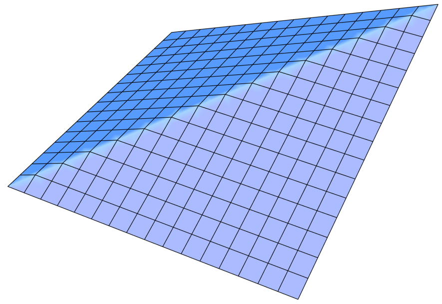

MATH 147: Calculus I (Advanced Level)

Prof. Kathryn Hare • Fall 2012 • University of Waterloo
MWF 8:30–9:20AM in AL 105

Transcribed, LaTeXed, and edited by Rob Zimmerman

Note: These are personal lecture notes, not official course materials. They may contain errors (which by default you should attribute to me, Rob Zimmerman, rather than the course instructor). The permanent home of these — and all of my other lecture notes — is https://rob-zimmerman.github.io/coursenotes.

Contents

1 Properties of Real Numbers **3**
Mathematical Induction 3
Irrational Numbers and Absolute Value 6
Least Upper Bounds and Suprema 8

2 Sequences and Their Limits **14**
Sequences and Convergence 15

Limit Laws . . . 21
Subsequences . . . 25
Cauchy Sequences . . . 27

3 Limits, Continuity, and Inverse Functions **30**
Limits at a Point . . . 31
One-Sided Limits and Limit Laws . . . 35
Continuity . . . 40
The Intermediate and Extreme Value Theorems . . . 47
Inverse Functions . . . 52

4 Elementary Functions and Differentiation **55**
Elementary Functions and the Derivative . . . 56
Differentiation Rules . . . 64
Critical Points and Derivative Applications . . . 71

5 The Mean Value Theorem and Its Consequences **74**
Rolle's Theorem and the Mean Value Theorem . . . 74
Consequences of the Mean Value Theorem . . . 79
Concavity and Derivative Tests . . . 84

6 L'Hôpital's Rule **88**
Motivation and the Cauchy Mean Value Theorem . . . 88
Statement and Proof of L'Hôpital's rule . . . 89
Applications and Variants . . . 92

7 Taylor Polynomials and Taylor's Theorem **96**
Taylor Polynomials . . . 97
Taylor's Theorem and Higher Derivative Tests . . . 100

8 Newton's Method **105**
The Algorithm . . . 105
Error Bounds and Failure Modes . . . 109
When Newton's Method Can Fail . . . 110

9 Cardinality **114**
Finiteness and Countability . . . 114
Uncountability and the Irrational Numbers . . . 117
Unions of Countable Sets . . . 118

Appendix: Lecture and Tutorial Index **122**

Lecture 1 — Monday, September 10, 2012

1. Properties of Real Numbers

Mathematical Induction

Our first goal is to understand the real numbers $\mathbb{R}$: what they are and how we work with them. The familiar number systems

$$\begin{aligned} \mathbb{N} &= \{1, 2, 3, \dots\}, \\ \mathbb{Z} &= \{\dots, -2, -1, 0, 1, 2, \dots\}, \quad \text{and} \\ \mathbb{Q} &= \left\{ \frac{p}{q} \,\middle|\, p, q \in \mathbb{Z},\ q \neq 0 \right\}, \end{aligned}$$

are the **natural numbers**, the **integers**, and the **rational numbers** (respectively).

Proposition 1.1 (Well-ordering principle) The set $\mathbb{N}$ is **well-ordered**; that is, every nonempty subset of $\mathbb{N}$ has a least element.

We take the well-ordering principle as a foundational property of $\mathbb{N}$. It is equivalent to the principle of mathematical induction proved below, so trying to justify it by "checking $1, 2, 3, \dots$ until we reach an element" would only hide the same principle inside the argument.

$\mathbb{Q}$ is a field, while $\mathbb{N}$ and $\mathbb{Z}$ are not. Clearly the set $\mathbb{Z}$ is not well-ordered. Nor is $\mathbb{Q}^+$ well-ordered; indeed, if $H = \{x \in \mathbb{Q} \mid x > 0\}$ and $y \in H$, then $\frac{y}{2} \in H$ and $\frac{y}{2} < y$.

Theorem 1.2 Let $P(n)$ be a statement about $n \in \mathbb{N}$. Suppose that $P(1)$ is true, and suppose that whenever $P(k)$ is true for some $k \in \mathbb{N}$, $P(k+1)$ is also true. Then $P(n)$ is true for all $n \in \mathbb{N}$.

Proof. We prove the result by contradiction. Suppose the assumptions hold, but the conclusion fails. Then there exists $n \in \mathbb{N}$ such that $P(n)$ is false. Let $A = \{n \in \mathbb{N} \mid P(n) \text{ is false}\}$. By assumption, $A \neq \emptyset$. By the well-ordering principle, A has a least member, say $n_0 \in \mathbb{N}$. We have $n_0 \neq 1$ because $P(1)$ is true. Hence $n_0 \geq 2$, so $n_0 - 1 \in \mathbb{N}$. Since $n_0 - 1 \notin A$, $P(n_0 - 1)$ is true.

By the induction hypothesis, $P(n_0)$ is true, a contradiction. $\square$

Lecture 2 — Wednesday, September 12, 2012

Example 2.1 Let $P(n)$ be the statement that

$$r + r^2 + \cdots + r^n = \frac{r - r^{n+1}}{1 - r},$$

where $r \neq 1$. We will prove that $P(n)$ holds for all $n \in \mathbb{N}$.

Proof. The base case $P(1)$ is true, since

$$\frac{r - r^2}{1 - r} = r.$$

Now assume that $P(k)$ is true for some $k \in \mathbb{N}$. Then

$$r + r^2 + \cdots + r^{k+1} = \frac{r - r^{k+1}}{1 - r} + r^{k+1} = \frac{r - r^{k+1} + r^{k+1}(1 - r)}{1 - r} = \frac{r - r^{k+2}}{1 - r}.$$

This is exactly $P(k+1)$. Hence, by mathematical induction, $P(n)$ is true for all $n \in \mathbb{N}$. $\square$

It is essential to check the base case. For example, if $P(n)$ is the statement $n \geq 100$, then $P(k)$ implies $P(k+1)$, but $P(n)$ is certainly not true for all $n \in \mathbb{N}$.

Example 2.2 Prove that $2^n > n^2$ for $n \geq 5$.

Proof. The base case is $P(5)$, and indeed

$$2^5 = 32 > 25 = 5^2.$$

For the induction step, assume that $P(k)$ is true and aim to show that $P(k+1)$ is true. Then

$$2^{k+1} = 2 \cdot 2^k > 2k^2.$$

For $k \geq 5$,

$$2k^2 - (k+1)^2 = k^2 - 2k - 1 > 0,$$

so $2^{k+1} > (k+1)^2$. Therefore $P(k+1)$ holds, and induction proves the claim for every $n \geq 5$. $\square$

The same contradiction argument works when the base case starts at some $n_0 \in \mathbb{N}$. Assume the result is false, and let

$$A = \{n \geq n_0 \mid P(n) \text{ is false}\}.$$

Then $A \neq \emptyset$. By the well-ordering principle, A has a least member, say N. Since $P(n_0)$ is true, we have $N \neq n_0$. Because $N \in A$, we have $N \geq n_0 + 1$, so $N - 1 \geq n_0$ and therefore $P(N-1)$ is true. But then $P((N-1)+1) = P(N)$ is true, which contradicts $N \in A$.

We restate the principle of mathematical induction in the following form:

Theorem 2.3 (Principle of strong induction) Let $P(n)$ be a statement about $n \in \mathbb{N}$. Suppose that $P(1)$ is true. Suppose also that, whenever $P(j)$ is true for all $j \in \mathbb{N}$ with $j \leq k$, it follows that $P(k+1)$ is true. Then $P(n)$ is true for all $n \in \mathbb{N}$.

Proof. Assume the result is false. Let $A = \{n \in \mathbb{N} \mid P(n) \text{ is false}\}$. Then $A \neq \emptyset$. By the well-ordering principle, A has a least member; call it N. By assumption, $N \neq 1$, so $N \geq 2$. Therefore $P(1), P(2), \ldots, P(N-1)$ are true. By the second assumption, this implies that $P(N)$ is true. But $N \in A$, so $P(N)$ is false. This is a contradiction. $\square$

Lecture 3 — Friday, September 14, 2012

Irrational Numbers and Absolute Value

Example 3.1 Suppose f is defined on $\mathbb{N}$ by

$$f(1) = 1, \qquad f(2) = 2, \qquad \text{and} \qquad f(n+2) = \frac{1}{2}(f(n) + f(n+1)) \quad \text{for } n \geq 1.$$

For instance,

$$f(3) = \frac{1}{2}(f(1) + f(2)) = \frac{3}{2}.$$

We will show that the range of f is contained in $\mathbb{Q}$ and that $1 \leq f(n) \leq 2$ for all $n \in \mathbb{N}$.

Proof. We prove the two claims separately.

First, we show that $f(n) \in \mathbb{Q}$ for all $n \in \mathbb{N}$. The base cases are immediate, since $f(1) = 1$ and $f(2) = 2$ are rational. Now assume that $f(j) \in \mathbb{Q}$ for every integer j with $1 \leq j \leq k+1$. Then

$$f(k+2) = \frac{1}{2}\big(f(k) + f(k+1)\big).$$

By the induction hypothesis, both $f(k)$ and $f(k+1)$ are rational, so their average is rational as well. Hence $f(k+2) \in \mathbb{Q}$. By strong induction, every term of the sequence is rational.

Next, we show that $1 \leq f(n) \leq 2$ for all $n \in \mathbb{N}$. Again, the base cases are immediate: $f(1) = 1$ and $f(2) = 2$. Assume that $1 \leq f(j) \leq 2$ for every integer j with $1 \leq j \leq k+1$. Then

$$1 \leq f(k), f(k+1) \leq 2.$$

Taking averages preserves these inequalities, so

$$1 \leq \frac{1}{2}\big(f(k) + f(k+1)\big) = f(k+2) \leq 2.$$

Thus $1 \leq f(k+2) \leq 2$. By strong induction, the inequality holds for every $n \in \mathbb{N}$. $\square$

Definition 3.2 A real number is called **irrational** if it does not belong to $\mathbb{Q}$.

For example, $\sqrt{2}$ and π are irrational.

Proposition 3.3 $\sqrt{2}$ is irrational.

Proof. Suppose, for a contradiction, that $\sqrt{2} = \frac{p}{q}$ for some coprime integers p and q with $q \neq 0$. Squaring both sides gives

$$2q^2 = p^2.$$

It follows that p^2 is even, and therefore p is even. Write $p = 2m$ for some $m \in \mathbb{Z}$. Substituting this into the equation above gives

$$2q^2 = (2m)^2 = 4m^2,$$

so

$$q^2 = 2m^2.$$

Hence q^2 is even, and therefore q is even. Thus both p and q are even, which contradicts the assumption that they are coprime. Therefore $\sqrt{2}$ is irrational. $\square$

Definition 3.4 The **absolute value** of a real number a is defined by

$$|a| = \begin{cases} a & \text{if } a \geq 0, \\ -a & \text{if } a < 0. \end{cases}$$

For example, $|-3| = 3$. We also have $|a|^2 = a^2$ and $\sqrt{a^2} = |a|$. If $r \geq 0$, then

$$|x| \leq r \iff -r \leq x \leq r,$$

so the inequality $|x| \leq r$ describes the interval $[-r, r]$. More generally, if $r \geq 0$, then

$$|a - b| \leq r \iff -r \leq a - b \leq r \iff b - r \leq a \leq b + r.$$

Equivalently,

$$|a - b| \leq r \iff a - r \leq b \leq a + r.$$

In particular, the distance from a to b is the same as the distance from b to a.

Proposition 3.5 (Triangle inequality) For all $a, b \in \mathbb{R}$,

$$|a + b| \leq |a| + |b|.$$

Proof. We have $-|a| \leq a \leq |a|$ and $-|b| \leq b \leq |b|$. Adding these gives

$$-(|a| + |b|) \leq a + b \leq |a| + |b|.$$

The conclusion follows. $\square$

The triangle inequality also yields the **reverse triangle inequality**. Indeed,

$$|a| = |(a - b) + b| \leq |a - b| + |b|$$

and

$$|b| = |(b - a) + a| \leq |a - b| + |a|.$$

Rearranging these inequalities gives

$$|a - b| \geq ||a| - |b||.$$

Least Upper Bounds and Suprema

The real numbers $\mathbb{R}$ are characterized by four important properties:

1. $\mathbb{R}$ is a field.
2. $\mathbb{R}$ is an ordered field: its order is total and is compatible with addition and multiplication.
3. $\mathbb{R}$ contains $\mathbb{N}$ (equivalently, $\mathbb{R}$ contains $\mathbb{Q}$).
4.

It is our goal to determine the missing property.

Definition 3.6 Let A be a nonempty subset of an ordered set S. We say that A is **bounded above** if there exists $x \in S$ such that

$$a \leq x \quad \text{for all } a \in A.$$

Similarly, A is **bounded below** if there exists $x \in S$ such that

$$x \leq a \quad \text{for all } a \in A.$$

Any such number x is called an **upper bound** or **lower bound**, respectively, for A.

Not every set is bounded above; for example, $\mathbb{Z}$ is not. On the other hand, the set $\mathbb{Q}^+ = \{x > 0 \mid x \in \mathbb{Q}\}$ is bounded below, but it has no least element.

Definition 3.7 A number x is called a **least upper bound** of A, or the **supremum** of A, if x is an upper bound for A and $x \leq y$ for every upper bound y of A. Similarly, we define the **greatest lower bound**, or **infimum**, of A.

If the least upper bound of A exists, then it is unique. Indeed, if both x and y were least upper bounds and $x \neq y$, then either $x < y$ or $y < x$, contradicting the minimality of one of them.

Definition 3.8 A set A is **bounded** if it is both bounded above and below.

Example 3.9 The set

$$A = \left\{ \frac{n}{n+1} \,\middle|\, n = 1, 2, 3, \ldots \right\} = \left\{ \frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \ldots \right\}$$

is bounded. For instance, 0 is a lower bound and 2 is an upper bound. Its greatest lower bound is $\frac{1}{2}$, which belongs to A. Its least upper bound is 1, and $1 \notin A$.

Lecture 4 — Monday, September 17, 2012

Recall that $\mathbb{R}$ is an ordered field containing $\mathbb{Q}$. The additional property that distinguishes $\mathbb{R}$ from $\mathbb{Q}$ is the completeness property, which we will state below. Before doing so, it is useful to record a practical characterization of least upper bounds.

Example 4.1 If

$$A = (1, 2] = \{x \in \mathbb{R} \mid 1 < x \leq 2\},$$

then the least upper bound of A is 2, while the greatest lower bound is 1, which does not belong to A.

Theorem 4.2 (Characterization of the least upper bound) Let A be a nonempty subset of an ordered set, and let x be an element of that ordered set. Then x is the least upper bound of A if and only if x is an upper bound for A and, for every $z < x$, there exists $a \in A$ such that

$$z < a \leq x.$$

The ordering used in this characterization is pictured in Figure 1.

FIGURE 1

Proof. First suppose that x is the least upper bound of A. Then x is certainly an upper bound for A. Let $z < x$. If there were no element $a \in A$ with $z < a$, then every element of A would satisfy $a \leq z$, so z would also be an upper bound for A. This would contradict the fact that x is the **least** upper bound. Therefore there exists $a \in A$ with $z < a$, and since x is an upper bound we automatically have $a \leq x$.

Conversely, suppose that x is an upper bound for A and that for every $z < x$ there exists $a \in A$ with $z < a \leq x$. To show that x is the least upper bound, let y be any upper bound for A. If $y < x$, then applying the hypothesis with $z = y$ produces an element $a \in A$ such that $y < a$, contradicting the fact that y is an upper bound. Hence every upper bound y satisfies $y \geq x$, and therefore x is the least upper bound of A. $\square$

The characterization above is often restated in ε-language: an upper bound x is the least upper bound of A if and only if, for every $\varepsilon > 0$, there exists $a \in A$ such that

$$x - \varepsilon < a \leq x.$$

Indeed, every number $z < x$ can be written as $x - \varepsilon$ for some $\varepsilon > 0$, and conversely $x - \varepsilon < x$ whenever $\varepsilon > 0$.

Definition 4.3 $\mathbb{R}$ has the **completeness property**; that is, every nonempty subset of $\mathbb{R}$ that is bounded above has a least upper bound in $\mathbb{R}$.

Proposition 4.4 Every nonempty subset of $\mathbb{R}$ that is bounded below has a greatest lower bound in $\mathbb{R}$.

Proof. Let $A \subseteq \mathbb{R}$ be nonempty and bounded below, and define

$$B = \{-a \mid a \in A\}.$$

Then B is nonempty and bounded above. By the completeness property, B has a least upper bound; call it s. We claim that $-s$ is the greatest lower bound of A. If $a \in A$, then $-a \in B$, so $-a \leq s$. Multiplying by -1 gives $-s \leq a$, so $-s$ is a lower bound for A. Now let ℓ be any lower bound for A. Then $\ell \leq a$ for every $a \in A$, so $-a \leq -\ell$ for every $a \in A$. Thus $-\ell$ is an upper bound for B. Since s is the least upper bound of B, we have $s \leq -\ell$, and therefore $\ell \leq -s$. Hence $-s$ is the greatest lower bound of A. □

We can now complete the definition of $\mathbb{R}$:

Definition 4.5 The set of **real numbers** $\mathbb{R}$ is a set with the following properties:

1. $\mathbb{R}$ is a field.
2. $\mathbb{R}$ is an ordered field: its order is total and is compatible with addition and multiplication.
3. $\mathbb{R}$ contains $\mathbb{N}$ (equivalently, $\mathbb{R}$ contains $\mathbb{Q}$).
4. $\mathbb{R}$ has the completeness property.

It is the completeness of $\mathbb{R}$ that distinguishes it from $\mathbb{Q}$. In more advanced mathematics courses, we show that $\mathbb{R}$ is well-defined: up to isomorphism, it is the unique complete ordered field.

Many important results follow from completeness. One of the first is the Archimedean property.

Proposition 4.6 (Archimedean property) Given any $x \in \mathbb{R}$, there exists $n \in \mathbb{N}$ with $n > x$.

Proof. Suppose, for a contradiction, that there exists $x \in \mathbb{R}$ such that $n \leq x$ for every $n \in \mathbb{N}$. Then $\mathbb{N}$ is nonempty and bounded above, so by the completeness property it has a least upper bound, say $r = \sup \mathbb{N}$. Since $r - 1 < r$, the characterization of the least upper bound shows that

there exists $m \in \mathbb{N}$ such that
$$r - 1 < m \leq r.$$
It follows that $m+1 > r$. But $m+1 \in \mathbb{N}$, which contradicts the fact that r is an upper bound for $\mathbb{N}$. Therefore, for every real number x, there exists $n \in \mathbb{N}$ with $n > x$. □

Corollary 4.7 The set $\{\frac{1}{n} \mid n \in \mathbb{N}\}$ has greatest lower bound 0. This means that given any $\varepsilon > 0$, there exists $n \in \mathbb{N}$ such that $\frac{1}{n} < \varepsilon$.

Proof. Let
$$A = \left\{ \frac{1}{n} \,\middle|\, n \in \mathbb{N} \right\}.$$
This set is nonempty and bounded below, so it has a greatest lower bound. Clearly 0 is a lower bound for A. Now let $z > 0$. By the Archimedean property, there exists $n \in \mathbb{N}$ such that $n > \frac{1}{z}$. Taking reciprocals gives
$$\frac{1}{n} < z.$$
Since $\frac{1}{n} \in A$, the number z cannot be a lower bound for A. Thus no positive number is a lower bound for A, and it follows that $\inf A = 0$. □

The completeness property is sometimes informally described as the "no holes" property of the real numbers.

Lecture 5 — Wednesday, September 19, 2012

Recall that $\mathbb{R}$ is an ordered field containing $\mathbb{Q}$ and satisfying the completeness property.

The Archimedean property states that for any $r \in \mathbb{R}$, there exists $n \in \mathbb{N}$ with $r < n$. A useful corollary is that for any $\varepsilon > 0$, there exists $n \in \mathbb{N}$ such that $1/n < \varepsilon$.

Theorem 5.1 ($\mathbb{Q}$ is dense in $\mathbb{R}$) If $x < y$ in $\mathbb{R}$, then there exists $q \in \mathbb{Q}$ such that $x < q < y$.

Proof. *Case 1:* $x \geq 0$. Let $\varepsilon = y - x > 0$. By Corollary 4.7, there exists $n \in \mathbb{N}$ such that

$$\frac{1}{n} < \varepsilon = y - x.$$

Multiplying by n gives

$$1 + nx < ny.$$

Now consider the set

$$S = \{m \in \mathbb{N} \mid m > nx\}.$$

This set is nonempty by the Archimedean property, so by the well-ordering principle it has a least element; call it M. Since M is the least natural number greater than nx, we have

$$M - 1 \leq nx < M.$$

Therefore

$$M \leq nx + 1 < ny.$$

Dividing by n yields

$$x < \frac{M}{n} < y.$$

Because $\frac{M}{n} \in \mathbb{Q}$, this proves the claim in the case $x \geq 0$.

Case 2: $x < 0$. By the Archimedean property, we may choose $N \in \mathbb{N}$ with $N > -x$. Then $x + N > 0$, and of course $x + N < y + N$. By Case 1, there exists $q \in \mathbb{Q}$ such that

$$x + N < q < y + N.$$

Since $N \in \mathbb{Z} \subseteq \mathbb{Q}$, the number $q - N$ is rational. Subtracting N from the displayed inequality gives

$$x < q - N < y.$$

Thus there is always a rational number strictly between x and y. $\square$

Proposition 5.2 (Existence of $\sqrt{2}$) $\sqrt{2} \in \mathbb{R}$, that is, there exists $r > 0$ with $r^2 = 2$.

Proof. Let

$$S = \{x \in \mathbb{R} \mid x^2 < 2\}.$$

This set is nonempty, since $1 \in S$, and it is bounded above, since every $x \geq 2$ satisfies $x^2 \geq 4 > 2$.

By the completeness property, S has a least upper bound; write

$$w = \sup S.$$

Since $1 \in S$, we have $w \geq 1$, so in particular $w > 0$. We now show that $w^2 = 2$. First suppose that $w^2 < 2$. Set

$$\delta = 2 - w^2 > 0.$$

By the Archimedean property, there exists $N \in \mathbb{N}$ such that

$$\frac{1}{N} < \min\left\{1, \frac{\delta}{2w+1}\right\}.$$

Then

$$\left(w + \frac{1}{N}\right)^2 = w^2 + \frac{2w}{N} + \frac{1}{N^2} \leq w^2 + \frac{2w+1}{N} < w^2 + \delta = 2.$$

Thus $w + \frac{1}{N} \in S$, which contradicts the fact that w is an upper bound for S. Next suppose that $w^2 > 2$. Set

$$\delta = w^2 - 2 > 0.$$

Again by the Archimedean property, there exists $N \in \mathbb{N}$ such that

$$\frac{1}{N} < \min\left\{w, \frac{\delta}{2w}\right\}.$$

Then $w - \frac{1}{N} > 0$, and

$$\left(w - \frac{1}{N}\right)^2 = w^2 - \frac{2w}{N} + \frac{1}{N^2} > w^2 - \frac{2w}{N} > w^2 - \delta = 2.$$

We claim that $w - \frac{1}{N}$ is an upper bound for S. Indeed, if $x \in S$ and $x > w - \frac{1}{N}$, then $x > 0$ and hence

$$x^2 > \left(w - \frac{1}{N}\right)^2 > 2,$$

which contradicts $x \in S$. Therefore $w - \frac{1}{N}$ is an upper bound for S, but it is strictly smaller than w, contradicting the fact that w is the least upper bound of S. Both alternatives lead to contradictions. Hence $w^2 = 2$, so $\sqrt{2}$ exists as a real number. $\square$

2. Sequences and Their Limits

Sequences and Convergence

Definition 5.3 A **sequence** in $\mathbb{R}$ is an ordered list of real numbers

$$x_1, x_2, x_3, \ldots$$

We often write a sequence as $(x_n)_{n=1}^{\infty}$, or simply as (x_n) when the indexing is clear from context.

Example 5.4 Common examples include the constant sequence $x_n = 1$ for all n, the reciprocal sequence $x_n = \frac{1}{n}$, the alternating sequence $x_n = (-1)^n$, and recursively defined sequences such as

$$x_1 = 1, \qquad x_2 = 2, \qquad \text{and} \qquad x_{n+2} = \frac{x_{n+1} + x_n}{2}.$$

Definition 5.5 We say that the sequence $(x_n)_{n=1}^{\infty}$ **converges** to a real number L if, for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$|x_n - L| < \varepsilon \quad \text{for all } n \geq N.$$

In that case, L is called the **limit** of the sequence. We write

$$\lim_{n \to \infty} x_n = L \qquad \text{or} \qquad x_n \to L \text{ as } n \to \infty.$$

If a sequence does not converge, then it **diverges**.

Observe that the inequality

$$|x_n - L| < \varepsilon$$

is equivalent to

$$L - \varepsilon < x_n < L + \varepsilon.$$

Thus convergence means that, once n is large enough, every term x_n lies inside the interval $(L - \varepsilon, L + \varepsilon)$. See Figure 2.

Example 5.6 Consider the sequence $x_n = \frac{1}{n}$. We claim that $x_n \to 0$ as $n \to \infty$.

