SOLUTIONS TO EXERCISES FOR

MATHEMATICS 205A

Fall 2014

General Remarks

The main objective of the course is to present basic graduate level material, but an important secondary objective of many point set topology courses is to is to build the students' skills in writing proofs and communicating them to others. Higher skill levels are needed because proofs in graduate level mathematics courses are frequently longer, more abstract and less straightforward than their counterparts in undergraduate level courses.

The files with solutions to exercises are named `solutions*.pdf`, where $*$ is some number, perhaps followed by a letter to indicate supplementary content like drawings to accompany arguments. Exercises marked with one or two asterisks should be viewed as having lower priorities unless their solutions are specifically assigned as readings for the course.

These solutions are posted mainly for students to compare with their own efforts and to determine whether their solutions are correct or can be improved upon (compensating in part for the lack of homework grading personnel), but in some cases the solutions might also be useful if and when a student reaches an impasse. At some points, solutions to some of the more complicated exercises may be parts of reading assignments. However, as a general rule **the solutions should not be viewed as an excuse for not trying to work the exercises, especially those that are specifically assigned in course announcements** (*i.e.*, the `aabUpdate`$*$ files in the course directory). Problems at the level of those in the assigned exercises will appear on both course and qualifying examinations, so it is important for students to be able to work out solutions on their own.

Suggestions for working exercises

The directory file `polya.pdf` contains a systematic and general list of suggestions for approaching and solving mathematical problems, and the file `mathproofs.pdf` — which was written for undergraduate level courses but is still relevant at higher levels — discusses the more formal aspects of mathematical proofs. Here are a few additional comments:

1. Many exercises can be solved by imitating arguments in the course text or notes, so it is usually worthwhile to see if an exercise can be analyzed by modifying a previously seen argument. The following quotation from the *Poetics* of Aristotle (384–322 B.C.E.; see Section I, Part IV) seems appropriate:

> *The instinct of imitation is implanted in man from childhood, one difference between him and other animals being that he is the most imitative of living creatures, and through imitation learns his earliest lessons.*

The differences between human and animal behavior in some species might not be as vast as they seemed to the ancient Greeks, but the passage still reflects the importance of imitation in human thought and action.

2. The first efforts (and in many cases subsequent efforts!) at solving exercises will not necessarily be as clear or polished as some proofs and solutions in textbooks or the files posted to the course directory. Often the first attempts to find solutions are at least somewhat messy, and they usually get better as a result of increased experience, skills, and trial and error. The following quotation due to A. S. Besicovitch (1891–1970) summarizes everything in a somewhat ironic manner:

> *A mathematician's reputation rests on the number of bad proofs he has given.*

3. It is usually good to try anything that might work rather than not getting started with work on a problem. This advice reflects a frequently repeated quotation due to L. R. ("Yogi") Berra (1925–):

> *When you come to a fork in the road ... Take it.*

Likewise, if a solution to a problem is not apparent after some thought, it is often worthwhile to be systematic and look at *everything* that can be said about the given situation, no matter how insignificant it might seem at first. This is summarized in a quotation from one of the *Sherlock Holmes* stories by A. C. Doyle (1859–1930):

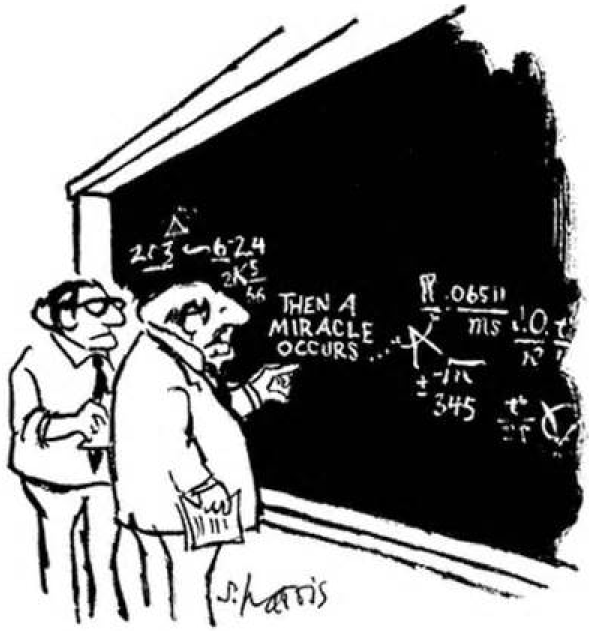
> *You know my method. It is founded upon the observation of trifles.*

4. Although rigorous mathematical proofs must be expressed in words and symbols rather than pictures, in many cases good (or even not so good) drawings are extremely helpful — sometimes even indispensable — sources for insights which can suggest an approach to proving a mathematical statement. **This is particularly true for courses with substantial amounts of geometric content like Mathematics 205A.** The following two quotations reflect this point:

> *Geometry is the science of correct reasoning [based] on incorrect figures.* — G. Polyá (1887–1985)
>
> *For me, following a geometrical argument purely logically, without a picture for it constantly in front of me, is impossible.* — F. Klein (1849–1925)

5. The first step in putting together a logical argument is to come up with something that appears to be correct, but one important test of an argument's clarity and validity is to *refine the ideas so that they become equally clear and convincing to someone else who is reasonably well-informed about the subject.* This generally requires healthy doses of skepticism and self-criticism; these can certainly be overdone, but initially most of us need to be more concerned about not going far enough.

The next page contains two well-known satirical cartoons by Sidney Harris (1933–) on convincing arguments. We should also note that that there are several anthologies of Harris' excellent cartoons on humorous aspects of scientists and their work, and they are definitely worth reading and viewing.



"I think you should be more explicit here in step two."

SOLUTIONS TO EXERCISES FOR

MATHEMATICS 205A — Part 1

Fall 2014

I. Foundational material

I.1 : Basic set theory

Problems from Munkres, § 9*, p.* 64

2. (a) – (c) For each of the first three parts, choose a 1–1 correspondence between the integers or the rationals and the positive integers, and consider the well-orderings that the latter inherit from these maps. For each nonempty subset, define the choice function to be the first element of that subset with respect to the given ordering.∎

TECHNICAL FOOTNOTE. (*This uses material from a graduate level measure theory course.*) In Part (d) of the preceding problem, one cannot find a choice function without assuming something like the Axiom of Choice. The following explanation goes beyond the content of this course but is hopefully illuminating. The first step involves the results from Section I.3 which show that the set of all functions from $\{0, 1\}$ to the nonnegative integers is in 1–1 correspondence with the real numbers. If one could construct a choice function over all nonempty subsets of the real numbers, then among other things one can prove that that there is a subset of the reals which is not Lebesgue measurable without using the Axiom of Choice (see any graduate level book on Lebesgue integration; for example, Section 3.4 of Royden). On the other hand, there are models for set theory in which every subset of the real numbers is Lebesgue measurable (see R. Solovay, *A model of set theory in which every set of reals is Lebesgue measurable*, Annals of Math. (2) **92** (1970), pp. 1–56). — It follows that one cannot expect to have a choice function for arbitrary families of nonempty subsets of the reals unless one makes some extra assumption related to the Axiom of Choice.

5. (a) For each $b \in B$ pick $h(b) \in A$ such that $f(a) = b$; we can find these elements by applying the Axiom of Choice to the family of subsets $f^{-1}[\{b\}]$ because surjectivity implies each of these subsets is nonempty. It follows immediately that $b = f(\,h(b)\,)$.∎

(b) For each $x \in A$ define a function $g_x : B \to A$ whose graph consists of all points of the form

$$\big(\,f(a),\, a\,\big) \quad \in \quad B \times A$$

togther with all points of the form (b, x) if b does not lie in the image of A. The injectivity of f implies that this subset is the graph of some function g_x, and by construction we have $g_x \circ f(a) = a$ for all $a \in A$. This does NOT require the Axiom of Choice; for each $x \in A$ we have constructed an EXPLICIT left inverse to f. — On the other hand, if we had simply said that one should pick some element of A for each element of $B - f[A]$, then we WOULD have been using the Axiom of Choice.∎

1. The commutativity law for $\oplus$ holds because

$$B \oplus A = (B - A) \ \cup \ (A - B)$$

by definition and the commutativity of the set-theoretic union operation. The identity $A \oplus A = \emptyset$ follows because

$$A \oplus A = (A - A) \ \cup \ (A - A) = \emptyset$$

and $A \oplus \emptyset = A$ because

$$A \oplus \emptyset = (A - \emptyset) \ \cup \ (\emptyset - A) = A \ \cup \ \emptyset = A \ .$$

In order to handle the remaining associative and distributive identities it is necessary to write things out explicitly, using the fact that every Boolean expression involving a finite list of subsets can be written as a union of intersections of subsets from the list. It will be useful to introduce some algebraic notation in order to make the necessary manipulations more transparent. Let $X \supset A \cup B \cup C$ denote the complement of $Y \subset X$ by $\widehat{Y}$ (or by $Y^{\frown}$ if Y is some compound algebraic expression), and write $P \cap Q$ simply as PQ. Then the symmetric difference can be rewritten in the form $(A\widehat{B}) \cup (B\widehat{A})$. It then follows that

$$(A \oplus B) \oplus C = \left(A\widehat{B} \cup B\widehat{A}\right) \widehat{C} \ \cup \ C \left(A\widehat{B} \cup B\widehat{A}\right)^{\frown} =$$

$$A\widehat{B}\widehat{C} \cup B\widehat{A}\widehat{C} \cup C\left((\widehat{A} \cup B)(\widehat{B} \cup A)\right) = A\widehat{B}\widehat{C} \cup B\widehat{A}\widehat{C} \cup C\left(\widehat{A}\widehat{B} \cup AB\right) =$$

$$A\widehat{B}\widehat{C} \cup \widehat{A}B\widehat{C} \cup \widehat{A}\widehat{B}C \cup ABC \ .$$

Similarly, we have

$$A \oplus (B \oplus C) = A\left(B\widehat{C} \cup C\widehat{B}\right)^{\frown} \cup \left(B\widehat{C} \cup C\widehat{B}\right)\widehat{A} =$$

$$A\left((\widehat{B} \cup C)(\widehat{C} \cup B)\right) \cup B\widehat{C}\widehat{A} \cup C\widehat{B}\widehat{A} = A(\widehat{B}\widehat{C} \cup BC) \cup B\widehat{C}\widehat{A} \cup C\widehat{B}\widehat{A} =$$

$$A\widehat{B}\widehat{C} \cup \widehat{A}B\widehat{C} \cup \widehat{A}\widehat{B}C \cup ABC \ .$$

This proves the associativity of $\oplus$ because both expressions are equal to the last expression displayed above. The proof for distributivity is similar but shorter (the left side of the desired equation has only one $\oplus$ rather than two, and we only need to deal with monomials of degree 2 rather than 3):

$$A\,(B \oplus C) = A(B\widehat{C} \cup C\widehat{B}) = AB\widehat{C} \cup A\widehat{B}C$$

$$AB \oplus AC = (AB)(AC)^{\frown} \cup (AC)(AB)^{\frown} =$$

$$\left(AB(\widehat{A} \cup \widehat{C})\right) \cup \left(AC(\widehat{A} \cup \widehat{B})\right) = AB\widehat{C} \cup A\widehat{B}C$$

Thus we have shown that both of the terms in the distributive law are equal to the same set.■

I.2 : Products, relations and functions

Problem from Munkres, § 4, p. 44

4. (*a*) This is outlined in the course notes.

Additional exercises

1. (*i*) Suppose that (x, y) lies in the left hand side. Then $x \in A$ and $y \in B \cap D$. Since the latter means $y \in B$ and $y \in D$, this means that

$$(x, y) \in (A \times B) \cap (A \times D) \ .$$

Now suppose that (x, y) lies in the set displayed on the previous line. Since $(x, y) \in A \times B$ we have $x \in A$ and $y \in B$, and similarly since $(x, y) \in A \times D$ we have $x \in A$ and $y \in D$. Therefore we have $x \in A$ and $y \in B \cap D$, so that $(x, y) \in A \times (B \cap D)$. Thus every member of the the first set is a member of the second set and vice versa, and therefore the two sets are equal.■

(*ii*) Suppose that (x, y) lies in the left hand side. Then $x \in A$ and $y \in B \cup D$. If $y \in B$ then $(x, y) \in A \times B$, and if $y \in D$ then $(x, y) \in A \times D$; in either case we have

$$(x, y) \in (A \times B) \cup (A \times D) \ .$$

Now suppose that (x, y) lies in the set displayed on the previous line. If $(x, y) \in A \times B$ then $x \in A$ and $y \in B$, while if $(x, y) \in A \times D$ then $x \in A$ and $y \in D$. In either case we have $x \in A$ and $y \in B \cup D$, so that $(x, y) \in A \times (B \cap D)$. Thus every member of the the first set is a member of the second set and vice versa, and therefore the two sets are equal.■

(*iii*) Suppose that (x, y) lies in the left hand side. Then $x \in A$ and $y \in Y - D$. Since $y \in Y$ we have $(x, y) \in A \times Y$, and since $y \notin D$ we have $(x, y) \notin A \times D$. Therefore we have

$$A \times (Y - D) \ \subset \ (A \times Y) - (A \times D) \ .$$

Suppose now that $(x, y) \in (A \times Y) - (A \times D)$. These imply that $x \in A$ and $y \in Y$ but $(x, y) \notin A \times D$; since $x \in A$ the latter can only be true if $y \notin D$. Therefore we have that $x \in A$ and $y \in Y - D$, so that

$$A \times (Y - D) \ \supset \ (A \times Y) - (A \times D) \ .$$

This proves that the two sets are equal.■

(*iv*) Suppose that (x, y) lies in the left hand side. Then we have $x \in A$ and $y \in B$, and we also have $x \in C$ and $y \in D$. The first and third of these imply $x \in A \cap C$, while the second and fourth imply $y \in B \cap D$. Therefore $(x, y) \in (A \cap C) \times (B \cap D)$ so that

$$(A \times B) \cap (C \times D) \ \subset \ (A \cap C) \times (B \cap D) \ .$$

Suppose now that (x, y) lies in the set on the right hand side of the displayed equation. Then $x \in A \cap C$ and $y \in B \cap D$. Since $x \in A$ and $y \in B$ we have $(x, y) \in A \times B$, and likewise since $x \in C$ and $y \in D$ we have $(x, y) \in C \times D$, so that

$$(A \times B) \cap (C \times D) \ \supset \ (A \cap C) \times (B \cap D) \ .$$

Therefore the two sets under consideration are equal.■

(v) Suppose that (x, y) lies in the left hand side. Then either we have $x \in A$ and $y \in B$, or else we have $x \in C$ and $y \in D$. The first and third of these imply $x \in A \cup C$, while the second and fourth imply $y \in B \cup D$. Therefore (x, y) is a member of $(A \cup C) \times (B \cup D)$ so that

$$(A \times B) \cup (C \times D) \ \subset \ (A \cup C) \times (B \cup D) \ .$$

Supplementary note: To see that the sets are not necessarily equal, consider what happens if $A \cap C = B \cap D = \emptyset$ but all of the four sets A, B, C, D are nonempty. Try drawing a picture in the plane to visualize this.■

(vi) Suppose that (x, y) lies in the left hand side. Then $x \in X$ and $y \in Y$ but $(x, y) \notin A \times B$. The latter means that the statement

$$x \in A \textbf{ and } y \in B$$

is false, which is logically equivalent to the statement

$$\textbf{either } x \notin A \textbf{ or } y \notin B \ .$$

If $x \notin A$, then it follows that $(x, y) \in \big((X - A) \times Y\big)$, while if $y \notin B$ then it follows that $(x, y) \in \big(X \times (Y - B)\big)$. Therefore we have

$$\big(X \times Y\big) - \big(A \times B\big) \ \subset \ \big(X \times (Y - B)\big) \cup \big((X - A) \times Y\big) \ .$$

Suppose now that (x, y) lies in the set on the right hand side of the containment relation on the displayed line. Then we have $(x, y) \in X \times Y$ and also

$$\textbf{either } x \notin A \textbf{ or } y \notin B \ .$$

The latter is logically equivalent to

$$x \in A \textbf{ and } y \in B$$

and this in turn means that $(x, y) \notin A \times B$ and hence proves the reverse inclusion of sets.■

2. Since the object of this exercise is to address ambiguities in notation, the key to solving this exercise is to find unambiguous ways of distinguishing between the two interpretations of the symbolism $f^{-1}[C]$. The formulation in the exercise goes part way towards doing this, but one must still be careful.

We shall systematically use the inverse function identities $g \circ f = \mathrm{id}_A$ and $f \circ g = \mathrm{id}_B$ which relate a 1–1 onto function $f : A \to B$ with its inverse function $g : B \to A$. As in the statement of the exercise, let $C \subset B$.

Suppose that $a \in IMAGE\ g[C]$. Then $a = g(b)$ for some $b \in C$, which implies that $f(a) = f(g(b)) = b$, so that $f(a) \in C$ and hence $a \in INVERSE\ IMAGE\ f^{-1}[C]$. Conversely, suppose that $a \in INVERSE\ IMAGE\ f^{-1}[C]$. Then $f(a) \in C$, so that $g(f(a)) = a \in C$, which means that $a \in IMAGE\ g[C]$. Therefore each of the sets $IMAGE\ g[C]$ and $INVERSE\ IMAGE\ f^{-1}[C]$ is contained in the other, so the two sets are equal.■

3. We shall first prove that $\mathcal{R}^\#$ is an equivalence relation.

The relation is reflexive. The definition of $\mathcal{R}^\#$ stipulates that $x\ \mathcal{R}^\#\ x$.

The relation is symmetric. By the preceding step we need only consider the case where $x \neq y$. If we are given a finite sequence $\{v_0\ \cdots\ , v_m\}$ as described in the definition such that $v_0 = x$ and $v_m = y$, then the reverse sequence with $w_j = v_{m-j}$ satisfies the criterion in the definition implies with $w_0 = y$ and $w_m = x$.

The relation is transitive. Suppose that we have a finite sequence $\{v_0 \ \cdots \ , v_m\}$ as described in the definition such that $v_0 = x$ and $v_m = y$, and another finite sequence $\{u_0 \ \cdots \ , u_n\}$ as described in the definition such that $u_0 = y$ and $u_n = z$, then we can concatenate (string together) the original sequences into a new sequence $\{t_0 \ \cdots \ , t_{m+n}\}$ such that $t_j = v_j$ if $j \leq m$ and $t_j = u_{j-m}$ if $j \geq m$ (the formulas are consistent at the overlapping value m, for which $t_m = y$). This new sequence still satisfies the criterion in the definition with $t_0 = x$ and $t_{m+n} = z$.

To complete the proof, we need to verify that if $\mathcal{S}$ is an equivalence relation such that $x \ \mathcal{S} \ y$ whenever $x \ \mathcal{R} \ y$, then $u \ \mathcal{R}^{\#} \ v$ implies that $u \ \mathcal{S} \ v$. If $u = v$ then the latter follows because we are working with equivalence relations, which are reflexive. If $u \neq v$ and $u \ \mathcal{S} \ v$, let $\{t_0 \ \cdots \ , t_m\}$ be a sequence starting with u and ending with v such that the for each i we have either $t_i \ \mathcal{R} \ t_{i+1}$ or $t_{i+1} \ \mathcal{R} \ t_i$. Then by the hypotheses on $\mathcal{S}$ we know that $t_i \ \mathcal{S} \ t_{i+1}$ for all i, and therefore by repeated application of the transitivity property we have $u \ \mathcal{S} \ v$.■

4. We shall refer to the file `math205Asolutions01a.pdf` for drawings which may help explain the underlying ideas; as usual, the proof must be written so that it does not formally depend upon these drawings.

