Lecture Notes on Lebesgue Integration and Fourier Analysis

by Stephen New

.

Chapter 1. Lebesgue Measure

1.1 Definition: When I is equal to any one of the bounded intervals (a,b), $[a,b)$, $(a,b]$ or $[a,b]$, where $a,b\in\mathbf{R}$ with $a\le b$, we define $|I|=b-a$. When I is equal to any of the unbounded intervals $(-\infty,a)$, $(-\infty,a]$, (a,∞), $[a,\infty)$ or $(-\infty,\infty)$, where $a\in\mathbf{R}$, we define $|I|=\infty$.

1.2 Definition: For a bounded set $A\subseteq\mathbf{R}$, the **Jordan outer content** of A is

$$c^*(A)=\inf\left\{\sum_{k=1}^n|I_k|\,\middle|\,n\in\mathbf{Z}^+,\text{ each }I_k\text{ is a bounded open interval and }A\subseteq\bigcup_{k=1}^n I_k\right\}.$$

1.3 Theorem: *(Properties of Jordan Outer Content) Let $A,B\subseteq\mathbf{R}$ be bounded.*
(1) (Translation) If $a\in\mathbf{R}$ and $0\ne r\in\mathbf{R}$ then $c^*(a+A)=c^*(A)$.
(2) (Scaling) If $0\ne r\in\mathbf{R}$ then $c^*(rA)=|r|\,c^*(A)$.
(3) (Inclusion) If $A\subseteq B$ then $c^*(A)\le c^*(B)$.
(4) If A is finite then $c^*(A)=0$.
(5) If I is a bounded interval then $c^*(I)=|I|$.
(6) (Subadditivity) We have $c^*(A\cup B)\le c^*(A)+c^*(B)$.
(7) We have $c^*(\overline{A}\,)=c^*(A)$.

Proof: The proof is left as an exercise.

1.4 Exercise: Show that when $A\subseteq\mathbf{R}$ and I and J are bounded intervals with $A\subseteq I\subseteq J$ we have $|I|-c^*(I\setminus A)=|J|-c^*(J\setminus A)$.

1.5 Definition: For a bounded set $A\subseteq\mathbf{R}$, we say that A has (a well-defined) **Jordan content** when

$$c^*(A)=|I|-c^*(I\setminus A)$$

where I is any interval which contains A and, in this case, we define the **Jordan content** of A to be $c(A)=c^*(A)$.

1.6 Exercise: Show that $\mathbf{Q}\cap[0,1]$ does not have a well-defined Jordan content.

1.7 Theorem: *(Properties of Content) Let $A,B\subseteq\mathbf{R}$ be bounded.*
(1)(Translation) If $a\in\mathbf{R}$ then $a+A$ has Jordan content if and only if A does.
(2) (Scaling) If $0\ne r\in\mathbf{R}$ then rA has Jordan content if and only if A does.
(3) If $c^*(A)=0$ then A has Jordan content.
(4) If A and B have Jordan content then so do $A\cup B$, $A\cap B$ and $A\setminus B$.
(5) Every bounded interval has Jordan content.
(6) The set A has Jordan content if and only if $c^*(\overline{A}\setminus A^0)=0$.

Proof: The proof is left as an exercise.

1.8 Definition: For a set $A \subseteq \mathbf{R}$, the (Lebesgue) **outer measure** of A is

$$\lambda^*(A) = \inf\left\{ \sum_{n=1}^{\infty} |I_n| \,\middle|\, \text{each } I_n \text{ is a bounded open interval and } A \subseteq \bigcup_{n=1}^{\infty} I_n \right\}.$$

1.9 Theorem: *(Properties of Outer Measure) Let $A, B \subseteq \mathbf{R}$ and let $A_k \subseteq \mathbf{R}$ for $k \in \mathbf{Z}^+$.*
(1) (Translation) If $a \in \mathbf{R}$ then $\lambda^*(a + A) = \lambda^*(A)$.
(2) (Scaling) If $0 \neq r \in \mathbf{R}$ then $\lambda^*(rA) = |r|\lambda^*(A)$.
(3) (Inclusion) If $A \subseteq B$ then $\lambda^*(A) \leq \lambda^*(B)$.
(4) If A is finite or countable then $\lambda^*(A) = 0$.
(5) If I is an interval then $\lambda^*(I) = |I|$.
(6) (Subadditivity) We have $\lambda^*\left(\bigcup_{n=1}^{\infty} A_n \right) \leq \sum_{n=1}^{\infty} \lambda^*(A_n)$.

Proof: We leave the proofs of parts (1), (2) and (3) as an exercise. We prove Part (4) in the case that A is countable. Let $A = \{a_1, a_2, a_3, \cdots\}$. Let $\epsilon > 0$. For each $n \in \mathbf{Z}^+$, let $I_n = \left(a_n - \frac{\epsilon}{2^n}, a_n + \frac{\epsilon}{2^n}\right)$. Then $A \subseteq \bigcup_{n=1}^{\infty} I_n$ so we have $\lambda^*(A) \leq \sum_{n=1}^{\infty} |I_n| = 2\epsilon$. Since $0 \leq \lambda^*(A) < 2\epsilon$ for every $\epsilon > 0$, it follows that $\lambda^*(A) = 0$.

Let us prove Part (5). When I is a degenerate interval (so I is empty or has only one point) we know, from Part (4), that $\lambda^*(I) = 0$. Suppose that I is a nondegenerate bounded interval, say I is equal to one of the intervals (a, b), $[a, b)$, $(a, b]$ or $[a, b]$ where $a < b$. Let $\epsilon > 0$, let $I_1 = (a - \epsilon, b + \epsilon)$ and let $I_n = \emptyset$ for $n \geq 2$. Then $I \subseteq \bigcup_{n=1}^{\infty} I_n$ so we have $\lambda^*(I) \leq \sum_{n=1}^{\infty} |I_n| = b - a + 2\epsilon$. Since $\epsilon > 0$ was arbitrary, it follows that $\lambda^*(I) \leq b - a$. It remains to show that $\lambda^*(I) \geq b - a$. Let $I_1, I_2, I_3, \cdots$ be any bounded open intervals such that $I \subseteq \bigcup_{n=1}^{\infty} I_n$. Let $0 < \epsilon < \frac{b-a}{2}$ and consider the compact interval $K = [a + \epsilon, b - \epsilon] \subset I$. Note that $\mathcal{U} = \{I_1, I_2, I_3, \cdots\}$ is an open cover of K. Choose a finite subset $\mathcal{V} \subseteq \mathcal{U}$ so that $K \subseteq \bigcup_{J \in \mathcal{V}} J$. Choose $J_1 = (a_1, b_1) \in \mathcal{V}$ so that $a_1 < a - \epsilon < b_1$. If $b_1 \leq b - \epsilon$ then choose $J_2 = (a_2, b_2) \in \mathcal{V}$ so that $a_2 < b_1 < b_2$. If $b_2 \leq b - \epsilon$ then choose $J_3 = (a_3, b_3) \in \mathcal{V}$ so that $a_3 < b_2 < b_3$. Continue this procedure until we have chosen $J_\ell = (a_\ell, b_\ell) \in \mathcal{V}$ with $b_\ell > b - \epsilon$, and note that $K \subseteq J_1 \cup J_2 \cup \cdots \cup J_\ell$ and $\{J_1, J_2, \cdots, J_\ell\} \subseteq \mathcal{V} \subseteq \mathcal{U}$. We have

$$\begin{aligned}
\sum_{n=1}^{\infty} |I_n| \geq \sum_{n=1}^{\ell} |J_n| &= (b_1 - a_1) + (b_2 - a_2) + \cdots + (b_\ell - a_\ell) \\
&> \big(a_2 - (a + \epsilon)\big) + (a_3 - a_2) + (a_4 - a_3) + \cdots + (a_\ell - a_\ell - 1) + \big((b - \epsilon) - a_\ell\big) \\
&= b - a - 2\epsilon.
\end{aligned}$$

Since ϵ was arbitrarily small it follows that $\sum_{n=1}^{\infty} |I_n| \geq b - a$. Since this is true for all bounded open intervals $I_1, I_2, I_3, \cdots$ which cover I, it follows that $\lambda^*(I) \geq b - a$, as required.

When I is an unbounded interval, we must have $\lambda^*(I) = \infty$ because for every $R > 0$ we can choose a bounded interval $J \subseteq I$ with $|J| > R$ and then we have $\lambda^*(I) \geq \lambda^*(J) > R$.

To prove Part (6), let $A_1, A_2, A_3, \cdots \subseteq \mathbf{R}$. Let $\epsilon > 0$. For each $n \in \mathbf{Z}^+$, choose open bounded intervals $I_{n,1}, I_{n,2}, I_{n,3}, \cdots$ so that $A_n \subseteq \bigcup\limits_{k=1}^{\infty} I_{n,k}$ and $\sum\limits_{k=1}^{\infty} |I_{n.k}| \leq \lambda^*(A_n) + \frac{\epsilon}{2^n}$. Then we have $\bigcup\limits_{n=1}^{\infty} A_n \subseteq \bigcup\limits_{n,k=1}^{\infty} I_{n,k}$ so that

$$\lambda^*\Big(\bigcup_{n=1}^{\infty} A_n\Big) \leq \sum_{n,k=1}^{\infty} |I_{n,k}| \leq \sum_{n=1}^{\infty} \big(\lambda^*(A_n) + \tfrac{\epsilon}{2^n}\big) = \sum_{n=1}^{\infty} \lambda^*(A_n) + \epsilon.$$

Since $\epsilon > 0$ was arbitrary, we have $\lambda^*\Big(\bigcup\limits_{n=1}^{\infty} A_k\Big) \leq \sum\limits_{n=1}^{\infty} \lambda^*(A_n)$, as required.

1.10 Definition: For $A \subseteq \mathbf{R}$, we say that A is (Lebesgue) **measurable** when for every set $X \subseteq \mathbf{R}$ we have

$$\lambda^*(X) = \lambda^*(X \cap A) + \lambda^*(X \setminus A).$$

When A is measurable, we define the (Lebesgue) **measure** of A to be $\lambda(A) = \lambda^*(A)$. We let $\mathcal{M}$ denote the set of all measurable subsets of $\mathbf{R}$.

1.11 Note: For any sets $A, X \subseteq \mathbf{R}$, we have $X = (X \cap A) \cup (X \setminus A)$ and so (by subadditivity) $\lambda^*(X) \leq \lambda^*(X \cap A) + \lambda^*(X \setminus A)$. Thus a set $A \subseteq \mathbf{R}$ is measurable if and only if for every set $X \subseteq \mathbf{R}$ we have

$$\lambda^*(X) \geq \lambda^*(X \cap A) + \lambda^*(X \setminus A).$$

1.12 Theorem: *(Properties of Measure) Let $A, B, A_k \subseteq \mathbf{R}$ for $k \in \mathbf{Z}^+$.*
(1) If $a \in \mathbf{R}$ then A is measurable if and only if $a + A$ is measurable.
(2) If $0 \neq r \in \mathbf{R}$ then A is measurable if and only if rA is measurable.
(3) $\emptyset$ and $\mathbf{R}$ are measurable.
(4) If $\lambda^*(A) = 0$ then A is measurable.
(5) If A is measurable then so is $A^c = \mathbf{R} \setminus A$.
(6) If A and B are measurable then so are $A \cup B$, $A \cap B$ and $A \setminus B$.
(7) Every interval is measurable.
(8) If $A_1, A_2, A_3, \cdots$ are measurable then so are $\bigcup\limits_{k=1}^{\infty} A_k$ and $\bigcap\limits_{k=1}^{\infty} A_k$.
(9) If $A_1, A_2, A_3, \cdots$ are measurable and disjoint then $\lambda\Big(\bigcup\limits_{k=1}^{\infty} A_k\Big) = \sum\limits_{k=1}^{\infty} \lambda(A_k)$.

Proof: We leave the proofs of Parts (1) and (2) as an exercise. To prove Part (3), note that $\emptyset$ and $\mathbf{R}$ are measurable because for every set $X \subseteq \mathbf{R}$ we have

$$\lambda^*(X \cap \emptyset) + \lambda^*(X \setminus \emptyset) = \lambda^*(\emptyset) + \lambda^*(X) = \lambda^*(X), \text{ and}$$
$$\lambda^*(X \cap \mathbf{R}) + \lambda^*(X \setminus \mathbf{R}) = \lambda^*(X) + \lambda^*(\emptyset) = \lambda^*(X).$$

To prove Part (4), let $A \subseteq \mathbf{R}$ and suppose that $\lambda^*(A) = 0$. Let $X \subseteq \mathbf{R}$. Since $X \cap A \subseteq A$ and $X \setminus A \subseteq X$ we have

$$\lambda^*(X \cap A) + \lambda^*(X \setminus A) \leq \lambda^*(A) + \lambda^*(X) = \lambda^*(X).$$

Part (5) holds because if $A \subseteq \mathbf{R}$ is measurable and $X \subseteq \mathbf{R}$ then, since $X \cap A^c = X \setminus A$ and $X \setminus A^c = X \cap A$, we have

$$\lambda^*(X \cap A^c) + \lambda^*(X \setminus A^c) = \lambda^*(X \setminus A) + \lambda^*(X \cap A) = \lambda^*(X).$$

To prove Part (6), suppose that A and B are measurable and let $X \subseteq \mathbf{R}$. Then

$$\begin{aligned}\lambda^*(X) &= \lambda^*(X\cap A)+\lambda^*(X\setminus A)\ \text{, since } A \text{ is measurable}\\ &= \lambda^*(X\cap A)+\lambda^*\big((X\setminus A)\cap B\big)+\lambda^*\big((X\setminus A)\setminus B\big)\ \text{, since } B \text{ is measurable}\\ &= \lambda^*(X\cap A)+\lambda^*\big((X\setminus A)\cap B\big)+\lambda^*\big(X\setminus (A\cup B)\big)\\ &\geq \lambda^*\big(X\cap (A\cup B)\big)+\lambda^*\big(X\setminus (A\cup B)\big)\ \text{, by subadditivity}\end{aligned}$$

since $(X\cap A)\cup\big((X\setminus A)\cap B\big)=X\cap(A\cup B)$. This shows that $A\cup B$ is measurable. Using Part (5), it follows that $A\cap B$ is measurable because $A\cap B=(A^c\cup B^c)^c$ and hence that $A\setminus B$ is measurable because $A\setminus B=A\cap B^c$.

Let us prove Part (7) in the case of a nonempty bounded open interval. Let $I=(a,b)$ where $a<b$. Let $X\subseteq\mathbf{R}$. Let $\epsilon>0$. Choose open bounded intervals $I_1,I_2,I_3,\cdots$ so that $X\subseteq\bigcup\limits_{n=1}^{\infty} I_n$ and $\sum\limits_{n=1}^{\infty}|I_n|<\lambda^*(X)+\epsilon$. For $n\in\mathbf{Z}^*$, let $J_n=I_n\cap(a,b)$, $K_n=I_n\cap(-\infty,a)$ and $L_n=I_n\cap(b,\infty)$. Then $X\cap I\subseteq\bigcup\limits_{n=1}^{\infty}J_n$ so that $\lambda^*(X\cap I)\leq\sum\limits_{n=1}^{\infty}|J_n|$ and $X\setminus I\subseteq (a-\epsilon,a+\epsilon)\cup(b-\epsilon,b+\epsilon)\cup\bigcup\limits_{n=1}^{\infty}K_n\cup\bigcup\limits_{n=1}^{\infty}L_n$ so that $\lambda^*(X\setminus I)\leq 4\epsilon+\sum\limits_{n=1}^{\infty}|K_n|+\sum\limits_{n=1}^{\infty}|L_n|$ and so we have

$$\lambda^*(X\cap I)+\lambda^*(X\setminus I)\leq 4\epsilon+\sum_{n=1}^{\infty}\big(|I_n|+|J_n|+|K_n|\big)=4\epsilon+\sum_{n=1}^{\infty}|I_n|<\lambda^*(X)+5\epsilon.$$

Since $\epsilon>0$ was arbitrary, we have $\lambda^*(X\cap I)+\lambda^*(X\setminus I)\leq\lambda^*(X)$, Since $X\subseteq\mathbf{R}$ was arbitrary, we se that I is measurable.

Before proving Parts (8) and (9) we remark that for $A,B\subseteq\mathbf{R}$, if A is measurable and $A\cap B=\emptyset$ then for all $X\subseteq\mathbf{R}$ we have

$$\begin{aligned}\lambda^*\big(X\cap(A\cup B)\big)&=\lambda^*\big((X\cap(A\cup B))\cap A\big)+\lambda^*\big((X\cap(A\cup B)\setminus A\big)\\ &=\lambda^*(X\cap A)+\lambda^*(X\cap B)\end{aligned}$$

It follows, inductively, that if $A_1,A_2,\cdots,A_n\subseteq\mathbf{R}$ are measurable and disjoint then for all $X\subseteq\mathbf{R}$ we have

$$\lambda^*\big(X\cap\bigcup_{k=1}^{n}A_k\big)=\sum_{k=1}^{n}\lambda^*(X\cap A_k).$$

Now let $A_1,A_2,A_3,\cdots\subseteq\mathbf{R}$ be measurable and disjoint and let $X\subseteq\mathbf{R}$. For all $n\in\mathbf{Z}^+$ we have

$$\begin{aligned}\sum_{k=1}^{n}\lambda^*(X\cap A_k)&=\lambda^*\big(X\cap\bigcup_{k=1}^{n}A_k\big)\quad\text{, by the above remark,}\\ &\leq\lambda^*\big(X\cap\bigcup_{k=1}^{\infty}A_k\big)\quad\text{, since } X\cap\bigcup_{k=1}^{n}A_k\subseteq X\cap\bigcup_{k=1}^{\infty}A_k,\\ &=\lambda^*\big(\bigcup_{k=1}^{\infty}(X\cap A_k)\big)\text{, since } X\cap\bigcup_{k=1}^{\infty}A_k=\bigcup_{k=1}^{\infty}(X\cap A_k),\\ &\leq\sum_{k=1}^{\infty}\lambda^*(X\cap A_k)\quad\text{, by subadditivity.}\end{aligned}$$

Taking the limit as n tends to infinity gives

$$\lambda^*\big(X \cap \bigcup_{k=1}^{\infty} A_k\big) = \sum_{k=1}^{\infty} \lambda^*(X \cap A_k).$$

The special case $X = \mathbf{R}$ gives the formula $\lambda^*\big(\bigcup_{k=1}^{\infty} A_k\big) = \sum_{k=1}^{\infty} \lambda^*(A_k)$ for Part (9). For all $n \in \mathbf{Z}^+$ we have

$$\begin{aligned}\lambda^*(X) &= \lambda^*\big(X \cap \sum_{k=1}^{n} A_k\big) + \lambda^*\big(X \setminus \bigcup_{k=1}^{n} A_k\big) \\ &= \sum_{k=1}^{n} \lambda^*(X \cap A_k) + \lambda^*\big(X \setminus \bigcup_{k=1}^{n} A_k\big) \\ &\geq \sum_{k=1}^{n} \lambda^*(X \cap A_k) + \lambda^*\big(X \setminus \bigcup_{k=1}^{\infty} A_k\big)\end{aligned}$$

Taking the limit as n tends to infinity gives

$$\begin{aligned}\lambda^*(X) &\geq \sum_{k=1}^{\infty} \lambda^*(X \cap A_k) + \lambda^*\big(X \setminus \bigcup_{k=1}^{\infty} A_k\big) \\ &= \lambda^*\big(X \cap \bigcup_{k=1}^{\infty} A_k\big) + \lambda^*\big(X \setminus \bigcup_{k=1}^{\infty} A_k\big)\end{aligned}$$

so that $\bigcup_{k=1}^{\infty} A_k$ is measurable, proving Part (8) in the case that the sets A_k are disjoint. To complete the proof of Part (8) in the case that $A_1, A_2, A_3, \cdots \subseteq \mathbf{R}$ are measurable (but not necessarily disjoint) simply note that

$$\bigcup_{k=1}^{\infty} A_k = A_1 \cup (A_2 \setminus A_1) \cup (A_3 \setminus (A_1 \cup A_2)) \cup (A_4 \setminus (A_1 \cup A_2 \cup A_3)) \cup \cdots$$

which is a countable union of disjoint measurable sets.

Finally, we recall that we only proved Part (7) in the case of a bounded open interval. We note that every interval can be obtained from bounded open intervals by performing complements and countable unions or intersections, and so every interval is measurable.

1.13 Corollary: *Let $A_1, A_2, A_3, \cdots \subseteq \mathbf{R}$ be measurable sets.*

(1) If $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots$ then $\lambda\Big(\bigcup_{n=1}^{\infty} A_n\Big) = \lim_{n\to\infty} \lambda(A_n)$.

(2) If $A_1 \supseteq A_2 \supseteq A_3 \supseteq \cdots$ and A_m is finite for some $m \in \mathbf{Z}^+$ then $\lambda\Big(\bigcap_{n=1}^{\infty} A_n\Big) = \lim_{n\to\infty} \lambda(A_n)$.

Proof: To prove Part (1), suppose that $A_1 \subseteq A_2 \subseteq A_3 \subseteq \cdots$. Let $B_1 = A_1$ and $B_k = A_k \setminus A_{k-1}$ for $k \geq 2$, Then the sets B_k are measurable and disjoint and we have $A_n = \bigcup_{k=1}^{n} B_k$ for all $n \in \mathbf{Z}^+$ and also $\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} B_n$. Thus

$$\lambda\big(\bigcup_{n=1}^{\infty} A_n\big) = \lambda\big(\bigcup_{n=1}^{\infty} B_n\big) = \sum_{n=1}^{\infty} \lambda(B_n) = \lim_{n\to\infty} \sum_{k=1}^{n} \lambda(B_k) = \lim_{n\to\infty} \lambda\big(\bigcup_{k=1}^{n} B_k\big) = \lim_{n\to\infty} \lambda(A_n).$$

This proves Part (1), and Part (2) follows from Part (1) by taking complements in A_m.

1.14 Theorem: *All open and closed sets in* $\mathbf{R}$ *are measurable.*

Proof: Recall that every set in $\mathbf{R}^n$ (or any metric or topological space) is equal to the disjoint union of its connected components, and recall that the connected components of an open set are all open. Note that the set of connected components of an open set in $\mathbf{R}^n$ is at most countable because we can choose an element of $\mathbf{Q}^n$ inside each of the open connected components. Also recall that the connected sets in $\mathbf{R}$ are the intervals in $\mathbf{R}$. It follows that every nonempty open set in $\mathbf{R}$ is equal to the finite or countable disjoint union of its connected components, each of which is a nonempty open interval. Thus every open set in $\mathbf{R}$ is measurable. We also remark that when the connected components of the nonempty open set $U\subseteq\mathbf{R}$ are the disjoint open intervals $I_1,I_2,I_3,\cdots$ we have $\lambda(U)=\sum\limits_{k\geq 1}|I_k|$. Closed sets are also measurable because every closed set is the complement of an open set.

1.15 Corollary: *For* $A\subseteq\mathbf{R}$ *we have*

$$\lambda^*(A)=\inf\big\{\lambda(U)\,\big|\,U\subseteq\mathbf{R}\text{ is open with }A\subseteq U\big\}.$$

1.16 Example: The (standard) **Cantor set** is the set $C\subseteq[0,1]$ constructed as follows. Let $C_0=[0,1]$. Let I_1 be the open middle third of C_0, that is let $I_1=\left(\frac{1}{3},\frac{2}{3}\right)$, and let $C_1=A_0\setminus U_1=\left[0,\frac{1}{3}\right]\cup\left[\frac{2}{3},1\right]$. Let I_2 and I_3 be the open middle thirds of the two component intervals of C_1, that is let $I_2=\left(\frac{1}{9},\frac{2}{9}\right)$ and $I_3=\left(\frac{7}{9},\frac{8}{9}\right)$, and let $C_2=C_1\setminus(I_2\cup I_3)$. Having constructed the set C_k, which is the disjoint union of 2^k closed intervals each of length $\frac{1}{3^k}$, let $I_{2^k},I_{2^k+1},\cdots,I_{2^{k+1}-1}$ be the open middle thirds of these 2^k component intervals and let $C_{k+1}=C_k\setminus(I_{2^k},I_{2^k+1},\cdots,I_{2^{k+1}-1})$. Finally, we let

$$C=\bigcap_{k=1}^{\infty}C_k.$$

Since $C_0\supseteq C_1\supseteq C_2\supseteq\cdots$, and since each C_k is the disjoint union of 2^k closed intervals each of size $\frac{1}{3^k}$ so that $\lambda(C_k)=\left(\frac{2}{3}\right)^k$, we have

$$\lambda(C)=\lim_{k\to\infty}\lambda(C_k)=0.$$

Note that C_k is the set of all numbers $x\in[0,1]$ which can be written in base 3 such that the the first k digits of x are not equal to 1, and so C is the set of all numbers $x\in[0,1]$ which can be written in base 3 with none of the digits of x equal to 1, and it follows that the cardinality of C is $|C|=2^{\aleph_0}$.

1.17 Example: We can construct a (generalized) **Cantor set** $C\subseteq[0,1]$, having any desired value for the measure $\lambda(C)<1$ as follows. Let $0\leq m<1$. Choose a sequence of positive real numbers $a_1,a_2,\cdots$ with $\sum\limits_{k=1}^{\infty}a_k=1-m$. Let $C_0=[0,1]$ and note that $\lambda(C_0)=1$. Choose an open interval $I_1\subseteq C_0$ with $\lambda(I_1)=a_1$ such that $C_0\setminus I_1$ is the disjoint union of two nondegenerate closed intervals each of measure less than $\frac{1}{2}$. Let $C_1=C_0\setminus I_1$ and note that $\lambda(C_1)=1-a_1$. Having constructed the set C_k, which is the disjoint union of 2^k nondegenerate closed intervals each of measure less than $\frac{1}{2^k}$ and having total measure $\lambda(C_k)=1-(a_1+a_2+\cdots+a_k)$, we choose 2^k open intervals $I_{2^k},I_{2^k+1},\cdots,I_{2^{k+1}-1}$ which are contained in each of the 2^k component intervals of C_k so that the set $C_{k+1}=C_k\setminus(I_{2^k}\cup\cdots\cup I_{2^{k+1}-1})$ is the disjoint union of 2^{k+1} non-degenerate closed intervals each of measure less than $\frac{1}{2^{k+1}}$ and having total measure $\lambda(C_{k+1})=1-(a_1+\cdots+a_{k+1})$. Finally, we let $C=\bigcap\limits_{k=1}^{\infty}C_k$ and note that $\lambda(C)=\lim\limits_{k\to\infty}\lambda(C_k)=1-\sum\limits_{k=1}^{\infty}a_k=m$.

1.18 Theorem: *Let $\mathcal{M}$ be the set of all measurable subsets of* $\mathbf{R}$. *Then* $|\mathcal{M}| = 2^{2^{\aleph_0}}$.

