X1. Let X be any set. Define $d\colon X \times X \to [0,\infty)$ by

$$d(x,x') = \begin{cases} 1 & \text{if } x \neq x', \\ 0 & \text{if } x = x'. \end{cases}$$

Verify that d is a metric on X and determine which subsets of X are open with respect to D.

X2. Let (X,ρ) be a metric space. Define $\sigma\colon X \times X \to [0,1)$ by *due 1Th*

$$\sigma(x,x') = \frac{\rho(x,x')}{1+\rho(x,x')}.$$

(a) Let $a, b \in (-1,\infty)$. Prove that

$$a < b \qquad \text{if and only if} \qquad \frac{a}{1+a} < \frac{b}{1+b}.$$

Hint: Do not use calculus. Do not cross multiply. Just notice what

$$1 - \frac{1}{1+c}$$

is equal to when $c \in \mathbf{R} \setminus \{-1\}$.

(b) Prove that σ is a metric on X.

(c) Prove that each open ball for ρ is also an open ball for σ.

(d) Prove that each open ball for σ is either an open ball for ρ or is all of X.

(e) Prove that for each $G \subseteq X$, G is open with respect to ρ iff G is open with respect to σ.

Remark. The significance of problem X2 is that it shows that any metric has the same open sets as some bounded metric.

X3. Let X be a non-empty set, let (Y,d) be a metric space, and let Z be the set of bounded functions from X into Y. Define a function D on $Z \times Z$ by *due 1Th*

$$D(f,g) = \sup\{\, d\big(f(x), g(x)\big) : x \in X \,\}.$$

(a) Prove that D is a metric on Z. (Warning: Do not "calculate" with sup. Any inequalities you assert involving sup must be justified based on the definition of sup. Also, please do not use proof by contradiction when it is not needed and does not shorten the argument.)

(b) If $f \in Z$ and if (f_n) is a sequence in Z which converges to f with respect to the metric D, then clearly (f_n) converges pointwise to f because for each $x \in X$, we have $d\big(f(x), f_n(x)\big) \leq D(f, f_n)$. However, the converse is false in general. Pointwise convergence does not imply convergence with respect to the metric D. To see this, consider the special case where $X = Y = \mathbf{R}$ and $d(y,y') = |y - y'|$ for all $y, y' \in \mathbf{R}$. Let (f_n) be the sequence in Z defined by

$$f_n(x) = \begin{cases} 1 & \text{if } x > n, \\ 0 & \text{if } x \leq n. \end{cases}$$

Prove that (f_n) converges pointwise to the zero function but does not converge to any element of Z with respect to the metric D.

(c) Now consider the general case again. Suppose that Y has at least two points. Prove that each pointwise convergent sequence in Z is convergent with respect to the metric D if and only if the set X is finite.

Remark. In Chapter 7 of the text, the notion of uniform convergence of sequences of functions is introduced and studied. Let f be a bounded function from X to Y and let (f_n) be a sequence of bounded functions from X to Y. In the notation of problem X3, $f_n \to f$ uniformly on X iff $D(f, f_n) \to 0$. In other words, $f_n \to f$ uniformly on X iff (f_n) converges to f in the metric space (Z, D). This provides an example to illustrate why we introduced the notion of a general metric space. There are many important metric spaces besides $\mathbf{R}$, $\mathbf{C}$, and $\mathbf{R}^k$. Some of the most important ones are metric spaces of functions. The space (Z,D) is one of the simplest examples of a metric space of functions.

Let (X, d) be a metric space. If (x_n) is a sequence in X, then to say that (x_n) *is Cauchy* means that for each $\varepsilon > 0$, there exists $N \in \mathbf{N}$ such that for all $m, n \in \mathbf{N}$ with $m, n > N$, we have $d(x_m, x_n) < \varepsilon$. It is easy to see that if (x_n) is a convergent sequence in X, then (x_n) is Cauchy. To say that (X, d) *is complete* (as a metric space) means that each Cauchy sequence in X is convergent in X. Thus in a complete metric space, a sequence is convergent if and only if it is Cauchy.

As we know, each non-empty subset of $\mathbf{R}$ which is bounded above has a least upper bound in $\mathbf{R}$. (This property of $\mathbf{R}$ is sometimes called *Dedekind completeness.*) Using this fact, we shall soon show that $\mathbf{R}$ is also complete as a metric space, when we give it its usual metric, namely $d(x, y) = |x - y|$. For now, you may take it for granted that $\mathbf{R}$ is complete in this sense.

X4. As in problem X3, let X be a non-empty set, let (Y, d) be a metric space, let Z be the set of bounded functions from X into Y, and let D be the metric on Z defined by *due 2Th*

$$D(f, g) = \sup \{\, d\big(f(x), g(x)\big) : x \in X \,\}.$$

Prove that if the metric space (Y, d) is complete, then so is the metric space (Z, D).

Let (X, ρ) and (Y, σ) be metric spaces. An *isometry* from X into Y is a map $f \colon X \to Y$ such that $\sigma\big(f(x), f(x')\big) = \rho(x, x')$ for all $x, x' \in X$. Informally, an isometry is a map that preserves distances between points.

X5. Let (X, d) be a metric space. Give $\mathbf{R}$ its usual metric. Let Z be the set of bounded functions from X into $\mathbf{R}$. Give Z the metric defined by *due 2Th*

$$D(f, g) = \sup \{\, |f(\xi) - g(\xi)| : \xi \in X \,\}.$$

By problem X4, the metric space (Z, D) is complete, because $\mathbf{R}$ with its usual metric is complete as a metric space. Fix $x_0 \in X$. For each $x \in X$, define $f_x \colon X \to \mathbf{R}$ by

$$f_x(\xi) = d(\xi, x) - d(\xi, x_0).$$

(a) Prove that for each $x \in X$, we have $f_x \in Z$.
(b) Define $\Phi \colon X \to Z$ by $\Phi(x) = f_x$. Prove that Φ is an isometry from (X, d) into (Z, D).

Let (X, d) be a metric space and let $X_1 \subseteq X$. Let d_1 be the restriction of d to $X_1 \times X_1$. Then clearly d_1 is a metric on X_1. We call d_1 the *subspace metric* that X_1 inherits from (X, d) and we say that the metric space (X_1, d_1) is a *subspace* of the metric space (X, d).

Remark. The result of problem X5 shows that any metric space is isometric to a subspace of a complete metric space.

Let A and B be sets. To say that A *is equinumerous to* B means that there is a bijection[1] from A to B. Recall that $\boldsymbol{\omega}$ denotes the set $\{\, 0, 1, 2, \dots \}$. If $n \in \boldsymbol{\omega}$, then to say that A *has* n *elements* means that either $n = 0$ and A is empty or $n \in \mathbf{N}$ and A is equinumerous to $\{\, 1, \dots, n \,\}$. To say that A *is finite* means that there exists $n \in \boldsymbol{\omega}$ such that A has n elements. To say that A *is infinite* means that A is not finite. To say that A *is countably infinite* means that A is equinumerous to $\mathbf{N}$. To say that A *is countable* means that A is finite or countably infinite.[2] It is easy to show that if B is an infinite subset of $\mathbf{N}$, then B is equinumerous to $\mathbf{N}$. It follows that A is countable if and only if A is equinumerous to a subset of $\mathbf{N}$. To say that A *is uncountable* means that A is not countable. Cantor (1873) pointed out that the set of rational numbers is countable and proved that the set of real numbers is uncountable.

X6. Let X be a set. Prove that X contains a countably infinite subset if and only if X is equinumerous to a proper subset of itself. (Do not use the axiom of choice.)

---

[1] A bijection from A to B is a one-to-one map from A onto B. Another name for a bijection from A to B is a one-to-one correspondence between A and B.

[2] Warning: The definition of *countable* that I have given is the one accepted by most mathematicians, but you should watch out for the fact that Rudin uses the term *countable* to mean what I have chosen to call *countably infinite*. So for Rudin, a finite set is not countable. To me, that is just ridiculous, so I will not follow Rudin's use of the term *countable*.

X7. Let X be an infinite set. Prove that X contains a countably infinite subset.[3]

X8. Let B be a set, let C be a countable subset of B, and let $A = B \setminus C$. Suppose that A has a countably infinite subset. Prove that A is equinumerous to B. (Do not use the axiom of choice.)

Example. Let $\Sigma = \{0,1\}^{\mathbf{N}}$ be the set of infinite binary sequences. Let C be the subset of Σ consisting of binary sequences that end with an infinite string of ones. Then C is countable. Let $A = \Sigma \setminus C$. Then A has a countably infinite subset. For instance, if a_n is the element of Σ that has a one in the n-th place and zeroes in all the other places, then $\{a_1, a_2, a_3, \dots\}$ is a countably infinite subset of A. Therefore, by problem X8, A is equinumerous to Σ.

X9. Let (X, d) be a complete metric space. Suppose (A_n) is a sequence of non-empty subsets of X such that $A_n \supseteq \overline{A_{n+1}}$ for each n and $\operatorname{diam}(A_n) \to 0$ as $n \to \infty$. Show that $\bigcap_n A_n$ contains exactly one point.

Notation. The following notation is not standard but will be convenient to use in the next few exercises. For each $n \in \mathbf{N}$, let $S_n = \{0,1\}^n$, the set of finite binary sequences of length n. Let $S = \bigcup_{n \in \mathbf{N}} S_n$, the set of non-empty finite binary sequences. Finally, let $\Sigma = \{0,1\}^{\mathbf{N}}$, the set of infinite binary sequences.

Reminder. To say that x *is a dyadic rational number* means that there exist an integer k and a natural number ℓ such that $x = k/2^\ell$. For instance 0, 1, $1/2$, $1/4$, and $3/4$ are dyadic rationals but $2/3$ is not. The number $15/24$ is a dyadic rational because it is equal to $5/8$.

due 3Th revised 26 Oct 05

X10. Let S_n, S, and Σ be as above. For each $n \in \mathbf{N}$ and each $s = (s_1, \dots, s_n) \in S_n$, let $x(s) = \sum_{k=1}^{n} s_k 2^{-k}$, let $y(s) = x(s) + 2^{-n}$, and let $I(s)$ be the closed interval $[x(s), y(s)]$. Thus $I(0) = [0, 1/2]$, $I(1) = [1/2, 1]$, $I(0,0) = [0, 1/4]$, $I(0,1) = [1/4, 1/2]$, $I(1,0) = [1/2, 3/4]$, $I(1,1) = [3/4, 1]$, $I(0,0,0) = [0, 1/8]$, and so on. For each $\sigma = (s_1, s_2, s_3, \dots) \in \Sigma$, the sequence of sets

$$I(s_1, \dots, s_n), \qquad n = 1, 2, 3, \dots$$

is a decreasing sequence of non-empty closed subsets of $\mathbf{R}$ with diameters tending to zero, so by problem X9, the intersection of this sequence of sets contains exactly one point which we shall denote by $f(\sigma)$ and which by definition is the point represented in binary by

$$0.s_1 s_2 s_3 \cdots .$$

The parts below will lead you through a proof that each number in $[0,1]$ has at least one such binary representation, that each dyadic rational in $[0,1)$ has exactly two such representations, one ending with repeating zeroes and one ending with repeating ones, and that each other number in $[0,1)$ has exactly one such representation and that this unique representation does not end with repeating zeroes nor does it end with repeating ones. Of course 1 has the two binary representations $0.111\cdots$ and $1.000\cdots$ but we are not considering the latter representation here.

- **(a)** Prove that f maps Σ onto $[0,1]$.
- **(b)** Let $\sigma = (s_1, s_2, s_3, \dots) \in \Sigma$. Prove that $f(\sigma) = 1$ if and only if $s_n = 1$ for all n.
- **(c)** Let $\sigma = (s_1, s_2, s_3, \dots) \in \Sigma$ and let $\tau = (t_1, t_2, t_3, \dots) \in \Sigma$. Suppose that $\sigma \neq \tau$. Then $s_n \neq t_n$ for some n. Consider the least such n and suppose (without loss of generality) that $s_n = 0$ and $t_n = 1$. Prove that $f(\sigma) = f(\tau)$ if and only if for each $k > n$, we have $s_k = 1$ and $t_k = 0$.

Now let C be the subset of Σ consisting of binary sequences that end with an infinite string of ones. Let $A = \Sigma \setminus C$ and let g be the restriction of f to A.

- **(d)** Prove that g is a one-to-one map from A onto the interval $[0,1)$.
- **(e)** Deduce that Σ is equinumerous to $[0,1]$.

---

[3] You will have to use the axiom of choice to prove this, though not the full strength of the axiom of choice. The simplest proof uses the principle of dependent choice. The principle of dependent choice is the principle that justifies constructing a sequence by induction, where at each stage one must select the next term of the sequence from a non-empty set of possible choices which depends on the terms that have already been chosen. There is a slightly more complicated proof that uses only the countable axiom of choice. The countable axiom of choice states that each countable family of non-empty sets has a choice function. The countable axiom of choice is strictly weaker than the principle of dependent choice.

Let A be a subset of a topological space X. To say that A *is dense in* X means that the closure of A in X is all of X. For instance, the set of rational numbers is dense in $\mathbf{R}$. To say that A *is nowhere dense in* X means that the closure of A in X has empty interior in X. For instance, the set of integers is nowhere dense in $\mathbf{R}$. To say that p *is an isolated point of* A means that $p \in A$ and there exists a nhd U of p in X such that $U \cap A = \{p\}$. For instance, if $X = \mathbf{R}$ and $A = \{0\} \cup \{n^{-1} : n \in \mathbf{N}\}$, then for each $n \in \mathbf{N}$, n^{-1} is an isolated point of A but 0 is not an isolated point of A. If $p \in X$, then to say that p *is a limit point of* A means that for each nhd U of p, $U \cap (A \setminus \{p\})$ is non-empty. It is not hard to see that the closure of A is the disjoint union of the set of isolated points of A and the set of limit points of A. If $p \in X$, then to say that p *is a condensation point of* A means that for each nhd U of p, $U \cap A$ is uncountable. Clearly each condensation point of A is a limit point of A. But if $X = \mathbf{R}$ and $A = \{0\} \cup \{n^{-1} : n \in \mathbf{N}\}$, then 0 is a limit point of A but not a condensation point of A.

X11. Let S_n, S, and Σ be as above. For each $n \in \mathbf{N}$ and for each $s = (s_1, \ldots, s_n) \in S_n$, let $a(s) = \sum_{k=1}^{n} 2s_k 3^{-k}$, let $b(s) = a(s) + 3^{-n}$, and let $J(s)$ be the interval $[a(s), b(s)]$. Notice that $a(s,0) = a(s)$, $b(s,0) = a(s) + 3^{-(n+1)}$, $a(s,1) = a(s) + 2 \cdot 3^{-(n+1)}$, and $b(s,1) = b(s)$. Thus when we remove the open middle third from $J(s)$, we get $J(s,0) \cup J(s,1)$. Let $C_0 = [0,1]$. For each $n \in \mathbf{N}$, *due 3Th revised 13 Oct 05*

$$C_n = \bigcup_{s \in S_n} J(s).$$

Thus $C_1 = [0,1/3] \cup [2/3,1]$, $C_2 = [0,1/9] \cup [2/9,1/3] \cup [2/3,7/9] \cup [8/9,1]$, and in general, for each $n \in \mathbf{N}$, C_n consists of the union of the 2^n closed intervals, each of length 3^{-n}, that are obtained by removing the open middle third from each of the 2^{n-1} disjoint closed intervals that make up C_{n-1}. Let

$$C = \bigcap_{n \in \mathbf{N}} C_n.$$

The set C is called *the Cantor set.*

(a) Prove that C is closed and has empty interior in $\mathbf{R}$. Thus C is nowhere dense in $\mathbf{R}$.

Now for each $\sigma = (s_1, s_2, s_3, \ldots)$ in Σ, let $f(\sigma)$ be the unique point belonging to $\bigcap_{n \in \mathbf{N}} J(s_1, \ldots, s_n)$. (We can do this, by problem X9.)

(b) Prove that f is a one-to-one map from Σ onto C.

(c) Deduce that C is equinumerous to $[0,1]$. This is surprising, since by part (a), C is a topologically small subset of $[0,1]$. (Benoit Mandelbrot calls C the "Cantor dust" because its points are so sparsely scattered in the interval $[0,1]$.)

(d) Prove that each point of C is a condensation point of C. In particular, C has no isolated points. (Hint: Prove that for each $x \in C$ and each nhd U of x, $U \cap C$ contains a "scale model" of C.)

Remark. Let C_n and C be as in problem X11. Since C_n is the union of 2^n disjoint intervals each of which has length 3^{-n}, the total length of C_n is $(2/3)^n$. Since $C \subseteq C_n$ for each n, and since $(2/3)^n \to 0$ as $n \to \infty$, it seems reasonable to say that the total length of C is 0. (In spring quarter, we'll develop a theory of length that makes this rigorous.) This is another sense in which C is a very small set and another reason why it is surprising that C is equinumerous to $[0,1]$.

Remark. By modifying the construction of the Cantor set, it is possible to obtain other interesting Cantor-like sets, including ones which, while they still are nowhere dense in $\mathbf{R}$, have strictly positive total length. With this as motivation, in the next exercise we shall consider these Cantor-like sets in a very general setting. First we need a definition.

Definition. Let S_n, S, and Σ be as above. If (X,d) is a metric space, then by a *Cantor scheme*[4] in (X,d), we shall mean a family $(A(s))_{s \in S}$ of non-empty subsets of X, indexed by S, such that:

(a) $A(0)$ and $A(1)$ are disjoint;

(b) For each $n \in \mathbf{N}$ and each $s \in S_n$, $A(s_1, \ldots, s_n, 0)$ and $A(s_1, \ldots, s_n, 1)$ are disjoint and for $i = 0, 1$,

$$A(s_1, \ldots, s_n) \supseteq \overline{A(s_1, \ldots, s_n, i)};$$

(c) For each $\sigma = (s_1, s_2, s_3, \ldots) \in \Sigma$, $\operatorname{diam}(A(s_1, \ldots, s_n)) \to 0$ as $n \to \infty$.

[4] This is not standard terminology.

X12. Let S_n, S, and Σ be as above. Let (X,d) be a complete metric space and let $(A(s))_{s\in S}$ be a Cantor scheme in (X,d). For each $n \in \mathbf{N}$, let $B_n = \bigcup_{s\in S_n} A(s)$. Let $K = \bigcap_{n\in\mathbf{N}} B_n$. For each $\sigma = (s_1, s_2, s_3, \dots) \in \Sigma$, let $f(\sigma)$ be the unique point in $\bigcap_n A(s_1,\dots,s_n)$. (We can do this by problem X9.)

(a) Prove that h is a one-to-one map from Σ onto K.

(b) Deduce that K is equinumerous to the $[0,1]$.

Notation. Let (X,d) be a metric space. For each non-empty set $A \subseteq X$ and each $x \in X$, the *distance from x to A* is

$$d(x,A) = \inf\{\, d(x,a) : a \in A \,\},$$

by definition. (Warning: There need not exist $a \in A$ with $d(x,A) = d(x,a)$.)

X13. Let (X,d) be a metric space, let A be a non-empty subset of X, and let E be the closure of A. Prove that $E = \{\, x \in X : d(x,A) = 0 \,\}$. In particular, A is closed if and only if $A = \{\, x \in X : d(x,A) = 0 \,\}$.

X14. Let (X,d) be a metric space. Let A be a non-empty subset of X and define $f\colon X \to [0,\infty)$ by $f(x) = d(x,A)$.

(a) Prove that for all $x, y \in X$, we have

$$f(x) \le d(x,y) + f(y).$$

(Warning: Do not "calculate" with inf. Any inequalities you assert involving inf must be justified based on the definition of inf. Also, please do not use proof by contradiction when it is not needed and does not shorten the argument.)

(b) Deduce from part (a) that for all $x, y \in X$, we have $|f(x) - f(y)| \le d(x,y)$.

Definitions. Let (X,d) and (M,ρ) be metric spaces and let $g\colon X \to M$. To say that C *is a Lipschitz constant for g* means that $C \in [0,\infty)$ and that for all $x, y \in X$, we have $\rho(g(x), g(y)) \le C\, d(x,y)$. To say that g *is a Lipschitz function* means that there exists a Lipschitz constant for g.

Remark. Part (b) of problem X14 says that the function $x \mapsto d(x,A)$ is a Lipschitz function, with Lipschitz constant 1.

X15. Let (X,d) be a metric space, let A be a non-empty subset of X, let $r > 0$, and let

$$G = \{\, x \in X : d(x,A) < r \,\} \qquad \text{and} \qquad H = \{\, x \in X : d(x,A) \le r \,\}.$$

Prove that G is open and H is closed.

Definitions. Let X be a topological space. To say that A *is a $\mathscr{G}_\delta$ set in X* means that there exists a sequence (G_n) of open subsets of X such that $A = \bigcap_{n=1}^\infty G_n$. To say that A *is an $\mathscr{F}_\sigma$ set in X* means that there exists a sequence (F_n) of closed subsets of X such that $A = \bigcup_{n=1}^\infty F_n$. Less formally, a $\mathscr{G}_\delta$ set is a countable intersection of open sets and an $\mathscr{F}_\sigma$ set is a countable union of closed sets.

Example. The set of irrational numbers is a $\mathscr{G}_\delta$ set in $\mathbf{R}$ because it is equal to $\bigcap_{q\in\mathbf{Q}}(\mathbf{R}\setminus\{\, q \,\})$, a countable intersection of open sets. It can be shown that the set of irrational numbers is not an $\mathscr{F}_\sigma$ set. (This follows easily from the Baire category theorem, for those of you who know that theorem. We will discuss it later.)

X16. Let (X,d) be a metric space and let E be a closed subset of X. Prove that E is a $\mathscr{G}_\delta$ set in X. (By De Morgan's laws, it follows that if U is an open subset of X, then U is an $\mathscr{F}_\sigma$ set in X.)

X17. Let a and b be non-negative real numbers. The *geometric mean of a and b* is $\sqrt{ab}$. The *arithmetic mean of a and b* is $(a+b)/2$. Prove that these two means satisfy the inequality

$$\sqrt{ab} \le \frac{a+b}{2},$$

with equality if and only if $a = b$. (Do not use calculus. Just use simple algebra. If your proof is longer than one line, you are making it too complicated.)

X18. Let $\alpha \in (0, \infty)$. Define $f\colon (0,\infty) \to (0,\infty)$ by *due 4Th*

$$f(x) = \frac{1}{2}\left(x + \frac{\alpha}{x}\right).$$

(a) Prove that for each $x \in (0,\infty)$, we have $f(x) \geq \sqrt{\alpha}$, with equality if and only if $x = \sqrt{\alpha}$.

Now let $x_1 \in (\sqrt{\alpha}, \infty)$ and define $x_2, x_3, x_4, \ldots$ by the recursion formula $x_{n+1} = f(x_n)$.

(b) Prove that (x_n) is a strictly decreasing sequence in $(\sqrt{\alpha}, \infty)$ and that $x_n \to \sqrt{\alpha}$.

(c) Let $\varepsilon_n = x_n - \sqrt{\alpha}$, let $\beta = 2\sqrt{\alpha}$, and show that for each n,

$$\varepsilon_{n+1} = \frac{\varepsilon_n^2}{2x_n} < \frac{\varepsilon_n^2}{\beta}.$$

By induction, deduce that for each n,

$$\varepsilon_{n+1} < \beta\left(\frac{\varepsilon_1}{\beta}\right)^{2^n}.$$

(d) This is a good algorithm for computing square roots, because the recursion formula is simple and the convergence is swift. For example, suppose $\alpha = 3$ and $x_1 = 2$. Show that $\varepsilon_1/\beta < 10^{-1}$ and that therefore

$$\varepsilon_5 < 4 \cdot 10^{-16} \qquad \text{and} \qquad \varepsilon_6 < 4 \cdot 10^{-32}.$$

Definition. Let X be a topological space and let (x_n) be a sequence in X. To say that p *is a cluster point of* (x_n) *in* X means that $p \in X$ and that for each $N \in \mathbf{N}$ and for each nhd U of p in X, there exists $n \geq N$ such that $x_n \in U$.

Remark. Informally, p is a cluster point of (x_n) if and only if each tail of (x_n) meets each nhd of p.

Reminder. Each metric space is first countable.

X19. Let X be a topological space. Let (x_n) be a sequence in X and let $p \in X$.

(a) Suppose (x_n) has a subsequence $(y_k) = \left(x_{n_k}\right)$ which converges to p. Prove that p is a cluster point of (x_n).

(b) Conversely, suppose p is a cluster point of (x_n). Prove that if, in addition, X is first countable, then (x_n) has a subsequence which converges to p.

Definition. Let X be a set and let $\mathscr{B}$ and $\mathscr{C}$ be filter bases on X. To say that $\mathscr{C}$ *is finer than* $\mathscr{B}$ means that for each $B \in \mathscr{B}$, there exists $C \in \mathscr{C}$ such that $C \subseteq B$.

Example. Let X be a set, let (x_n) be a sequence in X, and let $(y_k) = \left(x_{n_k}\right)$ be a subsequence of (x_n). Then the filter base of tails of (y_k) is finer than the filter base of tails of (x_n). In this sense, the notion of finer filter base generalizes the notion of subsequence.

Example. Let X be a topological space, let $p \in X$, and let $\mathscr{B}$ be a nhd base at p for X. Then $\mathscr{B}$ is a filter base on X. Let $\mathscr{C}$ be another filter base on X. Then $\mathscr{C}$ converges to p if and only if $\mathscr{C}$ is finer than $\mathscr{B}$.

Definition. Let X be a topological space and let $\mathscr{B}$ be a filter base on X. To say that p *is a cluster point for* $\mathscr{B}$ *in* X means that $p \in X$ and that for each $B \in \mathscr{B}$ and for each nhd U of p in X, $B \cap U$ is non-empty.

Example. Let X be a topological space and let (x_n) be a sequence in X. Then clearly p is a cluster point for the filter base of tails of (x_n) if and only if p is a cluster point for (x_n).

Remark. Let X be a topological space, let $\mathscr{B}$ be a filter base on X, and let C be the set of cluster points for $\mathscr{B}$ in X. A moment's thought reveals that

$$C = \bigcap \left\{ \overline{B} : B \in \mathscr{B} \right\},$$

so C is closed because the intersection of any non-empty collection of closed sets is closed. In particular, the set of cluster points of any sequence is a closed set, because it is the same as the set of cluster points of the filter base of tails of the sequence.

The next exercise generalizes problem X19.

X20. Let X be a topological space, let $\mathscr{B}$ be a filter base on X, and let $p \in X$. Prove that p is a cluster point for $\mathscr{B}$ in X if and only if there exists a filter base $\mathscr{C}$ on X such that $\mathscr{C}$ is finer than $\mathscr{B}$ and $\mathscr{C}$ converges to p.

Reminder. Let (x_n) be a sequence in $[-\infty, \infty]$. For each n, let

$$a_n = \inf_{m \geq n} x_m \qquad \text{and} \qquad b_n = \sup_{m \geq n} x_m.$$

Then $a_n \leq a_{n+1}$ and $b_n \geq b_{n+1}$ for all n. Hence each of (a_n) and (b_n) has a limit in $[-\infty, \infty]$. By definition,

$$\liminf_{n \to \infty} x_n = \lim_{n \to \infty} a_n \qquad \text{and} \qquad \limsup_{n \to \infty} x_n = \lim_{n \to \infty} b_n.$$

Note that this is not the definition given in Rudin, *Principles of Mathemtical Analysis*, but I recommend that you use it because it is a more convenient working definition than the one in Rudin.

X21. Let (x_n) and (y_n) be bounded sequences of real numbers.

(a) Prove that for each n, we have

$$\sup_{m \geq n} (x_m + y_m) \leq \sup_{m \geq n} x_m + \sup_{m \geq n} y_m$$

and

$$\inf_{m \geq n} (x_m + y_m) \geq \inf_{m \geq n} x_m + \inf_{m \geq n} y_m.$$

(Warning: Do not "calculate" with sup or inf. Any inequalities you assert involving sup must be justified based on the definition of sup. Similarly for inf. Also, please do not use proof by contradiction when it is not needed and does not shorten the argument.)

(b) Prove that

$$\limsup(x_n + y_n) \leq \limsup x_n + \limsup y_n.$$

and

$$\liminf(x_n + y_n) \geq \liminf x_n + \liminf y_n.$$

(c) Prove that

$$\liminf(x_n + y_n) \leq \limsup x_n + \liminf y_n \leq \limsup(x_n + y_n)$$

and

$$\liminf(x_n + y_n) \leq \liminf x_n + \limsup y_n \leq \limsup(x_n + y_n).$$

(Hint: These inequalities follow from the ones in part (b) by routine calculations. For instance, try applying the second inequality in part (b) to the sequences $(-x_n)$ and $(x_n + y_n)$ instead of (x_n) and (y_n) and see what you get.)

(d) Use parts (b) and (c) to prove that if (x_n) is convergent, then

$$\limsup(x_n + y_n) = \lim x_n + \limsup y_n$$

and

$$\liminf(x_n + y_n) = \lim x_n + \liminf y_n.$$

X22. Give an example of two bounded sequences in $\mathbf{R}$ for which all of the inequalities in parts (a), (b), and (c) of problem X21 are strict. (Hint: There is an example in which each sequence just repeats its first four terms over and over.)

X23. Let (x_n) be a sequence in $\mathbf{R}$. For each n, let $s_n = (x_1 + \cdots + x_n)/n$ be the average of $x_1, \ldots, x_n$. *due 4Th*
Let $a = \liminf x_n$, $A = \liminf s_n$, $B = \limsup s_n$, and $b = \limsup x_n$.

(a) Prove that $B \le b$. (Hint: To show that $B \le b$, it suffices to show that for each $c \in (b, \infty)$, we have $B \le c$. This is even valid when $b = \infty$, since in that case there is nothing to show anyway!)

Similarly, $a \le A$. (Or this can be deduced by applying part (a) to the sequence $(-x_n)$ instead of (x_n).)

(b) Let $L \in [-\infty, \infty]$. Deduce from part (a) that if $x_n \to L$, then $s_n \to L$ too.

X24. Give an example of a sequence (x_n) in $\mathbf{R}$ such that $\limsup x_n = \infty$ and $\liminf x_n = -\infty$ but $(x_1 + \cdots + x_n)/n \to 0$.

Definition. Let (x_n) be a sequence in $\mathbf{R}$ and let $L \in \mathbf{R}$. To say that (x_n) *is Cesáro convergent to* L means that $(x_1 + \cdots + x_n)/n \to L$.

Remark. By problem X23(b), ordinary convergence implies Cesáro convergence. By problem X24, the converse does not hold in general.

I hope the last few exercises have convinced you that the definitions for lim inf and lim sup given above are good working definitions. The next exercise, in combination with problem X19, shows that these definitions are equivalent to the corresponding ones given in Rudin, *Principles of Mathemtical Analysis.*

X25. Let (x_n) be a sequence in $[-\infty, \infty]$.

(a) Let $a = \liminf x_n$ and let $b = \limsup x_n$. Prove that a and b are cluster points for (x_n) in $[-\infty, \infty]$.

(b) Let c be a cluster point for (x_n) in $[-\infty, \infty]$. Prove that $a \le c \le b$.

Thus a is the smallest cluster point for (x_n) and b is the largest cluster point for (x_n).

Definition. Let $\mathscr{B}$ be a filter base on $[-\infty, \infty]$. Then

$$\liminf \mathscr{B} = \sup \{ \inf B : B \in \mathscr{B} \} \qquad \text{and} \qquad \limsup \mathscr{B} = \inf \{ \sup B : B \in \mathscr{B} \},$$

by definition.

Example. Let (x_n) be a sequence in $[-\infty, \infty]$ and let $\mathscr{B}$ be the filter base of tails of (x_n). Then clearly

$$\liminf \mathscr{B} = \liminf x_n \qquad \text{and} \qquad \limsup \mathscr{B} = \limsup x_n.$$

The next exercise is a generalization of problem X25.

X26. Let $\mathscr{B}$ be a filter base on $[-\infty, \infty]$.

(a) Let $a = \liminf \mathscr{B}$ and let $b = \limsup \mathscr{B}$. Prove that a and b are cluster points for $\mathscr{B}$ in $[-\infty, \infty]$.

(b) Let c be a cluster point for $\mathscr{B}$ in $[-\infty, \infty]$. Prove that $a \le c \le b$.

Thus a is the smallest cluster point for $\mathscr{B}$ and b is the largest cluster point for $\mathscr{B}$.

Definition. Let (X, d) be a metric space and let $\mathscr{B}$ be a filter base on X. To say that $\mathscr{B}$ *is Cauchy* means that for each $\varepsilon > 0$, there exists $B \in \mathscr{B}$ such that $\operatorname{diam}(B) < \varepsilon$.

Remark. Let (X, d) be a metric space, let (x_n) be a sequence in X, and let $\mathscr{B}$ be the filter base of tails of (x_n). Then clearly $\mathscr{B}$ is Cauchy if and only if (x_n) is Cauchy.

The result in the next exercise may be viewed as a refinement of problem X9 and may be proved by a similar method.

X27. Let (X, d) be a complete metric space, so that each Cauchy sequence in X converges. Prove that each Cauchy filter base on X converges.

X28. Let (a_n) be a sequence of strictly positive real numbers. According to Theorem 3.37 in Rudin, *do before MT*
Principles of Mathematical Analysis, Third Edition,

$$\liminf \frac{a_{n+1}}{a_n} \le \liminf a_n^{1/n} \qquad \text{and} \qquad \limsup a_n^{1/n} \le \limsup \frac{a_{n+1}}{a_n}.$$

(a) Rudin gives a self-contained proof of these inequalities. Give a short alternative proof, based on problem X23.

(b) Compare the proof of Theorem 3.37 in Rudin with the solution of problem X23(a).

The next exercise really belongs right after the definition of lim inf and lim sup for a filter base on $[-\infty, \infty]$.

X29. Let $\mathscr{B}$ be a filter base on $[-\infty, \infty]$ and let $L \in [-\infty, \infty]$. Prove that $\mathscr{B} \to L$ if and only if $\liminf \mathscr{B} = L = \limsup \mathscr{B}$.

Reminder. Let X be a topological space and let $A \subseteq X$. To say that p *is a limit point of* A *in* X means that $p \in X$ and for each nhd U of p in X, $U \cap (A \setminus \{p\})$ is non-empty. It is clear that p is a limit point of A in X if and only if p belongs to the closure of $A \setminus \{p\}$ in X. Consequently, p is a limit point of A in X if and only if there exists a filter base $\mathscr{B}$ on $A \setminus \{p\}$ such that $\mathscr{B} \to p$.

X30. Let X be a first countable topological space.[5] Let $A \subseteq X$ and let $p \in X$. Prove that p is a limit point of A in X if and only if there is a sequence (a_n) in A such that $a_n \neq p$ for each n and $a_n \to p$ as $n \to \infty$.

Definition. Let X be a topological space. To say that X is T_1 means that for each $x \in X$ and each $y \in X$, if $x \neq y$, then there exists an open set G in X such that $x \in G$ and $y \notin G$.

Obviously each Hausdorff space is T_1. In particular, each metric space is T_1.

X31. Let X be a topological space. Prove that the following are equivalent:
- **(a)** X is T_1.
- **(b)** Each singleton in X is closed.
- **(c)** Each finite subset of X is closed.

X32. Let X be a first countable T_1 space.[6] Let $A \subseteq X$ and let p be a limit point of A in X. Prove that there is a sequence (a_n) of distinct points in $A \setminus \{p\}$ such that $a_n \to p$ as $n \to \infty$. Deduce that for each nhd U of p in X, $U \cap A$ is infinite.

The next three exercises may be solved with the help of problem X10.

X33. Prove that $[0,1]^2$ is equinumerous to $[0,1]$.

X34. Let $n \in \mathbf{N}$. Prove that $[0,1]^n$ is equinumerous to $[0,1]$.

X35. Prove that $[0,1]^{\mathbf{N}}$ is equinumerous to $[0,1]$.

Remark. In late 1873, Georg Cantor asked whether it was possible to define a one-to-one correspondence between $[0,1]$ and $\mathbf{N}$. He believed that this would not be possible and within a few days, he succeeded in proving this. In doing so, he became the first person in history to realize that just because two sets are both infinite, it does not follow that they have the same number of elements. About a month later, he asked whether it was possible to define a one-to-one correspondence between $[0,1]^2$ and $[0,1]$. Again, he believed that this would not be possible. This time it took him about three and a half years to answer the question and the answer that he found astonished him: As you were asked to show in problem X33, he found that such a one-to-one correspondence can be defined! This initially led him to believe that he had shown that the concept of dimension was meaningless. However, Cantor's friend Richard Dedekind pointed out that the one-to-one correspondence that Cantor had found was not continuous and conjectured that a one-to-one correspondence would preserve dimension if both it and its inverse were continuous. In the years that followed, many mathematicians, including Cantor himself, tried to prove this. The first correct proof was published by the Dutch mathematician L. E. J. Brouwer in 1911, some 34 years after Dedekind formulated the conjecture.

X36. Let $z \in \mathbf{C}$ with $z \neq 1$. For each $n \in \mathbf{N}$, let $S_n = \sum_{k=0}^{n-1} z^k$, let $T_n = \sum_{k=1}^{n} k z^k$, and let $U_n = \sum_{k=1}^{n} k^2 z^k$. As we know, *do before MT*

$$S_n = \frac{1 - z^n}{1 - z}.$$

This is a simpler expression for S_n than the expression that we used to define S_n, in the sense that it involves far fewer arithmetic operations when n is large.

[5] Remember that each metric space is first countable.

[6] Remember that each metric space is first countable and T_1.

(a) Find a simpler expression[7] for T_n. Then show that if $|z| < 1$, then the infinite series $\sum_{k=1}^{\infty} kz^k$ converges absolutely and find its sum.

(b) Find a simpler expression[8] for U_n. Then show that if $|z| < 1$, then the infinite series $\sum_{k=1}^{\infty} k^2z^k$ converges absolutely and find its sum.

Example. Here is an application of the results of problem X36. Suppose you toss a coin repeatedly. Let p be the probability that the coin comes up heads on any given toss. Assume that $p \in (0,1)$, but don't assume that $p = 1/2$. (For instance, perhaps there is a lump of clay stuck to one side of the coin.) Let $q = 1 - p$. Thus q is the probability that the coin comes up tails on any given toss. Let N be the number of tosses until the coin first comes up heads. For each $k \in \mathbf{N}$, $P(N = k)$ denotes the probability that N has the value k. We have $P(N = 1) = p$, $P(N = 2) = qp$, $P(N = 3) = q^2p$, and so on. In general, $P(N = k) = q^{k-1}p$, for each $k \in \mathbf{N}$. Using the formula for the sum of a geometric series, you can easily check that $\sum_{k\in\mathbf{N}} P(N = k) = 1$. So with probability 1, the coin comes up heads after only a finite number of tosses. Now $E(N)$ denotes the expected value of N, or in other words, the average value of N. We have $E(N) = \sum_{k\in\mathbf{N}} kP(N = k)$. Using the result from problem X36(a), you can easily check that $E(N) = 1/p$. (For instance, if $p = 1/2$, the average number of tosses until the coin first comes up heads is 2.) Similarly, $E(N^2) = \sum_{k\in\mathbf{N}} k^2P(N = k)$. Using the result from problem X36(b), you can easily find the value of $E(N^2)$. Let $\nu = E(N)$. Then $E[(N-\nu)^2]$ is called the variance of N. The square root of the variance of N is called the standard deviation of N. It is a measure of how much N deviates from its average value, on the average. We have $E[(N-\nu)^2] = E(N^2 - 2\nu N + \nu^2) = E(N^2) - 2\nu E(N) + \nu^2 = E(N^2) - 2\nu^2 + \nu^2 = E(N^2) - \nu^2$. You should find that the variance of N is q/p^2, so the standard deviation of N is $\sqrt{q}/p$. (For instance, if $p = 1/2$, then $q = 1/2$ too and the variance of N is 2, so the standard deviation of N is $\sqrt{2}$.)

X37. Let (c_k) be a sequence in $(0,\infty)$. Suppose $\sum_{k=1}^{\infty} c_k$ converges. *do before MT*

(a) (*The limit comparison test for convergence.*) Suppose (a_k) is a sequence in $\mathbf{C}$ and

$$\limsup_{k\to\infty} \frac{|a_k|}{c_k} < \infty.$$

Prove that $\sum_{k=1}^{\infty} a_k$ converges absolutely.[9]

(b) Let $b_k = c_k/\sqrt{r_k}$, where $r_k = \sum_{\ell=k}^{\infty} c_\ell$. Prove that

$$\lim_{k\to\infty} \frac{b_k}{c_k} = \infty$$

but that[10] $\sum_{k=1}^{\infty} b_k$ converges.[11]

X38. Let (d_k) be a sequence in $(0,\infty)$. Suppose $\sum_{k=1}^{\infty} d_k$ diverges. *do before MT*

(a) (*The limit comparison test for divergence.*) Suppose (a_k) is a sequence in $[0,\infty)$ and

$$\liminf_{k\to\infty} \frac{a_k}{d_k} > 0.$$

Prove that $\sum_{k=1}^{\infty} a_k$ diverges.[12]

---

[7] Hint: Consider $T_n - zT_n$ and use it to express T_n in terms of S_n. Another way involves the derivative dT_n/dz, but please don't use that method. It is not shorter and it is less elementary. Remember, we are pretending for the time being that we don't know calculus!

[8] Hint: Consider $U_n - zU_n$ and use it to express U_n in terms of T_n and S_n.

[9] Colloquially, this says that if the terms of a given series tend to zero as fast as the terms of some absolutely convergent series, then the given series is itself absolutely convergent.

[10] Hint: Show that $b_k < 2(\sqrt{r_k} - \sqrt{r_{k+1}})$.

[11] Colloquially, this implies that given any convergent series with positive terms, there is another convergent series whose terms are positive and tend to zero more slowly than the terms of the given series. In this sense, there is no most slowly convergent series with positive terms.

[12] Colloquially, this implies that if the terms of a given series are positive and tend to zero more slowly than the terms of some divergent series with positive terms, then the given series is itself divergent.

(b) Let $b_k = d_k / s_k$, where $s_k = \sum_{\ell=1}^{k} d_\ell$. Prove that

$$\lim_{k \to \infty} \frac{b_k}{d_k} = 0$$

but that[13] $\sum_{k=1}^{\infty} b_k$ diverges.[14]

[13] Hint: Show that if $n > m$, then

$$b_{m+1} + \cdots + b_n \geq 1 - \frac{s_m}{s_n}.$$

[14] Colloquially, this implies that given any divergent series with positive terms tending to zero, there is another divergent series whose terms are positive and tend to zero faster than the terms of the given series. In this sense, there is no most slowly divergent series with positive terms.

X39. Imagine a pile of n identical books stacked on a table in such a way that the top book overhangs the one below by as much as possible (clearly the overhang would be half a booklength), the next book from the top overhangs the one below it by as much as possible, and so on, and the bottom book overhangs the edge of the table by as much as possible. Let S_n be the amount (in booklengths) by which the top book overhangs the edge of the table. *due 6Th*

(a) Find a general formula for S_n.

(b) Determine how large n must be so that no part of the top book is over the table.

(c) Show that in principle, if we can use as many books as we like, then we can arrange things so that the top book overhangs the edge of the table by as many booklengths as we like.

X40.

(a) Prove that the series

$$\sum_{k=3}^{\infty} \frac{1}{(\log k)^{\log k}}$$

converges.

(b) Prove that the series

$$\sum_{k=3}^{\infty} \frac{1}{(\log k)^{\log \log k}}$$

diverges. (Hint: The inequality $\log x \leq x/e$ may help.[15])

X41. (*The monotone convergence theorem for series.*) Suppose that for each $k \in \mathbf{N}$, we have *due 6Th*

$$0 \leq a(1,k) \leq a(2,k) \leq a(3,k) \leq \cdots$$

and $a(j,k) \to b(k)$ as $j \to \infty$. Prove that

$$\sum_{k=1}^{\infty} a(j,k) \to \sum_{k=1}^{\infty} b(k) \qquad \text{as } j \to \infty.$$

(If $\sum_{k=1}^{\infty} b(k) < \infty$, then this could be deduced from Tannery's theorem. But please give a self-contained proof that works whether or not $\sum_{k=1}^{\infty} b(k)$ is finite. Warning: Do not take it for granted that if $n \in \mathbf{N}$, then $\sum_{k=1}^{n} a(j,k) \to \sum_{k=1}^{n} b(k)$ as $j \to \infty$. This is not covered by what we did in class, because $b(k)$ may be ∞ for some k, and even $a(j,k)$ may be ∞ for some j and k. So if you want to use this fact, you should prove it in detail. Note also that if $\varepsilon \in (0,\infty)$, you cannot conclude that $b(k) - \varepsilon < b(k)$ unless $b(k) < \infty$.)

Remark. Suppose $f(j,k) \geq 0$ for all $j,k \in \mathbf{N}$. Then

$$\sum_{j=1}^{\infty}\sum_{k=1}^{\infty} f(j,k) = \sum_{k=1}^{\infty}\sum_{j=1}^{\infty} f(j,k) \tag{1}$$

whether or not either side is finite. We have seen one proof of this in class, based on Tannery's theorem. For another proof, we can apply problem X41. Let $a(m,k) = \sum_{j=1}^{m} f(j,k)$ and let $b(k) = \sum_{j=1}^{\infty} f(j,k)$. Then $0 \leq a(1,k) \leq a(2,k) \leq a(3,k) \leq \cdots$ and $a(m,k) \to b(k)$ as $m \to \infty$. Hence

$$\sum_{k=1}^{\infty} a(m,k) \to \sum_{k=1}^{\infty} b(k)$$

as $m \to \infty$, by problem X41. But $\sum_{k=1}^{\infty} b(k) = \sum_{k=1}^{\infty} f(j,k)$ and for each $m \in \mathbf{N}$, $\sum_{k=1}^{\infty} a(m,k) = \sum_{k=1}^{\infty}\sum_{j=1}^{m} f(j,k) = \sum_{j=1}^{m}\sum_{k=1}^{\infty} f(j,k)$. Letting $m \to \infty$, we get (1).

[15] The proof of this inequality relies on calculus, so will justify it fully later. For now let's give a preview of the way to prove it. The function $x \mapsto \log x$ is concave down. Hence its graph lies below each of its tangent lines. But the line with equation $y = x/e$ is tangent to the graph of $y = \log x$ at the point $(e,1)$.

X42.

(a) (*Fatou's lemma for series.*) Suppose $c(j,k) \in [0,\infty)$ for all $j,k \in \mathbf{N}$. Let $b(k) = \liminf_{j\to\infty} c(j,k)$. Prove that

$$\sum_{k=1}^{\infty} b(k) \le \liminf_{j\to\infty} \sum_{k=1}^{\infty} c(j,k).$$

(Hint: Apply problem X41 with $a(j,k) = \inf_{j' \ge j} c(j',k)$.)

(b) (*A generalization of Tannery's theorem.*) Suppose that

$$a(j,k) \to b(k) \text{ in } \mathbf{C} \text{ as } j \to \infty \text{ for each } k \in \mathbf{N},$$

$$|a(j,k)| \le L(j,k) \text{ for all } j,k \in \mathbf{N},$$

$$\sum_{k=1}^{\infty} L(j,k) < \infty \text{ for each } j \in \mathbf{N},$$

$$L(j,k) \to M(k) \text{ as } j \to \infty \text{ for each } k \in \mathbf{N},$$

$$\sum_{k=1}^{\infty} M(k) < \infty,$$

and

$$\sum_{k=1}^{\infty} L(j,k) \to \sum_{k=1}^{\infty} M(k) \qquad \text{as } j \to \infty.$$

Prove that as $j \to \infty$,

$$\sum_{k=1}^{\infty} |a(j,k) - b(k)| \to 0$$

and

$$\sum_{k=1}^{\infty} a(j,k) \to \sum_{k=1}^{\infty} b(k).$$

(Hint: Note that $L(j,k) + M(k) - |a(j,k) - b(k)| \ge 0$ and as $j \to \infty$,

$$L(j,k) + M(k) - |a(j,k) - b(k)| \to 2M(k).$$

Apply part (a). You may find that problem X21(b) helps.)

Example. Let us mention an application of problem X42(b) to probability theory. For each $j \in \mathbf{N}$, let $(p(j,k))_{k\in\mathbf{N}}$ be a probability distribution on $\mathbf{N}$. This means that $p(j,k) \ge 0$ and $\sum_{k=1}^{\infty} p(j,k) = 1$. Suppose that $p(j,k) \to \pi(k)$ as $j \to \infty$. Then $\sum_{k=1}^{\infty} \pi(k) \le 1$. (This obviously follows from problem X42(a).) Suppose that $\sum_{k=1}^{\infty} \pi(k) = 1$, so that $(\pi(k))_{k\in\mathbf{N}}$ is also a probability distribution on $\mathbf{N}$. Then as $j \to \infty$,

$$\sum_{k=1}^{\infty} |p(j,k) - \pi(k)| \to 0$$

and

$$\sup_{A \subseteq \mathbf{N}} \left| \sum_{k\in A} p(j,k) - \sum_{k\in A} \pi(k) \right| \to 0.$$

It is clear that the second conclusion follows from the first. As for the first, it follows from problem X42(b) if we take $a(j,k) = L(j,k) = p(j,k)$ and $b(k) = M(k) = \pi(k)$.

Remark. Suppose $0 \le a(j,k) \le b(k)$ and as $j \to \infty$, $a(j,k) \to b(k)$. Then $\sum_{k=1}^{\infty} a(j,k) \to \sum_{k=1}^{\infty} b(k)$ as $j \to \infty$. This may be viewed as a generalization of problem X41. It obviously follows from problem X42(a).