The first step is to show that if $(i, j) \in \mathcal{E}$, then every point of the form $(i + t, j + t)$ in B — where t runs through all admissible integers such that the point in question belongs to B — also lies in $\mathcal{E}$. In other words, if $i' - j' = i - j$, then $(i', j') \ \mathcal{E} \ (i, j)$. For points in B the difference values $i - j$ are the 15 integers between ± 7, so this shows that there are at most 15 equivalence classes (in the first drawing, the squares with $i - j \ = \ $ CONSTANT are on the diagonal lines and have the same color). To prove the assertion in the first sentence, observe that $(i, j) \ \mathcal{R} \ (i + \varepsilon, j + \varepsilon)$ for $\varepsilon = \ \pm 1$ by definition, and by definition of $\mathcal{E}$ this yields $(i, j) \ \mathcal{E} \ (i + \varepsilon, j + \varepsilon)$. The statement for general values of t now follows by repeated application of the final assertion in the previous sentence and the transitivity of $\mathcal{E}$.

Next, let $\mathcal{F}$ be the binary relation with $(i', j') \ \mathcal{F} \ (i, j)$ if $i' + j'$ and $i + j$ are both even or both odd. This is an equivalence relation by one of the exercises in Sutherland and the fact that two ordered pairs are $\mathcal{F}$ are related if and only if they have the same values under the function $\varphi : B \to \{\text{EVEN}, \text{ODD}\}$ whose value is determined by whether $i+j$ is even or odd. The definition of $\mathcal{R}$ implies that if $(i, j) \ \mathcal{R} \ (p, q)$ then both $i+j$ and $p+q$ are even or odd, and therefore $(i, j) \ \mathcal{F} \ (p, q)$ whenever $(i, j) \ \mathcal{R} \ (p, q)$. By Exercise 0, it follows that $(i, j) \ \mathcal{F} \ (p, q)$ whenever $(i, j) \ \mathcal{E} \ (p, q)$, and since $\mathcal{F}$ has two equivalence classes the equivalence relation $\mathcal{E}$ must also have at least two equivalence classes.

Finally, we need to show that $\mathcal{E}$ has exactly two equivalence classes. The idea is similar to that of the first step; namely, if $(i, j) \in \mathcal{E}$, then every point of the form $(i + t, j - t)$ in B — where t runs through all admissible integers such that the point in question belongs to B — also lies in $\mathcal{E}$. The main difference in the argument is the need to observe that we also have $(i, j) \ \mathcal{R} \ (i + \varepsilon, j - \varepsilon)$ for $\varepsilon = \ \pm 1$ by the definition of $\mathcal{R}$. By the same reasoning as in the first step, this implies that if $i' + j' = i + j$, then $(i', j') \ \mathcal{E} \ (i, j)$. — To conclude the argument, it suffices to observe that the set of all $(i, j) \in B$ with $i + j = 9$ the difference $i - j$ takes all odd values between -7 and $+7$, while the set of all (i, j) with $i + j = 8$ takes all even values between -6 and $+6$ (in the second drawing, observe how the two lines with slope -1 cut through all the lines with slope $+1$). This proves that there are at most two equivalence classes for $\mathcal{E}$, and by the preceding paragraph there must be precisely two equivalence classes.■

5. It turns out that, in order to make things less repetitive, the best place to start is by observing that if $[x] = [y]$ then $x \mathcal{S} y$. This follows from the reflexive property of the equivalence relation $\mathcal{R}_2$. Note that this also yields the reflexive property for $\mathcal{S}$.

Suppose now that $x \mathcal{S} y$, so that $[x] \mathcal{R}_2 [y]$. Since $\mathcal{R}_2$ is an equivalence relation, this means that $[y] \mathcal{R}_2 [x]$, which in turn implies that $y \mathcal{S} x$. Finally, suppose that $x \mathcal{S} y$ and $y \mathcal{S} z$, so that $[x] \mathcal{R}_2 [y]$ and $[y] \mathcal{R}_2 [z]$. Since $\mathcal{R}_2$ is transitive we have $[x] \mathcal{R}_2 [z]$, and this yields $x \mathcal{S} z$, so that $\mathcal{S}$ is an equivalence relation on X.■

I.3 : Cardinal numbers

Problem from Munkres, § 7, p. 51

4. (a) Let $\mathbb{Q}[t]$ denote the ring of polynomials with rational coefficients, and for each integer $d > 0$ let $\mathbb{Q}[t]_d$ denote the set of polynomials with degree equal to d. There is a natural identification of $\mathbb{Q}[t]_d$ with the subset of $\mathbb{Q}^{d+1}$ consisting of n-tuples whose last coordinate is nonzero, and therefore $\mathbb{Q}[t]_d$ is countable. Since a countable union of countable sets is countable (Munkres, Theorem 7.5, pp. 48–49), it follows that $\mathbb{Q}[t]$ is also countable.

Given an algebraic number α, there is a unique **monic** rational polynomial $p(t)$ of least (positive) degree such that $p(\alpha) = 0$ (the existence of a polynomial of least degree follows from the well-ordering of the positive integers, and one can find a monic polynomial using division by a positive constant; uniqueness follows because if p_1 and p_2 both satisfy the condition then $p_1 - p_2$ is either zero or a polynomial of lower degree which has α as a root). Let p_α be the polynomial associated to α in this fashion. Then p may be viewed as a function from the set $\mathcal{A}$ of algebraic numbers into $\mathbb{Q}[t]$; if f is an arbitrary element of degree $d \geq 0$, then we know that there are at most d elements of $\mathcal{A}$ that can map to p (and if $f = 0$ the inverse image of $\{f\}$ is empty). Letting $\mathcal{A}_f$ be the inverse image of f, we see that $\mathcal{A} = \cup_f \mathcal{A}_f$, so that the left hand side is a countable union of finite sets and therefore is countable.■

(b) Since every real number is either algebraic or transcendental but not both, we clearly have

$$2^{\aleph_0} = |\mathbb{R}| = |\text{algebraic}| + |\text{transcendental}| \ .$$

We know that the algebraic numbers are countable, so if the transcendental numbers are also countable the right hand side of this equation reduces to $\aleph_0 + \aleph_0$, which is equal to $\aleph_0$, a contradiction. Therefore the set of transcendental numbers is uncountable (in fact, its cardinality is $2^{\aleph_0}$ but the problem did not ask for us to go any further).■

Problem from Munkres, § 9, p. 62

5. (a) For each $b \in B$ let $L_b \subset A$ be the inverse image $f^{-1}(\{b\})$. Using the axiom of choice we can find a function g that assigns to each set L_b a point $g^*(L_b) \in L_b$. Define $g(b) = g^*(L_b)$; by construction we have that $g(b) \in f^{-1}(\{b\})$ so that $f(g(b)) = b$. This means that $f \circ g = \mathrm{id}_B$ and that g is a right inverse to f.■

(b) Given an element $z \in A$ define a map $g_z : B \to A$ as follows: If $b = f(a)$ for some a let $g_z(b) = a$. This definition is unambiguous because there is at most one $a \in A$ such that $f(a) = b$. If b does not lie in the image of f, set $g_z(b) = z$. By definition we then have $g_z(f(a)) = a$ for all $a \in A$, so that $g_z \circ f = \mathrm{id}_A$ and g_z is a left inverse to f. Did this use the Axiom of Choice? No.

What we actually showed was that for each point of A there is an associated left inverse. However, if we had simply said, "pick some point $z_0 \in A$ and define g using z_0," then we would have used the Axiom of Choice.■

Problem from Munkres, § 11, p. 72

8. (*a*) Note first that $\beta \notin A$ for otherwise it would be a linear combination of elements in A for trivial reasons.

Suppose the set in question is not linearly independent; then some finite subset C is not linearly independent, and we may as well add β to that subset. It follows that there is a relation

$$x_\beta \beta + \sum_{\gamma \in A \cap C} x_\gamma \gamma = 0$$

where not all of the coefficients x_β or x_γ are equal to zero. In fact, we must have $x_\beta \neq 0$ for otherwise there would be some nontrivial linear dependence relationship in $A \cap C$, contradicting our original assumption on A. However, if $x_\beta \neq 0$ then we can solve for β to express it as a linear combination of the vectors in $A \cap C$, and this contradicts our assumption on β. Therefore the set in question must be linearly independent.■

(*b*) Let $\mathbf{X}$ be the partially ordered set of linearly independent subsets of V, with inclusion as the partial ordering. In order to apply Zorn's Lemma we need to know that an arbitrary linearly ordered subset $\mathbf{L} \subset \mathbf{X}$ has an upper bound in in $\mathbf{X}$. Suppose that $\mathbf{L}$ consists of the subsets A_t; it will suffice to show that the union $A = \cup_t A_t$ is linearly independent, for then A will be the desired upper bound.

We need to show that if C is a finite subset of A then C is linearly independent. Since each A_t is linearly independent, it suffices to show that there is some r such that $C \subset A_r$, and we do this by induction on $|C|$. If $|C| = 1$ this is clear because $\alpha \in A$ implies $\alpha \in A_t$ for some t. Suppose we know the result when $|C| = k$, and let $D \subset A$ satisfy $|D| = k + 1$. Write $D = D_0 \cup \gamma$ where $\gamma \notin D_0$. Then there is some u such that $D_0 \subset A_u$ and some v such that $\gamma \in A_v$. Since $\mathbf{L}$ is linearly ordered we know that either $A_u \subset A_v$ or vice versa; in either case we know that D is contained in one of the sets A_u or A_v. This completes the inductive step, which in turn implies that A is linearly independent and we can apply Zorn's Lemma.■(*c*) Let A be a maximal element of $\mathbf{X}$ whose existence was guaranteed by the preceding step in this exercise. We claim that every vector in V is a linear combination of vectors in A. If this were not the case and β was a vector that could not be expressed in this fashion, then by the first step of the exercise the set $A \cup \{\beta\}$ would be linearly independent, contradicting the maximality of A.■

Additional exercises

1. Let X be the set in question, and let $Y \subset X$ be the subset of all one point subsets. Since there is a 1–1 correspondence between $\mathbb{R}$ and Y it follows that $2^{\aleph_0} = |\mathbb{R}| = |Y| \leq |X|$. Now write X as a union of subfamilies X_n where $0 \leq n \leq \infty$ such that the cardinality of every set in X_n is n if $n < \infty$ and the cardinality of every set in X_∞ is $\aleph_0$.

Suppose now that $n < \infty$. Then X_n is in 1–1 correspondence with the set of all points $(x_1, \cdots, x_n)$ in $\mathbb{R}^n$ such that $x_1 < \cdots < x_n$ (we are simply putting the points of the subset in order). Therefore $|X_0| = 1$ and $|X_n| \leq 2^{\aleph_0}$ for $1 \leq n < \infty$, and it follows that $\cup_{n<\infty} X_n$ has cardinality at most

$$\aleph_0 \cdot 2^{\aleph_0} \;\leq\; 2^{\aleph_0} \cdot 2^{\aleph_0} \;=\; 2^{\aleph_0} \; .$$

So what can we say about the cardinality of X_∞? Let S be the set of all infinite sequences in $\mathbb{R}$ indexed by the positive integers. For each choice of a 1–1 correspondence between an element of X_∞ and $\mathbb{N}^+$ we obtain an element of S, and if we choose one correspondence for each element we obtain a 1–1 map from X_∞ into S. By definition $|S|$ is equal to $\left(2^{\aleph_0}\right)^{\aleph_0}$, which in turn is equal to $2^{\aleph_0 \times \aleph_0} = 2^{\aleph_0}$; therefore we have $|X_\infty| \leq 2^{\aleph_0}$. Putting everything together we have

$$|X| = |\cup_{n<\infty} \ X_n| + |X_\infty| \leq 2^{\aleph_0} + 2^{\aleph_0} = 2^{\aleph_0}$$

and since we have already established the reverse inequality it follows that $|X| = 2^{\aleph_0}$ as claimed.■

IMPORTANT FOOTNOTE. The preceding exercise relies on the generalization of the law of exponents for cardinal numbers

$$\gamma^{\alpha\beta} = \left(\gamma^\alpha\right)^\beta$$

that was stated at the end of Section I.3 of the course notes without proof. For the sake of completeness we shall include a proof.

Choose sets A, B, C so that $|A| = \beta$, $|B| = \alpha$ (note the switch!!) and $|C| = \gamma$, and let $\mathbf{F}(S,T)$ be the set of all (set-theoretic) functions from one set S to another set T. With this terminology the proof of the cardinal number equation reduces to finding a 1–1 correspondence

$$\mathbf{F}(A \times B, C) \longleftrightarrow \mathbf{F}\big(A, \ \mathbf{F}(B,C) \ \big) \ .$$

In other words, we need to construct a 1–1 correspondence between functions $A \times B \to C$ and functions $A \to \mathbf{F}(B,C)$. In the language of category theory this is an example of an *adjoint functor* relationship.

Given $f : A \times B \to C$, construct $f^* : A \to \mathbf{F}(B,C)$ by defining $f^*(a) : B \to C$ using the formula

$$\left[f^*(a)\right](b) = f(a,b) \ .$$

This construction is onto, for if we are given $h^* : A \to \mathbf{F}(B,C)$ and we define $f : A \times B \to C$ by the formula

$$f(a,b) = \left[h(a)\right](b)$$

then $f^* = h$ by construction; in detail, one needs to check that $f^*(a) = h(a)$ for all $a \in A$, which amounts to checking that $[f^*(a)](b) = [h(a)](b)$ for all a and b — but both sides of this equation are equal to $f(a,b)$. To see that the construction is 1–1, note that $f^* = g^* \iff f^*(a) = g*^(a)$ for all a, which is equivalent to $[f^*(a)](b) = [g^*(a)](b)$ for all a and b, which in turn is equivalent to $f(a,b) = g(a,b)$ for all a and b, which is equivalent to $f = g$. Therefore the construction sending f to f^* is 1–1 onto as required.■

For the record, the other exponential law

$$\left(\beta \cdot \gamma\right)^\alpha = \beta^\alpha \cdot \gamma^\alpha$$

may be verified by first noting that it reduces to finding a 1–1 correspondence between $\mathbf{F}(A, B \times C)$ and

$$\mathbf{F}(A,B) \times \mathbf{F}(A,C) \ .$$

This simply reflects the fact that a function $f : A \to B \times C$ is completely determined by the ordered pair of functions $p_B \circ f$ and $p_C \circ f$ where p_B and p_C are the coordinate projections from $B \times C$ to B and C respectively.■

2. The inequality means that there is a 1–1 mapping j from some set A_0 with cardinality α to a set B with cardinality β. Since the cardinality of X equals β it follows that there is a 1–1 correspondence $f : B \to X$. If we take $A = j(f(A_)))$, then $A \subset X$ and $|A| = \alpha$.■

3. We shall prove this using Zorn's Lemma (also known as the Kuratowski-Zorn Lemma) and the partial ordering of $\mathcal{F}$ by inclusion. It suffices to show that linearly ordered subfamilies of $\mathcal{F}$ have upper bounds in $\mathcal{F}$.

Suppose that $\mathcal{L} \subset \mathcal{F}$ is linearly ordered. Define

$$L^* \quad = \quad \bigcup_{L_\alpha \in \mathcal{L}} L_\alpha .$$

Clearly $L^* \supset L_\alpha$ for all α, so it is only necessary to prove that $L^* \in \mathcal{F}$. Suppose that $A = \{a_1, \ \cdots \ , a_k\}$ is a finite subset of L^*, and for $1 \leq j \leq k$ choose $\alpha(j)$ such that $a_j \in L_{\alpha(j)}$. Since $\mathcal{L}$ is linearly ordered there is some m such that $1 \leq m \leq k$ and $L_{\alpha(j)} \subset L_{\alpha(m)}$ for all m. The latter implies that $A \subset L_{\alpha(m)}$. Since $L_{\alpha(m)} \in \mathcal{F}$ the first defining property of finite character implies that $A \in \mathcal{F}$, and since $\mathcal{F}$ has finite character the second defining property implies that $L^* \in \mathcal{F}$, which is what we wanted to prove.■

I.4 : The real number system

Problem from Munkres, § 4, *p.* 35

9. (c) We shall follow the hint. Part (b) states that if a real number x is not an integer, then there is a unique integer n such that $n < x < n+1$.

The solution to (c) has two cases depending upon whether or not y is an integer. If $y \in \mathbb{Z}$, then the same is true for $y+1$ and we have

$$y \quad < \quad y \; + \; 1 \quad < \quad y \; + \; (y-x) \quad = \quad x$$

so that $y+1$ is an integer with the required properties. Suppose now that $y \notin \mathbb{Z}$, so that there is a unique integer m such that $m < y < m+1$. We then have

$$x \quad = \quad y \; + \; (x-y) \;\; m \; + \; 1$$

so that $y < m+1 < x$.■

Remark. Of course, the integer n in the conclusion of the preceding exercise is not necessarily unique; for example if we have the stronger inequality $x - y \geq 2$ then there are at least two integers between x and y.