Proof: Let C be the standard Cantor set. Because $\lambda(C) = 0$ it follows that every subset of C is measurable. Because $|C| = 2^{\aleph_0}$ we have

$$2^{2^{\aleph_0}} = \big|\{A|A \subseteq \mathbf{R}\}\big| \geq |\mathcal{M}| \geq \big|\{A|A \subseteq C\}\big| = 2^{2^{\aleph_0}}.$$

1.19 Theorem: *There exists a nonmeasurable set in* $\mathbf{R}$.

Proof: Define an equivalence relation on the set $[0,1]$ by defining $x \sim y$ when $y - x \in \mathbf{Q}$. Let C denote the set of equivalence classes. For each $c \in C$, choose an element $x_c \in c$ and let $A = \{x_c | c \in C\} \subseteq [0,1]$. We shall prove that the set A is not measurable. Let $\mathbf{Q} \cap [0,2] = \{a_1, a_2, a_3, \cdots\}$, with the a_k distinct. For each $k \in \mathbf{Z}^+$, let $A_k = a_k + A \subseteq [0,3]$. We claim that the sets A_k are disjoint. Let $k, \ell \in \mathbf{Z}^+$ and suppose that $A_k \cap A_\ell \neq \emptyset$. Choose $y \in A_k \cap A_\ell$, say $y = a_k + x_c = a_\ell + x_d$ where $c, d \in C$. Since $x_c - x_d = a_\ell - a_k \in \mathbf{Q}$ we have $x_c \sim x_d$ and hence $c = d$ (since we only chose one element from each class). Since $c = d$ we have $x_c = x_d$, hence $a_k = a_\ell$, and hence $k = \ell$. Thus the sets A_k are disjoint, as claimed. Next, we claim that $[1,2] \subseteq \bigcup_{k=1}^{\infty} A_k$. Let $y \in [1,2]$. Since $y - 1 \in [0,1]$ we have $y - 1 \in c$ for some $c \in C$. Since $y - 1 \in c$ we have $y - 1 - x_c \in \mathbf{Q}$ hence also $y - x_c \in \mathbf{Q}$. Since $y \in [1,2]$ and $x_c \in [0,1]$ we have $y - x_c \in [0,2]$. Since $y - x_c \in \mathbf{Q} \cap [0,2]$ we have $y - x_c = a_k$ for some $k \in \mathbf{Z}^+$ so that $y \in A_k$. This proves that $[1,2] \subseteq \bigcup_{k=1}^{\infty} A_k$.

Suppose, for a contradiction, that the set A is measurable. By translation, each of the sets $A_k = a_k + A$ is measurable with $\lambda(A_k) = \lambda(A)$. Since the sets A_k are disjoint and measurable, additivity gives

$$\lambda\Big(\bigcup_{k=1}^{\infty} A_k\Big) = \sum_{k=1}^{\infty} \lambda(A_k) = \sum_{k=1}^{\infty} \lambda(A) = \begin{cases} 0 & \text{, if } \lambda(A) = 0, \\ \infty & \text{, if } \lambda(A) > 0. \end{cases}$$

But since $[0,1] \subseteq \bigcup_{k=1}^{\infty} A_k \subseteq [0,3]$ we also have $1 \leq \lambda\big(\bigcup_{k=1}^{\infty} A_k\big) \leq 3$, giving the desired contradiction.

1.20 Notation: Let X be a set. For any set $\mathcal{C}$ of subsets of X we write

$$\mathcal{C}_\sigma = \Big\{ \bigcup_{k=1}^{\infty} A_k \Big| \text{ each } A_k \in \mathcal{C} \Big\} \text{ and } \mathcal{C}_\delta = \Big\{ \bigcap_{k=1}^{\infty} A_k \Big| \text{ each } A_k \in \mathcal{C} \Big\}.$$

Note that $\mathcal{C}_{\sigma\sigma} = \mathcal{C}_\sigma$ and $\mathcal{C}_{\delta\delta} = \mathcal{C}_\delta$.

1.21 Definition: Let X be a set. A **σ-algebra** in X is a set $\mathcal{C}$ of subsets of X such that
(1) $\emptyset \in \mathcal{C}$,
(2) if $A \in \mathcal{C}$ then $A^c = X \setminus A \in \mathcal{C}$, and
(3) if $A_1, A_2, A_3, \cdots \in \mathcal{C}$ then $\bigcup_{k=1}^{\infty} A_k \in \mathcal{C}$.

Note that when $\mathcal{C}$ is a σ-algebra in X we have $\mathcal{C}_\sigma = \mathcal{C}$ and $\mathcal{C}_\delta = \mathcal{C}$.

1.22 Notation: In a metric space (or topological space) X, we let $\mathcal{G}$ denote the set of all open sets in X and we let $\mathcal{F}$ denote the set of all closed subsets of X. Note that $\mathcal{G}_\sigma = \mathcal{G}$ and $\mathcal{F}_\delta = \mathcal{F}$.

1.23 Example: For any set X, the set $\{\emptyset, X\}$ and the set $\mathcal{P}(X)$ of all subsets of X are σ-algebras in X, The set $\mathcal{M} = \mathcal{M}(\mathbf{R})$ of all measurable sets in $\mathbf{R}$ is a σ-algebra in $\mathbf{R}$.

1.24 Note: Note that given any set $\mathcal{C}$ of subsets of a set X there exists a unique smallest σ-algebra in X which contains $\mathcal{C}$, namely the intersection of all σ-algebras in X which contain $\mathcal{C}$.

1.25 Definition: In a metric space (or topological space) X, the **Borel** σ-algebra $\mathcal{B}$ is the smallest σ-algebra in X which contains $\mathcal{G}$ (hence also $\mathcal{F}$). The elements of $\mathcal{B}$ are called **Borel** sets. Note that $\mathcal{B}$ contains all of the sets $\mathcal{G}, \mathcal{G}_\delta, \mathcal{G}_{\delta\sigma}, \mathcal{G}_{\sigma\delta\sigma}, \cdots$ and all of the sets $\mathcal{F}, \mathcal{F}_\sigma, \mathcal{F}_{\sigma\delta}, \mathcal{F}_{\sigma\delta\sigma}, \cdots$.

1.26 Exercise: Show that $\mathcal{F} \subseteq \mathcal{G}_\delta$ or, equivalently, that $\mathcal{G} \subseteq \mathcal{F}_\sigma$,

1.27 Theorem: *All Borel sets in* $\mathbf{R}$ *are measurable.*

Proof: The set $\mathcal{M}$ of all measurable subsets of $\mathbf{R}$ is a σ-algebra which contains $\mathcal{G}$, and the Borel σ-algebra $\mathcal{B}$ is the intersection of all σ-algebra in which contain $\mathcal{G}$, so we have $\mathcal{B} \subseteq \mathcal{M}$.

1.28 Remark: It can be shown, using transfinite induction, that in $\mathbf{R}$ we have $|\mathcal{B}| = 2^{\aleph_0}$. Since $|\mathcal{B}| < |\mathcal{M}|$, it follows that there exist measurable functions which are not Borel.

1.29 Theorem: *For every set* $A \subseteq \mathbf{R}$ *there exists a set* $B \in \mathcal{G}_\delta$ *with* $A \subseteq B$ *such that* $\lambda(B) = \lambda^*(A)$.

Proof: Let $A \subseteq \mathbf{R}$. For each $n \in \mathbf{Z}^+$, choose bounded open intervals $I_{n,1}, I_{n,2}, I_{n,3}, \cdots$ such that $A \subseteq \bigcup_{k=1}^{\infty} I_{n,k}$ and $\sum_{k=1}^{\infty} |I_{n,k}| \leq \lambda^*(A) + \frac{1}{n}$, then let $U_n = \bigcup_{k=1}^{\infty} I_{n,k}$. Note that for each $n \in \mathbf{Z}^+$ the set U_n is open with $A \subseteq U_n$, and we have $\lambda(U_n) \leq \sum_{k=1}^{\infty} |I_{n,k}| \leq \lambda^*(A) + \frac{1}{n}$. Let $B = \bigcap_{n=1}^{\infty} U_n$ and note that $B \in \mathcal{G}_\delta$. Since $A \subseteq U_n$ for all $n \in \mathbf{Z}^+$, we have $A \subseteq \bigcap_{n=1}^{\infty} U_n$, that is $A \subseteq B$, and hence $\lambda^*(A) \leq \lambda(B)$. For every $n \in \mathbf{Z}^+$ we have $B \subseteq U_n$ so that $\lambda(B) \leq \lambda(U_n) \leq \lambda^*(A) + \frac{1}{n}$, and it follows that $\lambda(B) \leq \lambda^*(A)$. Thus $\lambda(B) = \lambda^*(A)$, as required.

1.30 Theorem: *Let* $A \subseteq \mathbf{R}$. *Then the following statements are equivalent.*

(1) A *is measurable.*
(2) For every $\epsilon > 0$ *there exists an open set* U *with* $A \subseteq U \subseteq \mathbf{R}$ *such that* $\lambda(U \setminus A) < \epsilon$.
(3) There exists a set $B \in \mathcal{G}_\delta$ *with* $A \subseteq B \subseteq \mathbf{R}$ *such that* $\lambda(B \setminus A) = 0$.
(4) For every $\epsilon > 0$ *there exists a closed set* $K \subseteq A$ *such that* $\lambda(A \setminus K) < \epsilon$.
(5) There exists a set $C \in \mathcal{F}_\sigma$ *with* $C \subseteq A$ *such that* $\lambda(A \setminus C) = 0$.

Proof: We prove that (1) is equivalent to (3) and leave proofs of other equivalences as an exercise. To show that (3) implies (1), suppose that there exists a set $B \in \mathcal{G}_\delta$ with $A \subseteq B$ such that $\lambda^*(B \setminus A) = 0$. Since $\lambda^*(B \setminus A) = 0$ we know that $B \setminus A$ is measurable, and hence the set $A = B \setminus (B \setminus A)$ is also measurable.

Suppose, conversely, that A is measurable. If $\lambda(A) < \infty$ then, by Theorem 1.28, we can choose $B \in \mathcal{G}_\delta$ with $A \subseteq B$ such that $\lambda(B) = \lambda(A)$, and then we have $\lambda(B \setminus A) = \lambda(B) - \lambda(A) = 0$, as required. If $\lambda(A) = \infty$ then more care is needed. For each $n \in \mathbf{Z}^+$, let $A_n = A \cap [-n, n]$, note that A_n is measurable, and choose $B_n \in \mathcal{G}_\delta$ with $A_n \subseteq B_n$ such that $\lambda(B_n) = \lambda(A_n)$, and note that $\lambda(B_n \setminus A_n) = \lambda(B_n) - \lambda(A_n) = 0$. Let $B = \bigcup\limits_{n=1}^{\infty} B_n$. Then we have $A = \bigcup\limits_{n=1}^{\infty} A_n \subseteq \bigcup\limits_{n=1}^{\infty} B_n = B$ and we have $B \setminus A = \bigcup\limits_{n=1}^{\infty} B_n \setminus A \subseteq \bigcup\limits_{n=1}^{\infty} B_n \setminus A_n$ so that $\lambda(B \setminus A) \leq \sum\limits_{n=1}^{\infty} \lambda(B_n \setminus A_n) = 0$.

1.31 Theorem: *Let $A, B \subseteq \mathbf{R}$. Suppose that $A \subseteq B$ and B is measurable with $\lambda(B) < \infty$. Then A is measurable if and only if $\lambda(B) = \lambda^*(A) + \lambda^*(B \setminus A)$.*

Proof: If A is measurable then for all $X \subseteq \mathbf{R}$ we have $\lambda^*(X) = \lambda^*(X \cap A) + \lambda^*(X \setminus A)$ so that in particular (taking $X = B$) we have $\lambda^*(B) = \lambda^*(A) + \lambda^*(B \setminus A)$.

Suppose that $\lambda^*(B) = \lambda^*(A) + \lambda^*(B \setminus A)$, and let $X \subseteq \mathbf{R}$. By Theorem 1.28, we can choose $E \in \mathcal{G}_\delta$ with $X \cap B \subseteq E$ such that $\lambda(E) = \lambda^*(X \cap B)$. Let $C = E \cap B$ and note that C is measurable with $X \cap B \subseteq C \subseteq B$. Since $X \cap B \subseteq C$ we have $\lambda^*(X \cap B) \leq \lambda(C)$ and since $C \subseteq E$ we have $\lambda(C) \leq \lambda(E) = \lambda^*(X \cap B)$, and so $\lambda(C) = \lambda^*(X \cap B)$. Since C is measurable, and since $(B \setminus A) \cap C = C \setminus A$ and $(A \setminus C) \cup ((B \setminus A) \setminus C) = B \setminus C$ and $(C \cap A) \cup (C \setminus A) = C$ and $C = B \cap C$, we have

$$\begin{aligned}
\lambda^*(B) &= \lambda^*(A) + \lambda^*(B \setminus A) \\
&= \big(\lambda^*(A \cap C) + \lambda^*(A \setminus C)\big) + \big(\lambda^*((B \setminus A) \cap C) + \lambda^*((B \setminus A) \setminus C)\big) \\
&= \lambda^*(A \cap C) + \lambda^*(A \setminus C) + \lambda^*(C \setminus A) + \lambda^*((B \setminus A) \setminus C) \\
&= \big(\lambda^*(C \cap A) + \lambda^*(C \setminus A)\big) + \big(\lambda^*(A \setminus C) + \lambda^*((B \setminus A) \setminus C)\big) \\
&\geq \big(\lambda^*(C \cap A) + \lambda^*(C \setminus A)\big) + \lambda^*(B \setminus C) \\
&\geq \lambda^*(C) + \lambda^*(B \setminus C) = \lambda^*(B \cap C) + \lambda^*(B \setminus C) \\
&= \lambda^*(B)
\end{aligned}$$

Since the first and last terms above are equal, it follows that all terms must be equal, so in particular we have $\lambda^*(C \cap A) + \lambda^*(C \setminus A) + \lambda^*(B \setminus C) = \lambda^*(C) + \lambda^*(B \setminus C)$ hence $\lambda^*(C) = \lambda^*(C \cap A) + \lambda^*(C \setminus A)$. Thus

$$\begin{aligned}
\lambda^*(X \cap B) = \lambda^*(C) &= \lambda^*(C \cap A) + \lambda^*(C \setminus A) \\
&\geq \lambda^*((X \cap B) \cap A) + \lambda^*((X \cap B) \setminus A) \text{ , since } X \cap B \subseteq C, \\
&= \lambda^*(X \cap A) + \lambda^*((X \cap B) \setminus A) \text{ , since } (X \cap B) \cap A = X \cap A.
\end{aligned}$$

hence

$$\begin{aligned}
\lambda^*(X) &= \lambda^*(X \cap B) + \lambda^*(X \setminus B) \text{ , since } B \text{ is measurable,} \\
&= \lambda^*(X \cap A) + \lambda^*((X \cap B) \setminus A) + \lambda^*(X \setminus B) \\
&\geq \lambda^*(X \cap A) + \lambda^*(X \setminus A) \text{ , since } ((X \cap B) \setminus A) \cup (X \setminus B) = X \setminus A.
\end{aligned}$$

Thus A is measurable, as required.

1.32 Definition: Let X be a metric space and let $A \subseteq X$. We say A is **dense** (in X) when for every nonempty open ball $B \subseteq X$ we have $B \cap A \neq \emptyset$, or equivalently when $\overline{A} = X$. We say A is **nowhere dense** (in X) when for every nonempty open ball $B \subseteq \mathbf{R}$ there exists a nonempty open ball $C \subseteq B$ with $C \cap A = \emptyset$, or equivalently when $\overline{A}^0 = \emptyset$.

1.33 Example: The generalized Cantor sets are nowhere dense in $\mathbf{R}$.

1.34 Note: When $A \subseteq B \subseteq X$, note that if A is dense in X then so is B and, on the other hand, if B is nowhere dense in X then so is A.

1.35 Note: When $A, B \subseteq X$ with $B = A^c = X \setminus A$, note that A is nowhere dense $\iff \overline{A}^0 = \emptyset \iff \overline{B^0} = X \iff$ the interior of B is dense.

1.36 Definition: Let $A \subseteq X$. We say that A is **first category** (or that A is **meagre**) when A is equal to a countable union of nowhere dense sets. We say that A is **second category** when it is not first category. We say that A **residual** when A^c is first category.

1.37 Example: Every countable set in $\mathbf{R}$ is first category since if $A = \{a_1, a_2, a_3, \cdots\}$ then we have $A = \bigcup_{k=1}^{\infty} \{a_k\}$. In particular $\mathbf{Q}$ is first category and $\mathbf{Q}^c = \mathbf{R} \setminus \mathbf{Q}$ is residual.

1.38 Note: If $A \subseteq X$ is first category then so is every subset of A.

1.39 Note: If $A_1, A_2, A_3, \cdots \subseteq X$ are are all first category then so is $\bigcup_{k=1}^{\infty} A_k$.

1.40 Theorem: *(The Baire Category Theorem) Let X be a complete metric space.*
(1) Every first category set in X has an empty interior.
(2) Every residual set in X is dense.
(3) Every countable union of closed sets with empty interiors in X has an empty interior.
(4) Every countable intersection of dense open sets in X is dense.

Proof: Parts (1) and (2) are equivalent by taking complements, and Parts (3) and (4) are special cases of Parts (1) and (2), so it suffices to prove Part (1). We sketch a proof.

Let $A \subseteq X$ be first category, say $A = \bigcup_{n=1}^{\infty} C_n$ where each C_n is nowhere dense. Suppose, for a contradiction, that A has nonempty interior, and choose an open ball $B_0 = B(a_0, r_0)$ with $0 < r_0 < 1$ such that $\overline{B}_0 \subseteq A$. Since each C_n is nowhere dense, we can chose a nested sequence of open balls $B_n = B(a_n, r_n)$ with $0 < r_n < \frac{1}{2^n}$ such that $\overline{B}_n \subseteq B_{n-1}$ and $\overline{B}_n \cap C_n = \emptyset$. Because $r_n \to 0$, it folows that the sequence $\{a_n\}$ is Cauchy. Because X is complete, it follows that $\{a_n\}$ converges in X, say $a = \lim_{n \to \infty} a_n$. Note that $a \in \overline{B}_n$ for all n since $a_k \in \overline{B}_n$ for all $k \geq n$. Since $a \in \overline{B}_0$ and $\overline{B}_0 \subseteq A$ we have $a \in A$. But since $a \in \overline{B}_n$for al $n \geq 1$, and $\overline{B}_n \cap C_n = \emptyset$, we have $a \notin C_n$ for all $n \geq 1$ hence $a \notin \bigcup_{n=1}^{\infty} C_n$, that is $a \notin A$.

1.41 Example: Recall that $\mathbf{Q}$ is first category and $\mathbf{Q}^c$ is residual. The Baire Category Theorem shows that $\mathbf{Q}^c$ cannot be first category because if $\mathbf{Q}$ and $\mathbf{Q}^c$ were both first category then $\mathbf{R} = \mathbf{Q} \cup \mathbf{Q}^c$ would also be first category, but this is not possible since $\mathbf{R}$ does not have empty interior.

1.42 Exercise: For each $n \in \mathbf{Z}^+$, let $f_n : \mathbf{R} \to \mathbf{R}$ be continuous. Suppose that for all $x \in \mathbf{R}$ there exists $n \in \mathbf{Z}^+$ such that $f_n(x) \in \mathbf{Q}$. Prove that there exists $n \in \mathbf{Z}^+$ such that f_n is constant in some nondegenerate interval.

1.43 Exercise: Show that $\mathcal{F}_\sigma \neq \mathcal{G}_\delta$ and show that $G_\delta \neq \mathcal{G}_{\delta\sigma}$.

1.44 Remark: Each of the following sets $\mathcal{C}$ of subsets of $\mathbf{R}$

$$\mathcal{C} = \big\{A \subseteq \mathbf{R} \big| \, A \text{ is finite or countable}\big\}$$
$$\mathcal{C} = \big\{A \subseteq \mathbf{R} \big| \, \lambda(A) = 0\big\}$$
$$\mathcal{C} = \big\{A \subseteq \mathbf{R} \big| \, A \text{ is first category}\big\}$$

has the following properties:

(1) if $A \subseteq B$ and $B \in \mathcal{C}$ then $A \in \mathcal{C}$,

(2) if $A_1, A_2, A_3, \dots \in \mathcal{C}$ then $\bigcup\limits_{k=1}^{\infty} A_k \in \mathcal{C}$, and

(3) if $A \in \mathcal{C}$ then $A^0 = \emptyset$.

Because of this, it seems reasonable to consider the sets in $\mathcal{C}$ to be, in some sense, "small". The following theorem, then, states that every set in $\mathbf{R}$ is the union of two small sets.

1.45 Theorem: *Every subset of* $\mathbf{R}$ *is equal to the disjoint union of a set of measure zero and a set of first category.*

Proof: Let $\mathbf{Q} = \{a_1, a_2, a_3, \cdots\}$. For $k, \ell \in \mathbf{Z}^+$, let $I_{k,\ell} = \left(a_\ell - \frac{1}{2^{k+\ell}}\, , \, a_\ell + \frac{1}{2^{k+\ell}}\right)$ and for $k \in \mathbf{Z}^+$, let $U_k = \bigcup\limits_{\ell=1}^{\infty} I_{k,\ell}$. Note that $U_1 \supseteq U_2 \supseteq U_3 \supseteq \cdots$ and for each $k \in \mathbf{Z}^+$ we have $\mathbf{Q} \subseteq U_k$ and $\lambda(U_k) \leq \sum\limits_{\ell=1}^{\infty} |I_{k,\ell}| = \frac{1}{2^{k-1}}$ and we have $U_1 \supset U_2 \supseteq U_3 \supseteq \cdots$. Let $B = \bigcap\limits_{k=1}^{\infty} U_k$. Note that B is residual (it is a countable intersection of dense open sets) and we have $\lambda(B) = \lim\limits_{k \to \infty} \lambda(U_k) = 0$ since $\lambda(U_k) \leq \frac{1}{2^k}$ for all $k \in \mathbf{Z}^+$. Finally note that any set A is equal to the disjoint union $A = (A \cap B) \cup (A \cap B^c)$, and we have $\lambda(A \cap B) = 0$ and the set $A \cap B^c$ is first category.

.

Chapter 2. Lebesque Integration

2.1 Definition: For $E \subseteq A \subseteq \mathbf{R}$, the **characteristic function** for E on A is the function $\chi_E : A \to \{0,1\}$ given by
$$\chi_E(x) = \begin{cases} 1 \text{ , if } x \in E, \\ 0 \text{ , if } x \notin E. \end{cases}$$

2.2 Definition: Let $a, b \in \mathbf{R}$ with $a < b$. A **step function** on $[a,b]$ is a function $s : [a,b] \to \mathbf{R}$ of the form
$$s = \sum_{k=1}^{n} c_k \chi_{I_k}$$
where $n \in \mathbf{Z}^+$, each $c_k \in \mathbf{R}$, and the sets I_k are disjoint intervals with $\bigcup_{k=1}^{n} I_k = A$. The numbers c_k and the intervals I_k are uniquely determined from s if we require that I_{k-1} is to the left of I_k and $c_{k-1} \neq c_k$ for $1 < k \leq n$, and then we have $I_k = s^{-1}(c_k)$.

2.3 Definition: For the step function on $[a,b]$ given by $s = \sum_{k=1}^{n} c_k \chi_{I_k}$, we define the **Riemann integral** of s on $[a,b]$ to be
$$\int_a^b s = \int_a^b s(x)\,dx = \sum_{k=1}^{n} c_k |I_k|.$$
For a bounded function $f : [a,b] \to \mathbf{R}$ we define the **upper Riemann integral** and the **lower Riemann integral** of f on $[a,b]$ to be
$$U(f) = \inf\left\{ \int_a^b s \;\middle|\; s \text{ is a step function on } [a,b] \text{ with } s \geq f \right\},$$
$$L(f) = \sup\left\{ \int_a^b s \;\middle|\; s \text{ is a step function on } [a,b] \text{ with } s \leq f \right\}.$$
We say that f is **Riemann integrable** on $[a,b]$ when $U(f) = L(f)$, and in this case we define the **Riemann integral** of f on $[a,b]$ to be
$$\int_a^b f = \int_a^b f(x)\,dx = U(f) = L(f).$$

2.4 Theorem: *(Properties of the Riemann Integral) Let $a < b$ and let $f, g : [a,b] \to \mathbf{R}$ be bounded.*

(1) If f and g are Riemann integrable on $[a,b]$ and $f \leq g$ then $\int_a^b f \leq \int_a^b g$.

(2) If f and g are Riemann integrable on $[a,b]$ and $c \in \mathbf{R}$ then the functions cf and $f+g$ are Riemann integrable on $[a,b]$ and $\int_a^b (cf) = c\int_a^b f$ and $\int_a^b (f+g) = \int_a^b f + \int_a^b g$.

(3) If $c \in (a,b)$ then f is Riemann integrable on $[a,b]$ if and only if f is Riemann integrable both on $[a,c]$ and on $[c,b]$ and, in this case, $\int_a^b f = \int_a^c f + \int_c^b f$.

(4) If $f(x) = g(x)$ for all but finitely many $x \in [a,b]$ then f is Riemann integrable on $[a,b]$ if and only if g is Riemann integrable on $[a,b]$ and, in this case, $\int_a^b f = \int_a^b g$.

Proof: The proof is left as an exercise.

2.5 Theorem: *Let $a < b$ and let $f : [a,b] \to \mathbf{R}$ be bounded.*
(1) If f is continuous then f is Riemann integrable.
(2) If f is monotonic then f is Riemann integrable.

Proof: We omit the proof.

2.6 Theorem: *(Lebesgue) Let $a < b$ and let $f : [a,b] \to \mathbf{R}$. Then f is Riemann integrable on $[a,b]$ if and only if f is bounded and the set of all points in $[a,b]$ at which f is discontinuous has measure zero.*

Proof: We omit the proof.

2.7 Theorem: *(The Fundamental Theorem of Calculus) Let $f, g : [a,b] \to \mathbf{R}$. Suppose that g is differentiable with $g' = f$ in $[a,b]$ and that f is Riemann integrable on $[a,b]$. Then*

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

Proof: We omit the proof.

2.8 Example: The function $f : [0,1] \to [0,1]$ defined by $f(x) = 1$ when $x \in \mathbf{Q}$ and $f(x) = 0$ when $x \notin \mathbf{Q}$ is discontinuous everywhere in $[0,1]$, and is not Riemann integrable.

2.9 Example: The function $f : [0,1] \to [0,1]$ given by $f\left(\frac{a}{b}\right) = \frac{1}{b}$ when $a, b \in \mathbf{Z}$ with $0 \leq a \leq b$ and $\gcd(a,b) = 1$, and $f(x) = 0$ when $x \notin \mathbf{Q}$, is discontinuous at all rational points, and is Riemann integrable.

2.10 Example: Define $s : \mathbf{R} \to [0,1]$ by $s(x) = 0$ for $x \leq 0$ and $s(x) = 1$ for $x > 0$. Let $\mathbf{Q} \cap [0,1] = \{a_1, a_2, a_3, \cdots\}$ and define $f : [0,1] \to [0,1]$ by $f(x) = \sum_{k=1}^{\infty} \frac{s(x-a_k)}{2^k}$. Then f is increasing with jump discontinuities at all rational points, and f is Riemann integrable.