Remark. Suppose $0 \le a(j,k) \le b(k)$ and as $j \to \infty$, $a(j,k) \to b(k)$. Suppose also that $\sum_{k=1}^{\infty} a(j,k)$ has a limit in $[0,\infty)$ as $j \to \infty$. Then $\sum_{k=1}^{\infty} b(k) < \infty$. This obviously follows from problem X42(a).

X43. Suppose that $\sum_{k=0}^{\infty} a_k$ and $\sum_{\ell=0}^{\infty} b_\ell$ both converge absolutely. For $m = 0, 1, 2, \dots$, let

$$c_m = \sum_{k=0}^{m} a_k b_{m-k}.$$

Prove that $\sum_{m=0}^{\infty} c_m$ converges absolutely.

X44. Suppose that $\sum_{k=1}^{\infty} a_k$ converges absolutely and that (b_k) is bounded. Prove that $\sum_{k=1}^{\infty} a_k b_k$ converges.

X45. (*Toeplitz's Lemma.*) Let $a(j,k) \in \mathbf{C}$ for all $j, k \in \mathbf{N}$. Suppose that for each $k \in \mathbf{N}$, *due 7Th*

$$a(j,k) \to 0 \qquad \text{as } j \to \infty. \tag{2}$$

Suppose also that

$$M = \sup_{j \in \mathbf{N}} \sum_{k=1}^{\infty} |a(j,k)| < \infty. \tag{3}$$

(a) Let (w_k) be a sequence in $\mathbf{C}$. Suppose that $w_k \to 0$ as $k \to \infty$. Prove that

$$\sum_{k=1}^{\infty} |a(j,k) w_k| \to 0 \qquad \text{as } j \to \infty$$

and therefore

$$\sum_{k=1}^{\infty} a(j,k) w_k \to 0 \qquad \text{as } j \to \infty.$$

(Hint: Don't try to use Tannery's theorem. It does not apply here. Instead, write a proof from scratch.)

(b) Now suppose in addition that $\alpha \in \mathbf{C}$ and that

$$\sum_{k=1}^{\infty} a(j,k) \to \alpha \qquad \text{as } j \to \infty. \tag{4}$$

Let $z \in \mathbf{C}$ and let (z_k) be a sequence in $\mathbf{C}$ such that $z_k \to z$ as $k \to \infty$. Prove that

$$\sum_{k=1}^{\infty} a(j,k) z_k \to \alpha z \qquad \text{as } j \to \infty.$$

(Hint: This is a simple corollary of part (a).)

Remark. Suppose that for each $j \in \mathbf{N}$, $(a(j,k))_{k \in \mathbf{N}}$ is a probability distribution on $\mathbf{N}$. Then (3) holds with $M = 1$. Hence if (2) also holds, then whenever $z_k \to z$ in $\mathbf{C}$, we have

$$\sum_{j=1}^{\infty} a(j,k) z_k \to z \qquad \text{as } j \to \infty.$$

Colloquially, we may say that weighted averages of the z_k's tend to z as long as the weights for each fixed k tend to 0. Obviously, this generalizes part of our earlier homework problem on Cesàro averages (problem X23).

Remark. Not only does Tannery's theorem not apply in problem X45, but in fact, part (b) this exercise provides an example where the limit of the sum is not the sum of the limit (unless $\alpha = 0$ or $z = 0$) and indeed we would not want it to be.

X46. (*Kronecker's Lemma.*) Let (b_k) be an increasing sequence in $(0,\infty)$ such that $b_k \to \infty$ as $k \to \infty$. Let (y_k) be a sequence in $\mathbf{C}$ and suppose that the series

$$\sum_{k=1}^{\infty} \frac{y_k}{b_k}$$

converges. Prove that

$$\frac{1}{b_n} \sum_{k=1}^{n} y_k \to 0 \qquad \text{as } n \to \infty.$$

(Hint: Let $z = \sum_{k=1}^{\infty} y_k/b_k$, let $z_0 = 0$, and for $n \geq 1$, let $z_n = \sum_{k=1}^{n} y_k/b_k$. Then $y_k = b_k(z_k - z_{k-1})$. Use this to rewrite $(1/b_n)\sum_{k=1}^{n} y_k$ and thereby show, using Toeplitz's lemma, that it tends to $z - z$.)

Remark. A particular case of Kronecker's lemma is that if

$$\sum_{k=1}^{\infty} \frac{y_k}{k}$$

converges, then

$$\frac{1}{n} \sum_{k=1}^{n} y_k \to 0 \qquad \text{as } n \to \infty.$$

Kronecker's lemma, and especially this particular case of it, is used in probability theory, in the proof of the strong law of large numbers. Other particular cases of Kronecker's lemma are used in the proofs of results about the rate of convergence in the strong law of large numbers.

X47. (*A variation on Toeplitz's Lemma.*) Let $a(j,k) \in [0,\infty)$ for all $j,k \in \mathbf{N}$. Suppose that for each $k \in \mathbf{N}$,

$$a(j,k) \to 0 \qquad \text{as } j \to \infty.$$

Suppose also that for each j, we have $\sum_{k=1}^{\infty} a(j,k) < \infty$. Let

$$\alpha = \liminf_{j\to\infty} \sum_{k=1}^{\infty} a(j,k) \qquad \text{and} \qquad \beta = \limsup_{j\to\infty} \sum_{k=1}^{\infty} a(j,k)$$

and suppose that $\alpha, \beta \in [0,\infty)$. Let (x_k) be a bounded sequence in $\mathbf{R}$ and let

$$a = \liminf_{k\to\infty} x_k \qquad \text{and} \qquad b = \limsup_{k\to\infty} x_k.$$

For each j, let

$$s_j = \sum_{k=1}^{\infty} a(j,k) x_k.$$

(Be sure to explain why this sum is defined for each j.) Let

$$A = \liminf_{j\to\infty} s_j \qquad \text{and} \qquad B = \limsup_{j\to\infty} s_j.$$

Prove that $B \leq b\beta$. Similarly $a\alpha \leq A$.

Remark. Evidently, problem X47 is another generalization of part of our earlier homework problem on Cesàro averages (problem X23).

X48. Let (a_m) be a sequence in $\mathbf{C}$ and let $A \in \mathbf{C}$. Let (M_n) be a sequence of integers satisfying

$$0 = M_1 < M_2 < M_3 < \cdots .$$

For each n, let

$$c_n = \max\left\{ \left|a_{M_n+1} + \cdots + a_m\right| : M_n + 1 \le m \le M_{n+1} \right\}.$$

Prove that the series

$$a_1 + a_2 + a_3 + \ldots$$

converges and its sum is A if and only if $c_n \to 0$ and the series

$$\left(a_1 + \cdots + a_{M_2}\right) + \left(a_{M_2+1} + \cdots + a_{M_3}\right) + \left(a_{M_3+1} + \cdots + a_{M_4}\right) + \cdots$$

converges and its sum is A. In other words, letting

$$b_n = \sum_{m=M_n+1}^{M_{n+1}} a_m,$$

prove that the series $\sum_{m=1}^{\infty} a_m$ converges and its sum is A if and only $c_n \to 0$ and the series $\sum_{n=1}^{\infty} b_n$ converges and its sum is A.

X49. Consider the series

$$S = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + - \cdots$$

and its rearrangement

$$T = 1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + + - \cdots . \tag{5}$$

In class, we showed informally that $T = 3S/2$. Prove this rigorously. (Hint: The result of problem X48 should help. By definition,

$$S = \sum_{m=0}^{\infty} \frac{(-1)^m}{m+1}.$$

Show that we also have

$$S = \sum_{n=0}^{\infty} \left(\frac{1}{2n+1} - \frac{1}{2n+2} \right) \tag{6}$$

Use this to get an expression for $S/2$. Add this expression for $S/2$ to the expression for S in(6). You should get

$$\frac{3S}{2} = \sum_{n=0}^{\infty} \left(\frac{1}{4n+1} + \frac{1}{4n+3} - \frac{1}{2n+2} \right).$$

Explain how it follows from this that the series in (5) converges and its sum is $3S/2$.)

X50. Let E be a set, let $f, g \colon E \to \mathbf{C}$, let $\mathscr{B}$ be a filter base on E, and let $a, b, c \in \mathbf{C}$. Prove the following: *due 7Th*

- **(a)** If $f \to a$ along $\mathscr{B}$, then $cf \to ca$ along $\mathscr{B}$.
- **(b)** If $f \to a$ along $\mathscr{B}$ and $g \to b$ along $\mathscr{B}$, then $f + g \to a + b$ along $\mathscr{B}$.
- **(c)** If $g \to b$ along $\mathscr{B}$, then there exists $B \in \mathscr{B}$ such that g is bounded on B.
- **(d)** If $f \to 0$ along $\mathscr{B}$ and if there exists $B \in \mathscr{B}$ such that g is bounded on B, then $fg \to 0$ along $\mathscr{B}$.
- **(e)** If $f \to a$ along $\mathscr{B}$ and $g \to b$ along $\mathscr{B}$, then $fg \to ab$ along $\mathscr{B}$.
- **(f)** Suppose $g \to b$ along $\mathscr{B}$ and $b \ne 0$. Let $E_1 = \{ x \in E : g(x) \ne 0 \}$ and let

  $$\mathscr{B}_1 = \{ B \cap E_1 : B \in \mathscr{B} \}.$$

  Then $\mathscr{B}_1$ is a filter base on E_1 and

  $$\frac{1}{g} \to \frac{1}{b} \quad \text{along } \mathscr{B}_1.$$

X51. Let X and Y be topological spaces, let S be a dense subset of X, and let $f\colon S \to Y$ be continuous. Let g be the set of all points (p,q) belonging to $X \times Y$ such that for each nhd V of q in Y, there exists a nhd U of p in X such that $f[U \cap S] \subseteq V$. Let D be the set of all $p \in X$ such that there exists $q \in Y$ such that $(p,q) \in g$. *due 8Th*

- **(a)** For each $p \in X$, let $\mathscr{U}_p$ be the collection of nhds of p in X, let $\mathscr{B}_p = \{\, U \cap S : U \in \mathscr{U}_p \,\}$, and let $\mathscr{C}_p = \{\, f[B] : B \in \mathscr{B}_p \,\}$. Since S is dense in X, each $p \in X$ belongs to the closure of S, so $\mathscr{B}_p$ is a filter base on S, and hence $\mathscr{C}_p$ is a filter base on Y. Verify that g is the set of all points $(p,q) \in X \times Y$ such that $\mathscr{C}_p \to q$, or in other words, such that $f(x) \to q$ as x runs along $\mathscr{B}_p$.
- **(b)** Prove that $S \subseteq D$ and that if Y is Hausdorff, then g is a function[16] from D to Y and g is an extension of f.
- **(c)** Suppose Y is Hausdorff and regular.[17] Prove that g is continuous.
- **(d)** Suppose Y is a metric space, with metric d say. Then of course Y is Hausdorff and regular, so the conclusions of parts (b) and (c) hold. Let

$$E = \{\, p \in X : \text{the filter base } \mathscr{C}_p \text{ is Cauchy} \,\}.$$

For each $n \in \mathbf{N}$, let

$$G_n = \{\, p \in X : p \text{ has a nhd } U \text{ in } X \text{ such that } \operatorname{diam}(f[U \cap S]) \le 1/n \,\}.$$

Prove that $D \subseteq E$, that each G_n is open, and that $E = \bigcap_{n=1}^{\infty} G_n$. In particular, E is a $\mathscr{G}_\delta$ set.
- **(e)** Suppose Y is a complete metric space, with metric d. Prove that $D = E$. In particular, D is a $\mathscr{G}_\delta$ set.
- **(f)** Continue to suppose that Y is a complete metric space, with metric d. Now suppose in addition that X is also a metric space, with metric ρ say, and that f is uniformly continuous.[18] Prove that $D = X$ and that g is uniformly continuous.

Remark. It is worth noticing that in problem X51(c), g is the maximal continuous extension of f, in a sense that we shall now explain. Suppose that $S \subseteq T \subseteq X$ and $h\colon T \to Y$ is a continuous extension of f. Then $T \subseteq D$ and g is an extension of h. In particular, for each T satisfying $S \subseteq T \subseteq D$, the restriction of g to T is the unique continuous extension of f to T.

Proof. Let $p \in T$. Then h is continuous at p. We wish to show that $p \in D$ and $g(p) = h(p)$. Let $q = h(p)$ and let V be a nhd of q in Y. Since h is continuous at p, there exists a nhd N of p in T such that $h[N] \subseteq V$. But $N = U \cap T$ for some nhd U of p in X. Let $x \in U \cap S$. Then $x \in U \cap T = N$, so $h(x) \in V$. But $h(x) = f(x)$, because $x \in S$ and h is an extension of f. Thus $f(x) \in V$. This holds for each $x \in U \cap S$. Therefore $f[U \cap S] \subseteq V$. We have shown that for each nhd V of q in Y, there exists a nhd U of p in X such that $f[U \cap S] \subseteq V$. Therefore $(p,q) \in g$. In other words, $p \in D$ and $g(p) = q$. But $q = h(p)$. Therefore $g(p) = h(p)$. ∎

X52. Let $E \subseteq \mathbf{R}$ and let $f\colon E \to \mathbf{R}$ be increasing.[19] Let $I = f[E]$. Suppose that I is an interval. Prove that f is continuous. Base your proof directly on the definition of a continous function. (Warning: E need not be an interval. Hint: Let $x_0 \in E$. We wish to show that f is continuous at x_0. Let $y_0 = f(x_0)$. Now either y_0 is in the interior of I, or y_0 is the left endpoint of I, or y_0 is the right endpoint of I. You may find that it clarifies matters to consider these three cases separately. Of course, if I does not contain its left endpoint, then the second case cannot occur, but you need not make special allowances for this possibility. You just need to make sure you cover all cases that might occur. Whether or not they actually do occur is not important. So likewise, you need not make special allowances for the possibility that I does not contain its right endpoint.) *due 8Th*

---

[16] As is commonly done, for the purposes of this problem, we consider a function to be the same thing as the set of ordered pairs that is its graph.

[17] To say that a topological space Z is *regular* means that each point in Z has a neighbourhood base consisting of closed sets. For instance, if Z is a metric space, then Z is regular, because for each $z \in Z$, the closed balls $B[z,r]$, $0 < r < \infty$, form a neighbourhood base at z consisting of closed sets.

[18] To say that f is uniformly continuous means that for each $\varepsilon > 0$, there exists $\delta > 0$ such that for all $x, x' \in S$, if $\rho(x,x') < \delta$, then $d\big(f(x), f(x')\big) < \varepsilon$.

[19] To say that f *is increasing* means that for all $x, x' \in E$, if $x \le x'$, then $f(x) \le f(x')$. And by the way, to say that f *is strictly increasing* means that for all $x, x' \in E$, if $x < x'$, then $f(x) < f(x')$. So for instance, even a constant function is increasing, though not strictly increasing of course.

Remark. In general, a function whose range is an interval need not be continuous, as simple examples show. The assumption in problem X52 that f is increasing matters. Of course it would also work to assume that f is decreasing.

Remark. Let $a \in (1,\infty)$. From Hölder's theorem on ordered groups, we know that there is a unique increasing function $f\colon (0,\infty) \to \mathbf{R}$ such that $f(a) = 1$ and for all $x, x' \in (0,\infty)$, $f(xx') = f(x) + f(x')$. Furthermore, the range of f is all of $\mathbf{R}$ and f is strictly increasing. By definition, for each $x \in (0,\infty)$, $\log_a x$ is $f(x)$. In view of problem X52, the function $x \mapsto \log_a x$ is continuous on $(0,\infty)$. Since f is strictly increasing, f is one-to-one, so f^{-1} is a function. Since the domain of f is $(0,\infty)$ and the range of f is $\mathbf{R}$, the domain of f^{-1} is $\mathbf{R}$ and the range of f^{-1} is $(0,\infty)$. Since f is strictly increasing, so is f^{-1}. It is not hard to check that f^{-1} is the unique increasing function $g\colon \mathbf{R} \to \mathbf{R}$ such that $g(1) = a$ and for all $y, y' \in \mathbf{R}$, $g(y + y') = g(y)g(y')$. By definition, for each $y \in \mathbf{R}$, a^y is $f^{-1}(y)$. In view of problem X52, the function $y \mapsto a^y$ is continuous on $\mathbf{R}$. Of course if $a \in (0,1)$ instead, then by definition, for each $x \in (0,\infty)$, $\log_a x$ is $-\log_{1/a} x$, and for each $y \in \mathbf{R}$, a^y is $(1/a)^{-y}$. Trivially, if $a = 1$, then by definition, for each $y \in \mathbf{R}$, $a^y = 1$. (And of course, if $a = 1$, then $\log_a$ is undefined.)

X53. Let I be an interval in $\mathbf{R}$ and let $f\colon I \to \mathbf{R}$. Suppose that f is midpoint-convex.[20]

(a) Prove for each dyadic rational number[21] t between 0 and 1, and for all $x, x' \in I$, we have

$$f((1-t)x + tx') \le (1-t)f(x) + tf(x'). \tag{7}$$

(Hint: For $n = 0, 1, 2, \ldots$, let D_n be the set of numbers of the form $k/2^n$ where $k \in \{0, 1, \ldots, 2^n\}$. Prove by induction on n that for each integer $n \ge 0$, for each $t \in D_n$, for all $x, x' \in I$, (7) holds.)

(b) Now suppose in addition that f is continuous. Prove that f is convex.[22]

Remark. Let $a \in (0,\infty)$. Recall that for all $u, v \in [0,\infty)$, $\sqrt{uv} \le (u+v)/2$. Hence $a^{(x+x')/2} = (a^x a^{x'})^{1/2} \le (a^x + a^{x'})/2$. Thus the function $x \mapsto a^x$ is midpoint convex on $\mathbf{R}$. But as we have seen above, this function is continuous on $\mathbf{R}$. Therefore, by problem X53, it is convex on $\mathbf{R}$.

Definition. Let X be a topological space, let $p \in X$, and let $E \subseteq X$. To say that p *is near* E means that p belongs to the closure of E.

Example. Let X be a metric space, with metric d. Let $p \in X$ and let $E \subseteq X$. Then p is near E iff $d(p, E) = 0$.

X54. Let X and Y be topological spaces, let $f\colon X \to Y$, and let $p \in X$. Prove that f is continuous at p if and only if for each subset $E \subseteq X$, if p is near E, then $f(p)$ is near $f[E]$.

Let X be a topological space and let $f\colon X \to [-\infty,\infty]$. To say that f *is lower semicontinuous* means that for each $y \in \mathbf{R}$, the set $\{f > y\} = \{x \in X : f(x) > y\}$ is open in X. To say that f *is upper semicontinuous* means that for each $y \in \mathbf{R}$, the set $\{f < y\} = \{x \in X : f(x) < y\}$ is open in X. It is easy to see that f is continuous if and only if f is both lower and upper semicontinous.

Let X be a set and let $A \subseteq X$. The *indicator function for* A is the function 1_A on X defined by

$$1_A(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{if } x \in X \setminus A. \end{cases}$$

X55. Let X be a topological space and let $A \subseteq X$.

(a) Prove that 1_A is lower semicontinuous iff A is open.

(b) Prove that 1_A is upper semicontinuous iff A is closed.

(c) Let $D = \{p \in X : 1_A \text{ is not continuous at } p\}$. Prove that D is the frontier of A.

---

[20] To say that f *is midpoint-convex* means that for all $x, x' \in I$, $f((x+x')/2) \le (f(x) + f(x'))/2$.

[21] To say that t *is a dyadic rational number* means that $t = k/2^n$ for some integer k and some integer $n \ge 0$.

[22] To say that f *is convex* means that for each $t \in [0,1]$, for all $x, x' \in I$, $f((1-t)x + tx') \le (1-t)f(x) + tf(x')$.

Let X be a set, let $f\colon X \to [-\infty, \infty]$, and let $\mathscr{B}$ be a filter base on X. By definition, *the limit inferior of f along $\mathscr{B}$* is

$$\liminf_{\mathscr{B}} f = \sup_{B \in \mathscr{B}} \inf_{x \in B} f(x)$$

and the *limit superior of f along $\mathscr{B}$* is

$$\limsup_{\mathscr{B}} f = \inf_{B \in \mathscr{B}} \sup_{x \in B} f(x).$$

Sometimes we may write $\liminf_{x \to \mathscr{B}} f(x)$ for $\liminf_{\mathscr{B}} f$ and $\limsup_{x \to \mathscr{B}} f(x)$ for $\limsup_{\mathscr{B}} f$. The notation $\liminf_{x \to \mathscr{B}} f(x)$ may be read "the limit inferior of $f(x)$ as x runs along $\mathscr{B}$." The notation $\limsup_{x \to \mathscr{B}} f(x)$ may be read "the limit superior of $f(x)$ as x runs along $\mathscr{B}$."

X56. Let X be a topological space, let S be a dense subset of X, and let $f\colon S \to [-\infty, \infty]$. For each $p \in X$, let $\mathscr{U}_p$ be the collection of nhds of p in X and let $\mathscr{B}_p = \{\, U \cap S : U \in \mathscr{U}_p \,\}$. Define $[-\infty, \infty]$-valued functions f_* and f^* on X by

$$f_*(p) = \liminf_{\mathscr{B}_p} f \qquad \text{and} \qquad f^*(p) = \limsup_{\mathscr{B}_p} f.$$

- **(a)** Verify that $f_* \le f^*$ on X and that $f_* \le f \le f^*$ on S.
- **(b)** Prove that f_* is lower semicontinuous and that f^* is upper semicontinuous.
- **(c)** Let $S \subseteq T \subseteq X$, let $\varphi\colon T \to [-\infty, \infty]$ be lower semicontinuous, and suppose that $\varphi \le f$ on S. Prove that $\varphi \le f_*$ on T. Thus the restriction of f_* to T is the largest lower semicontinuous function on T which is less than or equal to f on S.
- **(d)** Let $S \subseteq T \subseteq X$, let $\psi\colon T \to [-\infty, \infty]$ be upper semicontinuous, and suppose that $f \le \psi$ on S. Prove that $f^* \le \psi$ on T. Thus the restriction of f^* to T is the smallest upper semicontinuous function on T which is greater than or equal to f on S.
- **(e)** Verify that the set of all $p \in S$ such that f is continuous at p is equal to $\{\, p \in S : f_*(p) = f^*(p) \,\}$.
- **(f)** Suppose that f is continuous on S. Let $D = \{\, p \in X : f_*(p) = f^*(p) \,\}$ and let g be the restriction of f_* (or equivalently, f^*) to D. Prove that D and g are the same as in problem X51 in the special case where the topological space Y there is $[-\infty, \infty]$.

Terminology. In the notation of problem X56, in the special case where $S = X$, f_* is called *the lower regularization of f* and f^* is called *the upper regularization of f*.

X57. Let X be a topological space, let $A \subseteq X$, and let $f = 1_A$. Prove that

$$f_* = 1_{A^\circ} \qquad \text{and} \qquad f^* = 1_{\overline{A}},$$

where f_* and f^* are the lower and upper regularizations of f and where A° is the interior of A and $\overline{A}$ is the closure of A.

Reminder. Let $\mathscr{B}$ and $\mathscr{C}$ be filter bases. To say that $\mathscr{B}$ is finer than $\mathscr{C}$ means that for each $C \in \mathscr{C}$, there exists $B \in \mathscr{B}$ such that $B \subseteq C$.

X58. Let $\mathscr{A}$, $\mathscr{B}$, and $\mathscr{C}$ be filter bases. Suppose that $\mathscr{A}$ is finer than $\mathscr{B}$ and $\mathscr{B}$ is finer than $\mathscr{C}$. Prove that $\mathscr{A}$ is finer than $\mathscr{C}$.

Definition. Let $\mathscr{B}$ and $\mathscr{C}$ be filter bases. To say that *$\mathscr{B}$ and $\mathscr{C}$ are equally fine* means that each is finer than the other.

X59. Let X be a set and let Φ be the set of all filter bases on X. Define a binary relation $\sim$ on X by $\mathscr{A} \sim \mathscr{B}$ iff $\mathscr{A}$ and $\mathscr{B}$ are equally fine. Prove that $\sim$ is indeed an equivalence relation on Φ.

X60. Let X be a topological space, let $\mathscr{B}$ and $\mathscr{C}$ be equally fine filter bases on X, and let $p \in X$.

- **(a)** Prove that $\mathscr{B} \to p$ iff $\mathscr{C} \to p$.
- **(b)** Prove that p is a cluster point for $\mathscr{B}$ iff p is a cluster point for $\mathscr{C}$.

X61. Let X and Y be sets, let $f\colon X \to Y$, and let $\mathscr{B}$ and $\mathscr{C}$ be equally fine filter bases on X. Prove that $f\{\mathscr{B}\}$ and $f\{\mathscr{C}\}$ are equally fine filter bases on Y.

X62. Let $\mathscr{B}$ be a filter base. Let $\mathscr{C} \subseteq \mathscr{B}$ such that for each $B \in \mathscr{B}$, there exists $C \in \mathscr{C}$ such that $C \subseteq B$.

- **(a)** Prove that $\mathscr{C}$ is a filter base.
- **(b)** Prove that the filter bases $\mathscr{B}$ and $\mathscr{C}$ are equally fine.

Example. Let X be a topological space, let $p \in X$, let $\mathscr{U}$ be the filter base of all nhds of p in X, and let $\mathscr{B}$ be a nhd base at p for X. Then $\mathscr{B}$ is a filter base too and the filter bases $\mathscr{B}$ and $\mathscr{U}$ are equally fine.

X63. Let $\mathscr{B}$ be a filter base. Suppose there is a countable filter base $\mathscr{D}$ such that $\mathscr{B}$ and $\mathscr{D}$ are equally fine. Prove that there is a countable filter base $\mathscr{C}$ such that $\mathscr{B}$ and $\mathscr{C}$ are equally fine and $\mathscr{C} \subseteq \mathscr{B}$.

Definition. Let $\mathscr{B}$ be a filter base and let (x_n) be a sequence. To say that (x_n) *is cofinal in* $\mathscr{B}$ means that for each $B \in \mathscr{B}$, there exists N such that for each $n \geq N$, we have $x_n \in B$.

Example. Let X be a topological space, let $\mathscr{B}$ be a nhd base at p for X, and let (x_n) be a sequence in X. Then (x_n) is cofinal in $\mathscr{B}$ iff $x_n \to p$.

X64. Let (x_n) be a sequence and let $\mathscr{B}$ be a filter base. Let $\mathscr{T}$ be the filter base of tails of (x_n). Prove that (x_n) is cofinal in $\mathscr{B}$ iff $\mathscr{T}$ is finer than $\mathscr{B}$.

X65. Let $\mathscr{B}$ and $\mathscr{C}$ be equally fine filter bases. Let (x_n) be a sequence. Prove that (x_n) is cofinal in $\mathscr{B}$ iff (x_n) is cofinal in $\mathscr{C}$.

X66. Let E be a set, let Y be a topological space, let $f\colon E \to Y$, let $\mathscr{B}$ be a filter base on E, and let $q \in Y$.

- **(a)** Prove that if $f \to q$ along $\mathscr{B}$, then for each sequence (x_n) which is cofinal in $\mathscr{B}$, we have $f(x_n) \to q$.
- **(b)** Suppose that $\mathscr{B}$ and some countable filter base are equally fine and that for each sequence (x_n) which is cofinal in $\mathscr{B}$, we have $f(x_n) \to q$. Prove that $f \to q$ along $\mathscr{B}$.

X67. Let X and Y be topological spaces. Let $f\colon X \to Y$ and let $p \in X$.

- **(a)** Prove that if f is continuous at p, then for each sequence (x_n) in X such that $x_n \to p$, we have $f(x_n) \to f(p)$.
- **(b)** Suppose that there exists a countable nhd base at p for X and that for each sequence (x_n) in X such that $x_n \to p$, we have $f(x_n) \to f(p)$. Prove that f is continuous at p.

X68. Let X be a connected metric space with at least two points. Prove that the cardinality of X is greater than or equal to the cardinality of $\mathbf{R}$.

Remark. Let X be a topological space. It is obvious that if A and B are separated subsets of X, then A and B are disjoint. It is also obvious that if E and F are disjoint closed subsets of X, then E and F are separated. But in general, disjoint sets need not be separated. For instance, in $\mathbf{R}$, the sets $\mathbf{Q}$ and $\mathbf{R} \setminus \mathbf{Q}$ are disjoint but not separated.

X69. Let X be a topological space and let U and V disjoint open subsets of X. Prove that U and V are separated.

X70. Let X be a metric space, with metric d. Let A and B be separated subsets of X. Prove that there exist disjoint open subsets U and V in X such that $A \subseteq U$ and $B \subseteq V$. (Hint: To avoid trivialities, suppose that A and B are non-empty. Consider the functions f and g on X defined by $f(x) = d(x, A)$ and $g(x) = d(x, B)$. Use f and g to define U and V somehow.)

X71. Let X be an infinite set. Let $\mathscr{C}$ be the collection of all subsets $C \subseteq X$ such that $X \setminus C$ is finite. Let $\mathscr{G} = \mathscr{C} \cup \{ \emptyset \}$.

- **(a)** Prove that $\mathscr{G}$ is a topology on X. ($\mathscr{G}$ is called the cofinite topology on X.)

For the remainder of this exercise, let X be endowed with the topology $\mathscr{G}$.

- **(b)** Prove that X is T_1 but not Hausdorff. (To review the basic facts about T_1 spaces, see problem X32.)
- **(c)** Give an example of two disjoint closed sets A and B in X for which there do not exist disjoint open sets U and V with $A \subseteq U$ and $B \subseteq V$. Contrast this with what we saw in problem X70.

X72. Let X be a metric space and let A be a non-empty bounded subset of X. Define a function f on X by $f(x) = \sup \{ d(x, a) : a \in A \}$. *due 9Tu*

- **(a)** Verify that $f(x) \geq 0$ for each $x \in X$.
- **(b)** Verify that $f(x) < \infty$ for each $x \in X$.
- **(c)** Prove that for all $x, x' \in X$, $|f(x) - f(x')| \leq d(x, x')$. In particular, f is a Lipschitz function. Still more particularly, f is continuous. (Warning: Do not "calculate" with sup. Any inequalities you assert involving sup must be justified based on the definition of sup. Also, please do not use proof by contradiction when it is not needed and does not shorten the argument.)
- **(d)** Now suppose that each closed, bounded subset of X is compact. Let $\mathscr{R}$ be the set of all $r \in [0, \infty)$ such that $A \subseteq B[x, r]$ for some $x \in X$, where $B[x, r]$ is the closed ball with center x and radius r. Let $r_0 = \inf \mathscr{R}$. Prove that $r_0 \in \mathscr{R}$. Thus there exists a closed ball in X that is "circumscribed" about A. (Hint: Apply part (c).)

X73. Let A be a non-empty bounded subset of $\mathbf{R}^d$. By problem X72, there exists a closed ball in $\mathbf{R}^d$ that is "circumscribed" about A. Prove that this ball is unique. (Hint: The proof depends on special properties of the metric on $\mathbf{R}^d$.)

Let X be a topological space and let $A \subseteq X$. To say that A *is nowhere dense* means that the closure of A has empty interior. To say that A *is meager* means that A is a countable union of nowhere dense sets. For instance, any countable subset of $\mathbf{R}$ is meager in $\mathbf{R}$. So is any closed subset of $\mathbf{R}$ which has empty interior. For instance, the Cantor set is meager in $\mathbf{R}$. You should think of meager sets as sets which are small in the sense of topology. Meager sets are also known as sets *of the first category.* Sets which are not meager are also known as sets *of the second category.* To say that X is a *Baire space* means that no non-empty open subset of X is meager.

X74. Let X be a topological space. Prove that the following are equivalent:

- **(a)** X is a Baire space.
- **(b)** For each meager set $A \subseteq X$, the set $X \setminus A$ is dense in X.
- **(c)** For each sequence (G_n) of dense open subsets of X, we have $\bigcap_{n=1}^{\infty} G_n$ is dense in X.

X75. Let X be a complete metric space. Prove that X is a Baire space. (Terminology: This result is known as *the Baire category theorem.* Hint: Let (G_n) be a sequence of dense open subsets of X and let $H = \bigcap_{n=1}^{\infty} G_n$. We wish to show that H is dense in X. Let $x \in X$ and let $r > 0$. Since G_1 is dense, $B(x,r) \cap G_1 \neq \emptyset$. Let $x_1 \in B(x,r) \cap G_1$. Since $B(x,r) \cap G_1$ is open, there exists $r_1 \in (0,1)$ such that $B[x_1, r_1] \subseteq B(x,r) \cap G_1$. Since G_2 is dense, $B(x_1, r_1) \cap G_2 \neq \emptyset$. Let $x_2 \in B(x_1, r_1) \cap G_2$. Since $B(x_1, r_1) \cap G_2$ is open, there exists $r_2 \in (0, 1/2)$ such that $B[x_2, r_2] \subseteq B(x_1, r_1) \cap G_2$. Continuing in this way, construct a Cauchy sequence (x_n) which converges to a point in $B(x,r) \cap H$.)

X76. Let A be a dense $\mathscr{G}_\delta$ set in $\mathbf{R}$. Prove that A is uncountable. (Hint: Suppose A is countable. Use the Baire category theorem to get a contradiction.)

Remark. In problem X76, $\mathbf{R}$ could be replaced by any non-empty complete metric space without isolated points.

X77. Prove that the set of rational numbers is not a $\mathscr{G}_\delta$ set in $\mathbf{R}$.

Let X be a topological space. To say that $\mathscr{B}$ *is a base for the topology of* X means that $\mathscr{B}$ is a collection of open subsets of X and each open subset of X is the union of some subcollection of $\mathscr{B}$. Since the union of any collection of open sets is open, if $\mathscr{B}$ is a base for the topology of X, then the collection of open subsets of X is precisely the collection of all possible sets that can be expressed as the union of some subcollection of $\mathscr{B}$. To say that X *is second countable* means that there exists a countable base for the topology of X. To say that A *is dense in* X means that A is a subset of X whose closure is all of X. For instance, $\mathbf{Q}$ is dense in $\mathbf{R}$. More generally, $\mathbf{Q}^d$ is dense in $\mathbf{R}^d$. To say that X *is separable* means that some countable subset of X is dense in X. For instance, $\mathbf{R}^d$ is separable, because $\mathbf{Q}^d$ is countable and dense in $\mathbf{R}^d$.

X78. Let X be a separable metric space. Prove that X is second countable.

Remark. It follows from problem X78 that $\mathbf{R}^d$ is second countable. Another way to see this is to notice that the collection of sets of the form $\prod_{k=1}^{d} (a_k, b_k)$, where $a_k, b_k \in \mathbf{Q}$ and $a_k < b_k$, is a countable base for the usual topology on $\mathbf{R}^d$.

X79. Let X be a second countable topological space and let X_1 be a subspace of X. Prove that X_1 is second countable.

X80. Let X be a second countable topological space. Prove that X is separable and so is each subspace of X.

X81. For each point $z = (x,y) \in \mathbf{R}^2$ and each $r > 0$, let $U(z,r)$ be the rectangle $[x, x+r) \times [y, y+r)$. Let $\mathscr{G}$ be the collection of all subsets $G \subseteq \mathbf{R}^2$ such that for $z \in G$, there exists $r > 0$ such that $U(z,r) \subseteq G$.

(a) Prove that $\mathscr{G}$ is a topology on $\mathbf{R}^2$.

Let Z denote $\mathbf{R}^2$ endowed with the topology $\mathscr{G}$, rather than the usual topology. Let

$$Z_1 = \{ (x, -x) : x \in \mathbf{R} \}.$$

Endow Z_1 with the subspace topology it inherits from Z.

(b) Prove that Z is first countable.
(c) Prove that Z is separable.
(d) Prove that Z_1 is discrete. In other words, prove that every subset of Z_1 is open in Z_1.
(e) Prove that Z_1 is not separable. Thus a subspace of a separable topological space need not be separable.
(f) Prove that Z is not second countable. Thus a separable first countable space need not be second countable.

X82. Let $Y = \mathbf{R}$. Give an example of a subset $X \subseteq \mathbf{R}^2$ and a one-to-one Lipschitz function f from X onto Y such that $f^{-1}\colon Y \to X$ is discontinuous at each point in Y.

X83. Give an example of subsets $X, Y \subseteq \mathbf{R}$ and a one-to-one continuous map f from X onto Y such that $f^{-1}\colon Y \to X$ is discontinuous at each point in Y. (Hint: Consider an enumeration of the rationals.)

X84. Let X be a totally bounded metric space.
- **(a)** Prove that X is separable.
- **(b)** Prove that X is second countable.

In particular, a compact metric space is separable and second countable.

There are a couple of concepts in topology that are related to compactness, namely countable compactness and sequential compactness. Countable compactness and sequential compactness are less important concepts than compactness, but by exploring these three notions and the connections between them, as you are asked to do in some of the following exercises, you may improve your understanding of compactness.

Reminder. To say that a topological space is countably compact means that each countable open cover of the space has a finite subcover. Obviously each compact space is countably compact. It is not hard to give examples of countably compact spaces which are not compact, though these are not particularly important in analysis.

X85. Let X be a topological space. In class, we have seen that if X is countably compact, then each sequence in X has a cluster point. Prove the converse of this.

X86. Let X be a non-empty topological space. In class, we have seen that if X is countably compact, then each upper semicontinuous function on X achieves a maximum. Prove the converse of this. *due 10Th*

X87. Let X be a topological space. Prove that X is compact iff each filter base on X has a cluster point in X.

Definition. Let X be a topological space. To say that X *is sequentially compact* means that each sequence in X has a convergent subsequence.

X88. Let X be a topological space. Suppose X is sequentially compact. Prove that X is countably compact.

It is not hard to give examples of countably compact spaces which are not sequentially compact, but as such spaces are not particularly important in analysis, we shall not bother to do so.

X89. Let X be a topological space. Suppose X is countably compact and first countable. Prove that X is sequentially compact.

Remark. Since each compact space is countably compact, it follows from problem X89 that each compact first countable space is sequentially compact.

X90. Let X be a topological space. Prove that X is compact iff for each filter base $\mathscr{B}$ on X, there exists a finer filter base $\mathscr{C}$ on X which converges to some point in X.

Reminder. Let X be a metric space. In class, we saw that X is totally bounded iff each sequence in X has a Cauchy subsequence.

X91. Let X be a metric space. Prove that the following are equivalent:
- **(a)** X is compact.
- **(b)** X is countably compact.
- **(c)** X is sequentially compact.

Remark. Let X be a topological space. By problem X87 and problem X90, the following are equivalent:
- **(d)** X is compact
- **(e)** Each filter base on X has a cluster point.
- **(f)** For each filter base $\mathscr{B}$ on X, there exists a finer filter base $\mathscr{C}$ on X which converges to some point in X.

X92. Let Γ be a set of filter bases. Suppose Γ is upwards directed by fineness. In other words, suppose that for each $\mathscr{A} \in \Gamma$ and each $\mathscr{B} \in \Gamma$, there exists $\mathscr{C} \in \Gamma$ such that $\mathscr{C}$ is finer than each of $\mathscr{A}$ and $\mathscr{B}$. Let $\mathscr{F} = \bigcup \Gamma$. Prove that $\mathscr{F}$ is a filter base.

Definition. Let X be a set. To say that $\mathscr{M}$ *is a maximal filter base on* X means that $\mathscr{M}$ is a filter base on X and for each filter base $\mathscr{F}$ on X, if $\mathscr{F}$ is finer than $\mathscr{M}$, then $\mathscr{M}$ and $\mathscr{F}$ are equally fine.

Remark. Some of you may know about filters and ultrafilters. If so, you may be wondering about the relationship between maximal filter bases and ultrafilters. If you don't care about this relationship, or if you don't know about these topics, then you may disregard the rest of this remark, because it is enough just to know about filter bases and maximal filter bases.

Let X be a set. To say that $\mathscr{F}$ *is a filter on* X means that $\mathscr{F}$ is a non-empty collection of non-empty subsets of X such that $\mathscr{F}$ is closed under finite intersections and for each $F \in \mathscr{F}$ and each $A \subseteq X$, if $A \supseteq F$, then $A \in \mathscr{F}$. Evidently a filter on X is also a filter base on X, but not conversely. If $\mathscr{B}$ is a collection of subsets of X and

$$\mathscr{F} = \{\, F \subseteq X : F \supseteq B \text{ for some } B \in \mathscr{B} \,\}$$

then $\mathscr{F}$ is a filter on X iff $\mathscr{B}$ is a filter base. In this case, $\mathscr{B}$ and $\mathscr{F}$ are equally fine filter bases and we say that $\mathscr{B}$ is a base for the filter $\mathscr{F}$ and that $\mathscr{F}$ is the filter on X generated by $\mathscr{B}$. If $\mathscr{A}$ and $\mathscr{B}$ are filter bases on X and if $\mathscr{E}$ and $\mathscr{F}$ are the filters on X generated by $\mathscr{A}$ and $\mathscr{B}$ respectively, then $\mathscr{B}$ is finer than $\mathscr{A}$ iff $\mathscr{F} \supseteq \mathscr{E}$, and $\mathscr{A}$ and $\mathscr{B}$ are equally fine iff $\mathscr{E} = \mathscr{F}$. To say that $\mathscr{U}$ *is an ultrafilter on* X means that $\mathscr{U}$ is a filter on X which is not properly contained in any other filter on X. If $\mathscr{M}$ is a filter base on X and $\mathscr{U}$ is the filter on X generated by $\mathscr{M}$, then $\mathscr{M}$ is a maximal filter base on X iff $\mathscr{U}$ is an ultrafilter on X.

One reason why I have chosen to emphasize filter bases rather than filters is that filter bases are more concrete objects than filters. For instance, if (x_n) is a sequence in $\mathbf{R}$, then the filter base of tails of (x_n) is a countable set of countable subsets of $\mathbf{R}$, whereas the filter that this generates on $\mathbf{R}$ has cardinality $2^{2^{\aleph_0}}$ and contains many uncountable sets. Another reason is that the notion of the image of a filter base under a map is simpler than the notion of the image of a filter under a map.

X93. Let X be a set and let $\mathscr{E}$ be a filter base on X. Prove that there exists a maximal filter base $\mathscr{M}$ on X such that $\mathscr{M}$ is finer than $\mathscr{E}$. (Hint: Let Φ be the set of all filter bases $\mathscr{A}$ on X such that $\mathscr{E} \subseteq \mathscr{A}$. Partially order Φ by inclusion. Use Zorn's lemma to show that Φ has an element $\mathscr{M}$ which is maximal with respect to inclusion. Since $\mathscr{E} \subseteq \mathscr{M}$, it is obvious that $\mathscr{M}$ is finer than $\mathscr{E}$. Prove that $\mathscr{M}$ is maximal in the sense of filter bases.)

X94. Let X be a topological space, let $\mathscr{M}$ be a maximal filter base on X, and let $p \in X$. Prove that $\mathscr{M} \to p$ iff $\mathscr{M}$ clusters at p.

X95. Let X be a topological space. Prove that X is compact iff each maximal filter base on X converges.

X96. Let X be a set, let $\mathscr{A}$ be a filter base on X, let $E \subseteq X$, let $\mathscr{B} = \{\, E \cap A : A \in \mathscr{A} \,\}$, let $E^c = X \setminus E$, and let $\mathscr{C} = \{\, E^c \cap A : A \in \mathscr{A} \,\}$. Prove that at least one of $\mathscr{B}$ and $\mathscr{C}$ is a filter base.

X97. Let X be a set and let $\mathscr{M}$ be a filter base on X. For each $E \subseteq X$, let us say that E *is big* iff E has a subset belonging to $\mathscr{M}$.

(a) Prove that for each $E \subseteq X$, at most one of E and $X \setminus E$ is big.

(b) Prove that $\mathscr{M}$ is a maximal filter base on X iff for each $E \subseteq X$, either E is big or $X \setminus E$ is big.

X98. Let X and Y be sets, let $f\colon X \to Y$, and let $\mathscr{M}$ be a maximal filter base on X. Prove that $f\{\mathscr{M}\}$ is a maximal filter base on Y. (Hint: Use problem X97.)

Example. If U and V are open intervals in $\mathbf{R}$, then the rectangle $U \times V$ is open in the Cartesian plane $\mathbf{R}^2 = \mathbf{R} \times \mathbf{R}$. Most open subsets of $\mathbf{R}^2$ are not rectangles, of course. For instance, the unit disk

$$D = \{\, (x, y) \in \mathbf{R} \times \mathbf{R} : x^2 + y^2 < 1 \,\}$$

is open but is not a rectangle. But a subset of $\mathbf{R}^2$ is open iff it is a union of some collection (usually an infinite collection) of open rectangles. This motivates the general definition of the product topology on a Cartesian product of topological spaces.

In the next exercise, the product topology is treated for the Cartesian product of any family (even an infinite family) of topological spaces. But before that, let us review the notion of the Cartesian product of a family of sets.

Notation and Terminology. An *indexed family* $(x_\alpha)_{\alpha\in A}$ is just a function whose domain is the set A and whose value at α is x_α. The term *family* is just a shorter name for an indexed family. A family $x = (x_\alpha)_{\alpha\in A}$ is not the same thing as the set $\{\, x_\alpha : \alpha \in A \,\}$. Rather, this set is the range of the function x. (If we think of a function as being the same thing as its graph, then a family $x = (x_\alpha)_{\alpha\in A}$ is the same thing as the set $\{\, (\alpha, x_\alpha) : \alpha \in A \,\}$.) A *family of sets* $(X_\alpha)_{\alpha\in A}$ is just a family with the property that for each $\alpha \in A$, X_α is a set. A *choice function* for a family of sets $(X_\alpha)_{\alpha\in A}$ is a function x whose domain is A and which has the property that for each $\alpha \in A$, we have $x(\alpha) \in X_\alpha$. In other words, a choice function for a family of sets $(X_\alpha)_{\alpha\in A}$ is an family $(x_\alpha)_{\alpha\in A}$ such that for each $\alpha \in A$, we have $x_\alpha \in X_\alpha$. Informally, a choice function for a family of sets chooses one element from each of the sets in the family. One form of *the axiom of choice* states that each family of non-empty sets has at least one choice function. If $(X_\alpha)_{\alpha\in A}$ is a family of sets, then *the Cartesian product* $\prod_{\alpha\in A} X_\alpha$ of the family is the set of all choice functions for the family. In other words,

$$\prod_{\alpha\in A} X_\alpha = \{\, (x_\alpha)_{\alpha\in A} : x_\alpha \in X_\alpha \text{ for each } \alpha \in A \,\}.$$

For each $\alpha \in A$, the projection map $\pi_\alpha \colon X \to X_\alpha$ is defined as follows: if $x = (x_\alpha)_{\alpha\in A} \in X$, then $\pi_\alpha(x) = x_\alpha$. In the special case where $X_\alpha = B$ for each $\alpha \in A$, the Cartesian product $\prod_{\alpha\in A} X_\alpha$ is usually denoted by B^A. Thus B^A is the set of all functions from A into B.

X99. Let $(X_\alpha)_{\alpha\in A}$ be a family of topological spaces and let X be the Cartesian product $\prod_{\alpha\in A} X_\alpha$ of this family. Let $\mathscr{U}$ be the collection of all sets of the form $\prod_{\alpha\in A} U_\alpha$, where U_α is open in X_α for each $\alpha \in A$ and $U_\alpha = X_\alpha$ for all but finitely many $\alpha \in A$. Let $\mathscr{G}$ be the set of all $G \subseteq X$ such that G is the union of some subcollection of $\mathscr{U}$.

- **(a)** Prove that $\mathscr{U}$ is closed under finite intersections.
- **(b)** Prove that $\mathscr{G}$ is a topology on X.

We call $\mathscr{G}$ the product topology. Endow X with the topology $\mathscr{G}$. For each $\alpha \in A$, let π_α denote the projection map from X to X_α.

- **(c)** Prove that for each $\alpha \in A$, $\pi_\alpha \colon X \to X_\alpha$ is continuous.
- **(d)** Let E be a topological space and let $f \colon E \to X$. For each $\alpha \in A$, let $f_\alpha = \pi_\alpha \circ f$. Evidently, we have $f(u) = (f_\alpha(u))_{\alpha\in A}$ for each $u \in E$. Let $p \in E$. Prove that f is continuous at p iff for each $\alpha \in A$, f_α is continous at p.

Remark. Unless otherwise stated, the Cartesian product of a family of topological spaces is understood to be endowed with the product topology.

X100. Let $(X_\alpha)_{\alpha\in A}$ be a family of topological spaces and let $X = \prod_{\alpha\in A} X_\alpha$. For each $\alpha \in A$, let π_α denote the projection map from X to X_α. Let $\mathscr{B}$ be a filter base on X and let $x = (x_\alpha)_{\alpha\in A}$ be an element of X. Prove that $\mathscr{B} \to x$ iff for each $\alpha \in A$, $\pi_\alpha\{\mathscr{B}\} \to x_\alpha$.

Remark. Since convergence of filter bases characterizes a topology, the property of the product topology described in problem X100 characterizes the product topology.

Remark. Let $(X_\alpha)_{\alpha\in A}$ be a countable family of sequentially compact topological spaces and let $X = \prod_{\alpha\in A} X_\alpha$. Then X is sequentially compact. This is almost trivial when A is finite. When A is countably infinite, it is proved in Theorem 7.23 in Rudin, *Principles of Mathematical Analysis*, Third Edition, in the special case where each $X_\alpha = \mathbf{C}$, and the proof for the general case is essentially the same. It is worth mentioning that the proof is by a "diagonal subsequence argument" similar to the argument that we used in class to show that in a totally bounded metric space, each sequence has a Cauchy subsequence.