Additional exercise

1. The answer is emphatically **NO,** and there are many counterexamples. Let $\{ x_\alpha \}$ be a basis for $\mathbb{R}$ as a vector space over $\mathbb{Q}$. If $f : \mathbb{R} \to \mathbb{R}$ is a $\mathbb{Q}$-linear map, then f satisfies the condition in the problem. Thus it is only necessary to find examples of such maps that are not multiplication by a constant. Since $\mathbb{R}$ is uncountable, a basis for it over the rationals contains infinitely many elements. Pick one element x_0 in the basis, and consider the unique $\mathbb{Q}$-linear transformation f which sends

x_0 to itself and all other basis vectors to zero. Then f is nonzero but is neither 1–1 nor onto. In contrast, a mapping of the form $T_c(x) = c \cdot x$ for some fixed real number c is 1–1 and onto if $c \neq 0$ and zero if $c = 0$. Therefore there is no c such that $f = T_c$.■

Generalization – more difficult to verify. In fact, the cardinality of the set of all mappings f with the given properties is $2^{\mathbf{c}}$, where as usual $\mathbf{c} = 2^{\aleph_0}$ (the same as the cardinality of all maps from $\mathbb{R}$ to itself). To see this, first note that the statement in parentheses follows because the cardinality of the set of all such mappings is

$$\mathbf{c}^{\mathbf{c}} \quad = \quad \left(2^{\aleph_0}\right)^{\mathbf{c}} \quad = \quad 2^{\aleph_0 \cdot \mathbf{c}} \quad = \quad 2^{\mathbf{c}} \ .$$

To prove the converse, we first claim that a basis for $\mathbb{R}$ over $\mathbb{Q}$ must contain $\mathbf{c}$ elements. Note that the definition of vector space basis implies that $\mathbb{R}$ is in 1–1 correspondence with the set of finitely supported functions from B to $\mathbb{Q}$, where B is a basis for $\mathbb{R}$ over $\mathbb{Q}$ and finite support means that the coordinate functions are nonzero for all but finitely many basis elements. Thus if β is the cardinality of B, then the means that the cardinality of $\mathbb{R}$ is β (the details of checking this are left to the reader). For each subset A of B we may define a $\mathbb{Q}$-linear map from $\mathbb{R}$ to itself which sends elements of B to themselves and elements of the difference set $B - A$ to zero. Different subsets determine different mappings, so this shows that the set of all f satisfying the given condition has cardinality at least $2^{\mathbf{c}}$. Since the first part of the proof shows that the cardinality is at most $2^{\mathbf{c}}$, this completes the argument.■

Remark. By the results of Unit II, if one also assumes that the function f is **continuous,** then the answer becomes **YES.** One can use the material from Unit II to prove this quickly as follows: If $r = f(1)$, then by induction and $f(-x) = -f(x)$ implies that $f(a) = r \cdot a$ for all $a \in \mathbb{Z}$; if $a = p/q \in \mathbb{Q}$, then we have

$$q \cdot f(a) \quad = \quad f(q \cdot a) \quad = \quad f(p) \quad = \quad r\,p \ , \qquad \text{yielding} \qquad f(a) \quad = \quad r \cdot \frac{p}{q}$$

so that $f(a) = r \cdot a$ for all $a \in \mathbb{Q}$. By the results of Section II.4, if f and g are two continuous functions such that $f(a) = g(a)$ for all rational numbers a, then $f = g$. Taking g to be multiplication by r, we conclude that f must also be multiplication by c.■

2. Follow the hint. If A has an upper bound and β is given as in the hint, then by construction β is an upper bound for A, and it is the least such upper bound.■

3. Since $n(k)$ is strictly increasing we must have $n(k) \geq k$ for all k (Proof by induction: $n(0) \geq 0$, and if $n(m) \geq n$ then $n(m+1) >\geq n(m) + 1 \geq m + 1$ because $n(k)$ is strictly increasing). By the definition of a limit for a sequence, for every $\varepsilon > 0$ there is a positive integer N such that $n \geq N$ implies $|a_n - L| < \varepsilon$. Therefore if $k \geq N$ then $n(k) \geq n$, so that $|a_{n(k)} - L| < \varepsilon$.■

NOTE. The contrapositive forms of this result are particularly useful for showing that a specific sequence has no limit. If either (a) there is a subsequence with no limit, or (b) there are two subsequences with different limits, then the original sequence cannot have a limit.■

4. (i) The sequence of left hand endpoints $\{a_n\}$ is bounded from above by b_k, where k is an arbitrary nonnegative integer, while the sequence of right hand endpoints $\{b_n\}$ is bounded from below by a_k, where k is an arbitrary nonnegative integer. Therefore the set of left hand endpoints has a least upper bound A, and $A \leq b_k$ for all k. Similarly, the set of right hand endpoints has a greatest lower bound B, and by the preceding sentence we must have $A \leq B$. Therefore every point p satisfying $A \leq p \leq B$ must lie in each of the intervals. — Note that $A = B$ is possible; for example, consider the intervals $\left[2^{-k}, 2^k\right]$.

(*ii*) The answer is definitely **NO**. We shall give a counterexample involving $\sqrt{2}$. Since there is always a rational number between two real numbers, for each n we can find rational number a_n and b_n such that

$$\sqrt{2} \; - \; \left(\tfrac{1}{2}\right)^n \; < \; a_n \; < \; \sqrt{2} \; < \; b_n \; < \; \sqrt{2} \; + \; \left(\tfrac{1}{2}\right)^n \; .$$

It follows that if $p \in [a_n, b_n]$ then $|\sqrt{2} - p| \leq \left(\frac{1}{2}\right)^n$, so if p lies in each interval then this inequality holds for every n. But the latter implies that $|p - \sqrt{2}| = 0$, so that $p = \sqrt{2}$ and the latter is the only point which lies on each interval. Since $\sqrt{2}$ is irrational, there is no rational number which lies on all of the intervals.■

II. Metric and topological spaces

II.1 : Metrics and topologies

Problem from Munkres, § 13, *p.* 83

3. X lies in the family because $X - X = \emptyset$ and the latter is finite, while $\emptyset$ lies in the family because $X - \emptyset = X$. Suppose U_α lies in the family for all $\alpha \in A$. To determine whether their union lies in the family we need to consider the complement of that union, which is

$$X - \bigcup_\alpha U_\alpha \quad = \quad \bigcap_\alpha X - U_\alpha \ .$$

Each of the sets in the intersection on the right hand side is either countable or all of X. If at least one of the sets is countable then the whole intersection is countable, and the only other alternative is if each set is all of X, in which case the intersection is X. In either case the complement satisfies the condition needed for the union to belong to $\mathbf{T}_c$. Suppose now that we have two sets U_1 and U_2 in the family. To decide whether their intersection lies in the family we must again consider the complement of $U_1 \cap U_2$, which is

$$(X - U_1) \cup (X - U_2) \ .$$

If one of the two complements in the union is equal to X, then the union itself is equal to X, while if neither is equal to X then both are countable and hence their union is countable. In either case the complement of $U_1 \cap U_2$ satisfies one of the conditions under which $U_1 \cap U_2$ belongs to $\mathbf{T}_c$.

What about the other family? Certainly $\emptyset$ and X belong to it. What about unions? Suppose that X is an infinite set and that U and V lie in this family. Write $E = X - U$ and $F = X - V$; by assumption each of these subsets is either infinite or empty. Is the same true for their intersection? Of course not! Take X to be the positive integers, let E be all the even numbers and let F be all the prime numbers. Then E and F are infinite but the only number they have in common is 2. — Therefore the family $\mathbf{T}_\infty$ is not necessarily closed under unions and hence it does not necessarily define a topology for X.■

Problems from Munkres, § 16, *pp.* 91 – 92

1. We are given $A \subset Y \subset X$. Given $P \subset Q$ and a topology $\mathbf{T}$ on P, then $\mathbf{T}|Q$ will denote the subspace topology on Q, and with this notation the conclusion to be verified is that $\mathbf{T}|A = (\mathbf{T}|Y)\,|A$.

If U is open in $\mathbf{T}|A$, then $U = W \cap A$ where W is open in X. Since $A \subset Y$, we have $A = Y \cap A$ and hence $U = W \cap Y \cap A$, where $W \cap Y$ is open in Y by definition of the subspace topology. Therefore we have $\mathbf{T}|A \subset (\mathbf{T}|Y)\,|A$.

Conversely, a subset in $(\mathbf{T}|Y)\,|A$ has the form $V \cap A$ where V is open in Y. The latter condition translates into the statement $V = W \cap Y$ where W is open in X. Therefore we have $U = V \cap A = W \cap Y \cap A$ where W is open in X, and since $Y \cap A = A$ the latter means $U = W \cap A$ where W is open in X. Therefore we also have $\mathbf{T}|A \supset (\mathbf{T}|Y)\,|A$.

Finally, since each topology contains the other, they are equal.■**3.** A is open in both Y and $\mathbb{R}$. B is open in Y but not in $\mathbb{R}$. Neither C not D is open in either of Y or $\mathbb{R}$. E is the union of the open intervals

$$\left(\frac{1}{n+1}, \frac{1}{n}\right)$$

where n runs over all positive integers; this set is open in both Y and $\mathbb{R}$.■

0. (i) If X is a finite set with n elements there are 2^n subsets of X and hence 2^{2^n} families of subsets of X. Since a topology on X is a family of subsets, this means there are at most 2^{2^n} topologies on X.■

(ii) Since the empty set and $\{0,1\}$ are in every topology $\mathbf{T}$ on $\{0,1\}$, the problem reduces to determining which subsets with one element belong to $\mathbf{T}$. The extreme cases — no such subsets in $\mathbf{T}$ or both $\{0\}$ and $\{1\}$ in $\mathbf{T}$ — correspond to the indiscrete and discete topologies respectively. There are two remaining posssibilities in which exactly one of $\{0\}$ and $\{1\}$ belongs to $\mathbf{T}$. Since we are dealing with finite sets, the test for a topology is whether the family of nonempty proper subsets is closed under twofold unions and intersections (twofold + induction $\Leftarrow$ finite unions and intersections, and every union of subsets must be a finite union because there are only finitely many subsets). But each of the families whose nonempty proper subsets are just $\{0\}$ and $\{1\}$ is closed under union and intersection ($A \cup A = A = A \cap A$), so this means that there are FOUR topologies — the discrete and indiscrete topologies, and two more topologies which contain exactly one subset with exactly one element (these are sometimes called *Sierpiński spaces*).■

Note. Clearly one can ask similar questions about topologies on other finite sets of the form $\{0, 1, \cdots, n\}$, but things quickly become fairly complicated. The case $n = 2$ is worked out completely in the following online document:

`http://fdslive.oup.com/www.oup.com/booksites/pdf/uk/companion/9780199563081/S.7.pdf`

It turns out that there are 29 different topologies on $\{0, 1, 2\}$.

1. In the definition of $\mathbf{d}_p$ we tacitly assume that $a \neq b$ and set $\mathbf{d}_p(a,a) = 0$ for all a. The nonnegativity of the function and its vanishing if and only if both variables are equal follow from the construction, as does the symmetry property $\mathbf{d}_p(a,b) = \mathbf{d}_p(b,a)$. The Triangle Inequality takes more insight. There is a special class of metric spaces known as **ultrametric spaces**, for which

$$\mathbf{d}(x,y) \leq \max\big\{\mathbf{d}(x,z), \mathbf{d}(y,z)\big\}$$

for all x, y, $z \in X$; the Triangle Inequality is an immediate consequence of this ultrametric inequality.

To establish this for the metric $\mathbf{d}_p$, we may as well assume that $x \neq y$ because if $x = y$ the ultrametric inequality is trivial (the left side is zero and the right is nonnegative). Likewise, we may as well assume that all three of x, y, z are distinct, for otherwise the ultrametric inequality is again a triviality. But suppose that $\mathbf{d}_p(x,y) = p^{-r}$ for some nonnegative integer r. This means that $x - y = p^r q$ where q is not divisible by p. If the ultrametric inequality is false, then p^{-r} is greater than either of either $\mathbf{d}_p(x,z)$ and $\mathbf{d}(y,z)$, which in turn implies that both $x - z$ and $y - z$ are divisible by p^{r+1}. But these two conditions imply that $x - y$ is also divisible by p^{r+1}, which is a contradiction. Therefore the ultrametric inequality holds for $\mathbf{d}_p$.

One curious property of this metric is that it takes only a highly restricted set of values; namely 0 and all fractions of the form p^{-r} where r is a nonnegative integer.■

2. Since U is open in A there is an open subset W in X such that $U = W \cap A$, and since $U \subset V$ we even have $U = V \cap U = V \cap W \cap A$. But $V \cap W$ is contained in the union of $U = V \cap W \cap A$ and $V - A$, and thus we have

$$U \cup (V - A) \ \subset\ (V \cap W) \cup (V - A) \ \subset\ \big(U \cup (V - A)\big) \cup (V - A) \ \subset\ U \cup (V - A)$$

so that $U \cup (V - A) = (V \cap W) \cup (V - A)$. Since A is closed the set $V - A$ is open, and therefore the set on the right hand side of the preceding equation is also open; of course, this means that the set on the left hand side of the equation is open as well.∎

3. $(\implies)$ If $A = E$ then E is open in itself, and therefore the first condition implies that E is open in X. $(\impliedby)$ If E is any subset of X and A is open in E then $A = U \cap E$ where U is open in X. But we also know that E is open in X, and therefore $A = U \cap E$ is also open in X.∎

4. (i) Suppose that $u, v \in N_\varepsilon(x)$. Then $\mathbf{d}(u,x) < \varepsilon$ and $\mathbf{d}(v,x) < \varepsilon$, so by the Triangle Inequality we have $\mathbf{d}(u,v) \leq \mathbf{d}(u,x) + \mathbf{d}(v,x) < \varepsilon + \varepsilon = 2\,\varepsilon$.∎

(ii) Suppose that $\mathbf{d}(a_1, a_2) \leq M$ for $a_1, a_2 \in A$, and suppose also that $B \subset A$. Then $b_1, b_2 \in B$ implies $b_1, b_2 \in A$ and therefore $\mathbf{d}(b_1, b_2) \leq M$, which means that B is also bounded.

Suppose now that A and B are bounded, with $\mathbf{d}(a_1, a_2) \leq K$ for $a_1, a_2 \in A$ and $\mathbf{d}(b_1, b_2) \leq L$ for $b_1, b_2 \in B$. If $u, v \in A \cup B$ there are essentially three possibilities:

$$u, v \ \in \ A\ , \qquad u, v \ \in \ B\ , \qquad u \ \in \ A \quad \text{and} \quad v \ \in \ B$$

(Strictly speaking, $u \in B$ and $v \in A$ is also possible, but it can be handled by reversing the roles of u and v in the third displayed case.) In the first two cases we have $\mathbf{d}(u,v) \leq K$ and $\mathbf{d}(u,v) \leq L$ respectively. In the third case, let $a_0 \in A$ and $b_0 \in B$. Then two applications of the Triangle Inequality imply that

$$\mathbf{d}(u,v) \quad \leq \quad \mathbf{d}(u, a_0) \ + \ \mathbf{d}(a_0, b_0) \ + \ \mathbf{d}(b-0, v) \quad \leq \quad K \ + \ \mathbf{d}(a_0, b_0) \ + \ L$$

and therefore the distance between two points in $A \cup B$ is bounded by the right hand side of the display in all cases.

An infinite union of bounded subsets can be unbounded. One of the simplest examples is given by $U_n = (-n, n) \subset \mathbb{R}$. By (i) we know that the distance between two points of U_n is bounded from above by $2n$, but the union of all the sets U_n is the entire real line, which is unbounded.

There are also plenty of other families $\{A_n\}$ such that each A_n is bounded but $\cup_n\ A_n$ is not, and it is even possible to find examples where there is a single constant M such that for each n we have $\mathbf{d}(u,v) \leq M$ for all $u, v \in A_n$. A specific example along these lines is $A_n = [n, n+1]$ for n a nonnegative integer. For this family one can take $M = 1$ for all n, but $\cup_n\ A_n \ = \ [0, \infty)$, which is unbounded.∎

5. (i) Since $\mathbb{R} \cap A = A$, $\mathbb{R} \cup A = \mathbb{R}$, $\emptyset \cap A = \emptyset$ and $\emptyset \cup A = A$ for all subsets $A \subset \mathbb{R}$, we only need to show that an arbitrary union or twofold intersection of nonempty proper subsets in $\mathbf{U}$ is also in $\mathbf{U}$ (in which case the intersection might be empty or the union might be all of $\mathbb{R}$).

Since the sets (b, ∞) are open in the metric topology, it follows that $\mathbf{U}$ is contained in the metric topology. The containment is proper because nonempty open subsets in $\mathbf{U}$ never have an upper bound in $\mathbb{R}$ and there are many metrically open subsets that do — for example, the interval $(-1, 1)$.

Twofold intersections. Given (b_1, ∞) and (b_2, ∞), their intersection is (c, ∞), where c is the larger of b_1 and b_2.

Arbitrary unions. Suppose that we are given nonempty proper open subsets (b_α, ∞) for $\alpha \in A$. We claim that

$$\bigcup_{\alpha \in A} (b_\alpha, \infty) \quad = \quad (b^*, \infty)$$

where b^* is either $-\infty$ or the greatest lower bound of $B = \{\ b_\alpha \mid \alpha \in A\}$. Since $b^* \leq b_\alpha$ for all α it is clear that the left hand side is contained in the right hand side. It remains to prove the reverse inclusion.

Unbounded case. If B is unbounded, then for each $x \in \mathbb{R}$ there is some b_γ such that $b_\gamma < x$, and therefore $x \in (b_t, \infty)$ and hence x lies in the union.

Bounded case. Suppose now that B is bounded, and let b^* be its greatest lower bound. If $x > b^*$, then x is not a lower bound for B, so there is some b_γ such that $b_\gamma < x$. We then have $x \in (b_\gamma, \infty)$ and hence once again x lies in the union.■

(ii) This is very similar to (i), and the simplest way to dispose of it is in two steps:

(a) If $(X, \mathbf{T})$ is a topological space and $f : X \to Y$ is 1–1 onto, then the family $f_*\mathbf{T}$ of all sets of the form $f[V]$, where V is open in X, is a topology on Y.

(b) If $\mathbf{U}$ is the topology in (i) and $g(x) = -x$ on $\mathbb{R}$, then $g_*\mathbf{U} = \mathbf{L}$.

The second statement is clear because g maps (b, ∞) to $(-\infty, -b)$ and vice versa. The first statement follows because $g[\emptyset] = \emptyset$, $g[X] = Y$, $g[U_1 \cap U_2] = g[U_1] \cap g[U_2]$ and $g[\cup_\alpha\ U_\alpha] = \cup_\alpha\ g[U_\alpha]$ (the first and last hold for arbitrary maps, while the middle two hold if g is a 1–1 onto mapping).■

(iii) The nonempty proper subsets in $\mathbf{U} \cup \mathbf{L}$ either do not have an upper bound or do not have a lower bound. Each set $(x - \varepsilon, x + \varepsilon)$ is an intersection of one set in $\mathbf{L}$ and one set in $\mathbf{U}$, and if $\mathbf{T}$ is a topology containing $\mathbf{U} \cup \mathbf{L}$ then this intersection must lie in $\mathbf{T}$. Since the constructed interval has both upper and lower bounds, it cannot belong to $\mathbf{U} \cup \mathbf{L}$, and therefore the latter cannot be a topology on $\mathbb{R}$.■

6. We shall first verify that if X is a topological space with base $\mathcal{B}$, then the subfamilies $\mathcal{B}_x$ satisfy the given conditions.

Verification of **(N1)**. By definition, if $u \in \mathcal{B}_x$ then $x \in U$.

Verification of **(N2)**. The set $U_1 \cap U_2$ is an open subset containing x. Since $\mathcal{B}$ is a base for X we know that $U_1 \cap U_2$ is a union of open subsets in $\mathcal{B}$; express $U_1 \cap U_2$ as such a union V_α, and choose V^* such that $x \in V^*$.

Verification of **(N3)**. Let $W = U$; by definition, this set belongs to $\mathcal{B}_y$.

Verification of **(N4)**. If the condition holds, then we know that each V_x is open, and we also know that

$$U \quad = \quad \bigcup_{x \in U} \{x\} \quad \subset \quad \bigcup_{x \in U} V_x \quad \subset \quad U$$

which implies that the two subsets in the middle are equal to U. The third set in the sequence is open because each V_x is open, and therefore U itself must be open.

We shall now prove that if we are given the data $\mathcal{B}_x$ for each $x \in X$ such that **(N1)**–**(N3)** are satisfied, then the union $\cup_{x \in X}\ \mathcal{B}_x$ is the base for some topology on the set X; by the definition of a base for a topology, property (**N4**) will also hold in this case.