2.11 Example: Given a Cantor set $C = [0,1] \setminus U$, where $U = \bigcup_{k=1}^{\infty} I_k$ with the sets I_k being the disjoint open intervals from Example 1.17, we can construct a corresponding **Cantor function** $f : [0,1] \to [0,1]$ with $f(x) = \frac{1}{2}$ on I_1, $f(x) = \frac{1}{4}$ on I_2, $f(x) = \frac{3}{4}$ on I_3, $f(x) = \frac{1}{8}$ on I_4, $f(x) = \frac{3}{8}$ on I_5, $f(x) = \frac{5}{8}$ on I_6, $f(x) = \frac{7}{8}$ on I_7 and so on, and then extending f to make it continuous on all of $[0,1]$. Then f is continuous and nondecreasing with $f'(x) = 0$ for all $x \in U$.

2.12 Example: When $C = [0,1] \setminus U$ is a Cantor set and $f : [0,1] \to [0,1]$ is the corresponding Cantor function (as in the previous example), the function $g : [0,1] \to [0,2]$ given by $g(x) = x + f(x)$ is a homeomorphism. Note that g sends each component interval of U to an interval of the same size, so that we have $\lambda\big(g(U)\big) = \lambda(U)$.

In the case that C is the standard Cantor set we have $\lambda\big(g(U)\big) = \lambda(U) = 1$. It follows that $\lambda\big(g(C)\big) = 2 - \lambda(U) = 1$, so g sends a set of measure zero to a set of measure 1. Also note that if we choose a nonmeasurable set $B \subseteq g(C)$ and let $A = g^{-1}(B)$, then $A \subseteq C$ so that A is a measurable set with measure zero, but g sends A to the nonmeasurable set $g(A) = B$.

2.13 Example: Given a Cantor set $C = [0,1]\setminus U$ where U is the disjoint union $U = \bigcup_{k=1}^{\infty} I_k$, choose intervals $J_k \subsetneqq I_k$ so that J_k has the same centre as I_k with $|J_k| = \frac{1}{2}|I_k|$, then choose continuous functions $f_k : [0,1] \to [0,1]$ such that $f(x) = 0$ outside J_k and $f(x) = 1$ at the midpoint of J_k and then let $f(x) = \sum_{k=1}^{\infty} f_k(x)$ for all $x \in [0,1]$. Then f is continuous in U and discontinuous in C. When $\lambda(C) > 0$, f is not Riemann integrable. If we define $g(x) = \sum_{k=1}^{\infty} \int_0^x f_k(t)\,dt$ then g is differentiable with $g' = f$ in $[a,b]$.

2.14 Example: Let $\mathbf{Q}\cap[0,1] = \{a_1, a_2, \cdots\}$. Define $f : [0,1] \to \mathbf{R}$ by $f(x) = \sum_{k=1}^{\infty} \frac{(x-a_k)^{1/3}}{2^k}$. Then f is increasing with $f'(x) = \sum_{k=1}^{\infty} \frac{(x-a_k)^{-2/3}}{3\cdot 2^k}$ when $x \notin \mathbf{Q}$ and $f'(x) = \infty$ when $x \in \mathbf{Q}$. Verify that $f'(x) \geq \frac{1}{3}$ for all x. The map f sends the interval $[0,1]$ homeomorphically to an interval $[a,b]$ and the inverse map $g : [a,b] \to [0,1]$ is increasing and differentiabe with $g'(x) = 0$ for all $x \in \mathbf{Q}$ and $g'(x) \leq 3$ for all x. Note that g' cannot be Riemann integrable because if it was then we would have $\int_a^b g' = g(b) - g(a) = 1$ but, because $g'(x) = 0$ for all $x \in \mathbf{Q}$, all of the lower Riemann sums are zero.

2.15 Definition: We shall find it useful on occasion to allow our functions to take the values $\pm\infty$ so we shall use the set of **extended real numbers** $[-\infty,\infty] = \mathbf{R}\cup\{-\infty,\infty\}$. In $[-\infty,\infty]$, the open balls are the open intervals $B(-\infty,r) = \big(-\infty,\frac{1}{r}\big)$, $B(\infty,r) = \big(\frac{1}{r},\infty\big)$ and $B(a,r) = (a-r,a+r)$ with $a \in \mathbf{R}$. For $A \subseteq [-\infty,\infty]$, we say that A is **open** in $[-\infty,\infty]$ when for every $a \in A$ there exists $r > 0$ such that $B(a,r) \subseteq A$. Verify that every open set in $[-\infty,\infty]$ is a finite or countable union of disjoint open intervals, where each open inerval is of one of the forms $\emptyset$, (a,b), $(-\infty,a)$, (a,∞), $(-\infty,\infty)$, $[-\infty,a)$, $(a,\infty]$ or $[-\infty,\infty]$ where $a,b \in \mathbf{R}$. We also use (partially-defined) addition and multiplication operations on $[-\infty,\infty]$, as usual, leaving certain sums and products undefined. We do not define the expressions $\infty + (-\infty)$, $-\infty + \infty$, $0\cdot(\pm\infty)$ and $(\pm\infty)\cdot 0$.

2.16 Definition: For $f : A \subseteq \mathbf{R} \to B \subseteq [-\infty,\infty]$, we say that f is **measurable** (in A) when $f^{-1}(U)$ is measurable for every open set U in $[-\infty,\infty]$ (or equivalently for every open set U in B). Note that in particular, in order for f to be measurable, the set A must be measurable because $A = f^{-1}\big([-\infty,\infty]\big)$.

2.17 Note: If $f : A \subseteq \mathbf{R} \to B \subseteq [-\infty,\infty]$ is measurable and $\varphi : B \subseteq [-\infty,\infty] \to [-\infty,\infty]$ is continuous, then the composite $\varphi\circ f : A \subseteq \mathbf{R} \to [-\infty,\infty]$ is measurable because, for every open set U in $[-\infty,\infty]$, $\varphi^{-1}(U)$ is open in B since φ is continuous, and hence the set $(\varphi\circ f)^{-1}(U) = f^{-1}\big(\varphi^{-1}(U)\big)$ is measurable since the function f is measurable.

2.18 Theorem: *Let $A \subseteq \mathbf{R}$ be measurable and let $f : A \to [-\infty,\infty]$, Then*

$$\begin{aligned} f \text{ is measurable} &\iff f^{-1}\big((a,\infty]\big) \text{ is measurable for all } a \in \mathbf{R} \\ &\iff f^{-1}\big([a,\infty]\big) \text{ is measurable for all } a \in \mathbf{R} \\ &\iff f^{-1}\big([-\infty,a)\big) \text{ is measurable for all } a \in \mathbf{R} \\ &\iff f^{-1}\big([-\infty,a]\big) \text{ is measurable for all } a \in \mathbf{R} \end{aligned}$$

Proof: We shall prove the first equivalence. If f is measurable then $f^{-1}(U)$ is measurable for every open set $U \subseteq [-\infty, \infty]$ so, in particular, $f^{-1}\big((a, \infty]\big)$ is measurable for every $a \in \mathbf{R}$. Suppose, conversely, that $f^{-1}\big((a, \infty]\big)$ is measurable for every $a \in \mathbf{R}$. Then for every $a, b \in \mathbf{R}$, each of the following sets is measurable.

$$\begin{aligned} f^{-1}\big([-\infty, a]\big) &= \mathbf{R} \setminus f^{-1}\big((a, \infty]\big), \\ f^{-1}\big([-\infty, a)\big) &= \bigcup_{n=1}^{\infty} f^{-1}\big([-\infty, a - \tfrac{1}{n}]\big), \\ f^{-1}(a, b) &= f^{-1}\big([-\infty, b)\big) \cap f^{-1}\big((a, \infty]\big). \end{aligned}$$

Since every open U set in $[-\infty, \infty]$ is a finite or countable union of sets U_k, each of which is of one of the forms $[-\infty, a)$, (a, b), $(a, \infty]$, and because $f^{-1}\big(\bigcup_{k=1}^{\infty} U_k\big) = \bigcup_{k=1}^{\infty} f^{-1}(U_k)$, it follows that $f^{-1}(U)$ is measurable for every open set U in $[-\infty, \infty]$.

2.19 Theorem: *Let $E \subseteq A \subseteq \mathbf{R}$ with A measurable, and let $f : A \to [-\infty, \infty]$.*
(1) The function χ_E is measurable if and only if the set E is measurable.
(2) If f is continuous then f is measurable.
(3) If f is monotonic then f is measurable.

Proof: To prove Part (1), note that if E is not measurable then neither is χ_E because ${\chi_E}^{-1}\big((0, \infty]\big) = E$, and if E is measurable then so is χ_E because for all sets U in $[-\infty, \infty]$, the set $f^{-1}(U)$ is equal to one of the measurable sets $\emptyset$, E, $A \setminus E$ or A.

To prove Part (2), suppose that f is continuous and let U be any open set in $[-\infty, \infty]$. Since f is continuous and U is open, the set $f^{-1}(U)$ is open in A. Since $f^{-1}(U)$ is open in A, we can choose an open set V in $\mathbf{R}$ such that $f^{-1}(U) = V \cap A$, which is measurable.

To prove Part (c), suppose that f is monotonic, say f is increasing. Let $a \in \mathbf{R}$. For all $x, y \in A$, if $x \in f^{-1}\big((a, \infty]\big)$ and $y \geq x$ then $f(y) \geq f(x) > a$ so that $y \in f^{-1}\big((a, \infty]\big)$. It follows that the set $f^{-1}\big((a, \infty]\big)$ must a set of one of the forms $\emptyset$, $A \cap (b, \infty]$, $A \cap [b, \infty]$ or A, and so $f^{-1}\big((a, \infty]\big)$ is measurable.

2.20 Definition: Given a function $f : A \subseteq \mathbf{R} \to [-\infty, \infty]$, we define $f^+ : A \to [-\infty, \infty]$ and $f^- : A \to [-\infty, \infty]$ by

$$f^+(x) = \begin{cases} f(x) & \text{, if } f(x) \geq 0, \\ 0 & \text{, if } f(x) \leq 0, \end{cases} \qquad f^-(x) = \begin{cases} 0 & \text{, if } f(x) \geq 0, \\ -f(x) & \text{, if } f(x) \leq 0. \end{cases}$$

2.21 Theorem: *(Operations on Measurable Functions) Let $f, g : A \subseteq \mathbf{R} \to [-\infty, \infty]$ be measurable functions, and let $c \in \mathbf{R}$. Then each of the following fuctions are measurable*

$$cf \,,\; f + g \,,\; fg \,,\; |f| \,,\; f^+ \,,\; f^-$$

provided they are well-defined.

Proof: The function cf is well-defined unless $c = 0$ and $f(x) = \pm\infty$ for some $x \in A$. When $c = 0$ and $f(x) \neq \pm\infty$, the function cf is the zero function, which is measurable. When $c \neq 0$ the function $\varphi : [-\infty, \infty] \to [-\infty, \infty]$ given by $\varphi(x) = cx$ is continuous and so the the function $cf = \varphi \circ f$ is measurable by Note 2.17.

The function $f+g$ is measurable because for all $a\in\mathbf{R}$ we have

$$\begin{aligned}(f+g)^{-1}\big((a,\infty]\big)&=\big\{x\in A\big|f(x)+g(x)>a\big\}\\&=\bigcup_{r\in\mathbf{Q}}\big\{f(x)>r\text{ and }g(x)>a-r\big\}\\&=\bigcup_{r\in\mathbf{Q}}\big(f^{-1}\big(r,\infty]\big)\cap g^{-1}\big(a-r,\infty]\big),\end{aligned}$$

which is measurable.

The function $\varphi:[-\infty,\infty]\to[0,\infty]$ given by $\varphi(x)=x^2$ is continuous so, by Note 2.17, for every measurable function $h:A\to[-\infty,\infty]$, the function $h^2=\varphi\circ h$ is also measurable. It follows that the function $fg=\frac{1}{4}\big((f+g)^2-(f-g)^2\big)$ is measurable.

The function $\varphi:[-\infty,\infty]\to[0,\infty]$ given by $\varphi(x)=|x|$ is continuous so, by Note 2.17, the function $|f|=\varphi\circ f$ is measurable, hence so are the functions $f^+=\frac{1}{2}\big(|f|+f\big)$ and $f^-=\frac{1}{2}\big(|f|-f\big)$.

2.22 Theorem: *Let $f_n:A\subseteq\mathbf{R}\to[-\infty,\infty]$ be measurable for each $n\in\mathbf{Z}^+$. Then each of the following functions are well-defined and measurable:*

$$\sup\{f_n\big|n\in\mathbf{Z}^+\}\,,\ \inf\{f_n\big|n\in\mathbf{Z}^+\}\,,\ \limsup_{n\to\infty}\{f_n\}\,,\ \liminf_{n\to\infty}\{f_n\}.$$

Proof: Let $g=\sup\{f_n|n\in\mathbf{Z}^+\}$. For $x\in A$ and $a\in\mathbf{R}$ we have

$$\begin{aligned}x\in g^{-1}\big((a,\infty]\big)&\iff g(x)>a\iff\sup\{f_n|n\in\mathbf{Z}^+\}>a\\&\iff f_n(x)>a\text{ for some }n\in\mathbf{Z}^+\iff x\in\bigcup_{n=1}^{\infty}{f_n}^{-1}\big((a,\infty]\big).\end{aligned}$$

Thus for all $a\in\mathbf{R}$ we have $g^{-1}\big((a,\infty]\big)=\bigcup\limits_{n=1}^{\infty}{f_n}^{-1}\big((a,\infty]\big)$, which is measurable. Similarly, when $h=\inf\{f_n\big|n\in\mathbf{Z}^+\}$ and $a\in\mathbf{R}$ we have $h^{-1}\big([a,\infty]\big)=\bigcap\limits_{n=1}^{\infty}{f_n}^{-1}\big([a,\infty]\big)$, which is measurable. Also, we have

$$\limsup_{n\to\infty}f_n=\inf\Big\{\sup\{f_n|n\ge1\},\sup\{f_n|n\ge2\},\sup\{f_n|n\ge3\},\cdots\Big\}\text{ and}$$
$$\liminf_{n\to\infty}f_n=\sup\Big\{\inf\{f_n|n\ge1\},\inf\{f_n|n\ge2\},\inf\{f_n|n\ge3\},\cdots\Big\}.$$

It follows that $\limsup\limits_{n\to\infty}f_n$ and $\liminf\limits_{n\to\infty}f_n$ are measurable.

2.23 Definition: Let $A\subseteq\mathbf{R}$ be measurable. We say that a property or statement holds for **almost every** (written a.e.) $x\in A$, or **almost everywhere** (written a.e.) in A, when the property or statement holds for every $x\in A\setminus E$ for some set $E\subseteq A$ with $\lambda(E)=0$. For example, for functions $f,g:A\to[-\infty,\infty]$, we say that $f(x)=g(x)$ for a.e. $x\in A$ (or $f=g$ a.e. in A) when $f(x)=g(x)$ for every $x\in A\setminus E$ for some set $E\subseteq A$ with $\lambda(E)=0$.

2.24 Theorem: *Let $A\subseteq\mathbf{R}$ be measurable and let $f,g:A\to[-\infty,\infty]$.*
(1) If $\lambda(A)=0$ then f is measurable.
(2) If $A=B\cup C$ where B and C are disjoint and measurable then f is measurable (in A) if and only if the restrictions of f to B and to C are both measurable (in B and in C).
(3) If $f=g$ a.e. in A then f is measurable if and only if g is measurable.

Proof: The proof is left as an exercise.

2.25 Definition: Let $A \subseteq \mathbf{R}$. A **simple function** on A is a function $s : A \to \mathbf{R}$ of the form

$$s = \sum_{k=1}^{n} c_k \chi_{A_k}$$

where $n \in \mathbf{Z}^+$, each $c_k \in \mathbf{R}$, and the sets A_k are disjoint measurable sets with $\bigcup_{k=1}^{n} A_k = A$. The numbers c_k and sets A_k are uniquely determined from the function s if we require that $c_1 < c_2 < \cdots < c_n$, and then we have $A_k = s^{-1}(c_k)$.

2.26 Definition: For the nonnegative simple function $s : A \subseteq \mathbf{R} \to [0, \infty)$ given by $s = \sum_{k=1}^{n} c_k \chi_{A_k}$, the (Lebesgue) **integral** of s on A is defined to be

$$\int_A f(x)\,dx = \int_A s = \int_A s\; d\lambda = \sum_{k=1}^{n} c_k \lambda(A_k).$$

Note that the value of the integral does not depend on whether or not the numbers c_k are distinct because if $c_k = c_l$ then $c_k\lambda(A_k) + c_l\lambda(A_l) = c_k\big(\lambda(A_k) + \lambda(A_l)\big) = c_k\lambda(A_k \cup A_l)$.

2.27 Theorem: *(Properties of Integration) Let $r, s : A \subseteq \mathbf{R} \to [0, \infty)$ be nonnegative simple functions, and let $c \in \mathbf{R}$.*

(1) If $r \leq s$ then $\int_A r \leq \int_A s$.

(2) We have $\int_A (cs) = c\int_A s$ and $\int_A (r + s) = \int_A r + \int_A s$.

(3) If $A = B \cup C$, where B and C are disjoint and measurable, then $\int_A s = \int_B s + \int_C s$.

(4) If $B \subseteq A$ is measurable then $\int_B s = \int_A s \cdot \chi_B$.

(5) If $\lambda(A) = 0$ then $\int_A s = 0$.

(6) If $r = s$ a.e. in A then $\int_A r = \int_A s$, and if $\int_A r = 0$ then $r = 0$ a.e. in A.

Proof: We shall prove Parts (1) and (2) and leave the proofs of the remaining parts as an exercise. Let $r = \sum_{k=1}^{n} a_k \chi_{A_k}$ and $s = \sum_{l=1}^{m} b_l \chi_{B_l}$ and let $C_{k,l} = A_k \cap B_l$. Note that the sets $C_{k,l}$ are disjoint with $\bigcup_{k=1}^{n} C_{k,l} = \bigcup_{k=1}^{n} (A_k \cap B_l) = \big(\bigcup_{k=1}^{n} A_k\big) \cap B_l = A \cap B_l = B_l$ and it follows that $\sum_{k=1}^{n} \chi_{C_{k,l}} = \chi_{B_l}$ and that $\sum_{k=1}^{n} \lambda(C_{k,l}) = \lambda(B_l)$. Similary, we have $\bigcup_{l=1}^{m} C_{k,l} = A_k$, $\sum_{l=1}^{n} \chi_{C_{k,l}} = \chi_{A_k}$ and $\sum_{l=1}^{m} \lambda(C_{k,l}) = \lambda(A_k)$.

To prove Part (1), suppose that $r \leq s$. For all pairs (k, l) with $C_{k,l} \neq \emptyset$, we can choose $x \in C_{k,l}$ and then we have $a_k = r(x) \leq s(s) = b_l$. It follows that

$$\begin{aligned}\int_A r &= \sum_{k=1}^{n} a_k\lambda(A_k) = \sum_{k=1}^{n} a_k \sum_{l=1}^{m} \lambda(C_{k,l}) = \sum_{k,l} a_k\lambda(C_{k,l}) = \sum_{k,l\,\ni\, C_{k,l}\neq\emptyset} a_k\lambda(C_{k,l}) \\ &\leq \sum_{k,l\,\ni\, C_{k,l}\neq\emptyset} b_l\lambda(C_{k,l}) = \sum_{k,l} b_l\lambda(C_{k,l}) = \sum_{l=1}^{m} b_l \sum_{k=1}^{n} \lambda(C_{k,l}) = \sum_{l=1}^{m} b_l\lambda(B_l) = \int_A s\,.\end{aligned}$$

The first formula in Part (2) is clear. Let us prove the second formula. We have

$$r+s=\sum_{k=1}^{n}a_k\chi_{A_k}+\sum_{l=1}^{m}b_l\chi_{B_l}=\sum_{k=1}^{n}a_k\sum_{l=1}^{m}\chi_{C_{k,l}}+\sum_{l=1}^{m}b_l\sum_{k=1}^{n}\chi_{C_{k,l}}=\sum_{k,l}(a_k+b_l)\chi_{C_{k,l}}$$

and so

$$\begin{aligned}\int_A(r+s)&=\sum_{k,l}(a_k+b_l)\lambda(C_{k,l})=\sum_{k,l}a_k\lambda(C_{k,l})+\sum_{k,l}b_l\lambda(C_{k,l})\\&=\sum_{k=1}^{n}a_k\sum_{l=1}^{m}\lambda(C_{k,l})+\sum_{l=1}^{m}b_l\sum_{k=1}^{n}\lambda(C_{k,l})\\&=\sum_{k=1}^{n}a_k\lambda(A_k)+\sum_{l=1}^{m}b_l\lambda(B_l)=\int_A r+\int_A s\,.\end{aligned}$$

2.28 Note: Given any nonnegative measurable function $f:A\subseteq\mathbf{R}\to[0,\infty]$, we can construct an increasing sequence $\{s_n\}$ of nonnegative simple functions $s_n:A\to[0,\infty)$ with $s_n\to f$ pointwise in A as follows. For $n\in\mathbf{Z}^+$, we let

$$s_n(x)=\begin{cases}\frac{k-1}{2^n}\ ,\text{ if }\frac{k-1}{2^n}\le f(x)<\frac{k}{2^n}\text{ with }k\in\{1,2,\cdots,n2^n\},\\ \;n\ \ ,\text{ if }f(x)\ge n,\end{cases}$$

that is $s_n=\sum\limits_{k=1}^{n2^n}\frac{k-1}{2^n}\chi_{A_k}$ where $A_k=f^{-1}\left[\frac{k-1}{2^n},\frac{k}{2^n}\right)$ for $1\le k<n2^n$ and $A_{n2^n}=f^{-1}[n,\infty]$. We remark that if f is bounded the $s_n\to f$ uniformly in A.

2.29 Definition: For a nonnegative measurable function $f:A\subseteq\mathbf{R}\to[0,\infty]$, we define the (Lebesgue) **integral** of f on A to be

$$\int_A f(x)\,dx=\int_A f=\int_A f\;d\lambda=\sup\left\{\int_A s\,\middle|\;s\text{ is a simple function on }A\text{ with }0\le s\le f\right\}.$$

We say that $f:A\to[0,\infty]$ is (Lebesgue) **integrable** (on A) when $\displaystyle\int_A f<\infty$.

2.30 Theorem: *(Properties of Integration) Let $f,g:A\subseteq\mathbf{R}\to[0,\infty]$ be non-negative measurable functions and let $c\in\mathbf{R}$. Then*

(1) If $f\le g$ on A then $\displaystyle\int_A f\le\int_A g$.

(2) We have $\displaystyle\int_A(cf)=c\int_A f$ and $\displaystyle\int_A(f+g)=\int_A f+\int_A g$.

(3) If $A=B\cup C$, where B and C are disjoint and measurable, then $\displaystyle\int_A f=\int_B f+\int_C f$.

(4) If $B\subseteq A$ is measurable then $\displaystyle\int_B f=\int_A f\cdot\chi_B$.

(5) If $\lambda(A)=0$ then $\displaystyle\int_A f=0$.

(6) If $f=g$ a.e. in A then $\displaystyle\int_A f=\int_A g$, and if $\displaystyle\int_A f=0$ then $f=0$ a.e. in A.

Proof: All parts follow fairly easily from the analogous parts of Theorem 2.27 except for the second formula in Part (2). We shall return to the proof of this formula later.

2.31 Theorem: *(Fatou's Lemma) Let* $f_n : A \subseteq \mathbf{R} \to [0,\infty]$ *be nonnegative measurable functions for* $n \in \mathbf{Z}^+$. *Then*

$$\int_A \liminf_{n\to\infty} f_n \leq \liminf_{n\to\infty} \int_A f_n\,.$$

Proof: By the definition of the integral on the left, it suffices to prove that for every nonnegative simple function s on A with $s \leq \liminf_{n\to\infty} f_n$ we have $\int_A s \leq \liminf_{n\to\infty} \int_A f_n$. Let s be any nonnegative simple function on A with $s \leq \liminf_{n\to\infty} f_n$. Write $s = \sum_{k=1}^m a_k \chi_{A_k}$. For all $x \in A_k$ we have $a_k = s(x) \leq \liminf_{n\to\infty} f_n(x)$, and it follows that for all $0 \leq r < 1$ there exists $n \in \mathbf{Z}^+$ such that for all $l \geq n$ we have $f_l(x) \geq r a_k$. Let $0 \leq r < 1$. For $k, n \in \mathbf{Z}^+$, let

$$B_{k,n} = \left\{x \in A_k \middle| f_l(x) \geq r a_k \text{ for all } l \geq n\right\} = \bigcap_{l \geq n} {f_l}^{-1}[r a_k, \infty].$$

Note that each set $B_{k,n}$ is measurable with $B_{k,1} \subseteq B_{k,2} \subseteq B_{k,3} \subseteq \cdots$ and $\bigcup_{n=1}^{\infty} B_{k,n} = A_k$. It follows that $\lambda(A_k) = \lim_{n\to\infty} \lambda(B_{k,n})$. For all $x \in B_{k,n}$ we have $f_l(x) \geq r a_k$ for all $l \geq n$ so that, in particular, $f_n(x) \geq r a_k$. It follows that $f_n \geq \sum_{k=1}^m r a_k \chi_{B_{k,n}}$ hence

$$\int_A f_n \geq \sum_{k=1}^m r a_k \lambda(B_{k,n}).$$

Taking the lim inf on both sides gives

$$\liminf_{n\to\infty} \int_A f_n \geq \lim_{n\to\infty} \sum_{k=1}^m r a_k \lambda(B_{k,n}) = \sum_{k=1}^m r a_k \lambda(A_k) = r \int_A s\,.$$

Since $0 \leq r < 1$ was arbitrary, it follows that $\liminf_{n\to\infty} \int_A f_n \geq \int_A s$, as required.

2.32 Corollary: *Let* $f_n : A \subseteq \mathbf{R} \to [0,\infty]$ *be nonnegative measurable functions for* $n \in \mathbf{Z}^+$. *Suppose that the pointwise limit* $\lim_{n\to\infty} f_n(x)$ *exists with* $f_n(x) \leq \lim_{n\to\infty} f_n(x)$ *for all* $x \in A$. *Then*

$$\int_A \lim_{n\to\infty} f_n = \lim_{n\to\infty} \int_A f_n\,.$$

Proof: For all $n \in \mathbf{Z}^+$, since $f_n \leq \lim_{n\to\infty} f_n$ we have $\int_A f_n \leq \int_A \lim_{n\to\infty} f_n$. Taking the lim sup gives

$$\limsup_{n\to\infty} \int_A f_n \leq \int_A \lim_{n\to\infty} f_n\,.$$

By Fatou's Lemma, we also have

$$\int_A \lim_{n\to\infty} f_n = \int_A \liminf_{n\to\infty} f_n \leq \liminf_{n\to\infty} \int_A f_n\,.$$

2.33 Corollary: *(Lebesgue's Monotone Convergence Theorem) Let* $f_n : A \subseteq \mathbf{R} \to [0,\infty]$ *be nonnegative measurable functions such that* $\{f_n(x)\}$ *is increasing for every* $x \in A$. *Then*

$$\int_A \lim_{n\to\infty} f_n = \lim_{n\to\infty} \int_A f_n\,.$$

Proof: This is a special case of the previous corollary.

