Introduction to Analysis

Part II: Math 172

Lecture Notes
by
Stefan Waner

(second printing: 2001)

Department of Mathematics, Hofstra University

1. The Riemann Integral

Definition 1.1 A **partition** of the closed interval $[a, b]$ is a sequence

$$P = \{x_0 = a, x_1, x_2, \ldots, x_n = b\}$$

with $x_{i-1} < x_i$ for $1 \le i \le n$,. Also define

$$\Delta x_i = x_i - x_{i-1} \qquad (1 \le i \le n)$$

The **norm** (**mesh**) of the partition P is defined as

$$|P| = \Delta P = \max\{\Delta x_i \mid 1 \le i \le n\}$$

Examples 1.2

(a) $\{0, \frac{1}{2}, 1\}$ is a partition of $[0, 1]$.

(b) $\{\frac{1}{2^i} \mid 0 \le i \le n\} \cup \{0\}$ is also a parition of $[0, 1]$.

Definition 1.3 A partition Q of $[a, b]$ is a **refinement** of the partition P of $[a, b]$ if $P \subset Q$.

Examples 1.4

(a) $P = \{0, \frac{1}{2}, 1\}$; $Q = \{0, \frac{1}{4}, \frac{1}{2}, \frac{4}{5}, 1\}$ is a refinement of P.

(b) if P and Q are any two partitions of $[a, b]$, then the set $P \cup Q$ (ordered as a sequence) is a refinement of both P and Q.

Remarks 1.5

(a) If $a < b$, then any partition P of $[a, b]$ has a refinement.

(b) If P is any partition of $[a, b]$ and if $\varepsilon > 0$, then there exists a refinement Q of P with $\Delta Q < \varepsilon$. (Exercises)

Recall that a function f: $D \to \mathrm{R}$ is **bounded** if $\{f(x) \mid x \in D\}$ is a bounded set; that is, there exists $M \ge 0$ with $|f(x)| \le M$ for every $x \in D$.

Definitions 1.6 Let f: $[a, b] \to \mathrm{R}$ be bounded and let P be a partition of $[a, b]$. The **upper and lower Darboux sums of f** are given by

$$U(f, P) = \sum_{i=1}^{n} M_i \, \Delta x_i$$

Upper Sum

$$L(f, P) = \sum_{i=1}^{n} m_i \, \Delta x_i$$

Lower Sum

where

$$M_i = \sup\{f(x) \mid x \in [x_{i-1}, x_i]\} \text{ and } m_i = \inf\{f(x) \mid x \in [x_{i-1}, x_i]\}.$$

Remarks 1.7

(a) $L(f, P) \le U(f, P)$ for all partitions P (why?).

(b) Definition 1.6 makes no sense if f is not bounded on $[a, b]$ (why?).

Example Take f: $[-1, 1]\to$R; $f(x) = x^2$ and $P = \{-1, 0, \frac{1}{2}, 1\}$.

Lemma 1.8
Suppose f: $[a, b]\to$R is bounded, and let P and Q be partitions of $[a, b]$. Then
(a) If Q is a refinement of P, then $L(f, Q) \geq L(F, P)$ and $U(f, Q) \leq U(f, P)$
(b) $L(f, P) \leq U(f, Q)$.

Proof
(a) If $Q = P$, then the result is immediate. Thus, assume $Q \neq P$. If Q is obtained from P by adding a single point y between x_{i-1} and x_i, then, with $m_i = \inf\{f(x) \mid x \in [x_{i-1}, x_i]\}$,

$$\begin{aligned} L(f, P) &= m_1\Delta x_1 + \dots + m_i\Delta x_i + \dots + m_n\Delta x_n \\ &= m_1\Delta x_1 + \dots + m_i(y-x_{i-1}) + m_i(x_i-y) + \dots + m_n\Delta x_n \\ &\leq m_1\Delta x_1 + \dots + t_i(y-x_{i-1}) + s_i(x_i-y) + \dots + m_n\Delta x_n = L(f, Q), \end{aligned}$$

where $s_i = \inf\{f(x) \mid x \in [y, x_i]\}$ and $t_i = \inf\{f(x) \mid x \in [x_{i-1}, y]\}$ (since $A \subset B \Rightarrow \inf B \leq \inf A$). If Q is obtained from P by adding m points, then repeat the above argument m times. A similar argument works for the upper sums.
(b) Using (a) we get:

$$L(f, P) \leq L(f, P\cup Q) \leq U(f, P\cup Q) \leq U(f, Q).$$ ❖

From now on, we continue to assume that f: $[a, b]\to$R is a *bounded* function.

Definition 1.9 Define the **upper and lower integrals of f on $[a, b]$** by

$$L(f) = \sup\{L(f, P) \mid P \text{ a partition of } [a, b]\}$$

and

$$U(f) = \inf\{U(f, P) \mid P \text{ a partition of } [a, b]\}.$$

In words, the lower integral is the sup of all the lower sums, and the upper integral is the inf of all the upper sums.

Note It follows from Lemma 1.8(b) that $L(f) \leq U(f)$ (Exercise Set 1). Thus, if P and Q are any two partitions of $[a, b]$, then

$$\boxed{L(f, P) \leq L(f) \leq U(f) \leq U(f, Q).}$$

Definition 1.10 The bounded function f: $[a, b]\to$R is **Riemann integrable** if $L(f) = U(f)$, and we then write

$$\int_a^b f(x)\, dx = L(f) = U(f).$$

We sometimes write $f \in \mathcal{R}[a, b]$, where $\mathcal{R}[a, b]$ is the set of all Riemann integrable functions f: $[a, b]\to$R.

Note If f is integrable on $[a, b]$, then, ∀ partitions P, Q of $[a, b]$,

$$L(f, P) \le \int_a^b f(x)\, dx \le U(f, Q).$$

Exercise Set 1

1. Prove that, if P is any partition of $[a, b]$ and if $\varepsilon > 0$, then there exists a refinement Q of P with $\Delta Q < \varepsilon$.

2. Prove that $L(f) \le U(f)$ for any bounded function $f: [a, b] \to \mathbb{R}$.

3. Show that $f: [0, 1] \to \mathbb{R}$ given by

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

is not Riemann integrable.

4. Show that constant functions are Riemann integrable.

5. Prove that, if f and g are bounded functions on $[a, b]$, then

$$L(f) + L(g) \le L(f+g) \le U(f+g) \le U(f) + U(g).$$

Deduce that, if f and g are Riemann integrable, then so is $f+g$, and

$$\int_a^b [f(x) + g(x)]\, dx = \int_a^b f(x)\, dx + \int_a^b g(x)\, dx .$$

2. Integrable Functions

where we are reassured by the fact that the functions we know and love (and even some we despise) are integrable.

Lemma 2.1 (Technical Criterion for Integrability)
$f: [a, b] \to \mathbb{R}$ is integrable iff for all $\varepsilon > 0$, there exists a partition P of $[a, b]$ such that

$$U(f, P) - L(f, P) < \varepsilon.$$

Proof

⇒ If f is integrable, then $U(f) = L(f) = I(f)$, say. In other words, $I(f) = \sup_P\{L(f, P)\} = \inf_P\{U(f, P)\}$. Thus, there exist partitions P_1 and P_2 of $[a, b]$ such that

$$I(f) - \varepsilon/2 < L(f, P_1) \le U(f, P_2) < I(f) + \varepsilon/2.$$

Using $P = P_1 \cup P_2$, we get

$$I(f) - \varepsilon/2 < L(f, P_1) \le L(f, P) \le U(f, P) \le U(f, P_2) < I(f) + \varepsilon/2,$$

from which it follows that $U(f, P) - L(f, P) < \varepsilon$.

⇐ Assume that for all $\varepsilon > 0$, there exists a partition P of $[a, b]$ such that

$$U(f, P) - L(f, P) < \varepsilon.$$

Then, since

$$L(f, P) \le L(f) \le U(f) \le U(f, P),$$

we get $U(f) - L(f) < \varepsilon$ for all ε, whence $U(f) = L(f)$. ❖

We now try to answer some questions such as "which functions are integrable?"

The following is a most surprising result.

Theorem 2.2 (Monotone Functions)
If f: $[a, b] \to \mathrm{R}$ is monotone, then f is integrable.

Proof Assume wlog that f is increasing. First note that f is automatically bounded, since $f(a) \le f(x) \le f(b)$ for all $x \in [a, b]$, so that the upper and lower integrals exist. Next notice that, if P is any partition of $[a, b]$, then

$$\begin{aligned} U(f, P) - L(f, P) &= \Sigma_i M_i \Delta x_i - \Sigma_i m_i \Delta x_i \\ &= \Sigma_i (M_i - m_i) \Delta x_i \\ &= \Sigma_i (f(x_i) - f(x_{i-1})) \Delta x_i \quad \text{(since } f\uparrow) \\ &\le \Sigma_i (f(x_i) - f(x_{i-1})) \Delta P \quad \text{(recall that } \Delta P \text{ is the mesh of } P) \\ &= (f(b) - f(a)) \Delta P. \end{aligned}$$

Let $\varepsilon > 0$. By the lemma, it suffices to find a partition P with $U(f, P) - L(f, P) < \varepsilon$. However, by the above inequality, any partition with mesh $< \varepsilon/[f(b) - f(a)]$ will do. ❖

Proposition 2.3 (Continuous Functions)
If f: $[a, b] \to \mathrm{R}$ is continuous, then f is integrable.

Proof Since f is continuous on the closed interval $[a, b]$, it is uniformly continuous there. Let $\varepsilon > 0$. Then there exists $\delta > 0$ such that $|x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon/(b-a)$. If P is any partition with mesh $< \delta$, then noting that, by the Extreme Value Theorem, f attains its bounds on each subinterval, we have

$$\begin{aligned} U(f, P) - L(f, P) &= \Sigma_i f(C_i) \Delta x_i - \Sigma_i f(c_i) \Delta x_i, \text{(where } f(c_i) = m_i \text{ and } f(C_i) = M_i) \\ &= \Sigma_i [f(C_i) - f(c_i)| \Delta x_i \\ &< \Sigma_i \frac{\varepsilon}{b-a} \Delta x_i \\ &= \frac{\varepsilon}{b-a}(b-a) = \varepsilon. \end{aligned}$$ ❖

Definition 2.4 f: $[a, b] \to \mathrm{R}$ is **piecewise continuous** if it is continuous at all but a finite number of points.

Theorem 2.5 (Piecewise Continuous Functions)
If f: $[a, b] \to \mathrm{R}$ is piecewise continuous, then f is integrable.

Proof We do induction on the number d of points of discontinuity. If $d = 0$, then the result follows from Proposition 2.3. Thus, assume the result true for functions with $\le d$ points of discontinuity, and assume that f has $d+1$ points of discontinuity. Let one of these points be c. Let $\varepsilon > 0$, and let M be an upper bound of $|f(x))|$.

***Case* 1** c is an interior point of $[a, b]$.

Choose partitions P_1 of $[a, c-\varepsilon/12M]$ and P_2 *of* $[c+\varepsilon/12M, b]$ such that

$$U(f, P_1) - L(f, P_1) < \varepsilon/3 \quad \text{and} \quad U(f, P_2) - L(f, P_2) < \varepsilon/3.$$

(Why can we do this?). Now let $P = P_1 \cup \{c\} \cup P_2$. Then

$$U(f, P) - L(f, P) = U(f, P_1) - L(f, P_1) + \frac{\varepsilon}{6M}(2M) + U(f, P_2) - L(f, P_2)$$

$$< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

***Case* 2** c is an end point of $[a, b]$.
This is left as an exercise. ❖

Exercise Set 2

1. For each of the following functions, find a partition P such that $U(f, P) - L(f, P) < 0.1$

(a) f: $[-1, 1] \to$'R; $f(x) = |x|$ **(b)** f: $[-1, 1] \to \mathbb{R}$; $f(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ -1 & \text{if } x < 0 \end{cases}$

(c) f: $[-\pi, \pi] \to \mathbb{R}$; $f(x) = \sin x$

2. A Sequential Criterion for Integrability
Prove that, if there exists a sequence (P_n) of partitions of $[a, b]$ with the property that

$$U(f, P_n) - L(f, P_n) \to 0 \text{ as } n \to \infty,$$

then $f \in \mathcal{R}[a, b]$ with

$$\int_a^b f(x)\, dx = \lim_{n \to \infty} U(f, P_n) = \lim_{n \to \infty} L(f, P_n).$$

3. Complete the proof of Theorem 2.5.

4. Dirichlet's Function Revisited (yes!) Prove that the function f: $[0, 1] \to \mathbb{R}$ given by

$$f(x) = \begin{cases} \frac{1}{q} & \text{if } x \in \mathbb{Q} \text{ and } x = \frac{p}{q} \text{ in lowest terms} \\ 0 & \text{if } x \in \mathbb{R}\text{-}\mathbb{Q} \end{cases}$$

is Riemann integrable, with $\int_0^1 f(x)\, dx = 0$.

5. Let (x_n) be a sequence in $[a, b]$ with $x_n \to x$ as $n \to \infty$. Suppose that f: $[a, b] \to \mathbb{R}$ is bounded and continuous except possibly at the points of (x_n). Prove that $f \in \mathcal{R}[a, b]$.

3. Some Technical Results

in which we discuss "convenient" partitions, left sums, right sums, and the so-called "Riemann" sums that fill our baby calculus textbooks, and which will also lead to numerous useful consequences.

Lemma 3.1 (Adding a Single Point to a Partition)
Let $|f(x)| < K$ for all $x \in [a, b]$, and let P and P' be partitions of $[a, b]$ such that P' is obtained from P by adding a singe point. Then

$$U(f, P) - U(f, P') < 2K\Delta x, \quad \text{and} \quad L(f, P') - L(f, P) < 2K\Delta x,$$

where Δx is the mesh of P.

Proof If $[c, d]$ is the subinterval containing the point x of $P' - P$, then these differences are

$$U(f, P) - U(f, P') = (d-c)M - [(x-c)M_1 + (d-x)M_2]$$

(where M, M_1 and M_2 are the supremum of f on the appropriate subinterval)

$$< (d-c)K + K[(x-c) + (d-x)] \quad \text{(since } |M_i| < K, \text{ so that } -M_i \le K)$$
$$= 2(d-c)K \le 2\Delta x K. \text{ ❖}$$

Corollary 3.2 Let $|f(x)| < K$ for all $x \in [a, b]$, and let P and P' be partitions of $[a, b]$ such that P' is obtained from P by adding n points. Then

$$U(f, P) - U(f, P') < 2nK\Delta x, \quad \text{and} \quad L(f, P') - L(f, P) < 2nK\Delta x,$$

where Δx is the mesh of P.

Proposition 3.3 (Any Small Enough Partition is OK)
If f: $[a, b]\to$R is Riemann integrable, then , given any $\varepsilon > 0$, there exists $\delta>0$ such that if P is any partition with $\Delta x < \delta$, one has $\left| U(f, P) - \int_a^b f(x)\, dx \right| < \varepsilon$, and similarly for the lower sum.

Proof First, since $f \in \mathcal{R}[a, b]$, we can choose a partition Q of $[a, b]$ with N subdivisions, say, such that

$$U(f, Q) - L(f, Q) < \frac{\varepsilon}{3}.$$

Since f is bounded, choose K with $|f(x)| < K$ for all $x \in [a, b]$. Now choose $\delta = \dfrac{\varepsilon}{6NK}$.

Let P be any partition such that $\Delta x < \delta = \dfrac{\varepsilon}{6NK}$. Then, with $R = Q \cup P$, certainly

$$U(f, R) - L(f, R) < \frac{\varepsilon}{3} \qquad \text{by Lemma 1.8(a)}$$

Now, R is obtained from P by adding at most N points. Thus, by the lemma,

$$U(f, P) - U(f, R) < 2NK\Delta x = 2NK\frac{\varepsilon}{6NK} = \frac{\varepsilon}{3},$$

and similarly for the lower sums. Thus,

$$U(f, P) - L(f, P) = U(f, P) - U(f, R) + U(f, R) - L(f, R) + L(f, R) - L(f, P)$$
$$< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

The result now follows. ❖

Note This justifies the use of partitions with evenly spaced subintervals in Riemann sums, and we obtain the following very nice corollary.

Corollary 3.4 (In Which Convergence of Numerical Calculations is Assured)
Let $f \in \mathcal{R}[a, b]$. Then
(a) If (P_n) is any sequence of partitions with $\Delta(P_n) \to 0$ as $n \to \infty$, then

$$\lim_{n \to \infty} U(f, P_n) = \lim_{n \to \infty} L(f, P_n) = \int_a^b f(x)\, dx .$$

In particular, we can take $P_n = \{a, a+\Delta_n, a+2\Delta_n, \ldots, a+n\Delta_n = b\}$, where $\Delta_n = (b-a)/n$.