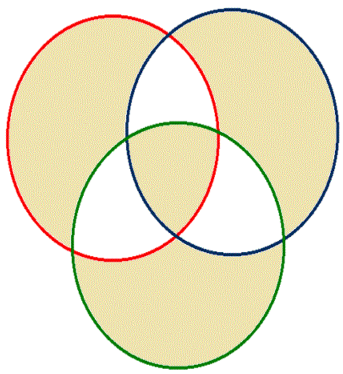
Let $\mathbf{T}$ be the topology generated by $\mathcal{U} = \cup_{x \in X}\ \mathcal{B}_x$; it will suffice to show that the latter is in fact a basis for $\mathbf{T}$. This will be true if each finite intersection of sets in $\mathcal{U}$ can be expressed as a union of subsets in $\mathcal{U}$. By induction and the distributivity properties of unions and intersections, this reduces to verifying the assertion in the preceding sentence for twofold intersections (because a threefold intersection is a union of twofold intersections, *etc.* — fill in the details!). Let $p, q \in X$, suppose that we are given open subsets $V_p \in \mathcal{B}_p$ and $V_q \in \mathcal{B}_q$, and suppose that $z \in V_p \cap V_q$. By **(N3)** there are sets $W_p, W_q \in \mathcal{B}_z$ such that $z \in W_p \subset V_p$ and $z \in W_q \subset V_q$. Therefore by **(N2)** there is some subset $W_z^* \subset W_p \cap W_q$ in $\mathcal{B}_z$. It follows that $W_z^* \subset W_p \cap W_q$ and that $V_p \cap V_q$ is the union of such sets W_z^*; therefore the intersection of two sets in $\mathcal{U}$ is a union of subsets in $\mathcal{U}$, and as noted before this implies that $\mathcal{U}$ is a base for $\mathbf{T}$.■

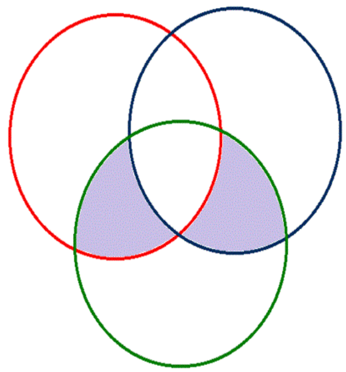
NOTE. If X is a metric space, the for each $x \in X$ the families of subsets $\{\ N_{1/n}(x) \mid n \in \mathbb{N}^+\ \}$ satisfy conditions **(N1)** – **(N4)**. The family $\mathcal{B}_x$ is called an *open neighborhood base at the point*

$x \in X$. — In this terminology, we can say that a topological space is completely determined by describing abstract open neighborhood bases at all the points of X.

Drawing to accompany Additional Exercise I.1.1

Since the hint for this exercise suggested drawing Venn diagrams, here are two such drawings. The first (on the left) illustrates the associativity law $(A \oplus B) \oplus C = A \oplus (B \oplus C)$ and the second one illustrates the distributive law $A \cap (B \oplus C) = (A \cap B) \oplus (A \cap C)$. In each of these drawings, the circles at the top represent A and B while circle at the bottom represents C, and one shows that the shaded regions depict both the set on the left and the set on the right of the asserted equations.





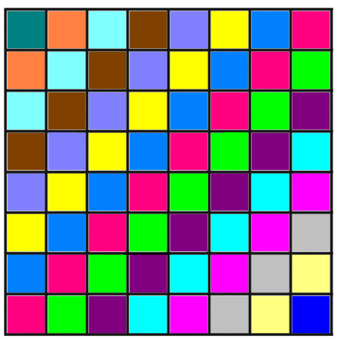
Note. If one defines an abstract Boolean algebra to be a set A with abstract union, intersection and complementation operations which satisfy the standard set – theoretic identities as in **http://en.wikipedia.org/wiki/Boolean_algebra_%28structure%29**, then one can define $\oplus$ exactly as in the exercise, and it satisfies the properties derived in the exercise. This fact has far – reaching implications. For example, it is central to the proof that every abstract Boolean algebra A is isomorphic to a subalgebra of the Boolean algebra $\mathcal{P}(X)$ of subsets of some set X. This result is due to M. H. Stone; here is an online reference for additional information (it uses concepts developed later in this course):

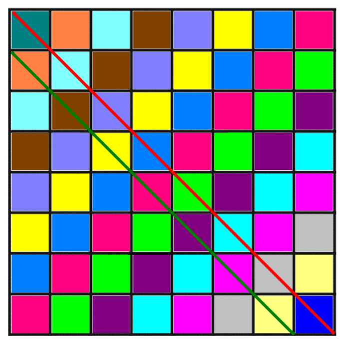
http://en.wikipedia.org/wiki/Stone%27s_representation_theorem_for_Boolean_algebras

Incidentally, despite the apparently abstract nature of the result, the motivation for proving it arose in connection with some sophisticated questions in functional analysis.

Drawing to accompany Additional Exercise I.2.4

Assume we label the squares by two positive integers on the board from left to right and from the bottom to the top. The first step in the argument is to show that a bishop can move diagonally to another square of the same color. In other words, if the bishop is located at the point whose horizontal coordinate is i and whose vertical coordinate is j, then the bishop can move one square up or down, to the square whose horizontal coordinate is $i+1$ and whose vertical coordinate is $j-1$, or to the square whose horizontal coordinate is $i-1$ and whose vertical coordinate is $j+1$, provided there are such squares on the board. Thus each diagonal lies in an equivalence class of points such that a bishop can move from one square to another in the class, and since there are exactly 15 diagonals in the drawing, this means there are at most 15 equivalence classes (see the drawing on the left).

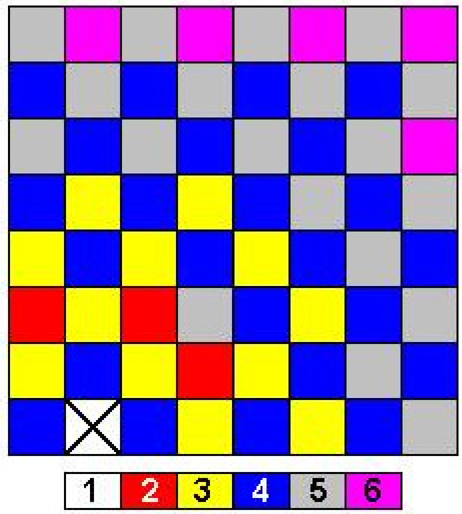




The final step in the argument is to note that the points on the red and green lines in the right hand drawing also lie in the same equivalence class. Since the two lines contain exactly one square of each color, it follows that there are at most two equivalence classes of squares, and they are distinguished by whether $i+j$ is even or odd. In fact, there are exactly two such equivalence classes, for if the bishop moves one square from the position with coordinates i and j to a square with coordinates p and q, then by construction the sums $i+j$ and $p+q$ are both even or both odd.

See the next page for remarks on the squares that a knight on a chessboard can reach.

Knight move illustrations



***<u>Explanation</u>*:** The color code indicates when a knight will get from the original position marked with an **X** to a given square on the chessboard. For example, the possible positions after one move are colored red, and the possible positions after two moves are colored yellow. Note that the knight can reach every square in at most five moves. It also follows that if the knight starts at an arbitrary square, then it can reach any other square within **10** moves.

SOLUTIONS TO EXERCISES FOR

MATHEMATICS 205A — Part 2

Fall 2014

II. Metric and topological spaces

II.2 : Closed sets and limit points

Problems from Munkres, § 17*, pp.* 101 – 102

2. Since A is closed in Y we can write $A = F \cap Y$ where F is closed in X. Since an intersection fo closed set is closed and Y is closed in X , it follows that $F \cap Y = A$ is also closed in x.■

8. (a) Since $C \subset Y$ implies $\overline{C} \subset \overline{Y}$ it follows that $\overline{A \cap B} \subset \overline{A}$ and $\overline{A \cap B} \subset \overline{B}$, which yields the inclusion

$$\overline{A \cap B} \quad \subset \quad \overline{A} \ \cap \ \overline{B} \ .$$

To see that the inclusion may be proper, take A and B to be the open intervals $(0, 1)$ and $(1, 2)$ in the real line. Then the left hand side is empty but the right hand side is the set $\{1\}$.■

(c) This time we shall first give an example where the first set properly contains the second. Take $A = [-1, 1]$ and $B = \{0\}$. Then the left hand side is A but the right hand side is $A - B$. We shall now show that $\overline{A - B} \supset \overline{A} - \overline{B}$ always holds. Given $x \in \overline{A} - \overline{B}$ we need to show that for eacn open set U containing x the intersection $U \cap (A - B)$ is nonempty. Given such an open set U, the condition $x \notin \overline{B}$ implies that $x \in U - \overline{B}$, which is open. Since $x \in \overline{A}$ it follows that

$$A \cap \left(U - \overline{B}\right) \quad \neq \quad \emptyset$$

and since $U - \overline{B} \subset U - B$ it follows that

$$(A - B) \cap U \quad = \quad A \cap (U - B) \quad \neq \quad \emptyset$$

and hence that $x \in \overline{A - B}$.■

19. (a) The interior of A is the complement of $\overline{X - A}$ while the boundary is contained in the latter, so the intersection is empty.■

(b) ($\Longleftarrow$) If A is open then $X - A$ is closed so that $X - A = \overline{X - A}$, and if A closed then $\overline{A} = A$. Therefore

$$\mathrm{Bd}(A) = \overline{A} \cap \overline{X - A} = A \cap (X - A) = \emptyset \ .■$$

($\Longrightarrow$) The set $\mathrm{Bd}(A)$ is empty if and only if $\overline{A}$ and $\overline{X - A}$ are disjoint. Since the latter contains $X - A$ it follows that $\overline{A}$ and $X - A$ are disjoint. Since their union is X this means that $\overline{A}$ must be contained in A, which implies that A is closed. If one reverses the roles of A and $X - A$ in the preceding two sentences, it follows that $X - A$ is also closed; hence A is both open and closed in X.■

(c) By definition the boundary is $\overline{U} \cap \overline{X - U}$.

($\Longrightarrow$) If U is open then $X-U$ is closed and thus equal to its own closure, and therefore the definition of the boundary for U reduces to $\overline{U}\cap(X-U)$, which is equal to $\overline{U}-U$.■

($\Longleftarrow$) We shall show that $X-U$ is closed. By definition $\mathrm{Bd}(U)=\mathrm{Bd}(X-U)$, and therefore $\mathrm{Bd}(X-U)=\overline{U}-U\subset X-U$. On the other hand, by part (a) we also know that $(X-U)\cup\mathrm{Bd}(X-U)=\overline{X-U}$. Since both summands of the left hand side are contained in $X-U$, it also follows that the right hand side is contained in $X-U$, which means that $X-U$ is closed in X.■

(*d*) NO. Take $U=(-1,0)\cup(0,1)$ as a subset of $\mathbb{R}$. Then the interior of the closure of U is $(-1,1)$. However, we always have $U\subset\mathrm{Int}(\overline{U})$ because $U\subset\overline{U} \implies U=\mathrm{Int}(U)\subset\mathrm{Int}(\overline{U})$.■

FOOTNOTE. Sets which have the property described in **19.(d)** are called *regular open sets*.

20. (*a*) We need to find the closures of A and its complement in $\mathbb{R}^2$. The complement of A is the set of all points whose second coordinate is nonzero. We claim it is open. But if $(x,y)\in\mathbb{R}^2-A$ then $y\neq 0$ and the set $N_{|y|}(\,(x,y)\,)\subset\mathbb{R}^2-A$. Therefore A is closed. But the closure of the complement of A is all of $\mathbb{R}^2$; one easy way of seeing this is that for all $x\in\mathbb{R}$ we have $\lim_{n\to\infty}(x,1/n)=(x,0)$, which means that every point of A is a limit point of $\mathbb{R}^2-A$. Therefore the boundary of A is equal to $A\cap\mathbb{R}^2=A$; *i.e.*, every point of A is a boundary point.■

(*b*) Again we need to find the closures of B and its complement. It will probably be very helpful to draw pictures of this set and the sets in the subsequent portions of this problem. The closure of B turns out to be the set of all points where $x\geq 0$ (find sequences converging to all points in this set that are not in B **!**) and the complement is just the complement of B because the latter is open in $\mathbb{R}^2$. Therefore the boundary of B is $\overline{B}-B$, and this consists of all points such that either $x=0$ or both $x>0$ and $y=0$ hold.■

(*c*) The first two sentences in part (*b*) apply here also. The set C consists of all points such that $x>0$ or $y=0$ The closure of this set is the set of all points such that $x\geq 0$ or $y=0$, the complement of C is the set of all points such that $x\geq 0$ and $y\neq 0$, and the closure of this complement is the set of all points such that $x\leq 0$. The intersection of the two sets will be the set of all point such that $x=0$ or both $x<0$ and $y=0$ hold.■

(*d*) The first two sentences in part (*b*) apply here also. Since every real number is a limit of irrational numbers, it follows that the closure of D is all of $\mathbb{R}^2$. The complement of D consists of all points whose first coordinates are rational, and since every real number is a limit of a sequence of rational numbers it follows that the closure of $\mathbb{R}^2-D$ is also $\mathbb{R}^2$. Therefore the boundary is $\mathbb{R}^2$.■

(*e*) The first two sentences in part (*b*) apply here also. In problems of this sort one expects the boundary to have some relationship to the curves defined by changing the inequalities into equations; for this example the equations are $x^2-y^2=0$ and $x^2-y^2=1$. The first of these is a pair of diagonal lines through the origin that make 45 degree angles with the coordinate axes, and the second is a hyperbola going through $(\pm 1,0)$ with asymptotes given by the lines $x^2-y^2=0$. The closure of E turns out to be the set of points where $0\leq x^2-y^2\leq 1$, and the closure of its complement is the set of all points where either $x^2-y^2\leq 0$ or $x^2-y^2\geq 1$. The intersection will then be the set of points where x^2-y^2 is either equal to 0 or 1.

For the sake of completeness, here is a proof of the assertions about closures: Suppose that we have a sequence $\{\,(x_n,y_n)\,\}$ in E and the sequence has a limit $(a,b)\in\mathbb{R}^2$. Then

$$a^2-b^2 \quad = \quad \lim_{n\to\infty} x_n^2-y_n^2$$

where for each n the n^{th} term on the right hand side lies in the open interval $(0,1)$, and therefore the limit value on the right hand side must lie in the closed interval $[0,1]$. Similarly, suppose that we have sequences in the complement of E that converge to some point (a, b). Since the sequence of numbers $z_n = x_n^2 - y_n^2$ satisfies $z_n \in \mathbb{R} - (0, 1)$ and the latter subset is closed, it also follows that $a^2 - b^2 \in \mathbb{R} - (0, 1)$. This proves that the closures of E and its complement are contained in the sets described in the preceding paragraph.

To complete the proof of the closure assertions we need to verify that every point on the hyperbola or the pair of intersecting lines is a limit point of E. Suppose that we are given a point (a, b) on the hyperbola, and consider the sequence of points

$$(x_n, y_n) \quad = \quad \left(a - \frac{\sigma(a)}{n}, \, b \right)$$

where $\sigma(x)$ is ± 1 depending on whether a is positive or negative (we know that $|a| > 1$ because $a^2 = 1 + b^2$). If we take $z_n = x_n^2 - y_n^2$ as before then $\lim_{n \to \infty} z_n = 1$ and moreover

$$z_n \quad = \quad \left(a - \frac{\sigma(a)}{n} \right)^2 - b^2 \quad = \quad 1 + \frac{1}{n^2} - \frac{2\, a\, \sigma(a)}{n}$$

where the expression on the right hand side is positive if

$$|a| \quad < \quad \frac{1}{2n} \; .$$

To see that this expression is less than 1 it suffices to note that $|a| = \sigma(a) \cdot a$ and

$$\frac{1}{n^2} - \frac{2\,|a|}{n} \quad \leq \quad \frac{1}{n^2} - \frac{2}{n} \quad < \quad 0$$

for all $n \geq 1$. This verifies that every point on the hyperbola is a limit point of E. — Graphically, we are taking limits along horizontal lines; the reader might want to draw a picture in order to visualize the situation.

Suppose now that (a, b) lies on the pair of intersecting lines, so that $a = \pm b$. How do we construct a sequence in E converging to (a, b)? Once again, we take sequences that live on a fixed horizontal line, but this time we choose

$$(x_n, y_n) \quad = \quad \left(a + \frac{\sigma(a)}{n}, \, b \right)$$

and note that $z_n = x_n^2 - y_n^2$ is equal to

$$\frac{1}{n^2} \; + \; \frac{2\,|a|}{n}$$

which is always positive and is also less than 1 if $n > 2|a| + 1$ (the latter follows because the expression in question is less than $(1 + 2|a|)/n$ by the inequality $n^{-2} < n^{-1}$ for all $n \geq 2$). Thus every point on the pair of intersecting lines also lies in $\mathbf{L}(E)$, and putting everything together we have verified that the closure of E is the set we thought it was.■

(f) Geometrically, this is the region underneath either the positive or negative branch of the hyperbola $y = 1/x$ with the x-axis removed, and the branches of the hyperbola are included in the set. We claim that the closure of F is the union of F with the y-axis (*i.e.*, the set of points where

$x = 0$). To see that the y-axis is contained in this set, consider a typical point $(0, y)$ and consider the sequence

$$\left(\frac{1}{n},\, y\right)$$

whose limit is $(0, y)$; the terms of this sequence lie in the set F for all $y \geq n$, and therefore the y-axis lies in $\mathbf{L}(F)$. Therefore the closure is at least as large as the set we have described. To prove that it is no larger, we need to show that there are no limit points of F such that $x \neq 0$ and $y > 1/x$. But suppose that we have an infinite sequence in F with terms of the form $(x_n.y_n)$ and limit equal to (a, b), where $a \neq 0$. There are two cases depending upon whether a is positive or negative. Whichever case applies, for all sufficiently large values of n the signs of the terms x_n are equal to the sign of a, so we may as well assume that all terms of the sequence have first coordinates with the same sign as a (drop the first finitely many terms if necessary). If $a > 0$ it follows that $x_n > 0$ and $y_n \leq 1/x_n$, which imply $x_n\, y_n \leq 1$ Taking limits we see that $ab \leq 1$ also holds, so that $b \leq 1/a$. On the other hand, if $a < 0$ then it follows that $x_n < 0$ and $y_n \leq 1/x_n$, which imply $x_n\, y_n \geq 1$ Taking limits we see that $ab \geq 1$ also holds, so that $b \leq 1/a$ holds in this case too.

Now we have to determine the closure of the complement of F; we claim it is the set of all points where either $x = 0$ or $x \neq 0$ and $y \geq 1/x$. By definition it contains all points where $x = 0$ or $y > 1/x$, so we need to show that the hyperbola belongs to the set of limit points and if we have a sequence of points of the form (x_n, y_n) in the complement of F with a limit in the plane, say (a, b), then $b \geq 1/a$. Proving the latter proceeds by the same sort of argument given in the preceding paragraph (it is not reproduced here, but it must be furnished in a complete proof). To see that the hyperbola belongs to the set of limit points take a typical point (a, b) such that $a \neq 0$ and $b = 1/a$ and consider the sequence

$$(x_n, y_n) \quad = \quad \left(a,\, \frac{1}{a} + \frac{1}{n}\right)$$

whose terms all lie in the complement of F and whose limit is (a, b).■

Additional exercises

0. Let U and V be the open intervals $(-1, 0)$ and $(0, 1)$ respectively. Then their closures are the closed intervals $[-1, 0]$ and $[0, 1]$ respectively, and the intersection of these two sets is $\{0\}$. This counterexample shows that the statement is false.