X101. (*Tychonoff's theorem*, 1930, 1935.) Let $(X_\alpha)_{\alpha\in A}$ be a family of compact topological spaces and let $X = \prod_{\alpha\in A} X_\alpha$. Prove that X is compact. (Hint: Prove that each maximal filter base on X converges.)

Remark on the Proof of Tychonoff's Theorem. The proof of Tychonoff's theorem via maximal filter bases is transparently clear and is a natural generalization of the proof that the product of countably many sequentially compact spaces is sequentially compact. There is another common proof of Tychonoff's theorem which is based a result called the *Alexandroff subbase theorem.* This approach does not require the theory of filter bases, so it has the advantage of being short and self-contained. Its disadvantage is that it is less motivated than the approach via the theory of filter bases.

Historical Remarks. It is fair to say that it was Tychonoff's theorem that clinched the modern definition of compactness. Prior to Tychonoff's theorem, countable compactness and sequential compactness were taken seriously as competitors for the distinction of being the "right" definition of compactness. But neither countable compactness nor sequential compactness behaves as well with respect to formation of product spaces as compactness does. By the remark immediately preceding problem X101, the product of countably many sequentially compact spaces is sequentially compact. The product of continuum many copies of $[0,1]$ is compact, by Tychonoff's theorem, but it is not sequentially compact. The behaviour of countable compactness with respect to formation of product spaces is even worse: a product of two countably compact spaces need not be countably compact. For a more detailed discussion of the history of notions of compactness, see the historical notes in Willard, *General Topology*, Addison-Wesley, 1970, pages 303 to 305.

Remark. Tychonoff's theorem depends on the axiom of choice through Zorn's lemma, which we used in problem X93. But more is true: It is not hard to show that Tychonoff's theorem is actually equivalent to the axiom of choice.

Let X and Y be topological spaces. To say that h *is a homeomorphism from* X *to* Y means that h is a one-to-one continuous map from X onto Y and $h^{-1}\colon Y \to X$ is also continuous. To say that X *is homeomorphic to* Y means that there exists a homeomorphism from X to Y. Informally, X and Y are homeomorphic when they look the same as topological spaces.

X102. Let $\Sigma = \{0,1\}^{\mathbf{N}}$. Note that Σ is the product of countably many copies of the two-point discrete space $\{0,1\}$. Give Σ the corresponding product topology.

(a) Explain why Σ is compact.

Now let C be the Cantor set, and let f be the one-to-one map from Σ onto C defined in problem X11.

(b) Prove that $f\colon \Sigma \to C$ is continuous.

(c) Prove that $f^{-1}\colon C \to \Sigma$ is continuous. (Hint: Σ is compact and C is Hausdorff.)

Thus Σ is homeomorphic to C. (In particular, Σ is not discrete, even though it is a product of discrete spaces.)

To say that a topological space is *metrizable* means that its topology arises from some metric. A metric space is more special than a metrizable space because a metric space has a metric already given on it. The topology on a metrizable space can always be induced by any one of an infinite number of metrics, provided that the space has at least two points.

X103. Prove that the product of countably many metrizable spaces is metrizable. (Hint: For a product of finitely many spaces, this is easy, so let's consider the countably infinite case. Suppose $(X_n)_{n\in\mathbf{N}}$ is a sequence of metrizable spaces. Let $X = \prod_{n\in\mathbf{N}} X_n$. For each n, let d_n be a metric on X_n which induces the given topology on X_n. By problem X2, may suppose that each $d_n \le 1$. Define $d\colon X\times X \to [0,1]$ by *due 10Th*

$$d(x,y) = \sum_{n\in\mathbf{N}} 2^{-n} d_n(x_n, y_n).$$

Prove that d is a metric on X and that d induces the product topology on X.)

X104. Let (X,ρ) be a metric space. Let $\sigma = \rho/(1+\rho)$. Then by problem X2, σ is a metric on X and the two metrics ρ and σ induce the same topology on X.

(a) Prove that the identity map on X is uniformly continuous from (X,ρ) to (X,σ).

(b) Prove that the identity map on X is uniformly continuous from (X,σ) to (X,ρ).

(c) Let (x_n) be a sequence in X. Prove that (x_n) is ρ-Cauchy iff (x_n) is σ-Cauchy.

(d) Prove that X is ρ-complete iff X is σ-complete.

To say that a topological space is *completely metrizable* means that its topology arises from some complete metric. For instance, $\mathbf{R}$ is completely metrizable, because the usual metric on $\mathbf{R}$ is complete. To mention another example, $(0,1)$ is completely metrizable, even though the usual metric on it is not complete, because $(0,1)$ is homeomorphic to $\mathbf{R}$.

X105. Prove that the product of countably many completely metrizable spaces is completely metrizable.

X106. Prove that the product of countably many second countable spaces is second countable.

X107. Prove that the product of countably many separable spaces is separable. (Warning: If $A_1, A_2, A_3, \ldots$ are countable, it does not follow that $\prod_{n\in\mathbf{N}} A_n$ is countable. For instance, $\{0,1\}^{\mathbf{N}}$ has the cardinality of the continuum.)

These problems are numbered consecutively with the ones from last quarter. (That is why the first of them is not numbered X1.)

X108. (*A discrete analog of part of l'Hospital's rule.*) Let (u_k) be a sequence in $\mathbf{R}$ and let (v_k) be a strictly increasing sequence in $(0,\infty)$ such that $v_k \to \infty$ as $k \to \infty$. Then

$$\limsup_{j\to\infty} \frac{u_j}{v_j} \le \limsup_{k\to\infty} \frac{u_k - u_{k-1}}{v_k - v_{k-1}}$$

and

$$\liminf_{j\to\infty} \frac{u_j}{v_j} \ge \liminf_{k\to\infty} \frac{u_k - u_{k-1}}{v_k - v_{k-1}}.$$

(Hint: Let $u_0 = v_0 = 0$. For all $j, k \in \mathbf{N}$, let

$$a(j,k) = \begin{cases} (v_k - v_{k-1})/v_j & \text{if } k \le j, \\ 0 & \text{if } k > j, \end{cases}$$

and

$$x_k = \frac{u_k - u_{k-1}}{v_k - v_{k-1}}.$$

Apply problem X47.)

X109. Let $f\colon I \to \mathbf{R}$ be twice-differentiable, where $I = (a,\infty)$ and $a \in [-\infty,\infty)$.

(a) Suppose $A, C \in [0,\infty)$ such that for each $x \in I$, $|f(x)| \le A$ and $|f''(x)| \le C$. Prove that for each $x \in I$ and each $h > 0$,

$$|f'(x)| \le \frac{A}{h} + Ch.$$

(Hint: Use Taylor's formula to express $f(x+2h)$ in terms of $f(x)$, $f'(x)$, and $f''(\xi)$ for a suitable ξ between x and $x+2h$.)

(b) Suppose A and C are as in part (a). Deduce from part (a) that for each $x \in I$,

$$|f'(x)|^2 \le 4AC.$$

(c) Suppose f'' is bounded and $f(x) \to 0$ as $x \to \infty$. Prove that $f'(x) \to 0$ as $x \to \infty$. (Hint: Let $\alpha \in I$ and apply part (b) with I replaced by (α,∞). Let $\alpha \to \infty$.)

X110. Let $f\colon I \to \mathbf{R}$, where I is an interval in $\mathbf{R}$. Suppose f is convex on I. Let I° be the interior of I. Then as we have seen in class, for each $x \in I^\circ$, the left derivative $D_L f(x)$ and the right derivative $D_R f(x)$ exist and satisfy $-\infty < D_L f(x) \le D_R f(x) < \infty$. As we have also seen in class, for all $x_1, x_2 \in I^\circ$ with $x_1 < x_2$, we have $D_R f(x_1) \le D_L f(x_2)$. *due 2Th*

(a) Prove that $D_R f$ is right-continuous on I° and $D_L f$ is left-continuous on I°.

(b) Prove that for all but countably many $x \in I^\circ$, f is differentiable at x.

By the way, you should not need (a) to prove (b) and you should not need (b) to prove (a).

Although we have not yet begun to discuss integration in this course, I think that should not hinder you in solving the next two problems.

X111. Let $f\colon [a,b] \to \mathbf{R}$, where $a, b \in \mathbf{R}$ with $a < b$. Suppose f is continuous on $[a,b]$ and f' and f'' exist on (a,b). Let $m = (a+b)/2$, the midpoint of the interval $[a,b]$. Prove that there exists $c \in (a,b)$ such that

$$\int_a^b f(x)dx = (b-a)f(m) + \frac{1}{24}f''(c)(b-a)^3$$

(Hint: Let F be an antiderivative for f. Expand F about m using Taylor's formula to get suitable expressions for $F(a)$ and $F(b)$. The fact that f'' has the intermediate value property, by Darboux's theorem, is relevant too.)

Remark. The result of problem X111 is the basis for the error estimate in the midpoint rule for numerical integration.

Remark. Let $T \in (0,\infty)$, let $L \in \mathbf{R}$, let $g,h\colon [0,T] \to \mathbf{R}$ be continuous, and suppose h is differentiable on $(0,T)$ and $h'(t) = g(t) + Lh(t)$ for all $t \in (0,T)$. In any introductory course on differential equations, students are taught how to solve for h in terms of g and $h(0)$ by the method of *integrating factors*, as follows: $h'(s) - Lh(s) = g(s)$, so $e^{-Ls}h'(s) - Le^{-Ls}h(s) = e^{-Ls}g(s)$, so $\big(e^{-Ls}h(s)\big)' = e^{-Ls}g(s)$, so $e^{-Lt}h(t) - e^{-L0}h(0) = \int_0^t e^{-Ls}g(s)\,ds$, so $h(t) = e^{Lt}\big(h(0) + \int_0^t e^{-Ls}g(s)\,ds\big)$.

X112. Let $T \in (0,\infty)$ and let $L \in [0,\infty)$. *due 2Th*

(a) (*Gronwall's Inequality).* Let $f,g\colon [0,T] \to \mathbf{R}$ be continuous, and suppose that for each $t \in [0,T]$,

$$f(t) \le g(t) + L\int_0^t f(s)\,ds. \tag{8}$$

Prove that for each $t \in [0,T]$,

$$f(t) \le g(t) + L\int_0^t e^{L(t-s)}g(s)\,ds. \tag{9}$$

(Hint: Let $h(t) = \int_0^t f(s)\,ds$. Then $h'(t) = f(t)$, so by (8), $h'(t) \le g(t) + Lh(t)$. Adapt the method illustrated in the remark that precedes this exercise to derive an inequality for $h(t)$ from which (9) follows.)

(b) Let $F\colon [0,T] \times \mathbf{R}^d \to \mathbf{R}^d$ be continuous and satisfy the following Lipschitz condition:

$$|F(t,x) - F(t,y)| \le L|x-y| \qquad \text{for all } x,y \in \mathbf{R}^d \text{ and all } t \in [0,T].$$

Suppose $u,v\colon [0,T] \to \mathbf{R}^d$ are continous and are differentiable on $(0,T)$ and satisfy $u'(t) = F(t,u(t))$ and $v'(t) = F(t,v(t))$ for all $t \in (0,T)$. Prove that $|u(t) - v(t)| \le |u(0) - v(0)|e^{Lt}$ for all $t \in [0,T]$. In particular, if $u(0) = v(0)$, then $u(t) = v(t)$ for all $t \in [0,T]$. (Hint: Notice that $u(t) = u(0) + \int_0^t F(s,u(s))\,ds$ and $v(t) = v(0) + \int_0^t F(s,v(s))\,ds$ for all $t \in [0,T]$, by the fundamental theorem of calculus (remember to verify its hypotheses!). Apply part (a) with $f(t) = |u(t) - v(t)|$ and $g(t) \equiv \gamma$ where $\gamma = |u(0) - v(0)|$. You will need the triangle inequality for integrals of vector-valued functions: $\big|\int_0^t \varphi(s)\,ds\big| \le \int_0^t |\varphi(s)|\,ds$ if $\varphi\colon [0,t] \to \mathbf{R}^d$ is continuous. We will prove this triangle inequality in class in due time. For now, feel free to use it without proving it.)

Remark. In problem X112(b), Gronwall's inequality is used to derive an estimate on how much the solution of an initial value problem

$$\begin{aligned} y' &= F(t,y),\\ y(0) &= y_0, \end{aligned}$$

changes as the initial value y_0 changes. Gronwall's inequality has other important applications in the theory of differential equations. For instance, if F is a function not just of t and y but also of a parameter w, then under suitable conditions, Gronwall's inequality can be used to derive estimates on how much the solution of the initial value problem

$$\begin{aligned} y' &= F(t,y;w),\\ y(0) &= y_0, \end{aligned}$$

changes as the parameter w changes.

X113. Let $f\colon [a,b] \to \mathbf{R}$, where $a, b \in \mathbf{R}$ with $a < b$. Suppose that f' is continuous on $[a,b]$ and that f'' exists on (a,b). Suppose in addition that $f'(a) = 0 = f'(b)$. Prove that there exists a number c in (a,b) such that

$$|f''(c)| \geq \frac{4}{(b-a)^2}|f(b) - f(a)|.$$

X114. Let Φ be the set of all twice continuously differentiable functions $f\colon [-1,1] \to \mathbf{R}$ such that $|f''(x)| \leq 1$ for all $x \in [-1,1]$ and $f(-1) = f'(-1) = f(1) = f'(1) = 0$. Let $E = \{\, f(0) : f \in \Phi \,\}$. Find $\sup E$.

X115. Let $f\colon [a,b] \to \mathbf{R}$, where $a, b \in \mathbf{R}$ with $a < b$. Let $c \in (a,b)$.

(a) Prove that if f is Riemann-integrable over $[a,b]$, then f is Riemann-integrable over each of $[a,c]$ and $[c,b]$ and

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

(Hint: If P is a subdivision of $[a,b]$, then either c is one of the points of P, or it is not. In the latter case, if we add c to P, we get a finer subdivision of $[a,b]$.)

(b) Prove that if f is Riemann-integrable over each of $[a,c]$ and $[c,b]$, then f is Riemann-integrable over $[a,b]$.

Review of the Integral Test. Let $h\colon (0,\infty) \to [0,\infty)$ be decreasing. According to the integral test, which you will have learned in calculus, since h is decreasing and non-negative, the series $\sum_{k=1}^{\infty} h(k)$ converges iff the improper integral $\int_1^{\infty} h(x)\,dx$ converges. Let us review why this is true and discuss some related matters.

Since h is monotone, h is Riemann-integrable over $[a,b]$ for all $a, b \in (0,\infty)$ with $a < b$. As a decreases, or as b increases, $\int_a^b h(x)\,dx$ increases, because $h \geq 0$. Hence if $\alpha, \beta \in [0,\infty]$ with $\alpha < \beta$, and if $\alpha = 0$ or $\beta = \infty$ (or both), then the improper Riemann integral

$$\int_\alpha^\beta h(x)\,dx = \lim_{a \downarrow \alpha,\, b \uparrow \beta} \int_a^b h(x)\,dx$$

exists in $[0,\infty]$. Since h is decreasing, if $k \in \mathbf{N}$, then whenever $0 < x_1 \leq k \leq x_2 < \infty$, we have $h(x_1) \geq h(k) \geq h(x_2)$, so

$$\int_{k-1}^{k} h(x)\,dx \geq \int_{k-1}^{k} h(k)\,dx = h(k) = \int_{k}^{k+1} h(k)\,dx \geq \int_{k}^{k+1} h(x)\,dx.$$

Hence for all $m, n \in \mathbf{N}$ with $m \leq n$,

$$\int_{m-1}^{n} h(x)\,dx = \sum_{k=m}^{n} \int_{k-1}^{k} h(x)\,dx \geq \sum_{k=m}^{n} h(k) \geq \sum_{k=m}^{n} \int_{k}^{k+1} h(x)\,dx = \int_{m}^{n+1} h(x)\,dx.$$

It follows that for each $m \in \mathbf{N}$,

$$\int_{m-1}^{\infty} h(x)\,dx \geq \sum_{k=m}^{\infty} h(k) \geq \int_{m}^{\infty} h(x)\,dx. \tag{10}$$

In particular, if $\sum_{k=1}^{\infty} h(k) < \infty$, then $\int_1^{\infty} h(x)\,dx < \infty$ and conversely, if $\int_1^{\infty} h(x)\,dx < \infty$, then $\sum_{k=2}^{\infty} h(k) < \infty$, so $\sum_{k=1}^{\infty} h(k) < \infty$. Furthermore, if we wish to estimate the value of $\sum_{k=1}^{\infty} h(k)$ to a specified degree of precision, then (10) gives us a way to determine how large m should be so that the difference $\sum_{k=m+1}^{\infty} h(k)$ between $\sum_{k=1}^{\infty} h(k)$ and the partial sum $\sum_{k=1}^{m} h(k)$ is as small as we like. Furthermore, it often happens that the difference between the upper and lower bounds on this difference that are furnished by (10) differ from each other by substantially less than the remainder $\sum_{k=m+1}^{\infty} h(k)$ itself, in which case the sum

$$\sum_{k=1}^{m} h(k) + \int_{m+1}^{\infty} h(x)\,dx$$

gives a much better approximation to $\sum_{k=1}^{\infty} h(k)$ than the partial sum $\sum_{k=1}^{m} h(k)$ alone. For example, this is the case if $h(x) = 1/x^2$.

Definition. Let X be a set, let $f, g\colon X \to \mathbf{C} \setminus \{0\}$, and let $\mathscr{B}$ be a filter base on X. To say that f *is asymptotic to* g *along* $\mathscr{B}$ (denoted $f(x) \sim g(x)$ as $x \to \mathscr{B}$) means that

$$\frac{f(x)}{g(x)} \to 1 \qquad \text{as } x \to \mathscr{B}.$$

For instance, if $X \subseteq \mathbf{R}$ and X is not bounded above, then $f(x) \sim g(x)$ as $x \to \infty$ iff $f(x)/g(x) \to 1$ as $x \to \infty$. (This should tell you what you need to know about this definition to solve the next exercise, even if you do not know about filter bases.)

X116. Let $a \in (1, \infty)$. Define $f\colon (0, \infty) \to (0, \infty)$ by *due 3Th*

$$f(x) = \sum_{k=1}^{\infty} \frac{1}{x + k^a}.$$

(a) Let $b = 1/a$. Prove that there exists $C(a) \in (0, \infty)$ such that

$$f(x) \sim C(a)x^{b-1} \qquad \text{as } x \to \infty$$

and find an expression for $C(a)$ in the form of a certain improper integral.[23] (Hint: Compare f with the function g defined by

$$g(x) = \int_0^{\infty} \frac{1}{x + t^a}\, dt.$$

The review above of the integral test should help you understand how to compare f and g.)

(b) Find $\displaystyle\lim_{x\to\infty} \frac{\log f(x)}{\log x}$. (Hint: Write $f(x)$ as $g(x)\dfrac{f(x)}{g(x)}$, where $g(x) = C(a)x^{b-1}$.)

Example. Let $a = 7/5$ and let f be as in problem X116. Then $f(x) \sim C(a)x^{-2/7}$ as $x \to \infty$. You may find it enlightening to compare this with problem 6 on the analysis qualifying from Autumn, 2004.

X117. Let φ be a continuous function from $\mathbf{R}$ to $\mathbf{C}$ such that $\varphi(0) = 1$ and $\varphi(t_1 + t_2) = \varphi(t_1)\varphi(t_2)$ for all $t_1, t_2 \in \mathbf{R}$. Since φ is continuous, we may fix $\delta > 0$ such that for each $t \in [0, \delta]$, $|\varphi(t) - \varphi(0)| \leq 1/2$. Let $c = \int_0^{\delta} \varphi(t)\, dt$.

(a) Prove that $c \neq 0$. (Make sure your proof does not assume that φ is real-valued! This goes for all parts of this problem.)

(b) Define $F\colon \mathbf{R} \to \mathbf{C}$ by $F(x) = \int_0^x \varphi(t)\, dt$. Prove that for each $x \in \mathbf{R}$, $F(x + \delta) = F(x) + c\varphi(x)$.

(c) Prove that φ is differentiable on $\mathbf{R}$.

(d) Let $k = \varphi'(0)$. Prove that $\varphi' = k\varphi$.

(e) Prove that $\varphi(t) = e^{kt}$ for all $t \in \mathbf{R}$. (Please give a detailed proof that $\varphi(t) = e^{kt}$. Please do not just quote a theorem on uniqueness of solutions of differential equations. And please do not give an argument that involves $\log \varphi(t)$. Since φ is complex-valued, the meaning of $\log \varphi(t)$ would have to be clarified to make such an argument work and this would be the hard way to do the proof.)

Remark. It is not difficult to extend the result of problem X117 to the case where φ takes values in $\mathbf{C}^{d \times d}$, the set of $d \times d$ matrices with complex entries. Then $k \in \mathbf{C}^{d \times d}$ and e^{kt} is defined as the sum of the series $\sum_{n=0}^{\infty} (kt)^n / n!$.

Reminder. If $f\colon [a, b] \to \mathbf{R}$, where $a, b \in \mathbf{R}$ with $a < b$, then the upper Riemann integral of f over $[a, b]$ is the infimum of all $U(f, P)$, where P varies over all subdivisions of $[a, b]$.

Definition. Let $f\colon [a, b] \to \mathbf{R}^d$, where $a, b \in \mathbf{R}$ with $a < b$. To say that f *is Riemann null* means that the upper Riemann integral of $|f|$ over $[a, b]$ is 0.

---

[23] You need not evaluate the integral, but be sure to prove that it is finite and strictly positive, so that $C(a) \in (0, \infty)$ as stated.

X118. Let C be the Cantor set and define $f\colon [0,1] \to \mathbf{R}$ by

$$f(x) = \begin{cases} 1 & \text{if } x \in C, \\ 0 & \text{if } x \in [0,1] \setminus C. \end{cases}$$

Prove that f is Riemann null.

X119. Let $f\colon [a,b] \to \mathbf{R}$, where $a,b \in \mathbf{R}$ with $a < b$. Prove that f is Riemann null iff f is Riemann-integrable and $\int_a^b |f(x)|\,dx = 0$.

X120. Let $f\colon [a,b] \to \mathbf{R}^d$, where $a,b \in \mathbf{R}$ with $a < b$. Suppose that f is Riemann null and that f is continuous. Prove that $f = 0$. In other words, prove that for each $x \in [a,b]$, $f(x) = 0$.

X121. Let $f_1, \ldots, f_d\colon [a,b] \to \mathbf{R}$, where $a,b \in \mathbf{R}$ with $a < b$. Define $f\colon [a,b] \to \mathbf{R}^d$ by $f(x) = (f_1(x), \ldots, f_d(x))$ for all $x \in [a,b]$. Prove that f is Riemann null iff $f_1, \ldots, f_d$ are all Riemann null.

X122. Let $f\colon [a,b] \to \mathbf{R}^d$, where $a,b \in \mathbf{R}$ with $a < b$. Suppose f is Riemann null. Let E be the range of f, let $\varphi\colon E \to \mathbf{R}$ be bounded, and suppose that for each $\varepsilon > 0$, there exists $\delta > 0$ such that for each $y \in E$, if $|y| < \delta$, then $|\varphi(y)| < \varepsilon$. Let $h = \varphi \circ f$. Prove that h is Riemann null. (Hint: There is a proof of this which is similar to but simpler than the proof of a result we discussed in class, namely a variation on Theorem 6.11 in Rudin, *Principles of Mathematical Analysis*, Third Edition.) *due 3Th*

Notation. Let $\mathscr{R}[a,b]$ be the vector space of Riemann integrable complex-valued functions on $[a,b]$, where $a,b \in \mathbf{R}$ with $a < b$. (This is not standard notation.) Let $p \in [1,\infty)$. Then for each $f \in \mathscr{R}[a,b]$, we define the p-norm of f to be

$$\|f\|_p = \left(\int_a^b |f(x)|^p \, dx \right)^{1/p}.$$

This is a slight abuse of terminology, since if $f \in \mathscr{R}[a,b]$ and $\|f\|_p = 0$, it does not follow that $f = 0$. It is obvious that if $f \in \mathscr{R}[a,b]$ and $C \in \mathbf{C}$, then $\|Cf\|_p = |C| \|f\|_p$.

X123. Let $f \in \mathscr{R}[a,b]$, where $a,b \in \mathbf{R}$ and $a < b$. Let $p \in (1,\infty)$. Prove that $\|f\|_p = 0$ iff f is Riemann null.

X124. Let $f \in C[a,b]$, where $a,b \in \mathbf{R}$ and $a < b$. Let $p \in (1,\infty)$. Prove that $\|f\|_p = 0$ iff $f = 0$.

X125.

(a) Let $B = \{ h \in \mathscr{R}[a,b] : \|h\|_p \leq 1 \}$. Show that B is a convex subset of the vector space $\mathscr{R}[a,b]$. In other words, for all $h_1, h_2 \in B$, for each $t \in [0,1]$, we have $(1-t)h_1 + th_2 \in B$. (Hint: Use the fact that since $p \in [1,\infty)$, the function $y \mapsto y^p$ is convex on $[0,\infty)$.)

(b) (*Minkowski's Inequality.*) Let $p \in [1,\infty)$ and let $f,g \in \mathscr{R}[a,b]$. Prove that $\|f+g\|_p \leq \|f\|_p + \|g\|_p$. (Hint: This can be deduced from (a). Let $a = \|f\|_p$ and $b = \|g\|_p$. Notice that if $a \neq 0$, then $f/a \in B$. Similarly, if $b \neq 0$, then $g/b \in B$. Note: Soon we shall consider Hölder's inequality. Many books give a proof of Minkowski's inequality that is based on Hölder's inequality. That is a much less enlightening proof. Please do not do it that way.)

Notice that Minkowski's inequality is the triangle inequality for p-norms. The next result develops some more theory related to convex functions and applies it to prove another famous inequality involving p-norms, namely Hölder's inequality.

X126.

(a) Let I be an interval in $\mathbf{R}$ and let $f\colon I \to \mathbf{R}$ be convex. Let $x_1, \ldots, x_n \in I$ and let $t_1, \ldots, t_n \in [0,1]$ with $\sum_{k=1}^n t_k = 1$. Prove in two ways that

$$\varphi(\textstyle\sum_{k=1}^n t_k x_k) \leq \sum_{k=1}^n t_k \varphi(x_k).$$

(Hint: For your first proof, use induction on n, starting from the definition of what it means for φ to be convex. For your second proof, proceed as follows. Let $X = \{ x_1, \ldots, x_n \}$, let $a = \min X$, let $b = \max X$, and let $x = \sum_{k=1}^n t_k x_k$. To avoid trivialities, assume that $a < b$ and that each

$t_k > 0$. Check that $a < x < b$, so x is an interior point of I. Then apply the fact that the graph of a convex function has a support line at each interior point of the interval in which the function is defined.)

(b) Let $y_1, \ldots, y_n \in [0, \infty)$ and let $t_1, \ldots, t_n \in [0, 1]$ with $\sum_{k=1}^n t_k = 1$. Let $G = \prod_{k=1}^n y_k^{t_k}$ and let $A = \sum_{k=1}^n t_k y_k$. (Note that G is a weighted geometric mean of $y_1, \ldots, y_n$ and A is a weighted arithmetic mean of $y_1, \ldots, y_n$.) Prove that $G \leq A$. (Hint: If some $y_k = 0$, then $G = 0$, so there is nothing to prove. In the case where each $y_k > 0$, apply the fact that the function $x \mapsto e^x$ is convex on $\mathbf{R}$.)

(c) Let $u, v \in [0, \infty)$ and let $p, q \in (1, \infty)$ with

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Prove that

$$uv \leq \frac{u^p}{p} + \frac{v^q}{q}.$$

(This follows from (b) with $n = 2$. The analogous statement for general n is true, but for the sake of focus, we refrain from considering it.)

(d) (*Hölder's Inequality.*) Again, let $p, q \in (1, \infty)$ with

$$\frac{1}{p} + \frac{1}{q} = 1.$$

Let $f, g \in \mathscr{R}[a, b]$. Prove that

$$\|fg\|_1 \leq \|f\|_p \|g\|_q.$$

Remark. We can generalize problem X126(d) as follows: Let $p_1, \ldots, p_n \in (1, \infty)$ with

$$\frac{1}{p_1} + \cdots + \frac{1}{p_n} = 1.$$

Let $f_1, \ldots, f_n \in \mathscr{R}[a, b]$. Then

$$\|f_1 \cdots f_n\|_1 \leq \|f_1\|_{p_1} \cdots \|f_n\|_{p_n}.$$

The proof is just like that of problem X126(d), but it uses the generalization of problem X126(c) to general n that we mentioned but did not pursue.

Remark. In problem X126(d), if $p = 2$, then $q = 2$. In this case, Hölder's inequality, in conjunction with the triangle inequality for integrals, implies that if $f \in \mathscr{R}[a, b]$, then

$$\left| \int_a^b \overline{f(x)} g(x)\, dx \right| \leq \|f\|_2 \|g\|_2.$$

This is a particular case of Schwartz's inequality from the theory of inner product spaces.

X127. Let $f\colon (1,\infty) \to \mathbf{R}$ be differentiable and define $g, h\colon (1,\infty) \to \mathbf{R}$ by $g(x) = f'(x)/x$ and $h(x) = f(x)/x$. Suppose g is bounded. Prove that h is uniformly continuous.

X128. Let f be an n times differentiable real-valued function on $[a,b]$, where $a, b \in \mathbf{R}$ with $a < b$. Suppose that the n-th derivative of f satisfies $f^{(n)}(x) > 0$ for each $x \in [a,b]$. Prove that f has at most n zeroes in $[a,b]$.

X129. Let $f\colon [0,\infty) \to \mathbf{R}$ be Riemann integrable over $[0,T]$ for each $T \in (0,\infty)$. For each $T \in (0,\infty)$, let

$$I(T) = \frac{1}{T}\int_0^T f(t)\,dt.$$

Let $a = \liminf_{t\to\infty} f(t)$, $A = \liminf_{T\to\infty} I(T)$, $B = \limsup_{T\to\infty} I(T)$, and $b = \limsup_{t\to\infty} f(t)$.

(a) Prove that $B \le b$.

Similarly, $a \le A$. (Or this can be deduced by applying part (a) to the function $-f$ instead of f.)

(b) Let $L \in [-\infty,\infty]$. Deduce from part (a) that if $f(t) \to L$ as $t \to \infty$, then $I(T) \to L$ as $T \to \infty$.

X130. Give an example of a continuous function $f\colon [0,\infty) \to \mathbf{R}$ such that $\limsup_{t\to\infty} f(t) = \infty$ and $\liminf_{t\to\infty} f(t) = -\infty$ but $I(T) \to 0$ as $T \to \infty$, where $I(T)$ is as in problem X129.

X131. Let $f \in \mathscr{R}[-1,1]$ and suppose f is continuous at 0. Prove that *due 4Th; put a simpler one first next time*

$$\int_{-1}^{1} \frac{u f(x)}{u^2 + x^2}\,dx \to \pi f(0) \qquad \text{as } u \to 0+.$$

X132. Let $\varphi\colon \mathbf{R} \to \mathbf{C}$ be 1-periodic[24] and N times continuously differentiable, where $N \in \mathbf{N}$. Prove that there is a constant $C \in [0,\infty)$ such that for each $k \in \mathbf{Z} \setminus \{0\}$, *due 4Th*

$$\left|\int_0^1 e^{2\pi i k x}\varphi(x)\,dx\right| \le \frac{C}{|k|^N}.$$

X133. Let $f\colon [a,b] \to \mathbf{R}^d$, where $a, b \in \mathbf{R}$ with $a < b$. Suppose f is Riemann integrable over $[a,b]$. Define $F\colon [a,b] \to \mathbf{R}^d$ by $F(x) = \int_a^x f(t)\,dt$. Prove that F is continuous.

X134. Let $f\colon [a,b] \to \mathbf{R}$ be bounded, where $a, b \in \mathbf{R}$ with $a < b$, and suppose that f is Riemann integrable over (α,β) for all $\alpha, \beta \in (a,b)$ with $\alpha < \beta$. Prove that f is Riemann integrable over $[a,b]$. In particular, if f is bounded on $[a,b]$ and continuous on (a,b), then f is Riemann integrable over $[a,b]$.

X135. Let $I = [0,1]$, let $f\colon I \to \mathbf{R}$, let C be the Cantor set, and let $D = I \setminus C$. Suppose f is bounded on I and continuous at each point in D. Prove that f is Riemann-integrable over I. (Hint: From the definition of C, you should be able to see that for each $\gamma > 0$, C can be covered by a finite number of intervals whose total length is less than γ.)

Notation. Let X be a set and let $f, g\colon X \to \overline{\mathbf{R}}$. Then the *join, or maximum, of f and g* is the function $f \vee g$ on X defined by $(f \vee g)(x) = \max\{f(x), g(x)\}$. The *meet, or minimum, of f and g* is the function $f \wedge g$ on X defined by $(f \wedge g)(x) = \min\{f(x), g(x)\}$. The *positive part of f* is $f \vee 0$. The *negative part of f* is $(-f) \vee 0$. (Yes, according to this definition, the negative part of f is positive.) Notice that $f^+ \ge 0$, $f^- \ge 0$, $f = f^+ - f^-$, and $|f| = f^+ + f^-$. Also, if $f = g - h$, where $g, h\colon X \to [0,\infty]$, then $f^+ \le g$ and $f^- \le h$.

X136. Let X be a set and let $f, g\colon X \to \mathbf{R}$. Prove that $f + g = (f \vee g) + (f \wedge g)$, $|f - g| = (f \vee g) - (f \wedge g)$,

$$f \vee g = \frac{f + g + |f - g|}{2} \qquad \text{and} \qquad f \wedge g = \frac{f + g - |f - g|}{2}.$$

X137. Let $f, g\colon [a,b] \to \mathbf{R}$ be Riemann integrable, where $a, b \in \mathbf{R}$. Prove that $f \vee g$ and $f \wedge g$ are Riemann integrable.

[24] To say that φ *is 1-periodic* means that for each $x \in \mathbf{R}$, $\varphi(x+1) = \varphi(x)$.

X138. Let $f, g\colon [a, b] \to \mathbf{R}^d$ be Riemann null and let $c \in \mathbf{R}$. Prove that $f + g$ and cf are both Riemann null.

X139. Let $f\colon [a, b] \to \mathbf{R}$, where $a, b \in \mathbf{R}$ with $a < b$. Prove that f is Riemann null iff f^+ and f^- are both Riemann null.

Definition. Let $f, g\colon [a, b] \to \mathbf{R}^d$, where $a, b \in \mathbf{R}$ with $a < b$. To say that $f = g$ *essentially* means that $f - g$ is Riemann null.

Exercise. Let $f, g, h\colon [a, b] \to \mathbf{R}^d$, where $a, b \in \mathbf{R}$ with $a < b$. Suppose that $f = g$ essentially and $g = h$ essentially. Prove that $f = h$ essentially.

Notation. Let $f\colon [a, b] \to \mathbf{R}$, where $a, b \in \mathbf{R}$ with $a < b$. Let $U(f, [a, b])$ denote the upper Riemann integral of f over $[a, b]$ and let $L(f, [a, b])$ denote the lower Riemann integral of f over $[a, b]$. If no confusion is likely to result, we may write $U(f)$ for $U(f, [a, b])$ and $L(f)$ for $L(f, [a, b])$.

X140. Let $f, g\colon [a, b] \to \mathbf{R}$, where $a, b \in \mathbf{R}$ with $a < b$.

(a) Prove that $U(f + g) \leq U(f) + U(g)$.

Similarly $L(f) + L(g) \leq L(f + g)$.

(b) Prove that for each $c \in (0, \infty)$, we have $U(cf) = cU(f)$ and $L(cf) = cL(f)$.

Similarly, for each $c \in (-\infty, 0)$, we have $U(cf) = cL(f)$ and $L(cf) = cU(f)$. Obviously $U(0 \cdot f) = 0 \cdot U(f) = L(0 \cdot f) = 0 \cdot L(f) = 0$.

Notation and Terminology. Let $a, b \in \mathbf{R}$ with $a < b$ and $E \subseteq [a, b]$. Let $E \subseteq [a, b]$. Then the *outer content of* E is $\gamma(E) = U(1_E, [a, b])$ and the *inner content of* E is $\gamma_\bullet(E) = L(1_E, [a, b])$ (It is not hard to check that these quantities do not depend on which interval $[a, b]$ containing E we consider, but we shall not need this fact.)

X141. Let X be a set and let $f, g\colon X \to \mathbf{R}$. Prove that $f = (f \wedge g) + (f - g)^+$ and $g = (f \vee g) - (f - g)^+$.

X142. Let $f\colon [a, b] \to \mathbf{R}^d$. Prove that f is Riemann null iff f is bounded and for each $\varepsilon > 0$, we have $\gamma(|f| > \varepsilon) = 0$. (Hints: Since only $|f|$ is involved, we may as well assume that $f\colon [a, b] \to [0, \infty)$. This will lighten the notation because then we can just write f instead of $|f|$. For the forward implication, notice that if $\varepsilon > 0$ then $\varepsilon 1_{\{f > \varepsilon\}} \leq f$. For the reverse implication, notice that $M = \sup f < \infty$ and that if $\varepsilon > 0$, then $f = (f \wedge \varepsilon) + (f - \varepsilon)^+$, $0 \leq f \wedge \varepsilon \leq \varepsilon$, and $(f - \varepsilon)^+ \leq M 1_{\{f > \varepsilon\}}$.)

The next exercise furnishes an example of a Riemann null function on $[0, 1]$ which is strictly positive on a dense subset of $[0, 1]$.

X143. Let $D = \mathbf{Q} \cap [0, 1]$. As we know, D is dense in $[0, 1]$. For each $n \in \mathbf{N}$, let

$$E_n = \left\{ \frac{0}{n}, \frac{1}{n}, \ldots, \frac{n}{n} \right\}.$$

Notice that $\bigcup_{n \in \mathbf{N}} E_n = D$. For each $x \in D$, let $N(x) = \min \{ n \in \mathbf{N} : x \in E_n \}$. Define $f\colon [0, 1] \to \mathbf{R}$ by

$$f(x) = \begin{cases} N(x)^{-1} & \text{if } x \in D, \\ 0 & \text{if } x \in [0, 1] \setminus D. \end{cases}$$

Prove that f is Riemann null. (Hint: This is very easy to do if you use problem X142. Or with a only little bit more work, it can be done directly from the definition of Riemann null.)

The next result is a variation on problem X122 and for most practical purposes is a satisfactory substitute for the result that exercise. The proof we suggest for it is different from the proof we suggested for problem X122 and you may feel that it is simpler than that proof.

X144. Let $f, g\colon [a, b] \to \mathbf{R}^d$ be bounded, where $a, b \in \mathbf{R}$ with $a < b$. Let $h\colon \mathbf{R}^d \to \mathbf{R}$ be continuous. Suppose that $f = g$ essentially. Prove that $h \circ f = h \circ g$ essentially. (Hint: Let K be the closure of the union of the ranges of f and g. Since f and g are bounded, K is a compact subset of $\mathbf{R}^d$, so h is uniformly continuous on K. Use problem X142.)

Example. Let $f_1, f_2, g_1, g_2\colon [a,b] \to \mathbf{C}$ be bounded and suppose that $f_1 = g_1$ essentially and $f_2 = g_2$ essentially. Define $f, g\colon [a,b] \to \mathbf{C}^2$ by $f(x) = (f_1(x), f_2(x))$ and $g(x) = (g_1(x), g_2(x))$. By problem X121, $f = g$ essentially. Then applying problem X144 with $h(z,w) = zw$, we find that $f_1 f_2 = g_1 g_2$ essentially.

Definition. Let $f, g\colon [a,b] \to \mathbf{R}$, where $a, b \in \mathbf{R}$ with $a < b$. To say that $f \le g$ *essentially* means that $(f-g)^+$ is Riemann null.

X145. Let $f, g\colon [a,b] \to \mathbf{R}$, where $a, b \in \mathbf{R}$ with $a < b$. Prove that the following are equivalent:
- **(a)** $f \le g$ essentially.
- **(b)** $f = f \wedge g$ essentially.
- **(c)** $g = f \vee g$ essentially.

X146. Let $f, g\colon [a,b] \to \mathbf{R}$, where $a, b \in \mathbf{R}$ with $a < b$. Prove that $f = g$ essentially iff $f \le g$ essentially and $g \le f$ essentially.

X147. Let $f, g, h\colon [a,b] \to \mathbf{R}$, where $a, b \in \mathbf{R}$ with $a < b$. Suppose that $f \le g$ essentially and $g \le h$ essentially. Prove that $f \le h$ essentially.

X148. Let $f, g\colon [a,b] \to \mathbf{R}$, where $a, b \in \mathbf{R}$ with $a < b$. Suppose that for each $\varepsilon > 0$, $f \le g + \varepsilon$ essentially. Prove that $f \le g$ essentially.

X149. Let $f_1, g_1, f_2, g_2\colon [a,b] \to \mathbf{R}$, where $a, b \in \mathbf{R}$ with $a < b$. Suppose that $f_1 \le g_1$ essentially and $f_2 \le g_2$ essentially. Prove that $f_1 + f_2 \le g_1 + g_2$ essentially.

Notation. Let $a, b \in \mathbf{R}$ with $a < b$. For each $f \in \mathscr{R}[a,b]$, we define the ∞-norm of f to be

$$\|f\|_\infty = \inf \{\, M \in [0,\infty) : |f| \le M \text{ essentially} \,\}.$$

Evidently $\|f\|_\infty \le \sup |f|$. It is not hard to see that if $f \in \mathscr{R}[a,b]$ and $c \in \mathbf{C}$, then $\|cf\|_\infty = |c| \|f\|_\infty$.

X150. Let $f\colon [a,b] \to \mathbf{C}$, where $a, b \in \mathbf{R}$ with $a < b$. Suppose f is continuous. Prove that $\|f\|_\infty = \sup |f|$.

X151. Let $f \in \mathscr{R}[a,b]$, where $a, b \in \mathbf{R}$ with $a < b$. Let $L = \|f\|_\infty$. Prove that $|f| \le L$ essentially.

X152. Let $f, g \in \mathscr{R}[a,b]$, where $a, b \in \mathbf{R}$ with $a < b$. Prove the following "limiting case" of Minkowski's inequality:

$$\|f + g\|_\infty \le \|f\|_\infty + \|g\|_\infty.$$

(Note: Even though we used the expression "limiting case," the most natural solution of this exercise does not involve taking a limit of p-norms as $p \to \infty$. But read on to learn what can be said about such a limit.)

Remark. In problem X126(d), on Hölder's inequality, one can easily check that $q = p/(p-1)$. Thus as $p \downarrow 1$, $q \uparrow \infty$. The next exercise justifies the notation $\|g\|_\infty$.

X153. Let $g \in \mathscr{R}[a,b]$, where $a, b \in \mathbf{R}$ with $a < b$. Prove that $\|g\|_q \to \|g\|_\infty$ as $q \to \infty$. (Hints: Since only $|g|$ is involved, we may as well assume that $g\colon [a,b] \to [0,\infty)$. This will lighten the notation because then we can just write g instead of $|g|$. Suppose $\infty > t > \|g\|_\infty$. Then $g \le t$ essentially, so by problem X144 and problem X145, for each $q \in [1,\infty)$, we have $g^q \le t^q$ essentially. Use this to prove that $\limsup_{q\to\infty} \|g\|_q \le t$. Next, suppose $0 \le s < \|g\|_\infty$. Then $(g-s)^+$ is not Riemann null. In other words, $U((g-s)^+) > 0$. Let $E = \{\, g > s \,\}$ and let $M = \sup g$. Then $(g-s)^+ \le M 1_E$, so $U(M 1_E) > 0$, so $U(1_E) > 0$. Now $g \ge s 1_E$. Use this to prove that $\liminf_{q\to\infty} \|g\|_q \ge s$.)

X154. Let $f, g \in \mathscr{R}[a,b]$, where $a, b \in \mathbf{R}$ with $a < b$. Prove the following "limiting case" of Hölder's inequality:

$$\|fg\|_1 \le \|f\|_1 \|g\|_\infty.$$

(Note: Even though we used the expression "limiting case," the most natural solution of this exercise does not involve taking a limit as $p \downarrow 1$ and $q = p/(p-1) \to \infty$.)

X155. Let $g\colon [0,1] \to \mathbf{R}$ be continuous. Suppose that

$$\int_0^1 x g(x)\, dx = 1 \qquad \text{and} \qquad \int_0^1 g(x)\, dx = 0.$$

Prove that $|g(x)| \ge 4$ for some $x \in [0,1]$.

Reminder. If E is a bounded subset of $\mathbf{R}$, then $\gamma(E)$ denotes the outer content of E and $\gamma_\bullet(E)$ denotes the inner content of E.

X156. Let $f\colon [a,b] \to \mathbf{R}$, where $a, b \in \mathbf{R}$ with $a < b$.

(a) Let $y, z \in \mathbf{R}$ with $y < z$. Obviously $\{\, f \geq z \,\} \subseteq \{\, f > y \,\}$, so $\gamma(f \geq z) \leq \gamma(f > y)$. But if f is Riemann integrable, then we can do better than this. Suppose that f is Riemann integrable. Prove that $\gamma(f \geq z) \leq \gamma_\bullet(f > y)$. (Hint: Let $h\colon \mathbf{R} \to \mathbf{R}$ be a continuous function such that $h = 0$ on $(-\infty, y]$, $0 \leq h \leq 1$ on $[y, z]$, and $h = 1$ on $[z, \infty)$. For instance, on $[y, z]$, we could take $h(t) = (t - y)/(z - y)$. Then $1_{\{f \geq z\}} \leq h \circ f \leq 1_{\{f > y\}}$.)

(b) Continue to suppose that f is Riemann integrable. Prove that $\gamma_\bullet(f > y) = \gamma(f \geq y)$ for all but countably many $y \in \mathbf{R}$. (Hint: This follows from part (a).) In particular, the set of all $y \in \mathbf{R}$ for which this equality holds is dense in $\mathbf{R}$.

X157. Let $f\colon [0,1] \to \mathbf{R}$ be continuously differentiable. Suppose that $f(0) = 0$. Show that

$$\int_0^1 |f(x)|^2 \, dx \leq \frac{1}{2} \int_0^1 |f'(x)|^2 \, dx.$$

Remark. We know that $\int_1^n 1/t \, dt = \log n$. Hence we expect that $\sum_{k=1}^n 1/k$ should be approximately $\log n$. The next exercise makes this precise.

X158. Prove that there is a unique real number γ such that for each $n \in \mathbf{N}$,

$$\gamma + \log n < \sum_{k=1}^{n} \frac{1}{k} < \frac{1}{2n} + \gamma + \log n.$$

Furthermore, $1/2 < \gamma < 1$. (Hint: Let

$$a_k = \frac{1}{k} - \int_k^{k+1} \frac{1}{t} \, dt.$$

Then

$$\sum_{k=1}^{n-1} \frac{1}{k} = \sum_{k=1}^{n-1} a_k + \log n.$$

Use the fact that the function $t \mapsto 1/t$ is strictly decreasing and strictly convex on $(0, \infty)$ to show that

$$\frac{1}{2}\left(\frac{1}{k} - \frac{1}{k+1}\right) < a_k < \frac{1}{k} - \frac{1}{k+1}.$$

You should find that $\gamma = \sum_{k=1}^\infty a_k$ works. Be sure to explain why this series converges. By the way, there is a picture that should help you understand how to do the proof. Notice that a_k is the area of an approximately triangular region of width 1 and height $1/k - 1/(k+1)$. If you slide each of these regions leftward, you can make them all fit, without overlapping, into the unit square $[0,1]^2$. It is visually clear that they then fill up a little more than half of this square.)

Remark. The number γ in problem X158 is called *Euler's constant.* To ten decimal places, its value is $\gamma = 0.5772156649\cdots$. It is not known whether γ is rational or irrational.

Remark. The function $x \mapsto e^x$ is convex on $\mathbf{R}$, so its graph lies above each of its tangent lines. Now the line with equation $y = 1 + x$ is tangent to the graph of $x \mapsto e^x$ at the point $(x, y) = (0, 1)$. Therefore

$$e^x \geq 1 + x \qquad \text{for each } x \in \mathbf{R}. \tag{11}$$

You should already be familiar with the inequality (11). This is just a reminder. Now let us observe some related inequalities that you may not have encountered before. The function $x \mapsto 2^x$ is convex on $\mathbf{R}$, so its graph on an interval $[a,b]$ lies below the chord joining the points $(a,2^a)$ and $(b,2^b)$. But when $a=0$ and $b=1$, this chord is a portion of the line with equation $y=1+x$. Thus

$$1+x \geq 2^x \qquad \text{for each } x \in [0,1]. \tag{12}$$

Similarly, the function $x \mapsto 4^{-x}$ is convex on $\mathbf{R}$, so its graph on an interval $[a,b]$ lies below the chord joining the points $(a,4^{-a})$ and $(b,4^{-b})$. When $a=0$ and $b=1/2$, this chord is a portion of the line with equation $y=1-x$. Therefore

$$1-x \geq 4^{-x} \qquad \text{for each } x \in [0,1/2]. \tag{13}$$

X159. Suppose that $a_k \in [0,\infty)$ for each $k \in \mathbf{N}$. Let $p_n = \prod_{k=1}^n (1+a_n)$ for each $n \in \mathbf{N}$. Notice that $1 \leq p_1 \leq p_2 \leq p_3 \leq \cdots < \infty$. Let $p = \lim_{n\to\infty} p_n$ and let $s = \sum_{k=1}^\infty a_k$. Obviously $p \in [1,\infty]$ and $s \in [0,\infty]$. Prove that $p<\infty$ iff $s<\infty$. (Hint: Use (11) and (12).)