(b) (Convergence of Left-and Right Sums)
If $P = \{x_0 = a, x_1, x_2, \ldots, x_n = b\}$ is any partition of $[a, b]$, define the **left** and **right Riemann sum** of f (associated with P) by

$$\boxed{Left(f, P) = \sum_{i=1}^{n} f(x_{i-1})\, \Delta x_i} \qquad \boxed{Right(f, P) = \sum_{i=1}^{n} f(x_i)\, \Delta x_i} .$$

left sum — right sum

Then, if (P_n) is any sequence of partitions with $\Delta(P_n) \to 0$ as $n \to \infty$, one has

$$\lim_{n \to \infty} Left(f, P_n) = \lim_{n \to \infty} Right(f, P_n) = \int_a^b f(x)\, dx .$$

(c) (Riemann Sum)
If $P = \{x_0 = a, x_1, x_2, \ldots, x_n = b\}$ is any partition of $[a, b]$, and if $c_i \in [x_{i-1}, x_i]$ is an arbitrary point for each i, the associated **Riemann sum** of f is given by

$$R(f, P, \{c_i\}) = \sum_{i=1}^{n} f(c_i)\, \Delta x_i .$$

Then, given any $\varepsilon > 0$, there exists $\delta > 0$ such that if P is any partition with $\Delta x < \delta$, one has $\left| R(f, P, \{c_i\}) - \int_a^b f(x)\, dx \right| < \varepsilon$. In particular, if (P_n) is any sequence of partitions with $\Delta(P_n) \to 0$ as $n \to \infty$, one has, for arbitrary choices of the associated points $\{c_i\}_n$,

$$\lim_{n \to \infty} R(f, P_n, \{c_i\}_n) = \int_a^b f(x)\, dx .$$

Proof
(a) follows directly from the proposition by choosing $N \in \mathbb{N}$ such that $\Delta(P_n) < \delta$ for $n \geq N$.
(b) follows from the inequalities

$$L(f, P_n) \leq Left(f, P_n) \leq U(f, P_n) \text{ and } L(f, P_n) \leq Right(f, P_n) \leq U(f, P_n),$$

and the sandwich rule.
(c) follows from the inequality

$$L(f, P) \leq R(f, P, \{c_i\}) \leq U(f, P),$$

so that

$$\left| R(f, P, \{c_i\}) - \int_a^b f(x)\, dx \right| \le \left| U(f, P) - \int_a^b f(x)\, dx \right| ,$$

giving the result. ❖

Exercise Set 3

1. The **Trapezoid sum** associated with a partition $P = \{x_0 = a, x_1, x_2, \ldots, x_n = b\}$ is given by

$$T(f, P) = \frac{1}{2}\sum_{i=1}^{n}\big(f(x_i) + f(x_{i-1})\big)\, \Delta x_i \quad .$$

Use a sketch to explain why it is called a "trapezoid sum," and show that, if $f \in \mathcal{R}[a, b]$, and if (P_n) is any sequence of partitions with $\Delta(P_n) \to 0$ as $n \to \infty$, one has

$$\lim_{n \to \infty} T(f, P_n) = \int_a^b f(x)\, dx \ .$$

2. (Riemann Sum Criterion for Integrability) Let f be bounded on $[a, b]$. Prove that $f \in \mathcal{R}[a, b]$ with integral $I = \int_a^b f(x)\, dx$. iff given any $\varepsilon > 0$, there exists $\delta > 0$ such that if P is any partition with $\Delta P < \delta$, one has

$$|R(f, P, \{c_i\}) - I| < \varepsilon$$

for all choices of $c_i \in [x_{i-1}, x_i]$.(Note that we have already proved one direction.)

3. Suppose that f is continuous on $[a, b]$ and differentiable on (a, b). Use the Mean Value Theorem to prove that, if P is any partition of $[a, b]$, then there exist $c_i \in [x_{i-1}, x_i]$ such that

$$R(f', P, \{c_i\}) = f(b) - f(a).$$

Deduce that if F is continuous on $[a, b]$ and differentiable on (a, b), with $F' = f$,

$$\int_a^b f(x)\, dx = F(b) - F(a).$$

(In other words: if F is any antiderivative of f on $[a, b]$, then the integral is given by the above formula.)

4. Let $[a, b]$ be an interval, and let $P_n = \{x_i\}$ be the partition mentioned in Corollary 3.4(a) , namely $x_i = a + \dfrac{i(b-1)}{n}$. Give an example of a bounded function f: $[a, b] \to \mathrm{R}$ with such that the left- and right sums associated with the partition P_n converge to the same number and yet f is not integable. Justify your assertions.

❖❖ Extra Credit ❖❖

A subset $A \subset \mathrm{R}$ has **measure zero** if, $\forall \varepsilon > 0$, there exist a countable collection $\{(a_n, b_n)\}$ of open intervals with $A \subset \cup_n (a_n, b_n)$ and $\Sigma_n (b_n - a_n) \le \varepsilon$ for all n. Prove that the bounded function f: $[a, b] \to \mathbb{R}$ is Riemann integrable if the set of discontinuities of f in $[a, b]$ has measure zero. (In fact, the converse is also true: f: $[a, b] \to \mathbb{R}$ is Riemann integrable iff the set of discontinuities of f in $[a, b]$ has measure zero.)

4. Properties of the Integral

which were listed for you in your previous calculus classes, but never justified.

Theorem 4.1 (Properties of the Integral)

(a) Linearity Let f and $g \in \mathcal{R}[a, b]$, and $c, d \in \mathbb{R}$. Then $cf + dg \in \mathcal{R}[a, b]$ and

$$\int_a^b (cf + dg)(x)\, dx = c\int_a^b f(x)\, dx + d\int_a^b g(x)\, dx .$$

(b) Preservation of Order If $f, g \in \mathcal{R}[a, b]$, with $f(x) \le g(x)\ \forall x \in [a, b]$, then

$$\int_a^b f(x)\, dx \le \int_a^b g(x)\, dx .$$

(c) Partitioning of the Range of Integration Let $a < c < b$. *Then* f is integrable on $[a, b]$ iff it is integrable on $[a, c]$ and $[c, b]$. When this happens, we have

$$\int_a^b f(x)\, dx = \int_a^c f(x)\, dx + \int_c^b f(x)\, dx .$$

Proof

(a) We use the criterion in Exercise Set 3 # 2. Thus, let $\varepsilon > 0$. Then since f and g are integrable, there exist δ_1 and δ_2 such that, for all partitions P_1 and P_2 of $[a, b]$ with $\Delta P_i < \delta_i$ (i = 1, 2) and for all choices of $\{x_i\}$, one has

$$\left| R(f, P_1, \{x_i\}) - \int_a^b f(x)\, dx \right| < \frac{\varepsilon}{2(|c|+1)} \quad \text{and} \quad \left| R(g, P_2, \{x_i\}) - \int_a^b g(x)\, dx \right| < \frac{\varepsilon}{2(|d|+1)} .$$

Now let $\delta = \min\{\delta_1, \delta_2\}$, and let P be any partition with $\Delta P < \delta$. Then

$$\left| R(cf+dg, P, \{y_i\}) - \left(c\int_a^b f(x)\, dx + d\int_a^b g(x)\, dx \right) \right|$$

$$=^{\dagger} \left| R(cf, P, \{y_i\}) + R(dg, P, \{y_i\}) - \left(c\int_a^b f(x)\, dx + d\int_a^b g(x)\, dx \right) \right|$$

$$= \left| c\left(R(f, P, \{y_i\}) - \int_a^b f(x)\, dx \right) + d\left(R(g, P, \{y_i\}) - \int_a^b g(x)\, dx \right) \right|$$

† That's what is nice about using Riemann sums instead of infs and sups here.

$$\leq |c| \left| R(f, P, \{y_i\}) - \int_a^b f(x)\, dx \right| + |d| \left| R(g, P, \{y_i\}) - \int_a^b g(x)\, dx \right|$$

$< \varepsilon$, by choice of δ.

This shows not only that $cf + dg$ is integrable, but also that its integral is $c\int_a^b f(x)\, dx + d\int_a^b g(x)\, dx$, as claimed.

(b) Since order is preserved by suprema and infima, it is also preserved by passage to Darboux sums, and hence to Riemann integrals.

(c) Assume $f \in \mathcal{R}[a, b]$. We show that f is integrable of $[a, c]$ and $[c, b]$, using the criterion in Lemma 2.1. Thus, let $\varepsilon > 0$. Since f is integrable on $[a, b]$, there exists a partition P of $[a, b]$ such that

$$U(f, P) - L(f, P) < \varepsilon.$$

We can assume wlog that $c \in P$. (If not, throw it in without effecting the inequality—why?) Then, we can write $P = Q \cup R$, where Q is a partition of $[a, c]$ and R is a partition of $[a, b]$. We have

$$\begin{aligned} U(f, Q) - L(f, Q) &\leq U(f, Q) - L(f, Q) + U(f, R) - L(f, R) \\ &= U(f, P) - L(f, P) < \varepsilon, \end{aligned}$$

showing that f is integrable on $[a, c]$. Similarly, it is integrable of $[c, b]$.

Conversely, if f is integrable of $[a, c]$ and $[c, b]$, then we shall use the criterion in Exercise Set 3 # 2 to show that it is integrable on $[a, b]$, with the integral as stated. Thus, let $\varepsilon > 0$. Since f is integrable on $[a, c]$ and $[c, b]$, there exist, there exist δ_1 and δ_2 such that, for all partitions P_1 of $[a, c]$ and P_2 of $[c, b]$ with $\Delta P_i < \delta_i$ (i = 1, 2) and for all choices of $\{x_i\}$, one has

$$\left| R(f, P_1, \{r_i\}) - \int_a^c f(x)\, dx \right| < \frac{\varepsilon}{3} \quad \text{and} \quad \left| R(f, P_2, \{s_i\}) - \int_c^b f(x)\, dx \right| < \frac{\varepsilon}{3}$$

for all choices of $\{r_i\}$ and $\{s_i\}$. Now let $\delta = \min\{\delta_1, \delta_2, \varepsilon/(6M)\}$, where M is an upper bound of $|f|$ on $[a, b]$, and let P be any partition of $[a, b]$ with $\Delta P < \delta$, and let $\{z_i\}$ be points with $z_i \in [x_{i-1}, x_i]$, as usual.

Case 1* $c \in P$.

In this case, P breaks up as two partitions P_1 of $[a, c]$ and P_2 of $[c, b]$ with $\Delta P_i < \delta \leq \partial_i$ for each i. Thus,

$$\begin{aligned} &\left| R(f, P, \{z_i\}) - \left(\int_a^c f(x)\, dx + \int_c^b f(x)\, dx \right) \right| \\ = &\left| R(f, P_1, \{z_i\}) + R(f, P_2, \{z_i\}) - \left(\int_a^c f(x)\, dx + \int_c^b f(x)\, dx \right) \right| \end{aligned}$$

* This seems to be the only case that Kosmala considers. He (conveniently) neglects the harder case. (The sum of his two Reimann sums is not necessarily the Riemann sum of the union of his two partitions if c is not already a partition point.)

$$\le \left| R(f, P_1, \{z_i\}) - \int_a^c f(x)\, dx \right| + \left| R(f, P_2, \{z_i\}) + \int_c^b f(x)\, dx \right|$$

$$< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} < \varepsilon.$$

Case 2. $c \notin P$.

Here, let $Q = P \cup \{c\}$, so that $\Delta Q < \delta$ as well. Assume $c \in [x_{i-1}, x_i]$, and let $P_1 = \{x_0, \ldots, x_{i-1}\}$—a partition of $[a, x_{i-1}]$, and let $P_2 = \{x_i, \ldots, x_n\}$—a partition of $[x_i, b]$. Then define a Riemann sum R associated with Q of f over $[a, b]$ as follows.

$$B = R(f, P_1, \{z_i\}) + R(f, P_2, \{z_i\}) + (c - x_{i-1})f(c) + (x_i - c)f(c)$$

$$= R(f, P_1, \{z_i\}) + R(f, P_2, \{z_i\}) + (x_i - x_{i-1})f(c)$$

Then

$$|R(f, P, \{z_i\}) - B| = (x_i - x_{i-1})|f(c) - f(z_i)| \le 2\delta M < \varepsilon/3,$$

by choice of δ. Further,

$$\left| R(f, P, \{z_i\}) - \left(\int_a^c f(x)\, dx + \int_c^b f(x)\, dx \right) \right|$$

$$= \left| R(f, P, \{z_i\}) - B + B - \left(\int_a^c f(x)\, dx + \int_c^b f(x)\, dx \right) \right|$$

$$\le |R(f, P, \{z_i\}) - B| + \left| B - \left(\int_a^c f(x)\, dx + \int_c^b f(x)\, dx \right) \right|$$

$$= |R(f, P, \{z_i\}) - B|$$

$$+ \left| R(f, P_1, \{z_i\}) + (c - x_{i-1})f(c) + R(f, P_2, \{z_i\}) + (x_i - c)f(c) - \left(\int_a^c f(x)\, dx + \int_c^b f(x)\, dx \right) \right|$$

$$\le |R(f, P, \{z_i\}) - B|$$

$$+ \left| R(f, P_1, \{z_i\}) + (c - x_{i-1})f(c) - \int_a^c f(x)\, dx \right|$$

$$+ \left| R(f, P_2, \{z_i\}) + (x_i - c)f(c) - \int_c^b f(x)\, dx \right|$$

$$< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

since the last two terms are Riemann sums for f over $[a, c]$ and $[c, b]$ respectively. (Whew!) ❖

Theorem 4.2 (Integral Mean Value Theorem)
Let f, g and $fg \in \mathcal{R}[a, b]$ with f continuous and either $g \geq 0$ or $g \leq 0$.[††] Then there exists $c \in [a, b]$ with

$$\int_a^b f(x)g(x)\,dx = f(c)\int_a^b g(x)\,dx .$$

Proof We do the case $g \geq 0$, leaving the other case to the exercises. Define H: $[a, b] \to \mathbb{R}$ by

$$H(x) = f(x)\int_a^b g(x)\,dx .$$

Then H, being a multiple of f, is continuous. If m and M are, respectively, a lower and upper bound of f, then $m \leq f(x) \leq M$ for all x, so that, since $g(x) \geq 0$,

$$mg(x) \leq f(x)g(x) \leq Mg(x)$$

for all x. Since f in continuous on $[a, b]$, there exist $r, s \in [a, b]$ with $f(r) = m$ and $f(s) = M$. Thus,

$$f(r)g(x) \leq f(x)g(x) \leq f(s)g(x).$$

Integrating, 4.1(b) gives

$$f(r)\int_a^b g(x)\,dx \leq \int_a^b f(x)g(x)\,dx \leq f(s)\int_a^b g(x)\,dx$$

That is,

$$H(r) \leq \int_a^b f(x)g(x)\,dx \leq H(s).$$

By the Intermediate Value Theorem applied to H, there exists $c \in [a, b]$ with

$$H(c) = \int_a^b f(x)g(x)\,dx ,$$

and this is the result. ❖

With $g(x) = 1$, Theorem 4.2 reduces to

Corollary 4.3 (Traditional IMVT)
Let f be continuous on $[a, b]$, then there exists $c \in [a, b]$ with

$$\int_a^b f(x)dx = f(c)\cdot(b-a).$$

Exercise Set 4

1. Show that, if f and g are integrable on $[a, b]$, then so are $\max\{f, g\}$ and $\min\{f, g\}$. (Hint: if $[x, y]$ is any interval, then $\min\{\inf(f), \inf(g)\} \leq \min\{f, g\} \leq \min\{\sup(f), \sup(g)\}$.)

2. (a) Let $f \in \mathcal{R}[a, b]$. Prove that $|f| \in \mathcal{R}[a, b]$, and

---

[††] Actually, we need only specify that f and g are integrable. The extra-credit problem would imply that fg is integrable, since the union of two sets with measure zero still has measure zero.

$$\left| \int_a^b f(x)\, dx \right| \le \int_a^b |f(x)|\, dx .$$

(b) Give an example of f with $|f|$ integrable, but not f.

3. Suppose that f and g are Riemann integrable on $[a, b]$ such that $f(x) = g(x)$ for all x except for finitely many values. Show that

$$\int_a^b f(x)\, dx = \int_a^b g(x)\, dx .$$

4. **(a)** Show by means of an example that the assumption that "$g \ge 0$ or $g \le 0$" is necessary in Theorem 4.2.

(b) Show by means of an example that the assumption that f is continuous is necessary in Theorem 4.2.

(c) Show that Theorem 4.2 continues to hold if $g \le 0$.

5. The Fundamental Theorem of Calculus & Other Results,

which we have already previewed in Exercise Set 3, number 3, and where, displeased by the odd approach in Wade, we take the liberty of deviating somewhat.

Definition 5.1 If f: $[a, b] \to \mathrm{R}$ is any function, then a **primitive,** or **antiderivative, of f** is a function F: $[a, b] \to \mathrm{R}$, such that F is continuous on $[a, b]$ and such that $F'(x)$ exists and equals $f(x)$ at all but at most a finite number of points in (a, b).

Note In the homework, you will prove that the primitive is unique up to a constant.

Theorem 5.2 (Fundamental Theorem of Calculus)
Let $f \in \mathcal{R}[a, b]$, and define F: $[a, b] \to \mathrm{R}$ by

$$F(x) = \int_a^x f(t)\, dt .$$

Then:

(a) F is continuous on $[a, b]$.