2.34 Note: We now return to the proof of the second formula in Part (2) of Theorem 2.30. We suppose that $f, g : A \subseteq \mathbf{R} \to [0, \infty]$ are nonnegative measurable functions, and we need to prove that

$$\int_A (f+g) = \int_A f + \int_A g.$$

Proof: Using the construction described in Note 2.28, choose increasing sequences $\{r_n\}$ and $\{s_n\}$ of nonnegative simple functions on A such that $\lim\limits_{n\to\infty} r_n = f$ and $\lim\limits_{n\to\infty} s_n = g$. Then the sequence $\{r_n + s_n\}$ is also increasing with $\lim\limits_{n\to\infty}(r_n + s_n) = f + g$. By the Monotone Convergence Theorem, along with Part (2) of Theorem 2.27, we have

$$\begin{aligned}\int_A (f+g) &= \int_A \lim_{n\to\infty}(r_n+s_n) = \lim_{n\to\infty}\int_A (r_n+s_n) = \lim_{n\to\infty}\left(\int_A r_n + \int_A s_n\right) \\ &= \lim_{n\to\infty}\int_A r_n + \lim_{n\to\infty}\int_A s_n = \int_A \lim_{n\to\infty} r_n + \int_A \lim_{n\to\infty} s_n = \int_A f + \int_A g\,.\end{aligned}$$

2.35 Corollary: *Let $A \subseteq \mathbf{R}$ be measurable and let $\{f_n\}$ be a sequence of nonnegative measurable functions $f_n : A \to [0, \infty]$. Then*

$$\int_A \sum_{n-1}^{\infty} f_n = \sum_{n=1}^{\infty} \int_A f_n\,.$$

Proof: This follows by applying Lebesgue's Monotone Convergence Theorem to the sequence of partial sums $S_n(x) = \sum\limits_{k=1}^{n} f_k(x)$.

2.36 Corollary: *Let $A = \bigcup\limits_{k=1}^{\infty} A_k$ where the sets A_n are measurable and disjoint, and let $f : A \to [0, \infty]$ be nonnegative and measurable. Then*

$$\int_A f = \sum_{n=1}^{\infty} \int_{A_n} f\,.$$

Proof: This follows from the above corollary using $f_n = f \cdot \chi_{A_n}$.

2.37 Remark: For a σ-algebra $\mathcal{C}$, a **measure** on $\mathcal{C}$ is a function $\mu : \mathcal{C} \to [0, \infty]$ such that
(1) $\mu(\emptyset) = 0$, and
(2) if $A_1, A_2, A_3, \cdots \in \mathcal{C}$ are disjoint then $\mu\left(\bigcup\limits_{k=1}^{\infty} A_k\right) = \sum\limits_{k=1}^{\infty} \mu(A_k)$.

When $\mathcal{M}$ is the σ-algebra of Lebesgue measurable sets in $\mathbf{R}$, and $f : \mathbf{R} \to [0, \infty]$ is any nonnegative measurable function on $\mathbf{R}$, the above corollary shows that we can define a measure μ on $\mathcal{M}$ by

$$\mu(A) = \int_A f\,.$$

2.38 Definition: For a measurable function $f : A \subseteq \mathbf{R} \to [-\infty, \infty]$, we say that f is (Lebesgue) **integrable** (on A) when the functions f^+ and f^- are both Lebesgue integrable on A and, in this case, we define the (Lebesgue) **integral** of f on A to be

$$\int_A f(x)\,dx = \int_A f = \int_A f\; d\lambda = \int_A f^+ - \int_A f^- .$$

In the case that $A = [a,b]$ we also write $\int_A f(x)\,dx$ as $\int_a^b f(x)\,dx$.

2.39 Note: For $f : A \subseteq \mathbf{R} \to [-\infty, \infty]$, f is integrable if and only if $|f|$ is integrable.

2.40 Theorem: *Let $f, g : A \subseteq \mathbf{R} \to [-\infty, \infty]$ be integrable and let $c \in \mathbf{R}$.*

(1) We have $\left| \int_A f \right| \leq \int_A |f|$.

(2) If $f \leq g$ then $\int_A f \leq \int_A g$.

(3) We have $\int_A (cf) = c \int_A f$ and $\int_A (f+g) = \int_A f + \int_A g$.

(4) If $A = B \cup C$ where B and C are disjoint and measurable then $\int_A f = \int_B f + \int_C f$.

(5) If $B \subseteq A$ is measurable then $\int_B f = \int_A f \cdot \chi_B$.

(6) If $\lambda(A) = 0$ then $\int_A f = 0$.

(7) If $f = g$ a.e. on A then $\int_A f = \int_A g$, and if $\int_A |f| = 0$ then $f = 0$ a.e. in A.

Proof: The proof is left as an exercise.

2.41 Theorem: *(Lebesgue's Dominated Convergence Theorem) Let $A \subseteq \mathbf{R}$ be a measurable set and let $f_n : A \to [-\infty, \infty]$ be measurable functions for $n \in \mathbf{Z}^+$. Suppose the pointwise limit $\lim_{n\to\infty} f_n(x)$ exists for all $x \in A$. Suppose there exists an integrable function $g : A \to [0, \infty]$ such that $|f_n(x)| \leq g(x)$ for all $n \in \mathbf{Z}^+$, $x \in A$. Then*

$$\int_A \lim_{n\to\infty} f_n = \lim_{n\to\infty} \int_A f_n .$$

Proof: Let $f = \lim_{n\to\infty} f_n$. Note that since $-g \leq f_n \leq g$ for all n we gave $-g \leq f \leq g$ so that f is integrable. By Fatou's Lemma, applied to the function $g + f_n$, we have

$$\int_A g + \int_A \lim_{n\to\infty} f_n = \int_A \liminf_{n\to\infty} (g + f_n) \leq \liminf_{n\to\infty} \int_A (g + f_n) = \int_A g + \liminf_{n\to\infty} f_n .$$

It follows, since $\int_A g < \infty$, that

$$\liminf_{n\to\infty} \int_A f_n \geq \int_A \lim_{n\to\infty} f_n .$$

By Fatou's Lemma, applied to the function $g - f_n$, we have

$$\int_A g - \int_A \lim_{n\to\infty} f_n = \int_A \liminf_{n\to\infty} (g - f_n) \leq \liminf_{n\to\infty} \int_A (g - f_n) = \int_A g - \limsup_{n\to\infty} \int_A f_n .$$

It follows, since $\int_A g < \infty$, that

$$\limsup_{n\to\infty} \int_A f_n \leq \int_A \lim_{n\to\infty} f_n .$$

2.42 Theorem: *Let $a, b \in \mathbf{R}$ with $a < b$ and let $f : [a, b] \to \mathbf{R}$ be bounded and Riemann integral. Then f is also measurable and Lebesgue integrable, and the two kinds of integral are equal.*

Proof: I may include a proof later.

2.43 Remark: I may include a discussion of complex-valued functions $f : A \subseteq \mathbf{R} \to \mathbf{C}$ later.

.

Chapter 3. The Lp Spaces

3.1 Definition: Let $F = \mathbf{R}$ or $\mathbf{C}$. Let W be a vector space over F. An **inner product** over F is a function $\langle\ ,\ \rangle : W \times W \to W$ (meaning that if $u, v \in W$ then $\langle u, v\rangle \in W$) such that for all $u, v, w \in W$ and all $t \in F$ we have

(1) (Sesquilinearity) $\langle u+v, w\rangle = \langle u, w\rangle + \langle v, w\rangle$, $\langle tu, v\rangle = t\,\langle u, v\rangle$,
$\langle u, v+w\rangle = \langle u, w\rangle + \langle v, w\rangle$, $\langle u, tv\rangle = \overline{t}\,\langle u, v\rangle$,
(2) (Conjugate Symmetry) $\langle u, v\rangle = \overline{\langle v, u\rangle}$, and
(3) (Positive Definiteness) $\langle u, u\rangle \geq 0$ with $\langle u, u\rangle = 0 \iff u = 0$.

For $u, v \in W$, $\langle u, v\rangle$ is called the inner product of u with v. An **inner product space** over F is a vector space over F equipped with an inner product. Given two inner product spaces U and V over F, a linear map $L : U \to V$ is called a **homomorphism** of inner product spaces (or we say that L **preserves inner product**) when $\big\langle L(x), L(y)\big\rangle = \langle x, y\rangle$ for all $x, y \in U$.

3.2 Theorem: *Let W be an inner product space over $F = \mathbf{R}$ or $\mathbf{C}$ and let $u, v \in W$. Then if $\langle x, u\rangle = \langle x, v\rangle$ for all $x \in U$, or if $\langle u, x\rangle = \langle v, x\rangle$ for all $x \in U$ then $u = v$.*

Proof: Suppose that $\langle x, u\rangle = \langle x, v\rangle$ for all $x \in U$. Then $\langle x, u-v\rangle = \langle x, u\rangle - \langle x, v\rangle = 0$ for all $x \in U$. In particular, taking $x = u - v$ we have $\langle u-v, u-v\rangle = 0$, so $u - v = 0$ hence u=v. Similarly, if $\langle u, x\rangle = \langle v, x\rangle$ for all $x \in U$ then $u = v$.

3.3 Definition: Let $F = \mathbf{R}$ or $\mathbf{C}$. Let W be a vector space over F. A **norm** on W is a map $||\ \ || : W \to \mathbf{R}$ such that for all $u, v \in W$ and all $t \in F$ we have

(1) (Scaling) $||tu|| = |t|\,||u||$,
(2) (Positive Definiteness) $||u|| \geq 0$ with $||u|| = 0 \iff u = 0$, and
(3) (Triangle Inequality) $||u+v|| \leq ||u|| + ||v||$.

For $u \in W$ the real number $||u||$ is called the **norm** (or **length**) of u, and we say that u is a **unit vector** when $||u|| = 1$. A **normed linear space** over F is a vector space over F equipped with a norm. Given two normed linear spaces U and V over F, a linear map $L : U \to V$ is called a **homomorphism** of normed linear spaces (or we say that L **preserves norm**) when $\big|\big|L(x)\big|\big| = ||x||$ for all $x \in U$.

3.4 Theorem: *Let $F = \mathbf{R}$ or $\mathbf{C}$. Let W be an inner product space over F. For $u \in W$ define $||u|| = \sqrt{\langle u, u\rangle}$. Then*

(1) (Scaling) $||tu|| = |t|\,||u||$,
(2) (Positive Definiteness) $||u|| \geq 0$ with $||u|| = 0 \iff u = 0$,
(3) $||u+v||^2 = ||u||^2 + 2\,\mathrm{Re}\langle u, v\rangle + ||v||^2$,
(4) (Pythagoras' Theorem) if $\langle u, v\rangle = 0$ then $||u+v||^2 = ||u||^2 + ||v||^2$,
(5) (Parallelogram Law) $||u+v||^2 + ||u-v||^2 = 2||u||^2 + 2||v||^2$,
(6) (Polarization Identity) if $F = \mathbf{R}$ then $\langle u, v\rangle = \frac{1}{4}\big(||u+v|| - ||u-v||\big)$ and if $F = \mathbf{C}$ then $\langle u, v\rangle = \frac{1}{4}\big(||u+v||^2 + i||u+iv||^2 - ||u-v||^2 - i||u-iv||^2\big)$,
(7) (The Cauchy-Schwarz Inequality) $|\langle u, v\rangle| \leq |u|\,|v|$ with $|\langle u, v\rangle| = ||u||\,||v||$ if and only if $\{u, v\}$ is linearly dependent, and
(8) (The Triangle Inequality) $\big|||u|| - ||v||\big| \leq ||u+v|| \leq ||u|| + ||v||$.

In particular, $||\ \ ||$ is a norm on W.

Proof: We only prove Part (7) and part of Part (8). To prove Cauchy's Inequality, suppose first that $\{u,v\}$ is linearly dependent. Then one of x and y is a multiple of the other, say $v=tu$ with $t\in F$. Then $|\langle u,v\rangle| = |\langle u,tu\rangle| = \left|\overline{t}\,\langle u,u\rangle\right| = |t|\,||u||^2 = ||u||\,||tu|| = ||u||\,||v||$.

Next we suppose that $\{u,v\}$ is linearly independent. Then $1\cdot v+t\cdot u\neq 0$ for all $t\in F$, so in particular $v-\frac{\langle v,u\rangle}{|u|^2}\,u\neq 0$. Thus we have

$$\begin{aligned}0 &< \left\|v-\tfrac{\langle v,u\rangle}{||u||^2}\,u\right\|^2 = \left\langle v-\tfrac{\langle v,u\rangle}{||u||^2}\,u\,,\,v-\tfrac{\langle v,u\rangle}{||u||^2}\,u\right\rangle\\
&= \langle v,v\rangle - \overline{\tfrac{\langle v,u\rangle}{||u||^2}}\,\langle v,u\rangle - \tfrac{\langle v,u\rangle}{||u||^2}\,\langle u,v\rangle + \tfrac{\langle v,u\rangle}{||u||^2}\,\overline{\tfrac{\langle v,u\rangle}{||u||^2}}\,\langle u,u\rangle\\
&= ||v||^2 - \tfrac{|\langle u,v\rangle|^2}{||u||^2}\end{aligned}$$

so that $\frac{|\langle u,v\rangle|^2}{||u||^2} < ||v||^2$ and hence $|\langle u,v\rangle|\leq ||u||\,||v||$. This proves Part (7).

Using Parts (3) and (7), and the inequality $|\operatorname{Re}(z)|\leq |z|$ for $z\in\mathbf{C}$ (which follows from Pythagoras' Theorem in $\mathbf{R}^2$), we have

$$\begin{aligned}||u+v||^2 &= ||u||^2 + 2\operatorname{Re}\langle u,v\rangle + ||v||^2 \leq ||u||^2 + 2|\langle u,v\rangle| + ||v||^2\\
&\leq ||u||^2 + 2||u||\,||v|| + ||v||^2 = \big(||u||+||v||\big)^2.\end{aligned}$$

Taking the square root on both sides gives $||u+v||\leq ||u||+||v||$.

3.5 Definition: A **metric** on a set X is a function $d:X\times X\to\mathbf{R}$ such that, for all $x,y,z\in X$ we have

(1) (Positive Definiteness) $d(x,y)\geq 0$ with $d(x,y)=0 \iff x=y$,
(2) (Symmetry) $d(x,y)=d(y,x)$ and
(3) (Triangle Inequality) $d(x,z)\leq d(x,y)+d(y,z)$.

A set with a metric is called a **metric space**.

3.6 Definition: A **topology** on a set X is a set $\mathcal{T}$ of subsets of X such that

(1) $\emptyset\in\mathcal{T}$ and $X\in\mathcal{T}$,
(2) if $U\in\mathcal{T}$ and $V\in\mathcal{T}$ then $U\cap V\in\mathcal{T}$, and
(3) if K is a set and $U_k\in\mathcal{T}$ for each $k\in K$ then $\bigcup\limits_{k\in K} U_k\in\mathcal{T}$.

For a subset $A\subseteq X$, we say that A is **open** (in X) when $A\in\mathcal{T}$ and we say that A is **closed** (in X) when $X\setminus A\in\mathcal{T}$. A set with a topology is called a **topological space**.

3.7 Note: Given an inner product on a vector space V over $F=\mathbf{R}$ or $\mathbf{C}$, Theorem 3.4 shows that we can define an associated norm on V by letting $||x||=\sqrt{\langle x,x\rangle}$ for $x\in V$.

Given a norm on a vector space V, verify that we can define an associated metric on any subset $X\subseteq V$ by letting $d(x,y)=||x-y||$ for $x,y\in X$.

Given a metric on a set X, verify that we can define an associated topology on X by stipulating that a subset $A\subseteq X$ is open when it has the property that for all $a\in A$ there exists $r>0$ such that $B(a,r)\subseteq A$, where $B(a,r)=\left\{x\in X\,\middle|\,d(x,a)<r\right\}$.

3.8 Definition: Let $\{x_n\}_{n\geq 1}$ be a sequence in a metric space X. We say that the sequence $\{x_n\}$ **converges** in X when there exists $a \in X$ such that $\lim\limits_{n\to\infty} x_n = a$, that is when

$$\exists\, a\in X\ \forall\, \epsilon>0\ \exists\, n\in \mathbf{Z}^+\ \forall k\in \mathbf{Z}^+\ \big(k\geq n \Longrightarrow d(x_n,a)<\epsilon\big).$$

We say that $\{x_n\}$ is **Cauchy** when

$$\forall\, \epsilon>0\ \exists\, n\in \mathbf{Z}^+\ \forall\, k,l\in \mathbf{Z}^+\ \big(k,l\geq n \Longrightarrow d(x_k,x_l)<\epsilon\big).$$

3.9 Note: Verify that, in a metric space, if a sequence converges then it is Cauchy.

3.10 Definition: A metric space X is called **complete** when, in X, every Cauchy sequence converges. A complete normed linear space is called a **Banach space** and a complete inner-product space is called a **Hilbert space**.

3.11 Theorem: *(The Completeness of $\mathbf{R}^n$) The metric space $\mathbf{R}^n$ is complete.*

Proof: We omit the proof.

3.12 Definition: Let $\mathbf{R}^\omega$ denote the set of all sequences $x = \{x_1, x_2, x_3, \cdots\}$ with each $x_k \in \mathbf{R}$. For $x \in \mathbf{R}^\omega$ and for $1 \leq p < \infty$ let

$$\begin{aligned} ||x||_p &= \Big(\sum_{k=1}^{\infty} |x_k|^p\Big)^{1/p} \text{ , and} \\ ||x||_\infty &= \sup\big\{|x_k|\,\big|\,k\in\mathbf{Z}^+\big\}. \end{aligned}$$

Let

$$\begin{aligned} \ell_p &= \big\{x\in\mathbf{R}^\omega\,\big|\,||x||_p<\infty\big\} \text{ , and} \\ \ell_\infty &= \big\{x\in\mathbf{R}^\omega\,\big|\,||x||_\infty<\infty\big\}. \end{aligned}$$

3.13 Definition: Let $A \subseteq \mathbf{R}$ be measurable. Let $\mathcal{M}(A)$ denote the set of all measurable functions $f : A \to [-\infty, \infty]$. For $f \in \mathcal{M}(A)$ and for $1 \leq p < \infty$, let

$$\begin{aligned} ||f||_p &= \left(\int_A |f|^p\right)^{1/p} \text{ , and} \\ ||f||_\infty &= \inf\Big\{a\geq 0\,\Big|\,\lambda\big(|f|^{-1}(a,\infty]\big)=0\Big\}. \end{aligned}$$

where $|f|^{-1}(a,\infty] = \big\{x \in A\,\big|\,|f(x)| > a\big\}$. Let

$$\begin{aligned} L_p(A) &= \Big\{f\in\mathcal{M}(A)\,\Big|\,||f||_p<\infty\Big\}\Big/\sim \quad \text{, and} \\ L_\infty(A) &= \Big\{f\in\mathcal{M}(A)\,\Big|\,||f||_\infty<\infty\Big\}\Big/\sim \end{aligned}$$

where $\sim$ is the equivalence relation given by $f \sim g \iff f = g$ a.e. in A.

3.14 Remark: The reason that we quotient by the equivalence relation in the above definition is that we want $||f||_p$ to define a norm on $L_p(A)$ and the quotient is necessary to ensure that $||f||_p$ is positive definite (see Part 6 of Theorem 2.30).

3.15 Lemma: *Let $f : A \subseteq \mathbf{R} \to [-\infty, \infty]$ be measurable. Then $\{x \in A \big| |f(x)| > ||f||_\infty\}$ has measure zero.*

Proof: We claim that for all $y > ||f||_\infty$ we have $\lambda\big(|f|^{-1}(y,\infty]\big) = 0$. Let $y > ||f||_\infty$. By the definition of $||f||_\infty$ we can choose a with $||f||_\infty \leq a < y$ such that $\lambda\big(|f|^{-1}(a,\infty]\big) = 0$. Since $a < y$ we have $(y,\infty] \subseteq (a,\infty]$, so $|f|^{-1}(y,\infty] \subseteq |f|^{-1}(a,\infty]$, hence $\lambda\big(|f|^{-1}(y,\infty]\big) = 0$, as claimed.

Let $B = \{x \in A \big| |f(x)| > ||f||_\infty\}$ and let $B_n = \{x \in A \big| |f(x)| > ||f||_\infty + \frac{1}{n}\}$ for $n \in \mathbf{Z}^+$. Then each B_n is measurable with $B_1 \subseteq B_2 \subseteq B_3 \subseteq \cdots$, and we have $\bigcup\limits_{n=1}^{\infty} B_n = B$. By the above claim, we have $\lambda(B_n) = 0$ for all $n \in \mathbf{Z}^+$ and so $\lambda(B) = \lim\limits_{n \to \infty} \lambda(B_n) = 0$.

3.16 Definition: For $p, q \in [1, \infty]$ we say that p and q are **conjugate** when $\frac{1}{p} + \frac{1}{q} = 1$ where we use the convention that $\frac{1}{\infty} = 0$ so that 1 and ∞ are conjugate.

3.17 Lemma: *Let $p, q \in (1, \infty)$ with $\frac{1}{p} + \frac{1}{q} = 1$. Then for all $a, b \geq 0$ we have $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$.*

Proof: Note that for $p, q \in (1, \infty)$ we have

$$\tfrac{1}{p} + \tfrac{1}{q} = 1 \iff \tfrac{1}{q} = 1 - \tfrac{1}{p} = \tfrac{p-1}{p} \iff q(p-1) = p \iff p(q-1) = q.$$

For $x, y \geq 0$ we have

$$y = x^{p-1} \iff y^q = x^{q(p-1)} \iff y^q = x^p \iff y^{p(q-1)} = x^p \iff y^{q-1} = x$$

so the functions $f(x) = x^{p-1}$ and $g(y) = y^{q-1}$ are inverses of each other. By considering the area under $y = f(x)$ with $0 \leq x \leq a$ and the area to the left of $y = f(x)$ with $0 \leq y \leq b$ we see that

$$ab \leq \int_{x=0}^{a} x^{p-1}\,dx + \int_{y=0}^{b} y^{q-1}\,dy = \Big[\tfrac{1}{p}\,x^p\Big]_{x=0}^{a} + \Big[\tfrac{1}{q}\,y^q\Big]_{y=0}^{b} = \tfrac{a^p}{p} + \tfrac{b^q}{q}\,.$$

3.18 Theorem: *(Hölder's Inequality) Let $p, q \in [1, \infty]$ with $\frac{1}{p} + \frac{1}{q} = 1$ and let $A \subseteq \mathbf{R}$ be measurable.*
(1) For all $x, y \in \mathbf{R}^\omega$ we have $||xy||_1 \leq ||x||_p ||y||_q$.
(2) For all $f, g \in \mathcal{M}(A)$ we have $||fg||_1 \leq ||f||_p ||g||_q$.

Proof: To prove Part (1) in the case that $p, q \in (1, \infty)$, let $x, y \in \mathbf{R}^\omega$. If $x = 0$ or $y = 0$ the equality holds, so suppose that $x, y \neq 0$. For each index k, apply the above lemma using $a = \frac{|x_k|}{||x||_p}$ and $b = \frac{|y_k|}{||y||_q}$ to get

$$\frac{|x_k y_k|}{||x||_p ||y||_q} \leq \frac{|x_k|^p}{p||x||_p^p} + \frac{|y|^q}{q||y||_q^q}.$$

Sum over k to get

$$||xy||_1 \leq \frac{||x||_p^p}{p||x||_p^p} + \frac{||y||_q^q}{q||y||_q^q} = \tfrac{1}{p} + \tfrac{1}{q} = 1\,.$$

To prove Part (2) in the case that $p, q \in (1, \infty)$, let $f, g \in \mathcal{M}(\mathbf{R})$. If $f = 0$ or $g = 0$ then the equality holds, so suppose that $f, g \neq 0$. For each $x \in A$, apply the above lemma using $a = \frac{|f(x)|}{||f||_p}$ and $b = \frac{|g(x)|}{||g||_q}$ to get

$$\frac{|f(x)g(x)|}{||f||_p ||g||_q} \leq \frac{|f(x)|^p}{p||f||_p} + \frac{|g(x)|^q}{q||g||_q}.$$

Integrate over A to get

$$||fg||_1 \leq \frac{||f||_p^p}{p||f||_p^p} + \frac{||g||_q^q}{q||g||_q^q} = \tfrac{1}{p} + \tfrac{1}{q} = 1\,.$$

To prove Part (1) in the case that $p = 1$ and $q = \infty$, let $x, y \in \mathbf{R}^\omega$. Note that $|y_k| \leq ||y||_\infty$ for all indices k and so

$$||xy||_1 = \sum_{k=1}^{\infty} |x_k||y_k| \leq \sum_{k=1}^{\infty} |x_k|\,||y||_\infty = ||x||_1 ||y||_\infty .$$

Finally, to prove Part (2) in the case that $p = 1$ and $q = \infty$, let $f, g \in \mathcal{M}(A)$. Let $B = \big\{x \in A \big| |g(x)| \leq ||g||_\infty\big\}$ and let $C = \big\{x \in A \big| g(x)| > ||g||_\infty\big\}$. Note that B and C are disjoint and measurable with $A = B \cup C$ and that $\lambda(C) = 0$ by Lemma 3.15. Thus

$$||fg||_1 = \int_A |f||g| = \int_B |f||g| \leq \int_B |f|\,||g||_\infty = \int_A |f|\,||g||_\infty = ||f||_1 ||g||_\infty .$$

3.19 Theorem: *(Minkowski's Inequality) Let $p \in [1, \infty]$ and let $A \subseteq \mathbf{R}$ be measurable.*
(1) For all $x, y \in \mathbf{R}^\omega$ we have $||x + y||_p \leq ||x||_p + ||y||_p$.
(2) For all $f, g \in \mathcal{M}(A)$ for which $f + g$ is defined, we have $||f + g||_p \leq ||f||_p + ||g||_p$.