X160. Suppose that $b_k \in [0,1)$ for each $k \in \mathbf{N}$. Let $q_n = \prod_{k=1}^n (1-b_n)$ for each $n \in \mathbf{N}$. Notice that $1 \geq q_1 \geq q_2 \geq q_3 \geq \cdots > 0$. Let $q = \lim_{n\to\infty} q_n$ and $t = \sum_{k=1}^\infty b_k$. Obviously $q \in [0,1]$ and $t \in [0,\infty]$. Prove that $q>0$ iff $t<\infty$. (Hint: Use (11) and (13).)

X161. Let P be the set of prime numbers. Prove that *due 5Th*

$$\sum_{p\in P} \frac{1}{p} = \infty.$$

Hint: Let $n \in \{2,3,4,\dots\}$. Let P_n be the set of all $p \in P$ such that there exists $k \in \{2,\dots,n\}$ such that p divides k. Now $2^n > n$. Hence for each $k \in \{1,\dots,n\}$, there exist numbers $m_p \in \{0,1,2,\dots,n\}$ such that $k = \prod_{p\in P_n} p^{m_p}$. Therefore

$$\sum_{k=1}^n \frac{1}{k} \leq \prod_{p\in P_n} (1+p^{-1}+p^{-2}+\cdots+p^{-n}) \leq \prod_{p\in P_n} \frac{1}{1-p^{-1}} = \left(\prod_{p\in P_n} (1-p^{-1})\right)^{-1}.$$

Now use this and (13) to show that $\sum_{p\in P_n} \frac{1}{p}$ is greater than or equal to a certain quantity which tends to infinity as n tends to infinity.

Remark. For each $n \in \mathbf{N}$, let p_n be the n-th prime number, so that $p_1=2$, $p_2=3$, $p_3=5$, and so on. In problem X161, you were asked to show that $\sum_{n=1}^\infty 1/p_n = \infty$. This shows that p_n tends to infinity fairly slowly. Now there is a famous, difficult result, called *the prime number theorem*, which gives much more exact information about p_n. The usual formulation of the prime number theorem is that

$$\pi(x) \sim \frac{x}{\log x} \qquad \text{as } x \to \infty, \tag{14}$$

where $\pi(x)$ denotes the number of primes which are less than or equal to x. In the next exercise, you are asked to show that an equivalent statement is that

$$p_n \sim n \log n \qquad \text{as } n \to \infty. \tag{15}$$

It is interesting to remark that (14) was conjectured by Carl Friedrich Gauss in 1792 or 1793, when he was around 16 years old, by perusing tables of prime numbers. Gauss did not prove (14). P. L. Chebyshev made progress toward proving it in the 1850s, and it was finally proved in 1896, by Jacques Hadamard and Charles de la Vallée-Poussin independently.

X162. Let (p_n) be a strictly increasing sequence in $(0,\infty)$. For each $x \in (0,\infty)$, let

$$P(x) = \{\, n \in \mathbf{N} : p_n \leq x \,\}$$

and let $\pi(x)$ be the number of elements in $P(x)$. Prove that the for the given sequence (p_n), (14) holds if and only if (15) holds. (Hints: For each $x \in [p_n, p_{n+1})$, we have $\pi(x) = n$. If (14) holds, then $\pi(x) < \infty$ for large x, and hence each $x \in (0,\infty)$, and it follows that $p_n \to \infty$.)

Remark. In view of problem X162, you may be wondering why the prime number theorem is usually expressed in the form (14) rather than the more easily grasped form (15). The reason is that in the case where (p_n) is the sequence of prime numbers listed in increasing order, although the error in each of these two approximations tends to zero, the error in the approximation (14) tends to zero faster.

Remark. If the famous *Riemann hypothesis* holds, then an even better approximation is obtained in (14) if we replace $x/\log x$ by $L(x) = \sum_{n=2}^{x}(\log n)^{-1}$. In this case, one has $\pi(x) = L(x) + O(x^{1/2}\log x)$ as $x \to \infty$. For a fascinating discussion of the prime number theorem and related matters, without proofs, but with heuristic explanations, see Serge Lang, *The Beauty of Doing Mathematics*, *Three Public Dialogues*, Springer-Verlag, 1985.

Remark. We know that as $h \to 0$, $(1+h)^{1/h} \to e$. It is sometimes of interest to know that

$$(1+h)^{1/h} \uparrow e \qquad \text{as } h \downarrow 0$$

and that

$$(1+h)^{1/h} \downarrow e \qquad \text{as } h \uparrow 0.$$

One way to show this would be to verify, by calculation, that

$$\frac{d}{dh}(1+h)^{1/h} < 0 \qquad \text{for } 0 < h < \infty.$$

However, there is a more enlightening and elegant way which is based on things we have already checked. We know that

$$\frac{d}{du}\log u = \frac{1}{u} \qquad \text{is strictly decreasing on the interval where } 0 < u < \infty.$$

Thus the function $u \mapsto \log u$ is strictly concave down on the interval $(0,\infty)$. In other words, the function $u \mapsto -\log u$ is strictly convex on $(0,\infty)$. Hence, if $0 < u < v < w < \infty$, then considering the slopes of suitable chords, we have

$$\frac{\log v - \log u}{v-u} > \frac{\log w - \log u}{w-u} > \frac{\log w - \log v}{w-v}. \tag{16}$$

(You should draw a picture!) If $0 < h_1 < h_2 < \infty$, then taking $u = 1$, $v = 1 + h_1$, and $w = 1 + h_2$, and applying the first inequality in (16), we get

$$\frac{\log(1+h_1)}{h_1} > \frac{\log(1+h_2)}{h_2},$$

or equivalently,

$$(1+h_1)^{1/h_1} > (1+h_2)^{1/h_2}.$$

Thus the map $h \mapsto (1+h)^{1/h}$ is strictly decreasing on the interval $(0,\infty)$. Similarly, if $-1 < h_1 < h_2 < 0$, then taking $u = 1 + h_1$, $v = 1 + h_2$, and $w = 1$, and applying the second inequality in (16), the we get

$$\frac{-\log(1+h_1)}{-h_1} > \frac{-\log(1+h_2)}{-h_2},$$

or again equivalently,

$$(1+h_1)^{1/h_1} > (1+h_2)^{1/h_2}.$$

Thus the map $h \mapsto (1+h)^{1/h}$ is also strictly decreasing on the interval $(-1,0)$. Putting all this together, we see that if we let

$$\psi(h) = \begin{cases} (1+h)^{1/h} & \text{if } h \in (-1,0) \cup (1,\infty), \\ e & \text{if } h = 0, \end{cases}$$

then ψ is continuous and strictly decreasing on the interval $(-1,\infty)$.

X163. *due 5Th*

(a) The gamma function is defined for $0 < x < \infty$ by

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

For $n = 1, 2, 3, \ldots$ and for $0 < x < \infty$, let

$$I_n(x) = \int_0^n t^{x-1} \left(1 - \frac{t}{n}\right)^n dt.$$

Prove that for each $x \in (0, \infty)$, we have $\lim_{n\to\infty} I_n(x) = \Gamma(x)$. (Hint: Fix $x \in (0, \infty)$. Let $f(t) = t^{x-1}e^{-t}$ for all $t \in (0, \infty)$. For $n = 1, 2, 3, \ldots$, let

$$f_n(t) = \begin{cases} t^{x-1}\left(1 - \frac{t}{n}\right)^n & \text{if } t \in (0, n), \\ 0 & \text{if } t \in [n, \infty). \end{cases}$$

Use the preceding Remark to show that for each $t \in (0, \infty)$, $f_n(t) \uparrow f(t)$ as $n \to \infty$. Then use Dini's theorem[25] to show that if $0 < a < b < \infty$, then $f_n \to f$ uniformly on $[a, b]$. This uniform convergence could be established by calculation rather than by appealing to Dini's theorem. However, by using Dini's theorem, we save ourselves the trouble of calculating.)

(b) For $-1 < z < \infty$, let $\Pi(z) = \Gamma(z + 1)$. Using integration by parts, one can easily see that for $z = 0, 1, 2, \ldots$, we have $\Pi(z) = z!$. For this reason, the function Π is called the factorial function. By explicitly evaluating $I_n(x)$, deduce from part (a) that

$$\begin{aligned}\Gamma(z+1) &= \lim_{n\to\infty} \frac{n!(n+1)^z}{(z+1)(z+2)\cdots(z+n)} \\ &= \lim_{n\to\infty} \prod_{k=1}^{n} \left[\left(\frac{k+1}{k}\right)^z \frac{k}{z+k}\right] \\ &= \left[\left(\frac{2}{1}\right)^z \frac{1}{z+1}\right]\left[\left(\frac{3}{2}\right)^z \frac{2}{z+2}\right]\left[\left(\frac{4}{3}\right)^z \frac{3}{z+3}\right]\cdots\end{aligned}$$

for each $z \in (-1, \infty)$. It was in the form of the latter infinite product that Euler (in 1729) invented the factorial function.[26]

X164. Let $f \in \mathscr{R}[0, 1]$ and suppose f is continuous at 1. Determine

$$\lim_{n\to\infty} n \int_0^1 x^n f(x)\, dx.$$

Justify your answer. (Hint: If you understand thoroughly how to solve problem X131, then you should be able to solve this problem. Vice versa too.)

---

[25] "Dini's theorem" is the customary name for a theorem like Theorem 7.13 in Rudin, *Principles of Mathematical Analysis*, Third Edition. Of course, the sequence of functions in question can just as well be increasing, rather than decreasing as in the theorem in Rudin.

[26] If you know some complex analysis, then you may find the following remarks interesting. It not hard to show that the product whose factors are the reciprocals of the factors in the product for $\Gamma(z+1)$ converges uniformly on each compact subset of $\mathbf{C}$. From this it follows that the reciprocal of the gamma function can be extended in a unique way to a function that is analytic in the whole complex plane. The set of zeroes of this extension is the set of non-positive integers and each of these is a simple zero. Thus the gamma function extends in a unique way to a function that is meromorphic in the whole complex plane. This extension has no zeroes and its set of poles is the set of non-positive integers and each of these is a simple pole. If you didn't understand some of the terms in this footnote, don't worry. You will learn them when you study complex analysis.

Remark. Let $u, v, w \in \mathbf{R}$ with $u < v < w$. As we know, if $f\colon [u,w] \to \mathbf{R}$ is Riemann integrable, then

$$\int_u^w f(x)\,dx = \int_u^v f(x)\,dx + \int_v^w f(x)\,dx.$$

The same argument that establishes this actually proves that the upper and lower Riemann integrals satisfy

$$U(f,[u,w]) = U(f,[u,v]) + U(f,[v,w]) \qquad \text{and} \qquad L(f,[u,w]) = L(f,[u,v]) + L(f,[v,w]).$$

Now let E be a bounded subset of $\mathbf{R}$. Suppose $E \subseteq [a_1,b_1]$ and $E \subseteq [a_2,b_2]$, where $a_1, b_1, a_2, b_2 \in \mathbf{R}$ with $a_1 < b_1$ and $a_2 < b_2$. Let $a = \max\{\,a_1, a_2\,\}$ and $b = \min\{\,b_1, b_2\,\}$. Then $a \le b$ and $E \subseteq [a,b]$. Notice that

$$U(1_E,[a_1,b_1]) = U(1_E,[a_1,a]) + U(1_E,[a,b]) + U(1_E,[b,b_1]) = 0 + U(1_E,[a,b]) + 0 = U(1_E,[a,b]).$$

Likewise, $U(1_E,[a_2,b_2]) = U(1_E,[a,b])$. Hence $U(1_E,[a_1,b_1]) = U(1_E,[a_2,b_2])$. Similarly, $L(1_E,[a_1,b_1]) = L(1_E,[a_2,b_2])$. In other words, $\gamma(E)$ and $\gamma_\bullet(E)$, the inner and outer content of E, do not depend on which compact interval containing E they are computed with respect to.

Definition. Let E be a subset of $\mathbf{R}$. To say that E *is contented* means that E is bounded and $\gamma_\bullet(E) = \gamma(E)$. (This somewhat whimsical terminology is not standard.)

Remark. If E is a bounded interval in $\mathbf{R}$, then E is contented and $\gamma(E)$ is equal to the length of E.

Remark. If $E \subseteq [a,b]$, where $a, b \in \mathbf{R}$ with $a < b$, then E is contented iff 1_E is Riemann integrable over $[a,b]$.

Remark. If E is a contented subset of $\mathbf{R}$, then $\gamma(E)$, the outer content of E, is also called simply *the content of* E.

X165. Let $E_1, \ldots, E_n$ be contented subsets of $\mathbf{R}$. Let $E = \bigcup_{k=1}^n E_k$ and $D = \bigcap_{k=1}^n E_k$. Prove that E and D are contented.

X166. Let A and B be contented subsets of $\mathbf{R}$. Prove that $A \setminus B$ is contented.

X167. Let $E_1, \ldots, E_n$ be contented subsets of $\mathbf{R}$ and let $E = \bigcup_{k=1}^n E_k$. Suppose that $E_1, \ldots, E_n$ are disjoint. Prove that $\gamma(E) = \sum_{k=1}^n \gamma(E_k)$.

X168. Let $E_1, \ldots, E_n$ be bounded subsets of $\mathbf{R}$.

- **(a)** Let $E \subseteq \bigcup_{k=1}^n E_k$. Prove that $\gamma(E) \le \sum_{k=1}^n \gamma(E_k)$.
- **(b)** Now suppose in addition that $E_1, \ldots, E_n$ are disjoint. Let F be a bounded subset of $\mathbf{R}$ such that $\bigcup_{k=1}^n E_k \subseteq F$. Prove that $\sum_{k=1}^n \gamma_\bullet(E_k) \le \gamma_\bullet(F)$.

X169. Let E be a bounded subset of $\mathbf{R}$.

- **(a)** Prove that $\gamma(E)$ is the infimum of sums of the form $\sum_{k=1}^n \gamma(I_k)$, where $n \in \mathbf{N}$, $I_1, \ldots I_n$ are bounded intervals in $\mathbf{R}$, and $E \subseteq \bigcup_{k=1}^n I_k$.
- **(b)** Prove that $\gamma_\bullet(E)$ is the supremum of sums of the form $\sum_{k=1}^n \gamma(I_k)$, where $n \in \mathbf{N}$ and where $I_1, \ldots I_n$ are disjoint intervals in $\mathbf{R}$ which are contained in E.

X170. Let E be a bounded subset of $\mathbf{R}$. Prove that

$$\gamma(E) = \inf\{\,\gamma(F) : E \subseteq F \text{ and } F \text{ is contented}\,\}$$

and that

$$\gamma_\bullet(E) = \sup\{\,\gamma(D) : D \subseteq E \text{ and } D \text{ is contented}\,\}.$$

X171. Let E be a contented subset of $\mathbf{R}$, let $A \subseteq E$, and let $B = E \setminus A$. Prove that $\gamma(E) = \gamma(A) + \gamma_\bullet(B)$. (Note that neither A nor B need be contented.)

Reminder. Let (Z, d) be a metric space. Let $f\colon I \to Z$, where I is an interval in $\mathbf{R}$. Let $a, b \in I$ with $a < b$. Recall that *a subdivision of* $[a, b]$ is an ordered n-tuple $(x_0, x_1, \ldots, x_n)$ of real numbers such that $a = x_0 < x_1 < \cdots < x_n = b$. Recall that if $P = (x_0, x_1, \ldots, x_n)$ is a subdivision of $[a, b]$, then *the variation of* f *on* P is

$$V(f, P) = \sum_{k=1}^{n} d\big(f(x_{k-1}), f(x_k)\big). \tag{17}$$

Let's write $\Pi[a, b]$ for the set of subdivisions of $[a, b]$. Recall that the variation of f on $[a, b]$ is

$$V(f, [a, b]) = \sup\{\, V(f, P) : P \in \Pi[a, b] \,\}.$$

Now let us specialize to the case where $Z = \mathbf{R}$ and $d(z, z') = |z' - z|$. Then (17) becomes

$$V(f, P) = \sum_{k=1}^{n} \big|f(x_k) - f(x_{k-1})\big|.$$

X172. Let $f\colon [a, b] \to \mathbf{R}$, where $a, b \in \mathbf{R}$ with $a < b$. Prove that $V(f, [a, b]) < \infty$ iff there exist increasing functions $\varphi, \psi\colon [a, b] \to \mathbf{R}$ such that $f = \varphi - \psi$. (Hint for the forward implication: Define φ and ψ by $\varphi(x) = V(f, [a, x])$ and $\psi(x) = \varphi(x) - f(x)$.)

Remark. Let $f\colon [a, b] \to \mathbf{R}$, where $a, b \in \mathbf{R}$ with $a < b$. In problem X172, we saw that $V(f, [a, b]) < \infty$ iff f is a difference of increasing functions on $[a, b]$. This decomposition of f is clearly not unique. Furthermore, the decomposition suggested in problem X172 is not optimal. The next exercise will explain what we mean by optimal here and will reveal the optimal way of choosing such a decomposition. First we need some preliminaries.

Reminder. Let y be a real number. Then $y^+ = \max\{\, y, 0 \,\}$ and $y^- = \max\{\, -y, 0 \,\}$. We have

$$y = y^+ - y^-, \qquad |y| = y^+ + y^-, \qquad y^+ = \frac{|y| + y}{2}, \qquad y^- = \frac{|y| - y}{2}. \tag{18}$$

Furthermore, if $y = u - v$, where $u, v \in [0, \infty)$, then $y^+ \le u$ and $y^- \le v$.

Definitions. Let $f\colon I \to \mathbf{R}$, where I is an interval in $\mathbf{R}$. Let $a, b \in I$ with $a < b$. If $P = (x_0, x_1, \ldots, x_n)$ is a subdivision of $[a, b]$, then by definition, *the positive variation of* f *on* P is

$$V^+(f, P) = \sum_{k-1}^{n} \big(f(x_k) - f(x_{k-1})\big)^+$$

and *the negative variation of* f *on* P is

$$V^-(f, P) = \sum_{k-1}^{n} \big(f(x_k) - f(x_{k-1})\big)^-.$$

In view of (18), it is obvious that

$$V^+(f, P) + V^-(f, P) = V(f, P)$$

and that

$$V^+(f, P) - V^-(f, P) = \sum_{k=1}^{n} \big(f(x_k) - f(x_{k-1})\big) = f(x_n) - f(x_0) = f(b) - f(a),$$

so that

$$V^+(f, P) = \frac{V(f, P) + f(b) - f(a)}{2}$$

and

$$V^-(f,P) = \frac{V(f,P) + f(a) - f(b)}{2}.$$

Next, by definition, *the positive variation of* f *on* $[a,b]$ is

$$V^+(f,[a,b]) = \sup\left\{\, V^+(f,P) : P \in \Pi[a,b] \,\right\}$$

and *the negative variation of* f *on* $[a,b]$ is

$$V^-(f,[a,b]) = \sup\left\{\, V^-(f,P) : P \in \Pi[a,b] \,\right\}.$$

X173. Let $f\colon [a,b] \to \mathbf{R}$, where $a, b \in I$ with $a < b$.

(a) Prove that

$$V^+(f,[a,b]) = \frac{V(f,[a,b]) + f(b) - f(a)}{2},$$

$$V^-(f,[a,b]) = \frac{V(f,[a,b]) + f(a) - f(b)}{2},$$

and

$$V(f,[a,b]) = V^+(f,[a,b]) + V^-(f,[a,b]).$$

In particular, if one of the quantities $V(f,[a,b])$, $V^+(f,[a,b])$, and $V^-(f,[a,b])$ is finite, then all three of them are finite.

(b) Prove that if $c \in (a,b)$, then

$$V^+(f,[a,c]) + V^+(f,[c,b]) = V^+(f,[a,b])$$

and

$$V^-(f,[a,c]) + V^-(f,[c,b]) = V^-(f,[a,b])$$

Define $F, g, h\colon [a,b] \to [0,\infty]$ by $F(x) = V(f,[a,x])$, $g(x) = V^+(f,[a,x])$, and $h(x) = V^-(f,[a,x])$. Obviously $F(a) = g(a) = h(a) = 0$. Notice that by part (a), with b replaced by x, we have

$$g(x) = \frac{F(x) + f(x) - f(a)}{2}, \qquad h(x) = \frac{F(x) + f(a) - f(x)}{2},$$

and $F(x) = g(x) + h(x)$ for each $x \in [a,b]$.

(c) We already know that F is increasing. Prove that g and h are increasing too.

(d) Prove that if $V(f,[a,b]) < \infty$, then $f = f(a) + g - h$.

(e) Suppose that $f = f(a) + G - H$, where $G, H\colon [a,b] \to \mathbf{R}$ are increasing and satisfy $G(a) = H(a) = 0$. Prove that $V(f,[a,b]) < \infty$, that $g \le G$ and that $h \le H$. Hence $F \le G + H$, with equality iff $G = g$ and $H = h$.

X174. Let $f\colon [0,1] \to \mathbf{R}$ be continuous. Suppose that *do before MT*

$$\int_0^1 f(x) g'(x)\, dx = 0$$

for each continuously differentiable function $g\colon [0,1] \to \mathbf{R}$ satisfying $g(0) = 0 = g(1)$. Prove that f must be a constant function. (Hint: One approach to solving this problem may be based on a thorough understanding of problem X131 and/or problem X164. I don't mean that the specific form of the functions considered in those problems is important. Rather, their qualitative behaviour is what is important.)

X175. Define complex-valued functions functions f and g by *do before MT*

$$f(t) = \sum_{k=1}^{\infty} \frac{e^{ikt}}{k^2} \qquad \text{and} \qquad g(t) = \sum_{k=1}^{\infty} \frac{e^{ikt}}{k}.$$

Let each of these functions be defined for all $t \in \mathbf{R}$ for which the corresponding series converges.

(a) Prove that the series for f converges uniformly on $\mathbf{R}$.

(b) Let's write $2\pi\mathbf{Z}$ for the set $\{ 2\pi n : n \in \mathbf{Z} \}$. Clearly for each $t \in 2\pi\mathbf{Z}$, the series for $g(t)$ reduces to $\sum_{k=1}^{\infty} 1/k$, which diverges. Let $A = \mathbf{R} \setminus 2\pi\mathbf{Z}$. (Notice that A is a dense open subset of $\mathbf{R}$.) Prove that the series for g converges locally uniformly in A.[27] (Hint: Recall that if $z \in \mathbf{C}$ with $|z| = 1$ but $z \neq 1$, then the series $\sum_{k=1}^{\infty} z^k$ has bounded partial sums. Review the proof of this fact and adapt it to prove a related uniform result. Then apply summation by parts.)

(c) Show that f is differentiable on A and that $f' = ig$ on A.

(d) Consider the functions defined by the series

$$\sum_{k=1}^{\infty} \frac{\cos(kt)}{k^2} \qquad \text{and} \qquad \sum_{k=1}^{\infty} \frac{\sin(kt)}{k^2}.$$

What does part (c) tell us about the derivatives of these functions?

Let X be a topological space, let Y be a metric space, and let Φ be a collection of functions from X into Y. If $a \in X$, then to say that Φ *is equicontinuous at* a means that for each $\varepsilon > 0$, there exists a nhd U of a in X such that for each $f \in \Phi$, for each $x \in U$, $d(f(x), f(a)) < \varepsilon$. To say that Φ *is equicontinuous on* X means that for each $a \in X$, Φ is equicontinuous at a. And by the way, although we don't need this concept right now, let us mention that if X is also a metric space, then to say that Φ *is uniformly equicontinuous on* X means that for each $\varepsilon > 0$, there exists $\delta > 0$ such that for each $f \in \Phi$, for all $x, x' \in X$, if $d_X(x, x') < \delta$, then $d_Y(f(x), f(x')) < \varepsilon$. (Warning: In Rudin, *Principles of Mathematical Analysis*, Third Edition, the term "equicontinuous on X" is used instead of "uniformly equicontinuous on X." When the metric space X is compact, which is the main case Rudin is interested in, one can show that if Φ is equicontinuous on X, then Φ is uniformly equicontinuous on X. Perhaps this is why Rudin does not bother to distinguish between these two notions. The proof of the fact just mentioned is almost word-for-word the same as the proof that a single continuous function from a compact metric space into a metric space is uniformly continuous.)

X176. Let X be a topological space, let Y be a complete metric space, let Φ be a collection of functions from X into Y, let $\mathscr{B}$ be a filter base on Φ, and let D be a dense subset of X. Suppose Φ is equicontinous on X and $\lim_{f \to \mathscr{B}} f(x)$ exist in Y for each $x \in D$. *do before MT*

(a) Show that $\lim_{f \to \mathscr{B}} f(x)$ exists for each $x \in X$.

(b) Define $g \colon X \to Y$ by $g(x) = \lim_{f \to \mathscr{B}} f(x)$ for each $x \in X$. Show that g is continuous.

(c) State what parts (a) and (b) tell us in the case of a sequence (f_n) of functions from X into Y.

[27] To say that a series converges *locally uniformly in* A means that each point in A has a nhd in A in which the series converges uniformly.

Reminder. Let E be a bounded subset of $\mathbf{R}$. Then $\gamma(E)$ denotes the outer content of E and $\gamma_\bullet(E)$ denotes the inner content of E. To say that E is contented means that $\gamma(E) = \gamma_\bullet(E)$. If E is contented, then $\gamma(E)$ is also called simply the content of E.

X177.

(a) Let E be a bounded subset of $\mathbf{R}$. Prove that $\gamma(E)$ is the infimum of the set of numbers of the form $\gamma(U)$, where U varies over all bounded open subsets of $\mathbf{R}$ containing E. (In fact, it is enough to consider sets U which are unions of finitely many bounded open intervals.)

(b) Let (K_n) be a decreasing sequence of compact subsets of $\mathbf{R}$ and let $K = \bigcap_{n=1}^{\infty} K_n$. Prove that $\gamma(K_n) \downarrow \gamma(K)$.

Notation. The following notation is not standard but will be convenient to use in the next few exercises. For each $n \in \mathbf{N}$, let $S_n = \{0,1\}^n$ be the set of finite binary sequences of length n. Let $S = \bigcup_{n\in\mathbf{N}} S_n$, the set of non-empty finite binary sequences. Finally, let $\Sigma = \{0,1\}^{\mathbf{N}}$ be the set of infinite binary sequences.

Reminder. Let S_n, S, and Σ be as above. If (X,d) is a metric space, then by a *Cantor scheme*[28] in (X,d), we shall mean a family $(A(s))_{s\in S}$ of non-empty subsets of X, indexed by S, such that:

(a) $A(0)$ and $A(1)$ are disjoint;

(b) For each $n \in \mathbf{N}$ and each $s \in S_n$, $A(s_1,\dots,s_n,0)$ and $A(s_1,\dots,s_n,1)$ are disjoint and for $i = 0, 1$,

$$A(s_1,\dots,s_n) \supseteq \overline{A(s_1,\dots,s_n,i)};$$

(c) For each $\sigma = (s_1, s_2, s_3, \dots) \in \Sigma$, $\operatorname{diam}(A(s_1,\dots,s_n)) \to 0$ as $n \to \infty$.

X178. Let S_n, S, and Σ be as above. Let (X,d) be a complete metric space and let $(A(s))_{s\in S}$ be a Cantor scheme in (X,d). For each $n \in \mathbf{N}$, let $B_n = \bigcup_{s\in S_n} A(s)$. Let $K = \bigcap_{n\in\mathbf{N}} B_n$. For each $\sigma = (s_1, s_2, s_3, \dots) \in \Sigma$, let $f(\sigma)$ be the unique point in $\bigcap_n A(s_0,\dots,s_n)$. (We can do this by problem X9.) We saw in problem X12 that h is a one-to-one map from Σ onto K and consequently that K is equinumerous to the interval $[0,1]$. Give Σ the product topology.

(a) Prove that h is a homeomorphism from Σ onto K.

(b) Deduce that K is homeomorphic to the ordinary Cantor set.

The ordinary Cantor set has zero total length, in the sense that its outer content is zero. In the extra problems last quarter, we mentioned that there are so-called *fat Cantor sets* which resemble the ordinary Cantor set topologically but which have strictly positive total length. The next exercise treats these sets in detail.

X179. Let S_n, S, and Σ be as above. For all $a, b \in \mathbf{R}$ with $a < b$ and for all $u \in (0,1)$, let $M(a,b,u)$ be the open interval of length $(b-a)(1-u)$ centered at the midpoint of $[a,b]$ and let $J([a,b],u,0)$ and $J([a,b],u,1)$ be the left and right portions of $[a,b]$ that remain after we delete $M(a,b,u)$ from $[a,b]$. Fix $t \in (0,1)$. Let (t_n) be a strictly decreasing sequence in $(t,1)$ such that $t_n \to t$ and let $t_0 = 1$. For each $n \in \mathbf{N}$, let $u_n = t_n/t_{n-1}$. Notice that $u_n \in (0,1)$ and $\prod_{k=1}^{n} u_k = t_n$ for each $n \in \mathbf{N}$. Let $I(0) = J([0,1],u_1,0)$ and let $I(1) = J([0,1],u_1,1)$. If $n \in \mathbf{N}$, $s \in S_n$, and $I(s)$ has already been defined, let $I(s,0) = J(I(s),u_{n+1},0)$ and let $I(s,1) = J(I(s),u_{n+1},1)$. For each $n \in \mathbf{N}$, let $F_n = \bigcup_{s\in S_n} I(s)$. Let $F = \bigcap_{n\in\mathbf{N}} F_n$.

(a) Prove that F is compact and has empty interior.

(b) Prove that F is homeomorphic to the ordinary Cantor set. In particular, F has no isolated points. In fact, each non-empty relatively open subset of F has the cardinality of the continuum.

(c) Prove that for each $n \in \mathbf{N}$, $\gamma(F_n) = t_n$.

(d) Prove that $\gamma(F) = t$.

(e) Prove that 1_F is not Riemann integrable over $[0,1]$.

Reminder. Let X be a topological space and let $f\colon X \to [-\infty,\infty]$. For each $p \in X$, let $\mathscr{U}_p$ be the collection of nhds of p in X. Of course $\mathscr{U}_p$ is a filter base on X. The lower regularization of f is the function f_* defined on X by

$$f_*(p) = \liminf_{\mathscr{U}_p} f = \sup_{U\in\mathscr{U}_p} \inf_{x\in U} f(x).$$

---

[28] This is not standard terminology.

The upper regularization of f is the function f^* defined on X by

$$f^*(p) = \limsup_{\mathscr{U}_p} f = \inf_{U \in \mathscr{U}_p} \sup_{x \in U} f(x).$$

We saw in problem X56 that f_* is the largest lower semicontinuous function on X which is less than or equal to f, that f^* is the smallest upper semicontinuous function on X which is greater than or equal to f, and that for each $p \in X$, f is continuous at p iff $f_*(p) = f^*(p)$.

X180. Let F be a fat Cantor set and let $f = 1_F$. Prove that $f_* = 0$ and $f^* = f$.

X181. Let $f\colon [a,b] \to \mathbf{R}$. *due 7Th*

(a) Prove that the lower Riemann integral of f over $[a,b]$ is the same as the lower Riemann integral of f_* over $[a,b]$ and that the upper Riemann integral of f over $[a,b]$ is the same as the upper Riemann integral of f^* over $[a,b]$.

(b) Prove that f is Riemann integrable over $[a,b]$ iff f is bounded and $f^* - f_*$ is Riemann null.

Remark. In a sense, problem X181 tells us that a Riemann integrable function must be nearly continuous. This is because the set where f is continuous is the set where $f^* - f_* = 0$. When we study Lebesgue measure, we'll make this more precise. In fact, once we have Lebesgue measure, it follows almost immediately from problem X181 that f is Riemann integrable iff f is bounded and the set where f is not continuous has Lebesgue measure zero.

Definition. Let Y be a set and let $\varphi\colon [a,b] \to Y$, where $a, b \in \mathbf{R}$ with $a < b$. To say that φ *is Riemann simple* means that the range of φ is a finite set and that for each $y \in Y$, the set

$$\{\, \varphi = y \,\} = \{\, x \in [a,b] : \varphi(x) = y \,\}$$

is contented.

X182. Let $\varphi\colon [a,b] \to \mathbf{R}$, where $a, b \in \mathbf{R}$ with $a < b$. Suppose that φ is Riemann simple. Prove that φ is Riemann integrable over $[a,b]$ and that $\int_a^b \varphi(x)\,dx = \sum_{y \in Y} y\gamma(\varphi = y)$, where Y denotes the range of φ.

X183. Let $f\colon [a,b] \to \mathbf{R}$, where $a, b \in \mathbf{R}$ with $a < b$. Prove that f is Riemann integrable iff there exists a sequence (φ_n) of Riemann simple functions $\varphi_n\colon [a,b] \to \mathbf{R}$ such that $\varphi_n \to f$ uniformly on $[a,b]$. (Hints: The reverse implication is the easy part, given the result of problem X182. To prove the forward implication, let $m = \inf f$ and $M = \sup f$, use problem X156 with problem X166 and partition the interval $[m, M]$ appropriately.)

Reminder. If $f\colon [a,b] \to \mathbf{R}$, where $a, b \in \mathbf{R}$ with $a < b$, then $U(f, [a,b])$ denotes the upper Riemann integral of f over $[a,b]$ and $L(f, [a,b])$ denotes the lower Riemann integral of f over $[a,b]$. If no confusion is likely to result, we may write $U(f)$ for $U(f, [a,b])$ and $L(f)$ for $L(f, [a,b])$.

X184. Let $f\colon [a,b] \to \mathbf{R}^d$, where $a, b \in \mathbf{R}$ with $a < b$. Prove that f is Riemann integrable over $[a,b]$ iff there exists a sequence of step functions (φ_n) such that $U(|f - \varphi_n|) \to 0$, and that in this case, $\int_a^b \varphi_n(t)\,dt \to \int_a^b f(t)\,dt$.

Reminder. $\mathbf{K}$ denotes either $\mathbf{R}$ or $\mathbf{C}$.

X185. Let $\varphi\colon [a,b] \to Y$, where $a, b \in \mathbf{R}$ with $a < b$ and where Y is a vector space over $\mathbf{K}$. Suppose φ is a step function. Then it is obvious what $\int_a^b \varphi(x)\,dx$ should be. The only question is whether this is well-defined (because there are many subdivisions of $[a,b]$ such that φ is constant on the interior of each subinterval associated with the subdivision). Show that $\int_a^b \varphi(x)\,dx$ is well-defined. (Hint: For any two subdivisions of $[a,b]$, we can consider a common refinement.)

X186. Let $\varphi, \psi\colon [a,b] \to Y$, where $a, b \in \mathbf{R}$ with $a < b$ and where Y is a vector space over $\mathbf{K}$. Suppose that φ and ψ are step functions.

(a) Prove that $\varphi + \psi$ is a step function and that

$$\int_a^b \varphi(x) + \psi(x)\,dx = \int_a^b \varphi(x)\,dx + \int_a^b \psi(x)\,dx.$$

(b) Let $c \in \mathbf{K}$. It is essentially obvious that $c\varphi$ is a step function. Check that

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

Definition. Let Y be a normed linear space over $\mathbf{K}$. Let $f\colon [a,b] \to Y$, where $a, b \in \mathbf{R}$ with $a < b$. To say that *y is a Riemann integral of f over* $[a,b]$ means that $y \in Y$ and there exists a sequence (φ_n) of step functions from $[a,b]$ to Y such that $U(|f - \varphi_n|) \to 0$ and $\int_a^b \varphi_n(x)\,dx \to y$. (Note that $|f - \varphi_n|$ is the function g_n on $[a,b]$ defined by $g_n(x) = |f(x) - \varphi_n(x)|$, where for each $v \in Y$, $|v|$ denotes the norm of v. Since g_n is real-valued, $U(g_n)$ makes sense.) To say that *f is Riemann integrable over* $[a,b]$ means that there exists $y \in Y$ such that y is a Riemann integral of f over $[a,b]$.

X187. Let Y be a normed linear space over $\mathbf{K}$. Let $f\colon [a,b] \to Y$, where $a, b \in \mathbf{R}$ with $a < b$.

(a) Suppose that y and z are Riemann integrals of f over $[a,b]$. Prove that $y = z$. Thus it makes sense to speak of *the Riemann integral of f over* $[a,b]$ and to write $\int_a^b f(x)\,dx = y$.

(b) Suppose that Y is a Banach space[29] and there exists a sequence (φ_n) of step functions from $[a,b]$ to Y such that $U(|f - \varphi_n|) \to 0$. Prove that f is Riemann integrable over $[a,b]$. (Hint: For each n, let $y_n = \int_a^b \varphi_n(x)\,dx$. Prove that (y_n) is a Cauchy sequence in Y.)

Reminder. Let $f\colon [a,b] \to Y$, where $a, b \in \mathbf{R}$ with $a < b$ and where Y is a metric space. To say that f *is regulated* means that for each $u \in (a,b]$, the left hand limit $f(u-) = \lim_{x \to u-} f(x)$ exists in Y, and for each $v \in [a,b)$, the right hand limit $f(v+) = \lim_{x \to v+} f(x)$ exists in Y. As we have seen, if f is regulated, then there is a sequence (φ_n) of step functions $\varphi_n\colon [a,b] \to Y$ such that $\varphi_n \to f$ uniformly on $[a,b]$. The converse is also true, provided Y is complete.

X188. Let Y be a Banach space and let $a, b \in \mathbf{R}$ with $a < b$. Let $f\colon [a,b] \to Y$ be regulated. Prove that f is Riemann integrable over $[a,b]$.

X189. Let Y be a metric space. Let $f\colon [a,b] \to Y$ be regulated. For each $\varepsilon > 0$, let

$$L(\varepsilon) = \{\, u \in (a,b] : d(f(u-), f(u)) > \varepsilon \,\} \qquad \text{and} \qquad R(\varepsilon) = \{\, v \in [a,b) : d(f(v+), f(v)) > \varepsilon \,\}.$$

Show that for each $\varepsilon > 0$, $R(\varepsilon)$ and $L(\varepsilon)$ are finite sets. Conclude that the set

$$D = \{\, c \in [a,b] : f \text{ is not continuous at } c \,\}$$

is countable.

X190. Let $D = \mathbf{C} \setminus \{0\}$. Define $g\colon D \to \mathbf{C}$ by $g(z) = 1/z$. Let $z_0 \in D$. Let $r = |z_0|$. Then $r > 0$ and $B(z_0, r) \subseteq D$. Find a power series representation $\sum_{k=0}^{\infty} c_k (z - z_0)^k$ for $g(z)$ that is valid for all $z \in B(z_0, r)$. (You should not need calculus for this.) Deduce that g is analytic.

Remark. If f and g are analytic, we would like to know that their composition $g \circ f$ is analytic. The next few exercises deal with this.

X191. Consider a power series *due 7Th*

$$g(z) = \sum_{\ell=0}^{\infty} b_\ell z^\ell$$

whose radius of convergence R_2 is strictly positive. Consider another power series

$$f(z) = \sum_{k=1}^{\infty} a_k z^k$$

[29] Recall that a Banach space is a normed linear space which is complete with respect to the metric induced by its norm. For instance, $\mathbf{R}^d$ with its usual norm is a Banach space. Here is another example. Let X be a topological space and let $C_b(X)$ be the vector space of all bounded continuous functions from X into $\mathbf{K}$. For each $f \in C_b(X)$, let $\|f\| = \sup|f|$. This defines a norm on $C_b(X)$. Under this norm, $C_b(X)$ is a Banach space, because a uniformly Cauchy sequence of $\mathbf{K}$-valued functions is uniformly convergent and because a uniform limit of continuous functions is continuous.

whose constant term is zero and whose radius of convergence R_1 is strictly positive. For each $k \in \mathbf{N}$, let $\alpha_k = |a_k|$. The series for f converges absolutely for $|z| < R_1$. Hence we may define $\varphi(z)$, for $|z| < R_1$, by

$$\varphi(z) = \sum_{k=1}^{\infty} \alpha_k z^k.$$

By the Cauchy product formula, there are coefficients $a_k^{(\ell)}$ and $\alpha_k^{(\ell)}$, where k and ℓ vary over $\boldsymbol{\omega} = \{0, 1, 2, \ldots\}$, such that for $|z| < R_1$, we have

$$f(z)^\ell = \sum_{k=\ell}^{\infty} a_k^{(\ell)} z^k$$

and

$$\varphi(z)^\ell = \sum_{k=\ell}^{\infty} \alpha_k^{(\ell)} z^k.$$

(a) Prove that $|a_k^{(\ell)}| \leq \alpha_k^{(\ell)}$.

Let $R_0 = \sup\{ r \in [0, R_1) : \varphi(r) < R_2 \}$.

(b) Prove that $R_0 > 0$ and that for $|z| < R_0$, we have

$$|f(z)| \leq \varphi(|z|) < R_2.$$

Hence $g(f(z))$ is defined for $|z| < R_0$.

(c) Prove that for $|z| < R_0$, we have

$$g(f(z)) = \sum_{k=0}^{\infty} c_k z^k,$$

where

$$c_k = \sum_{\ell=0}^{\infty} b_\ell a_k^{(\ell)}.$$

(Informally, the proof of this is just a computation involving interchanging the order of summation in a certain double series. Of course you are expected to justify this interchange in order of summation.)

X192. Let D_1 and D_2 be open subsets of $\mathbf{C}$. Let $f \colon D_1 \to D_2$ and $g \colon D_2 \to \mathbf{C}$ be analytic. Let $h = g \circ f$. Prove that h is analytic.

X193. Let D be an open subset of $\mathbf{C}$ and let $f \colon D \to \mathbf{C} \setminus \{0\}$ be analytic. Prove that $1/f$ is analytic. (Hint: This is trivial if you just combine the results of a couple of the exercises above.)

Reminder. For each $n \in \mathbf{N}$ and each $u \in \mathbf{C} \setminus \{1\}$, we know that

$$1 + u + u^2 + \cdots + u^{n-1} = \frac{1-u^n}{1-u}. \tag{19}$$

X194. For each $n \in \mathbf{N}$ and each $t \in \mathbf{R} \setminus \{-1\}$, if we let $u = -t$ in (19), we get

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

Use this (but no theorems about series) to show that for each $x \in (-1, 1]$,

$$\log(1+x) = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{x^k}{k},$$

and that for each $a \in (-1, 1)$, the series converges uniformly for $x \in [a, 1]$. In particular,

$$\log 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + - \cdots .$$

X195. For each $n \in \mathbf{N}$ and each $t \in \mathbf{R}$, if we let $u = -t^2$ in (19), we get

$$\frac{1}{1+t^2} = 1 - t^2 + t^4 - \cdots + (-1)^{n-1}t^{2n-2} + \frac{(-1)^n t^{2n}}{1+t^2}.$$

Use this (but no theorems about series) to show that for each $x \in [-1, 1]$,

$$\arctan x = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{x^{2k-1}}{2k-1},$$

and that the series converges uniformly for $x \in [-1, 1]$. In particular, since $\arctan 1 = \pi/4$,

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + - \cdots .$$

X196. *due 8Th*

(a) Let $a \in \mathbf{R}$. Prove Newton's binomial theorem: For each $x \in (-1, 1)$,

$$(1+x)^a = \sum_{k=0}^{\infty} \binom{a}{k} x^k, \tag{20}$$

where

$$\binom{a}{k} = \frac{a(a-1)(a-2)\cdots(a-k+1)}{k!}.$$

(When $k = 0$, $\binom{a}{k} = 1$ because a product of no factors is 1.) Hint: Show that the series on the right in (20) converges. Let $f(x)$ denote the sum of the series. Show that f satisfies the differential equation

$$(1+x)f'(x) = af(x).$$

Use the method of integrating factors to solve this differential equation.[30]

(b) Now suppose $a \in \mathbf{C}$. Then (20) still holds for each $x \in (-1, 1)$, provided we define $(1+x)^a$ as $e^{a \log(1+x)}$. (It should be clear that the argument you gave in part (a) still works. You need not repeat it.) Use the coincidence principal to show that for each $z \in \mathbf{C}$ with $|z| < 1$,

$$(1+z)^a = \sum_{k=0}^{\infty} \binom{a}{k} z^k,$$

where we take $(1+z)^a$ to be $e^{a \operatorname{Log}(1+z)}$. (This is the principal value of $(1+z)^a$. Recall that $\operatorname{Log}(1+z)$ is the principal logarithm of $1+z$, namely the one whose imaginary part lies in the interval $(-\pi, \pi]$.)

[30] Alternatively, one could prove (20) by applying Taylor's theorem, but this would involve more work, because it would be tedious to estimate the remainder.

X197. For each $n \in \mathbf{N}$, let $k_n \in \mathbf{N}$, let $z(n,j) \in \mathbf{C}$ for $j = 1, \ldots, k_n$, and let *due 8Th*

$$p_n = \prod_{j=1}^{k_n} (1 + z(n,j)) \qquad \text{and} \qquad s_n = \sum_{j=1}^{k_n} z(n,j).$$

Suppose that $\sum_{j=1}^{k_n} |z(n,j)|^2 \to 0$ as $n \to \infty$.

(a) Suppose $z \in \mathbf{C}$ with $z \neq -1$. Then $\mathrm{Log}(1+z) = z + \theta z^2$, where

$$\theta = \begin{cases} \frac{\mathrm{Log}(1+z) - z}{z^2} & \text{if } z \neq 0, \\ 0 & \text{if } z = 0. \end{cases}$$

Hence

$$1 + z = e^{z + \theta z^2}.$$

Show that if $|z| \leq 1/2$, then $|\theta| \leq 1$. Hint: If $|z| < 1$, then

$$\mathrm{Log}(1+z) = z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} + - \cdots .$$

(b) Show that $p_n \sim e^{s_n}$ as $n \to \infty$.

(c) Suppose in addition that $s_n \to s \in \mathbf{C}$ as $n \to \infty$. Show that $p_n \to e^s$ as $n \to \infty$.

Remark. One important application of problem X197 occurs in the proof of the central limit theorem in probability theory.

X198. Suppose that $a_k \in \mathbf{C} \setminus \{-1\}$ for each $k \in \mathbf{N}$. Let $p_n = \prod_{k=1}^{n} (1 + a_k)$ for each $n \in \mathbf{N}$. Notice that $p_n \neq 0$ for each $n \in \mathbf{N}$. Suppose that $\sum_{k=1}^{\infty} a_k$ converges and that $\sum_{k=1}^{\infty} |a_k|^2 < \infty$. Prove that (p_n) converges to a limit p in $\mathbf{C} \setminus \{0\}$.

Let E be a vector space over $\mathbf{R}$. If $v, w \in E$, then the *line segment from* v *to* w is the set $L(v,w) = \{(1-t)v + tw : t \in [0,1]\}$. Let S be a subset of E. To say that S *is convex* means that for all $v, w \in S$, $L(v,w) \subseteq S$. To say that S *is star-shaped with respect to* v means that for each $w \in S$, $L(v,w) \subseteq S$. If S is star-shaped with respect to v, then clearly $v \in S$. To say that S *is star-shaped* means that there exists v such that S is star-shaped with respect to v. Clearly S is convex iff S is star-shaped with respect to each $v \in S$. Now let us mention some examples. Considering $\mathbf{C}$ as a vector space over $\mathbf{R}$, each open or closed disc is a convex subset of $\mathbf{C}$, and so is each open or closed half-plane, while $\mathbf{C} \setminus (-\infty, 0]$ is star-shaped with respect to 1, but not convex.

Let X and Y be topological spaces. Let f_0 and f_1 be continuous maps from X into Y. To say that f_0 *is homotopic to* f_1 *in* Y means that there exists a continuous map H from $X \times [0,1]$ into Y such that for each $x \in X$, $H(x,0) = f_0(x)$ and $H(x,1) = f_1(x)$. Such a map H is called *a homotopy from* f_0 *to* f_1 *in* Y. Intuitively, such a homotopy describes a continuous deformation of the map f_0 into the map f_1, where each of the intermediate maps $H(\cdot, t)$, $t \in [0,1]$, is a continuous map from X into Y. It is not hard to show that the relation of homotopy between continuous maps from X into Y is an equivalence relation on $C(X,Y)$, the set of all continuous maps from X into Y.

Let X be a topological space. To say that X *is contractible* means that the identity map from X into itself is homotopic in X to some constant map.

X199. Let X be a star-shaped subset of $\mathbf{R}^d$. Show that X is contractible.

Let X, Y, and Z be topological spaces. Let $f\colon X \to Z$ be a continuous map. To say that f *can be factored through* Y means that there exist continuous maps $g\colon X \to Y$ and $h\colon Y \to Z$ such that $f = h \circ g$.

X200. Let $\mathbf{C}^\times = \mathbf{C} \setminus \{0\}$. Let X be a topological space and let f be a continuous map from X into $\mathbf{C}^\times$. Consider the following seven statements:

(a) There exists a branch of $\log f$.

(b) f can be factored through $\mathbf{C}$.

(c) f can be factored through a convex subset of $\mathbf{C}$.

(d) f can be factored through a star-shaped subset of $\mathbf{C}$.
(e) f can be factored through some contractible topological space.
(f) f is homotopic in $\mathbf{C}^\times$ to some constant map.
(g) f is homotopic in $\mathbf{C}^\times$ to the constant map $1\colon X \to \mathbf{C}^\times$.

Prove that (a)$\Rightarrow$(b)$\Rightarrow$(c)$\Rightarrow$(d)$\Rightarrow$(e)$\Rightarrow$(f)$\Rightarrow$(g). (It is an important non-trivial fact that (g)$\Rightarrow$(a) too, but you are not asked to prove this here.)