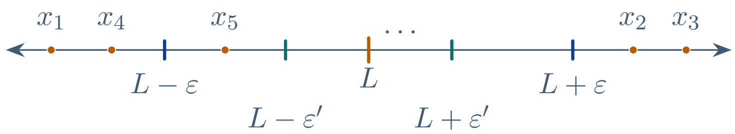

FIGURE 2

Proof. Let $\varepsilon > 0$. By the Archimedean property, there exists $N \in \mathbb{N}$ with $N > \frac{1}{\varepsilon}$. If $n \geq N$, then

$$\left|\frac{1}{n} - 0\right| = \frac{1}{n} \leq \frac{1}{N} < \varepsilon.$$

Therefore $x_n \to 0$ as $n \to \infty$. □

Lecture 6 — Friday, September 21, 2012

We continue by applying the definition of convergence to several examples.

Example 6.1 Let

$$x_n = \frac{(-1)^n}{n^2+1}.$$

We claim that $x_n \to 0$ as $n \to \infty$.

Proof. Let $\varepsilon > 0$. By the Archimedean property, choose $N \in \mathbb{N}$ with $N > \frac{1}{\varepsilon}$. If $n \geq N$, then

$$\left|\frac{(-1)^n}{n^2+1} - 0\right| = \frac{1}{n^2+1} \leq \frac{1}{n} \leq \frac{1}{N} < \varepsilon.$$

Therefore $x_n \to 0$. □

Example 6.2 The sequence $x_n = (-1)^n$ diverges.

Proof. Fix any real number L, and take $\varepsilon = \frac{1}{2}$. The open interval

$$\left(L - \frac{1}{2}, L + \frac{1}{2}\right)$$

has length 1, so it cannot contain both 1 and -1. However, no matter how large N is, there is an even integer $n \geq N$ with $x_n = 1$ and there is an odd integer $m \geq N$ with $x_m = -1$. Thus it is impossible for all sufficiently large terms of the sequence to lie inside the interval above. Since the definition of convergence fails for this choice of ε and for every L, the sequence diverges. □

Example 6.3 The sequence $x_n = n$ diverges.

Proof. Fix any real number L. By the Archimedean property, choose $N \in \mathbb{N}$ with $N > L+1$. If $n \geq N$, then

$$|x_n - L| = |n - L| = n - L \geq N - L > 1.$$

Therefore the defining inequality $|x_n - L| < 1$ can never hold for all sufficiently large n, so the sequence does not converge. □

Changing finitely many terms of a sequence does not affect its limit. For example, the sequence

$$1, 2, 3, 4, \frac{1}{5}, \frac{1}{6}, \frac{1}{7}, \dots$$

still converges to 0, because it agrees with $\frac{1}{n}$ for every $n \geq 5$.

Example 6.4 Suppose $|r| < 1$. Then the sequence $x_n = r^n$ converges to 0.

Proof. If $r = 0$, then $r^n = 0$ for all n, so the conclusion is immediate. Assume now that $0 < |r| < 1$. Then $\frac{1}{|r|} > 1$, so we may write

$$\frac{1}{|r|} = 1 + \delta$$

for some $\delta > 0$. By the binomial theorem,

$$\left(\frac{1}{|r|}\right)^n = (1+\delta)^n = 1 + n\delta + \binom{n}{2}\delta^2 + \cdots + \delta^n \geq n\delta.$$

Taking reciprocals gives

$$|r|^n \leq \frac{1}{n\delta}.$$

Now let $\varepsilon > 0$. By the Archimedean property, choose $N \in \mathbb{N}$ with $N > \frac{1}{\delta\varepsilon}$. If $n \geq N$, then

$$|r^n - 0| = |r|^n \leq \frac{1}{n\delta} \leq \frac{1}{N\delta} < \varepsilon.$$

Hence $r^n \to 0$ as $n \to \infty$. □

Proposition 6.5 (Uniqueness of limits) A sequence cannot converge to two different real numbers.

Proof. Suppose, for a contradiction, that $x_n \to L_1$ and $x_n \to L_2$ with $L_1 \neq L_2$. Without loss of generality, we may assume $L_1 < L_2$. Set

$$\varepsilon = \frac{L_2 - L_1}{3} > 0.$$

Since $x_n \to L_1$, there exists $N_1 \in \mathbb{N}$ such that

$$|x_n - L_1| < \varepsilon \quad \text{for all } n \geq N_1.$$

Thus

$$L_1 - \varepsilon < x_n < L_1 + \varepsilon \quad \text{for all } n \geq N_1.$$

Similarly, since $x_n \to L_2$, there exists $N_2 \in \mathbb{N}$ such that

$$L_2 - \varepsilon < x_n < L_2 + \varepsilon \quad \text{for all } n \geq N_2.$$

Now let $N = \max\{N_1, N_2\}$. For every $n \geq N$, both inequalities hold, so

$$L_2 - \varepsilon < x_n < L_1 + \varepsilon.$$

But our choice of ε gives

$$L_1 + \varepsilon = \frac{2L_1 + L_2}{3} < \frac{L_1 + 2L_2}{3} = L_2 - \varepsilon,$$

which is impossible. This contradiction shows that a sequence cannot have two distinct limits. □

Lecture 7 — Monday, September 24, 2012

Theorem 7.1 (Squeeze theorem) Suppose there exists $n_0 \in \mathbb{N}$ such that

$$x_n \leq y_n \leq z_n \quad \text{for all } n \geq n_0.$$

If $x_n \to L$ and $z_n \to L$, then $y_n \to L$.

Proof. Let $\varepsilon > 0$. Since $x_n \to L$, there exists $N_1 \in \mathbb{N}$ such that

$$|x_n - L| < \varepsilon \quad \text{for all } n \geq N_1.$$

Similarly, since $z_n \to L$, there exists $N_2 \in \mathbb{N}$ such that

$$|z_n - L| < \varepsilon \quad \text{for all } n \geq N_2.$$

Now let

$$N = \max\{n_0, N_1, N_2\}.$$

If $n \geq N$, then

$$L - \varepsilon < x_n \leq y_n \leq z_n < L + \varepsilon.$$

Hence $|y_n - L| < \varepsilon$ for all $n \geq N$, so $y_n \to L$. □

Example 7.2 The squeeze theorem is often the simplest way to evaluate limits. For example, if

$$x_n = \frac{n^2}{5^n},$$

then for $n \geq 4$ we have $n^2 \leq 2^n$, and therefore

$$0 \leq \frac{n^2}{5^n} \leq \left(\frac{2}{5}\right)^n.$$

Since $\left(\frac{2}{5}\right)^n \to 0$, the squeeze theorem gives $\frac{n^2}{5^n} \to 0$.

Example 7.3 Another quick application is

$$x_n = \frac{\cos^2(n)}{n}.$$

Because $0 \le \cos^2(n) \le 1$, we have

$$0 \le \frac{\cos^2(n)}{n} \le \frac{1}{n}.$$

Since $\frac{1}{n} \to 0$, the squeeze theorem shows that $\frac{\cos^2(n)}{n} \to 0$.

Definition 7.4 The sequence (x_n) is called **bounded** if there exists $C \in \mathbb{R}$ such that

$$|x_n| \le C \quad \text{for all } n \in \mathbb{N}.$$

Any such number C is called a **bound** for the sequence.

Example 7.5 The sequence $(\cos(n))$ is bounded. The constant sequence $(1, 1, 1, \dots)$ is bounded and convergent. The alternating sequence $(-1)^n$ is bounded but divergent, while the sequence (n) is unbounded and divergent.

Proposition 7.6 A convergent sequence is bounded.

Proof. Assume that $x_n \to L$. Taking $\varepsilon = 1$, we find $N \in \mathbb{N}$ such that

$$|x_n - L| < 1 \quad \text{for all } n \ge N.$$

If $n \ge N$, then the triangle inequality gives

$$|x_n| = |(x_n - L) + L| \le |x_n - L| + |L| < 1 + |L|.$$

Let C be the maximum of the finitely many numbers

$$|x_1|, |x_2|, \dots, |x_{N-1}|, 1 + |L|.$$

Then $|x_n| \le C$ for every $n \in \mathbb{N}$, so the sequence is bounded. $\square$

The converse is not necessarily true; for example, the bounded sequence in Example 6.2 is divergent.

Limit Laws

Theorem 7.7 (Limit laws) Assume that $x_n \to L$ and $y_n \to K$. Then

$$x_n \pm y_n \to L \pm K$$

and

$$x_n y_n \to LK.$$

If $K \neq 0$, then

$$\frac{x_n}{y_n} \to \frac{L}{K}.$$

Proof. We first prove the product rule.

Let $\varepsilon > 0$. Since (y_n) converges, it is bounded by Proposition 7.6; choose $C > 0$ such that

$$|y_n| \leq C \quad \text{for all } n \in \mathbb{N}.$$

Because $x_n \to L$, there exists $N_1 \in \mathbb{N}$ such that

$$|x_n - L| < \frac{\varepsilon}{2C} \quad \text{for all } n \geq N_1.$$

If $L = 0$, then for every $n \geq N_1$ we have

$$|x_n y_n - LK| = |x_n y_n| \leq C|x_n| = C|x_n - L| < \frac{\varepsilon}{2} < \varepsilon,$$

so the product rule follows immediately. Now assume $L \neq 0$. Since $y_n \to K$, there exists $N_2 \in \mathbb{N}$ such that

$$|y_n - K| < \frac{\varepsilon}{2|L|} \quad \text{for all } n \geq N_2.$$

Let $N = \max\{N_1, N_2\}$. If $n \geq N$, then

$$\begin{aligned} |x_n y_n - LK| &= |(x_n - L)y_n + L(y_n - K)| \\ &\leq |x_n - L||y_n| + |L||y_n - K| \\ &\leq C|x_n - L| + |L||y_n - K| < C \cdot \frac{\varepsilon}{2C} + |L| \cdot \frac{\varepsilon}{2|L|} \end{aligned}$$

$$= \varepsilon.$$

Therefore $x_n y_n \to LK$.

For the sum rule, given $\varepsilon > 0$, choose a common threshold after which

$$|x_n - L| < \frac{\varepsilon}{2} \qquad \text{and} \qquad |y_n - K| < \frac{\varepsilon}{2}.$$

Then $|(x_n + y_n) - (L + K)| < \varepsilon$. The difference rule follows by the same argument.

Finally, suppose $K \neq 0$. Eventually $|y_n| \geq |K|/2$, and hence

$$\left|\frac{1}{y_n} - \frac{1}{K}\right| = \frac{|y_n - K|}{|y_n||K|} \leq \frac{2}{|K|^2}|y_n - K| \longrightarrow 0.$$

Thus $1/y_n \to 1/K$, and the product rule gives $x_n/y_n \to L/K$. □

Definition 7.8 The sequence (x_n) is called **increasing** if

$$x_n \leq x_{n+1} \quad \text{for all } n \in \mathbb{N}.$$

Similarly, (x_n) is called **decreasing** if $x_n \geq x_{n+1}$ for all n. The sequence (x_n) is called **monotonic** if it is either increasing or decreasing.

Example 7.9 The sequence $x_n = n$ is increasing, while $\left(\frac{1}{n}\right)$ is decreasing. The convergent sequence $\left(\frac{(-1)^n}{n}\right)$ is neither increasing nor decreasing.

Lecture 8 — Wednesday, September 26, 2012

Theorem 8.1 (Monotone convergence theorem) If (x_n) is a bounded monotonic sequence, then (x_n) converges.

Proof. There are two cases to consider. First assume that (x_n) is increasing. Since the sequence is bounded, the set

$$A = \{x_1, x_2, x_3, \dots\}$$

is bounded above. By the completeness property, A has a least upper bound; write

$$L = \sup A.$$

We claim that $x_n \to L$. Let $\varepsilon > 0$. Since $L - \varepsilon < L$ and L is the least upper bound of A, the number $L - \varepsilon$ cannot be an upper bound for A. Therefore there exists $N \in \mathbb{N}$ such that

$$x_N > L - \varepsilon.$$

If $n \geq N$, then $x_n \geq x_N$ because the sequence is increasing, and also $x_n \leq L$ because L is an upper bound for A. Hence

$$L - \varepsilon < x_N \leq x_n \leq L < L + \varepsilon.$$

Thus $|x_n - L| < \varepsilon$ for all $n \geq N$, so $x_n \to L$.

Now assume that (x_n) is decreasing. Since the sequence is bounded, the set A is bounded below, so it has a greatest lower bound; write

$$\ell = \inf A.$$

Let $\varepsilon > 0$. Because $\ell + \varepsilon > \ell$ and ℓ is the greatest lower bound of A, the number $\ell + \varepsilon$ cannot be a lower bound for A. Hence there exists $N \in \mathbb{N}$ such that

$$x_N < \ell + \varepsilon.$$

If $n \geq N$, then $x_n \leq x_N$ because the sequence is decreasing, and also $x_n \geq \ell$ because ℓ is a lower bound for A. Therefore

$$\ell - \varepsilon < \ell \leq x_n \leq x_N < \ell + \varepsilon,$$

so $|x_n - \ell| < \varepsilon$ for all $n \geq N$. Hence $x_n \to \ell$. □

The proof of Theorem 8.1 shows slightly more than the statement: every increasing sequence that is bounded above converges, and every decreasing sequence that is bounded below converges. In particular, any increasing sequence is automatically bounded below by x_1, and any decreasing sequence is automatically bounded above by x_1.

Example 8.2 Let

$$a_1 = 1 \qquad \text{and} \qquad a_{n+1} = \frac{2a_n + 5}{6} \quad \text{for } n \geq 1.$$

Show that (a_n) converges and determine its limit.

Solution. First, (a_n) is increasing. Indeed, $a_1 \le a_2$, and if $a_{n-1} \le a_n$, then

$$a_n = \frac{2a_{n-1}+5}{6} \le \frac{2a_n+5}{6} = a_{n+1}.$$

By induction, $a_n \le a_{n+1}$ for all n.

Next, (a_n) is bounded above by 2. The base case $a_1 \le 2$ is clear. If $a_n \le 2$, then

$$a_{n+1} = \frac{2a_n+5}{6} \le \frac{2\cdot 2+5}{6} = \frac{9}{6} \le 2.$$

So $a_n \le 2$ for every n. By the monotone convergence theorem, (a_n) converges to some limit L.

To find L, note that the shifted sequence (a_{n+1}) has the same limit as (a_n). On the other hand,

$$a_{n+1} = \frac{2a_n+5}{6},$$

so the limit laws give

$$L = \frac{2L+5}{6}.$$

Solving this equation yields $L = \frac{5}{4}$. □

Observe that $L \le 2$ in Example 8.2. A general fact is that if $a_n \le M$ for all $n \in \mathbb{N}$ and $a_n \to L$, then $L \le M$. A useful strategy for many limit problems is to make some guesses for what the limit is, and then prove that one is correct.

However, it can sometimes be challenging to guess the limit of a sequence by inspection or trial and error, and we may just have to settle for a proof of convergence to *some* number.

Example 8.3 Consider the sequence of partial sums

$$x_n = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots + \frac{1}{n^2}.$$

Show that (x_n) converges.

Solution. This sequence is increasing, since each new term added is positive. We now show that it is bounded above. Group the terms dyadically:

$$x_n \le 1 + \left(\frac{1}{2^2} + \frac{1}{3^2}\right) + \left(\frac{1}{4^2} + \cdots + \frac{1}{7^2}\right) + \left(\frac{1}{8^2} + \cdots + \frac{1}{15^2}\right) + \cdots .$$

For each integer $m \ge 1$, the block from 2^m to $2^{m+1}-1$ contains 2^m terms, and each of those terms is at most $\frac{1}{(2^m)^2}$. Therefore that entire block is at most

$$2^m \cdot \frac{1}{(2^m)^2} = \frac{1}{2^m}.$$

Hence

$$x_n \le 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \cdots \le 2.$$

Thus (x_n) is increasing and bounded above, so the monotone convergence theorem implies that (x_n) converges. □

Determining the exact limit is substantially harder. This is the *Basel problem*; Euler showed in 1734 that

$$x_n \to \frac{\pi^2}{6}.$$

It is a good example of a nonconstructive argument: proving that a limit exists can be much easier than finding its value.

Subsequences

Definition 8.4 Given a sequence (x_n), choose indices

$$n_1 < n_2 < n_3 < \cdots .$$

The sequence $(y_k)_{k=1}^{\infty}$ defined by $y_k = x_{n_k}$ is called a **subsequence** of (x_n).

Example 8.5 Choosing the indices $n_1 = 3$, $n_2 = 6$, $n_3 = 7$, and $n_4 = 11$ produces the subsequence

$$y_1 = x_3, \qquad y_2 = x_6, \qquad y_3 = x_7, \qquad y_4 = x_{11}, \ldots$$

Example 8.6 If $x_n = (-1)^n$ and we choose $n_j = 2j$, then the resulting subsequence is

$$y_j = x_{2j} = 1 \quad \text{for all } j.$$

Proposition 8.7 If $(x_n) \to L$, then any subsequence of (x_n) also converges to L.

Proof. Let (y_k) be a subsequence of (x_n), say $y_k = x_{n_k}$. Fix $\varepsilon > 0$. Since $x_n \to L$, there exists $N \in \mathbb{N}$ such that

$$|x_n - L| < \varepsilon \quad \text{for all } n \geq N.$$

If $k \geq N$, then $n_k \geq k \geq N$, so

$$|y_k - L| = |x_{n_k} - L| < \varepsilon.$$

Therefore $y_k \to L$. $\square$

If two subsequences of the same sequence converge to different limits, then the original sequence cannot converge.

Proposition 8.8 If (x_n) is any sequence, it has a monotonic subsequence.

Proof. Call a term x_k a **peak point** if

$$x_k \geq x_n \quad \text{for all } n \geq k.$$

First suppose that (x_n) has infinitely many peak points. Choose them in increasing order of index:

$$n_1 < n_2 < n_3 < \cdots .$$

Then $x_{n_k} \geq x_{n_{k+1}}$ for every k, because $n_{k+1} > n_k$ and x_{n_k} is a peak point. Thus the subsequence (x_{n_k}) is decreasing.

Now suppose that (x_n) has only finitely many peak points. Then there exists $N \in \mathbb{N}$ such that no term x_n with $n \geq N$ is a peak point. Choose $n_1 \geq N$. Since x_{n_1} is not a peak point, there exists $n_2 > n_1$ with

$$x_{n_2} > x_{n_1}.$$

Similarly, because x_{n_2} is not a peak point, there exists $n_3 > n_2$ with

$$x_{n_3} > x_{n_2}.$$

Continuing inductively, we construct indices

$$n_1 < n_2 < n_3 < \cdots$$

such that

$$x_{n_1} < x_{n_2} < x_{n_3} < \cdots .$$

Hence (x_{n_k}) is an increasing subsequence. In either case, (x_n) has a monotonic subsequence. $\square$

Lecture 9 — Friday, September 28, 2012

The previous lecture showed that every sequence has a monotonic subsequence. Combining that fact with the monotone convergence theorem gives a powerful result:

Theorem 9.1 (Bolzano–Weierstrass theorem) Any bounded sequence has a convergent subsequence.

Proof. Let (x_n) be bounded. By Proposition 8.8, it has a monotonic subsequence. Since every subsequence of a bounded sequence is again bounded, that monotonic subsequence is bounded. The monotone convergence theorem therefore implies that the subsequence converges. $\square$

Cauchy Sequences

Definition 9.2 A sequence (x_n) is called a **Cauchy sequence** if, for every $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$|x_n - x_m| < \varepsilon \quad \text{whenever } n, m \geq N.$$

Proposition 9.3 Any convergent sequence is Cauchy.

Proof. Suppose that $x_n \to L$, and let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that

$$|x_n - L| < \frac{\varepsilon}{2} \quad \text{for all } n \geq N.$$

If $n, m \geq N$, then the triangle inequality gives

$$|x_n - x_m| = |(x_n - L) + (L - x_m)| \leq |x_n - L| + |x_m - L| < \varepsilon.$$

Therefore (x_n) is Cauchy. $\square$

Proposition 9.4 Any Cauchy sequence is bounded.

Proof. Since (x_n) is Cauchy, there exists $N \in \mathbb{N}$ such that

$$|x_n - x_m| < 1 \quad \text{whenever } n, m \geq N.$$

In particular, if $n \geq N$, then

$$|x_n| = |(x_n - x_N) + x_N| \leq |x_n - x_N| + |x_N| < 1 + |x_N|.$$

Let C be the maximum of the finitely many numbers

$$|x_1|, |x_2|, \ldots, |x_{N-1}|, 1 + |x_N|.$$

Then $|x_n| \leq C$ for all n, so (x_n) is bounded. $\square$

Theorem 9.5 (Cauchy criterion) Every Cauchy sequence converges.

Proof. Let (x_n) be a Cauchy sequence. By Proposition 9.4, the sequence is bounded, so by the Bolzano–Weierstrass theorem it has a convergent subsequence. Write that subsequence as (x_{n_k}), and suppose

$$x_{n_k} \to L.$$

We show that the whole sequence converges to L. Let $\varepsilon > 0$. Since (x_n) is Cauchy, there exists $N_1 \in \mathbb{N}$ such that

$$|x_n - x_m| < \frac{\varepsilon}{2} \quad \text{whenever } n, m \geq N_1.$$

Since $x_{n_k} \to L$, there exists $N_2 \in \mathbb{N}$ such that

$$|x_{n_k} - L| < \frac{\varepsilon}{2} \quad \text{for all } k \geq N_2.$$

Now fix any $n \geq N_1$. Choose $k \geq \max\{N_1, N_2\}$. Then $n_k \geq k \geq N_1$, so

$$|x_n - L| = |(x_n - x_{n_k}) + (x_{n_k} - L)| \leq |x_n - x_{n_k}| + |x_{n_k} - L| < \varepsilon.$$

Therefore $x_n \to L$. $\square$

Example 9.6 Suppose (x_n) satisfies

$$|x_{n+1} - x_n| \leq \frac{1}{2^n} \quad \text{for all } n \in \mathbb{N}.$$

Show that (x_n) converges.

Solution. It is enough to show that (x_n) is Cauchy. Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ such that

$$\frac{1}{2^{N-1}} < \varepsilon.$$

Now take integers $m > n \geq N$. Then

$$\begin{aligned}
|x_n - x_m| &= |(x_n - x_{n+1}) + (x_{n+1} - x_{n+2}) + \cdots + (x_{m-1} - x_m)| \\
&\leq |x_n - x_{n+1}| + |x_{n+1} - x_{n+2}| + \cdots + |x_{m-1} - x_m| \\
&\leq \frac{1}{2^n} + \frac{1}{2^{n+1}} + \cdots + \frac{1}{2^{m-1}} \\
&\leq \sum_{k=n}^{\infty} \left(\frac{1}{2}\right)^k \\
&= \frac{1}{2^{n-1}} \\
&\leq \frac{1}{2^{N-1}} \\
&< \varepsilon.
\end{aligned}$$

Therefore (x_n) is Cauchy, and hence it converges. $\square$

Lecture 10 — Monday, October 01, 2012

If a sequence is Cauchy, then its successive differences must tend to 0. The converse is false: the condition

$$|x_{n+1} - x_n| \to 0$$

does not by itself imply that (x_n) is Cauchy. A sufficiently strong bound on the differences does work. For example, if

$$|x_{n+1} - x_n| < \frac{1}{2^n}$$

for all $n \in \mathbb{N}$, then the same telescoping estimate used in the preceding example shows that (x_n) is Cauchy.

Example 10.1 Consider the harmonic partial sums

$$x_n = 1 + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n}.$$

Then

$$x_{n+1} - x_n = \frac{1}{n+1} \to 0,$$

but (x_n) is not Cauchy.

Proof. Indeed, the sequence is unbounded. For each integer $m \geq 1$,

$$x_{2^m} = 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \cdots + \frac{1}{8}\right) + \cdots + \left(\frac{1}{2^{m-1}+1} + \cdots + \frac{1}{2^m}\right).$$

For $j = 1, \ldots, m$, the block from $2^{j-1}+1$ through 2^j contains 2^{j-1} terms, each of which is at least $1/2^j$, so every block contributes at least $1/2$. Therefore

$$x_{2^m} \geq 1 + \frac{m}{2},$$

which tends to infinity as $m \to \infty$. Since every Cauchy sequence is bounded, (x_n) cannot be Cauchy. $\square$

3. Limits, Continuity, and Inverse Functions

Limits at a Point

Definition 10.2 If $f\colon A \to B$, then A is called the **domain** of f, B is called the **codomain** of f, and the set

$$f(A) = \{f(x) \mid x \in A\} \subseteq B$$

is called the **range** of f.

Definition 10.3 Let $f\colon A \to \mathbb{R}$, and let p be an accumulation point of A. We say that the **limit** of $f(x)$ as x approaches p is the real number L, and we write

$$\lim_{x\to p} f(x) = L,$$

if for every $\varepsilon > 0$ there exists $\delta > 0$ such that whenever $x \in A$ and

$$0 < |x - p| < \delta,$$

it follows that

$$|f(x) - L| < \varepsilon.$$

Example 10.4 We first verify that

$$\lim_{x\to 2}(3x + 2) = 8.$$

Proof. Let $\varepsilon > 0$, and choose $\delta = \frac{\varepsilon}{3}$. If $0 < |x - 2| < \delta$, then

$$|(3x + 2) - 8| = 3|x - 2| < 3\delta = \varepsilon.$$

Therefore $\lim_{x\to 2}(3x + 2) = 8$. □

Example 10.5 We next show that

$$\lim_{x\to 2} x^2 = 4.$$

Proof. Let $\varepsilon > 0$, and choose
$$\delta = \min\left\{1, \frac{\varepsilon}{5}\right\}.$$
If $0 < |x-2| < \delta$, then in particular $|x-2| < 1$, so $1 < x < 3$ and hence $3 < x+2 < 5$. Therefore
$$\left|x^2 - 4\right| = |x-2||x+2| \leq 5|x-2| < 5\delta \leq \varepsilon.$$
Thus $\lim_{x\to 2} x^2 = 4$. $\square$

Example 10.6 We also have
$$\lim_{x\to 1} \frac{3}{x} = 3.$$

Proof. Let $\varepsilon > 0$, and choose
$$\delta = \min\left\{\frac{1}{2}, \frac{\varepsilon}{6}\right\}.$$
If $0 < |x-1| < \delta$, then $|x-1| < \frac{1}{2}$, so $x > \frac{1}{2}$ and therefore $\frac{1}{|x|} \leq 2$. Hence
$$\left|\frac{3}{x} - 3\right| = \left|\frac{3-3x}{x}\right| = 3\frac{|1-x|}{|x|} \leq 6|x-1| < 6\delta \leq \varepsilon.$$
Thus $\lim_{x\to 1} \frac{3}{x} = 3$. $\square$

Example 10.7 Define
$$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q}, \\ 0 & \text{if } x \notin \mathbb{Q}. \end{cases}$$
(This function is often called **Dirichlet's function**.) Then $\lim_{x\to 5} f(x)$ does not exist.