Proof: To Prove Part (1) in the case that $p = 1$, note that when $x, y \in \mathbf{R}^\omega$ we have

$$||x + y||_1 = \sum_{k=1}^{\infty} |x_k + y_k| \leq \sum_{k=1}^{\infty} |x_k| + |y_k| = \sum_{k=1}^{\infty} |x_k| + \sum_{k=1}^{\infty} |y_k| = ||x||_1 + ||y||_1 .$$

To prove Part (2) in the case that $p = 1$, note that when $f, g, f + g \in \mathcal{M}(A)$ we have

$$||f + g||_1 = \int_A |f + g| \leq \int_A |f| + |g| = \int_A |f| + \int_A |g| = ||f||_1 + ||g||_1 .$$

To prove Part (1) in the case that $p \in (1, \infty)$, let $x, y \in \mathbf{R}^\omega$ and let q be the conjugate of p so that $\frac{1}{q} = 1 - \frac{1}{p} = \frac{p-1}{p}$. For each index k we have

$$\begin{aligned} |x_k + y_k|^p &= |x_k + y_k|\,|x_k + y_k|^{p-1} \leq \big(|x_k| + |y_k|\big)|x_k + y_k|^{p-1} \\ &= |x_k|\,|x_k + y_k|^{p-1} + |y_k|\,|x_k + y_k|^{p-1}. \end{aligned}$$

Sum over k then apply Hölder's Inequality to get

$$\begin{aligned} ||x + y||_p^p &\leq \Big\| |x|\,|x + y|^{p-1} \Big\|_1 + \Big\| |y|\,|x + y|^{p-1} \Big\|_1 \leq ||x||_p \Big\| |x + y|^{p-1} \Big\|_q + ||y||_p \Big\| |x + y|^{p-1} \Big\|_q \\ &= \Big(||x||_p + ||y||_q\Big) \Big\| |x + y|^{p-1} \Big\|_q = \Big(||x||_p + ||y||_q\Big) \Big(\sum_{k=1}^{\infty} |x + y|^{q(p-1)} \Big)^{1/q} \\ &= \Big(||x||_p + ||y||_q\Big) \Big(\sum_{k=1}^{\infty} |x + y|^p \Big)^{(p-1)/p} = \Big(||x||_p + ||y||_q\Big) ||x + y||_p^{p-1}. \end{aligned}$$

To prove Part (2) in the case that $p\in(1,\infty)$, let $f,g,f+g\in\mathcal{M}(A)$ and let q be the conjugate of p so that $\frac{1}{q}=1-\frac{1}{p}=\frac{p-1}{p}$. For each $x\in A$ we have

$$\begin{aligned}|f(x)+g(x)|^p&=|f(x)+g(x)|\,|f(x)+g(x)|^{p-1}\leq\big(|f(x)|+|g(x)|\big)|f(x)+g(x)|^{p-1}\\&=|f(x)|\,|f(x)+g(x)|^{p-1}+|g(x)|\,|f(x)+g(x)|^{p-1}.\end{aligned}$$

Integrate over A then apply Hölder's Inequality to get

$$\begin{aligned}||f+g||_p^p&\leq\Big\||f|\,|f+g|^{p-1}\Big\|_1+\Big\||g|\,|f+g|^{p-1}\Big\|_1\leq||f||_p\Big\||f+g|^{p-1}\Big\|_q+||g||_p\Big\||f+g|^{p-1}\Big\|_q\\&=\Big(||f||_p+||g||_q\Big)\Big\||f+g|^{p-1}\Big\|_q=\Big(||f||_p+||g||_q\Big)\Big(\int_A|f+g|^{q(p-1)}\Big)^{1/q}\\&=\Big(||f||_p+||g||_q\Big)\Big(\int_A|f+g|^p\Big)^{(p-1)/p}=\Big(||f||_p+||g||_q\Big)||f+g||_p^{p-1}.\end{aligned}$$

To prove Part (1) in the case that $p=\infty$, note that if $x,y\in\ell_\infty$ then we have

$$||x+y||_\infty=\sup_{k\geq1}|x_k+y_k|\leq\sup_{k\geq1}\big(|x_k|+|y_k|\big)\leq\sup_{k\geq1}|x_k|+\sup_{k\geq1}|y_k|=||x||_\infty+||y||_\infty.$$

To prove Part (2) in the case that $p=\infty$, let $f,g\in\mathcal{M}(A)$. For all $x\in A$, note that if $\big|f(x)+g(x)\big|>||f||_\infty+||g||_\infty$ then $|f(x)|+|g(x)|\geq\big|f(x)+g(x)\big|>||f||_\infty+||g||_\infty$ and hence either $|f(x)|>||f||_\infty$ or $|g(x)|>||g||_\infty$. This shows that

$$\big\{x\in A\big||f(x)+g(x)|>||f||_\infty+||g||_\infty\big\}\subseteq\big\{x\in A\big||f(x)|>||f||_\infty\big\}\cup\big\{x\in A\big||g(x)|>||g||_\infty\big\}.$$

By Lemma 3.15, the two sets on the right both have measure zero, and so the set on the left has measure zero. By the definition of $||f+g||_\infty$ it follows that $||f+g||_\infty\leq||f||_\infty+||g||_\infty$.

3.20 Corollary: *Let $p\in[1,\infty]$ and let $A\subseteq\mathbf{R}$ be measurable. Then ℓ_p and $L_p(A)$ are normed linear spaces using their p-norms.*

Proof: We prove that $L_p(A)$ is a normed linear space when $p\in[1,\infty)$. For $f,g\in\mathcal{M}(A)$ and $c\in\mathbf{R}$, we have $||f||_p=\Big(\int_A|f|^p\Big)^{1/p}\geq0$ and, by Part 6 of Theorem 2.30,

$$||f||=0\iff\int_A|f|^p=0\iff|f|^p=0\text{ a.e. in }A,$$

and we have $||cf||_p=\Big(\int_A|cf|^p\Big)^{1/p}=|c|\Big(\int_A|f|^p\Big)^{1/p}=|c|\,||f||_p$, and by Minkowski's Inequality we have $||f+g||_p\leq||f||_p+||g||_p$. This shows that $||f||_p$ satisfies the three properties which define a norm. We also need to verify that $L_p(A)$ is a vector space. Let $V=\big\{f\in\mathcal{M}(A)\big|||f||_p<\infty\big\}$. Then V is a vector space because if $f,g\in V$ and $c\in\mathbf{R}$ then we have $cf\in V$ because $||cf||_p=|c|||f||_p<\infty$ and we have $f+g\in V$ because $||f+g||_p\leq||f||_p+||g||_p<\infty$ by Minkowski's Inequality. Note that $L_p(A)$ is the quotient space of the vector space V by the subspace $W=\big\{f\in V\big|f=0\text{ a.e. in }A\big\}$.

3.21 Theorem: *Let $p \in [1,\infty]$ and let $A \subseteq \mathbf{R}$ be measurable. Then the normed linear spaces ℓ_p and $L_p(A)$ are complete.*

Proof: We leave the proof that ℓ_p is complete as an exercise. To prove that $L_p(A)$ is complete in the case that $p < \infty$, let $\{f_n\}$ be a Cauchy sequence in $L_p(A)$. This means that for all $\epsilon > 0$ there exists $m \in \mathbf{Z}^+$ such that $k, l \geq m \Longrightarrow ||f_k - f_l||_p < \epsilon$. Choose a subsequence $\{f_{n_k}\}$ with the property that $\left|\left|f_{n_{k+1}} - f_{n_k}\right|\right|_p \leq \frac{1}{2^k}$ for all $k \geq 1$. For each $\ell \in \mathbf{Z}^+$, let

$$g_\ell = \sum_{k=1}^{\ell} \left|f_{n_{k+1}} - f_{n_k}\right|$$

and let $g = \lim_{\ell\to\infty} g_\ell$ (note that the limit exists because $\{g_\ell(x)\}$ is increasing for all $x \in A$). By Minkowski's Inequality, for all $\ell \in \mathbf{Z}^+$ we have

$$||g_\ell||_p \leq \sum_{k=1}^{\ell} ||f_{n_{k+1}} - f_{n_k}||_p \leq \sum_{k=1}^{\ell} \tfrac{1}{2^k} < 1.$$

By Fatou's Lemma,

$$||g||_p^p = \int_A |g|^p = \int_A \lim_{\ell\to\infty} |g_\ell|^p \leq \liminf_{\ell\to\infty} \int_A |g_\ell|^p = \liminf_{\ell\to\infty} ||g_\ell||_p^p \leq 1$$

so that $g \in L_p(A)$. Because $||g||_p$ is finite, it follows that g is finite a.e. in A, so the sum $\sum \left|f_{n_{k+1}} - f_{n_k}\right|$ converges a.e. in A, hence the sum $\sum \left(f_{n_{k+1}} - f_{n_k}\right)$ converges a.e. in A, and hence the sequence $\{f_{n_\ell}\}$ converges a.e. in A because $f_{n_\ell} = f_{n_1} + \sum_{k=1}^{\ell-1} \left(f_{n_{k+1}} - f_{n_k}\right)$. We define $f : A \to \mathbf{R}$ by

$$f(x) = \begin{cases} \lim_{\ell\to\infty} f_{n_\ell}(x) & \text{, if the limit exists in } \mathbf{R}\text{, and} \\ 0 & \text{, otherwise.} \end{cases}$$

We claim that $f \in L_p(A)$ and that $\lim_{n\to\infty} f_n = f$ in $L_p(A)$. Let $\epsilon > 0$. Choose $m \in \mathbf{Z}^+$ so that for all $k, l \geq m$ we have $||f_k - f_l||_p \leq \epsilon$. Then for all k such that $n_k \geq m$ we have $||f_{n_k} - f_m||_p \leq \epsilon$. By Fatou's Lemma,

$$\begin{aligned} ||f - f_m||_p^p &= \int_A |f - f_m|^p = \int_A \lim_{k\to\infty} |f_{n_k} - f_m|^p \\ &\leq \liminf_{k\to\infty} \int_A |f_{n_k} - f_m|^p = \liminf_{k\to\infty} ||f_{n_k} - f_m||_p^p \leq \epsilon^p \end{aligned}$$

so that $||f - f_m||_p \leq \epsilon$. This shows that for all $\epsilon > 0$ there exists $m \in \mathbf{Z}^+$ such that for all $n \geq m$ we have $||f - f_n||_p \leq \epsilon$. It will follow that $\lim_{n\to\infty} f_n = f$ in $L_p(A)$ once we show that $f \in L_p(A)$. Taking $\epsilon = 1$ and choosing m as above so that $||f - f_m|| \leq 1$, Minkowski's Inequality gives $||f||_p \leq ||f - f_m||_p + ||f_m||_p \leq 1 + ||f_m||_p < \infty$ so that $f \in L_p(A)$, as required.

Now let us prove that $L_\infty(A)$ is complete. Let $\{f_n\}$ be a Cauchy sequence in $L_\infty(A)$. Let $B_n = \big\{x \in A \big| |f_n(x) > ||f_n||_\infty\big\}$ and let $C_{k,l} = \big\{x \in A \big| |f_k(x) - f_l(x)| > ||f_k - f_l||_\infty\big\}$. By Lemma 3.15, the sets B_n and $C_{k,l}$ all have measure zero. Let E be the union of all the sets B_n and $C_{k,l}$. Since E is a countable union of sets of measure zero, we have $\lambda(E) = 0$. Given $\epsilon > 0$, since $\{f_n\}$ is Cauchy in $L_\infty(A)$ we can choose $m \in \mathbf{Z}^+$ so that for all $k, l \geq m$ we have $||f_k - f_l||_\infty \leq \epsilon$. Then for all $k, l \geq m$ we have $|f_k(x) - f_l(x)| \leq ||f_k - f_l||_\infty \leq \epsilon$ for all $x \in A \setminus E$. It follows, by the Cauchy criterion for uniform convergence, that the sequence $\{f_n\}$ converges uniformly in $A \setminus E$. Define $f : A \to \mathbf{R}$ by

$$f(x) = \begin{cases} \lim\limits_{n\to\infty} f_n(x) & \text{, if } x \in A \setminus E \\ 0 & \text{, if } x \in E. \end{cases}$$

We claim that $f \in L_\infty(A)$ and that $\lim\limits_{n\to\infty} f_n = f$ in $L_\infty(A)$. Given $\epsilon > 0$, since $\{f_n\}$ converges uniformly to f in $A \setminus E$, we can choose $m \in \mathbf{Z}^+$ so that for all $n \geq m$ we have $|f_n(x) - f(x)| \leq \epsilon$ for all $x \in A \setminus E$ hence $||f_n - f||_\infty \leq \epsilon$ since $\lambda(E) = 0$. This shows that for all $\epsilon > 0$ there exists $m \in \mathbf{Z}^+$ such that for all $n \geq m$ we have $||f - f_n||_\infty \leq \epsilon$. Taking $\epsilon = 1$ and choosing m as above, we have $||f_m - f||_\infty \leq 1$ so by Minkowski's Inequality $||f||_\infty \leq ||f - f_m||_\infty + ||f_m||_\infty \leq 1 + ||f_m||_\infty$ and so $f \in L_\infty(A)$.

3.22 Theorem: *Let $1 \leq p < q \leq \infty$ and let $A \subseteq \mathbf{R}$ be measurable. Then*
(1) $\ell_p \subseteq \ell_q$, and
(2) if $\lambda(A) < \infty$ then $L_q(A) \subseteq L_p(A)$.

Proof: We leave the proof of Part (1) as an exercise. To prove Part (2), suppose that $\lambda(A) < \infty$. Consider first the case that $q < \infty$. Let $f \in L_q(A)$. Then by Hölder's Inequality, for any $u, v > 1$ with $\frac{1}{u} + \frac{1}{v} = 1$ we have

$$||f||_p^p = \int_A |f|^p = \Big\||f|^p\Big\|_1 \leq \Big\||f|^p\Big\|_u \Big\|1\Big\|_v = \left(\int_A |f|^{pu}\right)^{1/u} \lambda(A)^{1/v}.$$

Choose $u = \frac{q}{p}$ and, to get $\frac{1}{v} = 1 - \frac{1}{u} = 1 - \frac{p}{q} = \frac{q-p}{q}$, choose $v = \frac{q}{q-p}$. Then

$$||f||_p^p \leq \left(\int_A |p|^q\right)^{p/q} \lambda(A)^{(q-p)/q} = ||f||_q^p \, \lambda(A)^{(q-p)/q}$$

so that $||f||_p \leq ||f||_q \, \lambda(A)^{\frac{1}{p} - \frac{1}{q}}$. Thus $||f||_p < \infty$ so $f \in L_p(A)$.

Now consider the case that $q = \infty$. Let $f \in L_\infty(A)$. Let $B = \big\{x \in A \big| |f(x)| \leq ||f||_\infty\big\}$ and $C = \big\{x \in A \big| |f(x)| > ||f||_\infty\big\}$. By Lemma 3.15 we have $\lambda(C) = 0$, so

$$||f||_p^p = \int_A |f|^p = \int_B |f|^p \leq \int_B ||f||_\infty^p = ||f||_\infty^p \, \lambda(B) = ||f||_\infty^p \, \lambda(A)$$

so that $||f||_p \leq ||f||_\infty \, \lambda(A)^{1/p}$. Thus $||f||_p < \infty$ so $f \in L_p(A)$.

3.23 Theorem: *Let $1 \leq p < q < r \leq \infty$ and let $A \subseteq \mathbf{R}$ be measurable. Then*
(1) $\ell_p \cap \ell_r \subseteq \ell_q \subseteq \ell_p + \ell_r$, and
(2) $L_p(A) \cap L_r(A) \subseteq L_q(A) \subseteq L_p(A) + L_r(A)$.

Proof: Part (1) follows as an immediate corollary of Theorem 3.23. Let us prove Part (2). First we claim that $L_q(A) \subseteq L_p(A) + L_r(A)$. Let $f \in L_q(A)$. Let $B = \left\{x \in A \big| |f(x)| > 1\right\}$ and let $C = \left\{x \in A \big| |f(x)| \leq 1\right\}$. Let $g = f \cdot \chi_{_B}$ and $h = f \cdot \chi_{_C}$ so that $f = g + h$. Note that $g \in L_p(A)$ because

$$||g||_p^p = \int_A |g|^p = \int_B |f|^p \leq \int_B |f|^q \leq \int_A |f|^q = ||f||_q^q < \infty,$$

note that $h \in L_\infty(A)$ because $|f(x)| \leq 1$ for all $x \in A$ so that $||h||_\infty \leq 1$, and note that when $r < \infty$ we have $h \in L_r(A)$ because

$$||h||_r^r = \int_A |h|^r = \int_C |f|^r \leq \int_C |f|^q \leq \int_A |f|^q = ||f||_q^q < \infty.$$

Thus we have $L_q(A) \subseteq L_p(A) + L_r(A)$ as claimed.

Next we claim that $L_p(A) \cap L_r(A) \subseteq L_q(A)$. Let $f \in L_p(A) \cap L_r(A)$. Suppose first that $r < \infty$. Note that for any $0 < k, l \in \mathbf{R}$ with $k + l = q$ and for any $1 < u, v \in \mathbf{R}$ with $\frac{1}{u} + \frac{1}{v} = 1$, Hölder's Inequality gives

$$||f||_q^q = \int_A |f|^q \leq \big\||f|^k\big\|_u \big\||f|^l\big\|_v = \left(\int_A |f|^{ku}\right)^{1/u} \left(\int_A |f|^{lv}\right)^{1/v}.$$

We solve the equations $k + l = q$, $\frac{1}{u} + \frac{1}{v} = 1$, $ku = p$ and $lv = r$ to get

$$k = \tfrac{p(r-q)}{r-p}\ ,\ l = \tfrac{r(q-p)}{r-p}\ ,\ u = \tfrac{r-p}{r-q} \text{ and } v = \tfrac{r-p}{q-p}$$

and note that since $1 \leq p < q < r < \infty$ we have $k, l > 0$ and $1 < u, v < \infty$. Thus

$$||f||_q^q \leq \left(\int_A |f|^{ku}\right)^{1/u} \left(\int_A |f|^{lv}\right)^{1/v} = \left(\int_A |f|^p\right)^{k/p} \left(\int_A |f|^r\right)^{l/r} = ||f||_p^k ||f||_r^l < \infty.$$

When $r = \infty$, we let $B = \left\{x \in A \big| |f(x)| > ||f||_\infty\right\}$ and $C = \left\{x \in A \big| |f(x)| \leq ||f||_\infty\right\}$, and then by Lemma 3.15 we have $\lambda(B) = 0$, and so

$$||f||_q^q = \int_A |f|^q = \int_C |f|^q = \int_C |f|^p |f|^{q-p} \leq ||f||_\infty^{q-p} \int_C |f|^p \leq ||f||_p^p ||f||_\infty^{q-p} < \infty.$$

This proves that $L_p(A) \cap L_r(A) \subseteq L_q(A)$ as claimed.

3.24 Definition: A metrix space is called **separable** when it contains a countable dense subset.

3.25 Theorem: *Let $1 \le p < \infty$ and let $a < b$.*

(1) ℓ_p is separable but ℓ_∞ is not.
(2) $L_p\big([a,b]\big)$ is separable but $L_\infty\big([a,b]\big)$ is not.

Proof: We leave the proof of Part (1) as an exercise. We sketch a proof of Part (2) leaving the details as an exercise. To show that $L_p[a,b]$ is separable, we shall show that $\mathbf{Q}[x]$ is dense in $L_p[a,b]$ by showing that a given function $f \in L_p[a,b]$ can be approximated, arbitrarily closely in the p-norm, by a polynomial in $\mathbf{Q}[x]$. Since $f = f^+ - f^-$ it suffices to consider the case that f is nonnegative. By Note 2.28, together with the Monotone Convergence Theorem, we can approximate a given nonnegative function $f \in L_p[a,b]$, arbitrarily closely in the p-norm, using a nonnegative simple function since we can construct an increasing sequence of simple functions $s_n : [a,b] \to [0,\infty)$ with $s_n \to f$ pointwise on $[a,b]$. We can approximate a given nonnegative simple function $s : [a,b] \to [0,\infty)$, arbitrarily closely in the p-norm, using a nonnegative step function $r : [a,b] \to [0,\infty)$ because we can cover a measurable set $A \subseteq [a,b]$ by a disjoint union of intervals $J_k \subseteq [a,b]$ so that $\chi_{_A}$ is approximated by $\sum \chi_{_{J_k}}$. We can then approximate a given step function $r : [a,b] \to [0,\infty)$, arbitrarily closely in the p-norm, using a continuous function because for any interval J, the step function $\chi_{_J}$ can be approximated arbitrarily closely in the p-norm by a piecewise linear function. This shows that the set of continuous functions $C[a,b]$ is dense in $L_p[a,b]$, using the p-norm. On the other hand, using the ∞-norm (which agrees with the supremum norm for continuous functions), $\mathbf{Q}[x]$ is dense in $\mathbf{R}[x]$, and we know from the Stone-Weirstrass Theorem that $\mathbf{R}[x]$ is dense in $C[a,b]$. Since $\mathbf{Q}[x]$ is dense in $C[a,b]$ using the ∞-norm, it is also dense using the p-norm by the formula $||f||_p \le (b-a)^{1/p}||f||_\infty$ which is obtained in the proof of Theorem 3.22.

We claim that $L_\infty[a,b]$ is not separable. Let S be any dense subset of $L_\infty[a,b]$. We must show that S is uncountable. For each $k \in \mathbf{N}$ let $x_k = b - \frac{b-a}{2^k}$ so that we have $a = x_0 < x_1 < x_2 < \cdots < b$. Let $\{0,1\}^\omega$ denote the set of binary sequences $\alpha = (\alpha_1, \alpha_2, \cdots)$ where each $\alpha_k \in \{0,1\}$. For each $\alpha \in \{0,1\}^\omega$, let $s_\alpha = \sum\limits_{k=1}^{\infty} \alpha_k \chi_{_{[x_{k-1},x_k)}}$ and note that when $\alpha \ne \beta$ we have $||s_\alpha - s_\beta||_\infty = 1$. Since S is dense in $L_\infty[a,b]$, for each $\alpha \in \{0,1\}^\omega$ we can choose $f_\alpha \in S$ such that $||s_\alpha - f_\alpha||_\infty < \frac{1}{2}$. Define $F : \{0,1\}^\omega \to S$ by $F(\alpha) = f_\alpha$. Note that F is injective because when $\alpha \ne \beta$ we have

$$1 = ||s_\alpha - s_\beta||_\infty \le ||s_\alpha - f_\alpha||_\infty + ||f_\alpha - f_\beta||_\infty + ||f_\beta - s_\beta||_\infty < \tfrac{1}{2} + ||f_\alpha - f_\beta||_\infty + \tfrac{1}{2}$$

so that $||f_\alpha - f_\beta||_\infty > 0$. Since F is injective we have $|S| \ge \big|\{0,1\}^\omega\big| = 2^{\aleph_0}$, and so S is uncountable, as required.

3.26 Remark: I may include a discussion of the complex-valued L_p spaces $L_p(A, \mathbf{C})$ later.

.

Chapter 4. Banach and Hilbert Spaces

4.1 Definition: Let W be an inner product space over $F = \mathbf{R}$ or $\mathbf{C}$. For a subset $\mathcal{A} \subseteq W$, we say that $\mathcal{A}$ is **orthogonal** when $\langle u, v\rangle = 0$ for all $u, v \in \mathcal{A}$ with $u \neq v$, and we say that $\mathcal{A}$ is **orthonormal** when $\mathcal{A}$ is orthogonal with $||u|| = 1$ for every $u \in \mathcal{A}$.

4.2 Theorem: *Let W be an inner product space over $F = \mathbf{R}$ or $\mathbf{C}$. Let $\mathcal{A} \subseteq W$.*

(1) If $\mathcal{A}$ is an orthogonal set of nonzero vectors then for $x \in \text{Span}\mathcal{A}$ with say $x = \sum\limits_{k=1}^{n} c_k u_k$ where $c_k \in F$ and $u_k \in \mathcal{A}$, we have $c_k = \langle x, u_k\rangle / ||u_k||^2$ for all indices k, and in particular, $\mathcal{A}$ is linearly independent.

(2) If $\mathcal{A}$ is orthonormal then for $x \in \text{Span}\mathcal{A}$ with say $x = \sum\limits_{k=1}^{n} c_k u_k$ where $c_k \in F$ and $u_k \in \mathcal{A}$, we have $c_k = \langle x, u_k\rangle$ for all k, and in particular, $\mathcal{A}$ is linearly independent.

Proof: To prove Part (1), suppose that $\mathcal{A}$ is an orthogonal set of nonzero vectors and let $x = \sum\limits_{j=1}^{n} c_j u_j$ with each $c_j \in F$ and each $u_j \in \mathcal{A}$. Then for all indices k, since $\langle u_j, u_k\rangle = 0$ whenever $j \neq k$ we have $\langle x, u_k\rangle = \left\langle \sum\limits_{j=1}^{n} c_j u_j \,,\, u_k \right\rangle = \sum\limits_{j=1}^{n} c_j \langle u_j, u_k\rangle = c_k \langle u_k, u_k\rangle = c_k ||u_k||^2$ and so $c_k = \frac{\langle x, u_k\rangle}{||u_k||^2}$, as required. In particular, when $x = 0$ we find that $c_k = 0$ for all k, and this shows that $\mathcal{A}$ is linearly independent. This proves Part (1), and Part (2) follows immediately from Part (1).

4.3 Theorem: *(The Gram-Schmidt Procedure) Let W be a finite or countable dimensional inner product space over $F = \mathbf{R}$ or $\mathbf{C}$. Let $\mathcal{A} = \{u_1, u_2, \cdots\}$ be an ordered basis for W. Let $v_1 = u_1$ and for $n \geq 2$ let $v_n = u_n - \sum\limits_{k=1}^{n-1} \dfrac{\langle u_n, v_k\rangle}{||v_k||^2}\, v_k$. Then the set $\mathcal{B} = \{v_1, v_2, \cdots\}$ is an orthogonal basis for W with the property that for every index $n \geq 1$ we have $\text{Span}\{v_1, \cdots, v_n\} = \text{Span}\{u_1, \cdots, u_n\}$.*

Proof: We prove, by induction on n, that $\{v_1, v_2, \cdots, v_n\}$ is an orthogonal basis for $\text{Span}\{u_1, u_2, \cdots, u_n\}$. When $n = 1$ this is clear since $v_1 = u_1$. Let $n \geq 2$ and suppose, inductively, that $\{v_1, \cdots, v_{n-1}\}$ is an orthogonal basis for $\text{Span}\{u_1, \cdots, u_{n-1}\}$. Since $v_n = u_n - \sum_{k=1}^{n-1} \frac{\langle u_n, v_k\rangle}{||v_k||^2}\, v_k$, we see that u_n is equal to v_n plus a linear combination of the vectors $v_1, \cdots, v_{n-1}$, and so we have $\text{Span}\{v_1, \cdots, v_{n-1}, v_n\} = \text{Span}\{v_1, \cdots, v_{n-1}, u_n\}$. By the induction hypothesis, we have $\text{Span}\{v_1, \cdots, v_{n-1}\} = \text{Span}\{u_1, \cdots, u_{n-1}\}$ so we have

$$\text{Span}\{v_1, \cdots, v_{n-1}, v_n\} = \text{Span}\{v_1, \cdots, v_{n-1}, u_n\} = \text{Span}\{u_1, \cdots, u_{n-1}, u_n\}.$$

It remains to show that the set $\{v_1, v_2, \cdots, v_n\}$ is an orthogonal set. By the induction hypothesis, we have $\langle v_j, v_k\rangle = 0$ for all $1 \leq j, k < n$, so it suffices to show that $\langle v_n, v_k\rangle = 0$ for all indices $1 \leq k < n$ and indeed, for $1 \leq k < n$ we have

$$\begin{aligned}\langle v_n, v_k\rangle &= \left\langle u_n - \sum_{j=1}^{n-1} \tfrac{\langle u_n, v_j\rangle}{||v_j||^2}\, v_j \,,\, v_k \right\rangle = \langle u_n, v_k\rangle - \sum_{j=1}^{n-1} \tfrac{\langle u_n, v_j\rangle}{||v_j||^2}\, \langle v_j, v_k\rangle \\ &= \langle u_n, v_k\rangle - \frac{\langle u_n, v_k\rangle}{||v_k||^2}\, \langle v_k, v_k\rangle = 0.\end{aligned}$$

4.4 Corollary: *Every finite or countable dimensional inner product space W over $F = \mathbf{R}$ or $\mathbf{C}$ has an orthonormal basis.*

Proof: The proof is left as an exercise.