1. Take any set S with the discrete metric and let $\varepsilon = 1$. Then the set of all points whose distance from some particular $s_0 \in S$ is ≤ 1 is all of S, but the open disk of radius 1 centered at s_0 is just the one point subset $\{\, s_0 \,\}$.■

2. If X has the discrete topology then every subset is equal to its own closure (because every subset is closed), so the closure of a proper subset is always proper. Conversely, if X is the only dense subset of itself, then for every proper subset A its closure $\overline{A}$ is also a proper subset. Let $y \in X$ be arbitrary, and apply this to $X - \{y\}$. Then it follows that the latter is equal to its own closure and hence $\{y\}$ is open. Since y is arbitrary, this means that X has the discrete topology.■

3. Given a point $x \in X$ and and open subset W such that $x \in W$, we need to show that the intersection of W and $U \cap V$ is nonempty. Since U is dense we know that $W \cap U \neq \emptyset$; let y be a point in this intersection. Since V is also dense in X we know that

$$(U \cap V) \cap W \quad = \quad V \cap (U \cap W) \quad \neq \emptyset$$

and therefore $U \cap V$ is dense. — You should be able to construct examples in the real line to show that the conclusion is not necessarily true if U and V are not open.■

4. $(\Longleftarrow)$ If $A = E \cap U$ where E is closed and U is open then for each $a \in A$ one can take U itself to be the required open neighborhood of a. $(\Longrightarrow)$ Given $a \in A$ let U_a be the open set containing a such that $U_a \cap A$ is closed in U_a. This implies that $U_a - A$ is open in U_a and hence also in X. Let $U = \cup_a U_a$. Then by construction $A \subset U$ and

$$U - A \quad = \quad \bigcup_{\alpha} \, (U_a - A)$$

is open in X. If we take

$$E \quad = \quad X - \overline{U - A}$$

then E is closed in X and $A = U \cap E$ where U is open in X and E is closed in X.■

5. (a) Let X be the real numbers, let D be the rational numbers and let $A = X - D$. Then $A \cap D = \emptyset$, which is certainly not dense in X.■

(b) If $x \in \overline{B}$ and U is an open set containing x, then $U \cap B \neq \emptyset$. Let b be a point in this intersection. Since $b \in U$ and A is dense in B it follows that $A \cap U \neq \emptyset$ also. But this means that A is dense in $\overline{B}$.■

6. $(\Longrightarrow)$ If $A = E$ then E is closed in itself, and therefore the first condition implies that E is closed in X. $(\Longleftarrow)$ If E is any subset of X and A is closed in E then $A = U \cap E$ where U is closed in X. But we also know that E is closed in X, and therefore $A = U \cap E$ is also closed in X. (Does all of this sound familiar? The exercise is essentially a copy of an earlier one with with "closed" replacing "open" everywhere.) ■

7. Consider the subset A of $\mathbb{R}$ consisting of $(0,1) \cup \{2\}$. The closure of its interior is $[0,1]$.■

8. Suppose $x \in U \cap B \subset U \cap \overline{B}$. Then the inclusion $U \cap V \subset \overline{U \cap B}$ shows that $x \in \overline{U \cap B}$. Since $\overline{B} = B \cup \mathbf{L}(B)$ the resulting set-theoretic identity

$$U \cap \overline{B} \quad = \quad (U \cap B) \, \cap \, (U \cap \mathbf{L}(B))$$

implies that we need only verify the inclusion

$$(U \cap \mathbf{L}(B)) \quad \subset \quad \overline{U \cap B}$$

and it will suffice to verify the stronger inclusion statement

$$(U \cap \mathbf{L}(B)) \quad \subset \quad \mathbf{L}(U \cap B) \; .$$

Suppose that $x \in U \cap \mathbf{L}(B)$, and let W be an open subset containing x. Then $W \cap U$ is also an open subset containing X, and since $x \in \mathbf{L}(B)$ we know that

$$\Big(U \cap W - \{x\} \Big) \, \cap B \quad \neq \quad \emptyset \; .$$

But the expression on the left hand side of this display is equal to

$$\Big(W - \{x\} \Big) \; \cap \; U \; \cap \; B$$

and therefore the latter is nonempty, which shows that $x \in \mathbf{L}(U \cap B)$ as required.■

9. In order to define a topological space it is enough to define the family $\mathcal{F}$ of closed subsets that satisfies the standard properties: It contains the empty set and Y, it is closed under taking arbitrary intersections, and it is closed under taking the unions of two subsets. If we are given tne abstract operator **CL** as above on the set of all subsets of Y let $\mathcal{F}$ be the family of all subsets A such that $\mathbf{CL}(A) = A$. We need to show that this family satisfies the so-called standard properties mentioned in the second sentence of this paragraph.

The empty set belongs to $\mathcal{F}$ by **(C4)**, and Y does by **(C1)** and the assumption that $\mathbf{CL}(A) \subset Y$ for all $A \subset Y$, which includes the case $A = Y$. If A and B belong to $\mathcal{F}$ then the axioms imply

$$A \cup B \quad = \quad \mathbf{CL}(A) \cup \mathbf{CL}(B) \quad = \quad \mathbf{CL}(A \cup B)$$

(use **(C3)** to derive the second equality).

The only thing left to check is that $\mathcal{F}$ is closed under arbitrary intersections. Let $\mathcal{A}$ be a set and let $\{ A_\alpha \}$ be a family of subsets in $\mathcal{F}$ indexed by all $\alpha \in \mathcal{A}$; by assumption we have $\mathbf{CL}(A_\alpha) = A_\alpha$ for all α, and we need to show that

$$\mathbf{CL}\left(\bigcap_\alpha A_\alpha\right) \quad = \quad \bigcap_\alpha A_\alpha \ .$$

By **(C2)** we know that

$$\mathbf{CL}\left(\bigcap_\alpha \ A_\alpha\right) \quad \subset \quad \bigcap_\alpha \ \mathbf{CL}\left(A_\alpha\right)$$

and thus we have the chain of set-theoretic inclusions

$$\bigcap_\alpha A_\alpha \quad \subset \quad \mathbf{CL}\left(\bigcap_\alpha \ A_\alpha\right) \quad \subset \quad \bigcap_\alpha \ \mathbf{CL}\left(A_\alpha\right) \quad = \quad \bigcap_\alpha \ A_\alpha$$

which shows that all sets in the chain of inclusions are equal and hence that if $D = \cap_\alpha \ A_\alpha$, then $D = \mathbf{CL}(D)$.■

FOOTNOTE. Exercise 21 on page 102 of Munkres is a classic problem in point set topology that is closely related to the closure operator on subsets of a topological space: Namely, if one starts out with a fixed subset and applies a finite sequence of closure and (set-theoretic) complement operations, then one obtains at most 14 distinct sets, and there are examples of subsets of the real line for which this upper bound is realized. Some hints for working this exercise appear in the following web site:

`http://www.math.ou.edu/~nbrady/teaching/f02-5853/hint21.pdf`

10. (a) It suffices to show that $\mathbf{L}\left(\mathbf{L}(A)\right) \subset \mathbf{L}(A)$. Suppose that $x \in \mathbf{L}\left(\mathbf{L}(A)\right)$. Then for every open set U containing x we have $(U - \{x\}) \cap A \neq \emptyset$, so let y belong to this nonempty intersection. Since one point subsets are closed, it follows that $U - \{x\}$ is an open set containing y, and therefore we must have

$$\Big(U - \{x,\, y\}\Big) \ \cap \ A \quad \neq \quad \emptyset$$

and therefore the sets $U - \{x\}$ and A have a nonempty intersection, so that $x \in \mathbf{L}(A)$.■

(b) If all one point subsets of X are closed, then all finite subsets of X are closed, and hence the complements of all finite subsets of X are open. We shall need this to complete the proof.

Suppose that the conclusion is false; *i.e.*, the set $(U - \{b\}) \cap A$ is finite, say with exactly n elements. If F denotes this finite intersection, then by the preceding paragraph $V = U - F$ is an open set, and since $x \notin F$ we also have $x \in V$. Furthermore, we have $(V - \{x\}) \cap A = \emptyset$; on the other hand, since $b \in \mathbf{L}(A)$ we also know that this intersection is nonempty, so we have a contradiction. The contradiction arose from the assumption that $(U - \{b\}) \cap A$ was finite, so this set must be infinite.■

11. The first and second conditions respectively imply that X and the empty set both belong to $\mathbf{T}$. Furthermore, the fourth condition implies that the intersection of two sets in $\mathbf{T}$ also belongs to $\mathbf{T}$, so it remains to verify the condition on unions. Suppose that A is a set and $U_\alpha \in \mathbf{T}$ for all $\alpha \in A$. Then we have

$$\bigcap_\alpha U_\alpha \quad = \quad \bigcap_\alpha \mathbf{I}(U_\alpha) \quad \subset \quad \mathbf{I}\left(\bigcap_\alpha U_\alpha\right) \quad \subset \quad \bigcap_\alpha U_\alpha$$

where the third property implies the right hand containment; the chain of inequalities implies that $\cup_\alpha ! U_\alpha$ belongs to $\mathbf{T}$, and therefore it follows that the latter is a topology for X such that a set U is open if and only if $\mathbf{I}(U) = U$.■

12. (*a*) By definition $\mathrm{Ext}(A \cup B)$ is equal to

$$X - \overline{A \cup B} \quad = \quad X - \left(\overline{A} \cup \overline{B}\right) \quad = \quad \left(X - \overline{A}\right) \cap \left(X - \overline{B}\right)$$

which again by definition is equal to $\mathrm{Ext}(A) \cap \mathrm{Ext}(B)$.■

(*b*) Since $A \subset \overline{A}$ it follows that $X - \overline{A} \subset X - A$ and hence $\mathrm{Ext}(A) \cap A \subset (X - A) \cap A = \emptyset$.■

(*c*) The empty set is closed and therefore $\mathrm{Ext}(\emptyset) = X - \emptyset = X$.■

(*d*) What is the right hand side? It is equal to $X - \overline{B}$ where $B = X - \overline{X - A}$. Note that $B = \mathrm{Int}(A)$. Therefore the right hand side may be rewritten in the form

$$X - \overline{(\mathrm{Int}(A)\,)} \ .$$

We know that $\mathrm{Int}(A) \subset A$ and likewise for their closures, and thus the reverse implication holds for the complements of their closures. But the last containment relation is the one to be proved.■

13. We may write $B = A_i \cap F_i$ where F_i is closed in X. It follows that $B = B \cap F_2 = A_1 \cap F_1 \cap F_2$ and $B = B \cap F_1 = A_2 \cap F_2 \cap F_1 = A_2 \cap F_1 \cap F_2$. Therefore

$$B = B \cup B \quad = \quad \left(A_1 \cap F_1 \cap F_2\right) \cup \left(A_2 \cap F_1 \cap F_2\right) \quad - \quad \left(A_1 \cup A_2\right) \ \cap \ \left(F_1 \cap F_2\right)$$

which shows that B is closed in $A_1 \cup A_2$.■

Note that the statement and proof remain valid if "closed" is replaced by "open."■

14. The first follows because $\mathrm{Int}(A) \subset A$, closure preserves set-theoretic inclusion, and $A = \overline{A}$. To prove the second statement, begin by noting that the first set is contained in the second because $B \subset A$. The reverse inclusion follows because $B = \overline{\mathrm{Int}(A)} \supset \mathrm{Int}(A)$ implies

$$\mathrm{Int}(B) \supset \mathrm{Int}\,(\,\mathrm{Int}(A)\,) = \mathrm{Int}(A) \ .■$$

15. (*a*) The set $\mathrm{Int}_X(A)$ is an open subset of X and is contained in A, so it is also an open subet of Y that is contained in A. Since $\mathrm{Int}_Y(A)$ is the maximal such subset, it follows that $\mathrm{Int}_X(A) \subset \mathrm{Int}_Y(A)$.■

(b) It will be convenient to let $\mathrm{CL}_U(B)$ denote the closure of B in U in order to write things out unambiguously.

By definition $\mathrm{Bd}_Y(A)$ is equal to $\mathrm{CL}_Y(A) \cap \mathrm{CL}_Y(Y-A)$, and using the formula $\mathrm{CL}_Y(B) = \mathrm{CL}_X(B) \cap Y$ we may rewrite $\mathrm{Bd}_Y(A)$ as the subset $\mathrm{CL}_X(A) \cap \mathrm{CL}_X(Y-A) \cap Y$. Since $Y - A \subset X - A$ we have $\mathrm{CL}_X(Y-A) \subset \mathrm{CL}_X(X-A)$, and this yields the relation

$$\mathrm{Bd}_Y(A) = \mathrm{CL}_X(A) \cap \mathrm{CL}_X(Y-A) \cap Y \subset \mathrm{CL}_X(A) \cap \mathrm{CL}_X(X-A) = \mathrm{Bd}_X(Y)$$

that was to be established.■

(c) One obvious class of examples for **(a)** is given by taking A to be a nonempty subset that is not open and to let $Y = A$. Then the interior of A in X must be a proper subset of A but the interior of A in itself is simply A.

Once again, the best way to find examples where BOTH inclusions are proper is to try drawing a few pictures with pencil and paper. Such drawings lead to many examples, and one of the simplest arises by taking $A = [0,1] \times \{0\}$, $Y = \mathbb{R} \times \{0\}$ and $X = \mathbb{R}^2$. For this example the interior inclusion becomes $\emptyset \subset (0,1) \times \{1\}$ and the boundary inclusion becomes $\{0,1\} \times \{0\} \subset [0,1] \times \{0\}$. The details of verifying these are left to the reader.■

16. (i) Since $A = X - (X - A)$, by definition the boundary of $X - A$ is $\overline{X-A} \cap \overline{X - (X-A)} = \overline{X-A} \cap \overline{A}$, which is the definition of $\mathrm{Bdy}\,(A)$.■

(ii) Use the hint, applying it to A and $X - A$. A point x lies in the intersection $\overline{A} \cap \overline{X-A}$ if and only if every open neighborhood U of X contains at least one point in A and at least one point in $X - A$, and this is the condition which appears in the statement of part (ii).■

(iii) The statement in this part of the exercise will follow if we know that $\mathrm{Int}\,(A) = X - \overline{X-A}$, so we shall prove that identity instead. Since $X - A \subset \overline{X-A}$, it follows that $X - \overline{X-A}$ is an open set contained in $A = X - (X-A)$. Furthermore, if U is an open set such that $U \subset A$, then $X - U$ is a closed subset containing $X - A$ and hence $X - U \subset \overline{X-A}$. Taking complements, we find that U must be contained in $X - \overline{X-A}$. Therefore the latter is the unique maximal open subset contained in A.■

(iv) If A is closed in X, then $A = \overline{A}$ contains $\mathrm{Bdy}\,(A) = \overline{A} \cap \overline{X-A}$. Conversely, if $\mathrm{Bdy}\,(A) \subset A$ we shall show that the set of limit points $\mathbf{L}(A)$ is contained in A. Suppose that $x \in \mathbf{L}(A)$ but $x \notin \mathrm{Bdy}\,(A)$. The second condition implies that some open neighborhood U_0 of x does not contain any points of $X - A$, and it follows that U_0 must be contained in A. In particular, this means that $x \in A$. Therefore a limit point of A either lies in $\mathrm{Bdy}\,(A)$ or it lies in A, and therefore our assumptions imply that $\mathbf{L}(A) \subset A \cup \mathrm{Bdy}\,(A) \subset A$, which implies that A is closed in X.■

17. Since the intersection on the right hand side is an open subset of $\cap_i A_i$, it follows that the intersection of the interiors is contained in the interior of the intersection. To prove the converse, use the identity

$$\mathrm{Int}\,(A) \quad = \quad X \; - \; \overline{X-A}$$

established in part (iii) of the preceding exercise. We then have

$$\mathrm{Int}\,(\cap_i A_i) \quad = \quad X \; - \; \overline{(X - \cap_i A_i)} \quad = \quad X \; - \; \overline{\cup_i X - A_i} \quad =$$

$$X \; - \; \cup_i \overline{X - A_i} \quad = \quad \cap_i X - \overline{X - A_i}$$

and by the identity from (iii) this is just $\cap_i \mathrm{Int}\,(A_i)$.■

18. Suppose that $x \in \mathrm{Bdy}\,(C \cap Y, Y)$. Then if V is an open neighborhood of x in Y, we know that V contains at least one point of $C \cap Y$ and at least one point of $Y - (C \cap Y)$. Let U be an open neighborhood of x in X, and let $V = U \cap Y$. By the first sentence we know that there is at least one point of $C \cap Y \subset C$ in $V \subset U$ and at least one point of $Y - (C \cap Y) \subset X - C$ in $V \subset U$.

It is not difficult to find examples with proper containment when $X = \mathbb{R}$. For example, if we let $Y = \mathbb{R} - \{0\}$ and C is the open interval $(0, 1)$ then the boundary of $C = C \cap Y$ in Y is $\{1\}$ but the boundary in $\mathbb{R}$ is $\{0, 1\}$.■

II.3 : Continuous functions

Problems from Munkres, § 18, pp. 111 – 112

2. If $f : X \to Y$ is continuous with $A \subset X$, and x is a limit point of A, then $f(x)$ is NOT NECESSARILY a limit point of $f[A]$. For example, if f is constant then $f[A]$ has no limit points.■

5. It suffices to take the standard linear map sending 0 to a and 1 to b; namely, $f(t) = a+t(b-a)$. The inverse map is given by the formula $g(u) = (u - a)/(b - a)$.■

6. Let $f(x) = x$ if x is rational and 0 if x is irrational. Then f is continuous at 0 because $|x| < \varepsilon \implies |f(x| < \varepsilon$. We claim that f is not continuous anywhere else. What does it mean in terms of δ and ε for f to be discontinuous at x? For some $\varepsilon > 0$ there is no $\delta > 0$ such that $|t - x| < \delta \implies |f(t) - f(x)| < \varepsilon$. Another way of putting this is that for some ε and all $\delta > 0$ sufficiently small, one can find a point t such that $|t - x| < \delta$ and $|f(t) - f(x)| \geq \varepsilon$.

There are two cases depending upon whether $x \neq 0$ is rational or irrational.

The rational case. Let $\varepsilon = |x|/2$ and suppose that $\delta < |x|$. Then there is an irrational number y such that $|y - x| < \delta$, and $|f(y) - f(x)| = |x| > \varepsilon$. Therefore f is not continuous at x.■

The irrational case. The argument is nearly the same. Let $\varepsilon = |x|/2$ and suppose that $\delta < |x|/4$. Then there is a rational number y such that $|y - x| < \delta$, and $|f(y) - f(x)| = |f(y)| > 3|x|/4 > \varepsilon$. Therefore f is not continuous at x.■

8. [*Only for the special case $X = \mathbb{R}$ where the order topology equals the standard topology*.]