Proof. Suppose, for a contradiction, that $\lim_{x\to 5} f(x) = L$. Take $\varepsilon = \frac{1}{2}$. For every $\delta > 0$, the interval $(5-\delta, 5+\delta)$ contains both rational and irrational numbers different from 5. Choose $x_1 \in \mathbb{Q}$ and $x_2 \notin \mathbb{Q}$ with
$$0 < |x_1 - 5| < \delta \qquad \text{and} \qquad 0 < |x_2 - 5| < \delta.$$
Then
$$f(x_1) = 1 \qquad \text{and} \qquad f(x_2) = 0.$$
But no interval of radius $\frac{1}{2}$ around a real number L can contain both 0 and 1. Therefore at least

one of $|f(x_1) - L|$ or $|f(x_2) - L|$ is at least $\frac{1}{2}$, contradicting the definition of the limit. Hence the limit does not exist. $\square$

Example 10.8 If

$$f(x) = \sin\left(\frac{1}{x}\right) \quad \text{for } x \neq 0,$$

then $\lim_{x\to 0} f(x)$ does not exist. See Figure 3.

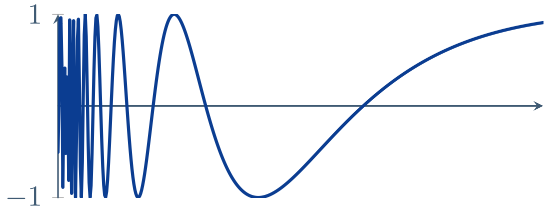

Figure 3

Proof. Suppose, for a contradiction, that $\lim_{x\to 0} f(x) = L$. Take $\varepsilon = \frac{1}{2}$. For any $\delta > 0$, choose n large enough that both

$$\frac{1}{2\pi n + \pi/2} < \delta \qquad \text{and} \qquad \frac{1}{2\pi n + 3\pi/2} < \delta.$$

Set

$$x_1 = \frac{1}{2\pi n + \pi/2}, \qquad \text{and} \qquad x_2 = \frac{1}{2\pi n + 3\pi/2}.$$

Then $0 < |x_1|, |x_2| < \delta$, while

$$f(x_1) = \sin\left(2\pi n + \frac{\pi}{2}\right) = 1 \qquad \text{and} \qquad f(x_2) = \sin\left(2\pi n + \frac{3\pi}{2}\right) = -1.$$

Again, no interval of radius $\frac{1}{2}$ around a single real number can contain both 1 and -1. This contradicts the definition of the limit, so $\lim_{x\to 0} \sin(1/x)$ does not exist. $\square$

Lecture 11 — Wednesday, October 03, 2012

Example 11.1 We have
$$\lim_{x\to 0}\frac{\sqrt{x+1}-1}{x}=\frac{1}{2}.$$

Proof. Let $\varepsilon > 0$, and choose $\delta = 2\varepsilon$. For $x \neq 0$,
$$\frac{\sqrt{x+1}-1}{x}=\frac{\sqrt{x+1}-1}{x}\cdot\frac{\sqrt{x+1}+1}{\sqrt{x+1}+1}=\frac{1}{\sqrt{x+1}+1}.$$

Hence
$$\begin{aligned}\left|\frac{\sqrt{x+1}-1}{x}-\frac{1}{2}\right| &= \left|\frac{1}{\sqrt{x+1}+1}-\frac{1}{2}\right|\\ &= \left|\frac{1-\sqrt{x+1}}{2(\sqrt{x+1}+1)}\right|\\ &= \left|\frac{(1-\sqrt{x+1})(1+\sqrt{x+1})}{2(\sqrt{x+1}+1)^2}\right|\\ &= \frac{|x|}{2(\sqrt{x+1}+1)^2}\\ &\leq \frac{|x|}{2}.\end{aligned}$$

Therefore, if $0 < |x| < \delta$, then
$$\left|\frac{\sqrt{x+1}-1}{x}-\frac{1}{2}\right| < \frac{\delta}{2}=\varepsilon.$$

$\square$

One-Sided Limits and Limit Laws

Definition 11.2 (One-sided limits) Let $f\colon A \to \mathbb{R}$. When p can be approached by points of A from the right, we write

$$\lim_{x \to p^+} f(x) = L$$

if for every $\varepsilon > 0$ there exists $\delta > 0$ such that whenever $x \in A$ and

$$0 < x - p < \delta,$$

we have $|f(x) - L| < \varepsilon$. When p can be approached by points of A from the left, we similarly write

$$\lim_{x \to p^-} f(x) = L$$

means that for every $\varepsilon > 0$ there exists $\delta > 0$ such that whenever $x \in A$ and

$$0 < p - x < \delta,$$

we have $|f(x) - L| < \varepsilon$.

Theorem 11.3 Let $f\colon A \to \mathbb{R}$ and let p be a point that can be approached from both sides within A. Then

$$\lim_{x \to p} f(x) = L$$

if and only if

$$\lim_{x \to p^+} f(x) = L \qquad \text{and} \qquad \lim_{x \to p^-} f(x) = L.$$

Proof. Suppose first that $\lim_{x \to p} f(x) = L$. Given $\varepsilon > 0$, choose $\delta > 0$ so that

$$0 < |x - p| < \delta \implies |f(x) - L| < \varepsilon.$$

This same δ works for both the right-hand and left-hand limits, because the conditions

$$0 < x - p < \delta \qquad \text{and} \qquad 0 < p - x < \delta$$

are special cases of $0 < |x - p| < \delta$. Conversely, suppose that both one-sided limits equal L. Given $\varepsilon > 0$, let $\delta_1 > 0$ correspond to the right-hand limit and $\delta_2 > 0$ correspond to the left-hand limit.

Set
$$\delta = \min\{\delta_1, \delta_2\}.$$
If $0 < |x - p| < \delta$, then either $x > p$ or $x < p$. In the first case,
$$0 < x - p < \delta \le \delta_1,$$
so $|f(x) - L| < \varepsilon$. In the second case,
$$0 < p - x < \delta \le \delta_2,$$
so again $|f(x) - L| < \varepsilon$. Therefore $\lim_{x\to p} f(x) = L$. $\square$

Example 11.4 Let
$$f(x) = \begin{cases} 3x + 1, & x \ge 2, \\ \frac{1}{x}, & 0 < x < 2. \end{cases}$$
Then
$$\lim_{x\to 2^+} f(x) = 7 \qquad \text{and} \qquad \lim_{x\to 2^-} f(x) = \frac{1}{2}.$$
Since these one-sided limits are different, the two-sided limit $\lim_{x\to 2} f(x)$ does not exist.

Limit Laws

Theorem 11.5 (Limit laws for functions) Assume that
$$\lim_{x\to p} f(x) = L \qquad \text{and} \qquad \lim_{x\to p} g(x) = K.$$
Then
$$\lim_{x\to p}(f(x) \pm g(x)) = L \pm K$$
and
$$\lim_{x\to p}(f(x)g(x)) = LK.$$
If $K \neq 0$, then
$$\lim_{x\to p} \frac{f(x)}{g(x)} = \frac{L}{K}.$$

The proofs are parallel to the corresponding limit laws for sequences. The only new point is that, in the quotient rule, we must know that $g(x)$ does not vanish near p.

Proposition 11.6 If $\lim_{x\to p} g(x) = K \neq 0$, then there exists $\delta > 0$ such that

$$0 < |x-p| < \delta \implies g(x) \neq 0.$$

Proof. Take

$$\varepsilon = \frac{|K|}{2} > 0.$$

Since $\lim_{x\to p} g(x) = K$, there exists $\delta > 0$ such that

$$0 < |x-p| < \delta \implies |g(x) - K| < \frac{|K|}{2}.$$

If $0 < |x-p| < \delta$, then

$$|g(x)| \geq |K| - |g(x) - K| > \frac{|K|}{2} > 0.$$

Hence $g(x) \neq 0$. □

Theorem 11.7 (Squeeze theorem) Suppose that

$$f(x) \leq g(x) \leq h(x)$$

for all x sufficiently close to p, and assume that

$$\lim_{x\to p} f(x) = \lim_{x\to p} h(x) = L.$$

Then $\lim_{x\to p} g(x) = L$.

Proof. Let $\varepsilon > 0$. Choose $\delta_1 > 0$ so that

$$0 < |x-p| < \delta_1 \implies |f(x) - L| < \varepsilon$$

and choose $\delta_2 > 0$ so that

$$0 < |x-p| < \delta_2 \implies |h(x) - L| < \varepsilon.$$

Also choose $\delta_3 > 0$ so that

$$0 < |x-p| < \delta_3$$

guarantees $f(x) \leq g(x) \leq h(x)$. Let

$$\delta = \min\{\delta_1, \delta_2, \delta_3\}.$$

If $0 < |x - p| < \delta$, then

$$L - \varepsilon < f(x) \leq g(x) \leq h(x) < L + \varepsilon.$$

Hence $|g(x) - L| < \varepsilon$, so $\lim_{x \to p} g(x) = L$. $\square$

Example 11.8 We can now evaluate

$$\lim_{x \to 0} x \sin\left(\frac{1}{x}\right).$$

Since

$$-|x| \leq x \sin\left(\frac{1}{x}\right) \leq |x|$$

for all $x \neq 0$, and since both $-|x|$ and $|x|$ tend to 0 as $x \to 0$, the squeeze theorem gives

$$\lim_{x \to 0} x \sin\left(\frac{1}{x}\right) = 0.$$

Proposition 11.9 $\lim_{x \to 0} \dfrac{\sin(x)}{x} = 1.$

Proof. We first consider the right-hand limit. For $0 < x < \frac{\pi}{4}$, Euclidean geometry gives

$$\sin(x) \leq x \leq \tan(x) = \frac{\sin(x)}{\cos(x)}.$$

Since $\sin(x) > 0$ in this range, dividing by $\sin(x)$ and then multiplying by $\cos(x)$ yields

$$\cos(x) \leq \frac{\sin(x)}{x} \leq 1.$$

Also, for $0 < x < \frac{\pi}{4}$ we have $0 \leq \sin(x) \leq x$, so by the squeeze theorem,

$$\lim_{x \to 0^+} \sin(x) = 0.$$

Therefore

$$\lim_{x \to 0^+} \sin^2(x) = 0,$$

and hence

$$\lim_{x\to 0^+} \cos^2(x) = \lim_{x\to 0^+} (1 - \sin^2(x)) = 1.$$

Because $\cos(x) > 0$ near 0, it follows that

$$\lim_{x\to 0^+} \cos(x) = 1.$$

Applying the squeeze theorem to

$$\cos(x) \le \frac{\sin(x)}{x} \le 1$$

shows that

$$\lim_{x\to 0^+} \frac{\sin(x)}{x} = 1.$$

Finally, the function $\frac{\sin(x)}{x}$ is even for $x \neq 0$, since

$$\frac{\sin(-x)}{-x} = \frac{\sin(x)}{x}.$$

Therefore the left-hand limit at 0 is also 1, and the two-sided limit exists with value 1. $\square$

Corollary 11.10 Using Proposition 11.9, we obtain

$$\lim_{x\to 0} \frac{1 - \cos(x)}{x} = 0.$$

Proof. For $x \neq 0$,

$$\frac{1-\cos(x)}{x} = \frac{1-\cos(x)}{x} \cdot \frac{1+\cos(x)}{1+\cos(x)} = \frac{1-\cos^2(x)}{x(1+\cos(x))} = \left(\frac{\sin(x)}{x}\right)\left(\frac{\sin(x)}{1+\cos(x)}\right).$$

As $x \to 0$, we have $\frac{\sin(x)}{x} \to 1$, $\sin(x) \to 0$ by Proposition 11.9, and $\cos(x) \to 1$. Hence

$$\frac{\sin(x)}{1+\cos(x)} \to \frac{0}{2} = 0.$$

The product law now gives

$$\lim_{x\to 0} \frac{1-\cos(x)}{x} = 1 \cdot 0 = 0.$$

$\square$

Continuity

Definition 11.11 Let $f\colon A \to \mathbb{R}$ with $A \subseteq \mathbb{R}$. We say that f is **continuous at** $a \in A$ if, for every $\varepsilon > 0$, there exists $\delta > 0$ such that whenever $x \in A$ and

$$|x - a| < \delta,$$

it follows that

$$|f(x) - f(a)| < \varepsilon.$$

For $\lim_{x\to a} f(x)$, we examine points with

$$0 < |x - a| < \delta,$$

so the function must be defined at all sufficiently nearby points of the domain, excluding a itself. In contrast, continuity at a compares $f(x)$ directly with $f(a)$, so we only consider points $x \in A$ satisfying $|x - a| < \delta$.

Lecture 12 — Friday, October 05, 2012

Examine the definition of continuity at a. If $A = \mathbb{R}$, then the phrase "and $x \in A$" is unnecessary, and the definition reduces to the familiar condition

$$\lim_{x\to a} f(x) = f(a).$$

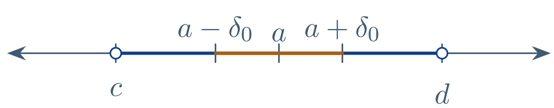

FIGURE 4

Now suppose

$$A = (c, d) = \{x \mid c < x < d\}.$$

Suppose $a \in (c, d)$; see Figure 4. Since $(a - \delta_0, a + \delta_0) \subseteq A$ for some sufficiently small $\delta_0 > 0$, the phrase "and $x \in A$" again adds nothing, and continuity at a is equivalent to

$$\lim_{x\to a} f(x) = f(a).$$

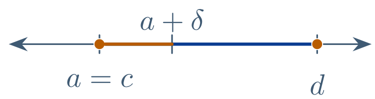

FIGURE 5

On the other hand, if

$$A = [c, d] = \{x \mid c \le x \le d\}$$

and $a = c$ (see Figure 5), then "and $x \in A$" adds that we check $x \ge a$, and continuity at a is equivalent to

$$\lim_{x \to a^+} f(x) = f(a).$$

Similarly, if $a = d$, then continuity at a is equivalent to

$$\lim_{x \to a^-} f(x) = f(a).$$

Thus continuity at an interior point is a two-sided limit condition, while continuity at an endpoint is naturally expressed using a one-sided limit.

Example 12.1 Define

$$f(x) = \begin{cases} \dfrac{2x^2 - 8}{x - 2}, & x \ne 2, \\ 3, & x = 2. \end{cases}$$

Then f is not continuous at $x = 2$.

Proof. For $x \ne 2$,

$$\frac{2x^2 - 8}{x - 2} = \frac{2(x-2)(x+2)}{x - 2} = 2x + 4.$$

Therefore

$$\lim_{x \to 2} f(x) = \lim_{x \to 2} (2x + 4) = 8.$$

But $f(2) = 3$, and $8 \ne 3$. Hence $\lim_{x \to 2} f(x) \ne f(2)$, so f is not continuous at 2. $\square$

Discontinuities can arise in different ways. A function may have a jump discontinuity, an oscillatory discontinuity such as

$$f(x) = \begin{cases} \sin(1/x), & x \ne 0, \\ 1, & x = 0, \end{cases}$$

or a removable discontinuity such as the previous example.

Example 12.2 The function

$$f(x) = \begin{cases} 1, & x \in \mathbb{Q}, \\ 0, & x \notin \mathbb{Q} \end{cases}$$

is discontinuous at every real number a.

Proof. At every point $a \in \mathbb{R}$, every interval around a contains both rational and irrational numbers. Thus the same argument used in Example 10.7 at $x = 5$ shows that $\lim_{x \to a} f(x)$ does not exist. Since continuity at a would require this limit to equal $f(a)$, the function is discontinuous everywhere. $\square$

Exercise 12.3 Show that a continuous function which is equal to 1 at every rational number is 1 everywhere. Also show that a continuous function which is equal to 0 at every irrational is 0 everywhere.

Example 12.4 Let

$$f(x) = \begin{cases} 0, & x \notin \mathbb{Q}, \\ 1, & x = 0, \\ \frac{1}{n}, & x \in \mathbb{Q} \setminus \{0\},\ x = \frac{m}{n} \text{ in lowest terms with } n \in \mathbb{N}. \end{cases}$$

This highly pathological function, often called **Thomae's function**, is discontinuous at every rational point and continuous at every irrational point.

Proof. First suppose that $a \in \mathbb{Q}$. If $a = 0$, then $f(a) = 1$. Take $\varepsilon = \frac{1}{2}$. Every interval around 0 contains irrational numbers x, and for such x we have $f(x) = 0$. Hence

$$|f(x) - f(0)| = 1 \not< \frac{1}{2},$$

so f is not continuous at 0. Now suppose that $a \neq 0$ is rational, say $a = \frac{m}{n}$ in lowest terms. Then $f(a) = \frac{1}{n}$. Take

$$\varepsilon = \frac{1}{2n}.$$

Every interval around a contains irrational numbers x, and for such x we have $f(x) = 0$. Therefore

$$|f(x) - f(a)| = \frac{1}{n} > \frac{1}{2n} = \varepsilon,$$

so f is not continuous at a. Now suppose that $a \notin \mathbb{Q}$. Then $f(a) = 0$. Let $\varepsilon > 0$. Choose $N \in \mathbb{N}$ so that

$$\frac{1}{N} < \varepsilon,$$

which is possible by the Archimedean property. Consider the set of rational numbers in the interval $(a-1, a+1)$ whose reduced denominator is less than N. This set is finite. Since a is irrational, it does not belong to that finite set, so its distance from that set is positive. Let $\delta > 0$ be smaller than both 1 and that positive distance. If $|x - a| < \delta$, then $x \in (a-1, a+1)$. If x is irrational, then $f(x) = 0$. If x is rational, then its reduced denominator cannot be less than N, by the choice of δ. Therefore its reduced denominator is at least N, and so

$$f(x) \leq \frac{1}{N} < \varepsilon.$$

In either case,

$$|f(x) - f(a)| = |f(x)| < \varepsilon.$$

Hence f is continuous at every irrational point. □

Lecture 13 — Wednesday, October 10, 2012

Exercise 13.1 (Challenge) Prove that there is no function on $\mathbb{R}$ that is continuous at every rational number and discontinuous at every irrational number.

This exercise is *extremely* challenging without the Baire category theorem or Cantor's intersection theorem.

Theorem 13.2 Let $f\colon A \to \mathbb{R}$ and let $a \in A$. Then f is continuous at a if and only if, for every sequence (x_n) in A with $x_n \to a$, the sequence $(f(x_n))$ converges to $f(a)$.

Proof. *Forward implication.* Assume that f is continuous at a, and let (x_n) be a sequence in A such that $x_n \to a$. We must show that $f(x_n) \to f(a)$.

Let $\varepsilon > 0$. By continuity at a, there exists $\delta > 0$ such that whenever $x \in A$ and $|x - a| < \delta$, we have

$$|f(x) - f(a)| < \varepsilon.$$

Since $x_n \to a$, there exists $N \in \mathbb{N}$ such that $|x_n - a| < \delta$ for all $n \geq N$. Therefore

$$|f(x_n) - f(a)| < \varepsilon \quad \text{for all } n \geq N,$$

and hence $f(x_n) \to f(a)$.

Reverse implication. Assume that whenever (x_n) is a sequence in A with $x_n \to a$, we have $f(x_n) \to f(a)$. We prove that f is continuous at a.

Suppose, to the contrary, that f is not continuous at a. Then there exists $\varepsilon_0 > 0$ such that for every $\delta > 0$ there is some $x \in A$ with

$$|x - a| < \delta \qquad \text{but} \qquad |f(x) - f(a)| \geq \varepsilon_0.$$

In particular, for each $n \in \mathbb{N}$, taking $\delta = 1/n$, we can choose $x_n \in A$ such that

$$|x_n - a| < \frac{1}{n} \qquad \text{and} \qquad |f(x_n) - f(a)| \geq \varepsilon_0.$$

Then $x_n \to a$, so by assumption $f(x_n) \to f(a)$. This is impossible, because the inequality

$$|f(x_n) - f(a)| \geq \varepsilon_0$$

holds for every n. The contradiction shows that f is continuous at a. □

Proposition 13.3 Let $f, g\colon A \to \mathbb{R}$ be continuous at $a \in A$, and let $c \in \mathbb{R}$. Then $f + g$, $f - g$, cf, and fg are continuous at a. If, in addition, $g(a) \neq 0$, then f/g is continuous at a.

Proof. Let (x_n) be a sequence in A with $x_n \to a$. Since f and g are continuous at a, we have

$$f(x_n) \to f(a) \qquad \text{and} \qquad g(x_n) \to g(a).$$

The limit laws for sequences now give

$$(f+g)(x_n) = f(x_n) + g(x_n) \to f(a) + g(a) = (f+g)(a),$$

and similarly

$$(f-g)(x_n) \to (f-g)(a), \qquad (cf)(x_n) \to (cf)(a), \qquad \text{and} \qquad (fg)(x_n) \to (fg)(a).$$

If $g(a) \neq 0$, then $g(x_n) \to g(a) \neq 0$, so the quotient law yields

$$\frac{f(x_n)}{g(x_n)} \to \frac{f(a)}{g(a)}.$$

Therefore each of these functions is continuous at a by Theorem 13.2. $\square$

Corollary 13.4 Every polynomial is continuous at every real number.

Proof. The identity function $x \mapsto x$ is continuous, and constant functions are continuous. By repeated use of Proposition 13.3, every monomial $c_k x^k$ is continuous, and therefore every finite sum

$$p(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_n x^n$$

is continuous. $\square$

Corollary 13.5 Every rational function is continuous at each point of its domain.

Proof. A rational function has the form P/Q, where P and Q are polynomials. Since polynomials are continuous by Corollary 13.4, the quotient proposition shows that P/Q is continuous at every point where Q is nonzero. $\square$

Example 13.6 Define

$$f(x) = \begin{cases} x-2, & x < 1, \\ 3x^2 - 1, & x \geq 1. \end{cases}$$

Then f is continuous at every point except $x = 1$.

Proof. On the interval $(-\infty, 1)$, the function is given by $x - 2$, and on $(1, \infty)$ it is given by $3x^2 - 1$. Each of these expressions defines a continuous function, so f is continuous at every point other than 1. At $x = 1$,

$$\lim_{x \to 1^-} f(x) = \lim_{x \to 1^-} (x - 2) = -1,$$

whereas

$$\lim_{x \to 1^+} f(x) = \lim_{x \to 1^+} (3x^2 - 1) = 2.$$

Since the one-sided limits are not equal, the limit $\lim_{x\to 1} f(x)$ does not exist. Therefore f is not continuous at 1. $\square$

Definition 13.7 If $f\colon A \to B$ and $g\colon B \to \mathbb{R}$, then the **composition** $g \circ f$ is the function from A to $\mathbb{R}$ defined by

$$(g \circ f)(x) = g(f(x)).$$

$$A \xrightarrow{f} B \xrightarrow{g} \mathbb{R}$$

Theorem 13.8 Let $f\colon A \to B$ and $g\colon B \to \mathbb{R}$. If f is continuous at $a \in A$ and g is continuous at $f(a)$, then $g \circ f$ is continuous at a.

Proof (using Theorem 13.2). Let (x_n) be a sequence in A such that $x_n \to a$. Since f is continuous at a, we have

$$f(x_n) \to f(a).$$

The sequence $(f(x_n))$ lies in B, and g is continuous at $f(a)$. Therefore

$$g(f(x_n)) \to g(f(a)).$$

That is,

$$(g \circ f)(x_n) \to (g \circ f)(a).$$

By Theorem 13.2, $g \circ f$ is continuous at a. $\square$

Proof (using the definition). Let $\varepsilon > 0$. Because g is continuous at $f(a)$, there exists δ_0 such that

$$|g(y) - g(f(a))| < \varepsilon$$

for all $y \in B$ such that $|y - f(a)| < \delta_0$. Moreover, because f is continuous at a, there exists $\delta > 0$ such that

$$|f(x) - f(a)| < \delta_0$$

for all $x \in A$ such that $|x - a| < \delta$. Choosing $y = f(x)$, we see that there exists $\delta > 0$ such that

$$|g(f(x)) - g(f(a))| < \varepsilon$$

whenever $|x - a| < \delta$, and so $g \circ f$ is continuous at a. $\square$

Lecture 14 — Friday, October 12, 2012

The Intermediate and Extreme Value Theorems

Theorem 14.1 (Intermediate value theorem) Suppose that $f\colon [a,b] \to \mathbb{R}$ is continuous and that

$$f(a) < 0 < f(b).$$

Then there exists $c \in [a,b]$ such that $f(c) = 0$.

The completeness of $\mathbb{R}$ is essential here. The proof uses the least-upper-bound property, so the theorem can fail in ordered fields with gaps.