(b) If f is continuous at $x \in (a, b)$, then F is differentiable at x, with $F'(x) = f(x)$. (It follows that, if f is continuous at all but a finite number of points, then F is a primitive of f.)

(c) If f is any integrable function (no assumptions about continuity here) and G is any primitive of f, then

$$\int_a^b f(x)\, dx = G(b) - G(a).$$

Proof

(a) Let M be such that $|f(x)| \le M$ for all $x \in [a, b]$. Then, for $x \le y$ in $[a, b]$, one has

$$|F(y) - F(x)| = \left|\int_x^y f(t)\,dt\right| \le \int_x^y |f(t)|\,dt \le \int_x^y M\,dt = M|y-x|,$$

showing that F is Lipschitz, and hence (uniformly) continuous, on $[a, b]$.

(b) Let $\varepsilon > 0$. If f is continuous at $x \in (a, b)$, then $\exists\ \delta>0$ such that $0 < |t-x| < \delta \Rightarrow |f(t) - f(x)| < \varepsilon$. Now, if $|h| < \delta$, then

$$\begin{aligned}\left|\frac{F(x+h) - F(x)}{h} - f(x)\right| &= \left|\frac{1}{h}\int_x^{x+h} f(t)\,dt - f(x)\right| \\ &= \left|\frac{1}{h}\int_x^{x+h} (f(t) - f(x))\,dt\right| \\ &\le \frac{1}{h}\int_x^{x+h} |f(t) - f(x)|\,dt \qquad \text{(this assumes } h > 0\text{)}^* \\ &< \frac{1}{h}\int_x^{x+h} \varepsilon\,dt \qquad \text{(since } h < \delta\text{)} \\ &= \varepsilon.\end{aligned}$$

(c) Let G be a primitive of f, so that $G'(x)$ exists and equals $f(x)$ for all except finitely many points in (a, b). Then, since f is integrable, there is a partition $P = \{x_0, \ldots, x_n\}$ of $[a, b]$ that contains all those questionable points, and is such that, for all choices of $z_i \in [x_{i-1}, x_i]$, one has

$$\left|R(f, P, \{z_i\}) - \int_a^b f(x)\,dx\right| < \varepsilon \qquad \ldots(1)$$

Since G is continuous on each $[x_{i-1}, x_i]$ and differentiable on each (x_{i-1}, x_i), we can apply the MVT to obtain points $c_i \in (x_{i-1}, x_i)$ with

$$G'(c_i)\Delta x_i = G(x_i) - G(x_{i-1}).$$

But $G'(c_i) = f(c_i)$, so that

$$f(c_i)\Delta x_i = G(x_i) - G(x_{i-1}).$$

Summing,

$$R(f, P, \{c_i\}) = G(b) - G(a).$$

Thus, by (1),

$$\left|G(b) - G(a) - \int_a^b f(x)\,dx\right| < \varepsilon.$$

Since ε is arbitrary, we obtain $\int_a^b f(x)\,dx = G(b) - G(a)$, as required. ❖

Examples 5.3

A. Let us find a primitive for f: $[0, \pi]\to\mathbb{R}$; $f(x) = \sin x$.

B. Now consider the unit step function $u(x) = \begin{cases} 0 & \text{if } x < 0 \\ 1 & \text{if } x \ge 0 \end{cases}$.

---

* We get the same result two steps later if $h < 0$.

C. f: $[-1, 1]\to\mathrm{R}$; $f(x) = \begin{cases} x^2 & \text{if } x \geq 0 \\ x^3 & \text{if } x < 0 \end{cases}$

D. f: $(0, +\infty)\to\mathrm{R}$; $f(x) = 1/x$.

Corollary 5.4 (Fancy Fundamental Theorem)
Let g: $[c, d]\to[a, b]$ be differentiable at x and let f: $[a, b]\to\mathrm{R}$ be integrable over $[a, b]$ and continuous at $g(x)$. Then, with

$$F(x) = \int_a^{g(x)} f(t)\, dt\ ,$$

F is differentiable at x, and

$$F'(x) = f(g(x))g'(x).$$

Proof Define H: $[a, b]\to\mathrm{R}$ by $H(u) = \int_a^u f(t)\, dt$. Then $F(x) = H(g(x)) = H\circ g(x)$. Since g is differentiable at x and H is also differentiable at $g(x)$ (since f is continuous there, and so its integral is differentiable there, by the theorem), we can use the chain rule to get

$$F'(x) = H'(g(x))g'(x) = f(g(x))g'(x),$$

as required. ❖

Example 5.6 Evaluate $\dfrac{d}{dx}\displaystyle\int_a^{x^2+1} \sin t\, dt$.

Theorem 5.7 (Change-of-Variables; Substitution)
Let g: $[c, d]\to[a, b]$ be continuous on $[c, d]$ and differentiable on (c, d). Let f: $[a, b]\to\mathrm{R}$ be continuous . Then

$$\int_a^b f(u)\, du = \int_c^d f(g(x))g'(x)\, dx\ ,$$

provided the integral on the right exists.

Proof Let F: $[a, b]\to\mathrm{R}$ be a primitive of f. Then F is differentiable everywhere (since f is continuous everywhere). By the chain rule, $(F\circ g)'(x)$ exists and equals $F'(g(x))g'(x) = f(g(x))g'(x)$. In other words, $(F\circ g)$ is a primitive of the integrand on the right. So,

$$\int_a^b f(u)\, du = F(b) - F(a) = F\circ g(d) - F\circ g(c) = \int_c^d f(g(x))g'(x)\, dx\ ,$$

as required. ❖

Example

Use substitution to evaluate $\displaystyle\int_0^1 \frac{x}{(x^2+1)^2}\, dx$.

Exercise Set 5

1. Find primitives for each of the following functions:

(a) f: $[-1, 1]\to R$; $f(x) = \lfloor x \rfloor$

(b) f: $[0, n]\to R$; $f(x) = [x]$ (the "floor" function)

(c) f: $[-1, 1]\to R$; $\delta_\varepsilon(x) = \begin{cases} 0 & \text{if } x < -\varepsilon \\ 1 & \text{if } -\varepsilon \le x \le \varepsilon \\ 0 & \text{if } x \ge \varepsilon \end{cases}$ where δ is a fixed positive number.

2. Uniqueness of the Primitive up to a Constant Use the Fundamental Theorem of Calculus to prove that any two primitives of f differ by a constant. That is, if F and G are primitives of f, then there exists a constant C such that $F(x) = G(x) + C$. (In particular, F and G are differentiable at the same points.)

3. The **error function**, erf: $R\to R$, is defined by the formula

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\, dt \ .$$

(Just believe in the existence of the exponential function for now—we'll prove it in the next section!) Use fundamental theorem of calculus to evaluate each of the following.

(a) $\dfrac{d}{dx}\left[\dfrac{\sqrt{\pi}}{2}\operatorname{erf}(x)\right]$ (b) $\displaystyle\int e^{-x^2}\, dx$ (c) $\dfrac{d}{dx}\left[\operatorname{erf}(x^2 - 1)\right]$

4. Integration by parts Study Wade's Theorem 5.31 (Integrtion by Parts), and use it to do p. 131 #1(d), (e)

5. Evaluate:

(a) $\displaystyle\frac{d}{dx}\int_1^x \sqrt{2+t^3}\, dt$ **(b)** $\displaystyle\frac{d}{dx}\int_{x^2}^{x^3} \sqrt{t}\sin t\ \, dt$

6. The Logarithmic and Exponential Functions,

which Wade scatters through the exercises and which,up to this point in time, have never existed except in the fantasies of certain high-school teachers who, eager for results at the expense of understanding, have urged their students to accept, on blind faith, the existence of functions that were never defined and yet endowed with remarkable properties.

Definition 6.1 Define the **natural logarithm function** ln: $(0, +\infty)\to R$ by

$$\ln(x) = \int_1^x \frac{1}{t}\, dt \ .$$

Proposition 6.1 (Properties of the Natural Logarithm Function)
The function ln: $(0, +\infty)\rightarrow$R has the following properties:

(a) The natural logarithm is differentiable on $(0, +\infty)$ with $\frac{d}{dx}(\ln x) = \frac{1}{x}$

(b) $\ln 1 = 0$

(c) For all $x > 0$, $\boxed{1-\frac{1}{x} \leq \ln x \leq x-1}$ (with equality only when $x = 1$) .

(d) The natural logarithm is strictly increasing and bijective.

(e) $\ln xy = \ln x + \ln y$

(f) $\ln\left(\frac{1}{x}\right) = -\ln x$ for all $x > 0$

(g) $\ln\left(\frac{x}{y}\right) = \ln x - \ln y$ for all $x, y > 0$

Proof (a), (b) in class. For **(c)**, we use the fact that $1/t^2 < 1/t < 1$ for $t > 1$, and the reverse is true if $t < 1$, thus negating the change in signs. For **(d)**, we use the fact that the natural derivative is positive to show the increasing part. We then do **(e)** in class, and the rest is in the homework. ❖

Note For further properties, see the exercise set.

Corollary 6.2 (Existence of the Exponential Function)
Since the natural logarithm function is strictly increasing and bijective, it has a differentiable strictly increasing inverse,

exp: $R\rightarrow(0, +\infty)$

with the following properties:

(a) $\frac{d}{dx}(\exp(x)) = \exp(x)$

(b) $\exp(0) = 1$

(c) $\exp(x+y) = \exp(x) \times \exp(y)$

Proof (a) By general properties of inverses, $(f^{-1})'(x) = 1/f'(f^{-1}(x))$, so

$$\frac{d}{dx}(\exp(x)) = \frac{1}{1/\exp(x)} = \exp(x).$$

Rest in class. ❖

Note For further properties, see the exercise set. In particular:

$\ln(x^r) = r\ln x$ for all $r \in$ Q.

Definition 6.3 The **base of the natural logarithm,** ***e*****,** is given by the formula

$e = \exp(1)$.

Proposition 6.4 (Some properties of *e*)

(a) $\ln(e) = 1$

(b) If $r \in Q$, then $e^r = \exp(r)$

(c) $e = \lim_{n\to\infty}\left(1 + \frac{1}{n}\right)^n$, and the sequence in question is strictly increasing.

(d) $\exp(-x) = \frac{1}{\exp(x)}$ for all $x \in R$

(e) $\exp(x-y) = \frac{\exp(x)}{\exp(y)}$ for all $x, y \in R$

(f) $\exp(r \ln x) = x^r$ for all $x > 0$ and $r \in Q$.

Proof (a) $\ln(e) = \ln(\exp(1)) = 1$.

(b) $\ln(e^r) = r\ln(e) = r$, by Exercise 1 and part (a) above. Taking exp of both sides gives the result.

(c) Let $a_n = \left(1 + \frac{1}{n}\right)^n$. Consider the sequence $(\ln(a_n))$. One has

$$\ln(a_n) = n\ln(1+1/n) = \frac{\ln\left(1 + \frac{1}{n}\right)}{\left(\frac{1}{n}\right)} = \frac{\ln\left(1 + \frac{1}{n}\right) - \ln(1)}{\left(\frac{1}{n}\right)} \to \ln'(1) = 1$$

as $n\to\infty$. Hence,

$$a_n = \exp(\ln(a_n)) \to \exp(1) = e \text{ as } n\to\infty.$$

Finally, to show that the sequence is increasing, let

$f\colon (0, +\infty)\to R$ be given by $f(x) = \frac{\ln(1+x)}{x}$. Then $f'(x) = \frac{1}{x^2}\left(1 - \frac{1}{x+1} - \ln(1+x)\right)$

. It suffices to show that the quantity in parentheses is always negative (for positive x). But this is Proposition 6.1(c) (with x replaced by $x+1$).

(d)-(f) are homework. ❖

Notes

1. Part **(f)** says, with x replaced by e, that $e^r = \exp(r\ln e) = \exp(r)$ for all $r \in Q$. This suggests that $\exp(r)$ should be e^r for *all* $r \in R$, but we do not yet have a definition of the letter!

2. Similarly, it says that $x^r = \exp(r \ln x)$ for all $r \in Q$, and thus suggests that they should be the same for al*l* r, although, again, we don't have a definition of x^r.

Thus, we do the following.

Definition 6.5

(a) If $x \in R$, then define $\boxed{e^x = \exp(x)}$.

(b) If $x \in R$ and $a > 0$, then define $\boxed{a^x = \exp(x \ln a) = e^{x\ln a}}$.

By the above notes, these definitions agree with the usual definitions for rational exponents, and since exp is continuous (in fact, it is infinitely differentiable), we get, if

(r_n) is a sequence of rational numbers converging to the real number x, then $a^{r_n} \to a^x$ as $n \to \infty$, as we would like. Further, all the usual rules for exponents are satisfied.

Exercise Set 6

1. Prove each of the following identities:

(a) $\ln\left(\frac{1}{x}\right) = -\ln x$ for all $x > 0$

(b) $\ln\left(\frac{x}{y}\right) = \ln x - \ln y$ for all $x, y > 0$

(c) $\ln(x^r) = r\ln x$ for all $r \in \mathrm{R}$.

2. Complete the proof of Proposition 6.2 by showing that ln: $(0, +\infty) \to \mathrm{R}$ is surjective.

3. Prove each of the following identities:

(a) $\exp(-x) = 1/\exp(x)$ for all $x \in \mathrm{R}$

(b) $\exp(x-y) = \exp(x)/\exp(y)$ for all $x, y \in \mathrm{R}$

(c) $\exp(r \ln x) = x^r$ for all $x > 0$ and $r \in \mathrm{Q}$.

4. (a) Prove that $e^{-1} = \lim\limits_{n \to \infty} \left(1 - \frac{1}{n}\right)^n$ and that the sequence in question is strictly increasing.

(b) Use part (a) together with Proposition 6.4(c) to prove that $2.7 < e < 2.8$.

5. (a) By considering Riemann sums or otherwise, prove that, if $n \in \mathrm{N}^*$, then

$$\frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} < \ln n < 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n-1}.$$

(b) Show that $\left(1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} - \ln n\right)$ is a positive monotone decreasing sequence, and deduce from part (a) that $\lim\limits_{n \varnothing \infty} \left(1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} - \ln n\right) < 1$.[†]

6. Prove that, if $a > 0$, then the function f: $\mathrm{R} \to (0, +\infty)$ given by $f(x) = a^x$ is differentiable and invertible. Writing its inverse as $\log_a$: $(0, +\infty) \to \mathrm{R}$, show that the logarithm satisfies properties (b), (e), (f) and (g) of Proposition 6.1. What is its derivative?

7. (Do not hand in) Now that you know what arbitrary powers of a real number are, complete the proof of the power rule: $\frac{d}{dx}(x^a) = ax^{a-1}$, where $a \in \mathrm{R}$.

7. Improper Integrals

where we deal with functions that are unbounded, or with intervals of integration that are unbounded. First, the latter.

Definition 7.1 Let f: $[a, +\infty) \to \mathrm{R}$ be bounded be integrable on every closed interval $[a, M]$. Then we define

---

[†] The limit of this seqauence is denoted by $\gamma \approx 0.57721566$, and is called either Euler's constant or Mascheroni's constant in honor of Lorenzo Mascheroni (1750-1800) It is not known whethere γ is rational or not.

$$\int_a^\infty f(x)\,dx = \lim_{M\longrightarrow+\infty} \int_a^M f(x)\,dx \quad ,$$

if the limit exists (and is finite), and say that the **improper integral** $\int_a^\infty f(x)\,dx$ **converges**. Otherwise, we say it **diverges** (if the limit does not exist at all) or **diverges to ±∞** if that is what the limit does.

Note If f: $(-\infty, b]\to$R is bounded and integrable on every interval of the form $[-M, b]$, then we make a similar definition of $\int_{-\infty}^b f(x)\,dx$.

Examples 7.2

A. If $a > 0$, then $\int_a^\infty \frac{1}{x^2}\,dx$ converges. **B.** If $a > 0$, then $\int_a^\infty \frac{1}{x}\,dx$ diverges.

C. In general, if $a > 0$, then $\int_a^\infty \frac{1}{x^p}\,dx$ converges iff $p > 1$.

D. A. $\int_0^\infty xe^{-x^2}\,dx$ converges.

Theorem 7.3 (Tests for Convergence of Improper Integrals)
Comparison Test
If f and g are integrable on $[a, M]$ for every $M > a$, and suppose there exists $K \geq a$ such that $0 \leq f(x) \leq g(x)$ for $x \geq K$. Then
(a) if $\int_a^\infty f(x)\,dx$ diverges, so does $\int_a^\infty g(x)\,dx$;
(b) if $\int_a^\infty g(x)\,dx$ converges, then so does $\int_a^\infty f(x)\,dx$.
Absolute Convergence Test
If f is integrable on $[a, M]$ for every $M > a$, and if $\int_a^\infty |f(x)|\,dx$ converges, then so does $\int_a^\infty f(x)\,dx$, and $\left|\int_a^\infty f(x)\,dx\right| \leq \int_a^\infty |f(x)|\,dx$.