Remark. Let $\mathbf{C}^\times$, X, and f be as in problem X200. Obviously f can always be factored through X, because $f = f \circ \mathrm{id}_X$, where id_X denotes the identity function on X. Hence if X is contractible, then there exists a branch of $\log f$. In particular, if X is a star-shaped subset of $\mathbf{C}$, then there exists a branch of $\log f$. Still more particularly, if X is a star-shaped subset of $\mathbf{C}^\times$, then there exists a branch of log in X. To see this, take f to be id_X. (If $X \subseteq \mathbf{C}^\times$, then to say that g *is a branch of* log *in* X means that for each $z \in X$, $e^{g(z)} = z$. In other words, if $X \subseteq \mathbf{C}^\times$, then a branch of log in X is a branch of $\log f$, where $f = \mathrm{id}_X$.)

Definition. Let $f\colon X \to Y$, where X is a topological space and Y is a metric space. Let $p \in X$ and let $\mathscr{U}_p$ be the collection of nhds of p in X. The *oscillation of* f *at* p is

$$\mathrm{osc}(f,p) = \inf_{U \in \mathscr{U}_p} \mathrm{diam}(f[U]).$$

X201. Let $f\colon X \to Y$, where X is a topological space and Y is a metric space. Define $\varphi\colon X \to [0,\infty]$ by

$$\varphi(p) = \mathrm{osc}(f,p)$$

for all $p \in X$.
(a) Let $p \in X$. Prove that f is continuous at p iff $\mathrm{osc}(f,p) = 0$.
(b) Prove that φ is upper semicontinous.
(c) Let $C = \{\, p \in X : f \text{ is continuous at } p \,\}$. Prove that C is a $\mathscr{G}_\delta$ subset of X.

X202. Let X be a topological space and let $f\colon X \to \mathbf{R}$ be bounded. Let $p \in X$. Prove that

$$\mathrm{osc}(f,p) = f^*(p) - f_*(p).$$

Reminder. Let X be a topological space and let $A \subseteq X$. To say that A *is nowhere dense* means that the closure of A has empty interior. To say that A *is meager* means that A is a countable union of nowhere dense sets. For instance, any countable subset of $\mathbf{R}$ is meager in $\mathbf{R}$. So is any closed subset of $\mathbf{R}$ which has empty interior. For instance, the Cantor set is meager in $\mathbf{R}$. You should think of meager sets as sets which are small in the sense of topology. Meager sets are also known as sets *of the first category.* Sets which are not meager are also known as sets *of the second category.* To say that X is a *Baire space* means that no non-empty open subset of X is meager. According to problem X74, X is a Baire space iff for each meager set $A \subseteq X$, the set $X \setminus A$ is dense in X iff for each sequence (G_n) of dense open subsets of X, we have $\bigcap_{n=1}^\infty G_n$ is dense in X. The Baire category theorem states that each complete metric space is a Baire space.

Reminder. Clearly $\mathbf{Q}$ is an $\mathscr{F}_\sigma$ subset of $\mathbf{R}$, so $\mathbf{R} \setminus \mathbf{Q}$ is a $\mathscr{G}_\delta$ subset of $\mathbf{R}$. By problem X77, $\mathbf{Q}$ is not a $\mathscr{G}_\delta$ subset of $\mathbf{R}$. Hence $\mathbf{R} \setminus \mathbf{Q}$ is not an $\mathscr{F}_\sigma$ subset of $\mathbf{R}$.

X203. Let (f_n) be a sequence of continuous functions from a topological space X to a metric space (Y,d), such that (f_n) converges pointwise to a function $f\colon X \to Y$. Prove that for each open set $G \subseteq Y$, $f^{-1}[G]$ is an $\mathscr{F}_\sigma$ subset of X.

X204. Prove that $1_{\mathbf{Q}}$, the indicator function of the set $\mathbf{Q}$ of rational numbers, is not the limit of a sequence of continuous functions from $\mathbf{R}$ to $\mathbf{R}$.

X205. Let X be a topological space. Prove that X is a Baire space iff for each sequence (F_n) of closed subsets of X, if $X = \bigcup_n F_n$, then $\bigcup_n \mathrm{Int}(F_n)$ is dense in X.

X206. Let f be a function from a Baire space X to a separable metric space (Y,d), such that for each open set $G \subseteq Y$, $f^{-1}[G]$ is an $\mathscr{F}_\sigma$ subset of X. Let C be the set of all $x \in X$ such that f is continuous at x. By problem X201, we know that C is a $\mathscr{G}_\delta$ subset of X. Prove that C is dense in X. (Hint: For each $n \in \mathbf{N}$, Y can be covered by a sequence of open sets of diameter at most $1/n$.)

X207. Define $f\colon \mathbf{R} \to \mathbf{R}$ by
$$f(x) = \begin{cases} e^{-1/x} & \text{if } x > 0, \\ 0 & \text{if } x \le 0. \end{cases}$$

(a) Prove that for each $n \in \mathbf{N}$, $f^{(n)}(x) \to 0$ as $x \to 0+$.

(b) Prove that f is infinitely differentiable on $\mathbf{R}$.

X208. Let $a, b \in \mathbf{R}$ with $a < b$. Find an infinitely differentiable function $g\colon \mathbf{R} \to \mathbf{R}$ such that $g(x) = 1$ for each $x \in (-\infty, a]$, $0 < g(x) < 1$ for each $x \in (a, b)$, and $g(x) = 0$ for each $x \in [b, \infty)$.

X209. Let $D = \mathbf{C} \setminus (-\infty, 0]$. Define $f\colon D \to \mathbf{C}$ by $f(z) = \operatorname{Log} z$. We have seen that f is continuous in D. Prove that f is analytic in D.

Reminder. In problem X187, we considered Riemann integration of functions taking values in a Banach space.

X210. (*The triangle inequality for a Riemann integrable function taking values in a Banach space.*) Let $f\colon [a, b] \to Y$, where $a, b \in \mathbf{R}$ with $a < b$ and where Y is a Banach space. Suppose that f is Riemann integrable over $[a, b]$. Let $y = \int_a^b f(t)\, dt$. Prove that $|y| \le \int_a^b |f(t)|\, dt$. (Hint: This is easy if f is a step function. For the general case, approximate f by step functions.)

X211. Let $f\colon [a, b] \to Y$, where $a, b \in \mathbf{R}$ with $a < b$ and where Y is a Banach space. Suppose f is Riemann integrable over $[a, b]$. From the discussion in problem X187, it is obvious that then f is integrable over $[u, v]$ for all $u, v \in [a, b]$ with $u < v$. Define $F\colon [a, b] \to Y$ by $F(x) = \int_a^x f(t)\, dt$. Let $c \in [a, b]$ such that f is continuous at c. Prove that F is differentiable at c and that $F'(c) = f(c)$. (By the way, in the course of your proof, you should notice that analogous statements hold for the one-sided derivatives of F at c and one-sided continuity of f at c.)

X212. (*The straddle lemma.*) Let Y be a normed linear space. Let $f\colon E \to Y$, where $E \subseteq \mathbf{R}$. Let E' be the set of limit points of E and let $a \in E' \cap E$. Suppose that f is differentiable[31] at a. Prove that for each $\varepsilon > 0$, there exists $\delta > 0$ such that for all $u, v \in E$ with $u \le a \le v$ and $0 < v - u < \delta$, we have
$$\left| \frac{f(v) - f(u)}{v - u} - f'(a) \right| < \varepsilon.$$

Remark. The mean value theorem does not apply to vector-valued functions. Even for a function taking values in $\mathbf{R}^2$, the points that the mean value theorem furnishes for the two components of the function are usually going to be different. The next result is an adequate substitute for the mean value theorem in many situations involving vector-valued functions.

X213. (*The mean value inequality.*) Let $f\colon [a, b] \to Y$, where $a, b \in \mathbf{R}$ with $a < b$ and where Y is a normed linear space.

(a) Suppose that f is differentiable on $[a, b]$. Prove that there exists $c \in [a, b]$ such that
$$|f'(c)| \ge \left| \frac{f(b) - f(a)}{b - a} \right|.$$
(Hint: First check that if $u, v, w \in [a, b]$ with $u < v < w$, then at least one of the two quantities
$$\left| \frac{f(v) - f(u)}{v - u} \right| \quad \text{and} \quad \left| \frac{f(w) - f(v)}{w - v} \right|$$
is greater than or equal to
$$\left| \frac{f(w) - f(u)}{w - u} \right|.$$
In particular, this holds if v is the midpoint of $[u, w]$. Then start with the whole interval $[a, b]$, apply a bisection argument, and finish by applying the straddle lemma.)

(b) Suppose that f is continuous on $[a, b]$ and differentiable on (a, b). Let
$$M = \sup\{\, |f'(t)| : t \in (a, b)\,\}.$$
Prove that $|f(b) - f(a)| \le M \cdot (b - a)$.

[31] Differentiability for a function taking values in a normed linear space is defined just like for a function taking values in $\mathbf{R}$.

X214. Let Y be a normed linear space and let $a, b \in \mathbf{R}$ with $a < b$.

(a) Let $H\colon [a,b] \to Y$ be continuous on $[a,b]$ and differentiable on (a,b). Suppose also that $H'(x) = 0$ for each $x \in (a,b)$. Prove that H is constant on $[a,b]$.

(b) Let $F, G\colon [a,b] \to Y$ be continuous on $[a,b]$ and differentiable on (a,b). Suppose also that $F'(x) = G'(x)$ for each $x \in (a,b)$. Prove that there is a constant $C \in Y$ such that $G = F + C$.

X215. (*The fundamental theorem of calculus for a continuous function taking values in a Banach space.*) Let $f\colon [a,b] \to Y$ be continuous, where $a, b \in \mathbf{R}$ with $a < b$ and where Y is a Banach space. Suppose that $G\colon [a,b] \to Y$ is continuous on $[a,b]$ and differentiable on (a,b). Suppose also that $G'(x) = f(x)$ for each $x \in (a,b)$. Prove that

$$\int_a^b f(t)\,dt = G(b) - G(a).$$

X216. Give an example of a differentiable function $f\colon \mathbf{R} \to \mathbf{R}$ and sequences (u_n) and (v_n) in $\mathbf{R}$ such that $u_n \neq v_n$ for all n, $u_n \to 0$ and $v_n \to 0$, but

$$\limsup_{n\to\infty} \frac{f(v_n) - f(u_n)}{v_n - u_n} = \infty \qquad \text{and} \qquad \liminf_{n\to\infty} \frac{f(v_n) - f(u_n)}{v_n - u_n} = -\infty.$$

Explain why your example does not contradict the straddle lemma.

X217. Define functions u and v by *due 9Th*

$$u(t) = \sum_{k=1}^{\infty} \frac{\cos(kt)}{k} \tag{21}$$

and

$$v(t) = \sum_{k=1}^{\infty} \frac{\sin(kt)}{k}. \tag{22}$$

Let each of these functions be defined for all $t \in \mathbf{R}$ for which the sum of the corresponding series is defined. We know from problem X175 that these series converge locally uniformly in $\mathbf{R} \setminus 2\pi\mathbf{Z}$. Clearly the series (21) diverges to $+\infty$ on $2\pi\mathbf{Z}$ while the series (22) converges on $2\pi\mathbf{Z}$, because each term is zero there.

(a) Find simple closed form expressions for $u(t)$ and $v(t)$. Hint: As in problem X175, define a function g by

$$g(t) = \sum_{k=1}^{\infty} \frac{e^{ikt}}{k}$$

for each $t \in \mathbf{R}$ for which the series converges in $\mathbf{C}$. Notice that $u(t)$ is the real part of $g(t)$ and $v(t)$ is the imaginary part of $g(t)$. Now for $|z| \le 1$ with $z \neq -1$, the principal logarithm of $1+z$ is given by

$$\operatorname{Log}(1+z) = z - \frac{z^2}{2} + \frac{z^3}{3} - + \cdots.$$

Hence for $|z| \le 1$ with $z \neq 1$,

$$-\operatorname{Log}(1-z) = z + \frac{z^2}{2} + \frac{z^3}{3} + \cdots.$$

Substituting $z = e^{it}$ in the preceding equation, we find that

$$g(t) = -\operatorname{Log}(1 - e^{it})$$

for each $t \in \mathbf{R} \setminus 2\pi\mathbf{Z}$. To get simple expressions from this for u and v, check that

$$1 - e^{it} = 2e^{i(t-\pi)/2} \sin\frac{t}{2}.$$

The expressions you find for u and v should involve only real numbers.

(b) Sketch the graph of v. If you found as simple an expression for v as you should have in (a), then this should be a piece of cake.

(c) Even though the series (22) does not converge uniformly,[32] we can still hope that each of the integrals,

$$\int_0^{2\pi} v(t)\cos(kt)\,dt, \qquad k = 0, 1, 2, \ldots$$

and

$$\int_0^{2\pi} v(t)\sin(kt)\,dt, \qquad k = 1, 2, 3, \ldots$$

can be calculated by expressing $v(t)$ as in (22) and integrating term-by-term. We shall soon see that this is indeed the case. Verify that the values so obtained for these integrals are consistent with your answer to (a).

X218. Let $a \in \mathbf{C}$. We saw in problem X196 that for $|z| < 1$, the principal value of $(1+z)^a$ is given by *due 9Th*

$$(1+z)^a = \sum_{k=0}^{\infty} \binom{a}{k} z^k. \tag{23}$$

If $a \in \boldsymbol{\omega} = \{0, 1, 2, \ldots\}$, then of course $\binom{a}{k} = 0$ for each $k \in \{a+1, a+2, \ldots\}$, so (23) reduces to the ordinary binomial expansion $(1+z)^a = \sum_{k=0}^{a} \binom{a}{k} z^k$, which is valid for all $z \in \mathbf{C}$. For the remainder of this exercise, let us assume that $a \notin \boldsymbol{\omega}$, so that the right side of (23) really is an infinite series. Then by the ratio test, it is easy to see that the radius of convergence of the series is exactly 1.

(a) Suppose that $\mathrm{Re}(a) \le -1$. Show that $\left|\binom{a}{k}\right| \ge 1$ for each $k \in \boldsymbol{\omega}$. Deduce that the series in (23) diverges for all z with $|z| = 1$.

(b) Consider general $a \in \mathbf{C}$ again. As we know, $\binom{a}{0} = 1$. Verify that for each $k \in \mathbf{N}$, $\binom{a}{k} = (-1)^k b_k$, where

$$b_k = \prod_{\ell=1}^{k} \Bigl(1 - \frac{a+1}{\ell}\Bigr). \tag{24}$$

Then use problem X197(a), together with problem X158, to show that for each $k \in \mathbf{N}$,

$$b_k = \frac{c_k}{k^{a+1}}, \tag{25}$$

where (c_k) is a sequence in $\mathbf{C} \setminus \{0\}$ which converges to a limit c belonging to $\mathbf{C} \setminus \{0\}$, and where we mean the principal value of k^{a+1}.

(c) Suppose that $\mathrm{Re}(a) > 0$. Show that the series in (23) converges absolutely for each z with $|z| = 1$ and that for each such z, the equation (23) holds.[33]

(d) Suppose that $-1 < \mathrm{Re}(a) \le 0$. Use (25) to show that $b_k \to 0$. Use (24) to show that

$$b_k - b_{k+1} = b_k d_k, \tag{26}$$

where d_k is a suitable simple factor that you should determine explicitly. Then use (25) and (26) to show that the sequence (b_k) is of bounded variation. Finally, deduce that if $|z| = 1$ but $z \ne -1$, then the series in (23) converges, although not absolutely, and that the equation (23) holds for each such z.

(e) Suppose again that $-1 < \mathrm{Re}(a) \le 0$. Suppose also that $z = -1$. Prove that the series in (23) diverges. (Hint: This is easy if you use a theorem that we proved in class recently. If the series did converge, what would that tell us about $\lim_{x \downarrow -1} (1+x)^a$? Remember that $a \ne 0$ because we are assuming that $a \notin \boldsymbol{\omega}$.)

---

[32] It cannot converge uniformly, because v is discontinous at each integer multiple of 2π, as should be clear from your answer to (a).

[33] For $w \in \mathbf{C} \setminus \{0\}$, if ζ is the principal value of w^a, then $|\zeta| = e^{\mathrm{Re}(a)\log|w| - \mathrm{Im}(a)\,\mathrm{Arg}(w)}$. Hence if $\mathrm{Re}(a) > 0$, then as $w \to 0$, the principal value of w^a tends to 0, because $\mathrm{Re}(a)\log|w|$ tends to negative infinity while $\mathrm{Arg}(w)$ stays between $-\pi$ and π. Hence when $\mathrm{Re}(a) > 0$, we consider the principal value of 0^a to be 0. In particular, when $\mathrm{Re}(a) > 0$ and $z = -1$, the left side in (23) has the value 0.

Remark. If $(a_k)_{k\in\mathbf{N}}$ is a sequence in $[0,\infty]$, then it is easy to check that

$$\sum_{k=1}^{\infty} a_k = \sup\left\{ \sum_{k\in B} a_k : B \text{ is a finite subset of } \mathbf{N} \right\}.$$

This motivates the following definition.

Definition. Let $(a_k)_{k\in A}$ be a family of elements of $[0,\infty]$, where A is an arbitrary index set. (A could even be uncountably infinite.) Then by definition,

$$\sum_{k\in A} a_k = \sup\left\{ \sum_{k\in B} a_k : B \text{ is a finite subset of } A \right\}$$

X219. Let V be an inner product space and let $(u_k)_{k\in A}$ be an orthonormal family in V. Let $f \in V$. For each $k \in A$, let $c_k = \langle u_k \,|\, f\rangle$.

(a) Let B be a finite subset of A and let $h = \sum_{k\in B} c_k u_k$. Show that

$$\|f-h\|^2 + \sum_{k\in B} |c_k|^2 = \|f\|^2.$$

(b) Use part (a) to prove *Bessel's inequality* in its abstract form:

$$\sum_{k\in A} |c_k|^2 \le \|f\|^2. \tag{27}$$

Reminder. $\mathscr{R}(\mathbf{T})$ denotes the (almost) inner product vector space of complex-valued 1-periodic functions on $\mathbf{R}$ which are Riemann integrable over $[0,1]$. For each $k \in \mathbf{Z}$, we define $e_k \colon \mathbf{R} \to \mathbf{C}$ by $e_k(t) = e^{2\pi i k t}$. We have seen that $(e_k)_k \in \mathbf{Z}$ is an orthonormal family in $\mathscr{R}(\mathbf{T})$.

Remark. Let $g \in \mathscr{R}(\mathbf{T})$. Since $(e_k)_{k\in\mathbf{Z}}$ is an orthonormal family in $\mathscr{R}(\mathbf{T})$, Bessel's inequality (27) tells us that

$$\sum_{k\in\mathbf{Z}} |\hat{g}(k)|^2 \le \|g\|_2^2. \tag{28}$$

Actually, according to Parseval's equation, we have equality in (28), though we don't need that fact for the application of (28) in the next exercise.

X220. Let $f \colon \mathbf{R} \to \mathbf{C}$ be 1-periodic and differentiable. Suppose f' is Riemann integrable over $[0,1]$. *due 10Th*

(a) Prove that $\sum_{k\in\mathbf{Z}} |\hat{f}(k)| < \infty$. (Hint: Let $g = f'$. Find an expression for $\hat{g}(k)$ in terms of $\hat{f}(k)$. Then use (28).)

(b) For each $n \in \boldsymbol{\omega} = \{0,1,2,\dots\}$, let $S_n = \sum_{k=-n}^{n} \hat{f}(k) e_k$. Prove that the sequence of functions (S_n) converges uniformly to f. (Hint: Prove first that (S_n) is uniformly Cauchy and hence converges uniformly to some continuous function $S \colon \mathbf{R} \to \mathbf{C}$. Then show that $\|f - S\|_2 = 0$ and use the continuity of f and of S to conclude that for each $t \in \mathbf{R}$, $S(t) = f(t)$.)

(c) Show that in fact, $\sum_{k\in\mathbf{Z}} \hat{f}(k) e_k = f$, where the series converges uniformly to f in the sense of unordered sums. In other words, show that for each $\varepsilon > 0$, there exists a finite subset $B \subseteq \mathbf{Z}$ such that for each finite subset $C \subseteq \mathbf{Z}$ with $B \subseteq C$, and for each $t \in \mathbf{R}$, we have

$$\Big| f(t) - \sum\nolimits_{k\in C} \hat{f}(k) e_k(t) \Big| \le \varepsilon.$$

X221. Let $P_1(t) = t - 1/2$ for $0 < t < 1$. Let $P_1(0) = 0$ and extend P_1 to $\mathbf{R}$ by periodicity.[34] Show that the Fourier series for P_1 may be written as

$$-\sum_{k=1}^{\infty} \frac{2\sin 2\pi k t}{2\pi k}. \tag{29}$$

Then deduce from Parseval's equation that $\sum_{k=1}^{\infty} k^{-2} = \pi^2/6$.

Remark. Let P_1 be as in problem X221. Let

$$P_2(t) = \sum_{k=1}^{\infty} \frac{2\cos 2\pi k t}{(2\pi k)^2},$$

let

$$P_3(t) = \sum_{k=1}^{\infty} \frac{2\sin 2\pi k t}{(2\pi k)^3},$$

and define P_4, P_5, P_6, P_7, and so on, by

$$P_{2r}(t) = (-1)^{r-1} \sum_{k=1}^{\infty} \frac{2\cos 2\pi k t}{(2\pi k)^{2r}}$$

and

$$P_{2r+1}(t) = (-1)^{r-1} \sum_{k=1}^{\infty} \frac{2\sin 2\pi k t}{(2\pi k)^{2r+1}}.$$

---

[34] By the way, for each $t \in \mathbf{R} \setminus \mathbf{Z}$, $P_1(t) = t - \lfloor t \rfloor - 1/2$, where $\lfloor t \rfloor$ is the greatest integer that is less than or equal to t. Of course for each $t \in \mathbf{Z}$, $P_1(t) = 0$.

From problem X175 and problem X217, after some elementary modifications, we know that the series (29) converges to $P_1(t)$ for each $t \in \mathbf{R}$ and that $P_2'(t) = P_1(t)$ for each $t \in \mathbf{R} \setminus \mathbf{Z}$. The series for P_2, P_3, and so on, all converge uniformly on $\mathbf{R}$ and so it is elementary that for each $r \in \mathbf{N}$ with $r \geq 2$, and for each $t \in \mathbf{R}$, $P_{r+1}'(t) = P_r(t)$. It follows that for each $r \in \mathbf{N}$, the restriction of the 1-periodic function P_r to the open interval $(0,1)$ agrees with a certain polynomial p_r of degree r, where the constant term in p_r is determined by the fact that $\int_0^1 p_r(t)\,dt = 0$. These polynomials are known as the *Bernoulli polynomials*, after James Bernoulli, who introduced essentially these polynomials[35] in 1713, in connection with the problem of determining the sums $\sum_{k=1}^n k^r$. Note that it is straightforward to determine the Bernoulli polynomials explicitly, by successive integration, beginning from p_1. Then, just as we applied Parseval's equation to P_1 to evaluate $\sum_{k=1}^\infty k^{-2}$, we may apply Parseval's equation to P_r to evaluate $\sum_{k=1}^\infty k^{-2r}$.

Definition. Let (x_j) be a sequence of real numbers. To say that (x_j) *is equidistributed modulo* 1 means that for each $f \in C(\mathbf{T})$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(x_j) = \int_0^1 f(x)\,dx. \tag{30}$$

X222. Let $\alpha \in \mathbf{R}$. Let $x_0 \in \mathbf{R}$ and let $x_j = x_0 + j\alpha$ for $j \in \mathbf{N}$. Prove that the sequence (x_j) is equidistributed modulo 1 iff α is irrational. (Hint: The forward implication is trivial. For the reverse implication, first prove it for f of the form $f(x) = e^{2\pi i k x}$ where $k \in \mathbf{Z}$, by explicit calculation of $\sum_{j=0}^{n-1} f(x_j)$. Then use the fact that the set of trigonometric polynomials is dense in $C(\mathbf{T})$.) *due 10Th*

X223. Let (x_j) be a sequence of real numbers which is equidistributed modulo 1. By definition, (30) holds for each $f \in C(\mathbf{T})$. Prove that in fact, (30) holds for each $f \in \mathscr{R}(\mathbf{T})$. (Hint: It suffices to consider the case where f is real-valued. Verify that then for each $\varepsilon > 0$, there exist functions $g, h \in C(\mathbf{T})$ with $g \leq f \leq h$ and $\int_0^1 h(x) - g(x)\,dx < \varepsilon$.)

Terminology. Let x be a real number. Then *the fractional part of* x is the number $x - \lfloor x \rfloor$, where $\lfloor x \rfloor$ is the greatest integer less than or equal to x.

X224. Let (x_j) be a sequence of real numbers. Prove that (x_j) is is equidistributed modulo 1 iff for each interval $I \subseteq [0,1)$, we have $n^{-1}\#_n(I) \to |I|$, where $|I|$ is the length of I and $\#_n(I)$ is the number of $j \in \{0, \ldots, n-1\}$ such that the fractional part of x_j is in I. (Hint: The forward implication is a particular case of the result of problem X223. To prove the reverse implication, remember that a continuous complex-valued function on $[0,1]$ is a uniform limit of step functions.)

Reminder. Let $f, g \in \mathscr{R}(\mathbf{T})$. Then *the convolution of* f *and* g is the function $f * g$ defined on $\mathbf{R}$ by

$$(f * g)(t) = \int_0^1 f(t-s)g(s)\,ds.$$

X225. Let $f, g \in \mathscr{R}(\mathbf{T})$. Prove that $\sup|f * g| \leq \|f\|_2 \|g\|_2$.

X226. Let $f, g \in \mathscr{R}(\mathbf{T})$ and let (f_n) and (g_n) be sequences in $\mathscr{R}(\mathbf{T})$ such that $\|f - f_n\|_2 \to 0$ and $\|g - g_n\|_2 \to 0$. Prove that $f_n * g_n \to f * g$ uniformly.

Reminder. Let $f \in \mathscr{R}(\mathbf{T})$ and let $k \in \mathbf{Z}$. By definition, $\hat{f}(k) = \langle e_k \,|\, f \rangle = \int_0^1 e^{-2\pi i k t} f(t)\,dt$. As we saw in class, it is easy to check that $f * e_k = \hat{f}(k) e_k$.

X227. Let $j, k \in \mathbf{Z}$. Find $\hat{e}_j(k)$ and $e_j * e_k$.

X228. Let $j, k, \ell \in \mathbf{Z}$. Check that $e_j * (e_k * e_\ell) = (e_j * e_k) * e_\ell$.

X229. Let $f, g, h \in \mathscr{R}(\mathbf{T})$. Prove that $f * (g * h) = (f * g) * h$. (Hint: For now, we can prove this by using the fact that the set of trigonometric polynomials is dense in $\mathscr{R}(\mathbf{T})$ with respect to the 2-norm. If we had the basic theory of double integrals at our disposal, then we could prove it by a routine calculation involving an interchange in order of integration.)

X230. Let $f, g \in \mathscr{R}(\mathbf{T})$. Prove that $f * g$ is continuous.

---

[35] Some authors adopt related but slightly different definitions of the Bernoulli polynomials.

These problems are numbered consecutively with the ones from last quarter. (That is why the first of them is not numbered X1.)

Remark. Our goal at present is to study notions such as length, area, volume, mass, charge, and so on, with a view to developing a theory of integration that will be superior to the Riemann theory of integration. Such notions are all special cases of the general notion of measure. In due course, we shall define precisely what we mean by measure. For now, let us just say in order to study the measure of fairly general sets, we must first study the measure of very special sets. The collection of these special sets usually forms what is known as a pre-ring.

Reminder. To say that $\mathscr{H}$ *is a pre-ring* (*of sets*) means that $\mathscr{H}$ is a collection of sets such that for all $A, B \in \mathscr{H}$, we have $A \cap B \in \mathscr{H}$ and $A \setminus B$ is a finite disjoint union of elements of $\mathscr{H}$. If X is a set, then to say that $\mathscr{H}$ *is a pre-ring of subsets of* X or, more briefly, that $\mathscr{H}$ *is a pre-ring on* X, means that $\mathscr{H}$ is a pre-ring of sets and each element of $\mathscr{H}$ is a subset of X.

Example. Let $\mathscr{H} = \{\, (a,b] : -\infty < a \le b < \infty \,\}$. Then $\mathscr{H}$ is a pre-ring on $\mathbf{R}$. Notice that $\mathscr{H}$ is not closed under relative complements. For instance $(1,4] \setminus (2,3]$ is not an element of $\mathscr{H}$, although it is the disjoint union of two elements of $\mathscr{H}$, namely $(1,2]$ and $(3,4]$

Example. The collection of all bounded intervals in $\mathbf{R}$ is a pre-ring on $\mathbf{R}$. However, this pre-ring is less convenient to take as a collection of basic sets than the one in the preceding example, because in our proofs, we would have more cases to consider.

X231. Let $\mathscr{H}_1$ and $\mathscr{H}_2$ be pre-rings on a set X. Let

$$\mathscr{H} = \{\, A \cap B : A \in \mathscr{H}_1 \text{ and } B \in \mathscr{H}_2 \,\}.$$

Prove that $\mathscr{H}$ is a pre-ring.

Reminder. Let $\mathscr{A}$ and $\mathscr{B}$ be collections of sets. Then by definition,

$$\mathscr{A} \odot \mathscr{B} = \{\, A \times B : A \in \mathscr{A} \text{ and } B \in \mathscr{B} \,\}.$$

Warning. You should not write $\mathscr{A} \times \mathscr{B}$ for $\mathscr{A} \odot \mathscr{B}$. The notation $\mathscr{A} \times \mathscr{B}$ denotes a different collection, namely the set of ordered pairs $\{\, (A,B) : A \in \mathscr{A} \text{ and } B \in \mathscr{B} \,\}$.

Reminder. As we saw in class, it follows from the result of problem X231 that if $\mathscr{H}_1$ and $\mathscr{H}_2$ are pre-rings on sets X_1 and X_2 respectively, and if $\mathscr{H} = \mathscr{H}_1 \odot \mathscr{H}_2$, then $\mathscr{H}$ is a pre-ring on $X_1 \times X_2$. For instance, the collection

$$\{\, (a,b] \times (c,d] : -\infty < a \le b < \infty \text{ and } -\infty < c \le d < \infty \,\}$$

is a pre-ring on $\mathbf{R}^2$. Of course the result of problem X231 can be generalized by induction to any finite number of pre-rings instead of just two. From this it follows that if $\mathscr{H}_k$ is a pre-ring on a set X_k for $k = 1, \ldots, n$, and if $\mathscr{H}$ is the collection of all Cartesian products of the form $\prod_{k=1}^n A_k$, where $A_k \in \mathscr{H}_k$ for $k = 1, \ldots, n$, then $\mathscr{H}$ is a pre-ring on $\prod_{k=1}^n X_k$. For instance, the collection of Cartesian products of the form $\prod_{k=1}^n (a_k, b_k]$, where $-\infty < a_k \le b_k < \infty$ for $k = 1, \ldots, n$, is a pre-ring on $\mathbf{R}^n$.

X232. Let Φ be a collection of pre-rings. Suppose that Φ is upwards directed by set-inclusion.[36] Let $\mathscr{H} = \bigcup \Phi$. Prove that $\mathscr{H}$ is a pre-ring.

Definition. Let X be a set. To say that $\mathscr{H}$ *is a pre-field of subsets of* X or, more briefly, that $\mathscr{H}$ *is a pre-field on* X, means that $\mathscr{H}$ is a pre-ring on X and $X \in \mathscr{H}$. (Another name for a pre-field is a *pre-algebra*.)

Example. The collection of all intervals in $\mathbf{R}$, including unbounded intervals, is a pre-field on $\mathbf{R}$.

Example. Let $\mathscr{H} = \{\, (a,b] : 0 \le a \le b \le 1 \,\}$. Then $\mathscr{H}$ is a pre-field on $(0,1]$.

[36] To say that Φ *is upwards directed by set-inclusion* means that for each $\mathscr{H}_1 \in \Phi$ and each $\mathscr{H}_2 \in \Phi$, there exists $\mathscr{H}_3 \in \Phi$ such that $\mathscr{H}_1 \subseteq \mathscr{H}_3$ and $\mathscr{H}_2 \subseteq \mathscr{H}_3$.

Remark. The next exercise is relevant to the study of measure in infinite dimensional spaces. (The theory of measure in infinite dimensional spaces is important in probability theory, but we shall not pursue this topic in this course.)

X233. Let $(X_\alpha)_{\alpha\in A}$ be a family of sets and let $X = \prod_{\alpha\in A} X_\alpha$. For each $\alpha \in A$, let $\mathscr{H}_\alpha$ be a pre-field on X_α. Let $\mathscr{H}$ be the collection of all sets of the form $\prod_{\alpha\in A} H_\alpha$, where $H_\alpha \in \mathscr{H}_\alpha$ for each $\alpha \in A$ and $H_\alpha = X_\alpha$ for all but finitely many $\alpha \in A$. Prove that $\mathscr{H}$ is a pre-field on X. (Hint: For each finite set $B \subseteq A$, let $\mathscr{H}_B$ be the collection of all sets of the form $\prod_{\alpha\in A} H_\alpha$, where $H_\alpha \in \mathscr{H}_\alpha$ for each $\alpha \in A$ and $H_\alpha = X_\alpha$ for each $\alpha \in A \setminus B$. Use earlier work to show that each $\mathscr{H}_B$ is a pre-field on X. Let

$$\Phi = \{\, \mathscr{H}_B : B \text{ is a finite subset of } A \,\}.$$

Notice that $\mathscr{H} = \bigcup \Phi$. To finish, use problem X232.)

Remark. After the basic sets that we start with in our study of measure, the next simplest kind of sets are the ones which are finite disjoint unions of the basic sets. If the collection of basic sets that we start with is a pre-ring, then the collection of finite disjoint unions of such basic sets will be a ring.

Reminder. To say that $\mathscr{R}$ *is a ring* (*of sets*) means that $\mathscr{R}$ is a collection of sets such that $\varnothing \in \mathscr{R}$ and for all $A, B \in \mathscr{R}$, we have $A \cup B \in \mathscr{R}$ and $A \setminus B \in \mathscr{R}$. If X is a set, then to say that $\mathscr{R}$ *is a ring of subsets of* X or, more briefly, that $\mathscr{R}$ *is a ring on* X, means that $\mathscr{R}$ is a ring of sets and each element of $\mathscr{R}$ is a subset of X.

X234. Let $\mathscr{R}$ be a ring of sets. Prove that $\mathscr{R}$ is a pre-ring. (Obviously, to prove this, it remains only to show that $\mathscr{R}$ is closed under pairwise intersections. In other words, you just need to prove that for all $A, B \in \mathscr{R}$, we have $A \cap B \in \mathscr{R}$.)

Example. Let X be a set with at least two elements. Let $\mathscr{S} = \big\{ \{ x \} : x \in X \big\} \cup \{ \varnothing \}$. Then $\varnothing \in \mathscr{S}$ and $\mathscr{S}$ is closed under pairwise intersections and relative complements, but $\mathscr{S}$ is not a ring of sets.

X235. (*If you attended recitation on Tuesday, March 27, then you are excused from turning in this exercise.*) *due 1F*
Let $\mathscr{H}$ be a pre-ring. Let $\mathscr{R}$ be the collection of finite disjoint unions of elements of $\mathscr{H}$. The object of this exercise is to lead you through a proof of the fact that $\mathscr{R}$ is a ring. Notice that $\varnothing \in \mathscr{R}$ because $\varnothing$ is the union of the empty collection of elements of $\mathscr{H}$. Notice also that $\mathscr{R}$ is obviously closed under finite *disjoint* unions.

- **(a)** Let $A, B \in \mathscr{R}$. Prove that $A \cap B \in \mathscr{R}$.
- **(b)** Let $A \in \mathscr{H}$ and let $B \in \mathscr{R}$. Prove that $A \setminus B \in \mathscr{R}$.
- **(c)** Let $A, B \in \mathscr{R}$. Prove that $A \setminus B \in \mathscr{R}$.
- **(d)** Let $A, B \in \mathscr{R}$. Prove that $A \cup B \in \mathscr{R}$.
- **(e)** Prove that $\mathscr{R}$ is the smallest ring of sets that contains $\mathscr{H}$.

Remark. Let A be an interval, let (B_k) be a finite sequence of intervals, and let $C = A \setminus \bigcup_k B_k$. Since the collection of all intervals is a pre-ring, it follows from problem X235 that C is a finite disjoint union of intervals. It is important to realize that the analog of this statement when we replace "finite" by "countable" is false!

X236. Give two examples of an interval A and an infinite sequence of intervals (B_k) such that the set $C = A \setminus \bigcup_k B_k$ is not a countable disjoint union of intervals. (Hint: For one example, think about the set of irrational numbers. For another example, think about the Cantor set.)

Definition. Let X be a set. To say that $\mathscr{F}$ *is a field of subsets of* X or, more briefly, that $\mathscr{F}$ *is a field on* X means that $\mathscr{F}$ is a ring on X and $X \in \mathscr{F}$. (Another name for a field in this context is *an algebra*.)

Remark. Let X be a set, let $\mathscr{H}$ be a pre-field on X, and let $\mathscr{F}$ be the set of finite disjoint unions of elements of $\mathscr{H}$. Then $\mathscr{F}$ is a field on X. This obviously follows from problem X235.

Reminder. Let $\mathscr{H} = \{\, (u, v] : -\infty < u \le v < \infty \,\}$. As we know, $\mathscr{H}$ is a pre-ring on $\mathbf{R}$. Let V be a commutative group. Consider any function $F\colon \mathbf{R} \to V$. Define $\tau\colon \mathscr{H} \to V$ by

$$\tau((u, v]) = F(v) - F(u) \tag{31}$$

for all $u, v \in \mathbf{R}$ with $u \le v$. (Notice that τ is well-defined because if $(u, v] \ne \emptyset$, then u and v are uniquely determined by $(u, v]$, while if $(u, v] = \emptyset$, then $u = v$, so $F(v) - F(u) = 0$.) As we saw in class, τ is additive. The most important special case is the one where $V = \mathbf{R}$ and $F(x) \equiv x$. In this case, $\tau((u, v]) = v - u$, the length of the interval $(u, v]$.

Remark. In the special case where $V = \mathbf{R}$ and where $F\colon \mathbf{R} \to \mathbf{R}$ is C^1, for all $u, v \in \mathbf{R}$ with $u \le v$, we have

$$F(v) - F(u) = \int_u^v \frac{dF}{dx}\, dx,$$

so in this case (31) may rewritten as

$$\tau((u, v]) = \int_u^v \frac{dF}{dx}\, dx. \tag{32}$$

We shall not make mathematical use of (32), which is just as well since we are in the process of redeveloping the theory of integration "from scratch." We only mention (32) because its generalizations to higher dimensions will be useful to help us remember the generalizations of (31) to higher dimensions.

X237. Let $\mathscr{H} = \{\, (u, v] : -\infty < u \le v < \infty \,\}$. Let V be a commutative group. Consider any additive function $\tau\colon \mathscr{H} \to V$. Prove that there exists a function $F\colon \mathbf{R} \to V$ such that for all $u, v \in \mathbf{R}$ with $u \le v$, we have

$$\tau((u, v]) = F(v) - F(u).$$

Although F is not unique, because we can add any constant to it, show that one suitable choice for F is given by

$$F(x) = \begin{cases} \tau((0, x]) & \text{if } x \ge 0, \\ -\tau((x, 0]) & \text{if } x < 0. \end{cases}$$

Remark. In problem X237, F is called *a distribution function for* τ.

X238. Let $\mathscr{H}$ be a pre-ring and let $\mathscr{E}$ be a finite collection of elements $\mathscr{H}$. Prove that there is a finite disjoint collection $\mathscr{A} \subseteq \mathscr{H}$ such that $\bigcup \mathscr{A} = \bigcup \mathscr{E}$ and for each $E \in \mathscr{E}$, we have $E = \bigcup \mathscr{A}_E$, where

$$\mathscr{A}_E = \{\, A \in \mathscr{A} : A \subseteq E \,\}.$$

X239. Prove that the analog of problem X238 with "finite" replaced by "countable" is false. (Hint: Consider the countable collection of sets $\mathscr{E} = \{\, (-\infty, x] : x \in \mathbf{Q} \,\}$. You can take $\mathscr{H}$ to be the pre-ring of all subsets of $\mathbf{R}$.)

X240. Let $\mathscr{H}_1$ and $\mathscr{H}_2$ be pre-rings. Let $\mathscr{H} = \{\, A \cap B : A \in \mathscr{H}_1 \text{ and } B \in \mathscr{H}_2 \,\}$. As we know from problem X231, $\mathscr{H}$ is a pre-ring. Let V be a commutative semigroup with 0 and let $\tau\colon \mathscr{H} \to V$. Suppose that for each $B \in \mathscr{H}_2$, the function $A \mapsto \tau(A \cap B)$ is additive on $\mathscr{H}_1$, and that for each $A \in \mathscr{H}_1$, the function $B \mapsto \tau(A \cap B)$ is additive on $\mathscr{H}_2$. Prove that τ is additive on $\mathscr{H}$. (Hint: problem X238 should help.)

X241. Let $\mathscr{H}_1$ and $\mathscr{H}_2$ be pre-rings on sets X_1 and X_2 respectively. Let $\mathscr{H} = \mathscr{H}_1 \odot \mathscr{H}_2$. As we saw in class, $\mathscr{H}$ is is a pre-ring on $X_1 \times X_2$. Let V be a commutative semigroup with 0 and let $\tau\colon \mathscr{H} \to V$. Suppose that for each $B \in \mathscr{H}_2$, the function $A \mapsto \tau(A \times B)$ is additive on $\mathscr{H}_1$, and that for each $A \in \mathscr{H}_1$, the function $B \mapsto \tau(A \times B)$ is additive on $\mathscr{H}_2$. Prove that τ is additive on $\mathscr{H}$. (Hint: This can be deduced easily from problem X240.)

Remark. Of course the results of problem X240 and problem X241 can be generalized from two pre-rings to any finite number of pre-rings.

X242. Let $\mathscr{H}_1 = \{ (u,v] : -\infty < u \le v < \infty \}$. As we know, $\mathscr{H}_1$ is a pre-ring on $\mathbf{R}$. Let $\mathscr{H}_2$ be any pre-ring and let $\mathscr{H} = \mathscr{H}_1 \odot \mathscr{H}_2$. Let V be a commutative group. Let W be the set of all finitely additive functions from $\mathscr{H}_2$ to V. If $\sigma, \tau \in W$, then of course $\sigma + \tau$ denotes the function from $\mathscr{H}_2$ to V defined by $(\sigma+\tau)(B) = \sigma(B) + \tau(B)$ for all $B \in \mathscr{H}_2$, and it is obvious that $\sigma + \tau \in W$. Notice that with this operation of addition, W is itself a commutative group. Consider any function $G\colon \mathbf{R} \to W$. Define $\tau\colon \mathscr{H} \to V$ by

$$\tau((u,v] \times B) = G(v)(B) - G(u)(B)$$

for all $u, v \in \mathbf{R}$ with $u \le v$ and all $B \in \mathscr{H}_2$. (You should explain why τ is well-defined.) Prove that τ is additive on $\mathscr{H}$. (Hint: Use problem X241.)

X243. Let $\mathscr{H}_1 = \mathscr{H}_2 = \{ (u,v] : -\infty < u \le v < \infty \}$ and let $\mathscr{H} = \mathscr{H}_1 \odot \mathscr{H}_2$. As we know, $\mathscr{H}_1$ and $\mathscr{H}_2$ are pre-rings on $\mathbf{R}$ and $\mathscr{H}$ is a pre-ring on $\mathbf{R}^2$. Let V be a commutative group. Consider any function $F\colon \mathbf{R}^2 \to V$. Define $\tau\colon \mathscr{H} \to V$ by

$$\tau((u_1,v_1] \times (u_2,v_2]) = F(v_1,v_2) - F(u_1,v_2) - F(v_1,u_2) + F(u_1,u_2) \tag{33}$$

for all $u_1, v_1, u_2, v_2 \in \mathbf{R}$ with $u_1 \le v_1$ and $u_2 \le v_2$. (You should explain why τ is well-defined.) Prove that τ is additive on $\mathscr{H}$. (Hint: Use problem X242.)

Remark. Notice that even if $V = \mathbf{R}$ in problem X243, the W that we consider in applying problem X242 to solve problem X243 is an infinite dimensional vector space over $\mathbf{R}$, namely the space of all additive real-valued functions on the pre-ring $\{ (u,v] : -\infty < u \le v < \infty \}$. Thus we finally see a significant payoff to our having taken the trouble from the outset to consider additive functions taking values in semigroups more general than $\mathbf{R}$.

Remark. In the special case where $V = \mathbf{R}$ and where $F\colon \mathbf{R}^2 \to \mathbf{R}$ is C^2, for all $u_1, v_1, u_2, v_2 \in \mathbf{R}$ with $u_1 \le v_1$ and $u_2 \le v_2$, we have

$$F(v_1,v_2) - F(u_1,v_2) - F(v_1,u_2) + F(u_1,u_2) = \int_{u_1}^{v_1} \int_{u_2}^{v_2} \frac{\partial^2 F}{\partial x_2 \partial x_1}\, dx_2\, dx_1,$$

so in this case (33) may rewritten as

$$\tau((u_1,v_1] \times (u_2,v_2]) = \int_{u_1}^{v_1} \int_{u_2}^{v_2} \frac{\partial^2 F}{\partial x_2 \partial x_1}\, dx_2\, dx_1.$$

Remark. Let us consider a particular case of the situation treated in problem X243. Let $V = \mathbf{R}$ and let $F\colon \mathbf{R}^2 \to \mathbf{R}$ satisfy

$$F(x_1,x_2) = F_1(x_1)F_2(x_2)$$

for all $x_1, x_2 \in \mathbf{R}$, where $F_1, F_2\colon \mathbf{R} \to \mathbf{R}$. Then for all $u_1, v_1, u_2, v_2 \in \mathbf{R}$ with $u_1 \le v_1$ and $u_2 \le v_2$, we have

$$\begin{aligned}
\tau((u_1,v_1] \times (u_2,v_2]) &= F_1(v_1)F_2(v_2) - F_1(u_1)F_2(v_2) - F_1(v_1)F_2(u_2) + F_1(u_1)F_2(u_2) \\
&= \big(F_1(v_1) - F_1(u_1)\big)F_2(v_2) - \big(F_1(v_1) - F_1(u_1)\big)F_2(u_2) \\
&= \big(F_1(v_1) - F_1(u_1)\big)\big(F_2(v_2) - F_2(u_2)\big).
\end{aligned}$$

The most important special case is the one where

$$F(x_1,x_2) = x_1 x_2$$

for all $x_1, x_2 \in \mathbf{R}$. In this case, $\tau((u_1,v_1] \times (u_2,v_2]) = (v_1 - u_1)(v_2 - u_2)$, the area of the rectangle $(u_1,v_1] \times (u_2,v_2]$.

Remark. The next exercise is a converse to problem X243.

X244. Let $\mathscr{H}_1 = \mathscr{H}_2 = \{\, (u,v] : -\infty < u \le v < \infty \,\}$ and let $\mathscr{H} = \mathscr{H}_1 \odot \mathscr{H}_2$. As we know, $\mathscr{H}_1$ and $\mathscr{H}_2$ are pre-rings on $\mathbf{R}$ and $\mathscr{H}$ is a pre-ring on $\mathbf{R}^2$. Let V be a commutative group. Let $\tau\colon \mathscr{H} \to V$ be additive. Prove that there exists a function $F\colon \mathbf{R}^2 \to V$ such that for all $u_1, v_1, u_2, v_2 \in \mathbf{R}$ with $u_1 \le v_1$ and $u_2 \le v_2$, we have

$$\tau((u_1, v_1] \times (u_2, v_2]) = F(v_1, v_2) - F(u_1, v_2) - F(v_1, u_2) + F(u_1, u_2)$$

(Hint: Let W be the group of all finitely additive functions from $\mathscr{H}_2$ to V. Define $\sigma\colon \mathscr{H}_1 \to W$ by $\sigma(A)(B) = \tau(A \times B)$ for all $A \in \mathscr{H}_1$ and all $B \in \mathscr{H}_2$. Check that σ is additive. Hence, by problem X237, there is a function $G\colon \mathbf{R} \to W$ such that for all $u_1, v_1 \in \mathbf{R}$ with $u_1 \le v_1$, $\sigma((u_1, v_1]) = G(v_1) - G(u_1)$. To finish, for each $x_1 \in \mathbf{R}$, apply problem X237 to the additive function $B \mapsto G(x_1)(B)$ on $\mathscr{H}_2$.)

Remark. In problem X244, F is called *a distribution function for* τ. Of course F is not unique. It is not even unique up to a constant. Rather, one may add to it any function C from $\mathbf{R}^2$ to V which satisfies

$$C(v_1, v_2) - C(v_1, u_2) - C(u_1, v_2) + C(u_1, v_1) = 0$$

for all $u_1, v_1, u_2, v_2 \in \mathbf{R}$ with $u_1 \le v_1$ and $u_2 \le v_2$. When $V = \mathbf{R}$ and $C\colon \mathbf{R}^2 \to \mathbf{R}$ is C^2, it can be shown that this condition on C is equivalent to the condition

$$\frac{\partial^2 C}{\partial x_2 \partial x_1} \equiv 0.$$

Remark. Now we wish to generalize problem X243 and problem X244 from $\mathbf{R}^2$ to $\mathbf{R}^n$. To formulate the generalization succinctly, we introduce some notation. But first, to prepare for interpreting a particular case of this generalization, we consider a generalization of the binomial theorem.