4.5 Remark: It is not the case that every uncountable dimensional inner product space has an orthonormal basis. For example, we shall see below that an infinite dimensional separable Hilbert space does not have an orthonormal basis.

4.6 Corollary: *Let W be a finite or countable dimensional inner product space over $F = \mathbf{R}$ or $\mathbf{C}$. Let $U \subseteq W$ be a finite dimensional subspace. Then every orthogonal (or orthonormal) basis $\mathcal{A}$ for U extends to an orthogonal (or orthonormal) basis for W.*

Proof: The proof is left as an exercise.

4.7 Remark: The above corollary does not hold in general in the case that the subspace U is countable dimensional, as we shall soon see in Example 4.18.

4.8 Corollary: *Let $F = \mathbf{R}$ or $\mathbf{C}$ and let U and V be finite or countable dimensional inner product spaces over F. Then U and V are isomorphic (as inner product spaces) if and only if $\dim(U) = \dim(V)$. In particular, if $\dim(U) = n$ then U is isomorphic to F^n and if $\dim(U) = \aleph_0$ then U is isomorphic to F^∞.*

Proof: The proof is left as an exercise.

4.9 Definition: Let W be an inner product space over $F = \mathbf{R}$ or $\mathbf{C}$. For a subspace $U \subseteq W$, we define the **orthogonal complement** of U in W to be the set

$$U^\perp = \left\{ x \in W \middle| \langle x, u \rangle = 0 \text{ for all } u \in U \right\}.$$

4.10 Theorem: *Let W be an inner product space over $F = \mathbf{R}$ or $\mathbf{C}$. Let $U \subseteq W$ be a subspace. Then*
(1) $U^\perp$ is a subspace of W,
(2) if $\mathcal{A}$ is a basis for U then $U^\perp = \left\{ x \in W \middle| \langle x, u \rangle = 0 \text{ for all } u \in U \right\}$,
(3) $U \cap U^\perp = \{0\}$, and
(4) $U \subseteq (U^\perp)^\perp$.
If U is finite dimensional, then we also have
(5) $U \oplus U^\perp = W$, and
(6) $U = (U^\perp)^\perp$.

Proof: We leave the proofs of Parts (1) to (4) as an exercise. To prove Parts (5) and (6), suppose that U is finite-dimensional. Let $\mathcal{A} = \{u_1, u_2, \cdots, u_n\}$ be an orthonormal basis for $\mathcal{A}$. To prove Part (5), we need to show that for every $x \in W$ there exist unique vectors $u, v \in W$ with $u \in U$, $v \in U^\perp$ and $u + v = x$. First we prove uniqueness. Let $x \in W$, and suppose that $u \in U$, $v \in U^\perp$ and $u + v = x$. Note that for all indices k we have

$$\langle x, u_k \rangle = \langle u + v, u_k \rangle = \langle u, u_k \rangle + \langle v, u_k \rangle = \langle u, u_k \rangle.$$

and so, by Theorem 4.2, we have

$$u = \sum_{k=1}^{n} \langle u, u_k \rangle u_k = \sum_{k=1}^{n} \langle x, u_k \rangle u_k.$$

This proves uniqueness, since given $x \in W$, the vector u must be given by $u = \sum_{k=1}^{n} \langle x, u_k \rangle u_k$ and then the vector v must be given by $v = x - u$.

To prove existence, let $x \in W$ and choose u and v to be the vectors $u = \sum_{k=1}^{n} \langle x, u_k \rangle u_k$ and $v = x - u$. Then we have $u \in U$ and $u + v = x$, so it suffices to show that $v \in U^\perp$. For all indices k we have

$$\begin{aligned}\langle v, u_k \rangle = \langle x - u, u_k \rangle = \langle x, u_k \rangle - \langle u, u_k \rangle &= \langle x, u_k \rangle - \Big\langle \sum_{j=1}^{n} \langle x, u_j \rangle u_j \,,\, u_k \Big\rangle \\ &= \langle x, u_k \rangle - \sum_{j=1}^{n} \langle x, u_j \rangle \langle u_j, u_k \rangle = \langle x, u_k \rangle - \sum_{j=1}^{n} \langle x, u_j \rangle \delta_{j,k} = \langle x, u_k \rangle - \langle x, u_k \rangle = 0.\end{aligned}$$

Since $\langle v, u_k \rangle = 0$ for all $1 \leq k \leq n$, from Part (2) we have $v \in U^\perp$. This proves Part (5).

Let us prove Part (6). From Part (4), we have $U \subseteq (U^\perp)^\perp$. Conversely, let $x \in (U^\perp)^\perp$. Using Part (5), we can choose $u, v \in W$ with $u \in U$, $v \in V$ and $u + v = x$. Since $x \in (U^\perp)^\perp$ and $v \in U^\perp$, we have $\langle x, v \rangle = 0$, and so $0 = \langle x, v \rangle = \langle u + v, v \rangle = \langle u, v \rangle + \langle v, v \rangle = \langle v, v \rangle$. Since $\langle v, v \rangle = 0$ we have $v = 0$ and so $x = u + v = u \in U$. Thus $(U^\perp)^\perp \subseteq U$, as required.

4.11 Remark: Parts (5) and (6) of the above theorem do not always hold when U is infinite dimensional, as the following example shows.

4.12 Example: Let $F = \mathbf{R}$ or $\mathbf{C}$. Let $W = F^\infty$ and let $U = \big\{ a \in F^\infty \big| \sum_{k=1}^{\infty} a_k = 0 \big\}$. Note that W is a countable-dimensional inner product space with standard basis $\{e_1, e_2, e_3, \cdots\}$ and U is a countable-dimensional proper subspace of W with basis $\mathcal{A} = \{u_1, u_2, u_3, \cdots\}$ where $u_k = e_1 - e_{k+1} = (1, 0, \cdots, 0, -1, 0, \cdots)$. We have

$$\begin{aligned} U^\perp &= \big\{ x \in W \big| \langle x, u_k \rangle = 0 \text{ for all } k \big\} = \big\{ x \in W \big| \langle x, e_1 - e_{k+1} \rangle = 0 \text{ for all } k \big\} \\ &= \big\{ x \in W \big| x_1 = x_{k+1} \text{ for all } k \big\} = \big\{ x \in W \big| x_1 = x_2 = x_3 = \cdots \big\} = \{0\} \end{aligned}$$

because for $x \in F^\infty$ we have $x_n = 0$ for all but finitely many indices n. Notice that in this example we have $U \subsetneqq U^T = W$ and we do not have $U \oplus U^\perp = W$. Also notice that, although we could apply the Gram-Schmidt Procedure to the basis $\mathcal{A}$ to obtain an orthogonal basis $\mathcal{B} = \{v_1, v_2, \cdots\}$ for U, the basis $\mathcal{B}$ cannot be extended to an orthogonal basis for W because there is no nonzero vector $0 \neq x \in W$ with $\langle x, v_k \rangle$ for all k.

4.13 Definition: Let W be an inner product space over $F = \mathbf{R}$ or $\mathbf{C}$. Let $U \subseteq W$ be a subspace such that $W = U \oplus U^\perp$. For $x \in W$, we define the **orthogonal projection** of x onto U, denoted by $\text{Proj}_U(x)$, as follows. Since $W = U \oplus U^\perp$, we can choose unique vectors $u, v \in W$ with $u \in U$, $v \in V$ and $u + v = x$. We then define

$$\text{Proj}_U(x) = u.$$

Since $U = (U^\perp)^\perp$, for u and v as above we have $\text{Proj}_{U^\perp}(x) = v$. When $y \in W$ and $U = \text{Span}\{y\}$, we also write $\text{Proj}_y(x) = \text{Proj}_U(x)$.

4.14 Note: Let W be an inner product space over $F = \mathbf{R}$ or $\mathbf{C}$. Let U be a finite dimensional subspace of W. Let $\mathcal{A} = \{u_1, u_2, \cdots, u_n\}$ be an orthogonal basis for U. Then for $x \in W$, as in the proof of Part (5) of Theorem 4.15, we see that

$$\text{Proj}_U(x) = \sum_{k=1}^{n} \frac{\langle x, u_1 \rangle}{||u_k||^2} \, u_k.$$

4.15 Example: As an exercise, show that for $A \in M_{n\times m}(\mathbf{C})$ and $U = \mathrm{Col}(A)$, given $x \in \mathbf{C}^n$ there exists $y \in \mathbf{C}^m$ such that $A^*Ay = A^*x$ and that for any such y we have $\mathrm{Proj}_U(x) = Ay$. In particular, when $\mathrm{rank}(A) = m$ show that A^*A is invertible so that $\mathrm{Proj}_U(x) = A(A^*A)^{-1}A^*x$.

4.16 Theorem: *Let W be an inner product space over $F = \mathbf{R}$ or $\mathbf{C}$. Let $U \subseteq W$ be a subspace of W such that $W = U \oplus U^\perp$. Let $x \in W$. Then $\mathrm{Proj}_U(x)$ is the unique point in U which is nearest to x.*

Proof: Let $u, v \in W$ be the vectors with $u \in U$, $v \in V$ and $u + v = x$, so that we have $\mathrm{Proj}_U(x) = u$. Let $w \in U$ with $w \neq u$. Since $\langle w-u, x-u\rangle = \langle w-u, v\rangle = \langle w, v\rangle - \langle u, v\rangle = 0$, Pythagoras' Theorem gives

$$||x - w||^2 = ||(x - u) - (w - u)||^2 = ||x - u||^2 + ||w - u||^2 > ||x - u||^2$$

and so $||x - w|| > ||x - u||$.

4.17 Definition: Let W be a vector space over $F = \mathbf{R}$ or $\mathbf{C}$. For a subset $S \subseteq W$, we say that S is **convex** when for all $a, b \in S$ we have $a + t(b - a) \in S$ for all $0 \leq t \leq 1$.

4.18 Theorem: *Let H be a Hilbert space over $F = \mathbf{R}$ or $\mathbf{C}$. Let $S \subseteq H$ be nonempty, closed and convex. Then for every $a \in H$ there exists a unique point $b \in S$ which is nearest to a, that is such that $||a - b|| \leq ||a - x||$ for all $x \in S$.*

Proof: Let $a \in H$. Let $d = \mathrm{dist}(a, S) = \inf\big\{||x - a||\,\big|\,x \in S\big\}$. Choose a sequence $\{x_n\}$ in S so that $||x_n - a|| \to d$, hence $||x_n - a||^2 \to d^2$. Let $\epsilon > 0$ and choose $m \in \mathbf{Z}^+$ so that for all $n \geq m$ we have $||x_n - a||^2 \leq d^2 + \frac{\epsilon^2}{4}$. Let $k, l \geq m$. By the Parallelogram Law we have

$$\big\|(x_k - a) + (x_l - a)\big\|^2 + \big\|(x_k - a) - (x_l - a)\big\|^2 = 2\big\|x_k - a\big\|^2 + 2\big\|x_l - a\big\|^2$$

Since S is convex, we have $\frac{x_k + x_l}{2} \in S$, hence $\big\|\frac{x_k+x_l}{2} - a\big\| \geq d$, and so

$$\begin{aligned}\big\|x_k - x_l\big\|^2 &= \big\|(x_k - a) - (x_k - a)\big\|^2\\ &= 2\big\|x_k - a\big\|^2 + 2\big\|x_l - a\big\|^2 - \big\|(x_k - a) + (x_k - a)\big\|^2\\ &= 2\big\|x_k - a\big\|^2 + 2\big\|x_l - a\big\|^2 - 4\big\|\tfrac{x_k+x_l}{2} - a\big\|^2\\ &\leq 2\big(d^2 + \tfrac{\epsilon^2}{4}\big) + 2\big(d^2 + \tfrac{\epsilon^2}{4}\big) - 4d^2 = \epsilon^2.\end{aligned}$$

so that $||x_k - x_l|| \leq \epsilon$. This shows that the sequence $\{x_n\}$ is Cauchy. Since H is complete, $\{x_n\}$ converges in H, and since S is closed in H, the limit lies in S. Let $b = \lim\limits_{n\to\infty} x_n \in S$. Since $b \in S$ we have $||d - a|| \geq d$, and we have $||b - a|| \leq ||b - x_n|| + ||x_n - a||$ for all $n \in \mathbf{Z}^+$ so that $||b - a|| \leq \lim\limits_{n\to\infty}\big(||b - x_n|| + ||x_n - a||\big) = d$, and so $||b - a|| = d$. This shows that $||d - a|| \geq ||x - a||$ for all $x \in S$. Finally, we note that the point b is unique because given $c \in S$ with $||c - a|| = d$, since S is convex we have $\frac{b+c}{2} \in S$ so that $\big\|\frac{b+c}{2} - a\big\| \geq d$, and so the Parallelogram Law gives

$$\begin{aligned}||b - c||^2 &= \big\|(b - a) - (c - a)\big\|^2 = 2||b - a||^2 + 2||c - a||^2 - \big\|(b - a) + (c - a)\big\|\\ &= 4d^2 - 4\big\|\tfrac{b-c}{2} - a\big\|^2 \leq 4d^2 - 4d^2 = 0\end{aligned}$$

so that $||b - c|| = 0$ hence $b = c$.

4.19 Theorem: *Let H be a Hilbert space over $F = \mathbf{R}$ or $\mathbf{C}$. Let $U \subseteq H$ be a closed subspace. Then we have $H = U \oplus U^\perp$. This means that for all $x \in H$ there exist unique points $u \in U$ and $v \in U^\perp$ such that $u + v = x$. In this case, the point u is the unique point in U nearest to x.*

Proof: Let $x \in H$. Since U is a vector space it is convex, so by the previous theorem there is a unique point $u \in U$ which is nearest to x. Let u be this nearest point and let $v = x - u$ so that $u + v = x$. We claim that $v \in U^\perp$. Suppose, for a contradiction, that $v \notin U^\perp$. Choose $u_1 \in U$ with $\langle v, u_1 \rangle \neq 0$. Write $\langle v, u_1 \rangle = re^{i\theta}$ with $r > 0$ and $\theta \in \mathbf{R}$ (when $F = \mathbf{R}$ we have $e^{i\theta} = \pm 1$) and let $u_2 = e^{i\theta} u_1$. Note that $u_2 \in U$ and $\langle v, u_2 \rangle = \langle v, e^{i\theta} u_1 \rangle = e^{-i\theta} \langle v, u_1 \rangle = e^{-i\theta} r\, e^{i\theta} = r > 0$. For all $t \in \mathbf{R}$ we have

$$\big\|x - (u + tu_2)\big\|^2 = ||v - tu_2||^2 = ||v||^2 - 2t\,\mathrm{Re}\langle v, u_2 \rangle + t^2 ||u_2||^2 = ||v||^2 - 2r\,t + ||u_2||^2 t^2.$$

It follows that for small $t > 0$ we have $\big\|x - (u + tu_2)\big\|^2 \leq ||v||^2 = ||x - u||^2$ which is not possible, since u is the point in U which is nearest to x.

It remains to show that the points $u \in U$ and $v \in U^\perp$ with $u + v = x$, which we found in the previous paragraph, are the only such points. Let $x \in H$. Suppose that $u \in U$, $v \in U^\perp$ and $u + v = x$. We claim that u must be equal to the (unique) point in U which is nearest to x. Let $u' \in U$ with $u' \neq u$. Since $v \in U^\perp$ and $u' - u \in U$ we have $\langle x - u, u' - u \rangle = \langle v, u' - u \rangle = 0$ and so

$$\begin{aligned} \big\|x - u'\big\|^2 = \big\|(x - u) - (u' - u)\big\|^2 &= ||x - u||^2 - 2\,\mathrm{Re}\langle x - u, u' - u > + ||u' - u||^2 \\ &= ||x - u|| + ||u' - u|| > ||x - u||^2 \end{aligned}$$

so that $||x - u'|| > ||x - u||$. Thus u is the point in U which is nearest to x, as required.

4.20 Theorem: *Every inner product space contains a maximal orthonormal set.*

Proof: Let W be an inner product space. Let S be the set of all orthonormal sets in W, ordered by inclusion. If C is a chain in S (that is a totally ordered subset of S) then $\bigcup C$ is an upper bound for C in S. Since every chain in S has an upper bound, it follows from Zorn's Lemma that S has a maximal element.

4.21 Theorem: *Let H be a Hilbert space over $F = \mathbf{R}$ or $\mathbf{C}$. Let $\mathcal{A}$ be an orthonormal set in H and let $U = \mathrm{Span}_F \mathcal{A}$. Then $\mathcal{A}$ is maximal if and only if U is dense in H.*

Proof: If $\mathcal{A}$ is not maximal then we can choose $v \in U^\perp$ with $||v|| = 1$ (so that $\mathcal{A} \cup \{v\}$ is orthonormal) and then for all $u \in U$, since $\langle v, u \rangle = 0$, we have $||u - v||^2 = ||u||^2 + ||v||^2 \geq ||v||^2 = 1$. Thus U is not dense in H, indeed there is no $u \in U$ with $||u - v|| \leq \frac{1}{2}$.

Suppose, conversely, that U is not dense in H, that is $\overline{U} \neq H$. Note that $\overline{U}$ is a vector space, indeed given $a, b \in \overline{U}$ we can choose $\{x_n\}$ and $\{y_n\}$ with $x_n \to a$ and $y_n \to b$ in H and then $(x_n + y_n) \to (a + b)$ so that $a + b \in \overline{U}$, and for $c \in F$ we have $cx_n \to ca$ so that $ca \in \overline{U}$. By the above theorem, we have $H = \overline{U} \oplus \overline{U}^\perp$. Since $H \neq \overline{U}$ we must have $\overline{U}^\perp \neq \{0\}$. Choose $v \in \overline{U}^\perp$ with $||v|| = 1$. Since $\langle v, u \rangle = 0$ for all $u \in \overline{U}$ we certainly have $\langle v, u \rangle = 0$ for all $u \in U$, so the set $\mathcal{A} \cup \{v\}$ is orthonormal. And we cannot have $v \in U$ since $U \cap U^\perp = \{0\}$, and so $\mathcal{A} \subsetneqq \mathcal{A} \cup \{v\}$ so that $\mathcal{A}$ is not maximal.

4.22 Theorem: *Let H be a Hilbert space over $F = \mathbf{R}$ or $\mathbf{C}$. Let $\mathcal{A}$ be a maximal orthonormal set in H. Then H is separable if and only if $\mathcal{A}$ is at most countable.*

Proof: Suppose that $\mathcal{A}$ is uncountable. Let S be any dense subset of H. For each $u \in \mathcal{A}$ choose $s_u \in S$ with $||s_u - u|| \leq \frac{\sqrt{2}}{4}$. For $u, v \in \mathcal{A}$ with $u \neq v$ we have $||u|| = 1$ and $||v|| = 1$ and $\langle u, v\rangle = 0$ so that $||u - v||^2 = ||u||^2 + ||v||^2 = 2$ and so

$$||s_u - s_v|| = \big\|(s_u - u) + (u - v) + (v - s_v)\big\| \geq ||u - v|| - \big(||s_u - u|| + ||s_v - v||\big) = \sqrt{2} - \tfrac{\sqrt{2}}{4} - \tfrac{\sqrt{2}}{4} > 0$$

so that $s_u \neq s_v$. Thus $\mathcal{A}$ is uncountable.

Suppose, conversely, that $\mathcal{A} = \{u_1, u_2, \cdots\}$ is finite or countable. By the above theorem, $U = \text{Span}_F\,\mathcal{A}$ is dense in H. Note that $\text{Span}_{\mathbf{Q}}\mathcal{A}$ is dense in $\text{Span}_{\mathbf{R}}\mathcal{A}$ and $\text{Span}_{\mathbf{Q}[i]}\mathcal{A}$ is dense in $\text{Span}_{\mathbf{C}}\mathcal{A}$. Indeed given $c_1, c_2, \cdots, c_n \in F$ (where $F = \mathbf{R}$ or $\mathbf{C}$) we can choose $r_1, r_2, \cdots, r_n \in R$ (where $R = \mathbf{Q}$ or $\mathbf{Q}[i]$) such that $|r_k - c_k| < \frac{\epsilon}{n}$ and then

$$\begin{aligned}\Big\|\sum_{k=1}^n r_k u_k - \sum_{k=1}^n c_k u_k\Big\| &= \Big\|\sum_{k=1}^n (r_k - c_k)u_k\Big\| \leq \sum_{k=1}^n \big\|(r_k - c_k)u_k\big\| \\ &= \sum_{k=1}^n |r_k - c_k|\,||u_k|| = \sum_{k=1}^n |r_k - c_k| < \epsilon.\end{aligned}$$

4.23 Theorem: *Let H be a separable Hilbert space over $F = \mathbf{R}$ or $\mathbf{C}$, let $\mathcal{A} = \{u_1, u_2, \cdots\}$ be a countable orthonormal set in H, and let $U = \text{Span}_F\,\mathcal{A}$. Then the following are equivalent.*

(1) $\mathcal{A}$ is maximal.
(2) U is dense in H.
(3) For every $x \in H$ we have $x = \sum_{k=1}^\infty \langle x, u_k\rangle u_k$ in H.
(4) For every $x \in H$ we have $||x||^2 = \sum_{k=1}^\infty \big|\langle x, u_k\rangle\big|^2$ in $\mathbf{R}$.
(5) For all $x, y \in H$ we have $\langle x, y\rangle = \sum_{k=1}^\infty \langle x, u_k\rangle\overline{\langle y, u_k\rangle}$.

Proof: The equivalence of Parts (1) and (2) follows from Theorem 4.26. Let us prove that (2) implies (3). Suppose that U is dense in H. Let $x \in H$. For each $n \in \mathbf{Z}^+$, let $U_n = \text{Span}\{u_1, u_2, \cdots, u_n\}$ and let $w_n = \text{Proj}_{U_n}(x) = \sum_{k=1}^n \langle x, u_k\rangle u_k$. Let $\epsilon > 0$. Since U is dense in H we can choose $u \in U$ with $||u - x|| < \epsilon$. Say $u = \sum_{k=1}^m c_k u_k$. For all $n \geq m$, since $u \in U_n$ and w_n is the point in U_n nearest to x we have $||w_n - x|| \leq ||u - x|| < \epsilon$. Thus $\lim_{n\to\infty} w_n = x$ in H. This means that $x = \sum_{k=1}^\infty \langle x, u_k\rangle u_k$ in H.

Let us prove that (3) implies (4). Suppose that for every $x \in H$ we have $x = \lim_{n\to\infty} w_n$ where $w_n = \sum_{k=1}^n \langle x, u_k\rangle u_k$. Note that

$$||w_n||^2 = \Big\langle \sum_{k=1}^n \langle x, u_k\rangle u_k\,,\ \sum_{\ell=1}^n \langle x, u_\ell\rangle u_\ell\Big\rangle = \sum_{k=1}^n\sum_{\ell=1}^n \langle x, u_k\rangle\overline{\langle x, u_\ell\rangle}\delta_{k,\ell} = \sum_{k=1}^n \big|\langle x, u_k\rangle\big|^2.$$

Let $\epsilon > 0$. Choose $m \in \mathbf{Z}^+$ such that for all $n \geq m$ we have $||w_n - x|| < \epsilon$. By the Triangle Inequality, for all $n \geq m$ we have $\Big| ||w_n|| - ||x|| \Big| \leq ||w_n - x|| < \epsilon$. This shows that $\lim\limits_{n\to\infty} ||w_n|| = ||x||$ in $\mathbf{R}$, hence $||x||^2 = \lim\limits_{n\to\infty} ||w_n||^2 = \sum\limits_{k=1}^{\infty} \big|\langle x, u_k\rangle\big|^2$ in $\mathbf{R}$.

Next we prove that (4) implies (5). Suppose that (4) holds. Let $x, y \in H$. Let $x_n = \sum\limits_{k=1}^{n} \langle x, u_k\rangle u_k$ and $y_n = \sum\limits_{k=1}^{n} \langle y, u_k\rangle u_k$. Note that

$$\langle x_n, y_n\rangle = \Big\langle \sum_{k=1}^{n} \langle x, u_k\rangle u_k\,,\ \sum_{\ell=1}^{n} \langle y, u_k\rangle u_k \Big\rangle = \sum_{k=1}^{n}\sum_{\ell=1}^{n} \langle x, u_k\rangle \overline{\langle y, u_\ell\rangle}\delta_{k,\ell} = \sum_{k=1}^{n} \langle x, u_k\rangle \overline{\langle y, u_k\rangle}$$

and note that for $c \in \mathbf{C}$ we have $x_n + cy_n = \sum\limits_{k=1}^{n} \langle x + cy, u_k\rangle u_k$. Since (4) holds, we have $\lim\limits_{n\to\infty} ||x_n||^2 = ||x||^2$, $\lim\limits_{n\to\infty} ||y_n||^2 = ||y||^2$, and $\lim\limits_{n\to\infty} ||x_n + cy_n||^2 = ||x + cy||^2$. By the Polarization Identity,

$$\begin{aligned}\langle x, y\rangle &= \tfrac{1}{4}\Big(||x+y||^2 + i\,||x+iy||^2 - ||x-y||^2 - i\,||x-iy||^2 \Big) \\ &= \lim_{n\to\infty} \tfrac{1}{4}\Big(||x_n+y_n||^2 + i\,||x_n+iy_n||^2 - ||x_n-y_n||^2 - i\,||x_n-iy_n||^2 \Big) \\ &= \lim_{n\to\infty} \langle x_n, y_n\rangle = \sum_{k=1}^{\infty} \langle x, u_k\rangle \overline{\langle y, u_k\rangle}.\end{aligned}$$

Note that (4) follows immediately from (5) by taking $y = x$. We finish the proof by proving that (4) implies (2). Suppose that for all $x \in H$ we have $||x||^2 = \sum\limits_{k=1}^{\infty} \big|\langle x, u_k\rangle\big|^2$. As above, let $w_n = \text{Proj}_{U_n} = \sum\limits_{k=1}^{n} \langle x, u_k\rangle u_k$ so that $||w_n||^2 = \sum\limits_{k=1}^{n} \big|\langle x, u_k\rangle\big|^2$. Then we have $\lim\limits_{n\to\infty} ||w_n||^2 = ||x||^2$. Given $\epsilon > 0$, choose $n \in \mathbf{Z}^+$ so that $||x||^2 - ||w_n||^2 < \epsilon^2$. Since $w_n = \text{Proj}_{U_n}(x)$ we have $w_n \in U_n$ and $x - w_n \in {U_n}^\perp$ so that $\langle x - w_n, w_n\rangle = 0$. By Pythagoras' Theorem, $||x - w_n||^2 = ||x||^2 - ||w_n||^2 < \epsilon^2$, hence $||x = w_n|| < \epsilon$. Since $\epsilon > 0$ was arbitrary and $w_n \in U$, this shows that U is dense in H.