Proof of Comparison Test Let F: $[a, +\infty)\to$R be given by $F(M) = \int_K^\infty f(x)\,dx$, and similarly for G. one has, by preservation of order,

$0 \leq F(M) \leq G(M)$,

where F and G are increasing functions. If $\int_a^\infty f(x)\,dx$ diverges, then F is unbounded above, and hence diverges to $+\infty$, whence so does G. If $\int_a^\infty g(x)\,dx$ converges, then G is bounded above, whence so if F, showing that F has a limit as $M\to+\infty$.

Proof of Absolute Convergence Test Just write $f = f^+ - f^-$, so that $|f| = f^+ + f^-$. Since $0 \leq f^+ \leq |f|$ and $0 \leq f^- \leq |f|$, we see that $\int_a^\infty f^+(x)\,dx$ and $\int_a^\infty f^-(x)\,dx$ converge by the

Comparison Test, whence so does their difference, $f = f^+ - f^-$. The remaining inequality follows by preservation of order when we take the limit as $M \to +\infty$. ❖

Examples 7.4

A. $\int_0^\infty e^{-x^2}\, dx$ converges.

B. If $|f(t)| \le Ke^{at}$ for some constants K, a, then its Laplace transform, $F(s) = \int_0^\infty f(t)e^{-st}\, dt$ exists for $s > a$

Definition 7.5 Suppose f: $[a, b) \to \mathrm{R}$ is unbounded, but is bounded on every interval $[a, c]$ with $c < b$. Then define

$$\int_a^b f(x)\, dx = \lim_{c \to b^-} \int_a^c f(x)\, dx \quad ,$$

assuming the limit exists. If it does, we say the **improper integral converges**. (See Definition 7.1 for the rest of the terminology.) We make a similar definition for unbounded functions $(a, b] \to \mathrm{R}$.

Examples

A. $\int_0^1 x^{-p}\, dx$ for various p.

B. $\int_7^8 (8-x)^{-2/3}\, dx$

Note Theorem 8.3 continues to hold for these types of improper integrals also.

Definition 7.6 An integral $\int_a^b f(x)\, dx$ (with a and/or b possibly infinite) is **improper** if there exist finitely many points $a = c_1 < c_2 < \ldots < c_n = b$ such that $\int_{ci}^{ci+1} f(x)\, dx$ is an improper integral of one of the above types. The improper integral $\int_a^b f(x)\, dx$ **converges** if each of the integrals $J_i = \int_{ci}^{ci+1} f(x)\, dx$ converges, and **diverges to** $+\infty$ if each of the J_i either converges, or diverges to $+\infty$ (similarly for divergence to $-\infty$). Otherwise, it just plain **diverges.**

Examples in class.

Exercise Set 7

1. Wade, p. 139 #2

2. Wade, p. 139 #4.

3. Give an example of an integrable function f: $\mathrm{R} \to \mathrm{R}$ with $\int_a^\infty f(x)\, dx$ convergent, but $f(x) \not\to 0$ as $x \to +\infty$. (Hint: You may need to define this function piecewise with infinitely many pieces.)

4. The Gamma Function We define the **gamma function**, Γ: $(0, +\infty)\to\mathbb{R}$ by

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}\, dt\ .$$

(a) Prove that the integral converges. (Note that it is improper both at 0 (if $x < 1$) and at $+\infty$.)
(b) Prove that Γ enjoys the following properties:
(i) For all $x > 0$, $\Gamma(x+1) = x\Gamma(x)$.
(ii) For all $n \in \mathbb{N}$ with $n \geq 1$, $\Gamma(n) = (n-1)!$
5 Prove:
(a) $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.

(b) $\Gamma\left(\frac{2n+1}{2}\right) = \frac{(2n)!\sqrt{\pi}}{4^n n!}$

(c) For $x < 0$, $x \in \mathbb{Z}$, we define Γ inductively as follows. Assume Γ has already been defined on $(-n, -n+1)\cup\ldots\cup(-1, 0)\cup(0, +\infty)$. We then extend the definition to $(-n-1, -n)\cup(-n, -n+1)\cup\ldots\cup(-1, 0)\cup(0, +\infty)$ by defining

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

for $x \in (-n-1, -n)$. Sketch the graph of Γ.

8. Infinite Series

Definitions 8.1 Let $(a_n)_{n\geq 1}$ be a sequence. Then the **infinite series** $\sum_{k=1}^{\infty} a_k$ is just the *expression* $a_1 + a_2 + \ldots + a_k + \ldots$. The a_k are called the **terms** of the series, and a_k is the **kth term of the series**. The associated **sequence of partial sums** is the sequence (S_n) where $S_n = a_1 + a_2 + \ldots + a_n = \sum_{k=1}^{n} a_k$. We say that the series $\sum_{k=1}^{\infty} a_k$ **converges to *S*** if the associated sequence of partial sums converges to S, and we write

$$\sum_{k=1}^{\infty} a_k = \lim_{n\to\infty} S_n.$$

Similarly, we say that the infinite sum **diverges to ±∞** if the associated sequence of partial sums diverges to $\pm\infty$.

> **Proposition 8.2 (Divergence Test)**
> If $\sum_{k=1}^{\infty} a_k$ converges, then $a_n\to 0$ as $n\to\infty$.

Thus, if $a_n \not\to 0$ as $n\to\infty$, the series $\Sigma_{k=1}^{\infty} a_k$ diverges, and so this is called the **divergence test.**

Proof The fact that the sequence of partial sums is Cauchy implies that, for all $\varepsilon>0$ there exists N such that $n \geq N$ implies $|S_n - S_{n-1}| < \varepsilon$. In other words, $|a_n| < \varepsilon$, showing that $a_n \to 0$ as $n \to \infty$. ❖

Note The converse is false: that is, if $a_n \to 0$, then it does not necessarily imply that $\Sigma_{k=1}^{\infty} a_k$ converges. (See Example 8.3(C) below). All this says is that if $a_n \not\to 0$ as $n \to \infty$, then the series $\Sigma_{k=1}^{\infty} a_k$ has no hope of converging. If $a_n \to 0$ as $n \to \infty$, then the series *does have* a *hope* of converging, but might still diverge.

Examples 8.3

A. The series $\Sigma_{k=1}^{\infty} k/(k+1)$ diverges.

B. Geometric Series $\displaystyle\sum_{k=0}^{\infty} ar^k$ Here, $S_n = a + ar + \ldots + ar^n = \dfrac{a(1-r^{n+1})}{1-r}$ converges iff $|r| < 1$, in which case it converges to $\dfrac{a}{1-r}$.

C. Harmonic Series $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k}$ diverges

D. Telescoping Series, such as $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k(k+1)}$.

E. Alternating harmonic series $\displaystyle\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}$ has $S_k \leq 1$ and S_{2k} increasing -- picture for now (better proof later)

Proposition 8.4 (Cauchy Criterion)

The series $\Sigma_{k=1}^{\infty} a_k$ converges iff, given any $\varepsilon > 0$, there exists $N \in$ N such that

$$m, n \geq N \Rightarrow \left| \sum_{k=m}^{n} a_k \right| < \varepsilon.$$

Proof in class.

Exercise Set 8

1. Discuss convergence of the following series.

(a) $\displaystyle\sum_{k=1}^{\infty} \frac{2^{k+1}}{3^{2k}}$ **(b)** $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^2+3k+2}$ **(c)** $\displaystyle\sum_{k=1}^{\infty} \frac{1}{4k^2+1}$ **(d)** $\displaystyle\sum_{k=1}^{\infty} (\sqrt{k+1} - \sqrt{k})$

(e) $\displaystyle\sum_{k=1}^{\infty} \ln \frac{k}{k+1}$ **(f)** $\displaystyle\sum_{k=1}^{\infty} k \sin \frac{1}{k}$ **(g)** $\displaystyle\sum_{k=1}^{\infty} \ln \frac{1}{k}$

2. Prove of give a counterexample with regard to each of the following claims.
(a) If both Σa_n and Σb_n diverge, then so does $\Sigma(a_n + b_n)$.
(b) If both Σa_n and Σb_n converge, then so does $\Sigma(a_n + b_n)$.
(c) If Σa_n converges and Σb_n diverges, then $\Sigma(a_n + b_n)$ diverges.
(d) If both Σa_n and Σb_n converge, then so does $\Sigma(a_n b_n)$.

3. Use the Cauchy criterion to show that $\sum_{k=1}^{\infty} \frac{1}{k^2}$ converges.

4. (Decimal Representation of the Real Numbers)
(a) Prove that, if $(n_1, n_2, \ldots, n_k, \ldots)$ is any sequence of integers with $0 \le n_i \le 9$ for all i, then the infinite series

$$\sum_{n=1}^{\infty} \frac{n_k}{10^k} = 0.n_1 n_2 \ldots n_k \ldots$$

converges.
(b) Prove that every real number in $(0, 1)$ can be expressed in the form $0.n_1 n_2 \ldots n_k \ldots$.(Hint: Construct a certain increasing sequence $S_n \rightarrow x$ with $0 \le x - S_n < 10^{-n}$)

9. Series with (Eventually) Positive Terms

Here, we consider only series of the form $\Sigma_{k=1}^{\infty} a_k$ with $a_k \ge 0$ for all k, and give a few little tests for convergence. More generally, we also allow series whose terms are *eventually positive*, that is, $a_k \ge 0$ for $k \ge$ some n. In what follows, bear in mind that we can use an eventually positive series as well, but will simply refer to a **series with positive terms.**

Proposition 9.1 (Series with Positive Terms)
If Σa_k is a series with positive terms, then it either converges, or diverges to $+\infty$.

Proof The reason for this is that the sequence of partial sums is monotone increasing, and so the result follows from Part I of these notes, which says that a monotone increasing sequence either converges or diverges to $+\infty$. ❖

Theorem 9.2 (Comparison Test for Series with Positive Terms)
If Σa_k and Σb_k are series, and if there exists $M \ge 0$ with $0 \le a_k \le M b_k$ for all $k \ge$ some n, then:

(a) Σb_k convergent $\Rightarrow \Sigma a_k$ convergent.
(b) Σa_k divergent $\Rightarrow \Sigma b_k$ divergent

Proof Again the result follows by looking at the partial sums. ❖

Examples 9.3

A. $\sum \frac{5}{3^{k-2}+1}$ **B.** $\sum \frac{5}{1+\sqrt{k}}$ **C.** $\sum \frac{5^n}{n!}$ **D.**

$\sum \left(\frac{k+10}{3k}\right)^k$

Theorem 9.3 (Integral Test)

If f: $[1, +\infty) \to \mathbb{R}$ is any decreasing function with $a_k = f(k)$ for all $k \geq 1$, then $\sum_{k=1}^{\infty} a_k$ converges iff $\int_1^{\infty} f(x)\, dx$ converges.

Proof This follows at once from the inequality

$$S_{k+1} - a_1 \leq \int_1^{k+1} f(x)\, dx \leq S_k$$

by taking limits. ❖

Notes
1. We can replace 1 in the above theorem with any integer n everywhere it occurs.
2. The integral and the series may not converge to the same value.

Corollary 9.4 (p Series)

The series $\sum_{k=1}^{\infty} \frac{1}{k^p}$ converges iff $p > 1$.

Proof We settle the case $p \leq 0$ using the divergence test, and the case $p > 0$ follows from the integral test. ❖

More Examples 9.5

A. $\sum_{k=2}^{\infty} \frac{1}{k \ln k}$ **B.** $\sum_{k=10}^{\infty} \frac{1}{k \ln k \ln(\ln k)}$, etc.

Definition 9.6 If Σa_k is an series, then a **rearrangement** of Σa_k is a series of the form $\Sigma a_{\phi(k)}$, where ϕ: $\mathbb{N}^* \to \mathbb{N}^*$ is a bijection.

Examples
In class.

Theorem 9.6 (Rearrangement of a Series with Positive Terms)
If $\Sigma a_{\phi(k)}$ is any rearrangement of Σa_k, then Σa_k converges iff $\Sigma a_{\phi(k)}$ converges, in which case they converge to the same sum.

Proof Call the two series in question Σa_k and Σb_k, so that $b_k = a_{\phi(k)}$, and $a_k = b_{\phi^{-1}(k)}$, and call the corresponding partial sums S_n and T_n respectively. Then, for all n,

$$0 \le T_n \le S_{\max\{\phi(1),\, \phi(2),\, \ldots,\, \phi(n)\}}$$

(write out what T_n is, and notice that the term on the right includes all the summands of T_n. This is where we use the fact that the series has positive terms.) Writing $\max\{\phi(1), \phi(2), \ldots, \phi(n)\}$ as $\Phi(n)$, we get

$$0 \le T_n \le S_{\Phi(n)}.$$

Similarly, for all m,

$$0 \le S_m \le T_{\Psi(m)},$$

where $\Psi(m) = \max\{\phi^{-1}(1), \ldots, \phi^{-1}(m)\}$. Putting them together yields

$$0 \le T_n \le S_{\Phi(n)} \le T_{\Psi(\Phi(n))}.$$

Since both $\Phi(n)$ and $\Psi(n) \to +\infty$ as $n \to \infty$, we can take limits and get the result. ❖

Exercise Set 9

1. Test for convergence:

(a) $\displaystyle\sum_{k=2}^{\infty} \frac{1}{\ln k}$ **(b)** $\displaystyle\sum_{k=2}^{\infty} \frac{1}{k \ln k}$ **(c)** $\displaystyle\sum_{k=2}^{\infty} \frac{1}{[\ln k]^2}$ **(d)** $\displaystyle\sum_{k=2}^{\infty} \frac{1}{[\ln k]^k}$ **(e)** $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^{1+1/k}}$ **(f)** $\displaystyle\sum_{k=e^2}^{\infty} \frac{1}{k \ln k \ln(\ln k)}$

2. (a) Give an example to show that Theorem 9.6 (on rearrangements) does not work if the series includes infinitely many negative terms.

(b) If a series has finitely many negative terms, does Theorem 11.6 work? Give a proof or counterexample.

3. (a) Prove the following.

Theorem 9X3 (Limit Comparison Test)

Let Σa_k and Σb_k be two series with positive terms with $\displaystyle\lim_{k \to \infty} \frac{a_k}{b_k}$ finite. Then:

(a) Σb_k convergent $\Rightarrow$ Σa_k convergent.

Thus, **(b)** Σa_k divergent $\Rightarrow$ Σb_k divergent.

(b) Use the limit comparison test to discuss convergence of the following series:

(i) $\displaystyle\sum_{k=1}^{\infty} \frac{2\sqrt{k}}{3k^2 - 2k + \sqrt{k}}$ **(ii)** $\displaystyle\sum_{k=1}^{\infty} \frac{k^2}{k!}$

[Hint for (ii): The "obvious" candidate won't work, so try something not quite so small.]

10. More Tests for Series with Positive Terms,

in which we look at ratio and root tests.

Theorem 10.1 (Ratio Comparison Test)
S'pose that Σa_k and Σb_k are series with positive terms such that $\exists n > 0$ such that

$$\frac{a_{k+1}}{a_k} \le \frac{b_{k+1}}{b_k}$$

for all $k \ge n$. Then

(a) Σb_k convergent $\Rightarrow$ Σa_k convergent.

Thus, **(b)** Σa_k divergent $\Rightarrow$ Σb_k divergent.

Proof The given inequality implies that

$$\frac{a_{k+1}}{b_{k+1}} \le \frac{a_k}{b_k}$$

for $k \ge n$, implying that the sequence (a_k/b_k) is eventually monotone decreasing and positive, whence convergent. The conclusion now follows from the Limit Comparison Test. ❖

Example 10.2
Use the test to prove that $\Sigma\, 1/k!$ converges, given that $\Sigma\, 1/k^2$ converges.

Note As the proof shows, this is just the Limit Comparison Test in disguise, but harder to verify, since we need the ratio a_k/b_k to *decrease* to a limit rather than just to converge as required in the Limit Comparison Test. In other words,

The Ratio Comparison Test is Useless:
The Limit Comparison Test is more powerful than the Ratio Comparison Test. That is, if the Ratio Comparison Test works, then so does the Limit Comparison Test (and it works with the same series in the test). In yet other words, the Ratio Comparison Test gives us nothing new, so don't bother to use it. What we need it for is other stuff.

Still, we can use it to prove a more interesting test that doesn't require comparison with another series:

Theorem 10.3 (d'Alembert's Ratio Test)†
S'pose that Σa_k is a series with positive terms. Then:

(a) If there exists $n > 0$ and $\alpha < 1$ with

$$\frac{a_{k+1}}{a_k} \le \alpha$$

for all $k \ge n$, then Σa_k converges.

(b) If there exists $n > 0$ with

$$\frac{a_{k+1}}{a_k} > 1$$

for all $k \ge n$, then Σa_k diverges.