Proof. Let

$$A = \{x \in [a,b] : f(x) < 0\}.$$

Since $f(a) < 0$, we have $a \in A$, so A is nonempty. Also, $A \subseteq [a,b]$, so A is bounded above by b. By the completeness property, A has a least upper bound; let

$$c = \sup A.$$

Then $a \leq c \leq b$, so $c \in [a,b]$. We first show that $f(c) \leq 0$. For each $n \in \mathbb{N}$, the number $c - 1/n$ is not an upper bound for A, so there exists $x_n \in A$ such that

$$c - \frac{1}{n} < x_n \leq c.$$

Hence $x_n \to c$. Since each x_n belongs to A, we have $f(x_n) < 0$ for all n. By continuity of f at c, it follows that

$$f(x_n) \to f(c),$$

and therefore $f(c) \leq 0$. Next we show that $f(c) \geq 0$. If $c = b$, then $f(c) = f(b) > 0$, and we are done. So suppose that $c < b$. By the Archimedean property, there exists $N \in \mathbb{N}$ such that

$$c + \frac{1}{n} < b \quad \text{for all } n \geq N.$$

Since c is an upper bound for A, none of the points $c + 1/n$ belongs to A. Because $c + 1/n \in [a,b]$,

this means
$$f\left(c+\frac{1}{n}\right) \geq 0 \quad \text{for all } n \geq N.$$
The sequence $c+1/n$ converges to c, so continuity gives
$$f\left(c+\frac{1}{n}\right) \to f(c).$$
Hence $f(c) \geq 0$. We have shown that $f(c) \leq 0$ and $f(c) \geq 0$. Therefore $f(c) = 0$. $\square$

Corollary 14.2 Suppose that $f\colon [a,b] \to \mathbb{R}$ is continuous and that $f(a) < f(b)$. If $z \in \mathbb{R}$ satisfies
$$f(a) < z < f(b),$$
then there exists $c \in [a,b]$ such that $f(c) = z$.

Proof. Define $g\colon [a,b] \to \mathbb{R}$ by $g(x) = f(x) - z$. Then g is continuous, and
$$g(a) = f(a) - z < 0, \qquad \text{and} \qquad g(b) = f(b) - z > 0.$$
By the intermediate value theorem, there exists $c \in [a,b]$ such that $g(c) = 0$. Thus $f(c) = z$. $\square$

Theorem 14.3 Every polynomial of odd degree has at least one real root.

Proof. Let
$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0,$$
where n is odd and $a_n \neq 0$. Dividing by a_n does not change the roots, so we may assume without loss of generality that $a_n = 1$. Then
$$p(x) = x^n \left(1 + \frac{a_{n-1}}{x} + \frac{a_{n-2}}{x^2} + \cdots + \frac{a_0}{x^n}\right).$$

Set
$$M = \max\{|a_0|, |a_1|, \ldots, |a_{n-1}|\}.$$

Choose $N \in \mathbb{N}$ so large that $N > 2nM$. If $|x| \geq N$, then

$$\left|\frac{a_{n-1}}{x} + \frac{a_{n-2}}{x^2} + \cdots + \frac{a_0}{x^n}\right| \leq \frac{M}{|x|} + \frac{M}{|x|^2} + \cdots + \frac{M}{|x|^n} \leq \frac{nM}{|x|} \leq \frac{1}{2}.$$

Therefore, for $|x| \geq N$,

$$1 + \frac{a_{n-1}}{x} + \frac{a_{n-2}}{x^2} + \cdots + \frac{a_0}{x^n} \geq \frac{1}{2}.$$

If $x \geq N$, then $x^n > 0$, so

$$p(x) \geq \frac{x^n}{2} > 0.$$

If $x \leq -N$, then $x^n < 0$ because n is odd, so

$$p(x) \leq \frac{x^n}{2} < 0.$$

In particular,

$$p(-N) < 0 < p(N).$$

Since polynomials are continuous, the intermediate value theorem yields some $c \in [-N, N]$ such that $p(c) = 0$. See Figure 6.

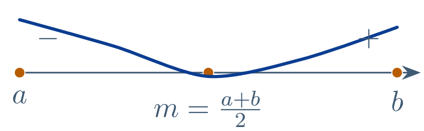

FIGURE 6

□

If $f(a) < 0 < f(b)$ and $m = (a+b)/2$, then either $f(m) = 0$, or else one of the intervals $[a, m]$ or $[m, b]$ still has opposite signs at its endpoints. Applying the intermediate value theorem to that smaller interval and repeating the procedure produces a nested sequence of intervals whose lengths are

$$\frac{b-a}{2}, \frac{b-a}{2^2}, \frac{b-a}{2^3}, \ldots.$$

See Figure 7. This is the **bisection method**. It may not locate the root exactly in finitely many steps, but it gives a rapidly improving approximation and requires continuity rather than differentiability.

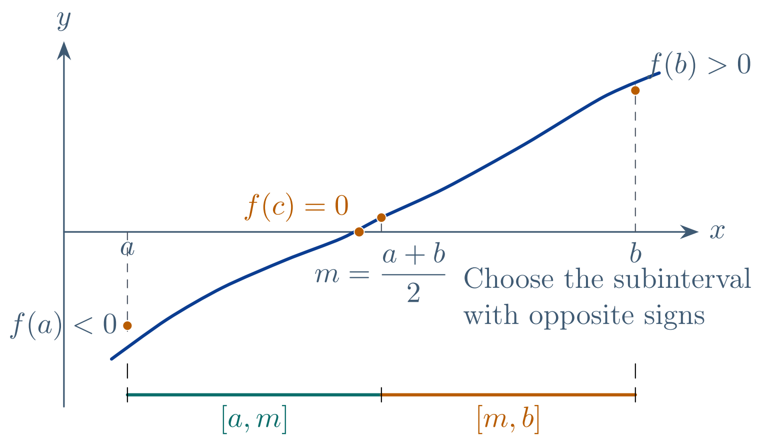

FIGURE 7

Definition 14.4 Let $f\colon A \to \mathbb{R}$. We say that f is **bounded above** if there exists $M \in \mathbb{R}$ such that

$$f(x) \leq M \quad \text{for all } x \in A.$$

Similarly, f is **bounded below** if there exists $m \in \mathbb{R}$ such that

$$m \leq f(x) \quad \text{for all } x \in A.$$

If both conditions hold, then f is **bounded**.

Example 14.5 The identity function $f(x) = x$ on $\mathbb{R}$ is neither bounded above nor bounded below. The function $\sin(x)$ is bounded above by 1 and below by -1, and both bounds are attained. The function $f(x) = -1/x^2$ on $\mathbb{R} \setminus \{0\}$ is bounded above by 0, but it does not attain a maximum value.

Lecture 15 — Monday, October 15, 2012

Theorem 15.1 (Extreme value theorem) Let $f\colon [a, b] \to \mathbb{R}$ be continuous. Then there exist points $c, d \in [a, b]$ such that

$$f(c) \leq f(x) \leq f(d) \quad \text{for every } x \in [a, b].$$

In particular, f is bounded on $[a, b]$ and attains both its minimum and its maximum there.

The hypotheses in the extreme value theorem are essential. A continuous function on an open interval need not attain a maximum or minimum. For example, $f(x) = x$ on $(0, 1)$ is bounded but attains neither an absolute maximum nor an absolute minimum. Likewise, $f(x) = \tan(x)$ on $(-\pi/2, \pi/2)$ is continuous but unbounded.

Proof of the extreme value theorem (Theorem 15.1). *Step 1: boundedness.* We first show that f is bounded above. Suppose not. Then for each $n \in \mathbb{N}$ there exists $x_n \in [a, b]$ such that

$$f(x_n) \geq n.$$

The sequence (x_n) lies in the bounded interval $[a, b]$, so the Bolzano–Weierstrass theorem yields a convergent subsequence (x_{n_k}). Its limit x_0 lies in $[a, b]$, because $a \leq x_{n_k} \leq b$ and these inequalities pass to the limit. Since f is continuous at x_0, we have

$$f(x_{n_k}) \to f(x_0).$$

But $f(x_{n_k}) \geq n_k \geq k$ for every k, so $(f(x_{n_k}))$ cannot even be bounded, let alone convergent. This contradiction shows that f is bounded above. Applying the same argument to $-f$ shows that f is bounded below.

Step 2: existence of a maximum. Let

$$A = f([a, b]) = \{f(x) : x \in [a, b]\}.$$

The set A is nonempty and bounded above, so the completeness property implies that A has a least upper bound. Write

$$z = \sup A.$$

For each $n \in \mathbb{N}$, the number $z - 1/n$ is not an upper bound for A, so there exists $y_n \in A$ such that

$$z - \frac{1}{n} < y_n \leq z.$$

Choose $x_n \in [a, b]$ with $f(x_n) = y_n$. Again, by the Bolzano–Weierstrass theorem, some subsequence (x_{n_k}) converges. As above, its limit x_0 belongs to the closed interval $[a, b]$. Continuity gives

$$f(x_{n_k}) \to f(x_0).$$

On the other hand, since
$$z - \frac{1}{n_k} < y_{n_k} \le z,$$
the squeeze theorem shows that $y_{n_k} \to z$. Because $y_{n_k} = f(x_{n_k})$, uniqueness of limits yields
$$f(x_0) = z.$$
Thus f attains its maximum at x_0.

Applying the same argument to $-f$ shows that f also attains its minimum on $[a, b]$. $\square$

Inverse Functions

Definition 15.2 A function $f\colon A \to B$ is called **one-to-one**, or **injective**, if
$$f(x_1) = f(x_2) \qquad \text{implies that} \qquad x_1 = x_2$$
for all $x_1, x_2 \in A$.

Graphically, a function is injective exactly when it passes the horizontal line test: every horizontal line intersects the graph at most once.

Definition 15.3 Let $f\colon A \to B$ be injective. The **inverse function** of f is the function
$$f^{-1}\colon f(A) \to A$$
defined by
$$f^{-1}(y) = x \quad \text{whenever } y = f(x).$$

If f is injective, then
$$f \circ f^{-1} = \mathrm{id}_{f(A)} \qquad \text{and} \qquad f^{-1} \circ f = \mathrm{id}_A\,.$$
In particular, the domain of f^{-1} is the range of f, and the range of f^{-1} is the domain of f.

Example 15.4 Consider $f(x) = x^3 + 1$ on $\mathbb{R}$. This function is injective, and its inverse is
$$f^{-1}(x) = \sqrt[3]{x-1}.$$

Indeed, if $y = f^{-1}(x)$, then $x = y^3 + 1$, so $y = \sqrt[3]{x-1}$.

Lecture 16 — Wednesday, October 17, 2012

Definition 16.1 Let $f\colon A \to \mathbb{R}$. We say that f is **strictly increasing** if

$$x_1 < x_2 \qquad \text{implies that} \qquad f(x_1) < f(x_2)$$

for all $x_1, x_2 \in A$. Similarly, f is **strictly decreasing** if

$$x_1 < x_2 \qquad \text{implies that} \qquad f(x_1) > f(x_2)$$

for all $x_1, x_2 \in A$. A function is called **strictly monotone** if it is either strictly increasing or strictly decreasing.

On an arbitrary domain, an injective function need not be monotone. The next theorem shows that continuity on an interval forces monotonicity.

Theorem 16.2 If $f\colon [a,b] \to \mathbb{R}$ is continuous and one-to-one, then f is strictly monotone.

Proof. Since f is one-to-one, we have $f(a) \neq f(b)$. Suppose first that

$$f(a) < f(b).$$

We will show that f is strictly increasing. We claim that for every $x \in (a,b)$,

$$f(a) < f(x) < f(b).$$

Indeed, if $f(x) \geq f(b)$, then the intermediate value theorem applied to f on $[a,x]$ shows that there exists $c \in [a,x]$ such that $f(c) = f(b)$. Since $x < b$, this gives $c \neq b$, contradicting injectivity. Similarly, if $f(x) \leq f(a)$, then the intermediate value theorem applied on $[x,b]$ produces some $c \in [x,b]$ with $f(c) = f(a)$, again contradicting injectivity. This proves the claim.

Now let $x, y \in [a,b]$ with $x < y$. We show that $f(x) < f(y)$. If $x = a$, then either the claim or the assumed inequality $f(a) < f(b)$ gives

$$f(a) < f(y),$$

so $f(x) < f(y)$. Similarly, if $y = b$, the claim gives $f(x) < f(b) = f(y)$. It remains to consider $a < x < y < b$. In this case,

$$f(a) < f(y) < f(b) \qquad \text{and} \qquad f(a) < f(x) < f(b).$$

Suppose, for a contradiction, that $f(x) \geq f(y)$. Since f is injective, this would imply $f(x) > f(y)$. The value $f(y)$ lies strictly between $f(a)$ and $f(x)$, so the intermediate value theorem applied on $[a, x]$ gives some $c \in (a, x)$ such that

$$f(c) = f(y).$$

This contradicts injectivity, since $c \neq y$. Therefore $f(x) < f(y)$. Hence f is strictly increasing. If instead $f(a) > f(b)$, then the same argument applied to $-f$ shows that f is strictly decreasing. In either case, f is strictly monotone. $\square$

Exercise 16.3 Prove that the same conclusion holds for a continuous one-to-one function $f\colon (a, b) \to \mathbb{R}$ by restricting f to each compact subinterval $[a + \varepsilon, b - \varepsilon]$.

Corollary 16.4 If $f\colon [a, b] \to \mathbb{R}$ is continuous and one-to-one, then the range of f is a closed interval.

Proof. By Theorem 16.2, f is strictly monotone. If f is strictly increasing, then

$$f([a, b]) = [f(a), f(b)].$$

Indeed, if $x \in [a, b]$, then $f(a) \leq f(x) \leq f(b)$, so $f([a, b]) \subseteq [f(a), f(b)]$. Conversely, if $z \in [f(a), f(b)]$, then the intermediate value theorem provides some $x \in [a, b]$ with $f(x) = z$. Thus every value between $f(a)$ and $f(b)$ is attained. The decreasing case is analogous, with range $[f(b), f(a)]$. $\square$

Theorem 16.5 If $f\colon [a, b] \to \mathbb{R}$ is continuous and one-to-one, then

$$f^{-1}\colon f([a, b]) \to [a, b]$$

is continuous.

Proof. Let (x_n) be a sequence in $f([a, b])$ such that $x_n \to x_0 \in f([a, b])$. To prove that f^{-1} is continuous, it suffices to show that

$$f^{-1}(x_n) \to f^{-1}(x_0).$$

Set

$$y_n = f^{-1}(x_n) \qquad \text{and} \qquad y_0 = f^{-1}(x_0).$$

Then each y_n lies in $[a, b]$, and

$$f(y_n) = x_n \qquad \text{and} \qquad f(y_0) = x_0.$$

Suppose, for a contradiction, that $y_n \not\to y_0$. Then there exists $\varepsilon > 0$ such that infinitely many terms of the sequence satisfy

$$|y_n - y_0| \geq \varepsilon.$$

Passing to a subsequence if necessary, we may assume that this inequality holds for every n. Since (y_n) lies in the bounded interval $[a, b]$, the Bolzano–Weierstrass theorem gives a convergent subsequence (y_{n_k}) with limit $t_0 \in [a, b]$. Because

$$|y_{n_k} - y_0| \geq \varepsilon$$

for every k, we must have $t_0 \neq y_0$. By continuity of f,

$$f(y_{n_k}) \to f(t_0).$$

But $f(y_{n_k}) = x_{n_k}$ and $x_{n_k} \to x_0 = f(y_0)$, so uniqueness of limits implies

$$f(t_0) = f(y_0).$$

Since f is one-to-one, this forces $t_0 = y_0$, a contradiction. Therefore $y_n \to y_0$, and hence f^{-1} is continuous. $\square$

Lecture 17 — Friday, October 19, 2012

4. Elementary Functions and Differentiation

Elementary Functions and the Derivative

Definition 17.1 Restrict the sine function to the interval $[-\pi/2, \pi/2]$. This restricted function is continuous and strictly increasing, so it is invertible. Its inverse is denoted by

$$\arcsin\colon [-1,1] \to \left[-\frac{\pi}{2}, \frac{\pi}{2}\right].$$

Thus $y = \arcsin(x)$ means that $\sin(y) = x$ and

$$y \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right].$$

Similarly, restricting $\cos(x)$ to $[0, \pi]$ gives the inverse function

$$\arccos\colon [-1,1] \to [0,\pi],$$

and restricting $\tan(x)$ to $(-\pi/2, \pi/2)$ gives the inverse function

$$\arctan\colon \mathbb{R} \to \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

Example 17.2 We have

$$\arcsin(1) = \frac{\pi}{2},$$

because $\sin(\pi/2) = 1$ and $\pi/2$ lies in the prescribed range of arcsin.

Example 17.3 If $y = \arcsin(x)$, then $\sin(y) = x$ and $y \in [-\pi/2, \pi/2]$. Therefore

$$\cos(\arcsin(x)) = \cos(y) = \sqrt{1 - \sin^2(y)} = \sqrt{1 - x^2}.$$

Likewise, if $y = \arctan(x)$, then $\tan(y) = x$ with $y \in (-\pi/2, \pi/2)$, so

$$\sec(\arctan(x)) = \sqrt{1 + \tan^2(y)} = \sqrt{1 + x^2}.$$

The identities involving inverse functions must respect the chosen branch. For example,

$$\sin(\arcsin(x)) = x$$

for every $x \in [-1, 1]$, but $\arcsin(\sin(x))$ is not always equal to x. Indeed,

$$\arcsin(\sin(2\pi)) = \arcsin(0) = 0.$$

Definition 17.4 For $x > 0$, let A_x denote the area bounded by the graph of $y = 1/t$, the t-axis, and the vertical lines $t = 1$ and $t = x$; see Figure 8. Define the **(natural) logarithm function**

$$\log(x) = \begin{cases} A_x, & x \geq 1, \\ -A_x, & 0 < x < 1. \end{cases}$$

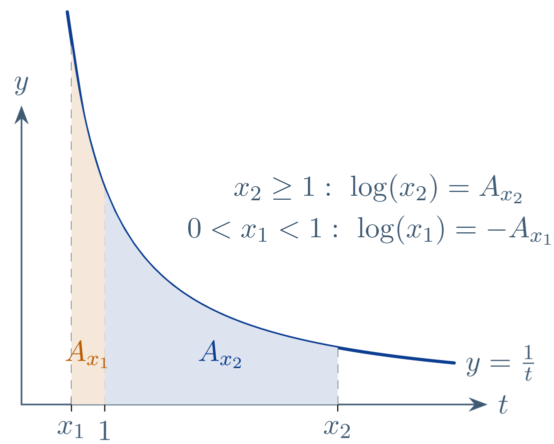

Figure 8

The natural logarithm $\log(x)$ is defined on $(0, \infty)$, satisfies $\log(1) = 0$, is strictly increasing, is negative on $(0, 1)$, and is positive on $(1, \infty)$. Moreover, if $x \geq N$, then

$$A_x \geq A_N \geq \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{N},$$

so

$$\log(x) \to \infty \quad \text{as } x \to \infty \qquad \text{and} \qquad \log(x) \to -\infty \quad \text{as } x \to 0^+.$$

Since log is strictly increasing, it is invertible.

Definition 17.5 The inverse of the natural logarithm is the **exponential function**

$$\exp \colon \mathbb{R} \to (0, \infty).$$

By definition, $\exp(1) = e \approx 2.71$, and therefore $\log(e) = 1$.

Definition 17.6 Let f be a real-valued function and let a be an interior point of the domain of f. We say that f is **differentiable at a** if the limit

$$\lim_{h\to 0}\frac{f(a+h)-f(a)}{h}$$

exists. When it does, that limit is denoted by $f'(a)$ and is called the **derivative** of f at a.

The derivative defines a new function

$$f'\colon \{x\in \operatorname{int}(\operatorname{Dom}(f)) : f \text{ is differentiable at } x\}\to\mathbb{R}.$$

Notation 17.7 Common notation for the derivative includes

$$f'(a),\qquad \left.\frac{\mathrm{d}f}{\mathrm{d}x}\right|_{x=a},\qquad y'(a),\qquad \frac{\mathrm{d}f}{\mathrm{d}x},\qquad \text{and}\qquad \frac{\mathrm{d}}{\mathrm{d}x}f(x).$$

The first three expressions denote derivative values, while the last two denote derivative functions.

Lecture 18 — Monday, October 22, 2012

The derivative represents the instantaneous rate of change of f at the point a, or equivalently the slope of the tangent line to the curve $y=f(x)$ at $x=a$.

The derivative at a can also be written as

$$f'(a)=\lim_{x\to a}\frac{f(x)-f(a)}{x-a},$$

since $x\to a$ is equivalent to writing $x=a+h$ with $h\to 0$. Geometrically, $f'(a)$ is the slope of the tangent line to the graph $y=f(x)$ at $x=a$, and in applications it represents an instantaneous rate of change. If f denotes position as a function of time, then $f'(a)$ is the velocity at time a.

Example 18.1 If $f(x)=c$ is constant, then

$$f'(a)=\lim_{h\to 0}\frac{c-c}{h}=0.$$

Example 18.2 If $f(x) = mx + b$, then

$$f'(a) = \lim_{h\to 0} \frac{m(a+h) + b - (ma+b)}{h} = \lim_{h\to 0} \frac{mh}{h} = m.$$

Thus the tangent line to a nonvertical line is the line itself.

Example 18.3 If $f(x) = x^3$, then

$$\frac{(x+h)^3 - x^3}{h} = \frac{3x^2h + 3xh^2 + h^3}{h} = 3x^2 + 3xh + h^2.$$

Letting $h \to 0$ gives

$$f'(x) = 3x^2.$$

Theorem 18.4 If f is differentiable at a, then f is continuous at a.

Proof. Observe that

$$f(x) - f(a) = \frac{f(x) - f(a)}{x - a}(x - a) \quad \text{for } x \neq a.$$

Since f is differentiable at a, the limit

$$\lim_{x\to a} \frac{f(x) - f(a)}{x - a} = f'(a)$$

exists. Therefore

$$\lim_{x\to a}(f(x) - f(a)) = \left(\lim_{x\to a} \frac{f(x) - f(a)}{x - a}\right)\left(\lim_{x\to a}(x - a)\right) = f'(a) \cdot 0 = 0.$$

Hence

$$\lim_{x\to a} f(x) = f(a),$$

so f is continuous at a. □

Proposition 18.5 The sine function is differentiable on $\mathbb{R}$, and

$$(\sin(x))' = \cos(x).$$

Proof.

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\sin(x) &= \lim_{h\to 0}\frac{\sin(x+h)-\sin(x)}{h}\\ &= \lim_{h\to 0}\frac{\sin(x)\cos(h)+\cos(x)\sin(h)-\sin(x)}{h}\\ &= \lim_{h\to 0}\left(\sin(x)\frac{\cos(h)-1}{h}+\cos(x)\frac{\sin(h)}{h}\right)\\ &= \sin(x)\cdot 0+\cos(x)\cdot 1\\ &= \cos(x).\end{aligned}$$

□

Proposition 18.6 The cosine function is differentiable on $\mathbb{R}$, and

$$(\cos(x))' = -\sin(x).$$

Proof.

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\cos(x) &= \lim_{h\to 0}\frac{\cos(x+h)-\cos(x)}{h}\\ &= \lim_{h\to 0}\frac{\cos(x)\cos(h)-\sin(x)\sin(h)-\cos(x)}{h}\\ &= \lim_{h\to 0}\left(\cos(x)\frac{\cos(h)-1}{h}-\sin(x)\frac{\sin(h)}{h}\right)\\ &= \cos(x)\cdot 0-\sin(x)\cdot 1\\ &= -\sin(x).\end{aligned}$$

□

Corollary 18.7 The functions $\sin(x)$ and $\cos(x)$ are continuous. Consequently, $\tan(x) = (\sin(x))/(\cos(x))$ is continuous on its domain. The inverse trigonometric functions $\arcsin(x)$, $\arccos(x)$, and $\arctan(x)$ are also continuous on their domains.

Proposition 18.8 The natural logarithm is differentiable on $(0, \infty)$, and

$$(\log(x))' = \frac{1}{x}.$$

Proof. Fix $x > 0$. For sufficiently small h, we have $x + h > 0$, and

$$\log(x + h) - \log(x)$$

is the signed area under the curve $y = 1/t$ between $t = x$ and $t = x + h$. Suppose first that $h > 0$. Since $t \mapsto 1/t$ is decreasing on $(0, \infty)$, the area lies between the areas of two rectangles:

$$\frac{h}{x+h} \leq \log(x + h) - \log(x) \leq \frac{h}{x}.$$

Dividing by h gives

$$\frac{1}{x+h} \leq \frac{\log(x + h) - \log(x)}{h} \leq \frac{1}{x}.$$

Now suppose that $h < 0$. Then

$$\log(x + h) - \log(x) = -(\text{area from } x + h \text{ to } x).$$

Again comparing with rectangles yields

$$\frac{-h}{x} \leq -(\log(x + h) - \log(x)) \leq \frac{-h}{x+h}.$$

Since $-h > 0$, dividing by $-h$ and then multiplying by -1 gives

$$\frac{1}{x} \leq \frac{\log(x + h) - \log(x)}{h} \leq \frac{1}{x+h}.$$

In both cases, the squeeze theorem shows that

$$\lim_{h \to 0} \frac{\log(x + h) - \log(x)}{h} = \frac{1}{x}.$$

Therefore $(\log(x))' = 1/x$ for every $x > 0$. $\square$

Corollary 18.9 The logarithm function is continuous on $(0, \infty)$. Since it is continuous, strictly increasing, and has range $\mathbb{R}$, its inverse $\exp(x)$ is continuous on $\mathbb{R}$.