(a) See the first proof of Additional Exercise 1 below.■

9. (c) The idea is to find an open covering by sets U_β such that each restriction $f|U_\beta$ is continuous; the continuity of f will follow immediately from this. Given $x \in X$, let U_x be an open subset containing x that is disjoint from all but finitely many closed subsets in the given family. Let $\alpha(1), \cdots \alpha(k)$ be the indices such that $U_x \cap A_\alpha = \emptyset$ unless $\alpha = \alpha(j)$ for some j. Then the subsets $A_{\alpha(j)} \cap U_x$ form a finite closed covering of the latter, and our assumptions imply that the restriction of f to each of these subsets is continuous. But this implies that the restriction of f to the open subset U_a is also continuous, which is exactly what we wanted to prove.■

Additional exercises

1. The easiest examples are those for which the image of $\mathbb{R}$ is neither open nor closed. One example of this sort is

$$f(x) = \frac{x^2}{x^2 + 1}$$

whose image is the half-open interval $(0, 1]$.■

2. FIRST SOLUTION. First of all, if $f : X \to \mathbb{R}$ is continuous then so is $|f|$ because the latter is the composite of a continuous function (absolute value) with the original continuous function and thus is continuous.

We claim that the set of points where $f \geq g$ is closed in X and likewise for the set where $g \geq f$ (reverse the roles of f and g to get this conclusion). But $f \geq g \Longleftrightarrow f - g \geq 0$, and the latter set is closed because it is the inverse image of the closed subset $[0, \infty)$ under the continuous mapping $f - g$.

Let A and B be the closed subsets where $f \geq g$ and $g \geq f$ respectively. Then the maximum of f and g is defined by f on A and g on B, and since this maximum function is continuous on the subsets in a finite closed covering of X, it follows that the global function (the maximum) is continuous on all of X. Similar considerations work for the minimum of f and g, the main difference being that the latter is equal to g on A and f on B.■

SECOND SOLUTION. First of all, if $f : X \to \mathbb{R}$ is continuous then so is $|f|$ because the latter is the composite of a continuous function (absolute value) with the original continuous function and thus is continuous. One then has the following formulas for $\max(f, g)$ and $\min(f, g)$ that immediately imply continuity:

$$\max(f,g) \quad = \quad \frac{f+g}{2} + \frac{|f-g|}{2}$$

$$\min(f,g) \quad = \quad \frac{f+g}{2} - \frac{|f-g|}{2}$$

Verification of these formulas is a routine exercises that is left to the reader to fill in; for each formula there are two cases depending upon whether $f(x) \leq g(x)$ or vice versa.■

3. (a) Suppose that f is open. Then $\mathrm{Int}(A) \subset A$ implies that $f[\,\mathrm{Int}(A)\,]$ is an open set contained in $f[A]$; since $\mathrm{Int}(\,f[A]\,)$ is the largest such set it follows that $f[\,\mathrm{Int}(A)\,] \subset \mathrm{Int}(\,f[A]\,)$.■

Conversely, if the latter holds for all A, then it holds for all open subsets U and reduces to $f[U] \subset \mathrm{Int}(\,f[U]\,)$. Since the other inclusion also holds (every set contains its interior), it follows that the two sets are equal and hence that $f[U]$ is open in Y.■

Suppose now that f is closed. Then $A \subset \overline{A}$ implies that $f[A] \subset f[\overline{A}]$, so that the latter is a closed subset containing $f[A]$. Since $\overline{f[A]}$ is the smallest such set, it follows that $\overline{f[A]} \subset f[\,\overline{A}\,]$.■

Conversely, if the latter holds for all A, then it holds for all closed subsets F and reduces to $\overline{f[F]} \subset f[F]$. Once again the other inclusion also holds (each set is contained in its closure), and therefore the two sets are equal and $f[F]$ is closed in Y.■

(b) To see the statement about continuous and closed maps, note that if f is continuous then for all $A \subset X$ we have $f[\,\overline{A}\,] \subset \overline{f[A]}$ (this is the third characterization of continuity in the course notes), while if f is closed then we have the reverse inclusion. This proves the ($\Longrightarrow$) direction. To prove the reverse implication, split the set-theoretic equality into the two containment relations given in the first sentence of this paragraph. One of the containment relations implies that f is continuous and the other implies that f is closed.■

To see the statement about continuous and open maps, note that f is continuous if and only if for all $B \subset Y$ we have $f^{-1}[\,\mathrm{Int}(B)\,] \subset \mathrm{Int}\left(f^{-1}[B]\right)$ (this is the sixth characterization of continuity in the course notes). Therefore it will suffice to show that f is open if and only if the reverse inclusion holds for all $B \subset Y$. Suppose that f is open and $B \subset Y$. Then by our characterization of open mappings we have

$$f\left(\,\mathrm{Int}\left(f^{-1}[B]\right)\,\right) \subset \mathrm{Int} f\left[\,f^{-1}[B]\,\right] \subset \mathrm{Int}(B)$$

and similarly if we take inverse images under f; but the containment of inverse images extends to a longer chain of containments:

$$\mathrm{Int}\left(f^{-1}[B]\right) \quad \subset \quad f^{-1}\left[\, f\left[\mathrm{Int}\left(f^{-1}[B]\right)\,\right]\,\right] \quad \subset \quad \mathrm{Int}\left(f^{-1}[B]\right) \quad \subset$$

$$f^{-1}\left[f\left[\mathrm{Int}\left(f^{-1}[B]\,\right)\,\right]\,\right] \quad \subset f^{-1}\left[\,\mathrm{Int}(B)\,\right]$$

This proves the ($\Longrightarrow$) implication. What about the other direction? If we set $B = f[A]$ the hypothesis becomes

$$\mathrm{Int}\left(f^{-1}[\,f[A]\,]\right) \quad \subset \quad f^{-1}\left[\mathrm{Int}(\,f[A]\,)\right]$$

and if we take images over f the containment relation is preserved and extends to yield

$$f\left[\,\mathrm{Int}\left(\,f^{-1}[\,f[A]\,]\right)\,\right] \quad \subset \quad f\left[\,f^{-1}\left[\,\mathrm{Int}(\,f[A]\,)\right]\,\right] \quad \subset \quad \mathrm{Int}(\,f[A]\,)\ .$$

Since $A \subset f^{-1}[\,f[A]\,]$ the left hand side of the previous inclusion chain contains $f[\,\mathrm{Int}(A)\,]$, and if one combines this with the inclusion chain the condition $f[\,\mathrm{Int}(A)\,] \subset \mathrm{Int}(\,f]A]\,)$, which characterizes open mappings, is an immediate consequence.■

4. Suppose that $f : X \to Y$ and $g : Y \to Z$ are both light mappings. For each $z \in Z$ let

$$E_z \quad = \quad (g^\circ f)^{-1}\ \left[\,\{z\}\,\right] \quad = \quad f^{-1}\ \left[\,g^{-1}[\{z\}]\,\right]\ .$$

Likewise, let $F_z = g^{-1}[\{z\}]$. We need to prove that E_z is discrete in the subspace topology; our hypotheses guarantee that F_z is discrete in the subspace topology. Likewise, if we let $H_y = f^{-1}[\{y\}]$, then H_y is also discrete in the subspace topology.

Let $x \in E_z$, and let $y = f(x)$, so that $y \in F_z$. Since F_z is discrete in the subspace topology, the subset $\{y\}$ is both open and closed, and hence one can find an open set $V_y \subset Y$ and a closed set $A_y \subset Y$ such that $V_y \cap F_z = A_y \cap F_z = \{y\}$. If we take inverse images under f and apply standard set-theoretic identities for such subsets, we see that

$$f^{-1}\,[V_y]\ \cap\ E_z \quad = \quad f^{-1}\,[A_y]\ \cap\ E_z \quad = \quad H_y$$

and by the continuity of f we know that $f^{-1}\,[V_y]\ \cap\ E_z$ and $f^{-1}\,[A_y]\ \cap\ E_z$ are respectively open and closed in E_z. Since each of these intersections is H_y, it follows that the set H_y is open and closed in E_z. As noted at the end of the previous paragraph we also know that $\{x\}$ is open and closed in H_y. Now if $C \subset B \subset D$ such that C is open (*resp., closed*) in B and B is also is open (*resp., closed*) in D, then C is open (*resp., closed*) in D by the standard properties of the subspace topology. Therefore we have shown that $\{x\}$ is both open and closed in the subspace topology for E_z, and since x was arbitrary this means that E_z must be discrete in the subspace topology.■

POSTSCRIPT. If X is a topological space such that one-point subsets are always closed (for example, if X comes from a metric space), then of course F_z and E_z are discrete closed subsets and have no limit points.■

5. Consider the line defined by the parametric equations $x(t) = at$, $y(t) = bt$ where a and b are not both zero; the latter is equivalent to saying that $a^2 + b^2 > 0$. The value of the function $f(at, bt)$ for $t \neq 0$ is given by the following formula:

$$f(at, bt) = \frac{a^2t^2 - b^2t^2}{a^2t^2 + b^2t^2} = \frac{a^2 - b^2}{a^2 + b^2}$$

If $a = 0$ or $b = 0$ then this expression reduces to 1, while if $a = b = 1$ this expression is equal to 0. Therefore we know that for every $\delta > 0$ and $t < \delta/4$ the points $(t, 0)$ and (t, t) lie in the open disk of radius δ about the origin and the values of the function at these points are given by $f(t, 0) = 1$ and $f(t, t) = 0$. If the function were continuous at the origin and its value was equal to L, then we would have that $|L|$, $|L - 1| < \varepsilon$ for all $\varepsilon > 0$. No such number exists, and therefore the function cannot be made continuous at the origin.∎

6. (*a*) Direct computation shows that $f(t, at)$ is equal to

$$\frac{2\,a\,t^3}{t^4 + a^2t^2} \quad = \quad \frac{2\,a\,t}{t^2 + a^2}$$

and the limit of this expression as $t \to 0$ is zero **provided** $a \neq 0$. Strictly speaking, this is not enough to get the final conclusion, for one also has to analyze the behavior of the function on the x-axis and y-axis. But for the nonzero points of the x-axis one has $f(t, 0) = 0$ and for the nonzero points of the y-axis one has $f(0, t) = 0$.∎

(*b*) Once again, we can write out the composite function explicitly:

$$f(t, t^2) = \frac{2\,t^4}{t^4 + t^4} \quad = \quad 1 \qquad \text{(provided } \; t \neq 0)$$

The limit of this function as $t \to 0$ is clearly 1.

One could give another $\varepsilon - \delta$ proof to show the function is not continuous as in the preceding exercise, but here is another approach by contradiction: Suppose that f is continuous at the origin. Since φ and ψ are continuous functions, it follows that the composites $f \circ \varphi$ and $f \circ \psi$ are continuous at $t = 0$, and that their values at zero are equal to $f(0, 0)$. What can we say about the latter? Using $f \circ \varphi$ we compute it out to be zero, but using $f \circ \psi$ it computes out to $+1$. This is a contradiction, and it arises from our assumption that f was continuous at the origin.∎

7. (*i*) If $\varepsilon > 0$, then $\mathbf{d}_X(u, v) < \varepsilon/r$ implies that $\mathbf{d}_Y(f(u), f(v)) = r \cdot \mathbf{d}_X(u, v) < r \cdot (\varepsilon/r) = \varepsilon$.∎

(*ii*) If $u = f^{-1}(a)$ and $v = f^{-1}(b)$, then we have $r \cdot \mathbf{d}_X(u, v) = \mathbf{d}_Y(a, b)$ because $f(u) = a$ and $f(v) = b$. Dividing by r, we see that $\mathbf{d}_X\left(f^{-1}(a), f^{-1}(b)\right) = \mathbf{d}_X(u, v) = r^{-1}\,\mathbf{d}_Y(a, b)$, which means that f^{-1} is a similarity transformation with ratio of similitude r^{-1}.∎

(*iii*) We have

$$\mathbf{d}_Z\left(g \circ f(u), g \circ f(v)\right) \quad = \quad s \cdot \mathbf{d}_Y\left(f(u), f(v)\right) \quad = \quad s \cdot r \cdot \mathbf{d}_X(u, v)$$

so that $g \circ f$ is a similarity transformation with ration of similitude equal to sr.∎

8. Since f is 1–1 and $f[A] = B$, it follows that f maps $X - A$ into $Y - B$, and general considerations about subspaces imply that the induced map $g : X - A \to Y - B$ is continuous and 1–1 onto. We need to show that g is also an open map.

Let U be an open subset of $X - A$, and write $U = V \cap (X - A)$ where U is open in X. Since f is 1–1 we then have

$$g[U] \quad = \quad f[U] \quad = \quad f[V \cap (X - A)] \quad = \quad f[V] \cap (Y - B) \ .$$

Since f is a homeomorphism, $f[V]$ is open in Y and therefore the right hand side of the display is an open subset of $Y - B$. Thus g is open, and as noted above this shows that g is a homeomorphism.∎

9. (i) $(\implies)$ If $f : (X, \mathbf{T}) \to (\mathbb{R}, \mathbf{U})$ is continuous, then for each $b \in \mathbb{R}$ the inverse image W_b of (b, ∞) is open in X. If we apply this to $(f(x) - \varepsilon, \infty)$ then for every $x \in X$ this inverse image is an open neighborhood of x on which $f(t) > f(x) - \varepsilon$. $(\impliedby)$ Let $b \in \mathbb{R}$, and define W_b as in (i); we want to show W_b is open. By the hypothesis, if $x \in W_b$ and $\varepsilon = f(x) - b > 0$ then there is some open neighborhood U_x of x such that $f(t) > f(x) - \varepsilon = b$ on U_x. Thus for each $x \in W_b$ we have constructed an open neighborhood of x which is contained in U. But if this condition holds, then we know that W_b is an open set.

(ii) We shall prove that the inverse image W_c of $(c\infty)$ is open for all $c \in \mathbb{R}$. There are three cases. (1) If $c \geq 1$ then $W_c = \emptyset$ and hence W_b is open. (2) If $1 > c \geq 0$ then W_c is equal to the open interval (a, b). (3) If $0 \geq c$ then $W_c = \mathbb{R}$.■

(iii) Let $x \in \mathbb{R}$ and let $\varepsilon > 0$. Since $f(x)$ is the least upper bound for the image of $(-infty, x)$, we can find some $y_\varepsilon < x$ such that $f(y_\varepsilon) > f(x) - \varepsilon$, and if t lies in the open set $U_{x,varepsilon} - (y_\varepsilon, \infty)$, then $f(t) > f(x) - \varepsilon$ because f is monotonically increasing. By (i), this means that f is lower semicontinuous. In fact, we have shown that $f : (\mathbb{R}, \mathbf{U}) \to (\mathbb{R}, \mathbf{U})$ is continuous (a stronger conclusion since $\mathbf{M}$ is strictly larger than $\mathbf{U}$).■

(iv) If $a = 0$ then we can factor the map in the form $(\mathbb{R}, \mathbf{U}) \to (\{b\}, \mathbf{U}^*) \to (\mathbb{R}, \mathbf{U})$ where $\mathbf{U}^*$ is the only possible topology on a one point set; both factors of this composite are continuous, and therefore f is continuous. In the more interesting case where $a > 0$, let W_c be the inverse image of (c, ∞) as in (i) and (ii). Then $x \in W_c \quad \Leftrightarrow ax + b > c$, and the condition on the right can be rewritten in the form

$$x \quad > \quad \frac{c - b}{a} \ .$$

This shows that W_c is $\mathbf{U}-$ open and hence that f is continuous.■

(v) If W_c is defined as in (iv), then $W_c = (-\infty, -c)$, which is not $\mathbf{U}-$ open. Therefore f is not continuous.■

II.4 : Cartesian products

Problems from Munkres, § 18, *pp.* 111 – 112

4. It will suffice to prove the first conclusion because the second is obtained by interchanging the roles of X and Y and replacing y_0 by x_0.

The map $f : X \to X \times \{y_0\}$ sending x to (x, y_0) is continuous because its projection onto the x-coordinate is the identity on X and its projection onto the y-coordinate is a constant mapping. If g is the restriction of the coordinate projection $X \times Y \to X$ to $X \times \{y_0\}$, then g is a composite of continuous mappings and hence is continuous. Since $g \circ f$ and $f \circ g$ are identity mappings, it follows that g is a continuous inverse to f.■

10. This is worked out and generalized in the course notes.■

11. Consider the maps $A(y_0) : X \to X \times Y$ and $B(x_0) : Y \to X \times Y$ defined by $[A(y_0)](x) = (x, y_0)$ and $[B(x_0)](y) = (x_0, y)$. Each of these maps is continuous because its projection onto one factor is an identity map and its projection onto the other is a constant map. The maps h and k are composites $F \circ A(y_0)$ and $f \circ B(x_0)$ respectively; since all the factors are continuous, it follows that h and k are continuous.■

FOOTNOTE. The next problem in Munkres gives the standard example of a function from $\mathbb{R}^2 \to \mathbb{R}$ that is continuous in each variable separately but not continuous at the origin. See also the solution to Additional Exercise 5 below.

Problem from Munkres, § 20, pp. 126 – 129

3. (*b*)
Let $\mathbf{U}$ be the metric topology, and let $\mathbf{V}$ be a topology for which

$$d : (X, \mathbf{V}) \;\times\; (X, \mathbf{V}) \;\longrightarrow\; \mathbb{R}$$

is continuous. We need to show that $\mathbf{V}$ contains every subset of the form $N_\varepsilon(x_0)$ where $\varepsilon > 0$ and $x_0 \in X$. The solution to this exercise will use the solution to Munkres, Exercise 18.4.

Consider the composite

$$(X, \mathbf{V}) \;\cong\; (X, \mathbf{V}) \;\times\; \{x_0\} \;\subset\; (X, \mathbf{V}) \;\times\; (X, \mathbf{V}) \;\longrightarrow\; \mathbb{R}\ .$$

This map is continuous, and therefore the inverse image of $(0, \varepsilon)$ is open. One can check directly that this inverse image is just $N_\varepsilon(x_0)$, and hence by the reasoning of the first paragraph we know that $\mathbf{V}$ contains $\mathbf{U}$.■

Note. The containment may be proper. For example, if $\mathbf{U}$ is the usual metric topology in $\mathbb{R}$ then we can take $\mathbf{V}$ to be the discrete topology.■

Additional exercises

1. The basic idea is to give axioms characterizing cartesian products and to show that they apply in this situation.

LEMMA. *Let $\{A_\alpha \mid \alpha \in \mathcal{A}\}$ be a family of nonempty sets, and suppose that we are given data consisting of a set P and functions $h_\alpha : P \to A_\alpha$ such that for* **EVERY** *collection of data $(S, \{f_\alpha : S \to A_\alpha\})$ there is a unique function $f : S \to P$ such that $h_\alpha \circ f = f_\alpha$ for all α. Then there is a unique $1-1$ correspondence $\Phi : \prod_\alpha A_\alpha \to P$ such that $h_\alpha \circ \Phi$ is the projection from $\prod_\alpha X_\alpha$ onto A_α for all α.*

If we suppose in addition that each A_α is a topological space, that P is a topological space, that the functions h_α are continuous and the unique map f is always continuous, then Φ is a homeomorphism to $\prod_\alpha A_\alpha$ with the product topology.