---

† named after Jeán Le Rond d'Alembert (1717-1783)

Proof For (a), apply the Ratio Comparison Test using the geometric series $\Sigma\alpha^k$. For (b), use, as Σa_k the series $\Sigma 1^k = \Sigma 1$, and as Σb_k the given series Σa_k. ❖

Corollary 10.4 (Cauchy's Ratio Test) (which you learned in Calc II)

S'pose that Σa_k is a series with positive terms, and that $\lim\limits_{k\to\infty} \frac{a_{k+1}}{a_k}$ exists and equals α.

Then if $\alpha < 1$, the series Σa_k converges, and if $\alpha > 1$, the series Σa_k diverges.

Note If the sequence of ratios a_{k+1}/a_k fails to converge, then the Cauchy test is useless, but d'Alembert's test may still work.

One More Test

Theorem 10.5 (Fancy Root Test)
S'pose that Σa_k is a series with positive terms. Then:
(a) If there exists $n > 0$ and $\alpha < 1$ with

$$\sqrt[k]{a_k} \le \alpha$$

for all $k \ge n$, then Σa_k converges.
(b) If there exists $n > 0$ with

$$\sqrt[k]{a_k} > 1$$

for all $k \ge n$, then Σa_k diverges.

Proof in the exercises

Corollary 10.6 (Limit Root Test) (which you also learned in Calc II)

S'pose that Σa_k is a series with positive terms, and that $\lim\limits_{k\to\infty} \sqrt[k]{a_k}$ exists and equals α.

Then if $\alpha < 1$, the series Σa_k converges, and if $\alpha > 1$, the series Σa_k diverges.

Examples 10.7
A. Any geometric Series **B.** $(a_n) = (1/2, 1/3, 1/2^2, 1/3^2, \ldots)$

We see that Example B above fails d'Alembert's test but passes the fancy root test.

Notes
1. The fancy root test is stronger than d'Alembert's ratio test: if the ratio test works, then so does the root test. (Proof in the exercises) Thus, if the ratio test fails, there may still be a chance that the root test works. However, the root test may be harder to use. (Just try taking the kth root of k factorial.)

Exercise Set 10
1. Discuss convergence of the following series:

(a) $\sum_{k=1}^{\infty} \frac{k^2}{2^k}$ **(b)** $\sum_{k=2}^{\infty} \frac{k!}{k^k}$ **(c)** $\sum_{k=2}^{\infty} (\sqrt[k]{k} - 1)^k$

2.* Discuss the convergence of $\sum_{k=1}^{\infty} \frac{k^k}{e^k\, k!}$ and $\sum_{k=1}^{\infty} \frac{e^k\, k!}{k^k}$. [Hint: Neither the ratio test nor the root test will work here. What will get you started is this: take the natural log of the terms, see what you get, and use a Riemann sum to compare the messy part with an integral…]
3. Illustrate the necessity of the strict inequality $\alpha < 1$ in d'Alembert's Ratio Test by giving two series; one convergent, one divergent, which satisfy the hypothesis of part (a) of the theorem, except that $\alpha = 1$.
4. Prove the Root test.
5. (a) Prove that if Σa_k satisfies hypothesis (a) of d'Alembert's ratio test, it also satisfies the corresponding hypothesis of the fancy root test. (Hence, if the ratio test shows convergence, so does the root test.)
(b) Give two divergent series: one that satisfies hypothesis (b) of d'Alembert's ratio test but not hypothesis (b) of the fancy root test, and *vice-versa*.

11. Absolute and Conditional Convergence,

where we look at series that are not eventually positive.

Definition 11.1 An (**eventually**) **alternating** series is a series of the form $\Sigma(-1)^k a_k$ or $\Sigma(-1)^{k+1} a_k$, where (a_k) is an (eventually) positive sequence.

Examples 11.2

A. The **alternating harmonic series,** $\sum_{k=1}^{\infty} \frac{(-1)^k}{k}$.

B. $\sum_{k=1}^{\infty} \sin\left(\frac{(2n+1)\pi}{2}\right)$ **C.** Any geometric series with negative r.

The following theorem shows that it is easy to tell when an alternating series converges. First, some notation.

Notation If (a_n) is a sequence, we write $a_n \searrow 0$ as $n \to \infty$ if (a_n) is eventually decreasing, and $a_n \to 0$ as $n \to \infty$.

Theorem 11.3 (Leibniz' Alternating Series Test)
If (a_n) is a sequence with $a_n \searrow 0$ as $n \to \infty$, then $\Sigma(-1)^k a_k$ and $\Sigma(-1)^{k+1} a_k$ converge.

Proof Let us do the proof for $\Sigma(-1)^k a_k$. We shall prove convergence of the sequence of partial sums (S_n). To do this, look at the subsequence of even terms, (S_{2n}). We can write

---

* This problem was given as an extra credit problem to a Math 20 class of mine some time in the past, and had all the students running running around the Math department in a frenzy.

$$S_{2(n+1)} = S_{2n} - a_{2n+1} + a_{2n+2} \le S_{2n},$$

since (a_n) is decreasing. Thus (S_{2n}) is a decreasing sequence. On the other hand, for all n,

$$\begin{aligned} S_{2n} &= S_1 + a_2 - a_3 + a_4 - a_5 + \ldots + a_{2n-2} - a_{2n-1} + a_{2n} \\ &\ge S_1 + (a_2 - a_2) + (a_4 - a_4) + \ldots + (a_{2n-2} - a_{2n-2}) + a_{2n} \\ &= S_1 + a_{2n} \ge S_1. \end{aligned}$$

Thus, since the sequence of even terms is decreasing and bounded below (by S_1), it converges to S, say. Finally, since for all n one has

$$\begin{aligned} |S_{2n+1} - S| &\le |S_{2n+1} - S_{2n}| + |S_{2n} - S| \\ &= a_{2n+1} + |S_{2n} - S| \to 0 \text{ as } n \to \infty, \end{aligned}$$

it also follows that the odd terms also converge to S. ❖

Examples 11.4

A. Alternating harmonic series $1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{5} + \ldots = \ln 2$

B. $\dfrac{\sin(n\pi/2)}{\ln n}$

Definition 11.5 The series Σa_k is **absolutely convergent** if the series $\Sigma|a_k|$ converges. If Σa_k converges, but $\Sigma|a_k|$ diverges, then the series is **conditionally convergent**.

Examples 11.6

A. $\displaystyle\sum_{k=1}^{\infty} \frac{(-1)^k}{k^2}$ is absolutely convergent. **B**. $\displaystyle\sum_{k=1}^{\infty} \frac{(-1)^k}{k}$ is conditionally convergent.

Proposition 11.7 (Absolute Convergence Implies Convergence)
Let Σa_k be absolutely convergent. Then:
(a) Σa_k is convergent
(b) $\Sigma \max\{a_k, 0\}$ is convergent
(c) $\Sigma \min\{a_k, 0\}$ is convergent

Proof (a) Let Σa_k be absolutely convergent, with partial sums S_n, and let T_n be the partial sums of $\Sigma|a_k|$. Let $\varepsilon > 0$. Then, since $\Sigma|a_k|$ converges, there exists N such that $n \ge m \ge N$ implies $|T_n - T_m| < \varepsilon$. But

$$\begin{aligned} |S_n - S_m| &= |a_{m+1} + a_{n+2} + \ldots + a_n| \\ &\le |a_{m+1}| + |a_{n+2}| + \ldots + |a_n| = |T_n - T_m| < \varepsilon, \end{aligned}$$

showing that Σa_k also converges.
(b) and **(c)** follow by the comparison test. ❖

Now, we look at rearrangements.

Theorem 11.8 (Rearrangement of an Absolutely Convergent Series)
If $\Sigma a_{\phi(k)}$ is any rearrangement of the absolutely convergent series Σa_k, then $\Sigma a_{\phi(k)}$ converges absolutely, and to the same sum as Σa_k.

Proof[†] Call the two series in question Σa_k and Σb_k, so that $b_k = a_{\phi(k)}$ (and $a_k = b_{\phi^{-1}(k)}$) for some permutation ϕ: N*→N*. First, we show that the series Σb_n converges absolutely by the Cauchy criterion. Thus, let $\varepsilon > 0$ and choose N such that

$$m \geq n \geq N \Rightarrow |a_n| + \ldots + |a_m| < \varepsilon \qquad \ldots \qquad (1)$$

Now, $\{a_1, a_2, \ldots, a_N\} = \{b_{\phi^{-1}(1)}, b_{\phi^{-1}(2)}, \ldots, b_{\phi^{-1}(N)}\}$. So, if we choose M to be the largest index of the set on the right, then $p \geq M$ implies $b_p = a_r$ for some $r \geq N$. Thus, if $p \geq q \geq M$, then

$$|b_q| + \ldots + |b_p| = |a_{\phi(p)}| + \ldots + |a_{\phi(q)}|,$$

where each $\phi(p) \geq N$. But all the indices on the righ-hand side are at least N, so the right-hand side is $< \varepsilon$ by (1), showing that the series Σb_n is absolutely convergent.

Let the corresponding partial sums of the series Σa_k and Σb_k be S_n and T_n respectively. Define a subsequence (T_{n_k}) of (T_n) by requiring that T_{n_k} be the sum of all the terms in S_k and then some. (There are lots of possible choices for this subsequence. All we require is that n_k be at least as big as the largest of $\phi^{-1}(1), \ldots, \phi^{-1}(k)$.) In symbols,

$$T_{n_k} = S_k + \text{terms } a_r \text{ with } r > k.$$

Thus, for every k,

$$|T_{n_k} - S_k| \text{ is the absolute value of a sum of terms } a_r \text{ with } r > k \qquad \ldots \qquad (2)$$

We now claim that the subsequence (T_{n_k}) converges to $S = \lim S_n$ (whence, since the original sequence (T_n) also converges—we showed above that Σb_n is (absolutely) convergent—it too must converge to S, completing the proof of the theorem). Indeed, let $\varepsilon > 0$. Since Σa_k converges *absolutely*, there exists N such that

$$m \geq k \geq N \Rightarrow |a_k| + \ldots + |a_m| < \varepsilon/2 \text{ and also } |S_k - S| < \varepsilon/2.$$

But then, for $k \geq N$, one has

$$|T_{n_k} - S| \leq |T_{n_k} - S_k| + |S_k - S|$$
$$\leq \varepsilon/2 + \varepsilon/2$$

by (2) and the choice of N, showing that (T_{n_k}) converges to S. ❖

On the other hand,

Theorem 11.9 (Rearrangement of a Conditionally Convergent Series)
If Σa_k is a conditionally convergent series, and if s is either an arbitrary real number or $\pm\infty$, there exists a rearrangement $\Sigma a_{\phi(k)}$ of Σa_k with infinite sum s.

Proof For each n, let $a_k^+ = \max\{a_k, 0\}$ and $a_n^- = \min\{a_k, 0\}$.

Claim 1 We assert that Σa_k^+ diverges to $+\infty$ and Σa_k^- diverges to $-\infty$.

Indeed, if Σa_k^+ converged, then , noting that

$$a_k = a_k^+ + a_k^-,$$

we see that two of the three corresponding series converge, whence so does the third. (See the exercises.) Thus, Σa_k^- must also converge. But now

$$|a_k| = a_k^+ - a_k^-,$$

[†] Kosmala doesn't seem to bother with this proof or even the easier one about series with positive terms. Is he scared?

and now the series corresponding to both terms on the right converge, whence so does $\Sigma|a_n|$, contradicting the fact that Σa_k is conditionally convergent. Thus Σa_k must diverge, and hence must diverge to $+\infty$, by the result on series with positive terms. *Mutatis mutandis*, it follows that Σa_k^- also diverges to $-\infty$.

Claim 2 There is a rearrangement converging to any finite number s.

Assume, without loss of generality, that $s \geq 0$. Let $n(1)$ be the least integer such that

$$a_1^+ + a_2^+ + \ldots + a_{n(1)}^+ > s,$$

and let $m(1)$ be the least integer such that

$$a_1^+ + a_2^+ + \ldots + a_{n(1)}^+ + a_1^- + a_2^- + \ldots + a_{nm1)}^- < s.$$

(Such integers exist by Claim 1.)Then choose $n(2)$ to be the least integer such that

$$\begin{aligned} &a_1^+ + a_2^+ + \ldots + a_{n(1)}^+ \\ &\quad + a_1^- + a_2^- + \ldots + a_{nm1)}^- \\ &\qquad + a_{n(1)+1}^+ + a_{n(1)+2}^+ + \ldots + a_{n(2)}^+ > s. \end{aligned}$$

Notice that $|s - \text{this sum}| \leq a_{n(2)}$ by choice of $n(2)$. Continuing in this vein and leaving out all zero terms of the form a_k+ and a_k^- (where the corresponding original terms were not zero), we are obtaining a rearrangement that converges to s.

Clam 3 There are rearrangements that diverge to $\pm\infty$.

Proof in the exercise set. ❖

Exercise Set 11

1. Determine whether the following series converge absolutely, conditionally, or neither. Give the reasons.

(a) $\displaystyle\sum_{k=1}^{\infty} \frac{k^2}{(-2)^k}$ **(b)** $\displaystyle\sum_{k=2}^{\infty} (-1)^k\left(\frac{3k+2}{4k^2-3}\right)$ **(c)** $\displaystyle\sum_{k=2}^{\infty} (-1)^k\frac{\arctan k}{k^2}$ **(d)** $\displaystyle\sum_{k=2}^{\infty} \frac{k!}{(-k)^k}$

2. Prove in three lines or less that, if $a_n = b_n + c_n$, and if two of the three corresponding series converge, so does the third.

3. Prove Claim 3 in the proof of Theorem 11.8: There are rearrangements of any conditionally convergent series that diverge to $\pm\infty$.

12. Sequences of Functions: Pointwise Convergence and Uniform Convergence

Definition 12.1 Let (f_n) be a sequence of functions f_n:$D\to$R. Then we say that $\boldsymbol{f_n \to f}$ **pointwise** if, for all $x \in D$, the sequence $(f_n(x))$ converges to $f(x)$.

Examples 12.2

A. f_n: $(-1, 1]\to$R; $f_n(x) = x^n$ (pictures of graphs in class)

B. f_n: R$\to$R; $f_n(x) = 1 + x + \dfrac{x^2}{2!} + \ldots + \dfrac{x^n}{n!}$

C. f_n: R$\to$R; $f_n(x) = x^n$

D. f_n: R$\to$R; $f_n(x) = u_n(x) - u_{n+1}(x)$, where, for $c \in$ R,

$$u_c(x) = \begin{cases} 0 & \text{if } x < c \\ 1 & \text{if } x \geq c \end{cases}.$$

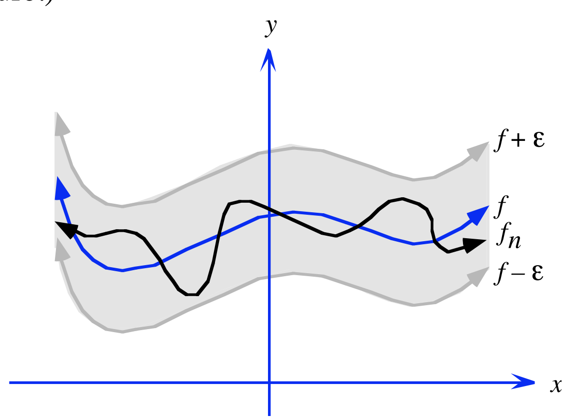
E. (An unintuitive example) f_n: R→R; $f_n(x) = n[u_n(x) - u_{2n}(x)]$

Question Is the limit of the integral equal to the integral of the limit?
Answer Not on your life! Let f_x: $[0, 1]$→R; $f_n(x) = (n+1)x^n$. Then the integral is 1 for each n, whereas the limit is zero. Also look at $\int_0^{+\infty} f_n(x)\, dx$ in Examples D and E above, where the functions also approach zero, but not the integrals. Clearly then, we need a better behaved (or more intuitive) form of convergence. Here is one.

Definition 12.3 Let (f_n) be a sequence of functions f_n:D→R. Then we say that $\boldsymbol{f_n \to f}$ **uniformly** if, for all $\varepsilon > 0$, there exists $N \in$ N such that

$$n \geq N \Rightarrow |f_n(x) - f(x)| < \varepsilon$$

for all $x \in D$. (See figure.)



Note Uniform convergence implies pointwise convergence (Exercise Set 14) but the converse need not hold (see below).

Proposition 12.4 (Sequential Criterion—Oh Yes There Is One!—For Uniform Convergence)
The sequence (f_n: D→R) converges uniformly iff the sequence

$$r_n = \sup\{|f_n(x) - f(x)| : x \in D\}$$

exists for large enough n and → 0 as $n\to\infty$.

❖

Examples 12.5
A. f_n: $[0, 1]$→R; $f_n(x) = x/n$ converges uniformly to 0, since $|x/n| \leq 1/n$ for all x in the domain, so we simply choose $N = 1/\varepsilon$.
B. Examples 12.2 (D) and (E) are not uniformly convergent.
C. Example 12.2 (B) is not uniformly convergent, since $e^x \to +\infty$, whereas the odd sums go off to $-\infty$. (Even the even sums behave badly: look at the difference $|e^x - f_n(x)|$ and let $x \to +\infty$.)