Example 18.10 The function $f(x) = |x|$ is not differentiable at $x = 0$, because

$$\lim_{h \to 0^+} \frac{|h| - 0}{h} = 1 \qquad \text{while} \qquad \lim_{h \to 0^-} \frac{|h| - 0}{h} = -1.$$

Thus the two-sided derivative does not exist at 0. Nevertheless, $|x|$ is continuous everywhere.

Continuity does not imply differentiability. A function may be continuous everywhere and fail to be differentiable at one point, or even at infinitely many points. For example, the **triangular wave function**

$$f(x) = \begin{cases} x - 2k, & 2k \le x \le 2k+1 \\ 2k + 2 - x, & 2k+1 \le x \le 2k+2 \end{cases}, \qquad k \in \mathbb{Z}$$

is non-differentiable at every integer, but is continuous everywhere. See Figure 9.

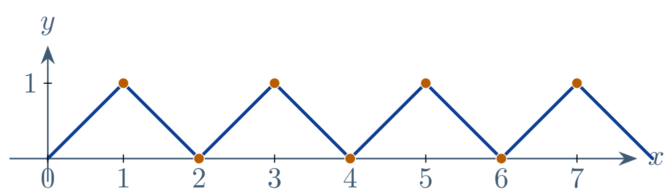

FIGURE 9

Lecture 19 — Wednesday, October 24, 2012

Example 19.1 Define

$$f(x) = \begin{cases} x \sin\left(\dfrac{1}{x}\right), & x \ne 0, \\ 0, & x = 0. \end{cases}$$

Then f is continuous at 0 but not differentiable there.

Proof. Since

$$0 \le \left| x \sin\left(\frac{1}{x}\right) \right| \le |x|,$$

the squeeze theorem gives

$$\lim_{x\to 0} f(x) = 0 = f(0).$$

Thus f is continuous at 0. However,

$$\frac{f(h)-f(0)}{h} = \frac{h\sin(1/h)}{h} = \sin(1/h)$$

for $h \neq 0$, and $\sin(1/h)$ has no limit as $h \to 0$. Therefore f is not differentiable at 0. $\square$

Example 19.2 Define

$$g(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right), & x \neq 0, \\ 0, & x = 0. \end{cases}$$

Then g *is* differentiable at 0, and $g'(0) = 0$.

Proof. For $h \neq 0$,

$$\frac{g(h)-g(0)}{h} = \frac{h^2\sin(1/h)}{h} = h\sin\left(\frac{1}{h}\right).$$

Since

$$\left|h\sin\left(\frac{1}{h}\right)\right| \leq |h|,$$

the squeeze theorem shows that the previous quotient tends to 0 as $h \to 0$. Hence $g'(0) = 0$. $\square$

The derivative of g exists everywhere, but it is not continuous everywhere. This shows that differentiability of f does not imply continuity of f'.

Definition 19.3 If f' is differentiable, then its derivative is denoted by

$$f'' = (f')'$$

and is called the **second derivative** of f. More generally, if $f^{(n-1)}$ is differentiable, then

$$f^{(n)} = \left(f^{(n-1)}\right)'$$

is called the ***n*th derivative** of f.

Example 19.4 If $f(x) = ax + b$, then $f'(x) = a$, $f''(x) = 0$, and in fact

$$f^{(n)}(x) = 0 \quad \text{for every } n \geq 2.$$

Differentiation Rules

Theorem 19.5 (Differentiation rules) Suppose that f and g are differentiable at a, and let $c \in \mathbb{R}$.

Then $f + g$, $f - g$, cf, and fg are differentiable at a, with

$$(f \pm g)'(a) = f'(a) \pm g'(a), \qquad \text{and} \qquad (cf)'(a) = cf'(a),$$

and

$$(fg)'(a) = f'(a)g(a) + f(a)g'(a).$$

This is the **product rule**.

If $g(a) \neq 0$, then $1/g$ and f/g are differentiable at a, and

$$\left(\frac{1}{g}\right)'(a) = \frac{-g'(a)}{(g(a))^2}, \qquad \text{and} \qquad \left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{(g(a))^2},$$

which are the **reciprocal rule** and **quotient rule**, respectively.

Proof. For the sum rule,

$$\begin{aligned}(f+g)'(a) &= \lim_{h\to 0} \frac{(f+g)(a+h) - (f+g)(a)}{h} \\ &= \lim_{h\to 0} \left(\frac{f(a+h) - f(a)}{h} + \frac{g(a+h) - g(a)}{h}\right) \\ &= f'(a) + g'(a).\end{aligned}$$

The proof for $f - g$ is identical. The constant multiple rule follows immediately:

$$(cf)'(a) = \lim_{h\to 0} \frac{cf(a+h) - cf(a)}{h} = c \lim_{h\to 0} \frac{f(a+h) - f(a)}{h} = cf'(a).$$

For the product rule,

$$\begin{aligned}(fg)'(a) &= \lim_{h\to 0}\frac{f(a+h)g(a+h)-f(a)g(a)}{h}\\ &= \lim_{h\to 0}\frac{f(a+h)g(a+h)-f(a)g(a+h)+f(a)g(a+h)-f(a)g(a)}{h}\\ &= \lim_{h\to 0}\left(g(a+h)\frac{f(a+h)-f(a)}{h}+f(a)\frac{g(a+h)-g(a)}{h}\right)\\ &= g(a)f'(a)+f(a)g'(a),\end{aligned}$$

because differentiability implies continuity, so $g(a+h)\to g(a)$. To prove the reciprocal rule, note first that differentiability of g implies continuity of g at a. Since $g(a)\neq 0$, we have $g(a+h)\neq 0$ for all sufficiently small h. Hence

$$\begin{aligned}\left(\frac{1}{g}\right)'(a) &= \lim_{h\to 0}\frac{\frac{1}{g(a+h)}-\frac{1}{g(a)}}{h}\\ &= \lim_{h\to 0}\frac{g(a)-g(a+h)}{h\,g(a+h)g(a)}\\ &= -\frac{1}{(g(a))^2}\lim_{h\to 0}\frac{g(a+h)-g(a)}{h}\\ &= \frac{-g'(a)}{(g(a))^2}.\end{aligned}$$

Finally,

$$\frac{f}{g}=f\cdot\frac{1}{g},$$

so the quotient rule follows from the product rule and the reciprocal rule. □

Corollary 19.6 (Power rule) For each $n\in\mathbb{N}$,

$$\frac{\mathrm{d}}{\mathrm{d}x}x^n = nx^{n-1}.$$

Proof. We argue by induction on n. The case $n=1$ is immediate. Assume the formula holds for $n-1$. Then

$$x^n = x^{n-1}\cdot x,$$

so the product rule and the induction hypothesis give

$$\frac{\mathrm{d}}{\mathrm{d}x}x^n = \frac{\mathrm{d}}{\mathrm{d}x}(x^{n-1}) \cdot x + x^{n-1} \cdot \frac{\mathrm{d}}{\mathrm{d}x}x = (n-1)x^{n-2} \cdot x + x^{n-1} = nx^{n-1}.$$

This completes the proof. $\square$

Theorem 19.7 (Chain rule) Let $f\colon A \to B \subseteq \mathbb{R}$ and let $g\colon B \to \mathbb{R}$. If f is differentiable at $a \in A$ and g is differentiable at $f(a)$, then

$$(g \circ f)'(a) = g'(f(a))\, f'(a).$$

Proof. Define a function Φ on B by

$$\Phi(y) = \begin{cases} \dfrac{g(y) - g(f(a))}{y - f(a)}, & y \neq f(a), \\ g'(f(a)), & y = f(a). \end{cases}$$

Because g is differentiable at $f(a)$, the function Φ is continuous at $f(a)$ (this fact is called **Carathéodory's theorem on differentiability**). For every $x \in A$, we have

$$g(f(x)) - g(f(a)) = \Phi(f(x))(f(x) - f(a)).$$

Therefore, for $x \neq a$,

$$\frac{g(f(x)) - g(f(a))}{x - a} = \Phi(f(x)) \cdot \frac{f(x) - f(a)}{x - a}.$$

Now let $x \to a$. Since f is differentiable at a, it is continuous at a, so $f(x) \to f(a)$. Hence

$$\Phi(f(x)) \to \Phi(f(a)) = g'(f(a)).$$

Also,

$$\frac{f(x) - f(a)}{x - a} \to f'(a).$$

Taking limits on both sides gives

$$(g \circ f)'(a) = g'(f(a))\, f'(a).$$

$\square$

Lecture 20 — Monday, October 29, 2012

Example 20.1 The derivative of $\tan(x)$ is $\sec^2(x)$ wherever $\cos(x) \neq 0$.

Proof. Since $\tan(x) = \frac{\sin(x)}{\cos(x)}$, the quotient rule gives

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}x}\tan(x) &= \frac{\cos(x)\cdot\cos(x) - \sin(x)\cdot(-\sin(x))}{\cos^2(x)} \\ &= \frac{\cos^2(x) + \sin^2(x)}{\cos^2(x)} \\ &= \frac{1}{\cos^2(x)} \\ &= \sec^2(x). \end{aligned}$$

$\square$

Example 20.2 If $y = (3x^2 + 1)^{17}$, then

$$y' = 17(3x^2+1)^{16} \cdot 6x = 102x(3x^2+1)^{16}.$$

Proof. This is an immediate application of the chain rule with outer function $u \mapsto u^{17}$ and inner function $u = 3x^2 + 1$. $\square$

Example 20.3 If $y = \log(\sin(x))$, then $y' = \cot(x)$ at every point where $\sin(x) > 0$.

Proof. Applying the chain rule gives

$$y' = \frac{1}{\sin(x)} \cdot \cos(x) = \cot(x).$$

$\square$

Suppose f is invertible and differentiable at a. Since $f^{-1}(f(x)) = x$, it is tempting to differentiate and conclude immediately that

$$(f^{-1})'(f(a)) \cdot f'(a) = 1,$$

so that $(f^{-1})'(f(a)) = 1/f'(a)$. The difficulty is that this argument already assumes f^{-1} is differentiable at $f(a)$, which is the point that must still be proved. There is also a genuine obstruction when $f'(a) = 0$: for example, the inverse of $f(x) = x^3$ is $x^{1/3}$, whose graph has a vertical tangent at the origin and hence is not differentiable there.

Theorem 20.4 Let $f : (c, d) \to \mathbb{R}$ be continuous and one-to-one. Suppose f is differentiable at $a \in (c, d)$ and $f'(a) \neq 0$. Then f^{-1} is differentiable at $b = f(a)$, and

$$(f^{-1})'(b) = \frac{1}{f'(a)} = \frac{1}{f'(f^{-1}(b))}.$$

This is essentially a one-dimensional, localized version of the inverse function theorem. The more general theorem proves local invertibility rather than assuming that f is one-to-one.

Proof. Let $b = f(a)$. We examine

$$\lim_{h\to 0} \frac{f^{-1}(b+h) - f^{-1}(b)}{h}.$$

Whenever h is small enough that $b + h$ lies in the range of f, write

$$b + h = f(a + k).$$

Then

$$h = f(a + k) - f(a)$$

and

$$f^{-1}(b + h) - f^{-1}(b) = (a + k) - a = k.$$

Therefore

$$\frac{f^{-1}(b+h) - f^{-1}(b)}{h} = \frac{k}{f(a+k) - f(a)}.$$

Choose α, β with $c < \alpha < a < \beta < d$. The restriction of f to $[\alpha, \beta]$ is continuous and one-to-one, so Theorem 16.5 shows that its inverse is continuous. Moreover, $b = f(a)$ is an interior point of the range of this restriction. Thus, once h is sufficiently small, $b + h$ lies in that range and $h \to 0$

implies

$$f^{-1}(b+h) \to f^{-1}(b) = a,$$

so $k \to 0$. It follows that

$$\begin{aligned}\lim_{h\to 0} \frac{f^{-1}(b+h) - f^{-1}(b)}{h} &= \lim_{k\to 0} \frac{k}{f(a+k) - f(a)} \\ &= \frac{1}{\lim_{k\to 0} \frac{f(a+k)-f(a)}{k}} \\ &= \frac{1}{f'(a)}.\end{aligned}$$

Since $f'(a) \neq 0$, this limit exists. Thus f^{-1} is differentiable at b, and

$$(f^{-1})'(b) = \frac{1}{f'(a)} = \frac{1}{f'(f^{-1}(b))}.$$

□

Example 20.5 Let $n \in \mathbb{N}$ and $g(x) = x^{1/n}$. Then

$$g'(x) = \frac{1}{n} x^{1/n-1}$$

for every $x \neq 0$ at which the function is defined.

Proof. The function g is the inverse of $f(x) = x^n$, after restricting the domain of f when necessary to make it one-to-one. Since $f'(a) = na^{n-1}$, Theorem 20.4 gives

$$g'(x) = \frac{1}{f'(g(x))} = \frac{1}{n(g(x))^{n-1}} = \frac{1}{n} x^{1/n-1},$$

provided $g(x) \neq 0$, or equivalently $x \neq 0$. □

Corollary 20.6 Let $p \in \mathbb{Z}$ and $q \in \mathbb{N}$ satisfy $\gcd(|p|, q) = 1$, and define $x^{p/q} = (x^{1/q})^p$ on its natural real domain. Then

$$\frac{\mathrm{d}}{\mathrm{d}x} x^{p/q} = \frac{p}{q} x^{p/q-1}$$

at every nonzero point in the interior of the real domain of $x^{p/q}$.

Proof. Write $x^{p/q} = (x^{1/q})^p$. The integer power rule (together with the reciprocal rule when $p < 0$), the chain rule, and Example 20.5 yield

$$\frac{\mathrm{d}}{\mathrm{d}x}x^{p/q} = p(x^{1/q})^{p-1} \cdot \frac{1}{q}x^{1/q-1} = \frac{p}{q}x^{p/q-1}.$$

□

Example 20.7 The exponential function satisfies

$$\frac{\mathrm{d}}{\mathrm{d}x}e^x = e^x$$

for all $x \in \mathbb{R}$.

Proof. Since e^x is the inverse of $\log(x)$, Theorem 20.4 gives

$$\frac{\mathrm{d}}{\mathrm{d}x}e^x = \frac{1}{(\log(x))'|_{x=e^x}} = \frac{1}{1/e^x} = e^x.$$

□

Example 20.8 For $-1 < x < 1$,

$$\frac{\mathrm{d}}{\mathrm{d}x}\arcsin(x) = \frac{1}{\sqrt{1-x^2}}.$$

Proof. Let $y = \arcsin(x)$. Then $x = \sin(y)$ with $y \in [-\pi/2, \pi/2]$, so $\cos(y) \geq 0$. By Theorem 20.4,

$$\frac{\mathrm{d}}{\mathrm{d}x}\arcsin(x) = \frac{1}{\cos(\arcsin(x))}.$$

Since $\sin(y) = x$, we have

$$\cos(y) = \sqrt{1-\sin^2(y)} = \sqrt{1-x^2}.$$

Therefore

$$\frac{\mathrm{d}}{\mathrm{d}x}\arcsin(x) = \frac{1}{\sqrt{1-x^2}}.$$

□

Exercise 20.9 Show that
$$\frac{\mathrm{d}}{\mathrm{d}x}\arccos(x) = -\frac{1}{\sqrt{1-x^2}}$$
for $-1 < x < 1$.

Example 20.10 For every $x \in \mathbb{R}$,
$$\frac{\mathrm{d}}{\mathrm{d}x}\arctan(x) = \frac{1}{1+x^2}.$$

Proof. The function $\arctan(x)$ is the inverse of $\tan(x)$ on $(-\pi/2, \pi/2)$. Therefore
$$\frac{\mathrm{d}}{\mathrm{d}x}\arctan(x) = \frac{1}{\sec^2(\arctan(x))}.$$
If $y = \arctan(x)$, then $\tan(y) = x$, so $\sec^2(y) = 1 + \tan^2(y) = 1 + x^2$. Hence
$$\frac{\mathrm{d}}{\mathrm{d}x}\arctan(x) = \frac{1}{1+x^2}.$$
□

Lecture 21 — Wednesday, October 31, 2012

Critical Points and Derivative Applications

The derivative formulas
$$\frac{\mathrm{d}}{\mathrm{d}x}e^x = e^x, \qquad \frac{\mathrm{d}}{\mathrm{d}x}\arcsin(x) = \frac{1}{\sqrt{1-x^2}}, \qquad \text{and} \qquad \frac{\mathrm{d}}{\mathrm{d}x}\arctan(x) = \frac{1}{1+x^2}$$
will be used repeatedly in what follows.

Example 21.1 Differentiate the following functions:
$$y = e^{\cos(3x)}, \qquad y = (\arctan(x)+1)^{1/5}, \qquad \text{and} \qquad y = \left(\arcsin\left(\sqrt{x}\right)\right)\sin^2(x).$$

Solution. For $y = e^{\cos(3x)}$, the chain rule gives

$$y' = e^{\cos(3x)}(-3\sin(3x)).$$

For $y = (\arctan(x) + 1)^{1/5}$, we again apply the chain rule; at points where $\arctan(x) + 1 \neq 0$,

$$y' = \frac{1}{5}(\arctan(x) + 1)^{-4/5} \cdot \frac{1}{1+x^2}.$$

Finally, if $y = (\arcsin(\sqrt{x}))\sin^2(x)$, then for $0 < x < 1$, the product rule and chain rule give

$$y' = \frac{1}{\sqrt{1-(\sqrt{x})^2}} \cdot \frac{1}{2\sqrt{x}}\sin^2(x) + (\arcsin\left(\sqrt{x}\right))(2\sin(x)\cos(x)).$$

□

Definition 21.2 A point $x \in \mathrm{Dom}(f)$ is a **local maximum** of f if there exists $\delta > 0$ such that

$$f(y) \leq f(x)$$

whenever $y \in (x-\delta, x+\delta) \cap \mathrm{Dom}(f)$. A **global maximum** of f is a point $x \in \mathrm{Dom}(f)$ such that

$$f(y) \leq f(x)$$

for every $y \in \mathrm{Dom}(f)$. **Local** and **global minima** are defined analogously. Every global extremum is automatically a local extremum, but the converse need not hold.

Definition 21.3 A point $x \in \mathrm{Dom}(f)$ such that $f'(x) = 0$ is called a **critical point** of f.

Theorem 21.4 If f has a local maximum or local minimum at $x \in (a, b) \subseteq \mathrm{Dom}(f)$ and f is differentiable at x, then $f'(x) = 0$.

Proof. Assume first that x is a local maximum of f. Then there exists $\delta > 0$ such that

$$f(x+h) \leq f(x)$$

whenever $|h| < \delta$ and $x + h \in \mathrm{Dom}(f)$. For $0 < h < \delta$, dividing by the positive number h gives

$$\frac{f(x+h) - f(x)}{h} \leq 0.$$

For $-\delta < h < 0$, dividing by the negative number h reverses the inequality and gives

$$\frac{f(x+h)-f(x)}{h} \geq 0.$$

Taking one-sided limits, we obtain

$$\lim_{h\to 0^+} \frac{f(x+h)-f(x)}{h} \leq 0 \qquad \text{and} \qquad \lim_{h\to 0^-} \frac{f(x+h)-f(x)}{h} \geq 0.$$

If f is differentiable at x, these one-sided limits are equal, so both must be 0. Hence $f'(x) = 0$. The proof for a local minimum is identical, with all inequalities reversed. $\square$

If $f : [a, b] \to \mathbb{R}$ is continuous, then the extreme value theorem guarantees that f has both a maximum and a minimum on $[a, b]$. To find them in practice, we check all candidate points: the endpoints a and b, all interior critical points, and all interior singular points where f is not differentiable.

Example 21.5 Let

$$f(x) = x - x^{2/3}, \quad -1 \leq x \leq 8.$$

Find the global maximum and minimum values of f.

Solution. The function f is continuous on $[-1, 8]$, so it has a global maximum and minimum on this interval. The candidates are the endpoints, the singular point $x = 0$, and any interior critical points.

For $x \neq 0$,

$$f'(x) = 1 - \frac{2}{3}x^{-1/3}.$$

Setting $f'(x) = 0$ gives

$$1 = \frac{2}{3}x^{-1/3},$$

so

$$x = \left(\frac{2}{3}\right)^3 = \frac{8}{27}.$$

We now evaluate f at all candidate points:

$$f(-1) = -2, \qquad f(0) = 0, \qquad f\left(\frac{8}{27}\right) = -\frac{4}{27}, \qquad \text{and} \qquad f(8) = 4.$$

Therefore the global minimum value is -2, attained at $x = -1$, and the global maximum value is 4, attained at $x = 8$. □

5. The Mean Value Theorem and Its Consequences

Rolle's Theorem and the Mean Value Theorem

The next major result is the key special case needed for the proof of the mean value theorem.

Theorem 21.6 (Rolle's theorem) If f is continuous on $[a, b]$, differentiable on (a, b), and satisfies $f(a) = f(b)$, then there exists $c \in (a, b)$ such that $f'(c) = 0$.

Figure 10 shows why none of the three hypotheses in Rolle's theorem can be dropped.

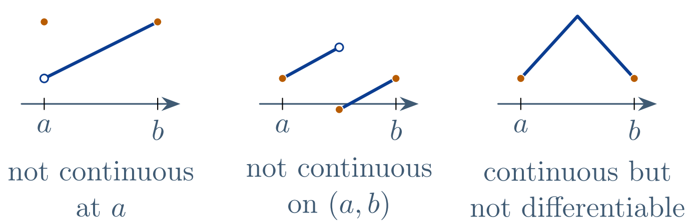

FIGURE 10

Proof. By the extreme value theorem, f attains a global maximum and a global minimum on $[a, b]$. If either extremum occurs at an interior point $c \in (a, b)$, then Theorem 21.4 gives $f'(c) = 0$, and we are done. It remains to consider the case in which both the maximum and minimum occur only at the endpoints. Since $f(a) = f(b)$, the endpoint values are equal. Therefore the maximum and minimum values are equal as well, so f is constant on $[a, b]$. Hence $f'(c) = 0$ for every $c \in (a, b)$. □

Lecture 22 — Friday, November 02, 2012

Theorem 22.1 (Mean value theorem) Let $a < b$, and let f be continuous on $[a, b]$ and differentiable on (a, b). Then there exists $c \in (a, b)$ such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Observe that Rolle's theorem is the special case of the mean value theorem in which $f(a) = f(b)$.

Proof. Let $L(x)$ be the equation of the secant line joining $(a, f(a))$ and $(b, f(b))$. Then

$$L'(x) = \frac{f(b) - f(a)}{b - a}$$

for every x. Define

$$g(x) = f(x) - L(x).$$

Since f is continuous on $[a, b]$ and differentiable on (a, b), the same is true of g. Also,

$$g(a) = g(b) = 0,$$

because L passes through the endpoints of the graph of f. By Rolle's theorem, there exists $c \in (a, b)$ such that $g'(c) = 0$. Hence

$$0 = g'(c) = f'(c) - L'(c),$$

so

$$f'(c) = L'(c) = \frac{f(b) - f(a)}{b - a}.$$

□

Corollary 22.2 If $f'(x) = 0$ for all x in an interval I, then f is constant on I.

Proof. Let $a, b \in I$ with $a < b$. Since differentiability implies continuity, f is continuous on $[a, b]$

and differentiable on (a, b). By the mean value theorem, there exists $c \in (a, b)$ such that

$$\frac{f(b) - f(a)}{b - a} = f'(c) = 0.$$

Therefore $f(b) = f(a)$. Since a and b were arbitrary, f is constant on I. $\square$

Corollary 22.3 If $f'(x) = g'(x)$ for all $x \in I$, then $f - g$ is constant on I.

Proof. Let $h(x) = f(x) - g(x)$. Then $h'(x) = 0$ for every $x \in I$, so the previous corollary shows that h is constant on I. $\square$

The interval hypothesis in the previous corollary is necessary. Consider

$$f(x) = \begin{cases} 1, & x > 0, \\ -1, & x < 0. \end{cases}$$

Then $f'(x) = 0$ at every point of its domain, but the domain $(-\infty, 0) \cup (0, \infty)$ is not an interval, and f is not constant.

Proposition 22.4 For $a, b, x > 0$ and $r \in \mathbb{Q}$, the following identities hold:

$$\log(ab) = \log(a) + \log(b), \qquad \log\left(\frac{a}{b}\right) = \log(a) - \log(b), \qquad \text{and} \qquad \log(x^r) = r\log(x).$$

Proof. Fix $b > 0$, and define $f(x) = \log(bx) - \log(x)$ for $x > 0$. Then

$$f'(x) = \frac{b}{bx} - \frac{1}{x} = 0$$

for all $x > 0$, so f is constant. Evaluating at $x = 1$ gives

$$f(x) = f(1) = \log(b),$$

hence

$$\log(ab) - \log(a) = \log(b),$$

which proves $\log(ab) = \log(a) + \log(b)$. Applying this with a/b and b, we get

$$\log(a) = \log\left(\frac{a}{b}\right) + \log(b),$$

so $\log(a/b) = \log(a) - \log(b)$. Now define $g(x) = \log(x^r) - r\log(x)$ for $x > 0$. Then

$$g'(x) = \frac{rx^{r-1}}{x^r} - \frac{r}{x} = 0,$$

so g is constant. Since $g(1) = 0$, we conclude that $g(x) = 0$ for all $x > 0$, and therefore

$$\log(x^r) = r\log(x).$$

□

Corollary 22.5 If $r \in \mathbb{Q}$, then

$$e^{rx} = (e^x)^r.$$

Proof. Let $y = e^{rx}$ and $z = (e^x)^r$. Then

$$\log(y) = rx$$

and, by Proposition 22.4,

$$\log(z) = \log((e^x)^r) = r\log(e^x) = rx.$$

Therefore $\log(y) = \log(z)$, so $y = z$. □

Definition 22.6 For $a > 0$ and $x \in \mathbb{R}$, define

$$a^x = e^{x\log(a)}.$$

Then $\log(a^x) = x\log(a)$ for all $x \in \mathbb{R}$.