X245. Prove that for each $n \in \mathbf{N}$, for all $b_1, \ldots, b_n \in \mathbf{R}$, we have

$$\prod_{i=1}^{n} (1 + b_i) = \sum_I \prod_{i \in I} b_i,$$

where the index I of summation ranges over all subsets of the set $\{\, 1, \ldots, n \,\}$. (Note: If $I = \varnothing$, then $\prod_{i \in I} b_i = 1$ by convention. This makes sense because for each real number c, the product of c and $\prod_{i \in \varnothing} b_i$) should be c.)

Remark. By a slight refinement of the method of proof for problem X245, one can prove the *generalized binomial theorem*, which states that for each $n \in \mathbf{N}$, for all real numbers $a_1, b_1, \ldots, a_n, b_n$, we have

$$\prod_{i=1}^{n} (a_i + b_i) = \sum_I \left[\Big(\prod_{i \notin I} a_i \Big) \Big(\prod_{i \in I} b_i \Big) \right],$$

where again the index I of summation ranges over all subsets of the set $\{\, 1, \ldots, n \,\}$ and where for each such I, the i in $\prod_{i \notin I} a_i$ ranges over the set $\{\, 1, \ldots, n \,\} \setminus I$.

Notation. Let $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$ be elements of $\mathbf{R}^n$. To write $u \le v$ means that for $k = 1, \ldots, n$, we have $u_k \le v_k$. If $u \le v$, then by definition,

$$(u, v] = \prod_{k=1}^{n} (u_k, v_k].$$

If $I \subseteq \{\, 1, \ldots, n \,\}$, then we'll write $u : I : v$ for the n-tuple $w = (w_1, \ldots, w_n)$ defined by

$$w_k = \begin{cases} u_k & \text{if } k \in I, \\ v_k & \text{if } k \notin I. \end{cases}$$

(This is not standard notation.) For instance, if $I = \varnothing$, then $u : I : v = v$, while if $I = \{\, 1, \ldots, n \,\}$, then $u : I : v = u$.

X246. Let $\mathscr{H} = \{\,(u,v] : u,v \in \mathbf{R}^n \text{ and } u \le v\,\}$. As we know, $\mathscr{H}$ is a pre-ring on $\mathbf{R}^n$. Let V be a commutative group.

(a) Consider any function $F\colon \mathbf{R}^n \to V$. Define $\tau\colon \mathscr{H} \to V$ by

$$\tau((u,v]) = \sum_I (-1)^{|I|} F(u : I : v) \tag{34}$$

for all $u,v \in \mathbf{R}^n$ with $u \le v$, where the index I of summation ranges over all subsets of the set $\{\,1,\ldots,n\,\}$ and where $|I|$ denotes the number of elements in I. (You should explain why τ is well-defined.) Prove that τ is additive on $\mathscr{H}$.

(b) Conversely, consider any additive function $\tau\colon \mathscr{H} \to V$. Prove that there exists a function $F\colon \mathbf{R}^n \to V$ such that for all $u,v \in \mathbf{R}^n$ with $u \le v$, (34) holds.

Remark. In the special case where $V = \mathbf{R}$ and where $F\colon \mathbf{R}^n \to \mathbf{R}$ is C^n, it is easy to show by induction on n that for all $u = (u_1,\ldots,u_n) \in \mathbf{R}^n$ and all $v = (v_1,\ldots,v_n) \in \mathbf{R}^n$ with $u \le v$, we have

$$\sum_I (-1)^{|I|} F(u : I : v) = \int_{u_1}^{v_1} \cdots \int_{u_n}^{v_n} \frac{\partial^n F}{\partial x_n \cdots \partial x_1}\, dx_n \cdots dx_1,$$

so in this case (34) may rewritten as

$$\tau((u,v]) = \int_{u_1}^{v_1} \cdots \int_{u_n}^{v_n} \frac{\partial^n F}{\partial x_n \cdots \partial x_1}\, dx_n \cdots dx_1.$$

Remark. Let us consider a particular case of the situation treated in problem X246(a). Let $V = \mathbf{R}$ and let $F\colon \mathbf{R}^n \to \mathbf{R}$ satisfy

$$F(x_1,\ldots,x_n) = \prod_{k=1}^n F_k(x_k)$$

for all $x_1,\ldots,x_n \in \mathbf{R}$, where $F_1,\ldots,F_n\colon \mathbf{R} \to \mathbf{R}$. Then for all $u = (u_1,\ldots,u_n) \in \mathbf{R}^n$ and all $v = (v_1,\ldots,v_n) \in \mathbf{R}^n$ with $u \le v$, we have

$$\tau((u,v]) = \sum_I (-1)^{|I|} F(u : I : v) = \sum_I \prod_{k=1}^n (-1)^{1_I(k)} F((u : I : v)_k) = \prod_{k=1}^n \bigl(F_k(v_k) - F_k(u_k)\bigr),$$

where we used the generalized binomial theorem in the last step. The most important special case is the one where

$$F(x_1,\ldots,x_2) = \prod_{k=1}^n x_k$$

for all $x_1,\ldots,x_n \in \mathbf{R}$. In this case, $\tau((u,v]) = \prod_{k=1}^n (v_k - u_k)$, the n-dimensional measure of $(u,v]$. (Length is 1-dimensional measure, area is 2-dimensional measure, volume is 3-dimensional measure, and so on.)

Remark. In problem X246(b), F is called *a distribution function for* τ. Of course F is not unique. It is not even unique up to a constant. Rather, one may add to it any function C from $\mathbf{R}^n$ to V which satisfies

$$\sum_I (-1)^{|I|} C(u : I : v) = 0$$

for all $u \in \mathbf{R}^n$ with $u \le v$. When $V = \mathbf{R}$ and $C\colon \mathbf{R}^n \to \mathbf{R}$ is C^n, it can be shown that this condition on C is equivalent to the condition

$$\frac{\partial^n C}{\partial x_n \cdots \partial x_1} \equiv 0.$$

X247. Let $\mathscr{R}$ be a ring of sets and let $\tau\colon \mathscr{R} \to [-\infty,\infty]$. Suppose that for all disjoint $A,B \in \mathscr{R}$, we have $\tau(A \cup B) = \tau(A) + \tau(B)$. (It is implicit in this assumption that for all such A and B, $\tau(A) + \tau(B)$ is defined, or in other words, $\tau(A) + \tau(B)$ is not of the form $\infty + (-\infty)$, nor is it of the form $(-\infty) + \infty$.)

(a) Prove that for all $A,B \in \mathscr{R}$, if $A \subseteq B$ and $\tau(A)$ is finite, then $\tau(B \setminus A) = \tau(B) - \tau(A)$.

(b) Prove that for all $A,B \in \mathscr{R}$, if $A \subseteq B$ and $\tau(A) = \infty$, then $\tau(B) = \infty$.

(c) Prove that for all $A,B \in \mathscr{R}$, if $A \subseteq B$ and $\tau(A) = -\infty$, then $\tau(B) = -\infty$.

(d) Prove that τ takes on at most one of the values ∞ and $-\infty$.

(e) For definiteness, assume that τ does not take on the value $-\infty$. Thus τ takes values in $(-\infty,\infty]$. Of course under addition, $(-\infty,\infty]$ is a commutative semigroup with 0. Suppose in addition that $\tau(C) \ne \infty$ for some $C \in \mathscr{R}$. Prove that τ is additive.

Terminology. In the light of problem X247, we shall slightly extend the definition of additivity. Let $\mathscr{R}$ be a ring of sets and let $\tau\colon \mathscr{R} \to [-\infty, \infty]$. To say that τ is additive will mean that τ takes values in $(-\infty, \infty]$ or in $[-\infty, \infty)$ and that τ is additive in the sense already defined for a set function taking values in a commutative semigroup with 0.

X248. Give an example of a pre-ring $\mathscr{H}$ and a function $\tau\colon \mathscr{H} \to [-\infty, \infty]$ such that for all disjoint $A, B \in \mathscr{H}$ for which $A \cup B$ happens to belong to $\mathscr{H}$, we have $\tau(A \cup B) = \tau(A) + \tau(B)$, and yet τ takes on both of the values $-\infty$ and ∞. (Hint: There is an example where $\mathscr{H}$ is a finite set.)

X249. Let $\mathscr{R}$ be a ring of sets, let V be a commutative semigroup with 0, and let $\tau\colon \mathscr{R} \to V$ be additive. Prove that for all $A, B \in \mathscr{R}$, we have

$$\tau(A \cup B) + \tau(A \cap B) = \tau(A) + \tau(B).$$

(Warning: Your proof must not involve subtraction, such as $\tau(C) - \tau(D)$, where $C, D \in \mathscr{R}$. Since V is only a semigroup, $\tau(C) - \tau(D)$ need not be defined. For instance, if $V = [0, \infty]$ under addition, then $\tau(C) - \tau(D)$ could be $\infty - \infty$. You should be able to write your proof so that it only involves addition.)

X250. Let $\mathscr{R}$ be a ring of sets, let V be a commutative group, and let $\mu\colon \mathscr{R} \to V$ be additive. Then by problem X249, for all $A, B \in \mathscr{R}$, we have

$$\mu(A \cup B) + \mu(A \cap B) = \mu(A) + \mu(B). \tag{35}$$

Now since we are assuming that V is a commutative *group*, it follows from (35) that for all $A, B \in \mathscr{R}$, we have

$$\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B).$$

From this, it is a simple calculation to show that if $A, B, C \in \mathscr{R}$, then

$$\begin{aligned}\mu(A \cup B \cup C) = \mu(A) &+ \mu(B) + \mu(C) \\ &- \mu(A \cap B) - \mu(A \cap C) - \mu(B \cap C) \\ &+ \mu(A \cap B \cap C).\end{aligned}$$

The corresponding formula when we consider n sets $A_1, \dots, A_n \in \mathscr{R}$ is

$$\mu\big(\textstyle\bigcup_{i=1}^n A_i\big) = \displaystyle\sum_i \mu(A_i) - \sum_{i<j} \mu(A_i A_j) + \sum_{i<j<k} \mu(A_i A_j A_k) - + \cdots + (-1)^{n+1} \mu(A_1 \cdots A_n), \tag{36}$$

where to save space, we have written $A_i A_j$ for $A_i \cap A_j$ and so on. Equation (36) is called *the inclusion-exclusion formula*. It can be proved by induction on n. Do so.

Remark. Let us note that (36) can be written more succinctly as follows.

$$\mu\big(\textstyle\bigcup_{i=1}^n A_i\big) = \displaystyle\sum_{I \neq \varnothing} (-1)^{|I|+1} \mu\Big(\bigcap_{i \in I} A_i\Big),$$

where the index I of summation ranges over all non-empty subsets of the set $\{1, \dots, n\}$ and where $|I|$ denotes the number of elements in I. To see this, write $\mathscr{J}_m$ for the set of m-element subsets of $\{1, \dots, n\}$, and notice that the sum on the right in (36) is

$$\begin{aligned}\sum_{m=1}^n (-1)^{m+1} \sum_{I \in \mathscr{J}_m} \mu\Big(\bigcap_{i \in I} A_i\Big) &= \sum_{m=1}^n \sum_{I \in \mathscr{J}_m} (-1)^{m+1} \mu\Big(\bigcap_{i \in I} A_i\Big) \\ = \sum_{m=1}^n \sum_{I \in \mathscr{J}_m} (-1)^{|I|+1} \mu\Big(\bigcap_{i \in I} A_i\Big) &= \sum_{I \neq \varnothing} (-1)^{|I|+1} \mu\Big(\bigcap_{i \in I} A_i\Big),\end{aligned}$$

where the last step follows from the fact that each non-empty subset $I \subseteq \{1, \dots, n\}$ belongs to exactly one of the sets $\mathscr{J}_1, \dots, \mathscr{J}_n$.

X251. Suppose n men put their hats in a sack and then each man draws a hat from the sack at random. *due 2Th*

(a) Determine the probability that at least one man gets his own hat back. (Hint: Let $E = \{1, \ldots, n\}$ and let Ω be the set of one-to-one maps from E into E. Let $\mathscr{F}$ be the set of all subsets of Ω. Then $\mathscr{F}$ is a ring of sets on Ω. For each $A \in \mathscr{F}$, let $|A|$ denote the number of elements in A. Define $P\colon \mathscr{F} \to [0,1]$ by

$$P(A) = \frac{|A|}{|\Omega|}$$

for each $A \in \mathscr{F}$. Then P is additive on $\mathscr{F}$. We seek $P(B)$, where $B = \bigcup_{i=1}^n A_i$ and $A_i = \{\omega \in \Omega : \omega(i) = i\}$.)

(b) Find the limit that the probability in (a) approaches as $n \to \infty$. (The number e will play a role.)

Definition. A *topological semigroup* is a semigroup V endowed with a topology which makes the semigroup operation, which is a function from $V \times V$ into V, continuous.

Examples. $[0,\infty]$ under addition, and endowed with its usual topology, is a topological semigroup. So are $(-\infty,\infty]$, $\mathbf{R}$, and $\mathbf{R}^n$. All of these examples are commutative and have 0. The last two are groups.

Unordered sums. Let V be a Hausdorff topological commutative semigroup with 0. Let $(y_k)_{k\in K}$ be a family of elements of V. (The index set K is arbitrary. It could even be uncountable.) Let Λ be the collection of finite subsets of K. For each $\lambda \in K$, let

$$S_\lambda = \sum_{k\in\lambda} y_k.$$

Notice that Λ becomes a directed set when ordered by inclusion. This makes $(S_\lambda)_{\lambda\in\Lambda}$ into a net in V. To say that y *is a sum for* $(y_k)_{k\in K}$ *in* V means that the net (S_λ) converges to y in V. (Explicitly, this means that for each nhd U of y in V, there exists a finite set $\lambda_0 \subseteq K$ such that for each finite set $\lambda \subseteq K$ with $\lambda_0 \subseteq \lambda$, we have $\sum_{k\in\lambda} y_k \in U$.) Since V is Hausdorff, limits in V are unique, so there can be at most one such y. Hence if y is a sum for $(y_k)_{k\in K}$ in V, then it makes sense to write $\sum_{k\in K} y_k = y$ and to call y *the* sum of the family $(y_k)_{k\in K}$ in V. To say that $(y_k)_{k\in K}$ *is summable in* V means that there exists y in V such that y is a sum for $(y_k)_{k\in K}$ in V.

X252. Let $(y_k)_{k\in K}$ be a family in $[0,\infty]$. As we've mentioned, under addition, $[0,\infty]$ is a Hausdorff topological semigroup with 0. Let Λ be the collection of finite subsets of K. For each $\lambda \in \Lambda$, let $S_\lambda = \sum_{k\in\lambda} y_k$. Let $y = \sup\{S_\lambda : \lambda \in \Lambda\}$. Prove that $y = \sum_{k\in K} y_k$ in the sense discussed immediately above.

Remark. The next exercise is a sort of generalized associative law for unordered sums. While it is not the best result of its type, it will be sufficient for the application we have in mind, which will be the extension of a countably additive set function from a pre-ring to a ring.

X253. Let V be a Hausdorff topological commutative semigroup with 0. Let $(y_k)_{k\in K}$ be a family in V. Suppose this family is summable in V and let $y = \sum_{k\in K} y_k$. Let $(K_\ell)_{\ell\in L}$ be a disjoint family of finite subsets of K whose union is all of K. For each $\ell \in L$, let $z_\ell = \sum_{k\in K_\ell} y_k$. Prove that $\sum_{\ell\in L} z_\ell = y$ too.

Reminder. If S is a set, then $S^{<\mathbf{N}}$ denotes the set of finite sequences of elements of S, including the empty sequence, $S^{\mathbf{N}}$ denotes the set of infinite sequences of elements of S (in other words, the set of functions from $\mathbf{N}$ into S), and $S^{\le\mathbf{N}}$ denotes the set of finite or infinite sequences of elements of S (in other words, $S^{<\mathbf{N}} \cup S^{\mathbf{N}}$).

Reminder. Let $\mathscr{H}$ be a collection of sets, let V be a Hausdorff topological commutative semigroup with 0, and let $\tau\colon \mathscr{H} \to V$. To say that τ *is countably additive*, or that τ *is σ-additive*, means that for each disjoint $(H_k) \in \mathscr{H}^{\le\mathbf{N}}$, if $\bigcup_k H_k \in \mathscr{H}$, then $\tau(\bigcup_k H_k) = \sum_k \tau(H_k)$. If τ is countably additive, then τ is additive, because $\mathscr{H}^{<\mathbf{N}} \subseteq \mathscr{H}^{\le\mathbf{N}}$. In particular, if τ is countably additive and $\emptyset \in \mathscr{H}$, then $\tau(\emptyset) = 0$.

X254. Let $\mathscr{H}$ be a pre-ring and let $\mathscr{R}$ be the ring of sets generated by $\mathscr{H}$. Let τ be an additive set function on $\mathscr{H}$. As we've seen in class, τ has a unique extension to an additive set function σ on $\mathscr{R}$. Now suppose in addition that τ takes values in a Hausdorff topological commutative semigroup with 0 and that τ is countably additive. Prove that σ is countably additive.

X255. Let $\mathscr{R}$ be a ring of sets, let V be a Hausdorff commutative topological semigroup with 0, Let $\tau\colon \mathscr{R} \to V$. Prove that the following are equivalent.

(a) τ is countably additive on $\mathscr{R}$.

(b) τ is additive on $\mathscr{R}$ and for each increasing sequence (R_n) in $\mathscr{R}$, if $R = \bigcup_n R_n \in \mathscr{R}$, then $\tau(R_n) \to \tau(R)$ as $n \to \infty$.

X256. Let $\mathscr{H}$ be a pre-ring and let $\tau\colon \mathscr{H} \to [0,\infty]$. Prove that if τ is countably additive, then τ is countably subadditive. (Hint: Use problem X254 and mimic the proof, which we did in class, that if τ is additive, then τ is subadditive.)

X257. Let $\mathscr{R}$ be a ring of sets, let V be a Hausdorff topological commutative group, and let $\tau\colon \mathscr{R} \to V$.

(a) Prove that if τ is countably additive, if (R_n) is a decreasing sequence in $\mathscr{R}$, and if $R = \bigcap R_n \in \mathscr{R}$, then $\tau(R_n) \to \tau(R)$ as $n \to \infty$.

(b) Prove that if τ is additive and if for each decreasing sequence (R_n) in $\mathscr{R}$, with $\bigcap_n R_n = \emptyset$, we have $\tau(R_n) \to 0$ as $n \to \infty$, then τ is countably additive.

Example. Let $\mathscr{R}$ be the collection of all subsets of $\mathbf{N}$. Of course $\mathscr{R}$ is a ring of subsets of $\mathbf{N}$. Define $\tau\colon \mathscr{R} \to [0,\infty]$ as follows: For each $A \subseteq \mathbf{N}$, let $\tau(A)$ be the number of elements in A, if A is a finite set, and let $\tau(A) = \infty$ if A is an infinite set. (By the way, τ is called *counting measure on* $\mathbf{N}$.) It is easy to check that τ is countably additive. But τ does not satisfy the conclusion of problem X257(a). For let $R_n = \{\, n, n+1, n+2, \dots \,\}$ for each $n \in \mathbf{N}$ and let $R = \bigcap_n R_n$. Then (R_n) is a decreasing sequence of elements of $\mathscr{R}$ but $\tau(R_n)$ does not tend to $\tau(R)$ as $n \to \infty$, because $\tau(R_n) = \infty$ for each $n \in \mathbf{N}$ but $R = \emptyset$ so $\tau(R) = 0$. This does not contradict problem X257(a), because $[0,\infty]$ is just a semigroup under addition, not a group.

Remark. Let $\mathscr{R}$ be a ring of sets and let $\tau\colon \mathscr{R} \to [-\infty,\infty]$. Suppose that τ is additive. Then by problem X247, τ takes on at most one of the values ∞ and $-\infty$, so either τ takes values in $(-\infty,\infty]$ or τ takes values in $[-\infty,\infty)$. Hence it makes sense to talk about whether or not τ is countably additive. Suppose τ is countably additive. Let (R_n) be a decreasing sequence in $\mathscr{R}$ such that $-\infty < \tau(R_1) < \infty$. Let $R = \bigcap R_n$ and suppose $R \in \mathscr{R}$. Then $\tau(R_n) \to \tau(R)$ as $n \to \infty$. To see this, consider $\mathscr{S} = \{\, S \in \mathscr{R} : S \subseteq R_1 \,\}$. Then $\mathscr{S}$ itself is a ring of sets. Also, $R \in \mathscr{S}$ and each $R_n \in \mathscr{S}$. Let σ be the restriction of τ to $\mathscr{S}$. By parts (b) and (c) of problem X247, $\sigma\colon \mathscr{S} \to \mathbf{R}$. Of course σ is countably additive. Since $\mathbf{R}$ is a group under addition, it follows from problem X257(a) that $\sigma(R_n) \to \sigma(R)$ as $n \to \infty$. But this says that $\tau(R_n) \to \tau(R)$ as $n \to \infty$.

Remark. Length is countably additive on the collection of bounded intervals in $\mathbf{R}$. The example in the next exercise shows that the analogous statement with $\mathbf{R}$ replaced by $\mathbf{Q}$ is false. This example also shows that for problem X257(b), it matters that $\mathscr{R}$ is a ring. A pre-ring would not do.

X258. Let X be a totally ordered set.[37] For all $u, v \in X$ with $u \le v$, let $(u,v] = \{\, x \in X : u < x \le v \,\}$. *due 2Th*
Let $\mathscr{H} = \{\, (u,v] : u, v \in X,\ u \le v \,\}$.

(a) Prove that $\mathscr{H}$ is a pre-ring on X.

(b) Let $A, B \in \mathscr{H}$. By the definition of $\mathscr{H}$, $A = (s,t]$ and $B = (u,v]$ for some $s,t,u,v \in X$ with $s \le t$ and $u \le v$. Suppose that $\emptyset \ne A \subseteq B$. Prove that $u \le s < t \le v$. In particular, if $\emptyset \ne A = B$, then $s = u$ and $t = v$.

Now let Γ be a commutative group, let $F\colon X \to \Gamma$, and define $\tau\colon \mathscr{H} \to \Gamma$ by $\tau((u,v]) = F(v) - F(u)$ for all $u, v \in X$ with $u \le v$.

(c) Prove that τ is additive on $\mathscr{H}$. (Remember to check that τ is well-defined.)

For the remainder of this exercise, assume that $X = \mathbf{Q}$, the set of rational numbers, that Γ is $\mathbf{R}$ under addition, and that $F(x) = x$ for all $x \in X$.

(d) Prove that τ is not countably subadditive on $\mathscr{H}$.

(e) Prove that τ is not countably additive on $\mathscr{H}$.

(f) Prove that for each decreasing sequence (H_n) in $\mathscr{H}$, if $\bigcap_n H_n = \emptyset$, then $\tau(H_n) \to 0$ as $n \to \infty$.

---

[37] To say that X *is a totally ordered set* means that X is a partially ordered set in which any two elements are comparable. If X is a partially ordered set and $u, v \in X$, then to say that u *and* v *are comparable* means that $u \le v$ or $v \le u$.

About Measurability.

If X is a set and μ is the outer measure on X generated by a function $\tau\colon \mathscr{H} \to [0,\infty]$, where $\mathscr{H}$ is a set of subsets of X, we would like to identify a collection $\mathscr{E}$ of subsets of X such that $\mathscr{H} \subseteq \mathscr{E}$, μ is countably additive on $\mathscr{E}$, and $\mathscr{E}$ is as large as possible subject to the condition that it have reasonable properties as a set of sets. The next exercise shows that in a quite satisfying sense, $\mathscr{M}_\mu$ is a good choice for $\mathscr{E}$.

X259. Let $\mathscr{H}$ be a set of subsets of a set X, let $\tau\colon \mathscr{H} \to [0,\infty]$, and let μ be the outer measure on X generated by τ. Let $\mathscr{A}$ be a set of subsets of X. Since μ is countably subadditive on $\mathscr{P}(X)$, μ will be countably additive on $\mathscr{A}$ iff μ is superadditive on $\mathscr{A}$. Suppose $\mathscr{A}$ is a pre-ring on X such that $\mathscr{H} \subseteq \mathscr{A}$ and μ is countably additive on $\mathscr{A}$. Prove that $\mathscr{A} \subseteq \mathscr{M}_\mu$. Thus $\mathscr{M}_\mu$ is the largest pre-ring on X containing $\mathscr{H}$ on which μ is countably additive.

X260. Let X be a set and let $\mu\colon \mathscr{P}(X) \to [0,\infty]$. For each $T \subseteq X$, define $\mu_T\colon \mathscr{P}(X) \to [0,\infty]$ by

$$\mu_T(E) = \mu(E \cap T)$$

for all $E \subseteq X$.

(a) Let $E, S, T \subseteq X$ and let $\nu = \mu_T$. Prove that E ν-splits S iff E μ-splits $S \cap T$.

(b) Let $T \subseteq X$ and let $\nu = \mu_T$. Prove that $\mathscr{M}_\mu \subseteq \mathscr{M}_\nu$.

(c) Suppose μ is subadditive. Let $N = \{\, \mu_T : T \subseteq X \text{ and } \mu(T) < \infty \,\}$. Prove that

$$\mathscr{M}_\mu = \bigcap \{\, \mathscr{M}_\nu : \nu \in N \,\}.$$

X261. Let X be a set and let $\mu\colon \mathscr{P}(X) \to [0,\infty]$. *due 3Th*

(a) Suppose μ is subadditive, $E \subseteq X$, and $\mu(E) + \mu(X \setminus E) = \mu(X) < \infty$. Let $F \in \mathscr{M}_\mu$. Prove that E μ-splits F. (Suggestion: Draw a Venn diagram.)

(b) Suppose μ is subadditive, $E \subseteq H \in \mathscr{M}_\mu$, and $\mu(E) + \mu(H \setminus E) = \mu(H) < \infty$. Let $F \in \mathscr{M}_\mu$. Prove that E μ-splits F. (Hint: For part of the proof, let $\nu = \mu_H$, apply part (a) with μ replaced by ν, and apply parts (a) and (b) of problem X260.)

(c) Let $\mathscr{H}$ be a pre-ring on X and let $\tau\colon \mathscr{H} \to [0,\infty]$ be countably additive. Suppose μ is the outer measure on X generated by τ. Let $H \in \mathscr{H}$ with $\tau(H) < \infty$ and let $E \subseteq H$. Prove that $E \in \mathscr{M}_\mu$ iff $\mu(E) + \mu(H \setminus E) = \tau(H)$.

Remark. Intuitively, in problem X261(c), $\mu(E)$ is an overestimate for what $\tau(E)$ ought to be and $\mu(H \setminus E)$ is an overestimate for what $\tau(H \setminus E)$ ought to be. (We say "ought to be" because E and $H \setminus E$ need not belong to the domain of τ, namely $\mathscr{H}$, and because our goal is to extend τ in as natural a way as possible to a countably additive function on a collection of sets larger than the starting collection $\mathscr{H}$.) If $\mu(E) + \mu(H \setminus E) = \tau(H) < \infty$, then intuitively, these two "overestimates" are not really overestimates after all, but are actually exactly right.

Outer Content.

Paralleling the theory of measure is the theory of content. The theory of measure came later than the theory of content and is superior to it. In this sense, the theory of content is obsolete. However, by a little study of the theory of content, and by contrasting it with the theory of measure, you can improve your understanding of the theory of measure. For now, let's compare and contrast outer measure and outer content. Soon we'll also introduce, and compare and contrast, inner measure and inner content.

Let X be a set, let $\tau\colon \mathscr{H} \subseteq \mathscr{P}(X) \to [0,\infty]$, and define $\gamma\colon \mathscr{P}(X) \to [0,\infty]$ by

$$\gamma(A) = \inf \left\{ \sum_n \tau(H_n) : A \subseteq \bigcup_n H_n \text{ and } (H_n) \in \mathscr{H}^{<\mathbf{N}} \right\}.$$

We call γ the *outer content generated by* τ. Note that the definition of outer content involves only finite coverings of A, whereas the definition of outer measure involves countable coverings of A. Clearly if μ is the outer measure on X generated by τ, then $\mu \le \gamma$.

X262. Let $\mathscr{H} = \{\, (a,b] : -\infty < a \le b < \infty \,\}$ and define $\tau\colon \mathscr{H} \to [0,\infty)$ by $\tau\big((a,b]\big) = b-a$ for $-\infty < a \le b < \infty$. Let γ be the outer content on $\mathbf{R}$ generated by τ and let μ be the outer measure on $\mathbf{R}$ generated by τ. (Of course, μ is Lebesgue outer measure on $\mathbf{R}$.) It is obvious from the definitions that $\mu \le \gamma$.

(a) Let $S = \{\, x \in [0,1] : x \text{ is rational} \,\}$. Prove that $\gamma(S) = 1$ and $\mu(S) = 0$.

(b) Let K be a compact subset of $\mathbf{R}$. Prove that $\gamma(K) = \mu(K)$.

In the next few exercises, you are asked to establish some properties of outer content which are analogous to properties of outer measure that we proved in class. You may find this to be a good way to review the proofs of those properties of outer measure.

X263. Let X be a set, let $\tau\colon \mathscr{H} \subseteq \mathscr{P}(X) \to [0,\infty]$, and let γ be the outer content generated by τ. Prove the following statements.

(a) $\gamma(H) \le \tau(H)$ for all $H \in \mathscr{H}$.

(b) γ is subadditive on $\mathscr{P}(X)$.

(c) γ is an extension of τ iff τ is subadditive on $\mathscr{H}$.

Let X be a set. To say that γ *is an outer content on* X means that $\gamma\colon \mathscr{P}(X) \to [0,\infty]$ and γ is subadditive on $\mathscr{P}(X)$. By problem X263, if $\tau\colon \mathscr{H} \subseteq \mathscr{P}(X) \to [0,\infty]$ and γ is the outer content on X generated by τ, then γ is indeed an outer content on X.

Let X be a set. Recall that for any $\gamma\colon \mathscr{P}(X) \to [0,\infty]$, if $\gamma(\emptyset) = 0$, then $\mathscr{M}_\gamma$ is a field of sets on X and γ is additive on $\mathscr{M}_\gamma$. In particular, if γ is an outer content on X, then $\mathscr{M}_\gamma$ is a field of sets on X and γ is additive on $\mathscr{M}_\gamma$. Also, if γ is an outer content on X, and $\mathscr{N}_\gamma = \{ A \subseteq X : \gamma(A) = 0 \}$, then, with the help of the finite subadditivity of γ, it is easy to see that $\mathscr{N}_\gamma \subseteq \mathscr{M}_\gamma$.

X264. Let X be a set, let $\tau\colon \mathscr{H} \subseteq \mathscr{P}(X) \to [0,\infty]$, and let γ be the outer content on X generated by τ. Let $E \subseteq X$. Prove that the following are equivalent.

(a) $E \in \mathscr{M}_\gamma$.

(b) For each $H \in \mathscr{H}$, E γ-splits H.

(c) For each $H \in \mathscr{H}$, $\gamma(E \cap H) + \gamma(E^c \cap H) \le \tau(H)$.

X265. Let X be a set, let $\mathscr{H}$ be a pre-ring on X, let $\tau\colon \mathscr{H} \to [0,\infty]$ be superadditive, and let γ be the outer content on X generated by τ. Prove that $\mathscr{H} \subseteq \mathscr{M}_\gamma$.

This completes the group of exercises on the properties of outer content that are analogous to the properties of outer measure that we proved in class. The next result is analogous to problem X259.

X266. Let $\mathscr{H}$ be a set of subsets of a set X, let $\tau\colon \mathscr{H} \to [0,\infty]$, and let γ be the outer content on X generated by τ. Let $\mathscr{A}$ be a set of subsets of X. Since γ is subadditive on $\mathscr{P}(X)$, γ will be additive on $\mathscr{A}$ iff γ is superadditive on $\mathscr{A}$. Suppose $\mathscr{A}$ is a pre-ring on X such that $\mathscr{H} \subseteq \mathscr{A}$ and γ is additive on $\mathscr{A}$. Prove that $\mathscr{A} \subseteq \mathscr{M}_\gamma$. Thus $\mathscr{M}_\gamma$ is the largest pre-ring on X containing $\mathscr{H}$ on which γ is additive.

Note that parts (a) and (b) of problem X261 apply to any function $\mu\colon \mathscr{P}(X) \to [0,\infty]$, so in particular they apply if μ is an outer content. Part (c) of problem X261 applies to outer measures. Now here is the analog of that for outer contents.

X267. Let X be a set, let $\mathscr{H}$ be a pre-ring on X, and let $\tau\colon \mathscr{H} \to [0,\infty]$ be additive. Suppose μ is the outer content on X generated by τ. Let $H \in \mathscr{H}$ with $\tau(H) < \infty$ and let $E \subseteq H$. Prove that $E \in \mathscr{M}_\mu$ iff $\mu(E) + \mu(H \setminus E) = \tau(H)$.

Adapted Outer Measures and Adapted Outer Contents.

Let X be a set and let $\mu\colon \mathscr{P}(X) \to [0,\infty]$. To say that μ *is adapted*[38] means that $\mu(\emptyset) = 0$ and for each $A \subseteq X$, $\mu(A) = \inf\{ \mu(B) : A \subseteq B \in \mathscr{M}_\mu \}$.

X268. Let $\mathscr{H}$ be a pre-ring on a set X, let $\tau\colon \mathscr{H} \to [0,\infty]$ be superadditive, let γ be the outer content on X generated by τ, and let μ be the outer measure on X generated by τ. Prove that γ and μ are adapted.

Let X be a set and let $\mu\colon \mathscr{P}(X) \to [0,\infty]$. Let us write $\mu^\bullet$ for the function from $\mathscr{P}(X) \to [0,\infty]$ defined by

$$\mu^\bullet(A) = \inf\{ \mu(B) : A \subseteq B \in \mathscr{M}_\mu \}$$

for all $A \subseteq X$. Clearly μ is adapted iff $\mu(\emptyset) = 0$ and $\mu = \mu^\bullet$.

[38] This is not a standard use of the term *adapted*, so if you employ it outside this course, you should explain what you mean by it. Some authors do use the word *regular* in this context to mean the same thing that we mean by *adapted* here. However, the adjective *regular* is probably the most overworked word in mathematics. Even in measure theory, it is used in several different ways. You should always check its meaning when you see it used, and you should always explain what you mean by it when you use it yourself.

X269. Let X be a set, let $\mu\colon \mathscr{P}(X) \to [0,\infty]$ with $\mu(\varnothing) = 0$, and let τ be the restriction of μ to $\mathscr{M}_\mu$. As we know, $\mathscr{M}_\mu$ is a field of sets on X and τ is additive.

(a) Prove that $\mu^\bullet$ is the outer content on X generated by τ.

(b) Deduce that in particular, if μ is adapted, then μ is subadditive on $\mathscr{P}(X)$.

(c) Prove that $\mathscr{M}_\mu \subseteq \mathscr{M}_{\mu^\bullet}$.

(d) Suppose in addition that μ is subadditive and $\mu(X) < \infty$. Prove that $\mathscr{M}_\mu = \mathscr{M}_{\mu^\bullet}$.

X270. Let X be a set and let $\mu\colon \mathscr{P}(X) \to [0,\infty]$ be adapted. Let $E \subseteq X$ and suppose that E μ-splits each element of $\mathscr{M}_\mu$. Prove that $E \in \mathscr{M}_\mu$.

X271. Let X be a set, let μ be an outer measure on X, and let τ be the restriction of μ to $\mathscr{M}_\mu$. Prove that μ is adapted iff μ is the outer measure on X generated by τ.

X272. Let X be a set and let μ be an adapted outer measure on X. Let (A_n) be an increasing sequence of subsets of X and let $A = \bigcup_{n=1}^{\infty} A_n$. Prove that $\mu(A_n) \uparrow \mu(A)$ as $n \to \infty$. (Neither A nor the A_n's need be measurable.) *due 4M*

Inner Measure and Inner Content.

Let X be a set and let $\mu\colon \mathscr{P}(X) \to [0,\infty]$. Let us write $\mu_\bullet$ for the function from $\mathscr{P}(X)$ to $[0,\infty]$ defined by

$$\mu_\bullet(A) = \sup\{\, \mu(B) : B \in \mathscr{M}_\mu \text{ and } B \subseteq A \,\}.$$

If μ is the outer measure (respectively, the outer content) generated by a function $\tau\colon \mathscr{H} \subseteq \mathscr{P}(X) \to [0,\infty]$, then let us call $\mu_\bullet$ the *inner measure* (respectively, the *inner content*) on X generated by τ.

X273. Let X be a set, let $\mathscr{H}$ be a pre-ring on X, let $\tau\colon \mathscr{H} \to [0,\infty]$ be additive, let γ be the outer content generated by τ, and let $\gamma_\bullet$ be the inner content generated by τ. Let $A \subseteq X$ such that $\gamma(A) < \infty$. Prove that $\gamma_\bullet(A)$ is the supremum of sums of the form $\sum_n \tau(H_n)$ where (H_n) varies over finite disjoint sequences of elements of $\mathscr{H}$ with $\bigcup_n H_n \subseteq A$.

Note that problem X273 does not have an analog for inner measure. This is not a defect of inner measure. Rather, the result of problem X273 is a defect of inner content. For instance, if A is the set of irrational numbers in $[0,1]$, then the Lebesgue inner measure of A is 1, whereas if (H_i) is any disjoint family of intervals which are contained in A, then $\sum_i \text{length}(H_i) = 0$ because each H_i is either empty or a singleton.

By the way, originally, the result of problem X273 was taken as the definition of inner content. We have chosen instead to define inner content in a way that permits a unified treatment of inner measure and inner content.

X274. Let X be a set and let $\mu\colon \mathscr{P}(X) \to [0,\infty]$ be adapted.

(a) Let $A \subseteq H$ where $H \in \mathscr{M}_\mu$. Prove that $\mu_\bullet(A) + \mu(H \setminus A) = \mu(H)$.

(b) Let $A \subseteq X$ with $\mu(A) < \infty$. Prove that $A \in \mathscr{M}_\mu$ iff $\mu(A) = \mu_\bullet(A)$.

X275. Let $\mathscr{H}$ be a pre-ring on a set X and let $\tau\colon \mathscr{H} \to [0,\infty]$. Suppose either that τ is countably additive and μ is the outer measure on X generated by τ, or that τ is additive and μ is the outer content on X generated by τ.

(a) Let $A \subseteq X$ with $\mu(A) < \infty$. Prove that $A \in \mathscr{M}_\mu$ iff $\mu_\bullet(A) = \mu(A)$.

(b) Let $B \subseteq X$. Prove that $B \in \mathscr{M}_\mu$ iff for each $A \in \mathscr{M}_\mu$ with $\mu(A) < \infty$, we have $B \cap A \in \mathscr{M}_\mu$.

Reminder. Let X be a set, let Γ be a non-empty collection of σ-fields on X, and let $\mathscr{F} = \bigcap \Gamma$. In other words, let

$$\mathscr{F} = \{\, F : F \in \mathscr{G} \text{ for each } \mathscr{G} \in \Gamma \,\}.$$

Then $\mathscr{F}$ is a σ-field on X.

Reminder. Let X be a set and let $\mathscr{H}$ be a set of subsets of X. Then there is a smallest σ-field $\mathscr{F}$ on X containing $\mathscr{H}$. In fact, $\mathscr{F} = \bigcap \Gamma$ where Γ is the set of all σ-fields $\mathscr{A}$ on X such that $\mathscr{H} \subseteq \mathscr{A}$. (Note that Γ is non-empty, because $\mathscr{P}(X) \in \Gamma$.) By the way, it is clear that there can be at most one such $\mathscr{F}$, because if $\mathscr{F}_1$ and $\mathscr{F}_2$ are two such "smallest" elements of Γ, then $\mathscr{F}_1 \subseteq \mathscr{F}_2$ and $\mathscr{F}_2 \subseteq \mathscr{F}_1$.) We call $\mathscr{F}$ the *σ-field on X generated by $\mathscr{H}$*. We denote $\mathscr{F}$ by $\sigma(\mathscr{H})$. (Warning: $\sigma(\mathscr{H})$ depends on X as well as on $\mathscr{H}$. If $\mathscr{H}$ is a set of subsets of X and X is a proper subset of X', then the σ-field on X' generated by $\mathscr{H}$ is different from the σ-field on X generated by $\mathscr{H}$. For instance, X' belongs to the former σ-field but not to the latter. Thus the notation $\sigma(\mathscr{H})$ is context-dependent. Normally this does not cause confusion.)

Reminder. Let X be a set and let $\mathscr{G}$ and $\mathscr{H}$ be sets of subsets of X. Then $\sigma(\mathscr{G}) \subseteq \sigma(\mathscr{H})$ iff $\mathscr{G} \subseteq \sigma(\mathscr{H})$. It follows that $\sigma(\mathscr{G}) = \sigma(\mathscr{H})$ iff $\mathscr{G} \subseteq \sigma(\mathscr{H})$ and $\mathscr{H} \subseteq \sigma(\mathscr{G})$.

Reminder. Let X be a topological space, let $\mathscr{G}$ be the set of open subsets of X, and let $\mathscr{B} = \sigma(\mathscr{G})$. Then $\mathscr{B}$ is called the *Borel σ-field on X*. To say that E *is a Borel subset of* X means that $E \in \mathscr{B}$.

Reminder. Let $\mathscr{B}$ be the Borel σ-field on $\mathbf{R}$. Then each interval belongs to $\mathscr{B}$. For instance, if a is a real number, then although $[a, \infty)$ is not open, it is a countable intersection of open sets, because it is equal to $\bigcap_{n=1}^{\infty} (a - n^{-1}, \infty)$, so it is a Borel set.

X276. Let $\mathscr{B}$ be the Borel σ-field on $\mathbf{R}$.

- **(a)** Let $\mathscr{H} = \{\, (a, \infty) : a \in \mathbf{R} \,\}$. Prove that $\sigma(\mathscr{H}) = \mathscr{B}$. (Hint: Apply one of the reminders above.)
- **(b)** Let $\mathscr{I}$ be a set of intervals. Suppose that for each $a \in \mathbf{R}$, there exists a sequence (I_n) in $\mathscr{I}$ such that $(a, \infty) = \bigcup_{n=1}^{\infty} I_n$. Prove that $\sigma(\mathscr{I}) = \mathscr{B}$. (Similarly, if for each $b \in \mathbf{R}$, there exists a sequence (I_n) in $\mathscr{I}$ such that $(-\infty, b) = \bigcup_{n=1}^{\infty} I_n$, then $\sigma(\mathscr{I}) = \mathscr{B}$, although you are not asked to prove this.)

Example. Let X be a dense subset of $\mathbf{R}$ and let $\mathscr{B}$ be the Borel σ-field on $\mathbf{R}$. If

$$\mathscr{I} = \{\, (a, b] : a, b \in X \text{ and } a < b \,\},$$

then by problem X276(b), $\sigma(\mathscr{I}) = \mathscr{B}$. Similarly, if

$$\mathscr{I} = \{\, [a, a + 2006) : a \in X \,\},$$

then $\sigma(\mathscr{I}) = \mathscr{B}$.

Reminder. Let $\mathscr{B}$ be the Borel σ-field on $\mathbf{R}$. Let $F \colon \mathbf{R} \to \mathbf{R}$ be increasing and right-continuous. Let μ be the Lebesgue-Stieltjes outer measure on $\mathbf{R}$ with distribution function F. Then $\mathscr{B} \subseteq \mathscr{M}_\mu$. In particular, if we take $F(x) \equiv x$, we see that each Borel subset of $\mathbf{R}$ is Lebesgue measurable.

Remark. Let $\mathscr{M}$ be the σ-field of Lebesgue measurable subsets of $\mathbf{R}$ and let $\mathscr{B}$ be the Borel σ-field on $\mathbf{R}$. As we have just observed, $\mathscr{B} \subseteq \mathscr{M}$. So it is natural to ask whether $\mathscr{B} = \mathscr{M}$. The answer is no. In fact, it can be shown that the cardinality of $\mathscr{B}$ is the same as the cardinality of $\mathbf{R}$ and it is easy to show that the cardinality of $\mathscr{M}$ is the same as the cardinality of the power set of $\mathbf{R}$, so that the cardinality of $\mathscr{B}$ is strictly less than the cardinality of $\mathscr{M}$.

Now here is why the cardinality of $\mathscr{M}$ is the same as the cardinality of the power set of $\mathbf{R}$. Let C be the ordinary Cantor set (not a fat Cantor set). Then C has Lebesgue measure zero. Hence each subset of C is Lebesgue measurable. In other words, $\mathscr{P}(C) \subseteq \mathscr{M}$. Of course, $\mathscr{M} \subseteq \mathscr{P}(\mathbf{R})$. So to show that $\mathscr{M}$ has the same cardinality as $\mathscr{P}(\mathbf{R})$, it suffices[39] to show that $\mathscr{P}(C)$ has the same cardinality as $\mathscr{P}(\mathbf{R})$. But this follows immediately from the fact that the cardinality of C is the same as the cardinality of $\mathbf{R}$.

---

[39] by the Schroeder-Bernstein theorem, which states that if each of two given sets has the same cardinality as a subset of the other, then the two given sets have equal cardinality.

Remark. Let $\mathscr{M}$ be the σ-field of Lebesgue measurable subsets of $\mathbf{R}$. Since the cardinality of $\mathscr{M}$ is the same as the cardinality of the set of all subsets of $\mathbf{R}$, it is natural to ask whether $\mathscr{M}$ is actually equal to the set of all subsets of $\mathbf{R}$. Using the axiom of choice, one can show that the answer is no. There are subsets of $\mathbf{R}$ that are not Lebesgue measurable. We now turn to a proof of this.

Notation and Terminology. For each $A \subseteq \mathbf{R}$ and each $x \in \mathbf{R}$, let us write $A + x$ for $\{ a + x : a \in A \}$. We call $A + x$ the *translate of* A *by* x. Let μ be an outer measure on $\mathbf{R}$. To say that μ *is translation-invariant* means that for each $A \subseteq \mathbf{R}$ and each $x \in \mathbf{R}$, $\mu(A+x) = \mu(A)$. Clearly Lebesgue outer measure is translation-invariant. The next exercise tells in particular that not every subset of $\mathbf{R}$ is Lebesgue measurable.

X277. Let μ be a translation-invariant outer measure on $\mathbf{R}$. Suppose in addition that $\mu((0,1]) = 1$. Prove that there exists $E \subseteq \mathbf{R}$ such that $E \notin \mathscr{M}_\mu$. (Hint: If $x, y \in \mathbf{R}$ and $x - y$ is rational, then $\mathbf{Q} + x = \mathbf{Q} + y$. If $x, y \in \mathbf{R}$ and $x - y$ is irrational, then $\mathbf{Q} + x$ and $\mathbf{Q} + y$ are disjoint. Let $\Pi = \{ (\mathbf{Q} + x) \cap (0,1] : x \in \mathbf{R} \}$. Then Π is a set of non-empty subsets of $\mathbf{R}$. By the axiom of choice, there exists a family $(x_P)_{P \in \Pi}$ of real numbers such that for each $P \in \Pi$, $x_P \in P$. Let $E = \{ x_P : P \in \Pi \}$. Check that $\mathbf{R} = \bigcup \{ E + r : r \in \mathbf{Q} \}$. Deduce that $\mu(E) > 0$. Check that if $r, s \in \mathbf{Q}$ with $r \neq s$, then $E + r$ and $E + s$ are disjoint. Deduce that $(0,2]$ contains infinitely many disjoint translates of E. From this, deduce that if E were μ-measurable, then $\mu(E)$ would have to be zero.)

Remark. One might have hoped that the problem of defining the length of an arbitrary subset of $\mathbf{R}$ could be solved by producing a function $\mu \colon \mathscr{P}(\mathbf{R}) \to [0, \infty]$ such that μ is countably additive, μ is translation-invariant, and for each bounded interval I, $\mu(I)$ is the length of I. But from problem X277, it follows that no such function exists.

Remark. The construction in problem X277 is due to Vitaly (1905) and was one of the early, striking applications of the full axiom of choice. The first application of the full axiom of choice was Zermelo's theorem (1904) that every set can be well-ordered. The full axiom of choice was introduced by Zermelo for the express purpose of proving this well-ordering theorem.

Remark. Let $\mathscr{R}$ be the ring of all bounded subsets of $\mathbf{R}^d$. It can be shown that there does exist an additive (not countably additive), translation-invariant function $\mu \colon \mathscr{R} \to [0, \infty)$ such that $\mu((0,1]^d) = 1$. When $d \geq 2$, it is natural to ask whether we can also require that μ be invariant under rotations. It can be shown that this is possible when $d = 2$. But when $d \geq 3$, no such function that is also invariant under rotations exists. The reason is that Banach and Tarski (1924), building on work of Hausdorff (1914), showed that if A and B are bounded subsets of $\mathbf{R}^3$ having non-empty interior, then it is possible to partition A and B into finitely many pieces $A_1, \ldots, A_n$ and $B_1, \ldots, B_n$ respectively, in such a way that for each $k \in \{ 1, \ldots, n \}$, there is a proper rigid motion of $\mathbf{R}^3$ that maps A_k onto B_k. In colloquial terms, A can be disassembled into finitely many pieces which can be reassembled to form B. For instance, a ball of radius one can be cut up into finitely many pieces which can be reassembled to form two disjoint balls of radius one. This stunning result is known as the *Banach-Tarski paradox.* The reason why it is not contradictory is that the pieces are so bizarre that it is impossible to define their volume. The axiom of choice is used in the proof of the Banach-Tarski paradox. It is interesting to note that the way it is used is to show that one may choose one point from each coset of a suitable subgroup of the group of rotations of $\mathbf{R}^3$, just as Vitaly's construction of a subset of $\mathbf{R}$ that is not Lebesgue-measurable depended on choosing one point from each coset of $\mathbf{Q}$, considered as a subgroup of the group $(\mathbf{R}, +)$.