Question What about Example 12.2(A) and (C)?
Answer We can dispense with (A) and (C) once we show the following:

Proposition 12.5 (Uniform Limit of a Continuous Function)
If (f_n: $D \to \mathrm{R}$) is a sequence of continuous functions with uniform limit f, then f is also continuous

Proof Let $\varepsilon > 0$, and let $a \in D$. We show continuity at a. First, choose n such that

$$|f_n(x) - f(x)| < \varepsilon/3 \qquad (1)$$

for every $x \in D$. Then, choose δ such that $x \in D$ and $|x-a| < \delta$ implies

$$|f_n(x) - f_n(a)| < \varepsilon/3 \qquad (2).$$

Then, $x \in D$ and $|x-a| < \delta$ implies

$$\begin{aligned} |f(x) - f(a)| &= |f(x) - f_n(x) + f_n(x) - f_n(a) + f_n(a) - f(a)| \\ &\le |f(x) - f_n(x)| + |f_n(x) - f_n(a)| + |f_n(a) - f(a)| \\ &< \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon. \end{aligned}$$ ❖

This, plus other results, gives:

Proving the (f_n) does not converge uniformly to f
Either:
1. Show that each f_n is continuous, but not f
2. Prove that $\sup\{|f_n(x) - f(x)| : x \in D\}$ either does not exist for infinitely many n, or does not approach 0 as $n \to \infty$.

Examples 12.6

A. By the proposition (method 1 in the box) we can immediately rule out Examples 14.2 (A) and (C).

B. Does f_n: $[0, +\infty) \to \mathrm{R}$; $f_n(x) = x^2 e^{-nx}$ converge uniformly ? [Answer: Yes. Hint: locate the absolute maximum of this function and use it together with the sequential condition.]

C. f_n: $[0, +\infty) \to \mathrm{R}$; $f_n(x) = \dfrac{nx}{1+n^2x^2}$ [Answer: no. To see why, draw their graphs by plotting maximum points. Then invoke Method 2.]

Some more things preserved by uniform convergence:

Proposition 12.7 (Integral of a Limit)
If (f_n: $[a, b] \to \mathrm{R}$) is a sequence of continuous functions with uniform limit f, then f is integrable on $[a, b]$, and

$$\lim_{n \to \infty} \int_a^b f_n(x)\, dx \text{ exists and equals } \int_a^b f(x)\, dx .$$

Proof* By the sequential criterion,

$$M_n = \sup\{|f_n(x) - f(x)| : x \in D\}$$

exists for sufficiently large n and $\to 0$ as $n \to \infty$. Now, for such n,

$$0 \le |f_n(x) - f(x)| \le M_n.$$

Further, since every function in sight is continuous and hence integrable, one has

* See Kosmala, p. 387 for a less intersting proof.

$$0 \le \left| \int_a^b f_n(x)\,dx - \int_a^b f(x)\,dx \right| = \left| \int_a^b f_n(x) - f(x)\,dx \right| \le \int_a^b |f_n(x) - f(x)|\,dx$$

$$\le \int_a^b M_n\,dx = M_n(b-a) \to 0 \text{ as } n\to\infty,$$

from which the result follows by the sandwich rule. ❖

Note If we replace the requirement of continuity of the f_n by integrability, we still survive, since we can have:

$$0 \le |U(f_n, P) - U(f, P)| \le |U(f_n - f, P)| \le |U(M_n, P)| \to 0$$

and

$$0 \le |L(f_n, P) - L(f, P)| \le |L(f_n - f, P)| \le |U(M_n, P)| \to 0,$$

and hence so do the suprema taken over partitions P. But since

$$U(f_n) = L(f_n) = \int_a^b f_n(x)\,dx\ ,$$

we must have $U(f) = L(f)$, and everything works. (You will be asked to polish this proof up in the exercises.)

Finally, we have

Theorem 12.8 (U. Dini, *circa* 1935)
Let f_n: $[a, b]\to\mathbb{R}$ and s'pose that f_n ^f pointwise as $n\to\infty$. S'pose in addition that the f_n and f are continuous. Then $f_n\to f$ uniformly.

Proof is an extra credit exercise.†

Exercise Set 12

1. Discuss unform convergence of the following functions:

(a) f_n: $[-\pi, \pi]\to\mathbb{R}$; $f_n(x) = \dfrac{\sin(nx)}{\sqrt{n}}$

(b) f_n: $[0, +\infty)\to\mathbb{R}$; $f_n(x) = \dfrac{x^n}{1+x^{2n}}$

(c) f_n: $[0, 1]\to\mathbb{R}$; $f_n(x) = \dfrac{x^n}{n}$

(d) f_n: $[0, 2]\to\mathbb{R}$; $f_n(x) = \dfrac{x^n}{2^n}$

(e) f_n: $[0, \pi]\to\mathbb{R}$; $f_n(x) = \sin^n(x)$

---

† There is a proof in Wade, and if you can make head or tail of his argument (or decipher his bizarre notation), good luck to you. Alternatively, here is an easier plan of attack: By considering the functions $g_n = f - f_n$, observe that it suffices to prove the following special case: If (g_n: $[a,b\}\to\mathbb{R}$) is a decreasing sequence of continuous functions that converges pointwise to zero, then $g_n\to 0$ uniformly. To prove the latter, it suffeices in turn to show that $y_n = \max\{g_n(x) \mid x \in [a, b]\} \to 0$ as $n\to\infty$. That is where the work is...

2. Discuss uniform convergence of f_n: $(-1,\ 1)\to\mathrm{R}$ given by $f_n(x) = \dfrac{1-x^n}{1-x}$.

3. Prove that uniform convergence implies pointwise convergence.

4. Give examples of the following, justifying your claims.

(a) A sequence of discontinuous functions on [0, 1] that converges uniformly to a continuous function.

(b) A sequence of discontinuous functions on [0, 1] that converges uniformly to a discontinuous function.

(c) A sequence of continuous functions on [0, 1] that converges non-uniformly to a continuous function.

5. Polish up the note following Proposition 12.7.

Extra Credit Prove Dini's theorem.

13. Series of Functions

Recall that a series is nothing more than a sequence (of partial sums). But we will be somewhat interested in functions defined as limits of series (such as power series), and we desire to know some properties of these functions.

Definition 13.1 If (f_n) is a sequence of functions, then the associated **infinite series** $\sum_{k=1}^{\infty} f_k$ is just the sequence of partial sums $(s_n = \sum_{k=1}^{n} f_k)$. Thus, we can use the usual definitions to talk of **pointwise convergence,** and **uniform convergence.**

Note All the results and tests above still work here; for example, if the f_k are continuous, $\Sigma_{k=1}^{\infty} f_k$ converges uniformly, then its limit must be continuous.

Examples 13.1

A. f_n: $(-1,\ 1)\to\mathrm{R}$, $f_n(x) = x^n$ has $\Sigma_{k=0}^{\infty} f_k$ converging pointwise, but not uniformly, to $1/(1-x)$. (See the previous exercise set and/or use the sequential criterion.)

B. The same example as in (A), but with $(-1,\ 1)$ replaced by any closed subinterval. (Use the sequential criterion again.)

C. $f_n[0,\ 1]\to\mathrm{R}$; $f_n(x) = \dfrac{x}{(x+1)^{n-1}}$ has a discontinuous infinite sum (it is x times a geometric series away from 0).

We have the following theoretical result.

Proposition 13.2 (Cauchy Criterion for Uniform Convergence)
$\Sigma_{k=1}^{\infty} f_k$ converges uniformly iff, for every $\varepsilon>0$ there exists $N \in$ N such that $n, m \geq N$ implies $\left|\Sigma_{k=m}^{n} f_k(x)\right| < \varepsilon$.

Proof Exercise Set 13. ❖

Definition 13.3 The series $\Sigma_{k=1}^{\infty} f_k$ converges **pointwise** (resp. **uniformly**) **absolutely** if the series $\Sigma_{k=1}^{\infty} |f_k|$ converges pointwise (resp. uniformly). Absolute uniform convergence is also known as **normal convergence**, or **convergence in the sup norm**.

Theorem 13.3 (Weierstrass M-Test)
S'pose (f_n: $D \to$R) is a sequence of functions, and (M_n) is a sequence of real numbers with $|f_n(x)| \leq M_n$ for all $n \geq$ some k. If the series ΣM_n converges, then $\Sigma_{k=1}^{\infty} f_k$ converges uniformly absolutely.

Proof Cauchy criterion. ❖

Examples

A. $f_n(x) = \dfrac{\sin(4x-n)}{n^2}$

B. f_n: $(0, +\infty) \to$R; $f_n(x) = \dfrac{\sin(nx)}{xn^{2.1}}$

(note that $|\sin(nx)/x|$ is bounded.)

The following result follows from the analogous result on convergence of sequences of functions.

Proposition 13.4 (Term-by-Term Integration)
If Σf_k is a series of integrable functions that converges uniformly to f, then f is integrable, and

$$\sum_{k=1}^{\infty} \int_a^b f_k(x)\ dx \text{ exists and equals } \int_a^b f(x)\ dx\ .$$

In other words,

$$\int_a^b \sum_{k=1}^{\infty} f_k(x)\ dx = \sum_{k=1}^{\infty} \int_a^b f_k(x)\ dx\ .$$

Proposition 13.5 (Term-by-Term Differentiation)
If Σf_k is a series of differentiable functions that converges *pointwise* to f, and assume also that $\Sigma f_k\,'(x)$ converges *uniformly*. Then f is differentiable, and

$$\sum_{k=1}^{\infty} \frac{d}{dx}\left[f_k(x)\right] = \frac{d}{dx}\left[\sum_{k=1}^{\infty} f_k(x)\ dx\right].$$

Proof in the Exercises

Exercise Set 13

1. Analyze the following series for uniform and pointwise convergence [Hint: to chow that a given series does not converge, it sometimes helps to see if you can obtain the partial sums and refer to the method used in §12.]:

(a) $\sum_{k=1}^{\infty} \frac{x^2}{k^2}$; $x \in [0, 1]$

(b) $\sum_{k=1}^{\infty} \frac{1}{x^k+1}$; $x \in (1, +\infty)$

(c) $\sum_{k=1}^{\infty} e^{-kx}$; $x \in \mathrm{R}$

(d) $\sum_{k=1}^{\infty} e^{-kx}$; $x \in (0, +\infty)$

(e) $\sum_{k=1}^{\infty} k^r e^{-kx}$; $x \in [a, +\infty)$, $a > 0$, $r \in \mathrm{R}$

2. Prove the Cauchy criterion for uniform convergence (Proposition 13.2).

3. Prove Proposition 13.5.

14. Power Series

Definition 14.1 A **power series** is a series of functions of the form $\sum_{k=0}^{\infty} c_k(x-a)^k$, where each c_k and $a \in \mathrm{R}$.

Note We will only be interested here in *pointwise* convergence, so we can use the whole battery of convergence tests we know and love.

Examples 14.2 in class, including examples of which converge, and for which x they converge absolutely and conditionally.

Theorem 14.3 (Convergence of Power Series)

If $\Sigma c_k(x-a)^k$ is any power series, then there exists $R \in \overline{\mathrm{R}}$ such that:
(a) the series converges absolutely if $|x-a| < R$, and
(b) it diverges if $|x-a| > R$. Further,
(c) it converges *uniformly* on any closed subinterval of $(a-R, a+R)$.

Proof

For claim (a), note it suffices to prove that, if the series converges for $x = a \pm h$ (with h positive), then it converges absolutely for all x in the open interval $(a-h, a+h)$. (For then, just take $R = \sup\{h \mid \text{the series converges for } x = a \pm h\}$.) First, since the series converges for $x = a\pm h$, the nth term goes to zero, whence so does its absolute value. In particular, the terms $|c_i h^i|$ are bounded, so let M be an upper bound.

Now, let $x \in (a-h, a+h)$. Then

$$0 \leq |c_i(x-a)^i| = |c_i|\cdot|x-a|^i = |c_i| \left|\frac{x-a}{h}\right|^i h^i \leq M\left|\frac{x-a}{h}\right|^i .$$

But the series on the right is geometric with common ratio < 1, so this result follows by the comparison test.

Claim (b) follows from Claim (a) using an argument by contradiction.

For Claim (c), let $[a-k,\ a+k]$ be any subinterval of $(a+R,\ a-R)$, and let $x \in [a-k,\ a+k]$. Choose B between k and R, and an upper bound M of the terms $|c_i k^i|$. Then,

$$0 \le |c_i(x-a)^i| = |c_i|\cdot|x-a|^i = |c_i| \left|\frac{x-a}{B}\right|^i h^i \le M\left|\frac{x-a}{B}\right|^i \le M\left(\frac{k}{b}\right)^i ,$$

where $k/b < 1$. Hence the result by the Weierstrass M-test. ❖

Definition 14.4 We refer to the number R in the above result as the **radius of convergence** of the power series.

Corollary 14.5 (Properties of Power Series)
If $\Sigma c_k(x-a)^k$ is a power series with radius of convergence R, with $f(x) = \Sigma_{k=0}^{\infty} c_k(x-a)^k$, then:
(a) f is continuous on $(x-R,\ x+R)$.
(b) f is differentiable on $(x-R,\ x+R)$, and $f'(x) = \Sigma_{k=1}^{\infty} kc_k(x-a)^{k-1}$, where the latter series has the same radius of convergence as the original one.
(c) f is integrable on $(x-R,\ x+R)$ with integral $\Sigma_{k=0}^{\infty} c_k(x-a)^{k+1}/(k+1)$.

Proof Follows from various results, but needs the as yet unproved result that multiplying or dividing the terms by $k+1$ does not effect the radius of convergence (and this is in the exercise set).

Proposition 14.6 (Multiplication of Power Series)
If $f(x) = \Sigma_{k=0}^{\infty} a_k(x-a)^k$ and $g(x) = \Sigma_{k=0}^{\infty} b_k(x-a)^k$ are both power series with radius of convergence R. then $f(x)g(x) = \Sigma_{k=0}^{\infty} c_k(x-a)^k$, where

$$c_k = \Sigma_{i+j=k} a_i b_j.$$

Moreoever, this power series also has radius of convergence at least R.

Proof[*] First, a little notation which we will use in the proof. If S_n is the nth partial sum of a series Σa_k, write $\text{Abs}(S_n)$ for the nth partial sum of $\Sigma|a_k|$. Now for the proof itself. Let $I = (a-R,\ a+R)$, and write the partial sums of the three series in question as

$$f_n(x) = \Sigma_{k=0}^{n} a_k(x-a)^k,\ g_n(x) = \Sigma_{k=0}^{n} b_k(x-a)^k,\ h_n(x) = \Sigma_{k=0}^{n} c_k(x-a)^k.$$

How are they related? (Note that $f_n(x)g_n(x) = h_n(x)$ + other terms in powers of $(x-a)$ larger than n.) Specifically,

$$f_n(x)g_n(x) = h_n(x) + \Sigma_{k=n}^{2n} d_k(x-a)^k,$$

where each d_j is a product $a_p b_q$ with either p or $q > [n/2]$ (or else the power would be $\le n$) It suffices to show that $\Sigma_{k=n}^{2n} d_k(x-a)^k \to 0$ as $n\to\infty$ for each $x \in I$. However,

$$\Sigma_{k=n}^{2n} |d_k (x-a)^k| \le \Sigma_{k=n}^{2n} |d_k|\ |x-a|^k$$
$$\le \text{Abs}[f_n(x)-f_{[n/2]}(x)]\ \text{Abs}(g_n(x)) + \text{Abs}[g_n(x)-g_{[n/2]}(x)]\ Abs(fn(x)),$$

[*] Compare this with the (very complicated) proof in Wade, p. 201. Admittedly, his method of proof requires us only to use the fact that the convergence of *one* of the series is absolute, but that requires a lot more work, and is not necessary for this result.

by the properties of the d_k. But, since the series for f and g are abolutely convergent on I, it follows that Abs($f_n(x)$) and Abs($g_n(x)$) are convergent, and that Abs[$f_n(x)-f_{[n/2]}(x)$] and Abs[$g_n(x)-g_{[n/2]}(x)$] $\to 0$ as $n\to\infty$, giving the result. ❖

Exercise Set 14

1. Find the intervals of cenvergence of the following power series:

(a) $\sum_{k=0}^{\infty} (-1)^k(x-2)^k$ **(b)** $\sum_{k=0}^{\infty} \frac{2(x-1)^k}{k3^k}$ **(c)** $\sum_{k=3}^{\infty} (-1)^k\frac{(x+2)^k}{k \ln k}$ **(d)** $\sum_{k=0}^{\infty} \frac{k!x^k}{k^k}$ **(e)** $\sum_{k=1}^{\infty} \frac{k(x-1)^k}{2^k}$

2. Prove the claim in Corollary 14.5: that multiplying or dividing the terms in a power series by any polynomial in k does not effect its radius of convergence. [Hint: adapt the first inequality in the proof of Theorem 14.3 to show that is the original series converges for $x = a\pm h$, then the new series converges for $|x-a| < R$ for any $R < h$.

15. Analytic Functions, Taylor Series, and Taylor's Theorem,

actually, Maclaurin's Theorem (see below).