Theorem 22.7 If $r \in \mathbb{R}$ and $x > 0$, then

$$\frac{\mathrm{d}}{\mathrm{d}x}x^r = rx^{r-1}.$$

Proof. Write

$$x^r = e^{r\log(x)}.$$

Differentiating gives

$$\frac{\mathrm{d}}{\mathrm{d}x}x^r = e^{r\log(x)} \cdot \frac{r}{x} = x^r \cdot \frac{r}{x} = rx^{r-1}.$$

□

Example 22.8 If $y = 2^x$, then

$$y' = 2^x \log(2).$$

Proof. Since $2^x = e^{x\log(2)}$, the chain rule gives

$$y' = e^{x\log(2)} \log(2) = 2^x \log(2).$$

□

Theorem 22.9 (Logarithmic differentiation) Suppose f and g are differentiable and $f(x) > 0$ throughout their common domain. Then

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(f(x)^{g(x)}\right) = f(x)^{g(x)} \left(g'(x)\log(f(x)) + g(x)\frac{f'(x)}{f(x)}\right).$$

Proof. Write

$$f(x)^{g(x)} = e^{g(x)\log(f(x))}.$$

Differentiating with the chain rule gives

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(f(x)^{g(x)}\right) = e^{g(x)\log(f(x))} \left(g'(x)\log(f(x)) + g(x)\frac{f'(x)}{f(x)}\right),$$

which is the stated formula. □

Example 22.10 If $y = x^{\tan(x)}$, then

$$y' = x^{\tan(x)} \left(\sec^2(x)\log(x) + \frac{\tan(x)}{x}\right)$$

for $x > 0$.

Proof. Apply logarithmic differentiation with $f(x) = x$ and $g(x) = \tan(x)$. $\square$

Example 22.11 If $y = (\arcsin(x))^{\sqrt{x}}$, then

$$y' = (\arcsin(x))^{\sqrt{x}} \left(\frac{\log(\arcsin(x))}{2\sqrt{x}} + \frac{\sqrt{x}}{\arcsin(x)\sqrt{1-x^2}} \right)$$

where the expression is defined.

Proof. Apply logarithmic differentiation with $f(x) = \arcsin(x)$ and $g(x) = \sqrt{x}$. $\square$

Lecture 23 — Monday, November 05, 2012

Consequences of the Mean Value Theorem

Corollary 23.1

$$\arcsin(x) + \arccos(x) = \frac{\pi}{2}$$

for all $x \in [-1, 1]$.

Proof. Let

$$f(x) = \arcsin(x) + \arccos(x).$$

Then, for $-1 < x < 1$,

$$f'(x) = \frac{1}{\sqrt{1-x^2}} - \frac{1}{\sqrt{1-x^2}} = 0.$$

Hence f is constant on $(-1, 1)$ by Corollary 22.2, and therefore on $[-1, 1]$ by continuity. Since

$$f(1) = \frac{\pi}{2} + 0 = \frac{\pi}{2},$$

the constant value must be $\pi/2$. □

Corollary 23.2 Suppose f is continuous on $[a, b]$, differentiable on (a, b), and $f'(x) \geq M$ for all $x \in (a, b)$. Then
$$f(b) - f(a) \geq M(b - a).$$

Proof. By the mean value theorem, there exists $c \in (a, b)$ such that

$$\frac{f(b) - f(a)}{b - a} = f'(c).$$

Since $f'(c) \geq M$, multiplying by the positive number $b - a$ gives

$$f(b) - f(a) \geq M(b - a).$$

□

Corollary 23.3
$$|\sin(x) - \sin(y)| \leq |x - y|$$

for all real numbers x and y.

Proof. If $x = y$, the inequality is trivial, so assume $x \neq y$. Applying the mean value theorem to $f(\theta) = \sin(\theta)$ on the interval with endpoints x and y, we obtain some c between x and y such that

$$\frac{\sin(x) - \sin(y)}{x - y} = \cos(c).$$

Taking absolute values gives
$$\frac{|\sin(x) - \sin(y)|}{|x - y|} = |\cos(c)| \leq 1,$$
and therefore $|\sin(x) - \sin(y)| \leq |x - y|$. □

Theorem 23.4 (Increasing function theorem) If $f'(x) \geq 0$ for all $x \in (a, b)$, and f is continuous on $[a, b]$, then f is increasing on $[a, b]$.

Proof. Let
$$a \leq x < y \leq b.$$
By the mean value theorem, there exists $c \in (x, y)$ such that
$$\frac{f(y) - f(x)}{y - x} = f'(c) \geq 0.$$
Since $y - x > 0$, it follows that $f(y) \geq f(x)$. $\square$

The following is a partial converse:

Proposition 23.5 If f is differentiable on (a, b) and increasing on (a, b), then $f'(x) \geq 0$ for all $x \in (a, b)$.

Proof. Fix $x \in (a, b)$. For $h > 0$ small enough, increasingness gives $f(x + h) \geq f(x)$, so
$$\frac{f(x + h) - f(x)}{h} \geq 0.$$
For $h < 0$ small enough, we still have $f(x + h) \leq f(x)$, and dividing by the negative number h again yields
$$\frac{f(x + h) - f(x)}{h} \geq 0.$$
In either case, passing to the limit as $h \to 0$ from the relevant direction shows that $f'(x) \geq 0$. $\square$

Note that Proposition 23.5 cannot be strengthened to require $f'(x) > 0$. For example, the function $f(x) = x^3$ is strictly increasing on $\mathbb{R}$, but
$$f'(0) = 0.$$
Thus strict increase does not imply that the derivative is strictly positive at every point.

Example 23.6
$$f(x) = x^3 - 12x + 1.$$
Analyze the increasing and decreasing intervals of f.

Proof. We compute

$$f'(x) = 3x^2 - 12 = 3(x-2)(x+2).$$

The critical points are $x = -2$ and $x = 2$. The sign chart for f' is given in Figure 11. Therefore f

FIGURE 11

is increasing on $(-\infty, -2)$ and $(2, \infty)$, and decreasing on $(-2, 2)$. In particular, $x = -2$ is a local maximum and $x = 2$ is a local minimum. □

Example 23.7 Prove that

$$\log(x) \leq x - 1$$

for all $x \geq 1$.

Solution. Let

$$f(x) = \log(x) - (x - 1).$$

Then

$$f'(x) = \frac{1}{x} - 1 \leq 0$$

for all $x \geq 1$. Hence f is decreasing on $[1, \infty)$, so

$$f(x) \leq f(1) = 0$$

for all $x \geq 1$. Therefore $\log(x) \leq x - 1$. □

Theorem 23.8 (First derivative test) Assume f is continuous on $[a, b]$, and let $c \in (a, b)$ be either a critical point or a singular point.

1. If $f' \geq 0$ on (a, c) and $f' \leq 0$ on (c, b), then c is a local maximum.
2. If $f' \leq 0$ on (a, c) and $f' \geq 0$ on (c, b), then c is a local minimum.
3. If $f' > 0$ on both sides of c, or if $f' < 0$ on both sides, then c is neither a local maximum nor a local minimum.

Proof. For (1), the increasing function theorem shows that f is increasing on (a, c), so $f(x) \le f(c)$ for $x < c$ sufficiently close to c. Likewise, f is decreasing on (c, b), so $f(x) \le f(c)$ for $x > c$ sufficiently close to c. Hence c is a local maximum.

Statement (2) is proved in the same way, with the inequalities reversed.

For (3), suppose first that $f' > 0$ on both sides of c. The mean value theorem shows that f is strictly increasing on each side. Continuity at c then shows that values immediately to the left are strictly less than $f(c)$, while values immediately to the right are strictly greater than $f(c)$. Thus c is not a local extremum. The case $f' < 0$ on both sides is analogous. ☐

Example 23.9 Maximize $f(x) = xe^{-x}$ on $[0, \infty)$.

Proof. We have

$$f'(x) = e^{-x} - xe^{-x} = e^{-x}(1 - x).$$

There is a critical point at $x = 1$ and an endpoint at $x = 0$. The sign chart for f' is given in Figure 12.

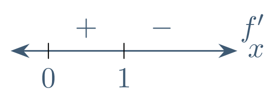

FIGURE 12

Thus f increases on $[0, 1]$ and decreases on $[1, \infty)$, so $f(1) = e^{-1}$ is the global maximum. Also, $f(0) = 0$ and $f(x) \ge 0$ for all $x \ge 0$, so 0 is the global minimum value. ☐

Example 23.10

$$g(x) = \frac{x^2 - 2x - 2}{x - 1}.$$

Analyze the graph qualitatively.

Proof. The domain is $x \neq 1$. Since

$$\lim_{x \to 1^-} g(x) = +\infty \qquad \text{and} \qquad \lim_{x \to 1^+} g(x) = -\infty,$$

the line $x = 1$ is a vertical asymptote. Differentiating gives

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

The critical points are $x = 0$ and $x = 2$, while $x = 1$ is not in the domain. The sign chart for g' is given in Figure 13.

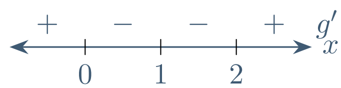

FIGURE 13

Thus g is increasing on $(-\infty, 0)$ and $(2, \infty)$, and decreasing on $(0, 1)$ and $(1, 2)$. Rewriting

$$g(x) = \frac{(x-1)^2 - 3}{x-1} = x - 1 - \frac{3}{x-1}$$

shows that $y = x - 1$ is an oblique asymptote. □

For rational functions $f(x) = p(x)/q(x)$, a few quick asymptote tests are useful:

1. If $q(x_0) = 0$ and $p(x_0) \neq 0$, then $x = x_0$ is a vertical asymptote.
2. If $\deg p < \deg q$, then $y = 0$ is a horizontal asymptote.
3. If $\deg p = \deg q$, then the horizontal asymptote is the ratio of the leading coefficients.
4. If $\deg p = \deg q + 1$, then there is an oblique asymptote, found by polynomial long division.

Lecture 24 — Wednesday, November 07, 2012

Concavity and Derivative Tests

Definition 24.1 A differentiable function f is **concave up** on an interval I if f' is increasing on I. It is **concave down** on I if f' is decreasing on I.

These terms are used in introductory calculus because they align with visual intuition and the sign of the second derivative. In higher math, the terms **convex** and **concave** are used instead for concave up and concave down, respectively; these are more formal, global definitions based on inequalities, which are conceptually heavier and less tied to local derivative behaviour.

Definition 24.2 We say that c is an **inflection point** of f if the concavity of f changes at c. Thus, for some $\delta > 0$, the function is concave up on $(c-\delta, c)$ and concave down on $(c, c+\delta)$, or vice versa.

Example 24.3 $y = x^3$ has an inflection point at $c = 0$.

Proposition 24.4 If $f''(x) > 0$ on an interval I, then f is concave up on I. If $f''(x) < 0$ on I, then f is concave down on I. If f has an inflection point at c and $f''(c)$ exists, then $f''(c) = 0$.

Proof. If $f''(x) > 0$ on I, then f' is increasing on I by the increasing function theorem, so f is concave up. The argument for $f''(x) < 0$ is the same, with the inequalities reversed. Now suppose f has an inflection point at c, and assume without loss of generality that f is concave up on $(c-\delta, c)$ and concave down on $(c, c+\delta)$. Then f' is increasing on $(c-\delta, c)$ and decreasing on $(c, c+\delta)$. Since the existence of $f''(c)$ makes f' continuous at c, for small $h > 0$,

$$\frac{f'(c+h) - f'(c)}{h} \leq 0,$$

while for small $h < 0$,

$$\frac{f'(c+h) - f'(c)}{h} \geq 0.$$

If $f''(c)$ exists, the one-sided limits must agree, so both must be 0. Hence $f''(c) = 0$. $\square$

The converse is false. For example, $y = x^4$ satisfies $y''(0) = 0$, but the graph is concave up everywhere, so 0 is not an inflection point.

Example 24.5 If $f'(x_0) = 0 = f''(x_0)$ and x_0 is not a local minimum or maximum, is it always true that x_0 is an inflection point?

Solution. No! Consider

$$f(x) = \begin{cases} x^4 \sin\left(\frac{1}{x}\right), & x \neq 0, \\ 0, & x = 0. \end{cases}$$

Since f takes both positive and negative values arbitrarily close to 0, the point 0 is neither a local maximum nor a local minimum. Also,

$$f'(x) = 4x^3 \sin\left(\frac{1}{x}\right) - x^2 \cos\left(\frac{1}{x}\right)$$

for $x \neq 0$, so $f'(0) = 0$. Moreover,

$$f''(0) = \lim_{h\to 0} \frac{f'(h) - f'(0)}{h} = \lim_{h\to 0} \left(4h^2 \sin\left(\frac{1}{h}\right) - h\cos\left(\frac{1}{h}\right)\right) = 0.$$

For $x \neq 0$, differentiating again gives

$$f''(x) = \sin\left(\frac{1}{x}\right)(12x^2 - 1) - 6x\cos\left(\frac{1}{x}\right).$$

Now fix any $\delta > 0$. Choose $\delta_0 \leq \delta$ so small that

$$\left|12x^2 \sin\left(\frac{1}{x}\right)\right| \leq \frac{1}{8} \qquad \text{and} \qquad \left|6x\cos\left(\frac{1}{x}\right)\right| \leq \frac{1}{8}$$

whenever $0 < |x| \leq \delta_0$. Next choose $x_1, x_2 \in (0, \delta_0)$ such that

$$\sin\left(\frac{1}{x_1}\right) = 1 \qquad \text{and} \qquad \sin\left(\frac{1}{x_2}\right) = -1.$$

By continuity of $\sin(1/x)$ at x_1 and x_2, there are small intervals I_1 and I_2 around these points such that

$$\sin\left(\frac{1}{x}\right) \geq \frac{1}{2} \text{ on } I_1 \qquad \text{and} \qquad \sin\left(\frac{1}{x}\right) \leq -\frac{1}{2} \text{ on } I_2.$$

For $x \in I_1$, the bounds above imply

$$f''(x) \leq \frac{1}{8} + \frac{1}{8} - \frac{1}{2} = -\frac{1}{4} < 0,$$

so f is concave down on I_1. For $x \in I_2$, they imply

$$f''(x) \geq -\frac{1}{8} - \frac{1}{8} + \frac{1}{2} = \frac{1}{4} > 0,$$

so f is concave up on I_2. Thus every interval $(0, \delta)$ contains subintervals on which f is concave up and others on which it is concave down. The same is true on the left of 0. Therefore 0 is not an inflection point, even though $f'(0) = f''(0) = 0$ and 0 is not a local extremum. □

This example shows that geometric intuition about stationary points is not enough; we still need genuine sign or monotonicity information.

Lecture 25 — Friday, November 09, 2012

Theorem 25.1 (Second derivative test) Suppose $f'(c) = 0$. If $f''(c) > 0$, then c is a local minimum. If $f''(c) < 0$, then c is a local maximum. If $f''(c) = 0$, no conclusion follows in general.

Proof. Assume $f''(c) > 0$. Since $f'(c) = 0$,

$$f''(c) = \lim_{h \to 0} \frac{f'(c+h) - f'(c)}{h} = \lim_{h \to 0} \frac{f'(c+h)}{h}.$$

Because this limit is positive, we have

$$\frac{f'(c+h)}{h} > 0$$

for all sufficiently small nonzero h. Thus $f'(c+h) > 0$ when $h > 0$, and $f'(c+h) < 0$ when $h < 0$. By the first derivative test, c is a local minimum. The proof when $f''(c) < 0$ is identical, with all inequalities reversed. □

Example 25.2

$$y = x(\log(x))^2, \quad x > 0.$$

Analyze the graph of y.

Proof. We compute

$$y' = (\log(x))^2 + 2\log(x) = (\log(x))(\log(x) + 2)$$

and
$$y'' = \frac{2\log(x)}{x} + \frac{2}{x} = \frac{2(\log(x)+1)}{x}.$$
The critical points satisfy $y' = 0$, so
$$\log(x) = 0 \qquad \text{or} \qquad \log(x) = -2,$$
hence $x = 1$ and $x = e^{-2}$. The second derivative vanishes when $\log(x) + 1 = 0$, so $x = e^{-1}$. From the sign of y', the function is increasing on $(0, e^{-2})$ and $(1, \infty)$, and decreasing on $(e^{-2}, 1)$. Thus $x = e^{-2}$ is a local maximum and $x = 1$ is a local minimum. From the sign of y'', the graph is concave down on $(0, e^{-1})$ and concave up on (e^{-1}, ∞), so $x = e^{-1}$ is an inflection point. Evaluating at these points gives
$$y(e^{-2}) = 4e^{-2}, \qquad y(e^{-1}) = e^{-1}, \qquad \text{and} \qquad y(1) = 0.$$
Since $y(1) = 0$ and $y(x) \geq 0$ for all $x > 0$, the point $x = 1$ is the global minimum. Together with
$$\lim_{x \to 0^+} x(\log(x))^2 = 0,$$
this information determines the qualitative graph. $\square$

6. L'Hôpital's Rule

Motivation and the Cauchy Mean Value Theorem

Definition 25.3 We write $\lim_{x\to a} f(x) = \infty$ if for every $N > 0$ there exists $\delta > 0$ such that
$$0 < |x - a| < \delta \qquad \text{implies that} \qquad f(x) > N.$$

L'Hôpital's rule predicts that, for indeterminate forms of type $0/0$ or ∞/∞, we should expect
$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}$$
when the relevant hypotheses hold. Why might this be true? If $f(a) = g(a) = 0$, both derivatives exist at a, and $g'(a) \neq 0$, then the difference quotients give
$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{h\to 0} \frac{f(a+h) - f(a)}{g(a+h) - g(a)} = \lim_{h \to 0} \frac{f(a+h)-f(a)}{h} \cdot \frac{h}{g(a+h) - g(a)} = \frac{f'(a)}{g'(a)}.$$

This special argument is too narrow for L'Hôpital's rule: the derivatives at a need not even exist. Applying the ordinary mean value theorem on the interval from a to $a+h$ instead gives

$$\frac{f(a+h)-f(a)}{h} = f'(c_h) \qquad \text{and} \qquad \frac{g(a+h)-g(a)}{h} = g'(d_h)$$

for some generally different points c_h and d_h between a and $a+h$. Although both points tend to a, this does not express the ratio using f' and g' at the same point. The next theorem supplies exactly that missing synchronization.

Theorem 25.4 (Cauchy mean value theorem) If $a < b$, and f and g are continuous on $[a, b]$ and differentiable on (a, b), then there exists $c \in (a, b)$ such that

$$(f(b) - f(a))g'(c) = (g(b) - g(a))f'(c).$$

Proof. Define

$$h(x) = f(x)(g(b) - g(a)) - g(x)(f(b) - f(a)).$$

Then h is continuous on $[a, b]$ and differentiable on (a, b). Also,

$$h(a) = f(a)g(b) - g(a)f(b) = h(b).$$

By Rolle's theorem, there exists $c \in (a, b)$ such that $h'(c) = 0$. Differentiating gives

$$0 = h'(c) = f'(c)(g(b) - g(a)) - g'(c)(f(b) - f(a)),$$

which is the desired conclusion. □

Statement and Proof of L'Hôpital's rule

Lecture 26 — Monday, November 12, 2012

Theorem 26.1 (L'Hôpital's rule) Assume f and g are differentiable on

$$I = [a - \delta^*, a + \delta^*]$$

except possibly at a. Suppose

$$\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0 \qquad \text{or} \qquad \lim_{x\to a} f(x) = \lim_{x\to a} g(x) = \infty.$$

Assume also that $g(x) \neq 0$ and $g'(x) \neq 0$ for all $x \in I \setminus \{a\}$. If

$$\lim_{x\to a} \frac{f'(x)}{g'(x)} = L \in \mathbb{R},$$

then

$$\lim_{x\to a} \frac{f(x)}{g(x)} = L.$$

Proof. *Case 1:* Assume $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0$.

Redefine $f(a) = g(a) = 0$. This does not change the derivatives away from a, nor does it change either limit. We will show that

$$\lim_{x\to a} \frac{f(x)}{g(x)} = L.$$

Let $\varepsilon > 0$. Since $\lim_{x\to a} f'(x)/g'(x) = L$, there exists $\delta_1 > 0$ such that

$$0 < |z - a| < \delta_1 \implies \left| \frac{f'(z)}{g'(z)} - L \right| < \varepsilon.$$

Set $\delta = \min\{\delta_1, \delta^*\}$, and let $0 < |x - a| < \delta$. The functions f and g are continuous on the interval with endpoints a and x, and differentiable in its interior. By the Cauchy mean value theorem, there exists a point c between a and x such that

$$\frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f'(c)}{g'(c)}.$$

Since $f(a) = g(a) = 0$, this becomes

$$\frac{f(x)}{g(x)} = \frac{f'(c)}{g'(c)}.$$

Because c lies between a and x, we have

$$0 < |c - a| < |x - a| < \delta \leq \delta_1.$$

Therefore

$$\left| \frac{f(x)}{g(x)} - L \right| = \left| \frac{f'(c)}{g'(c)} - L \right| < \varepsilon.$$

This proves the claim in the 0/0 case.

Case 2: Assume $\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = \infty$.

It is enough to prove the right-hand limit; the left-hand proof is identical. Let $\varepsilon > 0$. Since $f'(x)/g'(x) \to L$ as $x \to a^+$, there exists $\delta_1 > 0$ such that

$$0 < z - a < \delta_1 \implies \left|\frac{f'(z)}{g'(z)} - L\right| < \varepsilon.$$

Set $\delta_2 = \min\{\delta_1, \delta^*\}$, and choose

$$y = a + \frac{\delta_2}{2}.$$

If $x \in (a, y)$, then f and g are continuous on $[x, y]$ and differentiable on (x, y). By the Cauchy mean value theorem, there exists $c \in (x, y)$ such that

$$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(c)}{g'(c)}.$$

Since $c \in (x, y) \subset (a, a + \delta_2)$, we have

$$\left|\frac{f'(c)}{g'(c)} - L\right| < \varepsilon.$$

Hence

$$L - \varepsilon < \frac{f(x) - f(y)}{g(x) - g(y)} < L + \varepsilon.$$

Now $g(x) \to \infty$ as $x \to a^+$, so for x sufficiently close to a, the quantity $g(x)$ is positive and much larger than $g(y)$. Dividing the previous inequality by $g(x)$, we obtain

$$(L - \varepsilon)\left(1 - \frac{g(y)}{g(x)}\right) < \frac{f(x)}{g(x)} - \frac{f(y)}{g(x)} < (L + \varepsilon)\left(1 - \frac{g(y)}{g(x)}\right).$$

As $x \to a^+$, both $g(y)/g(x)$ and $f(y)/g(x)$ tend to 0. In particular, we may take x sufficiently close to a that

$$\left|\frac{f(y)}{g(x)}\right| < \varepsilon \qquad \text{and} \qquad \left|\frac{g(y)}{g(x)}\right| < \frac{\varepsilon}{|L| + \varepsilon + 1}.$$

The preceding two-sided bound then gives

$$L - 3\varepsilon < \frac{f(x)}{g(x)} < L + 3\varepsilon.$$

This proves that

$$\lim_{x \to a^+} \frac{f(x)}{g(x)} = L.$$

The left-hand limit is handled in the same way. □

Lecture 27 — Wednesday, November 14, 2012

Applications and Variants

Exercise 27.1

Prove the following useful variants of L'Hôpital's rule:

1. The rule applies equally to one-sided limits $x \to a^+$ and $x \to a^-$.
2. It is enough to assume $g(x) \to \pm\infty$. In particular, if $|f(x)| \leq M$ near a while $|g(x)| \to \infty$, then
$$\left|\frac{f(x)}{g(x)}\right| \leq \frac{M}{|g(x)|} \to 0,$$
so $f(x)/g(x) \to 0$ without any need for L'Hôpital's rule.
3. The same conclusions hold for limits as $x \to \pm\infty$. We reduce these to finite limits by substituting $x = 1/y$.
4. The conclusion also remains valid when the derivative quotient tends to $\pm\infty$.

L'Hôpital's rule should not be used when the hypotheses fail. For example,

$$\lim_{x \to 1} \frac{x-1}{x} = 0,$$

but

$$\lim_{x \to 1} \frac{(x-1)'}{x'} = 1.$$

The expression $(x-1)/x$ is not of indeterminate form at $x = 1$, so the rule does not apply.

The converse of L'Hôpital's rule also fails: if $\lim_{x \to a} f'(x)/g'(x)$ does not exist, it does not follow that $\lim_{x\to a} f(x)/g(x)$ fails to exist. For example,

$$\lim_{x\to\infty} \frac{x+\sin(x)}{x} = 1,$$

but

$$\lim_{x\to\infty} \frac{(x+\sin(x))'}{x'} = \lim_{x\to\infty} (1+\cos(x))$$

does not exist.