Sketch of the proof of the Lemma. The existence of Φ follows directly from the hypothesis. On the other hand, the data consisting of $\prod_\alpha A_\alpha$ and the coordinate projections π_α also satisfies the given properties. Therefore we have a unique map Ψ going the other way. By the basic conditions the two respective composites $\Phi \circ \Psi$ and $\Psi \circ \Phi$ are completely specified by the maps $\pi_\alpha \circ \Phi \circ \Psi$ and $h_\alpha \circ \Phi \circ \Psi$. Since $\pi_\alpha \circ \Phi = h_\alpha$ and $h_\alpha \circ \Phi = \pi_\alpha$ hold by construction, it follows that $\pi_\alpha \circ \Phi \circ \Psi = \pi_a lpha$ and $h_\alpha \circ \Phi \circ \Psi = h_\alpha$ and by the uniqueness property it follows that both of the composites $\Phi \circ \Psi$ and $\Psi \circ \Phi$ are identity maps. Thus Φ is a 1–1 correspondence.

Suppose now that everything is topologized. What more needs to be said? In the first place, The product set with the product topology has the unique mapping property for continuous maps. This means that both Φ and Ψ are continuous and hence that Φ is a homeomorphism.■

Application to the exercise. We shall work simultaneously with sets and topological spaces, and morphisms between such objects will mean set-theoretic functions or continuous functions in the respective cases.

For each β let $_\beta$ denote the product of objects whose index belongs to $\mathcal{A}_\beta$ and denote its coordinate projections by p_α. The conclusions amount to saying that there is a canonical morphism from $\prod_\beta P_\beta$ to $\prod_\alpha A_\alpha$ that has an inverse morphism. Suppose that we are given morphisms f_α from the same set S to the various sets A_α. If we gather together all the morphisms for indices α lying in a fixed subset $\mathcal{A}_\beta$, then we obtain a unique map $g_\beta : S \to P_\beta$ such that $p_\alpha \circ g_\beta = f_\alpha$ for all $\alpha in \mathcal{A}_\beta$. Let $q_\beta : \prod_\gamma P_\gamma \to P_\beta$ be the coordinate projection. Taking the maps g_β that have been constructed, one obtains a unique map $F : S \to \prod_\beta P_\beta$ such that $q_\beta \circ F = g_\beta$ for all β. By construction we have that $p_\alpha \circ q_\beta \circ F = f_\alpha$ for all α. If there is a unique map with this property, then $\prod_\beta P_\beta$ will be isomorphic to $\prod_\alpha A_\alpha$ by the lemma. But suppose that θ is any map with this property. Once again fix β. Then $p_\alpha \circ q_\beta \circ F = p_\alpha \circ q_\beta \circ \theta = f_\alpha$ for all $\alpha \in \mathcal{A}_\beta$ implies that $q_\beta \circ F = q_\beta \circ \theta$, and since the latter holds for all β it follows that $F = \theta$ as required.■

2. (a) First part: *X is irreducible if and only if every pair of nonempty open subsets has a nonempty intersection.* We shall show that the negations of the two statements in the second sentence are equivalent; *i.e.*, The space X is reducible (not irreducible) if and only if some pair of nonempty open subsets has a nonempty intersection. This follows because X is reducible $\iff$ we can write $X = A \cup B$ where A and B are nonempty proper closed subsets $\iff$ we can find nonempty proper closed subsets A and B such that $(X - A) \cap (X - B) = \emptyset$ $\iff$ we can find nonempty proper open subsets U and V such that $U \cap V = \emptyset$ (take $U = X - A$ and $V = Y - B$).

Second part: *An open subset of an irreducible space is irreducible.* The empty set is irreducible (it has no nonempty closed subsets), so suppose that W is a nonempty open subset of X where X is irreducible. But if U and V are nonempty open subspaces of W, then they are also nonempty open subspaces of X, which we know is irreducible. Therefore by the preceding paragraph we have $U \cap V \neq \emptyset$, which in turn implies that W is irreducible (again applying the preceding paragraph).■

(b) For the first part, it suffices to note that a space with an indiscrete topology has no nonempty proper closed subspaces. For the second part, note that if X is an infinite set with the finite complement topology, then the closed proper subsets are precisely the finite subsets of X, and the union of two such subsets is always finite and this is always a proper subset of X. Therefore X cannot be written as the union of two closed proper subsets.■

(c) If X is a Hausdorff space and $u,\ v \in X$ then one can find open subsets U and V such that $u \in U$ and $v \in V$ (hence both are nonempty) such that $U \cap V = \emptyset$. Therefore X is not irreducible because it does not satisfy the characterization of such spaces in the first part of **(a)** above.■

3. Let $\gamma : X \to \Gamma_f$ be the set-theoretic map sending x to $f(x)$. We need to prove that f is continuous $\iff$ γ is a homeomorphism. Let $j : \Gamma_f \to X \times Y$ be the inclusion map.

$(\implies)$ If f is continuous then γ is continuous. By construction it is 1–1 onto, and a continuous inverse is given explicitly by the composite $\pi_X \circ j$ where π_X denotes projection onto X.■

$(\impliedby)$ If γ is a homeomorphism then f is continuous because it may be written as a composite $\pi_Y \circ j \circ \gamma$ where each factor is already known to be continuous.■

4. Since A and B are closed in X we know that $A \times A$ and $B \times B$ are closed in $X \times X$. Since A and B are Hausdorff we know that the diagonals Δ_A and Δ_B are closed in $A \times A$ and $B \times B$ respectively. Since "a closed subset of a closed subset is a closed subset" it follows that Δ_A and Δ_B are closed in $X \times X$. Finally $X = A \cup B$ implies that $\Delta_X = \Delta_A \cup \Delta_B$, and since each summand on the right hand side is closed in $X \times X$ it follows that the left hand side is too. But this means that X is Hausdorff.■

EXAMPLE. Does the same conclusion hold if A and B are open? NO. Consider the topology on $\{1,\ 2,\ 3\}$ whose open subsets are the empty set and all subsets containing 2. Then both $\{1,\ 3\}$ and $\{2\}$ are Hausdorff with respect to the respective subspace topologies, but there union — which is

X — is not Hausdorff because all open sets contain 2 and thus one cannot find nonempty open subsets that are disjoint.■

5. For each α let $g_\alpha = f_\alpha^{-1}$. Then we have

$$\prod_\alpha f_\alpha \circ \prod_\alpha g_\alpha \quad = \quad \prod_\alpha (f_\alpha \circ g_\alpha) \quad = \quad \prod_\alpha \mathrm{id}(Y_\alpha) \quad = \quad \mathrm{id}\big(\textstyle\prod_\alpha Y_\alpha\big)$$

and we also have

$$\prod_\alpha g_\alpha \circ \prod_\alpha f_\alpha \quad = \quad \prod_\alpha (g_\alpha \circ f_\alpha) \quad = \quad \prod_\alpha \mathrm{id}(X_\alpha) \quad = \quad \mathrm{id}\big(\textstyle\prod_\alpha X_\alpha\big)$$

so that the product of the inverses $\prod_\alpha g_\alpha$ is an inverse to $\prod_\alpha f_\alpha$.■

6. Let π_i be projection onto the i^{th} factor for $i = 1, 2 3$. The map T is continuous if and only if each $\pi_i \circ T$ is continuous. But by construction we have $\pi_1 \circ T = \pi_3$, $\pi_2 \circ T = \pi_1$, and $\pi_3 \circ T = \pi_2$, and hence T is continuous.

We can solve directly for T^{-1} to obtain the formula $T^{-1}(u, v, w) = (v, w, u)$. We can prove continuity by looking at the projections on the factors as before, but we can also do this by checking that $T^{-1} = T^2$ and thus is continuous as the composite of continuous functions.■

7. (*a*) This is a special case of the estimates relating the $\mathbf{d}_1$, $\mathbf{d}_2$ and $\mathbf{d}_\infty$ metrics.■

(*b*) The set of points such that $|x|_\alpha < 1$ is open, and if $|x|_\alpha = 1$ is open and U is an open neighborhood of x, then U contains the points $(1 \pm t) \cdot x$ for all sufficiently small values of $|t|$.■

(*c*) Define $h(x) = |x|_\alpha \cdot |x|_\beta^{-1} \cdot x$ if $X \neq 0$ and $h(0) = 0$. By (a), the mapping h is continuous with respect to the α-norm if and only if if is continuous with respect to the β-norm, and in fact each norm is continuous with respect to the other. It follows immediately that h is continuous if $x \neq 0$. To see continuity at 0, it suffices to check that if $\varepsilon > 0$ then there is some $\delta > 0$ such that $|x|_\alpha < \delta$ implies $|h(x)|_\alpha = |x|_\beta < \varepsilon$. Since $|x|_\beta \leq |x|_\alpha / m$, we can take $\delta = m \cdot \varepsilon$.

To see that h is a homeomorphism, let k be the map constructed by interchanging the roles of α and β in the preceding discussion. By construction k is an inverse function to h, and the preceding argument together with the inequality $|x|_\alpha \leq M \cdot |x|_\beta$ imply that k is continuous.■

(*d*) In the preceding discussion we have constructed a homeomorphism which takes the set of all points satisfying $|x|_\beta \leq 1$ to the set defined by $|x|_\alpha \leq 1$, and likewise if the inequality is replaced by equality. If $\beta = \infty$ then the domain is the hypercube and the set of points with $|x|_\beta = 1$ is its frontier, and if $\alpha = 2$ then the codomain is the ordinary unit disk and the set of points where $|x|_\alpha = 1$ is the unit sphere which bounds that disk.■

8. We shall give a unified proof which works for all values of p (the main values of concern are $p = 1, 2, \infty$, but the same argument works for all p such that $1 \leq p \leq \infty$). Throughout this discussion W_1 and W_2 are metric spaces and $u, v \in W_1 \times W_2$ are expressed in terms of coordinates as (u_1, u_2) and (v_1, v_2) respectively.

Let $\|\cdots\|_p$ denote the $p-$norm on $\mathbb{R}^2$ for p as above (this actually works for all p such that $1 \leq p \leq \infty$). Then

$$\mathbf{d}^{\langle p \rangle}(u, v) \quad = \quad \|\,(\mathbf{d}_1(u_1, v_1),\, \mathbf{d}_2(u_2, v_2))\,\|_p \quad = \quad \|\mathbf{D}(u, v)\|_p$$

where the vector $\mathbf{D}(u, v)$ inside the norm sign has coordinates equal to the distances between the coordinates of u and v.

Now suppose that we have similarity transformations $f : X \to X'$ and $g : Y \to Y'$ whose ratios of similitude are both equal to $r > 0$. Let h be the product mapping $f \times g$. Then the definitions and similarity hypotheses imply that

$$\mathbf{D}(h(u), h(v)) \quad = \quad ((\mathbf{d}_1(f(u_1), f(v_1)),\ \mathbf{d}_2(g(u_2), g(v_2))) \quad =$$

$$(r\,\mathbf{d}_1(u_1, v_1),\ r\,\mathbf{d}_2(u_2, v_2)) \quad = \quad r \cdot \mathbf{D}(u, v)$$

and the assertion in the exercise follows if we take $p-$norms of the vectors at the left and right ends of this chain of equations.■

9. If X and Y are discrete, then every one point subset of each is open. Therefore the associated rectangular sets

$$\{x\} \times \{y\} \quad = \quad \{(x, y)\}$$

are open in $X \times Y$. But this means that every one point subset of $X \times Y$ is open, and therefore the product topology on $X \times Y$ is discrete.■

10. (i) Since $\mathrm{Int}\,(A) \times \mathrm{Int}\,(B)$ is an open subset which is contained in $A \times B$, we clearly have $\mathrm{Int}\,(A) \times \mathrm{Int}\,(B) \subset \mathrm{Int}\,(A \times B)$. Conversely, if $(a, b) \in \mathrm{Int}\,(A \times B)$ then there is some basic open set $U \times V$ such that $(a, b) \subset U \times V \subset A \times B$. Since U and V must be contained in A and B respectively, we must have $U \times V \subset \mathrm{Int}\,(A) \times \mathrm{Int}\,(B)$, which means that (a, b) must belong to the latter set.■

(ii) We shall use the identity $\mathrm{Bdy}\,(C) = \overline{C} - \mathrm{Int}\,(C)$ repeatedly. If we apply this to A and B we obtain decompositions

$$\overline{A} \quad = \quad \mathrm{Int}\,(A) \ \cup \ \mathrm{Bdy}\,(A)\ , \qquad \overline{B} \quad = \quad \mathrm{Int}\,(B) \ \cup \ \mathrm{Bdy}\,(B)$$

where in each case the summands are disjoint. By Proposition 7 and the preceding exercise, we know that the closure or interior of a product subset is the product of the closures or interiors of the factors, so that

$$\mathrm{Bdy}\,(A \times B) \quad = \quad \overline{A} \times \overline{B} \ - \ \mathrm{Int}\,(A) \times \mathrm{Int}\,(B)$$

and if we combine this with the preceding display we find that

$$\mathrm{Bdy}\,(A \times B) \quad = \quad \mathrm{Bdy}\,(A) \times \mathrm{Bdy}\,(B) \ \cup \ \mathrm{Bdy}\,(A) \times \mathrm{Int}\,(B) \ \cup \ \mathrm{Int}\,(A) \times \mathrm{Bdy}\,(B)\ .$$

Now the union of the first and second terms is $\mathrm{Bdy}\,(A) \times \overline{B}$ and the union of the first and third terms is $\overline{A} \times \mathrm{Bdy}\,(B)$, and therefore $\mathrm{Bdy}\,(A \times B)$ is equal to the union of $\mathrm{Bdy}\,(A) \times \overline{B}$ and $\overline{A} \times \mathrm{Bdy}\,(B)$.■

11. To simplify the notation we shall denote the closed unit interval by I and the open unit interval by J. Since $J \subset I$ with J open and I closed, it follows that J is contained in the interior of I and and the closure of J is contained in I. The difference $I - J$ is equal to $\{0, 1\}$, and both of these points are boundary points for both I and J, so this means that I is the closure of J (in $\mathbb{R}$) and J is the interior of I. As in the preceding exercise we then know that $I \times I$ is the closure of $J \times J$ (in $\mathbb{R}$) and $J \times J$ is the interior of $I \times I$. Furthermore, if E is either I or J, the second part of the preceding exercise implies that

$$\mathrm{Bdy}\,(E \times E) \quad = \quad \big(\mathrm{Bdy}\,(E) \times \overline{E}\big) \ \cup \ \big(\overline{E} \times \mathrm{Bdy}\,(E)\big) \quad = \quad \{0, 1\} \times I \ \cup \ I \times \{0, 1\}$$

which is equal to $I \times I - J \times J$.■

12. This exercise confirms that the concepts of closure, interior and boundary behave as expected for a fundamental class of plane subsets which arise in multivariable calculus. It might be helpful to look at the drawing in `math205Asolutions02b.pdf` when reading through this solutions.

As usual, we shall follow the hints. As in the preceding exercise, we shall use J and I to denote the open and closed unit intervals.

Let's begin by extending the continuous functions f and g to the entire real line as indicated. In both cases, the construction involves assembling continuous functions on the closed subsets $(-\infty, a]$, $[a, b]$ and $[b, \infty)$; the definitions agree on the overlapping pieces, and therefore both constructions yield continuous functions on all of $\mathbb{R}$.

Next, we shall justify the claim regarding homeomorphisms. Since a homeomorphism φ is a 1–1 correspondence which sends open sets to open sets and closed sets to closed sets (and likewise for its inverse), it follows that closures and interiors correspond: The closure of a subset A is mapped onto the closure of $\varphi[A]$ and the interior of a subset A is mapped onto the interior of $\varphi[A]$. Since boundaries can be characterized in terms of closures and interiors, it also follows that the boundary of a subset A is mapped onto the boundary of $\varphi[A]$. In particular, if $H : \mathbb{R}^2 \to \mathbb{R}^2$ is a homeomorphism which sends $J \times J$ to the subset V defined in the exercise and also sends $I \times I$ to the subset A, then we have the following:

> H maps the closure of $J \times J$, which is $I \times I$, to the closure of $V = H[J \times J]$, and hence $A = H[I \times I]$ must be the closure of V.
>
> H maps the interior of $I \times I$, which is $J \times J$, to the closure of $A = H[I \times I]$, and hence $V = H[J \times J]$ must be the interior of A.
>
> H maps the boundaries of $I \times I$ and $J \times J$, which are equal to $I \times I - J \times J$ into a subset which must be the boundary of both $A = H[I \times I]$ and $V = H[J \times J]$, and this set is equal to $H[I \times I - J \times J] = H[I \times I] - H[J \times J] = V - A$.

Therefore the proof reduces to constructing a homeomorphism H with the desired properties.

The homeomorphism is given by $H(s, t) = (x, y)$, where x and y are the functions of s and t which are defined by $x = a + s(b - a)$ and $y = g(x) + t(f(x) - g(x))$. We can show that H is a homeomorphism by solving these equations explicitly for s and t in terms of x and y:

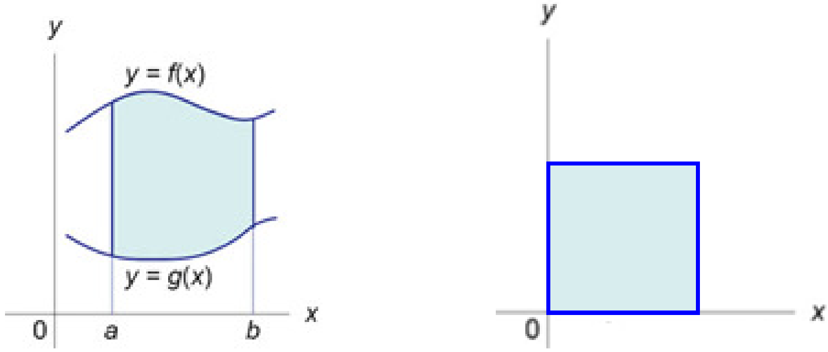
$$s \quad = \quad \frac{x - a}{b - a} \ , \qquad t \quad = \quad \frac{y - g(x)}{f(x) - g(x)}$$

At this point the only things left to check are that H maps $J \times J$ onto V and H maps $I \times I$ onto A. The formulas and elementary inequalities imply that if $(s, t) \in J \times J$ then $H(x, t)$ satisfies the strict inequalities which define V, and if $(s, t) \in I \times I$ then $H(x, t)$ satisfies the inequalities which define S. Conversely, we can use the formulas for the inverse function to show that if $(x, y) \in A$ then $(s, t) \in I \times I$ and if $(x, y) \in V$ then $(s, t) \in J \times J$.■

Drawing to accompany Additional Exercise II.4.12

Here is a drawing of a typical region being considered in this exercise. We are actually interested in two regions, one of which is the closed region A consisting of all points (x, y) where $a \leq x \leq b$ and $g(x) \leq y \leq f(x)$ and the other of which is the open region V consisting of all points (x, y) where $a < x < b$ and $g(x) < y < f(x)$.

Intuitively it probably seems clear that A should be the closure of V and V should be the interior of A, and that the boundaries of both regions should be the points in $A - V$. The purpose of the exercise is to justify this intuition.



(***Source:*** **http://www.math24.net/definite-integral.html**)

The idea is to set up a comparison with a fundamental example; namely the solid square region defined by $0 \leq x, y \leq 1$. In this case everything can be analyzed in a straightforward manner, and we generalize to regions like A and V by constructing a homomorphism from the square to A. More precisely, we construct a homeomorphism of the coordinate plane to itself which sends the solid square to A and its interior points to V (and its boundary points to $A - V$).

