambda \text{ and } \mu \text{ on } J_k,\ U(f, J_k, \mathcal{P}, \mu) - L(f, J_k, \mathcal{P}, \lambda) > \epsilon/10^k.$$

Let $c := c_1 c_2 \cdots$. Every open interval about c contains one of the intervals J_k. By the lemma, each interval J_k contains numbers w_1, w_2 such that $|f(w_1) - f(w_2)| > \frac{\epsilon}{2}$. Thus, there is no interval (a, b) about c such that $|f(x) - f(c)| < \frac{\epsilon}{4}$ for all $x \in (a, b) \cap [0, 1]$, so f is not continuous at c. /////

A.4. Other Theorems

The Least Upper Bound Principle and the fact that a continuous function on a closed bounded interval is uniformly continuous can be proved using the same pattern of reasoning employed above.

Theorem. *Let X be any subset of $[0, 1]$. Then the set*

$$U := \{\, y \in [0, 1] \mid \forall x \in X\,,\ x \le y \,\}$$

has a smallest element, c.

Proof. If X is empty, let $c = 0$. If X is not empty, then neither is U. Pick digits c_i so that

$$.c_1 c_2 \cdots c_k + 10^{-k}$$

is the smallest multiple of 10^{-k} in U. Let $c := .c_1 c_2 \cdots$. Now, $c \in U$, since if not then there is an integer k such that some $x \in X$, exceeds c by *more than* 10^{-k}, and then this x also properly exceeds $.c_1 c_2 \cdots c_k + 10^{-k}$. We must also show that no number properly smaller than c is in U. But if $c' < c$, then c' is less than some truncation of c—$.c_1 c_2 \cdots c_k$, say—and this is *not* in U, so neither is c'. /////

We will not write out a full proof of the theorem on uniform continuity, but only say enough to show how a proof along the lines of those we've already presented can go. Suppose f is a real-valued function on an interval I, and $\epsilon > 0$. Consider the following statement about f, I and ϵ:

$$\forall \delta > 0\ \exists x_1, x_2 \in I\ ,\ |x_2 - x_1| < \delta \text{ and } |f(x_2) - f(x_1)| > \epsilon. \qquad N(f, I, \epsilon)$$

If I is a union of finitely many intervals, $I = [x_0, x_1] \cup [x_1, x_2] \cup \cdots \cup [x_{n-1}, x_n]$ and f is continuous at each point $x_1, \ldots, x_{n-1}$, then

$$N(f, I, \epsilon) \text{ implies that for some } i,\ N(f, [x_{i-1}, x_i], \epsilon).$$

Using the pattern of reasoning exhibited above, this can be used to show: if $N(f, [0, 1], \epsilon)$ for some $\epsilon > 0$, then there is $c \in [0, 1]$ where f is not continuous.