Definitions 15.1 The function f: R ⇾ R is **smooth** (C^∞) on an open interval J if the derivatives of all orders exist on J. it is **analytic** on J if, for each $a \in J$, there is a power series $P(x)$ centered at a such $f(x) = P(x)$ on some open interval $(a-\delta, a+\delta)$. We say that f **has a power series expansion near *a*.**

Remark 15.2 By Corollary 14.5(b), it follows that every analytic function is C^∞. Is every C^∞ function analytic? See below...

Proposition 15.3 (Uniqueness of Power Series Expansion)

If $f(x) = \Sigma_{k=0}^{\infty} a_k(x-a)^k$, then

$$a_k = \frac{f^{(k)}(a)}{k!} .$$

Definition 15.4 If f: R ⇾ R is C^∞ on the interval J, then, for $a \in J$, define its **Macluarin Series**[**] **about *a*** to be the power series

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x-a)^k.$$

It follows from Proposition 15.3 that, if f is analytic on the interval J, then f is equal to its Maclaurin series expansion near every point of J.

[**] See the footnote to Theorem 15.5.

Examples 15.4

A. Some well-known Maclaurin Series are derived

B. Let $f(x) = \begin{cases} e^{-1/x^2} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$, then f is C^∞ on R but not analytic at 0.

Theorem 15.5 (Maclaurin)[††]

S'pose f is $n+1$ times differentiable on an interval J containing a and x. Then

$$f(x) = P_n(x) + R_n(x),$$

where

$$P_n(x) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2!} f''(a) + \dots + \frac{(x-a)^n}{n!} f^{(n)}(a),$$

and

$$R_n(x) = \int_a^x \frac{f^{(n+1)}(t)\,(x-t)^n}{n!} = \frac{(x-a)^{n+1}}{(n+1)!} f^{(n+1)}(\xi(x))$$

for some $\xi(x)$ strictly between a and x.

Proof We start by writing

$$f(x) = f(a) + \int_a^x f'(t)\, dt ,$$

which is just the Fundamental Theorem of Calculus. We integrate by parts using the following table (remember that x is regarded as fixed):

D	I
$f'(t)$	1
$f''(t)$	$(t-x)$
$f'''(t)$	$\frac{(t-x)^2}{2!}$
...	...
$f^{(n+1)}(t)$	$\frac{(t-x)^n}{n!}$

$t = x$

D	I
$f'(x)$	1
$f''(x)$	0
$f'''(x)$	0
...	...
$f^{(n+1)}(x)$	0

$t = a$

D	I
$f'(a)$	1
$f''(a)$	$(a-x)$
$f'''(a)$	$\frac{(a-x)^2}{2!}$
...	...
$f^{(n+1)}(a)$	$\frac{(a-x)^n}{n!}$

We now get:

$$f(x) - f(a) = \int_a^x f'(t)\, dt$$

$$= \Big[(t-x)f'(t)\Big]_{t=a}^{x} - \Big[\frac{(t-x)^2}{2!} f''(t)\Big]_{t=a}^{x} + \dots + (-1)^{n+1}\Big[\frac{(t-x)^n}{n!} f^{(n)}(t)\Big]$$

x

$t=a$

[††] Some people call it "Taylor's Theorem" but I am following the lead of my own (Scottish) applied mathematics teacher at Liverpool University, in attributing it—rightly or wrongly—to Maclaruin.

$$+ (-1)^n \int_a^x f^{(n+1)}(t) \frac{(t-x)^n}{n!} \, dx$$

= Middle Table − Right Table = What We Want.

Finally, to get the alternative form of the remainder, define a function H:$[a, x]\to$R (or $[a, x]\to$R if $x < a$) by

$$H(r) = f^{(n+1)}(r) \int_a^x \frac{(x-t)^n}{n!} \; .$$

Then H is the product of a continuous function and a differentiable one, so that it is continuous. Further, if $f^{(n+1)}$ achieves its minimum at p and maximum at q, tnen

$$H(p) \leq R_n(x) \leq H(p)$$

whence, by the IVT, we can find some $\xi(x)$ strictly between a and x with $H(\xi(x)) = R_n(x)$.

Corollary 15.6 (A Consequence: Sufficient Conditions for Analyticity)
Under the hypothesis of Theorem 15.5 **:**
(a) We can rewrite the remainder term as

$$|R_n(x)| \leq \frac{M\,|x-a|^{n+1}}{(n+1)!} \quad ,$$

where M is an upper bound of $|\, f^{(n+1)}(t)|$ on (a, x) $(or\ (x, a))$.
(b) If $|R_n(x)|\to 0$ as $n\to\infty$ then f is analytic at a. (It follows from one of the exercises that f is analytic *near* a as well.)
(c) If f is smooth on J and if there exists $K \in$ R with $|f^{(n+1)}(t)| \leq K^n$ for all $t \in J$, then f is analytic on J.

Proof Only part (c) is worthy of any explanation. However, choosing any x and a in J yields a Maclaurin sum with remainder

$$|R_n(x)| \leq \frac{K^{n+1}|x-a|^{n+1}}{(n+1)!} = \frac{L^{n+1}}{(n+1)!}$$

which is the nth term of a convergent series for every $L \in$ R, and hence approaches 0.

Examples 15.7
A. ins, cos, e^x, ln x, arctan(x)

Exercise Set 15
1. (Wade, p. 216 #1) Prove tht each of the following functions is analytic on R, and find its Maclaurin series about the indicated point.
(a) $\cos(3x)$; $a = 0$ **(b)** $3x^2 - 4x + 5$; $a = 1$ **(c)** $\sin^2 x$; $a = 0$ **(d)** $x^2 e^{x^2}$
2. Prove: If $f(x) = \Sigma_{k=0}^{\infty} a_k(x-a)^k$ converges on $(a-\delta,\ a+\delta)$, then f is analytic on $(a-\delta,\ a+\delta)$. [Hint: Solve $f(x) = \Sigma_{k=0}^{\infty} a_k(x-a)^k$ for $f(a)$, since we know that $a_0 = f(a)$, and then stare at what you have.]

3. Prove or give a counterexample: Let f be a smooth function on an open interval J, and s'pose that the Maclaurin series of f about $a \in J$ converges. Then f is analytic at a.

4. Binomial Series Prove: If $|x| < 1$, and r is any real number, then

$$(1+x)^r = 1 + rx + \frac{r(r-1)}{2!}x^2 + \frac{r(r-1)(r-2)}{3!}x^3 + ...$$

(Note, it is not enopugh simply to produce the Maclaurin series: you must prove that it converses.)

5. Prove that e is irrational.

We now switch texts to H.L. Royden, *Real Analysis* (Macmillan)

16. Sigma Algebras, Borel Sets and Outer Measure

First a comment: Let us look at our old friend,

$$f(x) = \begin{cases} 1 & \text{if } x \in Q \\ 0 & \text{if } x \,\square\, Q \end{cases} .$$

We can express f as the limit of an increasing sequence of integrable functions as follows: First enumerate all the rationals, writing Q = $\{r_1, r_2, ...\}$. Then, for each $n \geq 1$, define

$$f_n(x) = \begin{cases} 1 & \text{if } x \in \{r_1, r_2, ..., r_n\} \\ & \end{cases} .$$

ıerwise

Then the f_n are integrable (with integral 0) since they have only finitely many discontinuities. Further, f_n clearly increases to f, and yet f is not integrable. There is even a more compelling reason not to be happy with Riemann integrable functions: If we define the "distance" between two functions f and g by

$$||f - g|| = \int_a^b |f(x) - g(x)|\, dx$$

then there exist Cauchy sequences of Riemann integrable functions with respect to this distance function that do not converge to a Riemann integrable function, suggesting that we are missing something (what would the reals be without the irrationals?). In abstract terms (which could be made precise), the space of Riemann integrable functions is not complete. What is missing is a "good" theory of integration that permits us to integrate more functions than the Riemann integral.

We begin somewhat abstractly.

Definition 16.1 Let X be a set. A collection of subsets $\mathcal{A}$ of X is called an **algebra** if it is closed under the operations of (pairwise) union and complement.

Remark It follows from induction that algebras are closed under finite unions, and it follows by DeMargan's laws that they are closed under intersection:

$$(A \cap B) = (A \cap B)'' = (A' \cup B')'$$

Hence, by induction again, they are also closed under finite intersections.

Examples 16.2

A. Let X be arbitrary, and take $\mathcal{A} = \{X, \emptyset\}$

B. Let X be arbitrary, and take $\mathcal{A}$ to be the collection of all subsets of X.

C. Let $X = \mathbb{R}$, let $a \in \mathbb{R}$, and take $\mathcal{A} = \{\mathbb{R}, \emptyset, a, \mathbb{R}\text{-}\{a\}\}$

D. Let $X = \mathbb{R}$, and take $\mathcal{A}$ to be the set of all finite unions of intervals (including the degenerate intervals $\emptyset = (0, 0)$ and $\{a\} = [a, a]$).

We can get more examples from the following.

Proposition 16.3 (The Algebra Generated by a Collection of Subsets)
Given any collection C of subsets of X, there is an algebra $\mathcal{A}(C)$ of subsets of X with the following property:

If $\mathcal{A}$ is any algebra that contains C, then $\mathcal{A} \supset \mathcal{A}(C)$.

In other words, $\mathcal{A}(C)$ is the "smallest" algebra containing C. We call $\mathcal{A}(C)$ the **algebra generated by C.**

Proof Here is a constructive proof. Define sets $\mathcal{A} = \mathcal{A}_1 \subset \mathcal{A}_2 \subset .... \subset \mathcal{A}_n \subset ...$ inductively by taking the elements of $\mathcal{A}_{n+1}$ to be the finite unions of elements of $\mathcal{A}_n$ and their complements. Then take $\mathcal{A}(C) = \cup_n \mathcal{A}_n$. Then every element of $\mathcal{A}(C)$ is obtained from elements of $\mathcal{A}$ by taking a finite sequence of unions and complements, showing that every algebra containing C, being closed under these operations, must also contain $\mathcal{A}(C)$. Thus, it suffices to show that $\mathcal{A}(C)$ is an algebra. However, if S and T are in $\mathcal{A}(C)$, then they are in $\mathcal{A}_n$ for some n. But then S' and $S \cup T$ are in $\mathcal{A}_{n+1}$ and hence in $\mathcal{A}(C)$. ♣

Definition 16.4 A **σ-algebra** is an algebra $\mathcal{A}$ that is also closed under countable unions. That is, if $A_1, A_2, ..., A_n, ...$ are in $\mathcal{A}$, then so is $\cup_i A_i$.

Remark It now follows by De Morgan that σ-algebras are also closed under countable intersections.

The following proposition is proved in the exercises:

Proposition 16.5 (The σ-Algebra Generated by a Collection of Subsets)
Given any collection C of subsets of X, there is a σ-algebra $\mathcal{A}(C)$ of subsets of X with the following property:

If $\mathcal{A}$ is any σ-algebra that contains C, then $\mathcal{A} \supset \mathcal{A}(C)$.

In other words, $\mathcal{A}(C)$ is the "smallest" σ-algebra containing C. We call $\mathcal{A}(C)$ the **σ-algebra generated by C.**

Definition 16.6 The **collection $\mathcal{B}$ of Borel sets in R** is the σ-algebra generated by the collection of open intervals in R.

Examples 16.7 (of Borel Sets) The following subsets are in $\mathcal{B}$.
A. Single points **B.** Countable subsets of R (eg. Q)
C. Complements of countable sets (eg. R-Q)
D. The **Cantor Set** = $\cap_k J_k$ where the J_k are finite disjoint unions of *closed* intervals defined inductively as follows:

$J_0 = [0, 1]$

J_n is obtained from J_{n-1} by removing the middle 1/3 of each of its components.

The Cantor set consists of all real numbers between 0 and 1 that have a ternary decimal representation with all 0's and 2's (no 1's). It follows that C is uncountably infinite.

Definitions 16.8 From now on, we will work with the **extended real numbers,**

$\overline{\mathrm{R}} = \mathrm{R} \cup \{-\infty, \infty\}$.

The **length of the interval *I*** is the element of $\overline{\mathrm{R}}$ given by its usual length if it is finite, and ∞ if it is infinite. Similarly, the **sum of a series with positive terms** (whether or not it converges) is the element of $\overline{\mathrm{R}}$ defined in the natural way.
If $A \subset \mathrm{R}$, a **countable cover of *A* by open intervals** is a countable collection $\{I_1, I_2, ..., I_n, ...\}$ of open intervals whose union contains A.

OK now we are ready for the real definition

Definition 16.9 If $A \subset \mathrm{R}$, then the **outer measure, $\mu^*(A)$ of A**, is the element of $\overline{\mathrm{R}}$ given by

$$\mu^*(A) = \inf \Big\{ \sum_{k=1}^{\infty} |I_k| \ : \{I_1, I_2, ..., I_n, ...\} \text{ is a countable cover of } A \text{ by open intervals} \Big\}.$$

Examples 16.10
A. $\mu^*(\emptyset) = 0$ **B.** $\mu^*(\{a\}) = 0$
C. Countable sets have outer measure zero.

Proposition 16.11 (Properties of Outer Measure)
(a) If $A \supset B$, then $\mu^*(A) \geq \mu^*(B)$
(b) The outer measure of any interval is its length.

Proof Part (a) is in the exercise set. For part (b), let I be the interval in question. We can assume that I is non-degenerate by the above examples.
Case 1: I is a finite closed interval; $I = [a, b]$
Since $\{(a-\varepsilon, b+\varepsilon)\}$ is one of the possible countable covers in question, $\mu^*(I) \leq b-a + 2\varepsilon$ for all ε, so that $\mu^*(I) \leq b-a$. The hard part is to show that $\mu^*(I) \geq b-a$. Thus, let $\{I_1, I_2,$

..., I_n, ...} be any countable cover of I by open intervals. By Heine-Borel, $[a, b]$ is covered by finitely many of the I_i, say$\{I_1, I_2, ..., I_n\}$ (renumbering if necessary). It suffices to show that $\Sigma_{k=0}^{n} |I_k| \geq b-a$.

Since $a \in \cup_{k=0}^{n} I_k$, there is an I_k — call it I_1 by renumbering—with $a \in I_1 = (a_1, b_1)$, say. If $b_1 > b$, then we have $|I_1| = (b_1-a_1) \geq b-a = |I|$, and we are done. Otherwise, there is an interval—call it $I_2 = (a_2, b_2)$ containing b_1, so that

$$a_1 < a_2 < b_1 < b_2,$$

If $b \in I_2$, then $|I_1| + |I_2| > b_2-a_1 > b-a$, and again we are done. Continuing in this way, we eventually wind up with $b \in I_r$, and so we are eventually done.

Case 2: I is $(a, b]$, $[a, b)$, or (a, b).

The "easy" part ($\mu^*(I) \leq |I|$) still works here, and, for the "hard" part, notice that, for any $\varepsilon>0$, we can find a closed interval $J = [p, q] \subset I$ with $|J| \geq |I|.-\varepsilon$. Hence

$$\mu^*(I) \geq \mu^*(J) = |J| \geq |I|-\varepsilon$$

for all ε, showing that $\mu^*(I) \geq |I|$, as required for the hard part.

Case 3: I is an infinite interval.

In this case, I contains finite closed intervals or arbitrarily large length, and hence its outer measure is also arbitrarily large.

Proposition 16.12 (Outer Measure Is $\sigma-$ Subadditive)

If $\{A_n\}$ is any countable collection of subsets of R, then

$$\mu^*(\cup A_n) \leq \Sigma\mu^*(A_n).$$

Proof Let $\varepsilon>0$. By definition of $\mu^*(A_n)$ we can find, for each n, a countable cover C_n of A_n by open intervals such that the sum of the lengths of the intervals in C_n add to $\leq$ $\mu^*(A_n)+\varepsilon/2^n$. Since the $\cup C_n$ is now a cover of $\cup A_n$ by open intervals, it follows that $\mu^*(\cup a_n) \leq \Sigma_n(\mu^*(A_n)+\varepsilon/2^n) = \Sigma_n\mu^*(A_n) + \varepsilon$. Since ε is arbitrary, the result follows.

Note that we cannot expect there to be equality unless the union is a disjoint one. Even then, there are examples where the additivity is not strict.

Exercise Set 16

1. Describe explicitly (that is list all the elements) of the $\sigma-$algebras generated by the following collections of subsets of R.

(a) $\{\{0\}, \{1\}\}$ **(b)** $\{[0, 1]\}$ **(c)** $\{[0, 1], [2, 3]\}$ **(d)** $N = \{\{0\}, \{1\}, \{2\}, ... \}$

(you need not list all the elements in (d); just describe how to obtain them)

2. Prove: If $\mathcal{A}$ is any algebra of subsets, and $\{A_1, A_2, ..., A_n, ...\}$ is a sequence of subsets in $\mathcal{A}$, then there exists a sequence of *mutually disjoint* subsets $\{B_1, B_2, ..., B_n, ...\}$ such that $\cup_i A_i = \cup_i B_i$.