4.24 Definition: A maximal orthonormal set in a Hilbert space H (over $\mathbf{R}$ or $\mathbf{C}$) is called a **Hilbert basis** for H (over $\mathbf{R}$ or $\mathbf{C}$).

4.25 Theorem: *Let H be an infinite dimensional separable Hilbert space over F, where $F = \mathbf{R}$ or $\mathbf{C}$, and let $\mathcal{A} = \{u_1, u_2, u_3, \cdots\}$ be a countable Hilbert basis for H.*

(1) For all $x \in H$, if $x = \sum\limits_{k=1}^{\infty} a_k u_k$ and $x = \sum\limits_{k=1}^{\infty} b_k u_k$ then $a_k = b_k = \langle x, u_k\rangle$.

(2) For all $c_k \in F$, $\sum\limits_{k=1}^{\infty} c_k u_k$ converges in H if and only if $\sum\limits_{k=1}^{\infty} |c_k|^2$ converges in $\mathbf{R}$.

(3) The map $\phi : H \to \ell_2(F)$ given by $\phi(x) = \big(\langle x, u_1\rangle, \langle x, u_2\rangle, \cdots\big)$ is an inner product space isomorphism.

Proof: The proof is left as an exercise.

4.26 Definition: When X and Y are normed linear spaces over $F = \mathbf{R}$ or $\mathbf{C}$, a linear map $L : X \to Y$ is also called a **linear operator**, and a linear map $L : X \to F$ is also called a **linear functional** on X.

4.27 Definition: Let X and Y be normed linear spaces and let $L : X \to Y$ be a linear operator. The **operator norm** of L is given by

$$||L|| = \sup\left\{||Lx||\,\Big|\, x \in X \text{ with } ||x|| \leq 1\right\}$$

and we say that L is **bounded** when $||L| < \infty$. Since $Lx = ||x||\, L\big(\frac{x}{||x||}\big)$ for all $0 \neq x \in X$, it follows that

$$||Lx|| \leq ||L||\,||x|| \text{ for all } x \in X.$$

4.28 Theorem: *Let X and Y be normed linear spaces and let $L : X \to Y$ be a linear operator. Then the following are equivalent.*
(1) L is continuous at 0.
(2) L is bounded.
(3) L is uniformly continuous in X.

Proof: Suppose that L is continuous at 0. Choose $\delta > 0$ so that $||x|| \leq \delta \Longrightarrow ||Lx|| \leq 1$. Then $||x|| \leq 1 \Longrightarrow ||\delta x|| \leq \delta \Longrightarrow ||L(\delta x)|| \leq 1 \Longrightarrow ||L(x)|| = \frac{1}{\delta}\,||L(\delta x)|| \leq \frac{1}{\delta}$ so $||L|| \leq \frac{1}{\delta}$.

Now suppose that L is bounded. F$x, y \in X$ we have

$$\big\|Lx - Ly\big\| = \big\|L(x-y)\big\| = \left\|L\Big(\tfrac{x-y}{||x-y||}\Big)\right\|||x-y|| \leq ||L||\,||y-x||.$$

Thus given $\epsilon > 0$ we can choose $\delta = \frac{1}{||L||}$ and then $||x-y|| < \delta \Longrightarrow ||Lx - Ly|| < \epsilon$.

Finally, we note that if L is uniformly continuous in X then L is continuous at 0.

4.29 Theorem: *(The Uniform Boundedness Principle) Let X be a Banach space and let Y be a normed linear space. Let S be a set of bounded linear operators $L : X \to Y$. Suppose that for every $x \in X$ there exists $m_x \geq 0$ such that $||Lx|| \leq m_x$ for all $L \in S$. Then there exists $m \geq 0$ such that $||L|| \leq m$ for all $L \in S$.*

Proof: For each $n \in \mathbf{Z}^+$, let $A_n = \big\{x \in X \,\big|\, ||Lx|| \leq n \text{ for all } L \in S\big\}$. Note that A_n is closed because the sets $\big\{x \in X \,\big|\, ||Lx|| \leq n\big\}$ are closed for each $L \in S$, and A_n is the intersection of these sets. By the hypothesis of the theorem, we have $X = \bigcup_{n=1}^{\infty} A_n$. By the Baire Category Theorem (since X is complete), the sets A_n cannot all be nowhere dense. Choose $n \in \mathbf{Z}^+$ so that A_n is not nowhere dense. Chose $a \in A_n$ and $r > 0$ so that $\overline{B}(a,r) \subseteq A_n$. For all $x \in X$, if $x \in B(a,r)$ then $x \in A_n$ so we have $||L(x)|| \leq n$ for all $L \in S$. If $||x|| < r$ then $x + a \in B(a,r)$ and $a \in B(a,r)$ and so

$$||L(x)|| = ||L(x+a) - L(a)|| \leq ||L(x+a)|| + ||L(a)|| \leq 2n \text{ for all } L \in S.$$

For all $L \in S$ and $x \in X$, if $||x|| \leq 1$ then $||rx|| \leq r$ and so $||L(x)|| = \frac{1}{r}\,||L(rx)|| \leq \frac{2n}{r}$. Thus we have $||L|| \leq \frac{2n}{r}$ for all $L \in S$.

4.30 Theorem: *(Condensation of Singularities) Let X be a Banach space and let Y be a normed linear space. For each $m, n \in \mathbf{Z}^+$, let $L_{m,n} : X \to Y$ be a bounded linear operator. Suppose that for each $m \in \mathbf{Z}^+$ there exists $x_m \in X$ such that $\limsup_{n\to\infty} ||L_{mn}(x_m)|| = \infty$. Then the set $E = \left\{ x \in X \middle| \limsup_{n\to\infty} ||L_{mn}(x)|| = \infty \text{ for all } m \in \mathbf{Z}^+ \right\}$ is a dense $\mathcal{G}_\delta$ set.*

Proof: Fix $m \in \mathbf{Z}^+$. For each $\ell \in \mathbf{Z}^+$, let $A_\ell = \left\{ x \in X \,\middle|\, ||L_{n,m}(x)|| \le \ell \text{ for all } n \in \mathbf{Z}^+ \right\}$ and note that each set A_ℓ is closed. As in the proof of the Uniform Boundedness Principle, if one of the sets A_ℓ was not nowhere dense then we could choose $m \ge 0$ such that $||L_{m,n}|| \le m$ for all $n \in \mathbf{Z}^+$. But then for all $x \in X$ we would have $||L_{m,n}(x)|| \le m||x||$ for all n so that $\limsup_{n\to\infty} ||L_{m,n}(x)|| \le m||x||$, contradicting the hypothesis of the theorem. Thus all of the sets A_ℓ must be nowhere dense. Let $B_m = \bigcup_{\ell=1}^{\infty} A_\ell = \left\{ x \in X \,\middle|\, \limsup_{n\to\infty} ||L_{m,n}(x)|| < \infty \right\}$ and let $C = \bigcup_{m=1}^{\infty} B_m = \left\{ x \in X \,\middle|\, \limsup_{n\to\infty} ||L_{m,n}(x)|| < \infty \text{ for some } m \in \mathbf{Z}^+ \right\}$, and note that $E = X \setminus C$. Then C is a countable union of closed nowhere dense sets, so E is a countable intersection of open dense sets. By the Baire Category Theorem, E is dense.

.

Chapter 5. Fourier Series

5.1 Remark: We shall begin with an informal discussion of Fourier series and how they can be used in physics and engineering.

5.2 Definition: A real **trigonometric series** is a series of the form

$$a_0 + \sum_{n=1}^{\infty} a_n \cos nx + \sum_{n=1}^{\infty} b_n \sin nx$$

where $a_n, b_n \in \mathbf{R}$ and $x \in \mathbf{R}$. If the series converges, we say it is the real **Fourier series** of its sum

$$f(x) = a_0 + \sum_{n=1}^{\infty} a_n \cos nx + \sum_{n=1}^{\infty} b_n \sin nx,$$

which is a periodic function of the real variable x with period 2π, and the numbers a_n, b_n are called the **Fourier coefficients** of $f(x)$. If we are justified in integrating term by term then, using the formulas

$$\int_{-\pi}^{\pi} 1 = 2\pi \ , \ \int_{-\pi}^{\pi} \cos nx \ dx = 0 = \int_{-\pi}^{\pi} \sin nx \ dx \ , \ \int_{-\pi}^{\pi} \cos^2 nx \ dx = \pi = \int_{-\pi}^{\pi} \sin^2 nx \ dx$$

$$\int_{-\pi}^{\pi} \cos nx \ \cos mx \ dt = 0 = \int_{-\pi}^{\pi} \sin nx \ \sin mx \ dt \ , \text{ and } \int_{-\pi}^{\pi} \cos nx \ \sin mx \ dt = 0$$

where $n, n \in \mathbf{Z}^+$ with $n \neq m$, we find that the Fourier coefficients are given by

$$a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\, dx \ , \ a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \ dx \ , \text{ and } b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx \ dx.$$

5.3 Remark: For the moment, we shall blithely assume that, given a 2π-periodic function $f : \mathbf{R} \to \mathbf{R}$, the Fourier series with coefficients a_n and b_n given by the above formulas converges to the given function $f(x)$.

5.4 Example: Find the Fourier coefficients of the 2π-periodic function $f : \mathbf{R} \to \mathbf{R}$ with

$$f(x) = \begin{cases} \frac{\pi}{2} + x \text{ for } -\pi \leq x \leq 0, \\ \frac{\pi}{2} - x \text{ for } \ 0 \leq x \leq \pi. \end{cases}$$

Solution: Since $f(x)$ is even, we have $a_n = 0$ for all $n \in \mathbf{Z}^+$, and we have

$$\begin{aligned} a_0 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\, dx = \frac{2}{\pi} \int_0^{\pi} f(x)\, dx = \frac{2}{\pi} \int_0^{\pi} \tfrac{\pi}{2} - x \ dx = \tfrac{2}{\pi} \Big[\tfrac{\pi}{2} x - \tfrac{1}{2} x^2 \Big]_0^{\pi} = 0 \\ a_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx \ dx = \frac{2}{\pi} \int_0^{\pi} f(x) \cos nx \ dx = \frac{2}{\pi} \int_0^{\pi} \big(\tfrac{\pi}{2} - x \big) \cos nx \ dx \\ &= \int_0^{\pi} \cos nx \ dx - \frac{2}{\pi} \int_0^{\pi} x \cos nx \ dx = \Big[\tfrac{1}{n} \sin nx \Big]_0^{\pi} - \tfrac{2}{\pi} \Big[\tfrac{1}{n} x \sin nx + \tfrac{1}{n^2} \cos nx \Big]_0^{\pi} \\ &= 0 - \tfrac{2}{\pi} \big(\tfrac{1}{n^2}(-1)^n - 1 \big) = \begin{cases} 0 & \text{if } n \text{ is even,} \\ \frac{4}{\pi n^2} & \text{if } n \text{ is odd.} \end{cases} \end{aligned}$$

Thus, assuming that the Fourier series of $f(x)$ converges to $f(x)$, we have

$$f(x) = \tfrac{4}{\pi} \big(\tfrac{1}{1^2} \cos x + \tfrac{1}{3^2} \cos 3x + \tfrac{1}{5^2} \cos 5x + \cdots \big)$$

5.5 Remark: Assuming convergence, putting $x=0$ into the above function $f(x)$ gives $\frac{\pi}{2}=\frac{4}{\pi}\left(\frac{1}{1^2}+\frac{1}{3^2}+\frac{1}{5^2}+\cdots\right)$ so we obtain the formula

$$\sum_{k=1}^{\infty}\frac{1}{(2k+1)^2}=\frac{\pi^2}{8}.$$

5.6 Example: (Forced Damped Oscillations) Suppose an object of mass m is attached to a spring of spring-constant k and vibrates in a fluid of damping-constant c and let $x=x(t)$ be the displacement of the object from its equilibrium position at time t. Suppose, in addition, that the object is acted on by an external force $f(t)$. The total force $F(t)$ acting on the object consists of the force exerted by the spring, which is equal to $-kx(t)$, the resistive force exerted by the fluid, which is equal to $-cx'(t)$, and the external driving force, which is equal to $f(t)$. By Newton's Second Law of motion we have $F(t)=mx''(t)$ and so $x(t)$ satisfies the differential equation (the DE)

$$mx''(t)+cx'(t)+kx(t)=f(t).$$

5.7 Example: Use Fourier series to solve the above DE with $m=1$, $c=2$ and $k=5$, where $f(t)$ is the function from Example 5.4,

Solution: We need to solve the DE

$$x''(t)+2x'(t)+5x(t)=f(t).$$

To solve the associated homogeneous DE $x''+2x'+5x=0$, we look for a solution of the form $x=x(t)=e^{rt}$. Putting $x=e^{rt}$, $x'=re^{rt}$ and $x''=r^2e^{rt}$ into the homogeneous DE gives $(r^2+2r+5)e^{rt}=0$ hence $r=-1\pm 2i$. This gives us the two complex-valued solutions $x(t)=e^{(-1\pm 2i)t}=e^{-t}(\cos 2t\pm i\sin 2t)$. By taking suitable linear combinations of these two complex-valued solutions we obtain the two real-valued solutions $x_1(t)=e^{-t}\cos 2t$ and $x_2(t)=e^{-t}\sin 2t$. The general solution to the DE $x''+2x'+5x=0$ is given by

$$x(t)=Ae^{-t}\cos 2t+Be^{-t}\sin 2t\quad\text{, where } A,B\in\mathbf{R}.$$

For each $n\in\mathbf{Z}^+$, to find a particular solution to the DE $x''+2x'+5x=\cos nt$, we look for a solution of the form $x=x(t)=A_n\cos nt+B_n\sin nt$. Putting $x=A_n\cos nt+B_n\sin nt$, $x'=-nA_n\sin nx+nB_n\cos nt$ and $x''=-n^2A_n\cos nt-n^2B_n\sin nt$ into $x''+2x'+5=\cos nt$ gives $\big(-n^2A_n+2nB_n+5A_n\big)\cos nt+\big(-n^2B_n-2nA_n+5B_n\big)\sin nt=\cos nt$ for all $t\in\mathbf{R}$ and so we must have $(5-n^2)A_n+2nB_n=1$ and $(5-n^2)B_n-2nA_n=0$. We solve these two equations to get $A_n=\frac{5-n^2}{n^4-6n^2+25}$ and $B_n=\frac{2n}{n^4-6n^2+25}$ and so one solution to the DE $x''+2x'+5=\cos nt$ is given by

$$x(t)=A_n\cos nt+B_n\sin nx\ \text{, where } A_n=\tfrac{5-n^2}{n^4-6n^2+25}\text{ and } B_n=\tfrac{2n}{n^4-6n^2+25}.$$

Since $f(t)=\sum_{n\text{ odd}}\frac{4}{\pi n^2}\cos nt$, one particular solution, called the **steady state solution**, to the original DE $x''+2x'+5x=f(t)$ is given by

$$x(t)=\sum_{n\text{ odd}}\frac{4}{\pi n^2}\big(A_n\cos nt+B_n\sin nt\big)$$

and the general solution is

$$x(t)=Ae^{-t}\cos 2t+Be^{-t}\sin 2t+\sum_{n\text{ odd}}\frac{4}{\pi n^2}\big(A_n\cos nt+B_n\sin nt\big)\quad\text{, where } A,B\in\mathbf{R}.$$

5.8 Example: (The One-Dimensional Wave Equation) An elastic string is stretched to length π and is fixed at its two endpoints along the x-axis at $x = 0$ and $x = \pi$. The string is displaced so that it follows the curve $u = f(x)$ with $f(0) = 0$ and $f(\pi) = 0$, then at time $t = 0$ the string is released and allowed to vibrate. The problem is to determine the strings shape $u = u(x, t)$ at all points $0 \le x \le \pi$ and all times $t \ge 0$.

To formulate a differential equation (or DE) which models the situation, we consider a segment of string, at time t, between the points $p_1 = (x_1, u(x_1, t))$ and $p_2 = (x_2, u(x_2, t))$ where the difference $dx = x_2 - x_1$ is small. The slope of the curve $u = g(x) = u(x, t)$ at p_1 is $\frac{\partial u}{\partial x}(x_1, t)$ and the angle θ_1 from the horizontal is given by $\tan\theta_1 = \frac{\partial u}{\partial x}(x_1, t)$. Similarly, the angle θ_2 at p_2 is given by $\tan\theta_2 = \frac{\partial u}{\partial x}(x_2, t)$, and we have

$$\tan\theta_2 - \tan\theta_1 = \frac{\partial u}{\partial x_1}(x_1, t) - \frac{\partial u}{\partial x}(x_2, t) = \frac{\partial^2 u}{\partial x^2}\, dx.$$

Let T_1 be the magnitude of the force exerted on p_1 by the portion of the string which lies to the left of p_1, and let T_2 be the magnitude of the force exerted on p_2 by the portion of the string which lies to the right of p_2. Assuming that the segment of string moves only vertically (so the total horizontal component of the force is zero) we have $T_1\cos\theta_1 = T_2\cos\theta_2$. Let

$$T = T_1\cos\theta_1 = T_2\cos\theta_2$$

and note that T is a constant which we call the **tension** of the string. The total vertical component of the force is $F = T_2\sin\theta_2 - T_1\sin\theta_1$ and by Newton's Second Law of motion, we have

$$T_2\sin\theta_2 - T_2\sin\theta_1 = m\frac{\partial^2 u}{\partial t^2} = \rho\, dx\, \frac{\partial^2 u}{\partial t^2}$$

where ρ is the linear **density** of the string, that is its mass per unit length. From the equations $\tan\theta_2 - \tan\theta_1 = \frac{\partial^2 u}{\partial x^2}\, dx$, $T_1\cos\theta_1 = T_2\cos\theta_2$ and $T_2\sin\theta_2 - T_1\sin\theta_2 = \rho\, dx\, \frac{\partial^2 u}{\partial t^2}$ we obtain the **one-dimentional wave equation**

$$\frac{\partial^2 u}{\partial t^2} = c^2\frac{\partial^2 u}{\partial x^2}\ \text{, where } c^2 = \frac{T}{\rho}.$$

5.9 Example: Use Fourier series to solve the one-dimensional wave equation $\frac{\partial^2 u}{\partial t^2} = c^2\frac{\partial^2 u}{\partial x^2}$ subject to the boundary conditions $u(0, t) = 0$ and $u(0, \pi) = 0$ for all $t \ge 0$ and to the initial conditions $u(x, 0) = f(x)$ and $\frac{\partial u}{\partial t}(x, 0) = g(x)$ for all $0 \le x \le \pi$.

Solution: We use a method known as **separation of variables**. We look for a solution to the DE of the form $u(x, t) = y(x)s(t)$ which satisfies the given boundary conditions $0 = u(0, t) = y(0)s(t)$ and $0 = u(\pi, t) = y(\pi)s(t)$. If we had $y(x) = 0$ for all x or $s(t) = 0$ for all t then we would obtain the trivial solution $u(x, t) = 0$ for all x, t, so let us assume this is not the case, so the boundary conditions become $y(0) = y(\pi) = 0$. When $u(x, t) = y(x)s(t)$, the DE becomes $y(x)s''(t) = c^2 y''(x)s(t)$ which we can write as $\frac{y''(x)}{y(x)} = \frac{1}{c^2}\frac{s''(t)}{s(t)}$. Since the function on the left is a function of x (and is constant in t) and the function on the right is a function of t (and is constant in x), in order for these two functions to be equal for all x, t they must both be constant, say

$$\frac{y''(x)}{y(x)} = k = \frac{1}{c^2}\frac{s''(t)}{s(t)}$$

where k is constant.

First we solve the DE $\frac{y''(x)}{y(x)} = k$ subject to the boundary conditions $y(0) = y(\pi) = 0$. If $k = 0$ then the DE becomes $y'' = 0$, which has solution $y = Cx + D$, and the boundary conditions give $C = D = 0$, so we obtain the trivial solution. If $k > 0$, say $k = n^2$ where $n > 0$, then the DE becomes $y'' - n^2 y = 0$, which has solution $y = Ce^{nx} + De^{-nx}$, and the boundary conditions give $C + D = 0$ and $Ce^{n\pi} + De^{-n\pi} = 0$ which imply that $C = D = 0$, so again we obtain the trivial solution. Suppose that $k < 0$, say $k = -n^2$ where $n > 0$. The DE becomes $y'' + n^2 y = 0$ which has solution $y = C\cos nx + D\sin nx$. The boundary condition $y(0) = 0$ gives $C = 0$ so that $y = D\sin nx$, and the boundary condition $y(\pi) = 0$ gives $D = 0$ or $\sin n\pi = 0$. If $D = 0$ we obtain the trivial solution and if $\sin n\pi = 0$ then we must have $n \in \mathbf{Z}$. Thus in order to obtain a nontrivial solution to the DE which satisfies the boundary conditions we must have $k = -n^2$ for some $n \in \mathbf{Z}^+$ and, in this case,

$$y(x) = D_n \sin nx \text{ , where } D_n \in \mathbf{R}.$$

When $k = -n^2$ with $n \in \mathbf{Z}^+$, and $y(x) = D_n \sin nx$, the DE $\frac{1}{c^2}\,\frac{s''(t)}{s(t)} = k$. becomes $s''(t) + (cn)^2 s(t) = 0$, and the solution is $s(t) = A_n \cos(cn\,t) + B_n \sin(cn\,t)$. Thus, for each $n \in \mathbf{Z}^+$, and for all $A_n, B_n \in \mathbf{R}$, the function

$$u(x,t) = y(x)s(t) = \big(A_n \cos cnt + B_n \sin cnt\big)\sin nx$$

is a solution to the one-dimensional wave equation which satisfies the boundary conditions (we remark that it would be redundant to include the constants D_n as they could be amalgamated with the constants A_n and B_n).

In order to find a solution which satisfies the given initial conditions $u(x,0) = f(x)$ and $\frac{\partial u}{\partial t}(x,0) = g(x)$, we look for a solution of the form

$$u(x,t) = \sum_{n=1}^{\infty} \big(A_n \cos cnt + B_n \sin cnt\big)\sin nx\,.$$

In order to obtain $u(x,0) = f(x)$ we need $\sum\limits_{n=1}^{\infty} A_n \sin nx = f(x)$ and so we choose A_n to be equal to the Fourier coefficients of the odd 2π-periodic function $F(x)$ with $F(x) = f(x)$ for $0 \le x \le \pi$, that is we choose

$$A_n = \frac{1}{\pi}\int_{-\pi}^{\pi} F(x)\sin nx\;dx = \frac{2}{\pi}\int_0^{\pi} f(x)\sin nx\;dx\,.$$

Assuming that we can differentiate term-by-term, we have

$$\tfrac{\partial u}{\partial t}(x,t) = \sum_{n=1}^{\infty} \big(-cnA_n \sin cnt + cnB_n \cos nt\big)\sin nx.$$

In order to obtain $\frac{\partial u}{\partial t}(x,0) = g(x)$ we need $\sum\limits_{n=1}^{\infty} cnB_n \sin nx = g(x)$ and so we choose B_n to be equal to the Fourier coefficients of the odd 2π-periodic function $G(x)$ with $G(x) = g(x)$ for $0 \le x \le \pi$, that is

$$B_n = \frac{2}{cn\pi}\int_0^{\pi} g(x)\sin nx\;dx\,.$$

5.10 Remark: Let us now begin a more formal presentation of Fourier series in which we consider convergence issues more carefully.

5.11 Definition: A real-valued **trigonometric polynomial** is a function $f:\mathbf{R}\to\mathbf{R}$ of the form

$$f(x)=a_0+\sum_{n=1}^{m}a_n\cos nx+\sum_{n=1}^{m}b_n\sin nx$$

for some $a_n,b_n\in\mathbf{R}$, and we say that $f(x)$ is of degree m when either $a_m\neq 0$ or $b_m\neq 0$. A real-valued **trigonometric series** is a series of the form

$$a_0+\sum_{n=1}^{\infty}a_n\cos nx+\sum_{n=1}^{\infty}b_n\sin nx$$

which is given by its sequence of partial sums $s_m(x)=a_0+\sum\limits_{n=1}^{m}a_n\cos nx+\sum\limits_{n=1}^{m}b_n\sin nx$.

5.12 Remark: A trigonometric series may or may not converge and, indeed, we can consider several different notions of convergence, for example pointwise convergence, uniform convergence, or convergence with respect to a p-norm.

5.13 Definition: Every real-valued trigonometric polynomial is a smooth 2π-periodic function $f:\mathbf{R}\to\mathbf{R}$. Every 2π-periodic function $f:\mathbf{R}\to\mathbf{R}$ determines, and is determined by, a function $f:[-\pi,\pi]\to\mathbf{R}$ with $f(-\pi)=f(\pi)$, or equivalently by a function $f:T\to\mathbf{R}$ where $T=\mathbf{R}/2\pi\mathbf{Z}$, or equivalently by a function $f:S\to\mathbf{R}$ where $S=\left\{e^{it}\big|-\pi\leq t\leq\pi\right\}$ $=\left\{z\in\mathbf{Z}\big||z|=1\right\}$. A function $f:T\to\mathbf{R}$ is continuous (or differentiable, or $\mathcal{C}^k$) if and only if the corresponding 2π-periodic function $f:\mathbf{R}\to\mathbf{R}$ is continuous (or differentiable, or $\mathcal{C}^k$). We say that a function $f:T\to\mathbf{R}$ is **measurable** when the corresponding 2π-periodic function $f:\mathbf{R}\to\mathbf{R}$ is measurable, or equivalently when the corresponding function $f:[-\pi,\pi]\to\mathbf{R}$ with $f(-\pi)=f(\pi)$ is measurable. For a measurable function $f:T\to\mathbf{R}$ and for $1\leq p\leq\infty$ we define the **p-norm** $||f||_p$ of the function $f:T\to\mathbf{R}$ to be equal to the p-norm $||f||_p$ of the corresponding function $f:[-\pi,\pi]\to\mathbf{R}$. We define $L_p(T,\mathbf{R})$ to be the quotient of the set of measurable functions $f:T\to\mathbf{R}$ with $||f||_p<\infty$ under the equivalence relation in which $f\sim g$ when $f(x)=g(x)$ for a.e. $x\in[-\pi,\pi]$. Note that because $\lambda\big([-\pi,\pi]\big)=2\pi<\infty$, for $1\leq p\leq\infty$ we have $L_\infty(T)\subseteq L_p(T)\subseteq L_1(T)$.