Some Counterexamples to Uniqueness of Measures.

Reminder. A *measurable space* is an ordered pair $(X, \mathscr{A})$ such that X is a set and $\mathscr{A}$ is a σ-field on X.

Reminder. Let $(X, \mathscr{A})$ be a measurable space. To say that μ *is a measure on* $\mathscr{A}$ means that $\mu \colon \mathscr{A} \to [0, \infty]$ and μ is countably additive on $\mathscr{A}$.

Reminder. A *measure space* is an ordered triple $(X, \mathscr{A}, \mu)$ such that X is a set, $\mathscr{A}$ is a σ-field on X, and μ is a measure on $\mathscr{A}$. (In French, this is called "une espace mesurée," which may be literally translated as "a measured space." Thus, from the French point of view, when we put a measure μ on a measurable space $(X, \mathscr{A})$, then we have "measured" it and it becomes a "measured space." Unfortunately, in English the terminology is slightly less logical.)

Reminder. Let $\mathscr{H}$ be a pre-ring on a set X, let $\tau\colon \mathscr{H} \to [0,\infty]$ be countably additive, let τ^* be the outer measure on X generated by τ, let $\mathscr{A}$ be the σ-field on X generated by $\mathscr{H}$, and let μ be the restriction of τ^* to $\mathscr{A}$. Suppose X can be covered by a sequence (H_k) of elements of $\mathscr{H}$ with $\tau(H_k) < \infty$ for each k. Then μ is the unique extension of τ to a measure on $\mathscr{A}$.

Definition. Let $(X, \mathscr{A})$ be a measurable space. To say that μ *is a probability measure on* $\mathscr{A}$ means that μ is a measure on $\mathscr{A}$ and $\mu(X) = 1$.

X278.

- **(a)** Give an example of a measurable space $(X, \mathscr{A})$, a subset $\mathscr{H} \subseteq \mathscr{A}$ with $\mathscr{A} = \sigma(\mathscr{H})$, and probability measures μ and ν on $\mathscr{A}$, such that $\mu(H) = \nu(H)$ for all $H \in \mathscr{H}$ but $\mu \neq \nu$. (Hint: There is an example where X is a small finite set.)
- **(b)** Explain why the example you found in part (a) does not conflict with what we learned in class and recalled above about uniqueness of measures.

X279. Let X be an uncountable set, let $\mathscr{H} = \big\{ \{x\} : x \in X \big\} \cup \{\emptyset\}$, let $\mathscr{A}$ be the σ-field on X generated by $\mathscr{H}$, and let τ be the function on $\mathscr{H}$ which is identically zero. Clearly $\mathscr{H}$ is a pre-ring on X and τ is countably additive on $\mathscr{H}$.

- **(a)** Explain exactly which sets belong to $\mathscr{A}$.
- **(b)** Let $\alpha \in [0,\infty]$. Prove that there is a unique measure μ_α on $\mathscr{A}$ such that μ_α is an extension of τ and $\mu_\alpha(X) = \alpha$.
- **(c)** If $\alpha, \beta \in [0,\infty]$ with $\alpha \neq \beta$, then μ_α and μ_β are different measures on $\mathscr{A}$ which agree on $\mathscr{H}$. Explain why this does not conflict with what we learned in class and recalled above about uniqueness of measures.
- **(d)** Let μ be the restriction to $\mathscr{A}$ of the outer measure on X generated by τ. Find $\alpha \in [0,\infty]$ such that $\mu = \mu_\alpha$. Justify your answer.

Definition. Let $(X, \mathscr{A}, \mu)$ be a measure space. To say that $(X, \mathscr{A}, \mu)$ *is σ-finite*, or that μ *is σ-finite*, means that there exists a sequence (A_k) of elements of $\mathscr{A}$ such that $X = \bigcup_{k=1}^{\infty} A_k$ and for each k, $\mu(A_k) < \infty$.

X280. Let $\mathscr{H} = \{ (a,b] : -\infty < a \leq b < \infty \}$ and let $\mathscr{A}$ be the σ-field on $\mathbf{R}$ generated by $\mathscr{H}$. As we know, $\mathscr{H}$ is a pre-ring on $\mathbf{R}$ and $\mathscr{A}$ is the Borel σ-field on $\mathbf{R}$. Define $\mu\colon \mathscr{A} \to [0,\infty]$ by $\mu(\emptyset) = 0$ and $\mu(A) = \infty$ if $A \neq \emptyset$. Let τ be the restriction of μ to $\mathscr{H}$. It is easy to see that μ is a measure on $\mathscr{A}$ which is not σ-finite, and obviously μ is an extension of τ.

- **(a)** Define $\nu\colon \mathscr{A} \to [0,\infty]$ by letting $\nu(A)$ be the number of elements in $A \cap \mathbf{Q}$, if this is finite, and by letting $\nu(A) = \infty$ otherwise. Prove that ν is a σ-finite measure on $\mathscr{A}$ and that ν is an extension of τ.
- **(b)** The measures μ and ν are different extensions of τ. Explain why this does not conflict with what we learned in class and recalled above about uniqueness of measures.

More About Measurability.

We have seen that there are far more Lebesgue measurable subsets of $\mathbf{R}$ than Borel subsets of $\mathbf{R}$. However, it turns out that given any Lebesgue measurable set $E \subseteq \mathbf{R}$, there exist Borel sets $A, B \subseteq \mathbf{R}$ such that $A \subseteq E \subseteq B$ and $B \setminus A$ has Lebesgue measure zero. (This follows from the next exercise.) Thus each Lebesgue measurable subset of $\mathbf{R}$ differs from some Borel subset of $\mathbf{R}$ only by a set of Lebesgue measure zero.

X281. Let X be a set, let $\mathscr{H}$ be a pre-ring on X, let $\tau\colon \mathscr{H} \to [0,\infty]$ be countably additive, and let μ *due 4Th* be the outer measure on X generated by τ. Then we know that μ is an extension of τ, that $\mathscr{H} \subseteq \mathscr{M}_\mu$, that μ is countably additive on $\mathscr{M}_\mu$, and that $\mathscr{M}_\mu$ is a σ-field on X. Let $\mathscr{B} = \sigma(\mathscr{H})$. Since $\mathscr{H} \subseteq \mathscr{M}_\mu$ and $\mathscr{M}_\mu$ is a σ-field on X, $\mathscr{B} \subseteq \mathscr{M}_\mu$.

- **(a)** Let $E \subseteq X$. Prove that there exists $C \in \mathscr{B}$ such that $E \subseteq C$ and $\mu(E) = \mu(C)$. (Comment: Of course, this is only interesting when $\mu(E) < \infty$. When $\mu(E) = \infty$, we can just take $C = X$.)
- **(b)** Let $E \subseteq C \in \mathscr{M}_\mu$ with $\mu(E) = \mu(C) < \infty$. Prove that $E \in \mathscr{M}_\mu$ iff $\mu(C \setminus E) = 0$. (Hint: Recall that since μ is subadditive, $\mathscr{N}_\mu \subseteq \mathscr{M}_\mu$, where $\mathscr{N}_\mu$ is the set of all $N \subseteq X$ such that $\mu(N) = 0$.)
- **(c)** Suppose in addition that μ is σ-finite. Let $E \subseteq X$. Prove that $E \in \mathscr{M}_\mu$ iff there exist $A, B \in \mathscr{B}$ such that $A \subseteq E \subseteq B$ and $\mu(B \setminus A) = 0$. (Hint: The reverse implication is essentially trivial. To prove the forward implication, first consider the case where $\mu(E) < \infty$.)

Approximating Measurable Sets by Simpler Sets.

Recall that if A and B are sets, then *the symmetric difference of A and B* is

$$A \,\Delta\, B = (A \setminus B) \cup (B \setminus A).$$

It is worth noticing that $A \,\Delta\, B = \{\, 1_A \neq 1_B \,\}$, where 1_A is the function that is 1 on A and 0 elsewhere and 1_B is the function that is 1 on B and 0 elsewhere.

X282. Let X be a set and let $\mu\colon \mathscr{P}(X) \to [0,\infty]$ be subadditive. Let $A \subseteq X$ and suppose that for each $\varepsilon > 0$, there exists $E \in \mathscr{M}_\mu$ such that $\mu(A \,\Delta\, E) < \varepsilon$. Prove that $A \in \mathscr{M}_\mu$.

X283. Let X be a set, let $\mathscr{H}$ be a pre-ring on X, let $\tau\colon \mathscr{H} \to [0,\infty)$ be countably additive (note that τ does not assume the value ∞), and let μ be the outer measure on X generated by τ. Let $\mathscr{R}$ be the set of finite disjoint unions of elements of $\mathscr{H}$. (As we know, $\mathscr{R}$ is the smallest ring of sets on X containing $\mathscr{H}$.) Let $A \subseteq X$. Prove that the following are equivalent:

(a) $A \in \mathscr{M}_\mu$ and $\mu(A) < \infty$.

(b) For each $\varepsilon > 0$, there exists $E \in \mathscr{R}$ such that $\mu(A \,\Delta\, E) < \varepsilon$.

Example. Let $F\colon \mathbf{R} \to \mathbf{R}$ be increasing and right continuous. Let μ be the Lebesgue-Stieltjes outer measure on $\mathbf{R}$ with distribution function F. Let $\mathscr{R}$ be the set of finite disjoint unions of intervals of the form $(a,b]$ where $-\infty < a \le b < \infty$. Let $A \subseteq \mathbf{R}$. Then the following are equivalent:

(a) $A \in \mathscr{M}_\mu$ and $\mu(A) < \infty$.

(b) For each $\varepsilon > 0$, there exists $R \in \mathscr{R}$ such that $\mu(A \,\Delta\, R) < \varepsilon$.

To see this, apply problem X283 with $\mathscr{H} = \{\, (a,b] : -\infty < a \le b < \infty \,\}$ and $\tau((a,b]) = F(b) - F(a)$ for all $a, b \in \mathbf{R}$ with $a < b$.

X284. Let X be a set and let $\mu\colon \mathscr{P}(X) \to [0,\infty]$ be subadditive. Let $\mathscr{F} = \{\, A \subseteq X : \mu(A) < \infty \,\}$. Define $\rho\colon \mathscr{F} \times \mathscr{F} \to [0,\infty)$ by $\rho(A,B) = \mu(A \,\Delta\, B)$. Prove that ρ is a pseudometric on $\mathscr{F}$. (In other words, prove that ρ has all the properties of a metric on $\mathscr{F}$ except that if $\rho(A,B) = 0$, it need not follow that $A = B$.)

Remark. Let X, $\mathscr{H}$, τ, μ, and $\mathscr{R}$ be as in problem X283. Let $\mathscr{F} = \{\, A \subseteq X : \mu(A) < \infty \,\}$ and let $\mathscr{E} = \mathscr{F} \cap \mathscr{M}_\mu$. Let ρ be as in problem X284. In terms of the result of problem X284, problem X283 has the following interpretation: $\mathscr{E}$ is the closure of $\mathscr{R}$ with respect to the topology induced on $\mathscr{F}$ by the pseudometric ρ.

Subadditivity and Superadditivity Revisited.

With the help of the integral for non-negative simple functions, we can give an enlightening alternative proof that for a non-negative set function on a field of sets, additivity implies subadditivity and superadditivity.

X285. Let $\mathscr{A}$ be a field of subsets of a set X and let $\mu\colon \mathscr{A} \to [0,\infty]$ be additive.

(a) Let $A \in \mathscr{A}$ and let (B_k) be a finite sequence of elements of $\mathscr{A}$ such that $A \subseteq \bigcup_k B_k$. Use the theory of integrals of non-negative simple functions to prove that $\mu(A) \le \sum_k \mu(B_k)$. (Hint: Let $N = \sum_k 1_{B_k}$. Notice that for each $x \in X$, $N(x)$ is equal to the number of values of k for which $x \in B_k$. In particular, $1_A \le N$.)

(b) Let (A_k) be a finite disjoint sequence of elements of $\mathscr{A}$ and let $B \in \mathscr{A}$ such that $\bigcup_k A_k \subseteq B$. Use the theory of integrals of non-negative simple functions to prove that $\sum_k \mu(A_k) \le \mu(B)$.

The Inclusion-Exclusion Formula Revisited.

Recall from problem X250 that if X is a set, $\mathscr{R}$ is a ring of sets on X, V is a commutative group, $\mu\colon \mathscr{R} \to V$ is additive, and $A_1, \dots, A_n \in \mathscr{R}$, then the inclusion-exclusion formula states that

$$\mu\big(\textstyle\bigcup_{i=1}^n A_i\big) = \displaystyle\sum_i \mu(A_i) - \sum_{i<j} \mu(A_i A_j) + \sum_{i<j<k} \mu(A_i A_j A_k) - + \cdots + (-1)^{n+1} \mu(A_1 \cdots A_n),$$

where to save space, we have written $A_i A_j$ for $A_i \cap A_j$ and so on. By the remark which follows problem X250, the inclusion-exclusion formula may be written more succinctly as follows:

$$\mu\big(\textstyle\bigcup_{i=1}^n A_i\big) = \displaystyle\sum_{I \neq \varnothing} (-1)^{|I|+1} \, \mu\big(\bigcap_{i \in I} A_i\big), \tag{37}$$

where the index I of summation ranges over all non-empty subsets of the set $\{1, \ldots, n\}$ and where $|I|$ denotes the number of elements in I. My aim here is to lead you to a slick proof of the inclusion-exclusion formula in the form (37). As a small first step, note that it suffices to treat the case where $\mathscr{R}$ is a field of sets on X, because we are free to replace X by $X' = \bigcup_{k=1}^n A_k$, $\mathscr{R}$ by $\mathscr{A} = \{ R \in \mathscr{R} : R \subseteq X' \}$ (which is a field of sets on X'), and μ by the restriction of μ to $\mathscr{A}$.

In class, we have treated integrals of $[0,\infty)$-valued simple functions with respect to an additive $[0,\infty]$-valued function μ on a field $\mathscr{A}$. In the same way, we may treat integrals of integer-valued simple functions with respect to an additive function μ on a field $\mathscr{A}$ taking values in a commutative group, because just as we can multiply elements of $[0,\infty]$ by elements of $[0,\infty)$, we can multiply elements of a commutative group by integers. You may take this for granted in the next exercise.

X286. Let X be a set, let $\mathscr{A}$ be a field of sets on X, let V be a commutative group, and let $\mu\colon \mathscr{A} \to V$ be additive. Use the result of problem X245 to give a slick proof of the generalized inclusion-exclusion formula. (Hint: Given $n \in \mathbf{N}$ and sets $A_1, \ldots, A_n \in \mathscr{A}$, let $B = \bigcup_{i=1}^n A_i$. Verify that

$$1_B = 1 - 1_{B^c} = 1 - 1_{\bigcap_{i=1}^n A_i^c} = 1 - \prod_{i=1}^n 1_{A_i^c} = 1 - \prod_{i=1}^n \big(1 - 1_{A_i}\big).$$

Use the result of problem X245 to expand the last product. In the expression that you thereby obtain for 1_B, integrate over X with respect to μ. You should get the inclusion-exclusion formula in the form (37). (Notice that the functions you are integrating are all integer-valued, so it makes sense to integrate them with respect to μ.)

Remark. Perhaps this is a good place to mention a further generalization of the theory of integrals of simple functions. Let U, V, and W be commutative semigroups with 0. Suppose we are given an operation of "multiplication" $(u,v) \to uv$ from $U \times V$ to W, such that for all $u, u' \in U$ and all $v, v' \in V$, we have $(u+u')v = uv + u'v$, $u \cdot 0 = 0$, $u(v+v') = uv + uv'$, and $0 \cdot v = 0$. Let X be a set, let $\mathscr{A}$ be a field of subsets of X, and let $\mu\colon \mathscr{A} \to V$ be additive. Then we may develop a theory of integrals of U-valued $\mathscr{A}$-simple functions with respect to μ. If $\varphi\colon X \to U$ is $\mathscr{A}$-simple, then by definition, its integral with respect to μ is

$$I(\varphi) = \sum_u u\mu(\varphi = u).$$

The proofs of the basic properties of this integral that we gave in the special case where $U = [0,\infty)$ and $V = [0,\infty]$ all go through in this more general situation with essentially no change. We have already mentioned the case where $U = \mathbf{Z}$ under addition, V is a commutative group, and $W = V$. Here are some other cases:

- **(a)** $U = \boldsymbol{\omega}$, the set of non-negative integers under addition, V is any commutative semigroup with 0, and $W = V$.
- **(b)** $U = V = W = \mathbf{R}$ under addition.
- **(c)** U is a field, V is a vector space over U, and $W = V$. (Integral of a scalar-valued simple function with respect to a vector-valued μ.)
- **(d)** U is a vector space over a field V and $W = V$. (Integral of a vector-valued simple function with respect to a scalar-valued μ.)
- **(e)** U is an inner product space over $\mathbf{K}$, $V = U$, $W = \mathbf{K}$, and $uv = \langle u|v\rangle$, the inner product of u and v, for all $u, v \in U$.
- **(f)** $U = V = W = \mathbf{R}^3$ and $uv = u \times v$, the cross product of v and w, for all $u, v \in \mathbf{R}^3$.

More About Measurable Functions.

X287. Let $(X, \mathscr{A})$ be a measurable space. Let $f, g\colon X \to \overline{\mathbf{R}}$ be measurable. Prove that $\{ f < g \} \in \mathscr{A}$. (Hint: For each $x \in X$, we have $f(x) < g(x)$ iff there exists a rational number r such that $f(x) < r$ and $r < g(x)$.)

X288. Let $(X, \mathscr{A})$ be a measurable space. Let $f, g\colon X \to \overline{\mathbf{R}}$ be measurable. Prove that $\{ f = g \} \in \mathscr{A}$.

X289. Let $(X, \mathscr{A})$, $(Y, \mathscr{B})$, and $(Z, \mathscr{C})$ be measurable spaces. Let $f\colon X \to Y$ and $g\colon Y \to Z$ be measurable. Let $h = g \circ f$. Prove that $h\colon X \to Z$ is measurable.

X290. Let $(X, \mathscr{A})$ be a measurable space and let Y be a metric space. Let $f\colon X \to Y$. Prove that f is measurable iff for each continuous function $g\colon Y \to \mathbf{R}$, $g \circ f$ is measurable. (Hint for the reverse implication: For each non-empty set $C \subseteq Y$, the function $y \mapsto d(y, C) = \inf\{\, d(y, y') : y' \in C \,\}$ is continuous and if C is closed, then $d(y, C) = 0$ iff $y \in C$.)

X291. Let $(X, \mathscr{A})$ be a measurable space and let Y be a metric space. Let (f_n) be a sequence of measurable functions from X to Y. Let $f\colon X \to Y$ and suppose that for each $x \in X$, we have $f_n(x) \to f(x)$ as $n \to \infty$. Prove that f is measurable.

Pullbacks of σ-Fields.

X292. Let $(Y, \mathscr{B})$ be a measurable space, let X be a set, and let $g\colon X \to Y$. Let $\mathscr{A} = \{\, g^{-1}[B] : B \in \mathscr{B} \,\}$. *due 5Th*

- **(a)** Prove that $\mathscr{A}$ is the smallest σ-field on X which makes g measurable. (We call $\mathscr{A}$ the *σ-field on X generated by g (from $(Y, \mathscr{B})$)* and we denote $\mathscr{A}$ by $\sigma(g)$.)
- **(b)** Let $(W, \mathscr{E})$ be a measurable space and let $f\colon W \to X$. Prove that f is $\mathscr{E}/\mathscr{A}$-measurable iff $g \circ f$ is $\mathscr{E}/\mathscr{B}$-measurable.
- **(c)** Prove that $\mathscr{A}$ is the unique σ-field on X that has the property described in (b) for all measurable spaces $(W, \mathscr{E})$ and all maps $f\colon W \to X$.
- **(d)** Suppose $\mathscr{D}$ is a set of subsets of Y such that $\mathscr{B} = \sigma(\mathscr{D})$ on Y. Let $\mathscr{C} = \{\, g^{-1}[D] : D \in \mathscr{D} \,\}$. Prove that $\mathscr{A} = \sigma(\mathscr{C})$ on X.

Subspaces of Measurable Spaces.

Example. Let $(Y, \mathscr{B})$ be a measurable space. Let X be a subset of Y. (We do not assume that $X \in \mathscr{B}$.) Let $\mathscr{A} = \{\, B \cap X : B \in \mathscr{B} \,\}$. Notice that for each $T \subseteq Y$, we have $T \cap X = g^{-1}[T]$, where g is the inclusion map[40] from X to Y. Therefore from problem X292, it is clear that:

- **(a)** $\mathscr{A}$ is the smallest σ-field on X which makes the inclusion map from X to Y measurable. (We call $\mathscr{A}$ the *subspace σ-field that X inherits from $(Y, \mathscr{B})$*.)
- **(b)** If $(W, \mathscr{E})$ be a measurable space and $f\colon W \to X$, then f is $\mathscr{E}/\mathscr{A}$-measurable from W to X iff f is $\mathscr{E}/\mathscr{B}$-measurable from W to Y.
- **(c)** $\mathscr{A}$ is the unique σ-field on X that has the property described in (c) for all measurable spaces $(W, \mathscr{E})$ and all maps $f\colon W \to X$.
- **(d)** If $\mathscr{D}$ is a set of subsets of Y such that $\mathscr{B} = \sigma(\mathscr{D})$ on Y and if $\mathscr{C} = \{\, D \cap X : D \in \mathscr{D} \,\}$, then $\mathscr{A} = \sigma(\mathscr{C})$ on X.

Example. Let Y be a topological space and let $\mathscr{B}$ be the Borel σ-field on Y. Thus $\mathscr{B} = \sigma(\mathscr{G}(Y))$ on Y, where $\mathscr{G}(Y)$ is the collection of open subsets of Y. Let $X \subseteq Y$. Then the subspace topology that X inherits from Y is $\mathscr{G}(X) = \{\, G \cap X : G \in \mathscr{G}(Y) \,\}$. Let $\mathscr{A}$ be the subspace σ-field that X inherits from $(Y, \mathscr{B})$. Then by the preceding example, $\mathscr{A} = \sigma(\mathscr{G}(X))$ on X. In other words, $\mathscr{A}$ is equal to the Borel σ-field on X.

X293. Let $(Y, \mathscr{B})$ be a measurable space, let $X \subseteq Y$, and let $\mathscr{A}$ be the subspace σ-field that X inherits from $(Y, \mathscr{B})$. Let $(Z, \mathscr{C})$ be a measurable space and let $h\colon Y \to Z$ be $\mathscr{B}/\mathscr{C}$-measurable. Prove that the restriction of h to X is $\mathscr{A}/\mathscr{C}$-measurable.

X294. Let $(Y, \mathscr{B})$ be a measurable space and let $g\colon Y \to \overline{\mathbf{R}}$ be $\mathscr{B}$-measurable. Let $X = \{\, g \neq 0 \,\}$ and let $\mathscr{A}$ be the subspace σ-field that X inherits from $(Y, \mathscr{B})$. Prove that $1/g$ is $\mathscr{A}$-measurable from X to $\mathbf{R}$. (Take $1/\infty = 0 = 1/(-\infty)$.)

X295. Let $(Y, \mathscr{B})$ be a measurable space and let $f_n\colon Y \to \overline{\mathbf{R}}$ be $\mathscr{B}$-measurable for each $n \in \mathbf{N}$. Let

$$X = \Big\{\, x \in Y : \lim_{n\to\infty} f_n(x) \text{ exists in } \overline{\mathbf{R}} \,\Big\}.$$

- **(a)** Prove that $X \in \mathscr{B}$.
- **(b)** Define $f\colon X \to \overline{\mathbf{R}}$ by $f(x) = \lim_{n\to\infty} f_n(x)$ for all $x \in X$. Let $\mathscr{A}$ be the subspace σ-field that X inherits from $(Y, \mathscr{B})$. Prove that f is $\mathscr{A}$-measurable.

---

[40] Given sets X and Y with $X \subseteq Y$, to say that g *is the inclusion map from X to Y* means that g is the function from X to Y defined by $g(x) = x$ for all $x \in X$.

X296. Let $(Y, \mathscr{B})$ be a measurable space, let $X \subseteq Y$, let $\mathscr{A}$ be the subspace σ-field that X inherits from Y, and let $\mathscr{E} = \{ B \in \mathscr{B} : B \subseteq X \}$.

(a) Prove that $\mathscr{E} \subseteq \mathscr{A}$.

(b) Prove that $\mathscr{E} = \mathscr{A}$ iff $X \in \mathscr{E}$ iff $X \in \mathscr{B}$.

Cartesian Products of Measurable Spaces.

For simplicity, we shall consider just Cartesian products of two measurable spaces. By induction, our results extend to Cartesian products of finitely many measurable spaces. (It is also possible to consider Cartesian products of infinitely many measurable spaces.)

Notation. Let $(X, \mathscr{A})$ and $(Y, \mathscr{B})$ be measurable spaces and let $Z = X \times Y$. Recall that

$$\mathscr{A} \odot \mathscr{B} = \{ A \times B : A \in \mathscr{A} \text{ and } B \in \mathscr{B} \}.$$

We call the elements of $\mathscr{A} \odot \mathscr{B}$ *measurable rectangles*. Except in trivial cases, $\mathscr{A} \odot \mathscr{B}$ is not a σ-field on Z. The σ-field on Z generated by $\mathscr{A} \odot \mathscr{B}$ is called *the product of the σ-fields $\mathscr{A}$ and $\mathscr{B}$* and is denoted by $\mathscr{A} \otimes \mathscr{B}$.

X297. Let $(X, \mathscr{A})$ and $(Y, \mathscr{B})$ be measurable spaces. Let $Z = X \times Y$. Make Z into a measurable space by equipping it with the σ-field $\mathscr{C} = \mathscr{A} \otimes \mathscr{B}$.

(a) Let $\pi_1 : Z \to X$ and $\pi_2 : Z \to Y$ be the usual projection maps. Prove that π_1 and π_2 are measurable.

(b) Let $(W, \mathscr{E})$ be a measurable space, let $f : W \to Z$, and let $g = \pi_1 \circ f$ and $h = \pi_2 \circ f$, so that $f(w) = (g(w), h(w))$ for all $w \in W$. Prove that f is measurable iff g and h are measurable.

(c) Prove that the product σ-field on Z is the unique σ-field on Z having the property described in part (b) for all measurable spaces $(W, \mathscr{E})$ and all functions $f : W \to Z$.

X298. Let X and Y be topological spaces and let $Z = X \times Y$. Let $\mathscr{A}$ and $\mathscr{B}$ be the Borel σ-fields on X and Y respectively. Let $\mathscr{C}$ be the Borel σ-field on Z. *due 5Th*

(a) Prove that $\mathscr{A} \otimes \mathscr{B} \subseteq \mathscr{C}$. (By the way, this inclusion can be strict, although you are not asked to give an example where it is. Warning: No hand waving allowed! Hint: If $A \subseteq X$ and $B \subseteq Y$, then $A \times B = \pi_1^{-1}[A] \cap \pi_2^{-1}[B]$.)

(b) Suppose that X is second countable. Prove that $\mathscr{A} \otimes \mathscr{B} = \mathscr{C}$. (Hint: Let $\{ U_n : n \in \mathbf{N} \}$ be a countable base for the topology of X. Let G be open in Z. Prove that there exist open sets V_n in Y such that $G = \bigcup_{n=1}^{\infty} (U_n \times V_n)$. Remark: Similarly, if Y is second countable and X is arbitrary, then $\mathscr{A} \otimes \mathscr{B} = \mathscr{C}$, although you are not asked to prove this.)

Example. Let $\mathscr{B}$ be the Borel σ-field on $\mathbf{R}$ and let $\mathscr{C}$ be the Borel σ-field on $\mathbf{R} \times \mathbf{R}$. Then by problem X298(b), $\mathscr{C} = \mathscr{B} \otimes \mathscr{B}$, because $\mathbf{R}$ is a second countable topological space. Hence by problem X297(b), if $(W, \mathscr{E})$ is any measurable space, $f : W \to \mathbf{R} \times \mathbf{R}$, and $g = \pi_1 \circ f$ and $h = \pi_2 \circ f$, so that $f(w) = (g(w), h(w))$ for all $w \in W$, then f is measurable as a map into the topological space $\mathbf{R} \times \mathbf{R}$ iff g and h are measurable as maps into the topological space $\mathbf{R}$. To say the same thing in different words, if $f : W \to \mathbf{C}$, then f is measurable iff $\mathrm{Re}(f)$ and $\mathrm{Im}(f)$ are both measurable.

X299. Let X and Y be sets and let $Z = X \times Y$. Let $\mathscr{D}$ and $\mathscr{E}$ be sets of subsets of X and Y respectively. Let $\mathscr{A} = \sigma(\mathscr{D})$ on X and let $\mathscr{B} = \sigma(\mathscr{E})$ on Y.

(a) Let $\mathscr{F} = \{ D \times Y : D \in \mathscr{D} \} \cup \{ X \times E : E \in \mathscr{E} \}$. Prove that $\mathscr{A} \otimes \mathscr{B} = \sigma(\mathscr{F})$ on Z.

(b) Suppose that $X = \bigcup_{k=1}^{\infty} D_k$ for some $(D_k) \in \mathscr{D}^{\mathbf{N}}$ and $Y = \bigcup_{k=1}^{\infty} E_k$ for some $(E_k) \in \mathscr{D}^{\mathbf{N}}$. Prove that $\mathscr{A} \otimes \mathscr{B} = \sigma(\mathscr{D} \odot \mathscr{E})$ on Z.

Reminder. Let m^* be Lebesgue outer measure on $\mathbf{R}$. Let $\mathscr{M} = \mathscr{M}_{m^*}$. If $A \subseteq \mathbf{R}$, then to say that A *is Lebesgue measurable* means that $A \in \mathscr{M}$. Let m be the restriction of m^* to $\mathscr{M}$. The measure m is called *Lebesgue measure on* $\mathbf{R}$. Let $f\colon \mathbf{R} \to Y$, where Y is a topological space. To say that f *is Lebesgue measurable* means that f is measurable with respect to the σ-field $\mathscr{M}$. Now let $f\colon \mathbf{R} \to \mathbf{R}$. To say that f *is Lebesgue integrable* means that f is integrable with respect to the measure m. If f is Lebesgue integrable, then of course $\int f\,dm$ is called *the Lebesgue integral of* f.

The Riemann Integral versus the Lebesgue Integral.

We wish to characterize the functions f which are properly Riemann integrable and to show that for each such function f, the Lebesgue integral of f is the same as the Riemann integral of f. We may as well consider only functions $f\colon \mathbf{R} \to \mathbf{R}$, since if f is defined only on a subinterval of $\mathbf{R}$, we may extend it to all of $\mathbf{R}$ by defining it to be zero outside that subinterval. Now let us formulate a convenient definition of the Riemann integral. (It should be obvious that it is equivalent to the usual definition.) If Y is a set and $\varphi\colon \mathbf{R} \to Y$, then to say that φ *is a step function* means that $\mathbf{R}$ can be partitioned into finitely many intervals on each of which φ is constant. Just as for simple functions, any way of combining finitely many step functions yields another step function. For instance, if $\varphi_1, \varphi_2\colon \mathbf{R} \to \mathbf{R}$ are step functions, then so are $\varphi_1 + \varphi_2$ and $\varphi_1 \vee \varphi_2 = \max\{\,\varphi_1, \varphi_2\,\}$. Notice that if $\varphi\colon \mathbf{R} \to \mathbf{R}$ is a step function, then φ is Lebesgue integrable iff $\{\,\varphi \neq 0\,\}$ is bounded. Let us write Φ for the set of Lebesgue integrable step functions $\varphi\colon \mathbf{R} \to \mathbf{R}$. For each $\varphi \in \Phi$, let $I(\varphi) = \int \varphi\,dm$. (Of course if $\varphi \in \Phi$, then since $\mathbf{R}$ can be partitioned into finitely many intervals on each of which φ is constant, $I(\varphi)$ is really just a simple finite sum.) Now let $f\colon \mathbf{R} \to \mathbf{R}$. Then the *lower Riemann integral of* f is

$$R_*(f) = \sup\{\, I(\varphi) : \varphi \in \Phi \text{ and } \varphi \leq f \,\}$$

and the *upper Riemann integral of* f is

$$R^*(f) = \inf\{\, I(\varphi) : \varphi \in \Phi \text{ and } f \leq \varphi \,\}.$$

To say that f *is properly Riemann integrable* means that $-\infty < R_*(f) = R^*(f) < \infty$, and in this case, the (proper) Riemann integral of f is $R(f) = R_*(f) = R^*(f)$. Notice that if f is not bounded below or if $\{\, f < 0 \,\}$ is not bounded, then there is no $\varphi \in \Phi$ satisfying $\varphi \leq f$, so $R_*(f) = -\infty$. Similarly, if f is not bounded above or if $\{\, f > 0 \,\}$ is not bounded, then there is no $\varphi \in \Phi$ satisfying $f \leq \varphi$, so $R^*(f) = \infty$. Thus if f is properly Riemann integrable, then f must be a bounded function and $\{\, f \neq 0 \,\}$ must be a bounded set.

X300. Let $f\colon \mathbf{R} \to \mathbf{R}$ be bounded below and suppose $\{\, f < 0 \,\}$ is bounded. Let f_* be the lower regularization of f. The object of this exercise is to lead you to a proof that the lower Riemann integral of f is equal to the Lebesgue integral of f_*.

- **(a)** Prove that $\int f_*\,dm$ is defined and $\int f_*\,dm > -\infty$.
- **(b)** Prove that $R_*(f) \leq \int f_*\,dm$. (Hint: If $\varphi \in \Phi$ and $\varphi \leq f$ and if φ_* is the lower regularization of φ, then $\{\, \varphi_* < \varphi \,\}$ is a finite set and $\varphi_* \leq f_*$.)
- **(c)** Prove that $\int f_*\,dm \leq R_*(f)$. (Hint: If $x \in \mathbf{R}$, $y \in \mathbf{Q}$, and $y < f_*(x)$, then there exist $a, b \in \mathbf{Q}$ with $a < x < b$ such that for each $x' \in (a, b)$, we have $y < f_*(x')$. Deduce that there is a sequence (ψ_n) in Φ such that $f_* \leq \sup_n \psi_n \leq f$. Then let $\varphi_n = \max_{k \leq n} \psi_k$ to get an increasing sequence (φ_n) in Φ such that $f_* \leq \lim_n \varphi_n \leq f$.)

X301. Let $f\colon \mathbf{R} \to \mathbf{R}$ be bounded and suppose $\{\, f \neq 0 \,\}$ is bounded. Let f_* and f^* be the lower and upper regularizations of f respectively and let

$$D = \{\, x_0 \in \mathbf{R} : f \text{ is not continuous at } x_0 \,\}.$$

From problem X300, we see that $\int f_*\,dm$ is defined, $\int f_*\,dm > -\infty$, and $R_*(f) = \int f_*\,dm$. Similarly, $\int f^*\,dm$ is defined, $\int f^*\,dm < \infty$, and $R^*(f) = \int f^*\,dm$.

- **(a)** Prove that f is (properly) Riemann integrable iff $m(D) = 0$.
- **(b)** Suppose that f is (properly) Riemann integrable. Prove that f is Lebesgue measurable, that f is Lebesgue integrable, and that $R(f) = \int f\,dm$.

Riemann Null Functions versus Lebesgue Null Functions.

Definition. Let $f\colon \mathbf{R} \to \mathbf{R}$. To say that f *is Lebesgue null* means that $f = 0$ m-almost everywhere.

X302. Let $f\colon \mathbf{R} \to \mathbf{R}$. Prove that f is Riemann null iff f is Riemann integrable and Lebesgue null.

Reminder. Let $f = 1_{\mathbf{Q}}$. Then f is Lebesgue null. However, f is not Riemann null. Indeed, f is not even Riemann integrable.

X303. Let m be Lebesgue measure on $\mathbf{R}$. Let $f\colon \mathbf{R} \to \mathbf{C}$ be Lebesgue measurable.

- **(a)** Let C be the set of $p \in \mathbf{R}$ such that f is continuous at p. Prove that if $f = 0$ m-a.e., then for each $p \in C$, $f(p) = 0$. In particular, if f is continuous on $\mathbf{R}$ and $f = 0$ m-a.e., then for each $p \in \mathbf{R}$, $f(p) = 0$.
- **(b)** Suppose f is regulated. Prove that $f = 0$ m-a.e. iff $f(p) = 0$ for all but countably many $p \in \mathbf{R}$.
- **(c)** Give an example where $f = 0$ m-a.e., but $\{ f \neq 0 \}$ is uncountable.

Absolutely Convergent Improper Integrals.

As we have seen, a properly Riemann integrable function on $\mathbf{R}$ is Lebesgue integrable and its Riemann integral is equal to its Lebesgue integral. The next exercise treats the relation between absolutely convergent improper integrals over $[0,\infty)$ and Lebesgue integrals over $[0,\infty)$. Absolutely convergent improper integrals over other intervals may be treated similarly.

X304. Let $f\colon [0,\infty) \to \mathbf{R}$ be Lebesgue measurable. The integrals in this problem are to be understood as Lebesgue integrals.

- **(a)** Prove that $\lim_{b\to\infty} \int_0^b |f(x)|\,dx$ exists in $[0,\infty]$ and is equal to $\int_0^\infty |f(x)|\,dx$.
- **(b)** Prove that f is Lebesgue integrable over $[0,\infty)$ iff $\lim_{b\to\infty} \int_0^b |f(x)|\,dx < \infty$.
- **(c)** Prove that if f is Lebesgue integrable over $[0,\infty)$, then $\lim_{b\to\infty} \int_0^b f(x)\,dx$ exists and is equal to $\int_0^\infty f(x)\,dx$.

Warning. If $f\colon (a,b) \to \mathbf{R}$ is improperly Riemann-integrable over (a,b), but the improper Riemann integral of f over (a,b) is only conditionally convergent, then $\int_a^b f^+(x)\,dx = \infty = \int_a^b f^-(x)\,dx$ and consequently the Lebesgue integral of f over (a,b) is undefined. The Lebesgue theory of integration is a theory of absolutely convergent integrals. In the Lebesgue approach to integration, conditionally convergent integrals are viewed as limits of integrals, not as integrals per se.

Some Applications of the Integral Limit Theorems.

From now on, unless otherwise stated, we shall interpret integrals such as $\int_a^b f(x)\,dx$ as Lebesgue integrals, or in other words, as integrals with respect to Lebesgue measure.

X305. In problem X163, we saw that for each $x \in (0,\infty)$, we have

$$\lim_{n\to\infty} \int_0^n t^{x-1} \left(1 - \frac{t}{n}\right)^n \,dt = \Gamma(x) = \int_0^\infty t^{x-1} e^{-t}\,dt.$$

Use one of the Corollaries of Fatou's lemma to give a simpler proof of this fact. (The inequality $1 + u \leq e^u$ should be almost the only other ingredient that you will need for the proof.)

X306. Determine the following limits. Justify your answers.

- **(a)** $\lim_{n\to\infty} \int_0^1 (1 + nx^2)(1 + x^2)^{-n}\,dx$
- **(b)** $\lim_{n\to\infty} \int_0^\infty n \sin(x/n)[x(1 + x^2)]^{-1}\,dx$

X307.

- **(a)** Let $a \in (-1,\infty)$. Define $f\colon (0,\infty) \to (0,\infty)$ by $f(x) = \int_0^\infty t^a e^{-tx}\,dt$. Use an appropriate integral limit theorem, together with the definition of the derivative, to show that $f'(x) = -\int_0^\infty t^{a+1} e^{-tx}\,dt$.
- **(b)** Notice that $\int_0^\infty e^{-tx}\,dt = x^{-1}$. Use this, together with part (a), to show that $\int_0^\infty t^n e^{-t}\,dt = n!$.

A Nasty Lebesgue Integrable Function.

Reminder. Let X be a topological space and let $A \subseteq X$. To say that A *is nowhere dense* means that the closure of A has empty interior. To say that A *is meager* means that A is a countable union of nowhere dense sets.

For instance, any countable subset of $\mathbf{R}$ is meager in $\mathbf{R}$. So is any closed subset of $\mathbf{R}$ that has empty interior. For instance, the ordinary Cantor set is meager in $\mathbf{R}$ and so is any fat Cantor set. Thus the ordinary Cantor set, which is small in the sense of Lebesgue measure (we mean of Lebesgue measure zero) is also small in the sense of topology (we mean meager), whereas a fat Cantor set is not small in the sense of Lebesgue measure but is small in the sense of topology.

Meager sets are also known as sets *of the first category.* Sets that are not meager are also known as sets *of the second category.* To say that X is a *Baire space* means that no non-empty open subset of X is meager.

X308. Define $g\colon \mathbf{R} \to (0, \infty]$ by

$$g(x) = \begin{cases} |x|^{-1/2} e^{-|x|} & \text{if } x \in \mathbf{R} \setminus \{0\}, \\ \infty & \text{if } x = 0. \end{cases}$$

Let $r_1, r_2, r_3, \ldots$ be an enumeration of the rationals and define $f\colon \mathbf{R} \to (0, \infty]$ by $f(x) = \sum_{n=1}^{\infty} 2^{-n} g(x - r_n)$. Let m denote Lebesgue measure on $\mathbf{R}$.

- **(a)** Prove that f is Lebesgue integrable over $\mathbf{R}$. Hence $f < \infty$ m-a.e. on $\mathbf{R}$. (Thus $\{ f = \infty \}$ is small in the sense of Lebesgue measure.)
- **(b)** Prove that if $x_0 \in \mathbf{R}$ such that $f(x_0) < \infty$, then f is discontinuous at x_0. Thus f is discontinuous m-a.e. on $\mathbf{R}$.
- **(c)** Prove that f is lower semicontinuous.
- **(d)** Prove that if $x_0 \in \mathbf{R}$ such that $f(x_0) = \infty$, then f is continuous at x_0.
- **(e)** Prove that $\{ f < \infty \}$ is meager in $\mathbf{R}$. (Thus $\{ f < \infty \}$ is small in the sense of topology. You should contrast this with part (a).)
- **(f)** Prove that $\{ f = \infty \}$ is uncountable. Thus, although it is clear that $\{ f = \infty \}$ contains $\mathbf{Q}$, $\{ f = \infty \}$ is not equal to $\mathbf{Q}$. (In fact, $\{ f = \infty \}$ has the cardinality of the continuum, although you are not asked to prove this.)

Integration with Respect to Counting Measure.

X309. Let I be a set. Define $\gamma\colon \mathscr{P}(I) \to [0, \infty]$ by

$$\gamma(A) = \begin{cases} |A| & \text{if } A \text{ is a finite set}, \\ \infty & \text{otherwise}, \end{cases}$$

for all $A \subseteq I$, where $|A|$ denotes the number of elements in A.

- **(a)** Prove that γ is a measure on the σ-field $\mathscr{P}(I)$. (Reminder: γ is called *counting measure on I*.)
- **(b)** Let $f\colon I \to [0, \infty]$. Prove that $\int_I f\, d\gamma = \sum_{i \in I} f(i)$, where by definition,

$$\sum_{i \in I} f(i) = \sup \left\{ \sum_{i \in I_0} f(i) : I_0 \text{ is a finite subset of } I \right\}.$$

Notation. Let $(X, \mathscr{A}, \mu)$ be a measure space. Let f be a measurable function on X taking values in $\overline{\mathbf{R}}$ or $\mathbf{R}^d$ and let $A \in \mathscr{A}$. Then we write $\int_A f\,d\mu$ for $\int f 1_A\,d\mu$.

X310. Let $(X, \mathscr{A}, \mu)$ be a measure space, let $f\colon X \to \overline{\mathbf{R}}$, and suppose $\int f\,d\mu$ is defined. Prove that for each $A \in \mathscr{A}$, $\int_A f\,d\mu$ is defined.

Semifinite Measures.

Let $(X, \mathscr{A}, \mu)$ be a measure space. To say that μ *is semifinite* means that for each $A \in \mathscr{A}$, if $\mu(A) > 0$, then there exists $B \in \mathscr{A}$ such that $B \subseteq A$ and $0 < \mu(B) < \infty$.

X311. Let $(X, \mathscr{A}, \mu)$ be a measure space. Suppose that μ is σ-finite. Prove that μ is semifinite.

X312. Let X be a set and let γ be counting measure on X.

(a) Verify that γ is semifinite.

(b) Prove that γ is σ-finite iff X is countable. So for instance, if $X = \mathbf{R}$, then γ is semifinite but not σ-finite.

X313. Let $(X, \mathscr{A}, \mu)$ be a measure space. Let $f, g\colon X \to \overline{\mathbf{R}}$ such that $\int f\,d\mu$ and $\int g\,d\mu$ are both defined. Assume that μ is semifinite.

(a) Suppose that for each $A \in \mathscr{A}$, we have $\int_A f\,d\mu \le \int_A g\,d\mu$. Prove that $f \le g$ μ-almost everywhere. (Hint: $\{ f > g \}$ is the union of the sets of the form $\{ f > r \text{ and } s > g \}$ with r and s rational and $r > s$.)

(b) Suppose that for each $A \in \mathscr{A}$, we have $\int_A f\,d\mu = \int_A g\,d\mu$. Prove that $f = g$ μ-almost everywhere.

X314. Give an example of a measure space $(X, \mathscr{A}, \mu)$ and two measurable functions $f, g\colon X \to [0, \infty)$ such that $\int_A f\,d\mu = \int_A g\,d\mu$ for each $A \in \mathscr{A}$, but $\mu(f \ne g) > 0$. Of course μ cannot be semifinite. (Hint: There is an example where X has just one point.)

More Applications of the Integral Limit Theorems.

Let $f\colon \mathbf{R} \to \mathbf{C}$ be Lebesgue integrable. The *Fourier transform of* f is the function $\hat{f}\colon \mathbf{R} \to \mathbf{C}$ defined by

$$\hat{f}(y) = \int_{\mathbf{R}} f(t) e^{-2\pi i t y}\,dt$$

for all $y \in \mathbf{R}$. It is an important fact that if $f(t) = e^{-\pi t^2}$ for all $t \in \mathbf{R}$, then $\hat{f} = f$. The following exercise[41] outlines a particularly elegant proof of this.

X315. *due 7Th*

(a) Let $f\colon \mathbf{R} \to \mathbf{C}$ be Lebesgue measurable and suppose $\int_{\mathbf{R}} |f(t)| e^{xt}\,dt < \infty$ for all $x \in \mathbf{R}$. Let $z \in \mathbf{C}$. Justify the following calculation:

$$\int_{\mathbf{R}} f(t) e^{zt}\,dt = \int_{\mathbf{R}} f(t) \sum_{n=0}^{\infty} \frac{(zt)^n}{n!}\,dt = \sum_{n=0}^{\infty} c_n z^n,$$

where

$$c_n = \int_{\mathbf{R}} f(t) \frac{t^n}{n!}\,dt.$$

(b) Use the fact that $\int_{\mathbf{R}} e^{-\pi t^2}\,dt = 1$ (which you are not asked to prove here) to check that

$$\int_{\mathbf{R}} e^{-\pi t^2} e^{2\pi x t}\,dt = e^{\pi x^2}$$

[41] This exercise relies on a fact that we have not yet proved, namely that $\int_{\mathbf{R}} e^{-\pi t^2}\,dt = 1$. The most common way to prove this is to write the square of this integral as an integral over $\mathbf{R}^2$ and switch to polar coordinates to evaluate the double integral. Since we have not discussed double integrals, it is worth mentioning that another proof of this, in the form $\int_{\mathbf{R}} e^{-s^2}\,ds = \sqrt{\pi}$, can be found on page 194 in Rudin, *Principles of Mathematical Analysis*, Third Edition.

for each $x \in \mathbf{R}$. Then use part (a) and some complex analysis to explain why we are justified in replacing x by $-iy$ in the last equation to get

$$\int_{\mathbf{R}} e^{-\pi t^2} e^{-2\pi i y t}\,dt = e^{-\pi y^2}$$

for each $y \in \mathbf{R}$.

The next exercise generalizes part (a) of problem X315.

X316. Let $f\colon \mathbf{R} \to \mathbf{C}$ be Lebesgue measurable. Let $S = \{\, s \in \mathbf{R} : \int_{\mathbf{R}} e^{sx}|f(x)|\,dx < \infty \,\}$. Let $a = \inf S$ and $b = \sup S$. Assume that $a < b$. Let $u_0 = s_0 + it_0$, where $s_0, t_0 \in \mathbf{R}$ and $a < s_0 < b$ (and where of course $i = \sqrt{-1}$). Let $r = \min\{\, b - s_0,\ s_0 - a \,\}$ and let $D = \{\, u \in \mathbf{C} : |u - u_0| < r \,\}$. (Thus D is the largest open disk centered at u_0 and contained in the strip $\{\, u \in \mathbf{C} : a < \operatorname{Re} u < b \,\}$.) Prove that for all $u \in D$,

$$\int_{\mathbf{R}} e^{ux} f(x)\,dx = \sum_{n=0}^{\infty} c_n (u - u_0)^n,$$

where for each n,

$$c_n = \frac{1}{n!} \int_{\mathbf{R}} x^n e^{u_0 x} f(x)\,dx.$$

Don't forget to explain why the integrand in the expression for c_n is integrable.

Definitions. A *probability measure* on a measurable space $(X, \mathscr{A})$ is a measure μ on $\mathscr{A}$ such that $\mu(X) = 1$. A *Borel measure* on a topological space is simply a measure on the Borel σ-field on that space. If μ is a Borel probability measure on $\mathbf{R}$, then *the characteristic function of* μ is the function $\varphi\colon \mathbf{R} \to \mathbf{C}$ defined by $\varphi(t) = \int_{\mathbf{R}} e^{itx}\,d\mu(x)$ for all $t \in \mathbf{R}$.