3. Prove Proposition 16.5.

4. Prove Proposition 16.11 (a): If $A \supset B$, then $\mu^*(A) \geq \mu^*(B)$.

5. Prove in two lines: if $\mu^*(B) = 0$, then $\mu^*(A \cup B) = \mu^*(A)$

17. Measurable Sets and Lebesgue Measure

Definition 17.1[1] The set $E \subset R$ is **(Lebesgue) measurable** if, for every $A \subset R$, one has

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E').$$

If $\mathcal{M}$ is the collection of measurable sets, then the restriction μ of μ^* to $\mathcal{M}$ is called **Lebesgue measure**. In other words,, if $E \in \mathcal{M}$, then its Lebesgue measure is defined by

$$\mu(E) = \mu^*(E).$$

Notes 17.2
(a) One always has $\mu^*(A) \leq \mu^*(A \cap E) + \mu^*(A \cap E')$ as a consequence of the σ-subadditive property.
(b) E measurable $\Rightarrow$ E' measurable.

Examples 17.3
A. R and Ø are automatically measurable.
B. $\mu^*(E) = 0 \Rightarrow E$ measurable (since $\mu^*(A) \geq \mu^*(A \cap E')$ regardless of E)
C. The interval $(a, +\infty)$ is measurable for each $a \in R$.
Indeed, if $A \subset R$ and $\{I_n\}$ is a countable cover of A by open intervals, then, for every $\varepsilon > 0$, $\{I_n \cap (a, +\infty)\}$ and $\{I_n \cap (-\infty, a+\varepsilon/2^n)\}$ are, respectively, countable covers of $A \cap (a, +\infty)$ and $A \cap (a, +\infty)'$ by open intervals, respectively, so that

$$\mu^*(A \cap (a, +\infty)) + \mu^*(A \cap (a, +\infty)') \leq \Sigma(I_n) + \varepsilon \qquad \text{(summing the lengths involved)}$$

Since the LHS is independent of n, we have

$$\mu^*(A \cap (a, +\infty)) + \mu^*(A \cap (a, +\infty)') \leq \Sigma(I_n),$$

and hence, by taking infimum of the RHS, we get

$$\mu^*(A \cap (a, +\infty)) + \mu^*(A \cap (a, +\infty)') \leq \mu^*(A).$$

We get all the other intervals (and more) by the following.

Proposition 17.4
The collection $\mathcal{M}$ of measurable sets is a α-algebra.

Proof First we show closure under finite unions (for which it suffices to show closure under pairwise union). Let $A \subset R$. Since E is measurable,

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E') \qquad \ldots \quad (1)$$

Since F is measurable,

$$\mu^*(A \cap E') = \mu^*(A \cap E' \cap F) + \mu^*(A \cap E' \cap F')$$
$$= \mu^*(A \cap E' \cap F) + \mu^*(A \cap (E \cup F)') \qquad \ldots \quad (2) \text{ by De Morgan}$$

Substituting (2) in (1) yields

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A \cap E' \cap F) + \mu^*(A \cap (E \cup F)') \qquad \ldots \quad (3)$$

However, since E is measurable,

$$\mu^*(A \cap (E \cup F)) = \mu^*(A \cap (E \cup F) \cap E) + \mu^*(A \cap (E \cup F) \cap E')$$
$$= \mu^*(A \cap E) + \mu^*(A \cap E' \cap F)$$

---

[1] This definition is due to Carathéodory.

Substituting this into (3) gives

$$\mu^*(A) = \mu^*(A\cap(E\cup F)) + \mu^*(A\cap(E\cup F)'),$$

as claimed. Note that, since we already know that complements of measurable sets are measurable, it follows that the collection of measurable sets is an algebra.

Before showing σ-additivity, we first establish the following claim:

Claim: if E and F are disjoint measurable sets, then, for every $A \subset$ R, one has

$$\mu^*(A\cap(E\cup F)) = \mu^*(A\cap E) + \mu^*(A\cap F).$$

Indeed, the left-hand side is equal to

$$\mu^*(A\cap(E\cup F)\cap E) + \mu^*(A\cap(E\cup F)\cap E')$$

by measurability of E, and this in turn is equal to

$$\mu^*(A\cap E) + \mu^*(A\cap F)$$

by a direct check on the sets level, since E and F are disjoint.

Now for σ-additivity. S'pose that $\{E_n\}$ is a collection of measurable sets. Then we can assume the E_n are mutually disjoint (why?)

$$\mu^*(A) = \mu^*(A\cap[\cup_{k=1}^{n} E_n]) + \mu^*(A\cap[\cup_{k=1}^{n} E_n]')$$

$$\geq \mu^*(A\cap[\cup_{k=1}^{n} E_n]) + \mu^*(A\cap[\cup_{k=1}^{\infty} E_n]')$$

for every n. Since the E_i are mutually disjoint, we can rewrite this as

$$\mu^*(A) \geq \Sigma_{k=1}^{n}\mu^*(A\cap E_n) + \mu^*(A\cap[\cup_{k=1}^{\infty} E_n]')$$

so that, taking limits,

$$\mu^*(A) \geq \Sigma_{k=1}^{\infty}\mu^*(A\cap E_n) + \mu^*(A\cap[\cup_{k=1}^{\infty} E_n]').$$

Finally, an application of σ-subadditivity of μ^* gives the result we want:

$$\mu^*(\mathrm{A}) \geq \mu^*(A\cap[\cup_{k=1}^{\infty} E_n]) + \mu^*(A\cap[\cup_{k=1}^{\infty} E_n]'). \quad \square$$

Notes 17.6

In the proof of the above proposition (see the claim) we have shown that, if E and F are measurable sets, then

(a) If E and F are disjoint, then $\mu(E\cup F) = \mu(E) + \mu(F)$.

(b) Hence, if $F \subset E$, then $\mu(E-F) = \mu(E) - \mu(F)$.

Corollary 17.6

$$\mathcal{M} \supset \mathcal{B},$$

that is, all the Borel sets are measurable.

Proposition 17.7 (Properties of Lebesgue Measure)

Let $\{E_n\}$ be a sequence of measurable sets. Then:

(a) $\mu(\cup E_n) \leq \Sigma\mu(E_n)$

(b) If the E_n are mutually disjoint, then $\mu(\cup E_n) = \Sigma\mu(E_n)$ **(σ-additivity)**

(*c*) If $E_n \uparrow E$ or $E_n \downarrow E$ then $\mu(E) = \lim_{n\varnothing\infty} \mu(E_n)$.

Proof (a) is Proposition 16.12.

(b) In view of (a), it suffices to show $\mu(\cup E_n) \geq \Sigma\mu(E_n)$. By the Claim in the proof of Proposition17.4, we have, with A = R,

$$\mu(\cup_{k=1}^{n} E_k) = \Sigma\mu(_{k=1}^{n} E_k).$$

But then

$$\mu(\cup_{k=1}^{\infty} E_k) \geq \mu(\cup_{k=1}^{n} E_k) = \Sigma_{k=1}^{n} \mu(E_k)$$

for all n, showing the result by letting $n \to \infty$.

(c) $E_n \uparrow E$ means that $E_n \subset E_{n+1}$ for all n, and that $\cup E_n = E$. Take $E_0 = \emptyset$ and $F_n = E_n - E_{n-1}$ for all $n \geq 1$. Then

$$\mu(E) = \mu(\cup E_n) = \mu(\cup F_n) = \Sigma\mu(F_n)$$
$$= \Sigma\mu(E_n - E_{n-1}) = \Sigma\mu(E_n) - \mu(E_{n-1}) = \lim_{n \varnothing \infty} \mu(E_n),$$

as required.

On the other hand, if $E_n \downarrow E$, then take $F_1 = \emptyset$, and $F_n = E_1 - E_n$. Then $F_n \uparrow (E_1 - E)$, and so

$$\mu(E_1) - \mu(E) = \mu(E_1 - E) = \lim_{n \varnothing \infty} \mu(F_n) = \lim_{n \varnothing \infty} \mu(E_1 - E_n) = \lim_{n \varnothing \infty} \mu(E_1) - \mu(E_n)$$
$$= \mu(E_1) - \lim_{n \varnothing \infty} \mu(E_n),$$

giving the result.

Exercise Set 17

1. (a) If $A \subset$ R and $x \in$ R, define $A+x = \{a+x \mid a \in A\}$. Prove that $\mu^*(A+x) = \mu^*(A)$ for all $x \in$ R.

(b) Show that A is measurable iff $A+x$ is measurable for all $x \in$ R.

2. Generalizing (1), let $A \subset [0, 1)$, let $\mp$ denote addition modulo 1, (so that, for example, $0.8 \mp 0.3 = 0.1$) Prove that $\mu^*(A \mp x) = \mu^*(A)$ for all $x \in$ ™$[0, 1)$.

(b) Show that A is measurable iff $A \mp x$ is measurable for all $x \in [0, 1)$.

3. Show that the Cantor set has measure 0.

4. Constructing An Unmeasurable Set

Define an equivalence relation on $[0, 1)$ by $a \approx b$ if $a - b \in$ Q. Let T be obtained by choosing, as its elements, exactly one member of each equivalence class.

(a) Prove that the sets $\{T \mp r \mid r \in$ Q$\}$ are mutually disjoint subsets of $(0, 1]$.

(b) Prove that $\cup_r (T \mp r) = (0, 1]$.

(c) Deduce, using σ-additivity and the results of #2, that T cannot be measurable.

18. Measurable Functions and the Lebesgue Integral

Proposition 18.1

Let f: R $\to$ R have a measurable domain D. Then the following are equivalent.

(a) $f^{-1}(a, +\infty)$ is measurable for every $a \in$ R.

(b) $f^{-1}[a, +\infty)$ is measurable for every $a \in$ R.

(c) $f^{-1}(-\infty, a)$ is measurable for every $a \in$ R.

(d) $f^{-1}(-\infty, a]$ is measurable for every $a \in$ R.

(e) $f^{-1}I$ is measurable for every interval I of R.

When one (and hence all) of these conditions hold, we say that f is **measurable.**

Proof

(a) ⇒ (b)
$f^{-1}[a, +\infty) = f^{-1}(\cap_n (a-1/n, +\infty)) = \cap_n f^{-1}(a-1/n, +\infty)$, which is a countable intersection of measurable sets, and hence measurable.

(b) ⇒ (c) $f^{-1}(-\infty, a) = D - f^{-1}[a, +\infty)$

(c) ⇒ (d) $f^{-1}(-\infty, a] = f^{-1}(\cap_n (a, a+1/n)) = \cap_n f^{-1}(a, a+1/n)$, which is a countable intersection of measurable sets, and hence measurable.

(d) ⇒ (a) $f^{-1}(a, +\infty) = D - f^{-1}(-\infty, a]$

(a), (b), (c), or (d) ⇒ (e) Every interval I is either R, or can be expressed as the intersection of two intervals of the above types, so the result follows.

(e) ⇒ (a) *A fortiori*.

(In particular, $f^{-1}(a)$ is measurable for each $a \in$ R, since $\{a\} = [a, a]$.)

Examples 18.2

A. The identity function f: $D \to$R; $f(x) = x$ on any measurable set D

B. Constant functions f: $D \to$R; $f(x) = K$ on any measurable set D

C. If D is measurable, define χ_D: $D \to$R by $\chi_D(x) = \begin{cases} 1 & \text{if } x \in D \\ 0 & \text{if } x \notin D \end{cases}$, called the **characteristic function of D**. Notice that $\chi_D^{-1}(I)$ is either Ø, D, or R, depending on whether I contains 0 and/or 1.

Moreover, we can add, subtract, multiply, etc. measurable functions by the following result:

Proposition 18.3 (Sums, Products, Quotients, Sups, Infs, and Limits)
(a) If f and g are measurable, then so are $f+g$, $f-g$, cf, fg, f/g, $\max\{f, g\}$, and $\min\{f, g\}$ wherever these are defined.
(b) If $\{f_n\}$ is any sequence of measurable functions, then $\sup_n\{f_n\}$ and $\inf_n\{f_n\}$ are measurable.
(c) If the f_n are measurable, and $f_n \to f$ pointwise, then f is measurable.

Proof

Part (a) follows from the following facts, either immediate or proved in the exercises:

$$(f+g)^{-1}(-\infty, a) = \cup_{r \in \mathbb{Q}} f^{-1}(-\infty, r) \cap g^{-1}(-\infty, a-r)$$

$(cf)^{-1}(-\infty, a) = f^{-1}(-\infty, a/c)$ if $c > 0$, and $f^{-1}(a/c, +\infty)$ if $c < 0$.

$$(f^2)^{-1}(a, +\infty) = f^{-1}(\sqrt{a}, +\infty)$$

$$fg = \frac{1}{2}[(f+g)^2 - f^2 - g^2]$$

$$(1/f)^{-1}(a, +\infty) = \begin{cases} f^{-1}(1/a, +\infty) & \text{if } a > 0 \\ f^{-1}(0, +\infty) \cup f^{-1}(-\infty, 1/a) & \text{if } a < 0 \end{cases}$$

$$\max\{f, g\}^{-1}(a, +\infty) = f^{-1}(a, +\infty) \cup g^{-1}(a, +\infty)$$

$$\min\{f, g\} = -\max\{-f, -g\}$$

(b) follows from the fact that $\sup_n\{f_n\}^{-1}(a, +\infty) = \cup_n f_n^{-1}(a, +\infty)$ and a similar fact for inf.

(c) If $f(x) > a$, then $f_n(x) > a$ for every $n \geq$ some N, so that $x \in \cup_N \cap_{n\geq N} f_n^{-1}(a, +\infty)$. Conversely, if $x \in \cup_N \cap_{n\geq N} f^{-1}(a, +\infty)$, then there exists an N with $f_n(x) > a$ for every $n \geq N$ giving $f(x) > a$.

Examples 18.4
A. polynomials, rational functions, linear combinations of characteristic functions

Definition 18.5 A **simple** function φ is a measurable function that assumes only finitely many values.

Note that products, sums, etc. of simple functions are simple. Also, if φ is simple, then $\varphi = \alpha_1\chi_{A_1} + \alpha_2\chi_{A_2} + \ldots + \alpha_n\chi_{A_n}$, where $A_i = \varphi^{-1}(\alpha_i)$. Thus, simple functions are measurable. In particular

Definition of the Lebesgue Integral
If $\varphi = \alpha_1\chi_{A_1} + \alpha_2\chi_{A_2} + \ldots + \alpha_n\chi_{A_n}$ is simple, then define the **integral of $\boldsymbol{\varphi}$** by

$$\int \varphi \, d\mu = \alpha_1\mu(A_1) + \alpha_2\mu(A_2) + \ldots + \alpha_n\mu(A_n)$$

If E is any measurable set, also define

$$\int_E \varphi \, d\mu = \int \varphi\cdot\chi_E \, d\mu$$

If f is any *measurable bounded function,* define

$$\int_E f \, d\mu = \sup_{\varphi\leq f} \int_E \varphi \, d\mu ,$$

where the sup is taken over all simple functions $\varphi \leq f$. Note that if f is unbounded, then we can't approximate it too easily by simple functions. However:
If f is any *non-negative measurable function,* define

$$\int_E f \, d\mu = \sup_{h\leq f} \int_E h \, d\mu$$

where the sup is taken over all bounded measurable functions h whose support has finite measure. Finally, if f is any measurable functions such that both f^+ *and* f^- are **integrable** (that is, their integrals are zero). Then define

$$\int_E f \, d\mu = \int_E f^+ \, d\mu - \int_E f^- \, d\mu .$$

Some nice properties of the Lebesgue Integral:
(1) ("Monotone Convergence Theorem)") If $f_n \to f$ is an increasing sequence of measurable functions, then $\int_E f_n \, d\mu \to \int_E f \, d\mu$

(2) ("Dominated Convergence Theorem)") If $f_n \to f$ and such that $|f_n| \leq$ some fixed integrable function g, then $\int_E f_n \, d\mu \to \int_E f \, d\mu$

Exercise Set 18

1. Show that:

(a) $(f+g)^{-1}(-\infty, a) = \cup_{r \in Q} f^{-1}(-\infty, r) \cap g^{-1}(-\infty, a-r)$

(b) $(1/f)^{-1}(a, +\infty) = \begin{cases} f^{-1}(1/a, +\infty) & \text{if } a > 0 \\ f^{-1}(0, +\infty) \cup f^{-1}(-\infty, 1/a) & \text{if } a < 0 \end{cases}$

2. We say that $f = g$ **almost everywhere** (ae) if $f(x) = g(x)$ for every x outside some set of measure zero. Show that every function equal, almost everywhere, to a measurable function is measurable.

3. (Extremely hard!) Show that, if f: $[a, b] \to$R is measurable and $\varepsilon > 0$, then there exists a continuous function g with

$$\mu\{x : |f(x) - g(x)| > \varepsilon\} < \varepsilon.$$

[Hint: First construct a suitable simple function. Then note that a measurable set has measure very close to a union of intervals. The latter fact allows you to approximate the simple function with a suitable step function, and finally get rid of the discontinuities.]