5.14 Definition: When $f(x)=a_0+\sum\limits_{n=1}^{m}a_n\cos nx+b_n\sin nx$, where $a_n,b_n\in\mathbf{R}$, we have $f\in\mathcal{C}^\infty(T)$ and we know that the coefficients a_n and b_n are given by the formulas

$$a_0=\frac{1}{2\pi}\int_{-\pi}^{\pi}f(x)\,dx\ ,\ a_n=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\cos nx\ dx\ ,\ b_n=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\sin nx\ dx\,.$$

Note that the above integrals all exist and are finite for any function $f\in L_1(T,\mathbf{R})$. Given a function $f\in L_1(T,\mathbf{R})$, we define the real **Fourier coefficients** of f to be the real numbers $a_n=a_n(f)$ and $b_n=b_n(f)$ given by the above formulas, and we define the real **Fourier series** of f to be the corresponding real trigonometric series. Note that a real Fourier series is a real trigonometric series which arises, in this way, from some function $f\in L_1(T,\mathbf{R})$.

5.15 Definition: A complex-valued **trigonometric polynomial** is a function $f:\mathbf{R}\to\mathbf{C}$ of the form

$$f(x)=\sum_{n=-m}^{m} c_n e^{inx}$$

for some $c_n\in\mathbf{C}$, and we say that $f(x)$ is of degree m when either $c_m\neq 0$ or $c_{-m}\neq 0$. A complex-valued **trigonometric series** is a series of the form

$$\sum_{n=-\infty}^{\infty} c_n e^{inx}$$

which is given by its sequence of partial sums $s_m(x)=\sum\limits_{n=-m}^{m} c_n e^{inx}$.

5.16 Definition: Every complex-valued trigonometric polynomial is a smooth 2π-periodic function $f:\mathbf{R}\to\mathbf{C}$. Every 2π-periodic function $f:\mathbf{R}\to\mathbf{C}$ determines, and is determined by, a function $f:[-\pi,\pi]\to\mathbf{C}$ with $f(-\pi)=f(\pi)$, or equivalently by a function $f:T\to\mathbf{C}$ where $T=\mathbf{R}/2\pi\mathbf{Z}$, or equivalently by a function $f:S\to\mathbf{C}$ where $S=\big\{z\in\mathbf{Z}\big||z|=1\big\}$. For $1\leq p\leq\infty$, we define $L_p(T)=L_p(T,\mathbf{C})$ in the same way that we defined $L_p(T,\mathbf{R})$. For $f:T\to\mathbf{C}$ given by $f=u+iv$ where $u:T\to\mathbf{R}$ and $v:T\to\mathbf{R}$, f is measurable if and only if u and v are both measurable, and in this case we have $\int_T f=\int_T u+i\int_T v$, $\int_T|f|=\int_T\sqrt{u^2+v^2}$, $||f||_p=||\sqrt{u^2+v^2}||_p$ and $f\in L_p(T,\mathbf{C})$ if and only if $u\in L_p(T,\mathbf{R})$ and $v\in L_p(T,\mathbf{R})$.

5.17 Definition: When $f(x)=\sum\limits_{n=-m}^{m} c_n e^{inx}$, where $c_n\in\mathbf{C}$, because

$$\int_T e^{ikx}e^{-ilx}dx=\int_{-\pi}^{\pi}\cos(k-l)x\;dx+i\int_{-\pi}^{\pi}\sin(k-l)x\;dx=\begin{cases}2\pi \text{ if } k=l\\ 0 \;\text{ if } k\neq l,\end{cases}$$

it follows that the coefficients c_n are given by the formula

$$c_n=c_n(f)=\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx\,.$$

Note that the above integrals exist and are finite for any function $f\in L_1(T)=L_1(T,\mathbf{C})$. Given a function $f\in L_1(T)$, we define the (complex) **Fourier coefficients** of f to be the complex numbers $c_n=c_n(f)$ given by the above formulas, and we define the (complex) **Fourier series** of f to be the corresponding complex trigonometric series.

5.18 Note: Given $a_n,b_n\in\mathbf{R}$, we have

$$\begin{aligned}a_0+\sum_{n=1}^{m}a_n\cos nx+\sum_{n=1}^{m}b_n\sin nx&=a_0+\sum_{n=1}^{m}a_n\,\tfrac{e^{inx}+e^{-inx}}{2}+\sum_{n=1}^{m}b_n\,\tfrac{e^{inx}-e^{-inx}}{2i}\\&=a_0+\sum_{n=1}^{m}\big(\tfrac{a_0}{2}-i\,\tfrac{b_n}{2}\big)e^{inx}+\sum_{n=1}^{m}\big(\tfrac{a_0}{2}-\tfrac{b_n}{2i}\big)e^{-inx}=\sum_{n=-m}^{m}c_ne^{inx}\end{aligned}$$

where $c_0=a_0$ and $c_n=\frac{1}{2}(a_n-ib_n)$ and $c_{-n}=\overline{c_n}=\frac{1}{2}(a_n+ib_n)$ for $n>0$.

On the other hand, given $f \in L_1(T, \mathbf{R})$, for $n > 0$ we have

$$\begin{aligned} c_n &= \frac{1}{2\pi}\int_T f(x)e^{-inx}\,dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\big(\cos nx - i\sin nx\big)\,dx \\ &= \frac{1}{2\pi}\left(\int_{-\pi}^{\pi} f(x)\cos nx - i\int_{-\pi}^{\pi} f(x)\sin nx\;dx\right) = \tfrac{1}{2}(a_n - i\,b_n) \\ c_{-n} &= \frac{1}{2\pi}\int_T f(x)e^{inx}\,dx = \tfrac{1}{2}(a_n + i\,b_n) \end{aligned}$$

It follows that when $f \in L_1(T, \mathbf{R})$, the m^{th} partial sum of the real Fourier series for f is exactly equal to the m^{th} partial sum for the complex Fourier series for f.

5.19 Definition: For $f \in L_1(T) = L_1(T, \mathbf{C})$ we denote the m^{th} partial sum of the Fourier series of f by $s_m(f)$, so we have

$$s_m(f)(x) = \sum_{n=-m}^{m} c_n e^{inx} \quad \text{, where } c_n = c_n(f) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}dt.$$

5.20 Exercise: Show that if $f \in L_p(T)$ with $1 \le p \le \infty$, and $s_m(x) = \sum\limits_{n=-m}^{m} d_n\, e^{inx}$ with $s_m \to f$ in $L_p(T)$, then $d_n = c_n(f)$.

5.21 Theorem: *(The Stone-Weierstrass Theorem) Let X be a compact metric space and let $C(X) = C(X, F)$ be the set of continuous functions $f : X \to F$ where $F = \mathbf{R}$ or $\mathbf{C}$. Let A be an algebra in $C(X)$ which contains the constant functions and which separates points in X. Then A is uniformly dense in $C(X)$, which means that for all $f \in C(X)$ and for all $\epsilon > 0$ there exists $g \in A$ such that $||g - f||_\infty < \epsilon$.*

Proof: We omit the proof.

5.22 Corollary: *The set of polynomials $\mathbf{R}[x]$ is uniformly dense in $C\big([a, b]\big)$.*

5.23 Corollary: *The set of functions of the form*

$$u(x, y) = \sum_{k=1}^{n} f_k(x)g_k(y) \quad \text{, where } f_k \in C([a, b]) \text{ and } g_k \in C\big([c, d]\big)$$

is uniformly dense in $C\big([a, b] \times [c, d]\big)$.

5.24 Corollary: *The set of real trigonometric polynomials is uniformly dense in $C\big(T, \mathbf{R}\big)$, and the set of complex trigonometric polynomials is uniformly dense in $C(T) = C(T, \mathbf{C})$.*

5.25 Corollary: *(The Riemann-Lebesgue Lemma) Let $f \in L_1(T)$. Then $\lim\limits_{n\to\pm\infty} c_n(f) = 0$.*

Proof: Let $\epsilon > 0$. Choose a trigonometric polynomial $p(x) = \sum\limits_{n=-m}^{m} a_n e^{inx}$ with $||p - f||_\infty < \epsilon$. Then for $n > m$ we have $c_n(p) = 0$ and so

$$\begin{aligned} |c_n(f)| = \big|c_n(f) - c_n(p)\big| &= \left|\frac{1}{2\pi}\int_{-\pi}^{\pi} \big(f(x) - p(x)\big)\, e^{-inx}\,dx\right| \\ &\le \frac{1}{2\pi}\int_{-\pi}^{\pi} \big|f(x) - p(x)\big|\,dx \le ||f - p||_\infty < \epsilon. \end{aligned}$$

5.26 Note: Since real trigonometric polynomials are dense in $C(T,\mathbf{R})$, hence also in $L_2(T,\mathbf{R})$, it follows that the orthonormal set

$$\left\{\tfrac{1}{\sqrt{2\pi}}\,,\ \tfrac{1}{\sqrt{\pi}}\cos nx\,,\ \tfrac{1}{\sqrt{\pi}}\sin nx\,\middle|\,n\in\mathbf{Z}^+\right\}$$

is a Hilbert basis for the Hilbert space $L_2(T,\mathbf{R})$. For $f\in L_2(T,\mathbf{R})$ we have

$$a_0(f)=\tfrac{1}{2\pi}\langle f,1\rangle\ ,\ a_n(f)=\tfrac{1}{2\pi}\langle f,\cos nx\rangle\ ,\ b_n=\tfrac{1}{2\pi}\langle f,\sin nx\rangle.$$

Similarly, since complex trigonometric polynomials are dense in $L_2(T)=L_2(T,\mathbf{C})$, it follows that the orthonormal set

$$\left\{\tfrac{1}{\sqrt{2\pi}}e^{inx}\,\middle|\,n\in\mathbf{Z}\right\}$$

is a Hilbert basis for the Hilbert space $L_2(T,\mathbf{C})$. For $f\in L_2(T,\mathbf{C})$ we have

$$c_n(f)=\tfrac{1}{2\pi}\langle f,e^{inx}\rangle.$$

The following theorem is an immediate consequence of our earlier study of Hilbert spaces.

5.27 Theorem: *In the Hilbert space $L_2(T)=L_2(T,\mathbf{C})$, we have the following.*

(1) (Best Approximation) Given $f\in L_2(T)$, $s_m(f)$ is the unique trigonometric polynomial of degree at most m which best approximates f in $L_2(T)$.

(2) (Convergence) Given $f\in L_2(T)$ we have $s_m(f)\to f$ in $L_2(T)$.

(3) (Parseval's Identity) Given $f\in L_2(T)$ we have $||f||_2^2=2\pi\sum\limits_{n=-\infty}^{\infty}\left|c_n(f)\right|^2$.

(4) (Inner Product Formula) Given $f,g\in L_2(T)$ we have $\langle f,g\rangle=2\pi\sum\limits_{n=-\infty}^{\infty}c_n(f)\,\overline{c_n(g)}$.

(5) (The Riesz-Fischer Theorem) Given $c_n\in\mathbf{C}$, if $\sum\limits_{n=-\infty}^{\infty}|c_n|^2<\infty$ then there exists a unique $f\in L_2(T)$ such that $c_n=c_n(f)$.

Proof: These are immediate consequences of Theorems 4.23 and 4.24.

5.28 Exercise: Show that when $f\in L_2(T,\mathbf{R})$, Parseval's Identity becomes

$$||f||_2^2=2\pi\left|a_0(f)\right|^2+\pi\sum_{n=1}^{\infty}\left|a_n(f)\right|^2+\pi\sum_{n=1}^{\infty}\left|b_n(f)\right|^2.$$

5.29 Exercise: Use Parseval's Identity, together with the result of Example 5.4, to prove that $\sum\limits_{k=0}^{\infty}\frac{1}{(2k+1)^4}=\frac{\pi^4}{96}$ and use this result to calculate $\sum\limits_{n=1}^{\infty}\frac{1}{n^4}$.

5.30 Note: Let $f \in L_1(T)$. Then

$$s_m(f)(x) = \sum_{n=-m}^{m} c_n(f)\, e^{inx} = \sum_{n=-m}^{m} \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int}\, dt \right) e^{inx}$$
$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \sum_{n=-m}^{m} e^{in(x-t)}\, dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) D_m(x-t)\, dt$$

where

$$D_m(u) = \sum_{n=-m}^{m} e^{inu} = e^{-imu}\, \frac{e^{i(2m+1)u} - 1}{e^{iu} - 1} = \frac{e^{i(m+1)u} - e^{-imu}}{e^{iu} - 1} \cdot \frac{e^{-iu/2}}{e^{-iu/2}}$$
$$= \frac{e^{i(2m+1)u/2} - e^{i(2m+1)u/2}}{e^{iu/2} - e^{-iu/2}} = \frac{\sin \frac{(2m+1)u}{2}}{\sin \frac{u}{2}}\,.$$

5.31 Definition: The above function $D_m : T \to \mathbf{R}$ is called the m^{th} **Dirichlet kernel**.

5.32 Remark: For $f, g \in L_1(T)$, the **convolution** of f with g is the function $f \star g : T \to \mathbf{R}$ given by $(f \star g)(x) = \frac{1}{2\pi} \int_T f(t) g(x-t)\, dt$. Using this notation we have $s_m(f) = f \star D_m$.

5.33 Note: We have

$$\int_{-\pi}^{\pi} D_m(u)\, du = \int_{-\pi}^{\pi} \sum_{n=-m}^{m} e^{inu}\, du = \int_{-\pi}^{\pi} 1 + \sum_{n=1}^{m} 2\cos(nu)\ du = 2\pi$$

and

$$\int_{-\pi}^{\pi} \big|D_m(u)\big|\, du = \int_{-\pi}^{\pi} \left| \frac{\sin \frac{(2m+1)u}{2}}{\sin \frac{u}{2}} \right| du = 2 \int_0^{\pi} \left| \frac{\sin \frac{(2m+1)u}{2}}{\sin \frac{u}{2}} \right| du$$
$$\geq 2 \int_{u=0}^{\pi} \frac{\left| \sin \frac{(2m+1)u}{2} \right|}{\frac{u}{2}}\, du = 2 \int_{t=0}^{(m+\frac{1}{2})\pi} \frac{|\sin t|}{\frac{t}{2m+1}} \cdot \tfrac{2}{2m+1}\, dt$$
$$\geq 4 \sum_{n=1}^{m} \int_{(n-1)\pi}^{n\pi} \frac{|\sin t|}{t}\, dt \geq 4 \sum_{n=1}^{m} \int_{(n-1)\pi}^{n\pi} \frac{|\sin t|}{n\pi}\, dt$$
$$= \tfrac{8}{\pi} \sum_{n=1}^{m} \tfrac{1}{n} \geq \tfrac{8}{\pi} \int_{x=1}^{m+1} \tfrac{1}{x}\, dx = \tfrac{8}{\pi} \ln(m+1) \geq \tfrac{8}{\pi}\, \ln m\,.$$

5.34 Theorem: *(Pointwise Divergence) Let $C(T) = C(T, \mathbf{C})$ be the Banach space of continuous functions $f : T \to \mathbf{C}$ equipped with the supremum norm. There exists a dense $\mathcal{G}_\delta$ set $E \subseteq C(T)$ such that for every $f \in E$ the set of points $x \in T$ at which the Fourier series for f diverges is dense in T.*

Proof: First we fix $x = 0$. For $m \in \mathbf{Z}^+$, define $F_m : C(T) \to \mathbf{C}$ by

$$F_m(f) = s_m(f)(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)\, D_m(t)\, dt.$$

Note that

$$\big|F_m(f)\big| \le \frac{1}{2\pi} \int_{-\pi}^{\pi} \big|f(t)\big|\,\big|D_m(t)\big|\, dt \le \frac{1}{2\pi}\, ||f||_\infty \int_{-\pi}^{\pi} \big|D_m(t)\big|\, dt$$

so we have

$$||F_m|| \le \frac{1}{2\pi} \int_{-\pi}^{\pi} \big|D_m(t)\big|\, dt\,.$$

We claim that in fact $||F_m|| = \dfrac{1}{2\pi} \displaystyle\int_{-\pi}^{\pi} \big|D_m(t)\big|\, dt$. Fix m and define

$$s(t) = \begin{cases} 1 \text{ if } D_m(t) \ge 0, \\ -1 \text{ if } D_m(t) < 0. \end{cases}$$

Construct continuous functions $g_n : T \to \mathbf{R}$ with $|g_n| \le 1$ such that $g_n \to s$ in $L_1(T)$. By the Dominated Convergence Theorem, we have

$$F_m(g_n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} g_n(t) D_m(t)\, dt \longrightarrow \frac{1}{2\pi} \int_{-\pi}^{\pi} s(t) D_m(t)\, dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} \big|D_m(t)\big|\, dt.$$

It follows that $||F_m|| = \dfrac{1}{2\pi} \displaystyle\int_{-\pi}^{\pi} \big|D_m(t)\big|\, dt$, as claimed. By the above note, $||F_m|| \ge \frac{8}{\pi} \ln m$, so the set of linear operators $S = \big\{F_m \big| m \in \mathbf{Z}^+\big\}$ is not uniformly bounded. By the Uniform Boundedness Principle, applied to the set S, there exists a function $f \in C(T)$ such that for all $M > 0$ we have $\big\|F_m(f)\big\| > M$, that is $\big|s_m(f)(0)\big| > M$, for some $m \in \mathbf{Z}^+$. For this function $f \in C(T)$, the Fourier series for f diverges at 0 because $\limsup\limits_{m\to\infty} \big|s_m(f)(0)\big| = \infty$.

Let $Q = \{a_1, a_2, a_3, \cdots\}$ be a dense subset of $[0, 2\pi]$ and consider each a_n as an element in T. For each $n \in \mathbf{Z}^+$ let $f_n(x) = f(x - a_n)$ so that $\limsup\limits_{m\to\infty} \big|s_m(f_n)(a_n)\big| = \infty$. For $n, m \in \mathbf{Z}^+$, define $L_{n,m} : C(T) \to \mathbf{C}$ by $L_{n,m}(f) = s_m(f)(a_n)$. By Condensation of Singularities, the set

$$E = \Big\{ f \in C(T) \,\Big|\, \limsup_{m\to\infty} \big\|L_{n,m}(f)\big\| = \infty \text{ for all } n \in \mathbf{Z}^+ \Big\}$$

is a dense $\mathcal{G}_\delta$ in the Banach space $C(T)$. For each $f \in E$, we have $\limsup\limits_{m\ge 0} \big|s_m(f)(a_n)\big| = \infty$ for every $n \in \mathbf{Z}^+$, so the Fourier series for f diverges at every point a_n.

5.35 Theorem: *(Cesàro Convergence) Let $a_n \in \mathbf{C}$ for $n \geq 0$, let $s_m = \sum_{n=0}^{m} a_n$ and let*

$$\sigma_\ell = \tfrac{1}{\ell+1} \sum_{m=0}^{\ell} s_m \, .$$

If the sequence $\{s_m\}$ converges then so does the sequence $\{\sigma_\ell\}$ and, in this case, we have

$$\lim_{\ell\to\infty} \sigma_\ell = \lim_{m\to\infty} s_m .$$

Proof: The proof is left as an exercise.

5.36 Definition: For $f \in L_1(T) = L_1(T, \mathbf{C})$, we define the ℓ^{th} **Cesàro mean** of the Fourier series of f to be the function $\sigma_\ell(f) : T \to \mathbf{C}$ given by

$$\sigma_\ell(f) = \tfrac{1}{\ell+1} \sum_{m=0}^{\ell} s_m(f).$$

5.37 Note: For $f \in L_1(T) = L_1(T, \mathbf{C})$ we have

$$\begin{aligned}\sigma_\ell(f)(x) &= \frac{1}{\ell+1} \sum_{m=0}^{\ell} s_m(f)(x) = \frac{1}{\ell+1} \sum_{m=0}^{\ell} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) D_m(x-t)\, dt \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \, \tfrac{1}{\ell+1} \sum_{m=0}^{\ell} D_m(x-t)\, dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) K_\ell(x-t)\, dt\end{aligned}$$

where

$$\begin{aligned}K_\ell(u) &= \frac{1}{\ell+1} \sum_{m=0}^{\ell} D_m(u) = \frac{1}{\ell+1} \sum_{m=0}^{\ell} \frac{\sin \frac{(2m+1)u}{2}}{\sin \frac{u}{2}} \\ &= \frac{1}{(\ell+1)\sin\frac{u}{2}} \cdot \text{Im}\Big(\sum_{m=0}^{\ell} e^{i(2m+1)u/2} \Big) = \frac{1}{(\ell+1)\sin\frac{u}{2}} \cdot \text{Im}\Big(e^{iu/2} \sum_{m=0}^{\ell} e^{imu} \Big) \\ &= \frac{1}{(\ell+1)\sin\frac{u}{2}} \cdot \text{Im}\Big(e^{iu/2} \, \frac{e^{i(\ell+1)u} - 1}{e^{iu} - 1} \Big) = \frac{1}{(\ell+1)\sin\frac{u}{2}} \cdot \text{Im}\Big(\frac{e^{i(\ell+1)u} - 1}{e^{iu/2} - e^{\frac{-iu}{2}}} \Big) \\ &= \frac{1}{(\ell+1)\sin\frac{u}{2}} \cdot \text{Im}\Big(\frac{e^{i(\ell+1)u/2} - e^{-i(\ell+1)u/2}}{e^{iu/2} - e^{-iu/2}} \cdot e^{i(\ell+1)u/2} \Big) \\ &= \frac{1}{(\ell+1)\sin\frac{u}{2}} \cdot \frac{\sin\frac{(\ell+1)u}{2}}{\sin\frac{u}{2}} \cdot \sin\frac{(\ell+1)u}{2} = \frac{\sin^2 \frac{(\ell+1)u}{2}}{(\ell+1)\, \sin^2 \frac{u}{2}} \, .\end{aligned}$$

5.38 Definition: The above function $K_\ell : T \to \mathbf{R}$ is called the ℓ^{th} **Féjer kernel**.

5.39 Remark: Using convolution notation, for $f \in L_1(T)$ we have $\sigma_\ell(f) = f \star K_\ell$.

5.40 Lemma: *We have*

(1) For $0 < t \leq \pi$ we have $0 \leq K_\ell(t) \leq \frac{\pi^2}{(\ell+1)\, t^2}$.

(2) $\displaystyle\int_{-\pi}^{\pi} K_\ell(t)\, dt = 2 \int_0^{\pi} K_\ell(t)\, dt = 2\pi$.

(3) $\displaystyle\int_{\pi}^{\pi} f(t) K_\ell(x-t)\, dt = \int_{-\pi}^{\pi} f(x+t) K_\ell(t)\, dt = \int_{-\pi}^{\pi} f(x-t) K_\ell(t)\, dt$.

Proof: The proof is left as an exercise.

5.41 Theorem: *(Convergence of the Cesàro Means) Let $f \in L_1(T)$ and consider f as a 2π-periodic function $f : \mathbf{R} \to \mathbf{C}$.*

(1) If $a \in \mathbf{R}$ and the one-sided limits $f(a^-) = \lim\limits_{x\to a^-} f(x)$ and $f(a^+) = \lim\limits_{x\to a^+} f(x)$ both exist in $\mathbf{C}$, then

$$\lim_{\ell\to\infty} \sigma_\ell(f)(a) = \frac{f(a^-)+f(a^+)}{2}\,.$$

(2) If $a, b \in \mathbf{R}$ with $a \le b$ and f is continuous in $[a,b]$ then $\sigma_\ell \to f$ uniformly on $[a,b]$.

Proof: By Part 3 of the above lemma, we have

$$\sigma_\ell(f)(a) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)K_\ell(a-t)\,dt = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{f(a+t)+f(a-t)}{2}\,K_\ell(t)\,dt$$

and by Part 2 of the above lemma we have

$$\frac{f(a^+)+f(a^-)}{2} = \frac{1}{2\pi}\int_{-\pi}^{\pi}\frac{f(a^+)+f(a^-)}{2}\,K_\ell(t)\,dt$$

and so

$$\begin{aligned}
\left|\sigma_\ell(f)(a) - \frac{f(a^+)+f(a^-)}{2}\right| &= \left|\frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\frac{f(a+t)+f(a-t)}{2} - \frac{f(a^+)+f(a^-)}{2}\right)K_\ell(t)\;dt\right| \\
&= \left|\frac{1}{2\pi}\int_0^{\pi}\Big(\big(f(a+t)-f(a^+)\big)+\big(f(a-t)-f(a^-)\big)\Big)K_\ell(t)\;dt\right| \\
&\le \frac{1}{2\pi}\int_0^{\pi}\Big(\big|f(a+t)-f(a^+)\big| + \big|f(a-t)-f(a^-)\big|\Big)\,K_\ell(t)\;dt \\
&= I_\delta + J_\delta,
\end{aligned}$$

for any $0 < \delta \le \pi$, where

$$\begin{aligned}
I_\delta &= \frac{1}{2\pi}\int_0^{\delta}\Big(\big|f(a+t)-f(a^+)\big| + \big|f(a-t)-f(a^-)\big|\Big)\,K_\ell(t)\;dt \\
J_\delta &= \frac{1}{2\pi}\int_\delta^{\pi}\Big(\big|f(a+t)-f(a^+)\big| + \big|f(a-t)-f(a^-)\big|\Big)\,K_\ell(t)\;dt\,.
\end{aligned}$$

Let $\epsilon > 0$. Choose $\delta > 0$ so that for all $0 < t < \delta$ we have $\big|f(x+t)-f(a^+)\big| < \frac{\epsilon}{2}$ and $\big|f(x-t)-f(a^-)\big| < \frac{\epsilon}{2}$. Then, by Part 2 of the above lemma,

$$I_\delta \le \frac{1}{2\pi}\int_0^{\pi}\epsilon\cdot K_\ell(t)\,dt \le \tfrac{\epsilon}{2}.$$

By Part 1 of the above lemma, for $\delta \le t \le \pi$ we have $K_\ell(t) \le \frac{\pi^2}{(\ell+1)\,\delta^2}$ so for $\ell+1 \ge \frac{M}{\epsilon}$ where $M = \pi\big(||f||_1 + \pi|f(a^+)+f(a^-)|\big)/\delta^2$ we have

$$\begin{aligned}
J_\delta &\le \frac{1}{2\pi}\int_\delta^{\pi}\Big(|f(a+t)| + |f(a-t| + |f(a^+)+f(a^-)|\Big)\,\tfrac{\pi^2}{(\ell+1)\,\delta^2}\;dt \\
&\le \tfrac{1}{2\pi}\cdot\tfrac{\pi^2}{(\ell+1)\,\delta^2}\big(||f||_1 + \pi\big|f(a^+)+f(a^-)\big| = \tfrac{M}{2(\ell+1)} \le \tfrac{\epsilon}{2}.
\end{aligned}$$

This proves Part (1), and Part (2) can be proven using the same method noting that the estimates can be made uniformly.

5.42 Corollary: *Let $f \in L_1(T)$, consider f as a 2π-periodic function $f : \mathbf{R} \to \mathbf{C}$, and let $a \in \mathbf{R}$. If $f(a^+)$, $f(a^-)$ and $\lim_{m\to\infty} s_m(f)(a)$ all exist in $\mathbf{C}$ then*

$$\lim_{m\to\infty} s_m(f)(a) = \frac{f(a^+) + f(a^-)}{2}.$$

5.43 Remark: The above corollary justifies the argument given in Remark 5.5 where we showed that $\sum_{k=1}^{\infty} \frac{1}{(2k+1)^2} = \frac{\pi^2}{8}$.