X317. Let μ be a Borel measure on $\mathbf{R}$ and let $f\colon \mathbf{R} \to \mathbf{C}$ be μ-integrable. Define $g\colon \mathbf{R} \to \mathbf{C}$ by

$$g(t) = \int_{\mathbf{R}} e^{itx} f(x)\,d\mu(x)$$

for all $t \in \mathbf{R}$.

(a) Prove that g is uniformly continuous on $\mathbf{R}$.

(b) Suppose in addition that $\int_{\mathbf{R}} |xf(x)|\,d\mu(x) < \infty$. Prove that g is continuously differentiable and that for each $t \in \mathbf{R}$,

$$g'(t) = \int_{\mathbf{R}} ixe^{itx} f(x)\,d\mu(x).$$

Remark. Let μ be a Borel probability measure on $\mathbf{R}$ with characteristic function $\varphi(t) = \int_{\mathbf{R}} e^{itx}\,d\mu(x)$. Suppose that $\int_{\mathbf{R}} x^2\,d\mu(x) < \infty$. Then by applying problem X317(b) twice (first with $f(t) = 1$ and then with $f(t) = ix$), we find that φ is C^2 and that for each $t \in \mathbf{R}$,

$$\varphi''(t) = \int_{\mathbf{R}} -x^2 e^{itx}\,d\mu(x).$$

In particular,

$$-\varphi''(0) = \int_{\mathbf{R}} x^2\,d\mu(x).$$

Since $\varphi(0) = 1$, l'Hospital's rule plus the definition of $\varphi''(0)$ shows that as $t \to 0$,

$$\frac{2 - \varphi(t) - \varphi(-t)}{t^2} \to -\varphi''(0),$$

so

$$\frac{2 - \varphi(t) - \varphi(-t)}{t^2} \to \int_{\mathbf{R}} x^2\,d\mu(x).$$

This last fact still holds even without the assumption that $\int_{\mathbf{R}} x^2\,d\mu(x) < \infty$, as part (a) of the next exercise shows.

X318. *due 7Th*

(a) Let μ be a Borel probability measure on $\mathbf{R}$ with characteristic function φ. Prove that as $t \to 0$,

$$\frac{2 - \varphi(t) - \varphi(-t)}{t^2} \to \int_{\mathbf{R}} x^2 \, d\mu(x).$$

(This holds whether or not the integral on the right is finite. There is a simple proof that does not require considering cases and that does not depend on problem X317(b). Hint: Recall that $1 - \cos u = 2\sin^2(u/2)$.)

(b) For $0 < \alpha < \infty$, define φ_α on $\mathbf{R}$ by $\varphi_\alpha = \exp(-|t|^\alpha)$. Prove that if $\alpha > 2$, then there is no Borel probability measure on $\mathbf{R}$ with characteristic function φ_α.

Remark. What makes problem X318(b) interesting is that in contrast, if $0 < \alpha \le 2$, then φ_α is the characteristic function of a Borel probability measure on $\mathbf{R}$, a so-called symmetric stable distribution of index α. (You are not asked to prove this.) To those who know probability theory, it will be a familiar fact that if $\alpha = 2$, then φ_α is the characteristic function of the normal distribution with mean 0 and variance 2.

Reminder. If (a_n) and (b_n) are sequences of non-zero real or complex numbers, then to say that a_n *is asymptotic to* b_n *as* n *tends to infinity* (written $a_n \sim b_n$ as $n \to \infty$) means that $a_n/b_n \to 1$ as $n \to \infty$.

X319. Let $a > 1$. The question of the asymptotic behaviour of the sequence of numbers

$$I_n = \int_{na}^{\infty} x^{n-1} e^{-x} \, dx, \qquad n = 1, 2, 3, \ldots,$$

arises in probability theory.[42] Prove that

$$I_n \sim C n^{n-1} a^{n-1} e^{-na} \qquad \text{as } n \to \infty,$$

where $C \in (0, \infty)$ is a suitable constant, and determine the value of C. (Hint: $I_n = n^{n-1} a^{n-1} e^{-na} J_n$ where J_n is a certain integral over the interval (na, ∞). The constant C will be $\lim_{n\to\infty} J_n$. Perform a substitution in the integral J_n to get an integral over the interval $(0, \infty)$ to which you can apply a suitable limit theorem.)

Definitions. Let $(X, \mathscr{A})$ be a measurable space and let μ and ν be measures on $\mathscr{A}$. To say that ν *is absolutely continuous with respect to* μ (denoted by $\nu \ll \mu$) means that for each $A \in \mathscr{A}$, if $\mu(A) = 0$, then $\nu(A) = 0$. To say that μ *and* ν *are equivalent* means that $\mu \ll \nu$ and $\nu \ll \mu$.

X320. Let $(X, \mathscr{A}, \mu)$ be a measure space. Let $\rho\colon X \to [0, \infty]$ be measurable. Define $\nu\colon \mathscr{A} \to [0, \infty]$ by

$$\nu(A) = \int_A \rho \, d\mu$$

for all $A \in \mathscr{A}$.

(a) Prove that ν is a measure on $\mathscr{A}$. (We call ν the measure with density ρ with respect to μ.)

(b) Prove that $\nu \ll \mu$.

(c) Prove that $\nu \approx \mu$ iff $\mu(\rho = 0) = 0$.

(d) Prove that for each measurable function $f\colon X \to [0, \infty]$, we have $\int f \, d\nu = \int f\rho \, d\mu$. (Hint: First, check this if $f = 1_A$ where $A \in \mathscr{A}$. Then check it if $f\colon Y \to [0, \infty)$ is simple. Then check it in the general case. By the way, this sort of reasoning arises often in the theory of integration.)

(e) Let $f\colon X \to \overline{\mathbf{R}}$ be measurable. Prove that if either of the integrals $\int f \, d\nu$ and $\int f\rho \, d\mu$ exists, then so does the other and they are equal.

---

[42] For those who know probability theory, we mention that I_n is the probability that $(X_1 + \cdots + X_n)/n$ exceeds a, where $X_1, X_2, X_3, \ldots$ are independent exponentially distributed random variables each of which has expected value equal to 1.

The Radon-Nikodym Theorem. Let $(X, \mathscr{A}, \mu)$ be a measure space and let ν be another measure on $\mathscr{A}$. In the light of problem X320, it is natural to ask when ν will have a density with respect to μ. In other words, when will there exist a measurable function $\rho\colon X \to [0, \infty]$ such that for each $A \in \mathscr{A}$, we have

$$\nu(A) = \int_A \rho \, d\mu.$$

The obvious necessary condition is that ν must be absolutely continuous with respect to μ. The *Radon-Nikodym theorem* says that this condition is also sufficient, provided that μ is σ-finite. It is also natural to ask to what extent such a density of ν with respect to μ is unique. The next exercise deals with this.

X321. Let $(X, \mathscr{A}, \lambda)$ be a semifinite measure space. Let $g, h : X \to [0, \infty]$ be measurable. Let μ be the measure with density g with respect to λ and let ν be the measure with density h with respect to λ.

- **(a)** Suppose that $\mu \le \nu$. Prove that $g \le h$ λ-almost everywhere. (Hint: This follows immediately from an earlier exercise.)
- **(b)** Suppose that $\mu = \nu$. Prove that $g = h$ λ-almost everywhere.

X322. (*The abstract change of variables formula.*) Let $(X, \mathscr{A})$ and $(Y, \mathscr{B})$ be measurable spaces and let $T\colon X \to Y$ be measurable. Let λ be a measure on X. Define $\mu\colon \mathscr{B} \to [0, \infty]$ by $\mu(B) = \lambda(T^{-1}[B])$ for all $B \in \mathscr{B}$.

- **(a)** Prove that μ is a measure on $\mathscr{B}$.
- **(b)** Prove that for each measurable function $f\colon Y \to [0, \infty]$, we have $\int_Y f \, d\mu = \int_X f \circ T \, d\lambda$. (Hint: You can use the an approach similar to the one that was suggested for problem X320(d).)
- **(c)** Let $f\colon X \to \overline{\mathbf{R}}$ be measurable. Prove that if either of the integrals $\int_Y f \, d\mu$ and $\int_X f \, d\lambda$ exists, then so does the other, and they are equal.

Notation and Terminology. Let $(X, \mathscr{A})$ be a measure space and let $p \in X$. Define δ_p on $\mathscr{A}$ by

$$\delta_p(A) = \begin{cases} 1 & \text{if } p \in A, \\ 0 & \text{if } p \notin A, \end{cases}$$

for all $A \in \mathscr{A}$. It is easy to check that δ_p is a probability measure on $\mathscr{A}$. We call δ_p the *unit point mass at* p. Another name for δ_p is the *Dirac measure at* p. (The reason for this is that when $X = \mathbf{R}$ and $\mathscr{A}$ is the Borel σ-field on $\mathbf{R}$, δ_p is the mathematically rigorous realization of the notorious Dirac delta function centered at p.) By the way, the definition of δ_p may be written more succinctly as follows: $\delta_p(A) = 1_A(p)$ for all $A \in \mathscr{A}$.

Remark. Let $(X, \mathscr{A})$, $(Y, \mathscr{B})$, T, λ, and μ be as in problem X322. The measure μ is called *the image of* λ *under the map* T. It is easy to check that if $p \in X$ and λ is δ_p, the unit point mass at p, then μ is the unit point mass at $f(p)$. In this sense, the notion of the image of a measure under the map T is a generalization of the notion of the image of a point under T.

X323. Let $a, b \in \mathbf{R}$ with $a \ne 0$. Define $T\colon \mathbf{R} \to \mathbf{R}$ by $T(x) = ax + b$. Let m be Lebesgue measure on $\mathbf{R}$, let $\mathscr{M}$ be the σ-field of Lebesgue measurable subsets of $\mathbf{R}$, and let $\mathscr{B}$ be the Borel σ-field on $\mathbf{R}$.

- **(a)** Check that for each interval $I \subseteq \mathbf{R}$, we have $m(T^{-1}[I]) = |a^{-1}| m(I)$.
- **(b)** Prove that for each $B \in \mathscr{B}$, we have $m(T^{-1}[B]) = |a^{-1}| m(B)$. (Hint: Apply the uniqueness part of the Hahn extension theorem.)
- **(c)** Prove that for each $E \in \mathscr{M}$, we have $T^{-1}[E] \in \mathscr{M}$ and $m(T^{-1}[E]) = |a^{-1}| m(E)$.
- **(d)** Prove that for each Lebesgue measurable function $f\colon \mathbf{R} \to [0, \infty]$, $f \circ T$ is Lebesgue measurable and $\int_{\mathbf{R}} f(ax + b) \, dx = |a^{-1}| \int_{\mathbf{R}} f(x) \, dx$. (Hint: Apply problem X322.)
- **(e)** Let $f\colon \mathbf{R} \to \overline{\mathbf{R}}$ be Lebesgue measurable. Prove that if either of the integrals $\int_{\mathbf{R}} f(ax + b) \, dx$ and $\int_{\mathbf{R}} |a^{-1}| f(x) \, dx$ exists, then so does the other and they are equal.

X324. Let $(X, \mathscr{A}, \mu)$ be a probability space. Let $p, q \in [1, \infty]$ with $p \le q$ and let $f \in L^0(\mu)$. Prove that $\|f\|_p \le \|f\|_q$. (Hint: This can be deduced from Hölder's inequality. Take one of the functions to be identically 1.)

Remark. Let $(X, \mathscr{A}, \mu)$ be a finite measure space. (In other words, let $(X, \mathscr{A}, \mu)$ be a measure space for which $\mu(X) < \infty$.) Let $p, q \in [1, \infty]$ with $p \le q$. Then there is a constant $C \in [0, \infty)$, depending only on p, q, and $\mu(X)$, such that for each $f \in L^0(\mu)$, $\|f\|_p \le C\|f\|_q$. This can be proved in the same way as the result of problem X324. In particular, $L^q(\mu) \subseteq L^p(\mu)$.

X325. Let X be a set and let γ be counting measure on X. For each $p \in [1, \infty]$ and for each $f\colon X \to \mathbf{K}$, let $\|f\|_p$ denote the p-norm of f relative to the measure γ. In other words, for each $p \in [1, \infty]$ and for each $f\colon X \to \mathbf{K}$, let

$$\|f\|_p = \begin{cases} \left(\sum_{x \in X} |f(x)|^p\right)^{1/p} & \text{if } p < \infty, \\ \sup\{\, |f(x)| : x \in X \,\} & \text{if } p = \infty. \end{cases}$$

Let $p, q \in [1, \infty]$ with $p \le q$ and let $f\colon X \to \mathbf{K}$. Prove that $\|f\|_p \ge \|f\|_q$. (Hint: This is almost trivial. Reduce to the case where $\|f\|_p = 1$.)

Remark. Let X be a set. For each $p \in [1, \infty]$, $\ell^p(X)$ denotes $L^p(\gamma)$, where γ is counting measure on X. It follows immediately from problem X325 that if $p, q \in [1, \infty]$ with $p \le q$, then $\ell^p(X) \subseteq \ell^q(X)$.

Remark. Note that the inequalities in the conclusions of problem X324 and problem X325 are opposites of one another.

X326. Suppose μ is Lebesgue measure on $\mathbf{R}$. Fix $p, q \in [1, \infty]$ with $p < q$.
- **(a)** Show that $L^q(\mu)$ is not a subset of $L^p(\mu)$.
- **(b)** Show that $L^p(\mu)$ is not a subset of $L^q(\mu)$.

Reminder. $L^0(\mathbf{T})$ denotes the set of Lebesgue measurable 1-periodic functions from $\mathbf{R}$ to $\mathbf{C}$.

X327. Let $f \in L^0(\mathbf{T})$. Let $a \in \mathbf{R}$. Prove that if either of the integrals $\int_a^{a+1} f(t)\,dt$ and $\int_0^1 f(t)\,dt$ exists, then so does the other and they are equal.

Definition. Let $f, g \in L^0(\mathbf{T})$. The *convolution of f and g* is the function $f * g$ defined by

$$(f * g)(t) = \int_0^1 f(t - s)g(s)\,ds$$

for all $t \in \mathbf{R}$ for which the integral exists.

X328. Let $f, g \in L^0(\mathbf{T})$. Prove that $f * g = g * f$. (In other words, for each $t \in \mathbf{R}$, if either of the quantities $(f * g)(t)$ and $(g * f)(t)$ is defined, then so is the other and they are equal.)

X329. Let $p, q \in [1, \infty]$ with $p^{-1} + q^{-1} = 1$. Let $f \in L^p(\mathbf{T})$ and let $g \in L^q(\mathbf{T})$.
- **(a)** Prove that for each $t \in \mathbf{R}$, $(f * g)(t)$ is defined and $|(f * g)(t)| \le \|f\|_p \|g\|_q$.
- **(b)** If in addition $f \in C(\mathbf{T})$ or $g \in C(\mathbf{T})$, prove that $f * g \in C(\mathbf{T})$.
- **(c)** Without assuming that $f \in C(\mathbf{T})$ or $g \in C(\mathbf{T})$, prove that $f * g \in C(\mathbf{T})$. (Hint: Either $p \in [1, \infty)$ or $q \in [1, \infty)$. Say $p \in [1, \infty)$. Then there is a sequence f_n in $C(\mathbf{T}) \cap L^p(\mathbf{T})$ such that $\|f - f_n\|_p \to 0$. Prove that $f_n * g \to f * g$ uniformly on $\mathbf{R}$.)

Notation. For each $f \in L^1(\mathbf{T})$ and for each $k \in \mathbf{Z}$, $\hat{f}(k) = \int_0^1 e^{-2\pi i k x} f(x)\,dx$, by definition. (Earlier, for $f \in L^1(\mathbf{R})$, we defined $\hat{f}$ differently, namely by $\hat{f}(y) = \int_{\mathbf{R}} f(t)e^{-2\pi i t y}\,dt$ for all $y \in \mathbf{R}$. Which meaning is intended for $\hat{f}$ should be clear from the context.)

X330. Let $f, g \in C(\mathbf{T})$. Suppose $\hat{f} = \hat{g}$. Prove that $f(t) = g(t)$ for all $t \in \mathbf{R}$.

X331. Let $c \in \ell^1(\mathbf{Z})$.
- **(a)** Prove that the series $\sum_{k \in \mathbf{Z}} c(k)e_k$ converges in $C(\mathbf{T})$, in the sense of unordered sums, to some $f \in C(\mathbf{T})$.
- **(b)** Prove that $\hat{f}(k) = c(k)$ for each $k \in \mathbf{Z}$.

X332. Let $g \in C(\mathbf{T})$. Suppose that $\hat{g} \in \ell^1(\mathbf{Z})$. Let $f = \sum_{k\in\mathbf{Z}} \hat{g}(k)e_k$. By problem X331, this makes sense and $f \in C(\mathbf{T})$. Prove that $f(t) = g(t)$ for all $t \in \mathbf{R}$.

Reminder. From problem X315, we know that $\int_{\mathbf{R}} e^{-2\pi ikx}e^{-\pi x^2}\,dx = e^{-\pi k^2}$ for all $k \in \mathbf{R}$.

X333. Prove that *due 8Th*

$$(4\pi t)^{1/2}\sum_{k\in\mathbf{Z}} e^{-4\pi^2k^2t}e^{2\pi ikx} = \sum_{n\in\mathbf{Z}} e^{-(x-n)^2/(4t)} \qquad \text{for all } x \in \mathbf{R} \text{ and all } t > 0. \tag{38}$$

(Hint: Fix $t > 0$. Let φ be the function of x on the right side of the equation. Show that $\varphi \in L^2(\mathbf{T})$ and $\int_0^1 e^{-2\pi ikx}\varphi(x)\,dx = (4\pi t)^{1/2}e^{-4\pi^2k^2t}$ for all $k \in \mathbf{Z}$. Then explain how we can bridge the gap between convergence in $L^2(\mathbf{T})$ and convergence at x for all $x \in \mathbf{R}$ in this case. This is partly a special case of some of the preceding few exercises, but please make your proof reasonably self-contained.)

The identity (38) in problem X333 is one of Jacobi's theta function identities. Note that when t is close to 0, then a substantial number of terms are needed to calculate the left side of (38) accurately, but the right side is well-approximated by the few terms with n near x. Note also that the right side of the (38) is obviously positive, whereas it is not at all obvious that the left side is positive. The method suggested above to prove (38) can also be used to establish other interesting identities. In fact, there is a general result of which all such identities are special cases. It is called the Poisson summation formula. The Poisson summation formula is discussed in Folland, *Real Analysis: Modern Techniques and Their Applications*. But of course I do not want you to solve problem X333 merely by pointing out that it follows from the Poisson summation formula! I want you to give a reasonably self-contained proof.

Notation for Partial Derivatives. Let G be an open subset of $\mathbf{R}^d$ and let $f\colon G \to \mathbf{R}$. A *multi-index* is a d-tuple of non-negative integers. If $\alpha = (\alpha_1, \ldots, \alpha_d)$ is a multi-index and if f is sufficiently differentiable, then $D^\alpha f$ denotes the partial derivative

$$\frac{\partial^{\alpha_1+\cdots+\alpha_d} f}{\partial^{\alpha_1}x_1\cdots\partial^{\alpha_d}x_d}.$$

The identity (38) is the key to the next exercise, which is closely related to Weierstrass's own proof that the trigonometric polynomials are dense in $C(\mathbf{T})$.

X334. Let $f \in C(\mathbf{T})$. If we use separation of variables to find a solution u of the heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} \qquad \text{for } x \in \mathbf{R} \text{ and } t > 0, \tag{39}$$

subject to the initial condition

$$\lim_{t\downarrow 0} u(x,t) = f(x) \qquad \text{for } x \in \mathbf{R}, \tag{40}$$

then we are led to the function u defined by

$$u = \sum_{k\in\mathbf{Z}} u_k, \qquad \text{where} \qquad u_k(x,t) = e^{-4\pi^2k^2t}\hat{f}(k)e^{2\pi ikx} \qquad \text{for } k \in \mathbf{Z},\ x \in \mathbf{R} \text{ and } t > 0. \tag{41}$$

- **(a)** Prove that for each $t > 0$, $u(\cdot,t) = f * g_t$ where g_t is a certain function in $C(\mathbf{T})$ which you should be able to figure out from problem X333. Also, calculate $\|g_t\|_1$ for each $t > 0$.
- **(b)** For each $t > 0$, define $G_t\colon C(\mathbf{T}) \to C(\mathbf{T})$ by $G_t(h) = h * g_t$. Prove that each G_t is a bounded linear operator on $C(\mathbf{T})$ and that in fact, $\sup_{t>0}\|G_t\| < \infty$, where $\|G_t\|$ denotes the norm of G_t as a linear operator on $C(\mathbf{T})$.
- **(c)** Prove that $G_t(f) \to f$ uniformly as $t \downarrow 0$. (Hint: Show that $G_t(h) \to h$ uniformly in the special case where h is a trigonometric polynomial and then show that $G_t(f) \to f$ uniformly by using this special case in conjunction with part (b) and the fact that the trigonometric polynomials are dense in $C(\mathbf{T})$.)
- **(d)** Part (c) shows that u does satisfy the initial condition (40), and does so in as strong a sense as we could hope for given the fact that the initial temperature distribution f is merely continuous.

It remains to show that u does satisfy the heat equation. Let $\Omega = \mathbf{R} \times (0,\infty)$. Prove that $u \in C^\infty(\Omega)$, and that the series (41) converges to u in the following very strong sense: For each multi-index α, for each $t_0 > 0$, and for each $\varepsilon > 0$, there exists a finite subset $A_0 \subseteq \mathbf{Z}$ such that for each finite subset $A_1 \subseteq \mathbf{Z}$ with $A_0 \subseteq A_1$, and for each $(x,t) \in \mathbf{R} \times [t_0,\infty)$, $|(D^\alpha u)(x,t) - \sum_{k\in A_1}(D^\alpha u_k)(x,t)| \le \varepsilon$. Deduce in particular that u does satisfy (39).

The method outlined in the hint for part (c) of problem X334 is the model for a large variety of different results about convergence of Fourier series and related expressions. For instance, one can use it to show that if f is merely in $L^p(\mathbf{T})$, where $p \in [1,\infty)$, then the conclusion of part (d) still holds and $u(\cdot,t) \to f$ in $L^p(\mathbf{T})$, so that the initial condition (40) is still satisfied when suitably interpreted. The proof is not much different than for the case where $f \in C(\mathbf{T})$. One only needs to observe that $\|f * g_t\|_p \le \|f\|_p \|g_t\|_1$ and that the trigonometric polynomials are dense in $L^p(\mathbf{T})$. Note that the sense in which the initial condition (40) holds is adapted to the function space in which f is assumed to lie. Note also that the Fourier series $\sum_{k=-\infty}^{\infty} \hat{f}(k)e^{2\pi ikx}$ need not itself converge for the initial condition (40) to be satisfied. It is sufficient that the "regularized" series $\sum_{k\in\mathbf{Z}} e^{-4\pi^2k^2t}\hat{f}(k)e^{2\pi ikx}$ behave appropriately as $t \downarrow 0$, and this holds in much greater generality.[43] This is a pleasant feature of the heat equation. Laplace's equation is nicely behaved in a similar way, but the wave equation is not.

Notation. Let X be a set. Then $c_0(X)$ denotes the set of functions $g\colon X \to \mathbf{K}$ which tend to zero at infinity in the sense that for each $\varepsilon > 0$, the set $\{\, |g| \ge \varepsilon \,\}$ is finite. For instance, $c_0(\mathbf{Z})$ is the set of functions $g\colon \mathbf{Z} \to \mathbf{K}$ such that $g(k) \to 0$ as $|k| \to \infty$. When $c_0(\mathbf{Z})$ is mentioned in the context of Fourier series, it is understood that $\mathbf{K} = \mathbf{C}$.

X335. (*The Riemann-Lebesgue lemma for Fourier series.*) Let $f \in L^1(\mathbf{T})$. It is obvious that $\hat{f} \in \ell^\infty(\mathbf{Z})$ and $\|\hat{f}\|_\infty \le \|f\|_1$. Prove that $\hat{f} \in c_o(\mathbf{Z})$. (Hint: Let E be the set of all g in $L^1(\mathbf{T})$ such that $\hat{g} \in c_0(\mathbf{Z})$. Verify that E is closed in $L^1(\mathbf{T})$. This is easy. Then prove that E is dense in $L^1(\mathbf{T})$.)

Notation. Suppose $f \in L^1(\mathbf{T})$. For all integers $m,n \ge 0$, define $S_{m,n}f$ by

$$S_{m,n}f = \sum_{k=-m}^{n} \hat{f}(k)e_k.$$

We call $S_{m,n}f$ the (m,n)-th asymmetric partial sum of the Fourier series for f. For each integer $n \ge 0$, define S_nf by

$$S_nf = \sum_{k=-n}^{n} \hat{f}(k)e_k.$$

We call S_nf the n-th symmetric partial sum of the Fourier series for f. For each integer $n \ge 0$, define D_n by

$$D_n = \sum_{k=-n}^{n} e_k.$$

The function D_n is called *the Dirichlet kernel of order* n. As we know, for each $k \in \mathbf{Z}$, we have $e_k * f = \hat{f}(k)e_k$. It follows immediately that $S_nf = D_n * f$.

The elegant treatment outlined in the next couple of exercises is taken from an article by Paul R. Chernoff.

---

[43] Du Bois-Reymond (1873) gave an example of a function $f \in C(\mathbf{T})$ whose Fourier series fails to converge at a certain point. In fact, there is a function $f \in C(\mathbf{T})$ whose Fourier series converges only on a meager set. A meager set can be of full Lebesgue measure, but Kolmogorov showed that there is a function $f \in L^1(\mathbf{T})$ whose Fourier series diverges almost everywhere (1923) or even everywhere (1926). Marcel Riesz (1927) showed that if $f \in L^p(\mathbf{T})$, where $1 < p < \infty$, then the Fourier series of f converges to f in $L^p(\mathbf{T})$; however, unless $p = 2$, the convergence is in general not in the sense of unordered sums. Lusin (1912) had conjectured that if $f \in L^2(\mathbf{T})$, then the Fourier series of f converges almost everywhere. The evidence of the next few decades seemed to point in the opposite direction however and Zygmund (1959), in the preface of the second edition of his famous treatise on trigonometric series, wrote "The problem of the existence of a continuous function with an everywhere divergent Fourier series is still open." Carleson (1966) astonished the experts by proving Lusin's conjecture. R. A. Hunt (1968) simplified Carleson's proof and confirmed Carleson's claim that the conclusion of Lusin's conjecture holds for $f \in L^p(\mathbf{T})$ where $p > 1$, not just for $f \in L^2(\mathbf{T})$.

X336. Let $f \in L^1(\mathbf{T})$, let $a \in \mathbf{R}$, and suppose *due 8Th*

$$\int_{-1/2}^{1/2} \frac{|f(a+h) - f(a)|}{|h|}\, dh < \infty. \tag{42}$$

Prove that $(S_{m,n}f)(a) \to f(a)$ as $m, n \to \infty$. (Hint: To simplify the calculations, shift the origin and subtract a suitable constant from f to reduce to the case where $a = 0$ and $f(0) = 0$. Let $g(x) = f(x)/(e^{2\pi ix} - 1)$. Use (42) to show that $g \in L^1(\mathbf{T})$. Use the fact that $f(x) = (e^{2\pi ix} - 1)g(x)$ to relate $\hat{f}$ and $\hat{g}$. Conclude that $(S_{m,n}f)(0)$ is a telescoping sum and apply the Riemann-Lebesgue lemma to g to show that $(S_{m,n}f)(0) \to 0$ as $m, n \to \infty$.)

Remark. Let $f \in L^1(\mathbf{T})$ and let $a \in \mathbf{R}$. It is clear that if f satisfies any one of the following successively more general conditions, then (42) holds.

- **(a)** f is differentiable at a.
- **(b)** f has finite one-sided derivatives at a.
- **(c)** f satisfies a Lipschitz condition at a; in other words, there exists a constant $C \in [0, \infty)$ for that for each x in some nhd of a, $|f(x) - f(a)| \le C|x - a|$.
- **(d)** f satisfies a Hölder condition at a; in other words, there exist constants $C \in [0, \infty)$ and $\alpha \in (0, 1]$ such that for each x in some nhd of a, $|f(x) - f(a)| \le C|x - a|^{\alpha}$.

X337. Let $f \in L^1(\mathbf{T})$ and let $a \in \mathbf{R}$, and suppose

$$\int_{-1/2}^{1/2} \frac{|f(a+h) + f(a-h) - 2f(a)|}{|h|}\, dh < \infty. \tag{43}$$

Prove that $(S_n f)(a) \to f(a)$ as $n \to \infty$. (Hint: To simplify the calculations, shift the origin and subtract a suitable constant from f to reduce to the case where $a = 0$ and $f(0) = 0$. Then apply problem X336 to the function $(f(x) + f(-x))/2$.)

Remark. Let $f \in L^1(\mathbf{T})$, let $a \in \mathbf{R}$, and suppose $f(a+)$ and $f(a-)$ exist and are finite. Suppose also that f satisfies any one of the following successively more general conditions:

- **(a)** $\lim_{h\to 0+}(f(a+h) - f(a+))/h$ and $\lim_{h\to 0-}(f(a+h) - f(a-))/h$ both exist and are finite.
- **(b)** f satisfies one-sided Lipschitz conditions at a; in other words, there exists a constant $C \in [0, \infty)$ and there exists $\delta > 0$ such that for each $x \in (a, a+\delta)$, we have $|f(x) - f(a+)| \le C|x - a|$, and for each $x \in (a - \delta, a)$, we have $|f(x) - f(a-)| \le C|x - a|$.
- **(c)** f satisfies one-sided Hölder conditions at a; in other words, there exist constants $C \in [0, \infty)$ and $\alpha \in (0, 1]$, and there exists $\delta > 0$ such that for each $x \in (a, a+\delta)$, we have $|f(x) - f(a+)| \le C|x - a|^{\alpha}$, and for each $x \in (a - \delta, a)$, we have $|f(x) - f(a-)| \le C|x - a|^{\alpha}$.

Then $S_n f(a) \to (f(a+) + f(a-))/2$ as $n \to \infty$. I leave it to you to to use problem X337 to verify this. Note that the restriction to symmetric partial sums is essential here, as one sees by considering the "square wave" function f defined by $f(x) = \operatorname{sgn}(\sin 2\pi x)$, where $\operatorname{sgn}(u) = u/|u|$ for $u \neq 0$ and $\operatorname{sgn}(0) = 0$.

X338. Let X, Y, and Z be normed linear spaces. Define $f\colon \mathscr{L}(X,Y)\times\mathscr{L}(Y,Z)\to\mathscr{L}(X,Z)$ by

$$f(S,T)=T\circ S.$$

Prove that f is continuous.

X339. Define $f\colon \mathbf{R}^2\to\mathbf{R}$ by

$$f(x,y)=(x^2-y)(y-2x^2).$$

Let $Z=\{\,f=0\,\}$ and let $G=\{\,f\neq 0\,\}$.

(a) Sketch the set Z.

(b) Find all the components of G and determine the sign of f in each of them.

(c) Show that for each straight line L through the origin in $\mathbf{R}^2$, the restriction of f to L has a strict local maximum at the origin. (Do not use calculus for this.)

(d) Show that f does not have a local maximum at the origin. (Do not use calculus for this.)

X340. Let $f\colon \mathbf{R}\to\mathbf{R}$ be differentiable. Suppose that f achieves a local minimum at the origin. Suppose also that the origin is the only critical point for f. Show that f has a global minimum at the origin.

X341. Define $f\colon \mathbf{R}^2\to\mathbf{R}$ by *due 9Th*

$$f(x,y)=x^2(1+y)^3+y^2.$$

(a) Show that f achieves a strict local minimum at the origin. (Do not use calculus for this.)

(b) Show that the origin is the only critical point for f.

(c) Show that f does not achieve a global minimum at the origin.

X342. Let X and Y be Banach spaces. Define $f\colon \mathscr{GL}(X,Y)\to\mathscr{GL}(Y,X)$ by $f(S)=S^{-1}$. Let $S_0\in\mathscr{GL}(X,Y)$. Show that f is differentiable at S_0 and that $f'(S_0)$ is the linear map from $\mathscr{L}(X,Y)$ to $\mathscr{L}(Y,X)$ such that for each $T\in\mathscr{L}(X,Y)$,

$$f'(S_0)(T)=-S_0^{-1}TS_0^{-1}. \tag{44}$$

Remark. When $X=Y=\mathbf{K}$, then under the usual identification of $\mathscr{L}(X,Y)$ and $\mathscr{L}(Y,X)$ with $\mathbf{K}$, (44) reduces to the familiar formula

$$\left.\frac{d}{ds}s^{-1}\right|_{s=s_0}=-s_0^{-2}.$$

However, when $X=Y$ but $\dim(X)>1$, such a reduction is not possible, because S_0^{-1} and T need not commute.

X343. Let X and Y be Banach spaces, let Ω be an open subset of X, let $f\colon\Omega\to Y$, and let $a\in\Omega$. Show that changing the norms on X and Y to other equivalent norms does not affect whether f is differentiable at a, nor does it affect $f'(a)$.

X344. Let X and Y be normed linear spaces. Let ν be any norm on $\mathbf{R}^2$. Let

$$\|(x,y)\|=\nu(\|x\|,\|y\|) \tag{45}$$

for all $(x,y)\in X\times Y$.

(a) Show that (45) defines a norm on $X\times Y$ which induces the product topology on $X\times Y$.

(b) Show that if X and Y are Banach spaces, then $X\times Y$ is a Banach space under this norm.

Remark. Different choices of the norm ν in (45) yield different but equivalent norms on $X\times Y$. Thus confusion should not arise if we simply speak of $X\times Y$ as a normed linear space, without specifying the norm. Any norm that induces the product topology will do. And problem X344 allows us easily to give examples of such norms. For instance, $\|(x,y)\|=\|x\|+\|y\|$ defines such a norm. So does $\|(x,y)\|=(\|x\|^2+\|x\|^2)^{1/2}$. (A different but equivalent one, of course.)

X345. (*Coordinatewise differentiation.*) Let X, Y, and Z be Banach spaces. Let Ω be an open subset of X, let $f\colon\Omega\to Y$, let $g\colon\Omega\to Z$, and define $h\colon\Omega\to Y\times Z$ by $h(x)=\big(f(x),g(x)\big)$. Let $p\in\Omega$. Show that h is differentiable at p iff f and g are both differentiable at p, and that in this case, for each $u\in X$, we have $h'(p)(u)=\big(f'(p)(u),g'(p)(u)\big)$.

X346. (*The product rule.*) Let X, Y, and Z be Banach spaces. Let $\varphi\colon X \times Y \to Z$ be continuous and bilinear.

(a) Let $(a,b) \in X \times Y$. Prove that φ is differentiable at (a,b) and that for each $(x,y) \in X \times Y$,

$$\varphi'(a,b)(x,y) = \varphi(a,y) + \varphi(x,b).$$

(b) Let E be another Banach space, let Ω be an open subset of E, let $f\colon \Omega \to X$, let $g\colon \Omega \to Y$, let $p \in \Omega$, and suppose f and g are both differentiable at p. Define $h\colon \Omega \to Z$ by

$$h(u) = \varphi\big(f(u), g(u)\big).$$

Prove that h is differentiable at p and that for each $u \in E$,

$$h'(p)(u) = \varphi\big(f(p), g'(p)(u)\big) + \varphi\big(f'(p)(u), g(p)\big).$$

(Hint: Combine part (a), the chain rule, and problem X345.)

Example. For all $y, z \in \mathbf{R}^3$, let $y \times z \in \mathbf{R}^3$ be the usual cross product of the vectors y and z. Let E be a Banach space, let Ω be an open subset of E, let $f, g\colon E \to \mathbf{R}^3$, and define $h\colon \Omega \to \mathbf{R}^3$ by

$$h(u) = f(u) \times g(u).$$

Let $p \in \Omega$ and suppose that f and g are differentiable at p. Then h is differentiable at p and for each $u \in E$,

$$h'(p)(u) = f(p) \times g'(p)(u) + f'(p)(u) \times g(p).$$

To see this, just apply problem X346(b) with $\varphi(y,z) = y \times z$.

Remark. Similar considerations apply to the product of a scalar-valued function and a vector-valued function. In a Hilbert space, similar considerations apply to the inner product of two vector-valued functions.

X347. Let X be a real Hilbert space. Define $f, g\colon X \to \mathbf{R}$ by $f(x) = \|x\|^2$ and $g(x) = \|x\|$.

(a) Let $a \in X$. Show that f is differentiable at a and that for each $x \in X$, $f'(a)(x) = 2\,\langle a|x\rangle$.

(b) Let $a \in X \setminus \{0\}$. Show g is differentiable at a and that for each $x \in X$, $g'(a)(x) = \langle a|x\rangle / \|a\|$.

(c) Show that g is not differentiable at 0, except in the trivial case where $X = \{0\}$.

X348. Define $f, g\colon \mathbf{R}^2 \to \mathbf{R}$ by $f(0) = 0$, $g(0) = 0$, and for each $x = (x_1, x_2) \in \mathbf{R}^2 \setminus \{0\}$, *due 9Th*

$$f(x) = \frac{x_1 x_2^2}{x_1^2 + x_2^2} \qquad \text{and} \qquad g(x) = \frac{x_1^3 x_2}{x_1^4 + x_2^2}.$$

Clearly f and g are continuously differentiable on $\mathbf{R}^2 \setminus \{0\}$. What about at 0?

(a) Show that for each $v \in \mathbf{R}^2$, the directional derivative

$$\lim_{t \to 0} \frac{f(tv) - f(0)}{t} = A(v)$$

exists in $\mathbf{R}^2$. However, show that the function $v \mapsto A(v)$ is not linear. Deduce that f is not differentiable at 0.

(b) Show that for each $v \in \mathbf{R}^2$, the directional derivative

$$\lim_{t \to 0} \frac{g(tv) - g(0)}{t} = B(v)$$

exists in $\mathbf{R}^2$. Furthermore, show that the function $v \mapsto B(v)$ is linear. But show that nevertheless, g is not differentiable at 0.

Definition. Let X and Y be metric spaces, let $f, g\colon X \to Y$, and let $a \in X$. To say that *f and g are tangent at a* means that for each $\varepsilon > 0$, there exists $\delta > 0$ such that for each $x \in B_X(a, \delta)$, we have

$$d_Y(f(x), g(x)) \le \varepsilon d_X(x, a).$$

Notice that if a is not an isolated point of X, then f and g are tangent at a iff $f(a) = g(a)$ and

$$\lim_{x \to a} \frac{d_Y(f(x), g(x))}{d_X(x, a)} = 0.$$

X349. Let X and Y be Banach spaces, let Ω be an open subset of X, let $f\colon \Omega \to Y$, let $a \in \Omega$, let $T \in \mathscr{L}(X, Y)$, and define $g\colon \Omega \to Y$ by $g(x) = f(a) + T(x - a)$. Show that the following are equivalent:

(a) The function f is differentiable at a and $f'(a) = T$.

(b) The functions f and g are tangent at a.

X350. Let X and Y be metric spaces and let $a \in X$. Show that the relation of being tangent at a is an equivalence relation on the set of all functions from X into Y.

X351. Let X and Y be metric spaces, let $f, g\colon X \to Y$, and let $a \in X$. Suppose that f and g are tangent at a. Prove that f is continuous at a iff g is continuous at a.

Definition. Let X and Y be metric spaces, let $f\colon X \to Y$, and let $a \in X$. To say that *f satisfies Stepanoff's condition at a* means that there exist $C, \delta \in (0, \infty)$ such that for each $x \in B_X(a, \delta)$, we have $d_Y(f(x), f(a)) \le C d_X(x, a)$.

Evidently, if f is a Lipschitz function, then f satisfies Stepanoff's condition at each point. It is also obvious that if f satisfies Stepanoff's condition at a, then f is continuous at a.

X352. Let X and Y be metric spaces, let $f, g\colon X \to Y$, and let $a \in X$. Suppose that f and g are tangent at a. Prove that f satisfies Stepanoff's condition at a iff g satisfies Stepanoff's condition at a.

X353. Let X, Y, and Z be metric spaces and let $a \in X$. Let $f, F\colon X \to Y$ be tangent at a. Let $b = f(a)$. Note that $b = F(a)$ too. Let $g, G\colon Y \to Z$ be tangent at b. In addition, suppose that F satisfies Stepanoff's condition at a and that G is a Lipschitz function. Let $h = g \circ f$ and $H = G \circ F$. Prove that h and H are tangent at a.

X354. Use problem X353 to give a proof of the chain rule.

Remark. Let X be a Banach space over $\mathbf{C}$. Then we may view X as a Banach space over $\mathbf{R}$ simply by forgetting that we can multiply vectors in X by scalars that are not real.

Remark. Let X and Y be Banach spaces over $\mathbf{C}$, let Ω be an open subset of X, let $f\colon \Omega \to Y$, and let $a \in \Omega$. Then differentiability of f at a means differentiability in the complex sense, but we may also speak of differentiability of f at a in the real sense, just by viewing X and Y as Banach spaces over $\mathbf{R}$, as described in the previous remark. Suppose f is differentiable in the complex sense at a. Let T be the derivative of f at a in the complex sense. Then T is a continuous complex-linear map from X to Y. But then T is also real-linear and clearly f is differentiable in the real sense at a and the derivative of f at a in the real sense is also T.

X355. Let X and Y be Banach spaces over $\mathbf{C}$, let Ω be an open subset of X, let $f\colon \Omega \to Y$, and let $a \in \Omega$. Suppose that f is differentiable at a in the real sense. Let T be the derivative of f at a in the real sense. Note that T is a continuous real-linear map from X to Y. Prove that the following are equivalent:

(a) The map f is differentiable at a in the complex sense.

(b) For each $w \in X$, $T(iw) = iT(w)$, where $i = \sqrt{-1}$.

(c) The map T is complex-linear.

X356. View $\mathbf{C}$ as a two-dimensional vector space over $\mathbf{R}$. Let T be a real-linear map from $\mathbf{C}$ to $\mathbf{C}$. Prove that the following are equivalent:

(a) The map T is complex-linear.

(b) There is a complex number γ such that for each $w \in \mathbf{C}$, $T(w) = \gamma w$.

(c) The matrix of T relative to the basis

$$1,\ i$$

is of the form

$$[T] = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}$$

for suitable $\alpha, \beta \in \mathbf{R}$. (Hint: You should find that $\alpha = \operatorname{Re}(\gamma)$ and $\beta = \operatorname{Im}(\gamma)$, where γ is as in (b).)

X357. Let Ω be an open subset of $\mathbf{C}$, let $f \colon \Omega \to \mathbf{C}$, let $u = \operatorname{Re}(f)$, let $v = \operatorname{Im}(f)$, and let $p \in \Omega$. Prove that the following are equivalent:

(a) f is differentiable at p in the complex sense.

(b) The limit

$$\lim_{z \to p} \frac{f(z) - f(p)}{z - p}$$

exists in $\mathbf{C}$.

(c) f is differentiable at p in the real sense and f satisfies the Cauchy-Riemann equations at p:

$$\frac{\partial u}{\partial x}(p) = \frac{\partial v}{\partial y}(p) \qquad \text{and} \qquad \frac{\partial u}{\partial y}(p) = -\frac{\partial v}{\partial x}(p).$$

Show furthermore that if f is differentiable at p in the complex-sense, and if γ is the value of the limit in (b), then for each $w \in \mathbf{C}$, we have $f'(p)(w) = \gamma w$.

Should f be continuous? **X358.** (*Uniqueness in the implicit function theorem.*) Let X, Y, and Z be Banach spaces, let Ω be an open subset of $X \times Y$, let $f \colon \Omega \to Z$, and suppose $D_2 f$ exists and is continuous throughout Ω. Let *Should f be continuous?*

$$\Omega_1 = \{\, (a, b) \in \Omega : D_2 f(a, b) \in \mathscr{GL}(Y, Z) \,\}.$$

Note that Ω_1 is open in $X \times Y$. Suppose U is a connected open subset of X, $h_1, h_2 \colon U \to Y$ are continuous, $h_1(a) = h_2(a)$ for some $a \in U$, and for each $x \in U$, we have $(x, h_1(x)) \in \Omega_1$, $(x, h_2(x)) \in \Omega_1$,

$$f(x, h_1(x)) = 0 \qquad \text{and} \qquad f(x, h_2(x)) = 0.$$

Prove that $h_1 = h_2$.

X359. In problem X358, it is natural to ask what happens if instead of requiring that $(x, h_k(x)) \in \Omega_1$ for $k = 1, 2$, we require only that $(x, h_k(x)) \in \Omega$ for $k = 1, 2$. Investigate this question for the function $f \colon \mathbf{R}^2 \to \mathbf{R}$ defined by

$$f(x, y) = \big(x^2 + (y-1)^2 - 1\big)\big(x^2 + (y-2)^2 - 4\big).$$

You should find that there is more than one C^1 solution of the equation $f(x, y) = 0$ passing through the point $(x, y) = (0, 0)$.

Remark. In Chapter 9 of Rudin, *Principles of Mathematical Analysis*, Third Edition, the inverse function theorem (Theorem 9.24) is used to prove the implicit function theorem (Theorem 9.28). This approach requires stronger differentiability assumptions in the implicit function theorem than the approach we took in class. For this reason, it is better to reverse Rudin's order and use the implicit function theorem to prove the inverse function theorem. The next two exercises outline this approach to the inverse function theorem.

X360. (*The Inverse Function Theorem, Part 1 — Local Existence and Uniqueness; Continuity*). Let X and Y be Banach spaces, let E be an open subset of X, let $f\colon E \to Y$ be differentiable. Let $a \in E$ such that $f'(a) \in \mathscr{GL}(X,Y)$ and let $b = f(a)$. Suppose $f'\colon E \to \mathscr{L}(X,Y)$ is continuous at a. Prove that there exist open sets $U \subseteq X$ and $V \subseteq Y$ such that $a \in U \subseteq E$, $b \in V$, the restriction of f to U is a one-to-one map from U onto V, and the inverse g of this restriction is continuous. (Hint: Let $\Omega = E \times Y$. Then Ω is an open subset of $X \times Y$ and $(a,b) \in \Omega$. Let $Z = Y$. Define $\varphi\colon \Omega \to Z$ by $\varphi(x,y) = y - f(x)$. Apply the implicit function theorem to φ near the point (a,b).)

X361. (*The Inverse Function Theorem, Part 2 — Smoothness*). Let X and Y be Banach spaces, let E be an open subset of X, and let $f\colon E \to Y$ be continuous. Let V be an open subset of Y and let $g\colon V \to E$ such that for each $y \in V$, $f(g(y)) = y$. Let $U = g[V]$ and let φ be the restriction of f to U. Notice that φ is a one-to-one map from U onto V.

- **(a)** Let $b \in V$ such that g is continuous at b. Let $a = g(b)$. Suppose that f is differentiable at a and that $f'(a) \in \mathscr{GL}(X,Y)$. Prove that g is differentiable at b and that $g'(b) = f'(a)^{-1}$.

For the remainder of this exercise, suppose in addition that g is continuous in V, f is differentiable at each point of U, that the restriction of f' to U is continuous, and that for each $x \in U$, $f'(x) \in \mathscr{GL}(X,Y)$.

- **(b)** Prove that g is C^1 in V and that U is open in X.
- **(c)** Let $k \in \mathbf{N}$. Suppose that f is C^k in U. Prove that g is C^k in V.
- **(d)** Suppose that f is C^∞ in U. Prove that g is C^∞ in V

Definition. The *rank* of a linear map is the dimension of its range.

X362. Let Ω be an open subset of $\mathbf{R}^n$, let $f\colon \Omega \to \mathbf{R}^m$ be continuously differentiable, let $a \in \Omega$, and let r be the rank of $f'(a)$. Show that a has a nhd $U \subseteq \Omega$ such that for each $x \in U$, the rank of $f'(x)$ is at least r. Thus the rank of $f'(x)$ is a lower semicontinuous function of x. (Hint: Find linear maps $S\colon \mathbf{R}^r \to \mathbf{R}^n$ and $T\colon \mathbf{R}^m \to \mathbf{R}^r$ such that $T \circ f'(a) \circ S \in \mathscr{GL}(\mathbf{R}^r)$.)

About the Rank Theorem. Let $n, m, r \in \mathbf{N}$ with $r \le \min\{\, n, m \,\}$. Define $L\colon \mathbf{R}^n \to \mathbf{R}^m$ by

$$L(x_1, \ldots, x_r, x_{r+1}, \ldots, x_n) = (x_1, \ldots, x_r, 0, \ldots, 0).$$

Then L is a linear map and L has rank r. The *rank theorem* tells us that if Ω is an open subset of $\mathbf{R}^n$ and $f\colon \Omega \to \mathbf{R}^m$ is a C^k function, and if $a \in \Omega$ such that $f'(x)$ has rank r for all x in some nhd of a, then it is possible to introduce C^k curvilinear coordinates in $\mathbf{R}^n$ near a and in $\mathbf{R}^m$ near $f(a)$ in such a way that near a, in terms of these coordinates, the map f looks like the linear map L.

X363. Define $f\colon \mathbf{R}^2 \to \mathbf{R}^2$ by

$$f(x,y) = (x^2 - y^2, 2xy).$$

(By the way, under the usual identification of $\mathbf{R}^2$ with $\mathbf{C}$, f can also be described as the map $z \mapsto z^2$.)

- **(a)** Find the rank of $f'(a,b)$, if $(a,b) \in \mathbf{R}^2 \setminus \{\, (0,0) \,\}$.
- **(b)** Find the rank of $f'(0,0)$.
- **(c)** Explain how the conclusion of the rank theorem breaks down for f near the point $(0,0)$.