Example 27.2

$$\lim_{x\to 0} \frac{\tan(x)-x}{\sin^3(x)} = \frac{1}{3}.$$

Proof. Applying L'Hôpital's rule gives

$$\begin{aligned}\lim_{x\to 0} \frac{\tan(x)-x}{\sin^3(x)} &= \lim_{x\to 0} \frac{\sec^2(x)-1}{3\sin^2(x)\cos(x)} \\ &= \lim_{x\to 0} \frac{\tan^2(x)}{3\sin^2(x)\cos(x)} \\ &= \lim_{x\to 0} \frac{1}{3\cos^3(x)} \\ &= \frac{1}{3}.\end{aligned}$$

□

Example 27.3

$$\lim_{x\to\infty} x^2 e^{-x} = 0.$$

Proof. Rewrite the expression as x^2/e^x. Two applications of L'Hôpital's rule give

$$\lim_{x\to\infty} \frac{x^2}{e^x} = \lim_{x\to\infty} \frac{2x}{e^x} = \lim_{x\to\infty} \frac{2}{e^x} = 0.$$

□

Example 27.4

$$\lim_{x\to\infty} \frac{(\log(x))^3}{x^2} = 0.$$

Proof. Repeated applications of L'Hôpital's rule give

$$\begin{aligned}
\lim_{x\to\infty} \frac{(\log(x))^3}{x^2} &= \lim_{x\to\infty} \frac{3(\log(x))^2/x}{2x} \\
&= \lim_{x\to\infty} \frac{3(\log(x))^2}{2x^2} \\
&= \lim_{x\to\infty} \frac{6\log(x)/x}{4x} \\
&= \lim_{x\to\infty} \frac{6\log(x)}{4x^2} \\
&= \lim_{x\to\infty} \frac{6/x}{8x} \\
&= 0.
\end{aligned}$$

This is a standard instance of the fact that powers of $\log(x)$ grow more slowly than positive powers of x. □

Example 27.5

$$\lim_{x\to 0^+} x(\log(x))^2 = 0.$$

Proof. Write

$$x(\log(x))^2 = \frac{(\log(x))^2}{1/x}.$$

Applying L'Hôpital's rule twice gives

$$\begin{aligned}
\lim_{x\to 0^+} \frac{(\log(x))^2}{1/x} &= \lim_{x\to 0^+} \frac{2\log(x)\cdot\left(\dfrac{1}{x}\right)}{-1/x^2} \\
&= \lim_{x\to 0^+} -2x\log(x) \\
&= \lim_{x\to 0^+} \frac{-2\log(x)}{x^{-1}} \\
&= \lim_{x\to 0^+} \frac{-2/x}{-x^{-2}}
\end{aligned}$$

$$= \lim_{x\to 0^+} 2x$$
$$= 0.$$

□

Example 27.6

$$\lim_{x\to 0^+} x^{\sqrt{x}} = 1.$$

Proof. Write

$$x^{\sqrt{x}} = \exp\big(\sqrt{x}\log(x)\big).$$

From the previous example,

$$\begin{aligned}\lim_{x\to 0^+} \sqrt{x}\log(x) &= \lim_{x\to 0^+} \frac{\log(x)}{x^{-1/2}} \\ &= \lim_{x\to 0^+} \frac{1/x}{-(1/2)x^{-3/2}} \\ &= \lim_{x\to 0^+} (-2\sqrt{x}) \\ &= 0.\end{aligned}$$

By continuity of the exponential function,

$$\lim_{x\to 0^+} x^{\sqrt{x}} = e^0 = 1.$$

□

Example 27.7

$$\lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n = e.$$

Proof. Define

$$f(x) = \left(1 + \frac{1}{x}\right)^x$$

for $x > 0$. Then

$$f(x) = \exp\left(x\log\left(1 + \frac{1}{x}\right)\right).$$

Now

$$\lim_{x\to\infty} x\log\left(1+\frac{1}{x}\right) = \lim_{x\to\infty}\frac{\log(1+1/x)}{x^{-1}} = \lim_{x\to\infty}\frac{-(1/x^2)/(1+1/x)}{-x^{-2}} = \lim_{x\to\infty}\frac{1}{1+1/x} = 1.$$

Therefore,

$$\lim_{x\to\infty} f(x) = e,$$

and hence the same is true for the sequence $f(n) = (1+1/n)^n$. □

More generally, we have the following:

Proposition 27.8 Suppose $f(x) > 0$ eventually and

$$\lim_{x\to\infty} g(x)\log(f(x)) = L \in \mathbb{R}.$$

Then

$$\lim_{x\to\infty} f(x)^{g(x)} = e^L.$$

Proof. Since

$$f(x)^{g(x)} = \exp(g(x)\log(f(x))),$$

the conclusion follows immediately from continuity of the exponential function. □

Lecture 28 — Friday, November 16, 2012

7. Taylor Polynomials and Taylor's Theorem

Taylor Polynomials

Definition 28.1 Suppose that $f^{(k)}(a)$ exists for $k = 0, 1, \ldots, n$. The **Taylor polynomial of degree n centered at a** for f is

$$P_{n,a}(x) = \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k.$$

Proposition 28.2 For each $k = 0, 1, \ldots, n$,

$$P_{n,a}^{(k)}(a) = f^{(k)}(a).$$

Proof. The case $k = 0$ is immediate:

$$P_{n,a}(a) = f(a).$$

Differentiating once gives

$$P'_{n,a}(x) = \sum_{k=1}^{n} \frac{f^{(k)}(a)}{(k-1)!}(x-a)^{k-1},$$

so $P'_{n,a}(a) = f'(a)$. Differentiating twice gives

$$P''_{n,a}(x) = \sum_{k=2}^{n} \frac{f^{(k)}(a)}{(k-2)!}(x-a)^{k-2},$$

so $P''_{n,a}(a) = f''(a)$. The same pattern continues for every $k \leq n$. $\square$

Example 28.3 The first Taylor polynomials are

$$P_{1,a}(x) = f(a) + f'(a)(x-a),$$

which is the tangent line to the graph of f at a, and

$$P_{2,a}(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2.$$

More generally,

$$P_{n,a}(x) = P_{n-1,a}(x) + \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

Exercise 28.4 If f itself is a polynomial of degree at most n, prove that $P_{n,a}(x) = f(x)$ identically (use linear algebra).

Example 28.5 Let

$$f(x) = \sin(x), \quad a = 0.$$

Then

$$f(0) = 0, \qquad f'(0) = 1, \qquad f''(0) = 0,$$
$$f^{(3)}(0) = -1, \qquad \text{and} \qquad f^{(4)}(0) = 0,$$

so

$$P_{n,0}(x) = \sum_{j=0}^{\lfloor (n-1)/2 \rfloor} (-1)^j \frac{x^{2j+1}}{(2j+1)!},$$

where the sum is empty when $n = 0$. Thus the successive nonzero terms are $x - x^3/3! + x^5/5! - \cdots$.

The increasing order of contact at the expansion point is visible in Figure 14: the cubic approximation follows $\sin x$ farther from 0 than the tangent-line approximation does.

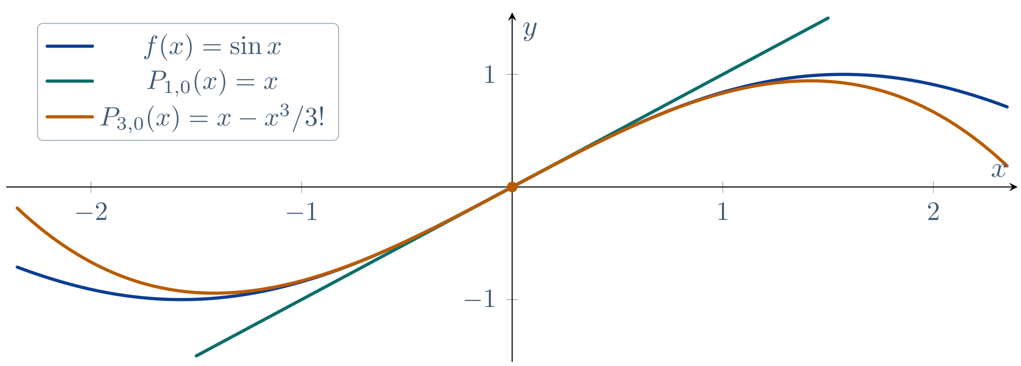

Figure 14

The central approximation question is whether $P_{n,a}(x)$ is a good approximation to $f(x)$, especially for x near a, and whether $P_{n,a}(x) \to f(x)$ as $n \to \infty$.

The general answer is no. Consider

$$f(x) = \begin{cases} e^{-1/x^2}, & x \neq 0, \\ 0, & x = 0. \end{cases}$$

For every $m \geq 0$,

$$\lim_{x \to 0} \frac{e^{-1/x^2}}{|x|^m} = 0;$$

indeed, after setting $t = 1/x^2$, this is the familiar fact that $t^{m/2}/e^t \to 0$, which follows from repeated applications of L'Hôpital's rule. Repeated differentiation away from zero gives

$$f^{(k)}(x) = p_k(1/x)e^{-1/x^2} \qquad (x \neq 0)$$

for some polynomial p_k. The displayed limit then shows inductively, through the difference quotient at zero, that $f^{(k)}(0) = 0$ for every k. Therefore every Taylor polynomial of f at 0 is identically zero:

$$P_{n,0}(x) = 0.$$

Hence $P_{n,0}(x) \not\to f(x)$ for any fixed $x \neq 0$.

For many important elementary functions, however, Taylor polynomials do converge to the original function.

Theorem 28.6 Let $n \in \mathbb{N}$. If f is n-times differentiable on an open interval containing a, and $P_{n,a}$ is the Taylor polynomial of degree n at a, then

$$\lim_{x \to a} \frac{f(x) - P_{n,a}(x)}{(x-a)^n} = 0.$$

Proof. Let

$$Q_n(x) = \sum_{k=0}^{n-1} \frac{f^{(k)}(a)}{k!}(x-a)^k = P_{n-1,a}(x).$$

Then

$$P_{n,a}(x) = Q_n(x) + \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

Hence

$$\lim_{x \to a} \frac{f(x) - P_{n,a}(x)}{(x-a)^n} = \lim_{x \to a} \frac{f(x) - Q_n(x)}{(x-a)^n} - \frac{f^{(n)}(a)}{n!}.$$

So it suffices to prove
$$\lim_{x\to a}\frac{f(x)-Q_n(x)}{(x-a)^n}=\frac{f^{(n)}(a)}{n!}.$$

For $k=0,1,\ldots,n-1$, we have $Q_n^{(k)}(a)=f^{(k)}(a)$. Since the numerator and denominator both vanish at $x=a$, we may apply L'Hôpital's rule repeatedly:
$$\lim_{x\to a}\frac{f(x)-Q_n(x)}{(x-a)^n}=\lim_{x\to a}\frac{f'(x)-Q_n'(x)}{n(x-a)^{n-1}}=\cdots=\lim_{x\to a}\frac{f^{(n-1)}(x)-Q_n^{(n-1)}(x)}{n!(x-a)}.$$

Now $Q_n^{(n-1)}(x)=Q_n^{(n-1)}(a)=f^{(n-1)}(a)$, because Q_n has degree $n-1$. Therefore
$$\lim_{x\to a}\frac{f^{(n-1)}(x)-Q_n^{(n-1)}(x)}{n!(x-a)}=\frac{1}{n!}\lim_{x\to a}\frac{f^{(n-1)}(x)-f^{(n-1)}(a)}{x-a}=\frac{f^{(n)}(a)}{n!}.$$

This is exactly the desired conclusion. □

Lecture 29 — Monday, November 19, 2012

Taylor's Theorem and Higher Derivative Tests

Recall that the Taylor polynomial of degree n centered at a is
$$P_{n,a}(x)=f(a)+f'(a)(x-a)+\frac{f''(a)}{2!}(x-a)^2+\cdots+\frac{f^{(n)}(a)}{n!}(x-a)^n.$$

Theorem 29.1 (General n-th derivative test) Let $n\in\mathbb{N}$, and suppose f is n-times differentiable on an open interval containing a, with
$$f^{(k)}(a)=0\quad\text{for } k=1,\ldots,n-1$$
and $f^{(n)}(a)\neq 0$. (When $n=1$, the first condition is vacuous.)

1. If n is even and $f^{(n)}(a)>0$, then f has a local minimum at a. If $f^{(n)}(a)<0$, then f has a local maximum at a.

2. If n is odd, then f has neither a local minimum nor a local maximum at a.

Proof. Subtracting the constant $f(a)$ does not change the derivatives or the local extremum behavior, so we may assume without loss of generality that $f(a) = 0$. Under this assumption,

$$P_{n,a}(x) = \frac{f^{(n)}(a)}{n!}(x-a)^n,$$

because all lower-order derivatives vanish at a. By Theorem 28.6,

$$\lim_{x\to a} \frac{f(x)}{(x-a)^n} = \frac{f^{(n)}(a)}{n!}.$$

If n is even and $f^{(n)}(a) > 0$, then this limit is positive, so for x sufficiently close to a,

$$\frac{f(x)}{(x-a)^n} > 0.$$

Since $(x-a)^n > 0$ for $x \neq a$, it follows that $f(x) > 0 = f(a)$, so a is a local minimum. If $f^{(n)}(a) < 0$, the same argument shows that a is a local maximum.

If n is odd and $f^{(n)}(a) > 0$, then $f(x)/(x-a)^n > 0$ near a. For $x > a$, this forces $f(x) > 0$, while for $x < a$, it forces $f(x) < 0$. Thus $f(a)$ is neither a local minimum nor a local maximum. The case $f^{(n)}(a) < 0$ is again analogous. $\square$

If $f^{(n)}(a) = 0$ for every n, then anything can happen. For example,

$$f(x) = \begin{cases} e^{-1/x^2}, & x \neq 0, \\ 0, & x = 0 \end{cases}$$

is continuous at 0 and has all derivatives equal to 0 at 0, as noted above, and yet 0 is a global minimum. The function $-f$ has the same derivatives at 0, but 0 is a global maximum. If we instead define

$$g(x) = \begin{cases} e^{-1/x^2}, & x > 0, \\ 0, & x = 0, \\ -e^{-1/x^2}, & x < 0, \end{cases}$$

then again every derivative at 0 vanishes, but 0 is neither a minimum nor a maximum. In all of these cases, the Taylor polynomials at 0 are identically zero.

Theorem 29.2 (Taylor's theorem) Suppose $f, f', \dots, f^{(n)}$ are continuous on the closed interval with endpoints a and x, and $f^{(n+1)}$ exists throughout its interior. Then

$$f(x) - P_{n,a}(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}$$

for some c between a and x.

Corollary 29.3 If $|f^{(n+1)}(t)| \leq M_n$ for all t between a and x, then

$$|f(x) - P_{n,a}(x)| \leq \frac{M_n|x-a|^{n+1}}{(n+1)!}.$$

Proof. By Taylor's theorem, there exists c between a and x such that

$$f(x) - P_{n,a}(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}.$$

Taking absolute values gives

$$|f(x) - P_{n,a}(x)| = \frac{|f^{(n+1)}(c)|}{(n+1)!}|x-a|^{n+1} \leq \frac{M_n|x-a|^{n+1}}{(n+1)!},$$

since $|f^{(n+1)}(c)| \leq M_n$. □

Corollary 29.4 If $|f^{(k)}(t)| \leq M$ for every $k \geq 0$ and every t between a and x, then $P_{n,a}(x) \to f(x)$ as $n \to \infty$.

Proof. Apply Corollary 29.3 with $M_n = M$. Then

$$|f(x) - P_{n,a}(x)| \leq \frac{M|x-a|^{n+1}}{(n+1)!}.$$

For fixed x, the right-hand side tends to 0 as $n \to \infty$, because factorial growth dominates any fixed power. Hence

$$|f(x) - P_{n,a}(x)| \to 0,$$

so $P_{n,a}(x) \to f(x)$. □

Example 29.5 For $f(x) = \sin(x)$, all derivatives are among $\pm\sin(x)$ and $\pm\cos(x)$, so $|f^{(n+1)}(t)| \le 1$ for every n and every t. Therefore

$$|\sin(x) - P_{n,0}(x)| \le \frac{|x|^{n+1}}{(n+1)!}.$$

For each fixed x, the right-hand side tends to 0, so the Taylor polynomials of $\sin(x)$ converge to $\sin(x)$. Recall that in this case,

$$P_{n,0}(x) = \sum_{j=0}^{\lfloor (n-1)/2 \rfloor} (-1)^j \frac{x^{2j+1}}{(2j+1)!}.$$

If $x \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, then

$$|\sin(x) - P_{n,0}(x)| \le \frac{2^{n+1}}{(n+1)!}.$$

If $x \in [-\frac{\pi}{4}, \frac{\pi}{4}]$, then the convergence is even more rapid; when $n = 30$, the remainder bound gives about 30 decimal places of accuracy.

Calculators need uniform accuracy over a prescribed range, not just accuracy near 0, so a raw Taylor series is usually not the best implementation. Range reduction, minimax polynomial approximations, and algorithms such as CORDIC are common alternatives.

Lecture 30 — Wednesday, November 21, 2012

Proof of Taylor's theorem (Theorem 29.2). Assume first that $a < x$; the case $x < a$ is identical with the interval reversed. Define, for $t \in [a, x]$,

$$R(t) = f(x) - f(t) - \sum_{k=1}^{n} \frac{f^{(k)}(t)}{k!}(x-t)^k.$$

Then $R(x) = 0$, while

$$R(a) = f(x) - P_{n,a}(x).$$

Also define

$$F(t) = \frac{(x-t)^{n+1}}{(n+1)!}.$$

Again $F(x) = 0$, and

$$F(a) = \frac{(x-a)^{n+1}}{(n+1)!}.$$

Both R and F are continuous on $[a, x]$ and differentiable on (a, x), so the Cauchy mean value theorem applies. Therefore there exists $c \in (a, x)$ such that

$$(R(a) - R(x))F'(c) = R'(c)(F(a) - F(x)).$$

Now

$$F'(t) = -\frac{(x-t)^n}{n!}.$$

For R, differentiating and simplifying the resulting telescoping sum gives

$$R'(t) = -\frac{f^{(n+1)}(t)}{n!}(x-t)^n.$$

Since $R(x) = F(x) = 0$, the Cauchy mean value theorem identity becomes

$$R(a)\left(-\frac{(x-c)^n}{n!}\right) = \left(-\frac{f^{(n+1)}(c)}{n!}(x-c)^n\right)F(a).$$

Canceling the common factor $-(x-c)^n/n!$, we obtain

$$R(a) = f^{(n+1)}(c)F(a) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}.$$

Since $R(a) = f(x) - P_{n,a}(x)$, this is exactly Taylor's theorem. □

Example 30.1 e is irrational.

Proof. Apply Taylor's theorem to $f(x) = e^x$ at $a = 0$. Since every derivative of e^x is again e^x, we have

$$P_{n,0}(x) = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!},$$

and

$$e^x - P_{n,0}(x) = \frac{e^c}{(n+1)!}x^{n+1}$$

for some c between 0 and x. Setting $x = 1$, we obtain

$$e = 1 + 1 + \frac{1}{2!} + \cdots + \frac{1}{n!} + E_n,$$

where

$$E_n = \frac{e^c}{(n+1)!}$$

for some $c \in (0,1)$. In particular,

$$0 < E_n < \frac{e}{(n+1)!}.$$

Suppose for a contradiction that $e = p/q$ with $p, q \in \mathbb{N}$. Choose an integer $n \geq q$ such that $n + 1 > e$. Multiplying the previous identity by $n!$ gives

$$n!e = n!\left(2 + \frac{1}{2!} + \cdots + \frac{1}{n!}\right) + n!E_n.$$

The left-hand side is an integer because $q \mid n!$, and the sum on the right excluding $n!E_n$ is also an integer. Therefore $n!E_n$ must be an integer. But

$$0 < n!E_n < \frac{e}{n+1} < 1,$$

so $n!E_n$ is a positive integer strictly between 0 and 1, which is impossible. Hence e is irrational. □

The number $\sin(1)$ is also irrational, although proving that fact requires a more delicate argument than this direct remainder estimate.

Lecture 31 — Friday, November 23, 2012

8. Newton's Method

The Algorithm

Recall the bisection method: if $f : [a,b] \to \mathbb{R}$ is continuous and $f(a) < 0 < f(b)$, then the intermediate value theorem guarantees a root $c \in (a,b)$. Newton's method seeks the same root by using tangent lines instead of repeated interval bisection.

Definition 31.1 Suppose f is differentiable on an interval containing a root c, and assume $f'(x) \neq 0$ at the points under consideration. Starting from an initial guess x_0, define

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \quad \text{for every integer } n \geq 0.$$

The numbers x_n are called the **Newton iterates** for f.

The tangent line to the graph of f at x_n has equation

$$y - f(x_n) = f'(x_n)(x - x_n).$$

Setting $y = 0$ gives the x-intercept of this tangent line:

$$x = x_n - \frac{f(x_n)}{f'(x_n)} = x_{n+1}.$$

Thus Newton's method replaces f near x_n by its tangent line and uses the zero of that tangent line as the next approximation.

For $f(x) = x^2 - 2$, the first two tangent steps are shown in Figure 15. Each new iterate is read from the horizontal-axis intercept of the previous tangent.

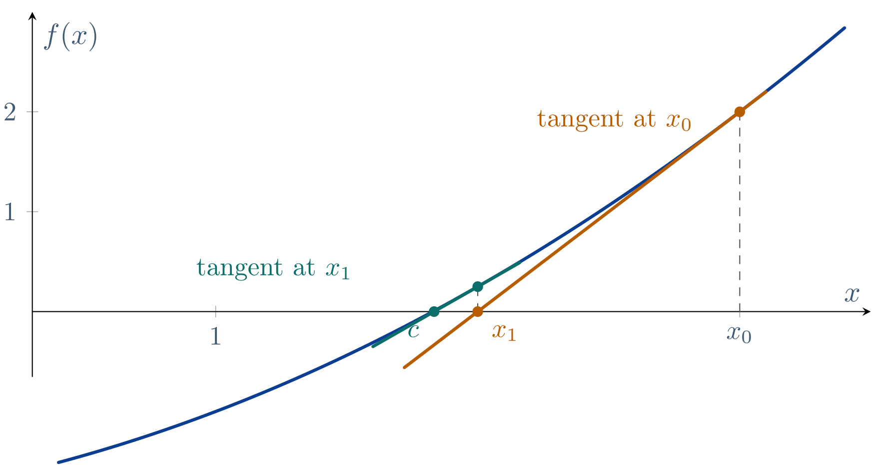

FIGURE 15

Theorem 31.2 (Newton's method) Let $f : [a, b] \to \mathbb{R}$ be such that f, f', and f'' are continuous on $[a, b]$. Assume that

$$f(a) < 0 < f(b) \qquad \text{and} \qquad f'(x), f''(x) > 0 \text{ for all } x \in [a, b].$$

Let $c \in (a, b)$ be the unique root of f, and choose $x_0 \in (c, b]$. Define the Newton iterates by

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Then each x_n lies in $(c, b]$, the sequence (x_n) is decreasing, and $x_n \to c$.

Proof. Since $f' > 0$ on $[a, b]$, the function f is strictly increasing, so it has exactly one root $c \in (a, b)$. We first show that $c < x_1 \le x_0$. Because $x_0 > c$ and f is strictly increasing, we have $f(x_0) > f(c) = 0$. Since $f'(x_0) > 0$,

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} < x_0.$$

Apply the mean value theorem to f on $[c, x_0]$. There exists $t_0 \in (c, x_0)$ such that

$$\frac{f(x_0) - f(c)}{x_0 - c} = f'(t_0).$$

Since $f(c) = 0$, we obtain

$$f(x_0) = f'(t_0)(x_0 - c),$$

and hence

$$c = x_0 - \frac{f(x_0)}{f'(t_0)}.$$

Because $f'' > 0$, the derivative f' is strictly increasing, so $f'(t_0) < f'(x_0)$. Since $f(x_0) > 0$, it follows that

$$c = x_0 - \frac{f(x_0)}{f'(t_0)} < x_0 - \frac{f(x_0)}{f'(x_0)} = x_1.$$

Thus $c < x_1 \le x_0$. Now assume inductively that

$$c < x_n \le x_{n-1} \le \cdots \le x_0 \le b.$$

Then $x_n \in [a, b]$, so x_{n+1} is defined. As before, $f(x_n) > 0$, and therefore

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} < x_n.$$

Applying the mean value theorem to f on $[c, x_n]$, we find $t_n \in (c, x_n)$ such that

$$f(x_n) = f'(t_n)(x_n - c).$$

Hence
$$c = x_n - \frac{f(x_n)}{f'(t_n)}.$$
Since $t_n < x_n$ and f' is strictly increasing, we have $f'(t_n) < f'(x_n)$, which implies
$$c < x_n - \frac{f(x_n)}{f'(x_n)} = x_{n+1}.$$
Thus $c < x_{n+1} < x_n$. By induction, every iterate lies in $(c, b]$, and the sequence (x_n) is decreasing and bounded below by c. By the monotone convergence theorem, (x_n) converges to some $p \in [c, b]$. Passing to the limit in
$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$
and using continuity of f and f', together with $f'(p) > 0$, we obtain
$$p = p - \frac{f(p)}{f'(p)}.$$
Therefore $f(p) = 0$. Since f has only one root, we must have $p = c$. □

How fast do the Newton iterates converge? Very fast:

Theorem 31.3 Under the hypotheses of Theorem 31.2, let
$$M_1 = \max_{x\in[a,b]} f''(x), \qquad M_2 = \min_{x\in[a,b]} f'(x), \qquad \text{and} \qquad M = \frac{M_1}{M_2}.$$
Then
$$|x_n - c| \le \frac{1}{M}\big(M|x_0 - c|\big)^{2^n} \quad \text{for every integer } n \ge 0.$$

Proof. From the previous proof, for each n there exists $t_n \in (c, x_n)$ such that
$$c = x_n - \frac{f(x_n)}{f'(t_n)}.$$
Therefore
$$|x_{n+1} - c| = \left| x_n - \frac{f(x_n)}{f'(x_n)} - x_n + \frac{f(x_n)}{f'(t_n)} \right| = |f(x_n)| \left| \frac{1}{f'(t_n)} - \frac{1}{f'(x_n)} \right| = \frac{|f(x_n)|\,|f'(x_n) - f'(t_n)|}{|f'(x_n)f'(t_n)|},$$

and by the mean value theorem applied to f', there exists $u_n \in (t_n, x_n)$ such that

$$f'(x_n) - f'(t_n) = f''(u_n)(x_n - t_n).$$

Lecture 32 — Monday, November 26, 2012

Proof continued. Substituting this identity gives

$$|x_{n+1} - c| = \frac{|f(x_n)|\,|f''(u_n)|\,|x_n - t_n|}{|f'(x_n)f'(t_n)|}.$$

Since $f(x_n) = f'(t_n)(x_n - c)$, we obtain

$$|x_{n+1} - c| = \frac{|x_n - c|\,|f''(u_n)|\,|x_n - t_n|}{|f'(x_n)|}.$$

Now $c < t_n < x_n$, so $0 < x_n - t_n < x_n - c$. Also, $f''(u_n) \le M_1$ and $f'(x_n) \ge M_2$. Therefore

$$|x_{n+1} - c| \le \frac{M_1}{M_2}|x_n - c|^2 = M|x_n - c|^2.$$

Multiplying by M yields

$$M|x_{n+1} - c| \le \big(M|x_n - c|\big)^2.$$

Iterating this estimate gives

$$M|x_n - c| \le \big(M|x_0 - c|\big)^{2^n},$$

and dividing by M proves the stated bound. □

Error Bounds and Failure Modes

If $b - c$ is large, then it is often efficient to apply a few steps of the bisection method first. Once we have a smaller interval $[a', b']$ containing the root and satisfying $|b' - a'| \le \frac{1}{2M}$, choosing $x_0 = b'$ gives

$$M|x_0 - c| \le M|b' - c| \le M|b' - a'| \le \frac{1}{2}.$$

The quadratic estimate in Theorem 31.3 then shows that the Newton iterates converge extremely rapidly.

Example 32.1 To approximate $\sqrt{2}$, take $f(x) = x^2 - 2$. On $[1, 2]$ we have $f(1) < 0 < f(2)$, $f'(x) = 2x > 0$, and $f''(x) = 2 > 0$. The positive root is $c = \sqrt{2}$.

If we instead start on the smaller interval $[1.4, 1.5]$, then $c \in [1.4, 1.5]$, $M_1 = 2$, and

$$M_2 = \min_{x \in [1.4,1.5]} 2x = 2.8, \qquad \text{and} \qquad M = \frac{2}{2.8} = \frac{5}{7}.$$

With $x_0 = 1.5$, we obtain

$$M|x_0 - c| \leq M|1.5 - 1.4| = \frac{5}{7} \cdot \frac{1}{10} = \frac{1}{14}.$$

The first two Newton iterates are

$$x_1 = x_0 - \frac{f(x_0)}{f'(x_0)} = \frac{3}{2} - \frac{1/4}{3} = \frac{17}{12},$$

and

$$x_2 = \frac{17}{12} - \frac{(17/12)^2 - 2}{2(17/12)} = \frac{577}{408} \approx 1.414215687.$$

Since

$$\sqrt{2} \approx 1.414213562,$$

we already obtain several correct decimal places after only two iterations.

The same reasoning works with minor sign changes in other geometric situations. For instance, if $f'' < 0$ on $[a, b]$, then it is natural to begin with $x_0 \in [a, c)$ instead. Likewise, the method can be adapted when $f' < 0$ on $[a, b]$.

When Newton's Method Can Fail

Without the convexity hypothesis, or with a poor starting point, Newton's method can behave badly. It may oscillate, diverge, or leave the region in which our previous theorem applies. For example, the function $f(x) = x^3 - 2x + 2$ started at $x_0 = 0$ produces the 2-cycle $0 \mapsto 1 \mapsto 0 \mapsto \cdots$ rather than convergence to the real root r. See Figure 16.

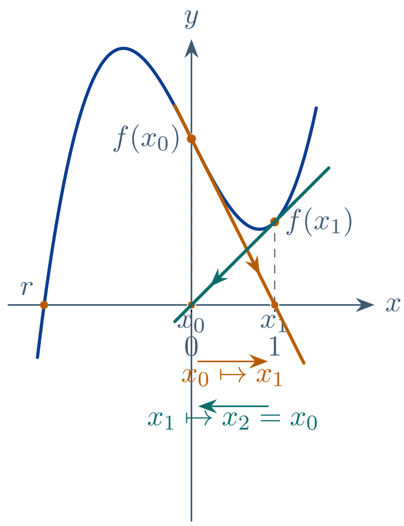

FIGURE 16

Proposition 32.2 Let f be differentiable on an interval I, and let $(x_n) \subset I$ be a Newton sequence:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Assume that $x_n \to p \in I$, that f' is continuous at p, and that $f'(p) \neq 0$. Then $f(p) = 0$.

Proof. Rearranging the Newton update gives

$$\frac{f(x_n)}{f'(x_n)} = x_n - x_{n+1}.$$

Since both x_n and x_{n+1} converge to p, the right-hand side converges to 0. Continuity of f and f' at p therefore gives

$$\frac{f(p)}{f'(p)} = 0.$$

Because $f'(p) \neq 0$, it follows that $f(p) = 0$. $\square$

Thus a convergent Newton sequence can fail to converge to a root only if the derivative misbehaves near the limit.

Example 32.3 Newton's method can cycle. Consider

$$f(x) = x^3 - 2x + 2.$$

Then

$$f'(x) = 3x^2 - 2,$$

so the Newton map is

$$N(x) = x - \frac{x^3 - 2x + 2}{3x^2 - 2}.$$

If we start at $x_0 = 0$, then

$$x_1 = N(0) = 1, \qquad \text{and} \qquad x_2 = N(1) = 0.$$

Thus the iterates alternate between 0 and 1, and the method never converges.

Example 32.4 Newton's method can also diverge. Take

$$f(x) = \sqrt[3]{x}.$$

For $x \neq 0$,

$$f'(x) = \frac{1}{3}x^{-2/3},$$

so for $x_n \neq 0$,

$$x_{n+1} = x_n - \frac{x_n^{1/3}}{\frac{1}{3}x_n^{-2/3}} = x_n - 3x_n = -2x_n.$$

Hence $|x_n| = 2^n|x_0|$ for every nonzero starting value, so the iterates move away from the root rather than toward it.

Proposition 32.5 There exists a function $f : \mathbb{R} \to \mathbb{R}$, differentiable everywhere, for which Newton's method converges to a point which is not a root of f.

Proof. For each integer $n \geq 0$, cubic Hermite interpolation gives a polynomial P_n satisfying

$$P_n\left(\frac{1}{2^n}\right) = -2\pi n \qquad \text{and} \qquad P_n\left(\frac{1}{2^{n+1}}\right) = -2\pi(n+1),$$

and

$$P_n'\left(\frac{1}{2^n}\right) = 2^{3n+1} \qquad \text{and} \qquad P_n'\left(\frac{1}{2^{n+1}}\right) = 2^{3(n+1)+1}.$$

Define $g : (0, \infty) \to \mathbb{R}$ by

$$g(x) = \begin{cases} P_n(x), & x \in \left(\dfrac{1}{2^{n+1}}, \dfrac{1}{2^n}\right], \quad n = 0, 1, 2, \ldots, \\ 2x - 2, & x > 1 \end{cases}.$$

The prescribed values and derivatives agree at every junction 2^{-n}. They also agree with the value and derivative of $2x - 2$ at $x = 1$. Consequently, g is continuously differentiable on $(0, \infty)$.

Now define

$$f(x) = \begin{cases} x^2 \sin(g(x)) + 1, & x > 0 \\ 1, & x = 0 \\ -x^2 + 1, & x < 0 \end{cases}$$

The bound $\left|x^2 \sin(g(x))\right| \leq x^2$ shows that f is continuous at 0, and it is differentiable away from 0. To check differentiability at 0, observe that

$$\lim_{h \to 0^-} \frac{f(h) - f(0)}{h} = \lim_{h \to 0^-} \frac{-h^2 + 1 - 1}{h} = 0$$

and

$$\lim_{h \to 0^+} \frac{f(h) - f(0)}{h} = \lim_{h \to 0^+} h \sin(g(h)) = 0.$$

So f is differentiable everywhere, and $f'(0) = 0$. For every integer $n \geq 0$,

$$f\left(\frac{1}{2^n}\right) = \frac{1}{2^{2n}} \sin\left(g\left(\frac{1}{2^n}\right)\right) + 1 = \frac{1}{2^{2n}} \sin(-2\pi n) + 1 = 1$$

and

$$f'\left(\frac{1}{2^n}\right) = \frac{2}{2^n} \sin\left(g\left(\frac{1}{2^n}\right)\right) + \frac{1}{2^{2n}} \cos\left(g\left(\frac{1}{2^n}\right)\right) g'\left(\frac{1}{2^n}\right) = \frac{1}{2^{2n}} \cos(-2\pi n) 2^{3n+1} = 2^{n+1}.$$

Now, take $x_0 = 1$. The Newton iterates are

$$\begin{aligned} x_1 &= x_0 - \frac{f(x_0)}{f'(x_0)} = 1 - \frac{1}{2^1} = \frac{1}{2} \\ x_2 &= x_1 - \frac{f(x_1)}{f'(x_1)} = \frac{1}{2} - \frac{1}{2^2} = \frac{1}{4}. \end{aligned}$$

We claim that $x_n = \frac{1}{2^n}$. The base cases are already done. Now assume $x_n = \frac{1}{2^n}$ for some integer $n \geq 0$. Then

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = \frac{1}{2^n} - \frac{1}{2^{n+1}} = \frac{1}{2^{n+1}},$$

where we used the formulas $f\left(\frac{1}{2^n}\right) = 1$ and $f'\left(\frac{1}{2^n}\right) = 2^{n+1}$ established above. Therefore, by induction, $x_n = \frac{1}{2^n}$ for every integer $n \geq 0$.

Thus $x_n \to 0$, but $f(0) = 1 \neq 0$, so the Newton iterates converge to a point that is not a root of f. $\square$

Lecture 33 — Wednesday, November 28, 2012

9. Cardinality

Finiteness and Countability

Definition 33.1 A function $f\colon A \to B$ is called **onto**, or **surjective**, if for every $b \in B$, there exists some $a \in A$ such that $f(a) = b$. A **bijection** is a function that is both one-to-one and onto.

A bijection pairs each element of the domain with exactly one element of the codomain, and every codomain element appears. Equivalently, $f : A \to B$ is a bijection precisely when it has an inverse $f^{-1} : B \to A$ satisfying $f(f^{-1}(b)) = b$ and $f^{-1}(f(a)) = a$ for all $a \in A$ and $b \in B$.

Definition 33.2 A set E is finite if there is a bijection $f : E \to \{1, 2, \ldots, n\}$ for some integer $n \geq 0$, where the target is empty when $n = 0$. In this case, n is the **cardinality** of E.

Definition 33.3 A set E is **countable** if there is a bijection $f : E \to \mathbb{N}$. Thus, in these notes, "countable" means countably infinite; finite sets are treated separately. We say that the cardinality of a countable set is $\aleph_0$.

Examples of countable sets include $\mathbb{N}$, $2\mathbb{N} = \{2n : n \in \mathbb{N}\}$, $\mathbb{Z}$, $\mathbb{N} \times \mathbb{N}$, and $\mathbb{Q}$.

Example 33.4 There exists a set of intervals whose total length can be made arbitrarily small, yet whose union contains $\mathbb{Q}$.

Proof. Let $\varepsilon > 0$. Enumerate $\mathbb{Q}$ as $\mathbb{Q} = \{r_1, r_2, \ldots\}$ and let

$$I_j = \left(r_j - \frac{\varepsilon}{2^{j+1}}, r_j + \frac{\varepsilon}{2^{j+1}}\right), \quad j \in \mathbb{N}.$$

Now, let

$$X = \bigcup_{j=1}^{\infty} I_j = \{x : x \in I_j \text{ for some } j\}.$$

Because each $r_j \in I_j$, we must have that $r_j \in X$ for any $j \in \mathbb{N}$, and hence $\mathbb{Q} \subseteq X$. However, the total length of the intervals $I_1, I_2, \ldots$ is

$$\sum_{j=1}^{\infty} \text{length}(I_j) = \sum_{j=1}^{\infty} \frac{\varepsilon}{2^j} = \varepsilon.$$

$\square$

In the construction above, observe that $\mathbb{R} \setminus X$ contains no intervals, since every interval contains a rational number.

Proposition 33.5 A set E is countable if and only if its elements can be put into an infinite ordered list.

Proof. Suppose first that the elements of E are ordered as $E = \{e_1, e_2, \ldots\}$. Define $f : E \to \mathbb{N}$ by $e_n \mapsto n$; this is a bijection between E and $\mathbb{N}$.

Conversely, if E is countable, choose a bijection $f : \mathbb{N} \to E$. Then $E = \{f(1), f(2), \ldots\}$ is an infinite ordered list. $\square$

Theorem 33.6 A union of two countable sets is countable.

Proof. Write $A = \{a_1, a_2, \ldots\}$ and $B = \{b_1, b_2, \ldots\}$. Interleave the two lists:

$$a_1, b_1, a_2, b_2, a_3, b_3, \ldots.$$

Delete each entry that has already appeared earlier in the list. Every element of $A \cup B$ still appears, and no element appears twice. The resulting list is infinite because it contains all the distinct elements of A. It is therefore an infinite ordered list of $A \cup B$, so the preceding proposition shows that $A \cup B$ is countable. $\square$

Lecture 34 — Friday, November 30, 2012

Theorem 34.1 A nonempty set E is either countable or finite if and only if there exists a map $g : \mathbb{N} \to E$ that is onto.

Proof. *Forward implication.* If E is countable, then by definition there is a bijection $f : E \to \mathbb{N}$. Hence $f^{-1} : \mathbb{N} \to E$ is a bijection, and in particular it is onto.

If E is finite, say $E = \{e_1, e_2, \ldots, e_k\}$, define $g : \mathbb{N} \to E$ by

$$g(n) = e_n \text{ for } 1 \leq n \leq k \qquad \text{and} \qquad g(n) = e_k \text{ for } n > k.$$

Then every element of E appears as some value of g, so g is onto.

Reverse implication. Suppose there exists an onto map $g : \mathbb{N} \to E$. Let

$$S = \{n \in \mathbb{N} : g(n) = g(m) \text{ for some } m < n\}$$

and let $T = \mathbb{N} \setminus S$. Thus $n \in T$ exactly when $g(n)$ appears for the first time at index n.

The map $h : T \to E$ defined by $h(n) = g(n)$ is one-to-one by construction and onto because g is onto, so h is a bijection from T onto E.

If T is finite, then E is finite. If T is infinite, list its elements increasingly as $t_1 < t_2 < \cdots$ and define $\varphi : \mathbb{N} \to T$ by $\varphi(j) = t_j$. Then φ is a bijection, so E is countable (via $h \circ \varphi : \mathbb{N} \to E$).

Therefore, E is either finite or countable. $\square$

Corollary 34.2 If $h : E \to \mathbb{N}$ is one-to-one, then E is either countable or finite.

Proof. If $E = \emptyset$, then E is finite. Assume $E \neq \emptyset$ and choose $e^* \in E$. Define $g : \mathbb{N} \to E$ as follows. If $n \in \text{Range}(h)$, then there is a unique $e \in E$ with $h(e) = n$ because h is one-to-one, so define $g(n) = e$. If $n \notin \text{Range}(h)$, define $g(n) = e^*$.

By construction, g is onto, and therefore E is either countable or finite by Theorem 34.1. □

Uncountability and the Irrational Numbers

Definition 34.3 A set that is neither countable nor finite is **uncountable**.

Corollary 34.4

1. If $A \subseteq B$ and B is countable, then A is either countable or finite.
2. If $A \subseteq B$ and A is uncountable, then B is uncountable.

Proof. (1) Since B is countable, Theorem 34.1 gives an onto map $g : \mathbb{N} \to B$.

If $A = \emptyset$, then A is finite. So assume $A \neq \emptyset$, and choose $a_0 \in A$. Define $q : \mathbb{N} \to A$ by

$$q(n) = \begin{cases} g(n), & g(n) \in A, \\ a_0, & g(n) \notin A. \end{cases}$$

For any $a \in A$, since g is onto B and $a \in B$, there exists $n \in \mathbb{N}$ with $g(n) = a$. Then $g(n) \in A$, so $q(n) = a$. Hence q is onto A. Applying Theorem 34.1 again, A is either countable or finite.

(2) If B were countable, then because $A \subseteq B$, part (1) would imply that A is either countable or finite, a contradiction. Therefore B is uncountable. □

Theorem 34.5 $(0, 1)$ is uncountable.

The proof uses Cantor's famous **diagonalization argument**.

Proof. Suppose for a contradiction that $(0, 1)$ is countable, and enumerate it as $(0, 1) = \{a_1, a_2, \ldots\}$. Write the decimal expansions in alignment, always choosing the expansion that does not end in an infinite string of 9s:

$$\begin{aligned} a_1 &= 0.a_{1,1}a_{1,2}a_{1,3}\dots \\ a_2 &= 0.a_{2,1}a_{2,2}a_{2,3}\dots \\ a_3 &= 0.a_{3,1}a_{3,2}a_{3,3}\dots \\ &\vdots \end{aligned}$$

where each $a_{i,j} \in \{0, 1, \dots, 9\}$. Now, define $r \in (0, 1)$ as follows:

$$r = 0.r_1r_2r_3\dots, \qquad \text{with} \qquad r_j = \begin{cases} 5, & a_{j,j} \neq 5, \\ 4, & a_{j,j} = 5. \end{cases}.$$

Then $r \notin \{a_1, a_2, \dots\}$ by construction, so $(0, 1) \not\subseteq \{a_1, a_2, \dots\}$, a contradiction. Hence $(0, 1)$ is uncountable. $\square$

Corollary 34.6 $\mathbb{R}$ is uncountable.

Proof. Suppose for a contradiction that $\mathbb{R}$ is countable, so that there exists a bijection $f : \mathbb{R} \to \mathbb{N}$. There is already a bijection $h : \mathbb{R}$ to $(0, 1)$ given by

$$h(x) = \frac{1}{\pi}\arctan(x) + \frac{1}{2}.$$

Then $(h \circ f^{-1}) : \mathbb{N} \to (0, 1)$ is a bijection as well, which contradicts the fact that $(0, 1)$ is uncountable. Hence $\mathbb{R}$ is uncountable. $\square$

Lecture 35 — Saturday, December 01, 2012

Unions of Countable Sets

Recall that a nonempty set E is either countable or finite if and only if there is an onto function $f : \mathbb{N} \to E$.

Corollary 35.1 A countable union of countable or finite sets is countable or finite.

Proof. Let $A_1, A_2, \ldots$ be countably many sets, each countable or finite, and set

$$A = \bigcup_{j=1}^{\infty} A_j.$$

If $A = \emptyset$, the conclusion is immediate. Otherwise choose $a^* \in A$. For each nonempty A_j, Theorem 34.1 gives an onto map $f_j : \mathbb{N} \to A_j$; if $A_j = \emptyset$, define $f_j(k) = a^*$ instead. Define

$$h : \mathbb{N} \times \mathbb{N} \to A \qquad \text{by} \qquad h(j,k) = f_j(k).$$

The map h is onto: if $a \in A$, then $a \in A_j$ for some nonempty A_j, and $f_j(k) = a$ for some $k \in \mathbb{N}$.

Now, take any bijection $g : \mathbb{N} \to \mathbb{N} \times \mathbb{N}$ and form the composition $(h \circ g) : \mathbb{N} \to A$, which must be onto. By Theorem 34.1, A is either countable or finite. □

More generally, if A_α is either countable or finite for each $\alpha \in I$, where I is countable, then

$$A = \bigcup_{\alpha \in I} A_\alpha = \{x : x \in A_\alpha \text{ for some } \alpha \in I\}$$

is either countable or finite.

Definition 35.2 The **algebraic numbers** are the numbers (either real or complex) which are roots of polynomials with integer coefficients. A number that is not algebraic is **transcendental**.

Example 35.3 The set of algebraic numbers is countable, and therefore the set of transcendental numbers is uncountable.

Proof. For reference, the **minimal polynomial** of an algebraic number is the unique primitive, irreducible polynomial in $\mathbb{Z}[x]$ with positive leading coefficient that has the number as a root.

We do not need uniqueness here. For each $n \geq 1$, let P_n be the set of integer polynomials of degree exactly n. The coefficient map

$$a_n x^n + \cdots + a_1 x + a_0 \longmapsto (a_n, \ldots, a_1, a_0)$$

injects P_n into $\mathbb{Z}^{n+1}$.

Every finite Cartesian power of $\mathbb{Z}$ is countable. Indeed, $\mathbb{Z}$ is countable. If $\mathbb{Z}^m$ is countable, choose bijections $f : \mathbb{Z}^m \to \mathbb{N}$ and $g : \mathbb{Z} \to \mathbb{N}$. Then

$$(w, z) \longmapsto (f(w), g(z))$$

is a bijection from $\mathbb{Z}^{m+1}$ to $\mathbb{N} \times \mathbb{N}$, which is countable. Induction proves the claim, and hence each P_n is countable.

Each polynomial in P_n has at most n complex roots. Thus the set of all algebraic numbers is contained in a countable union, over n and over P_n, of finite root sets. By Corollary 35.1, the algebraic numbers are finite or countable; because every integer is algebraic, the set is infinite and therefore countable. If the real transcendental numbers were finite or countable as well, their union with the real algebraic numbers would make $\mathbb{R}$ countable, contradicting the preceding corollary. Therefore the real transcendental numbers, and hence the full set of transcendental numbers, are uncountable. $\square$

The construction in the proof shows that $\mathbb{Z}^n$ and $\mathbb{N}^n$ are countable, for any $n \in \mathbb{N}$. What about "$\mathbb{Z}^\infty$"? Consider

$$\mathbb{N}^\mathbb{N} = \{(a_1, a_2, \dots) : a_j \in \mathbb{N} \text{ for each } j\} = \text{the set of all functions from } \mathbb{N} \text{ to } \mathbb{N}.$$

Proposition 35.4 $\mathbb{N}^\mathbb{N}$ is uncountable.

Proof. The set $\{1, 2\}^\mathbb{N}$ is plainly infinite. Suppose, for a contradiction, that every sequence in $\{1, 2\}^\mathbb{N}$ appears in a list

$$a^{(1)}, a^{(2)}, a^{(3)}, \dots \qquad \text{and} \qquad a^{(i)} = (a_1^{(i)}, a_2^{(i)}, \dots).$$

Define a new sequence b by

$$b_j = \begin{cases} 1, & a_j^{(j)} = 2, \\ 2, & a_j^{(j)} = 1. \end{cases}$$

For every j, the sequence b differs from $a^{(j)}$ in its jth entry, so it does not appear in the list. Hence $\{1, 2\}^\mathbb{N}$ is uncountable. Since $\{1, 2\}^\mathbb{N} \subseteq \mathbb{N}^\mathbb{N}$, the latter set is uncountable as well. $\square$

Replacing each 1 by 0 and each 2 by 1 gives a bijection from $\{1,2\}^{\mathbb{N}}$ to $\{0,1\}^{\mathbb{N}}$. There is, in turn, a bijection from $\{0,1\}^{\mathbb{N}}$ to $2^{\mathbb{N}}$, the set of all subsets of $\mathbb{N}$: send $(a_1, a_2, \ldots)$ to the set of indices j for which $a_j = 1$. Therefore $2^{\mathbb{N}}$ is uncountable as well. More generally, the power set of any set A has strictly larger cardinality than A, producing an infinite hierarchy of cardinalities.

Appendix: Lecture and Tutorial Index

1. Lecture 1 — Monday, September 10, 2012
2. Lecture 2 — Wednesday, September 12, 2012
3. Lecture 3 — Friday, September 14, 2012
4. Lecture 4 — Monday, September 17, 2012
5. Lecture 5 — Wednesday, September 19, 2012
6. Lecture 6 — Friday, September 21, 2012
7. Lecture 7 — Monday, September 24, 2012
8. Lecture 8 — Wednesday, September 26, 2012
9. Lecture 9 — Friday, September 28, 2012
10. Lecture 10 — Monday, October 01, 2012
11. Lecture 11 — Wednesday, October 03, 2012
12. Lecture 12 — Friday, October 05, 2012
13. Lecture 13 — Wednesday, October 10, 2012
14. Lecture 14 — Friday, October 12, 2012
15. Lecture 15 — Monday, October 15, 2012
16. Lecture 16 — Wednesday, October 17, 2012
17. Lecture 17 — Friday, October 19, 2012
18. Lecture 18 — Monday, October 22, 2012
19. Lecture 19 — Wednesday, October 24, 2012
20. Lecture 20 — Monday, October 29, 2012
21. Lecture 21 — Wednesday, October 31, 2012
22. Lecture 22 — Friday, November 02, 2012
23. Lecture 23 — Monday, November 05, 2012
24. Lecture 24 — Wednesday, November 07, 2012
25. Lecture 25 — Friday, November 09, 2012
26. Lecture 26 — Monday, November 12, 2012

27. Lecture 27 — Wednesday, November 14, 2012
28. Lecture 28 — Friday, November 16, 2012
29. Lecture 29 — Monday, November 19, 2012
30. Lecture 30 — Wednesday, November 21, 2012
31. Lecture 31 — Friday, November 23, 2012
32. Lecture 32 — Monday, November 26, 2012
33. Lecture 33 — Wednesday, November 28, 2012
34. Lecture 34 — Friday, November 30, 2012
35. Lecture 35 — Saturday, December 01, 2012