SOLUTIONS TO EXERCISES FOR

MATHEMATICS 205A — Part 3

Fall 2014

III. Spaces with special properties

III.1 : Compact spaces – I

Problems from Munkres, § 26, *pp.* 170 – 172

3. Suppose that $A_i \subset X$ is compact for $1 \leq i \leq n$, and suppose that $\mathcal{U}$ is a family of open subsets of X whose union contains $\cup_i A_i$. Then for each i there is a finite subfamily $\mathcal{U}_i$ whose union contains A_i. If we take $\mathcal{U}^*$ to be the union of all these subfamilies then it is finite and its union contains $\cup_i A_i$. Therefore the latter is compact.■

7. We need to show that if $F \subset X \times Y$ is closed then $\pi_X[F]$ is closed in X, and as usual it is enough to show that the complement is open. Suppose that $x \notin \pi_X[F]$. The latter implies that $\{x\} \times Y$ is contained in the open subset $X \times Y - F$, and by the Tube Lemma one can find an open set $V_x \subset X$ such that $x \in V$ and $V_x \times Y \subset X \times Y - F$. But this means that the open set $V_x \subset X$ lies in the complement of $\pi_X[F]$, and since one has a conclusion of this sort for each such x it follows that the complement is open as required.■

8. ($\Longrightarrow$) We shall show that the complement of the graph is open, and this does not use the compactness condition on Y (although it does use the Hausdorff property). Suppose we are given (x, y) such that $y \neq x$. Then there are disjoint open sets V and W in Y such that $y \in V$ and $f(x) \in W$. Since $f^{-1}[W]$ is open in X and contains x, there is an open set U containing x such that $f[U] \subset W$. It follows that $U \times V$ is an open subset of $X \times Y$ that is disjoint from Γ_f. Since we have such a subset for each point in the complement of Γ_f it follows that $X \times Y - \Gamma_f$ is open and that Γ_f is closed.■

($\Longleftarrow$) As in a previous exercises let $\gamma : X \to \Gamma_f$ be the graph map and let $j : \Gamma_f \to X \times Y$ be inclusion. General considerations imply that the map $\pi_X \circ j$ is continuous and 1–1 onto, and γ is the associated set-theoretic inverse. If we can prove that $\pi_X \circ j$ is a homeomorphism, then γ will also be a homeomorphism and then f will be continuous by a previous exercise. The map in question will be a homeomorphism if it is closed, and it suffices to check that it is a composite of closed mappings. By hypothesis j is the inclusion of a closed subset and therefore j is closed, and the preceding exercise shows that π_X is closed. Therefore the composite is a homeomorphism as claimed and the mapping f is continuous.■

Problem from Munkres, § 27, *pp.* 177 – 178

2. (a) The function of x described in the problem is continuous, so its set of zeros is a closed set. This closed set contains A so it also contains $\overline{A}$. On the other hand if $x \notin \overline{A}$ then there is an $\varepsilon > 0$ such that $N_\varepsilon(x) \subset X - \overline{A}$, and in this case it follows that $\mathbf{d}(x, A) \geq \varepsilon > 0$.■

(b) The function $f(a) = \mathbf{d}(x,a)$ is continuous and $\mathbf{d}(x,A)$ is the greatest lower bound for its set of values. Since A is compact, this greatest lower bound is a minimum value that is realized at some point of A.■

(c) The union is contained in $U(A,\varepsilon)$ because $\mathbf{d}(x,a) < \varepsilon$ implies $\mathbf{d}(x,A) < \varepsilon$. To prove the reverse inclusion suppose that y is a point such that $\delta = \mathbf{d}(y,A) < \varepsilon$. It then follows that there is some point $a \in A$ such that $\mathbf{d}(y,a) < \varepsilon$ because the greater than the greatest lower bound of all possible distances. The reverse inclusion is an immediate consequence of the existence of such a point a.■

(d) Let $F = X - U$ and consider the function $g(a) = \mathbf{d}(a,F)$ for $a \in A$. This is a continuous function and it is always positive because $A \cap F = \emptyset$. Therefore it takes a positive minimum value, say ε. If $y \in A \subset U(A,\varepsilon)$ then $\mathbf{d}(a,y) < \varepsilon \leq \mathbf{d}(a,F)$ implies that $y \notin F$, and therefore $U(A,\varepsilon)$ is contained in the complement of F, which is U.■

(e) Take X to be all real numbers with positive first coordinate, let A be the points of X satisfying $y = 0$, and let U be the set of all points such that $y < 1/|x|$. Then for every $\varepsilon > 0$ there is a point not in U whose distance from A is less than ε. For example, consider the points $(2n, 1/n)$.■

Additional exercises

1. Each of the subsets X_n is compact by an inductive argument, and since X is Hausdorff each one is also closed. Since each set in the sequence contains the next one, the intersection of finitely many sets $X_{k(1)}, \cdots, X_{k(n)}$ in the collection is the set $X_{k(m)}$ where $k(m)$ is the maximum of the $k(i)$. Since X is compact, the Finite Intersection Property implies that the intersection A of these sets is nonempty. We need to prove that $f(A) = A$. By construction A is the set of all points that lie in the image of the k-fold composite $\circ^k f$ of f with itself. To see that f maps this set into itself note that if $a = [\circ^k f](x_k)$ for each positive integer k then $f(a) = [\circ^k f](\,f(x_k)\,)$ for each k. To see that f maps this set onto itself, note that $a = [\circ^k f](x_k)$ for each positive integer k implies that

$$a = f\left([\circ^k f](x_{k+1})\right)$$

for each k.■

2. (a) In this situation it is convenient to work with topologies in terms of their closed subsets. Let $\mathcal{F}$ be the family of closed subsets of X associated to $\mathbf{T}$ and let $\mathcal{F}^*$ be the set of all subsets E such that $E \cap C$ is closed in X for every compact subset $C \subset X$. If E belongs to $\mathcal{F}$ then $E \cap C$ is always closed in X because C is closed, so $\mathcal{F} \subset \mathcal{F}^*$. We claim that $\mathcal{F}^*$ defines a topology on X and X is a Hausdorff k-space with respect to this topology.

The empty set and X belong to $\mathcal{F}^*$ because they already belong to $\mathcal{F}$. Suppose that E_α belongs to $\mathcal{F}^*$ for all α; we claim that for each compact subset $C \subset X$ the set $C \cap \bigcap_\alpha E_\alpha$ is $\mathcal{F}$-closed in X. This follows because

$$C \;\cap\; \bigcap_\alpha E_\alpha \;=\; \bigcap_\alpha (C \cap E_\alpha)$$

and all the factors on the right hand side are $\mathcal{F}$-closed (note that they are compact). To conclude the verification that $\mathcal{F}^*$ is a topology, suppose that E_1 and E_2 belong to $\mathcal{F}^*$. Once again let $C \subset X$ be compact, and observe that the set-theoretic equation

$$C \cap (E_1 \cup E_2) \;=\; (C \cap E_1) \;\cup\; (C \cap E_2)$$

implies the right hand side is $\mathcal{F}$-closed if E_1 and E_2 are.

Therefore $\mathcal{F}^*$ defines the closed subspaces of a topological space; let $\mathbf{T}^\kappa$ be the associated family of open sets. It follows immediately that the latter contains $\mathbf{T}$, and one obtains a Hausdorff space by the following elementary observation: *If* $(X, \mathbf{T})$ *is a Hausdorff topological space and* $\mathbf{T}^*$ *is a topology for* X *containing* $\mathbf{T}$*, then* $(X, \mathbf{T}^*)$ *is also Hausdorff.* This is true because the disjoint open sets in $\mathbf{T}$ containing a pair of disjoint points are also (disjoint) open subsets with respect to $\mathbf{T}^*$ containing the same respective points.

We now need to show that $\mathcal{K}(X) = (X, \mathbf{T}^\kappa)$ is a k-space and the topology is the unique minimal one that contains $\mathbf{T}$ and has this property. Once again we switch over to using the closed subsets in all the relevant topologies. The most crucial point is that a subset $D \subset X$ is $\mathcal{F}$-compact if and only if it is $\mathcal{F}^*$-compact; by construction the identity map from $[X, \mathcal{F}^*]$ to $[X, \mathcal{F}]$ is continuous (brackets are used to indicate the subset families are the closed sets), so if D is compact with respect to $\mathcal{F}^*$ it image, which is simply itself, must be compact with respect to $\mathcal{F}$. How do we use this? Suppose we are given a subset $B \subset X$ such that $B \cap D$ is $\mathcal{F}^*$-closed for every $\mathcal{F}^*$-compact subset D. Since the latter is also $\mathcal{F}$-compact and the intersection $B \cap D$ is $\mathcal{F}^*$ compact (it is a closed subspace of a compact space), we also know that $B \cap D$ is $\mathcal{F}$-compact and hence $\mathcal{F}$-closed. Therefore it follows that $\mathcal{F}^*$ is a k-space topology. If we are given the closed subsets for any Hausdorff k-space topology $\mathcal{E}$ containing $\mathcal{F}$, then this topology must contain all the closed sets of $\mathcal{F}^*$. Therefore the latter gives the unique minimal k-space topology containing the topology associated to $\mathcal{F}$.$\blacksquare$

(b) Suppose that $F \subset Y$ has the property that $F \cap C$ is closed in Y for all compact sets $C \subset Y$. We need to show that $f^{-1}[F] \cap D$ is closed in X for all compact sets $D \subset X$.

If F and D are as above, then $f[D]$ is compact and by the assumption on f we know that

$$f^{-1}\left[\, F \cap f[D] \,\right] = f^{-1}[F] \cap f^{-1}\left[\, f[D] \,\right)$$

is closed in X with respect to the original topology. Since $f^{-1}\left[\, f(D] \,\right)$ contains the closed compact set D we have

$$f^{-1}[F] \cap f^{-1}\left(f[D]\right) \cap D \quad = \quad f^{-1}[F] \cap D$$

and since the left hand side is closed in X the same is true of the right hand side. But this is what we needed to prove.$\blacksquare$

3. (a) ($\Longrightarrow$) The sequence of open subsets has a maximal element; let U_N be this element. Then $n \geq N$ implies $U_N \subset U_n$ by the defining condition on the sequence, but maximality implies the reverse inclusion. Thus $U_N = U_n$ for $n \geq N$.$\blacksquare$

($\Longleftarrow$) Suppose that the Chain Condition holds but there is a nonempty family $\mathcal{U}$ of open subsets with no maximal element. If we pick any open set U_1 in this family then there is another open set U_2 in the family that properly contains U_2. Similarly, there is another open subset U_3 in the family that properly contains U_2, and we can inductively construct an ascending chain of open subspaces such that each properly contains the preceding ones. This contradicts the Ascending Chain Condition. Therefore our assumption that $\mathcal{U}$ had no maximal element was incorrect.$\blacksquare$

(b) ($\Longrightarrow$) Take an open covering $\{\, U_\alpha \,\}$ of U for which each open subset in the family is nonempty, and let $\mathcal{W}$ be the set of all finite unions of subsets in the open covering. By definition this family has a maximal element, say W. If $W = U$ then U is compact, so suppose W is properly contained in U. Then if $u \in U - W$ and U_0 is an open set from the open covering that contains u, it will follow that the union $W \cap U_0$ is also a finite union of subsets from the open covering and it properly contains the maximal such set W. This is a contradiction, and it arises from our assumption that W was properly contained in U. Therefore U is compact.$\blacksquare$

($\Longleftarrow$) We shall show that the Ascending Chain Condition holds. Suppose that we are given an ascending chain

$$U_1 \;\subset\; U_2 \;\subset\; \cdots$$

and let $W = \cup_n U_n$. By our hypothesis this open set is compact so the open covering $\{U_n\}$ has a finite subcovering consisting of $U_{k(i)}$ for $1 \le i \le m$. If we take N to be the maximum of the $k(i)$'s it follows that $W = U_N$ and $U_n = U_N$ for $n \ge N$.■

(c) We begin by verifying the statement in the hint. If U is open in a noetherian Hausdorff space X, then U is compact and hence U is also closed (since X is Hausdorff). Since U is Hausdorff, one point subsets are closed and their complements are open, so the complements of one point sets are also closed and the one point subsets are also open. Thus a noetherian Hausdorff space is discrete. On the other hand, an infinite discrete space does not satisfy the Ascending Chain Condition (pick an infinite sequence of distinct points x_k and let U_n be the first n points of the sequence. Therefore a noetherian Hausdorff space must also be finite.■

(d) Suppose $Y \subset X$ where X is noetherian. Let $\mathcal{V} = \{V_\alpha\}$ be a nonempty family of open subspaces of Y, write $V_\alpha = U_\alpha \cap Y$ where U_α is open in X, and let $\mathcal{U} = \{U_\alpha\}$. Since X is noetherian, this family has a maximal element U^*, and the intersection $V^* = U^* \cap Y$ will be a maximal element of $\mathcal{V}$.■

4. Since X is Hausdorff every one point set is closed, and this implies that $\mathbf{L}(A)$ is closed in X. We are assuming that $\overline{A}$ is compact, and since the closed subset $\mathbf{L}(A)$ is contained in $\overline{A}$ it follows that $\mathbf{L}(A)$ is also compact.■

5. (i) Since X is compact we know that $f[X]$ is compact in $\mathbb{R}$ with respect to the lower semicontinuity topology. Every subset $C \subset \mathbb{R}$ has an open covering in this topology consisting of proper open subsets $(b_\alpha, \infty) \cap C$ because every point lies in a proper open subset (with respect to the lower semicontinuity topology). If C is compact, this means that C is contained in a finite union of these sets:

$$C \quad \subset \quad \bigcup_{j=1}^{k} (b_j, \infty)$$

The right hand side is equal to (b^*, ∞) where b^* is the smallest in the finite collection of numbers $\{b_j\}$, and this implies that $x > b^*$ for all $x \in C$. In particular, this applies to $f[X]$ if f is lower semicontinuous and X is compact, and hence we have shown that $f[X]$ has a lower bound.■

(ii) To continue the discussion from (i), let m be the greatest lower bound of $f[X]$. Then for each $n > 0$ the set $F_n = f^{-1}\left[(-\infty, m + \frac{1}{n}]\right]$ is nonempty, for there will be some $x \in X$ such that $f(x) < m - \frac{1}{n}$; since $(-\infty, m + \frac{1}{n}]$ is closed in the lower semicontinuity topology, it follows that F_n is also closed. Since $F_n \supset F_{n+1}$ for all n, this yields a nested sequence of closed subspaces $F_1 \supset \cdots F_n \supset F_{n+1} \cdots$ such that each F_n is nonempty. By the compactness of X we know that their intersection must also be nonempty.

Let $x_0 \in \cap_n F_n$. Then since $x_0 \in F_k$ for each k we have $f(x) \le m + \frac{1}{k}$ for each k. This implies that $f(x) \le m$. However, we also know that m is a lower bound for $f[X]$, and therefore we must have $f(x_0) = m$; in other words, f takes a minimum value at x_0.■

III.2 : Complete metric spaces

Problems from Munkres, § 43, pp. 270 – 271

(a) Let $\{x_n\}$ be a Cauchy sequence in X and choose M so large that $m, n \ge M$ implies $\mathbf{d}(x_m.x_n) < \varepsilon$. Then all of the terms of the Cauchy sequence except perhaps the first $M - 1$ lie in the closure of $N_\varepsilon(x_M)$, which is compact. Therefore it follows that the sequence has a convergent subsequence $\{ x_{n(k)} \}$. Let y be the limit of this subsequence; we need to show that y is the limit of the entire sequence.

Let $\eta > 0$ be arbitrary, and choose $N_1 \geq M$ such that $m,\ n \geq N_1$ implies $\mathbf{d}(x_m, x_n) < \eta/2$. Similarly, let $N_2 \geq M$ be such that $n(k) > N_2$ implies $\mathbf{d}(x_{n(k)}, y) < \eta/2$. If we take N to be the larger of N_1 and N_2, and application of the Triangle Inequality shows that $n \leq N$ implies $\mathbf{d}(x_n, y) < \eta$. Therefore y is the limit of the given Cauchy sequence and X is complete.■

(*b*) Take $U \subset \mathbb{R}^2$ to be the set of all points such that $xy < 1$. This is the region "inside" the hyperbolas $y = \pm 1/x$ that contains the origin. It is not closed in $\mathbb{R}^2$ and therefore cannot be complete. However, it is open and just like all open subsets U of $\mathbb{R}^2$ if $x \in X$ and $N_\varepsilon(x) \subset U$ then $N_{\varepsilon/2}(x)$ has compact closure in U.■

3. (*b*) By the symmetry of the problem it is enough to show that if $(X, \mathbf{d})$ is complete then so is $(X, \mathbf{e})$. Suppose that $\{ x_n \}$ is a Cauchy seuence with respect to $\mathbf{e}$; we claim it is also a Cauchy sequence with respect to $\mathbf{d}$. Let $\varepsilon > 0$, and take $\delta > 0$ such that $\mathbf{e}(u, v) < \delta$ implies $\mathbf{d}(u, v) < \varepsilon$. If we choose M so that $m,\ n \geq M$ implies $\mathbf{e}(x_n, x_m) < \delta$, then we also have $\mathbf{d}(x_n, x_m) < \varepsilon$. Therefore the original Cauchy sequence with respect to $\mathbf{e}$ is also a Cauchy sequence with respect to $\mathbf{d}$. By completeness this sequence has a limit, say y, with respect to $\mathbf{d}$, and by continuity this point is also the limit of the sequence with respect to $\mathbf{e}$.■

6. (*a*) This follows because a closed subspace of a complete metric space is complete.■

(*c*) The image of the function f is the graph of ϕ, and general considerations involving graphs of continuous functions show that f maps U homeomorphically onto its image. If this image is closed in $X \times \mathbb{R}$ then by **(a)** we know that U is topologically complete, so we concentrate on proving that $f(U) \subset X \times \mathbb{R}$ is closed. The latter in turn reduces to proving that the image is closed in the complete subspace $\overline{U} \times \mathbb{R}$, and as usual one can prove this by showing that the complement of $f[U]$ is open. The latter in turn reduces to showing that if $(x, t) \in \overline{U} \times \mathbb{R} - f[U]$ then there is an open subset containing (x, t) that is disjoint from U.

Since ϕ is continuous it follows immediately that the graph of ϕ is closed in the open subset $U \times \mathbb{R}$. Thus the open set $U \times \mathbb{R} - f[U]$ is open in $\overline{U} \times \mathbb{R}$, and it only remains to consider points in $\overline{U} - U \times \mathbb{R}$. Let (x, t) be such a point. In this case one has $\mathbf{d}(x, X - U) = 0$. There are three cases depending on whether t is less than, equal to or greater than 0. In the first case we have that $(x, t) \in \overline{U} \times (-\infty, 0)$ which is open in $\overline{U} \times \mathbb{R}$ and contains no points of $f[U]$ because the second coordinates of points in the latter set are always positive. Suppose now that $t = 0$. Then by continuity of distance functions there is an open set $V \subset \overline{U}$ such that $x \in V$ and $\mathbf{d}(y, X - U) < 1$ for all $y \in V$. It follows that $V \times (-1, 1)$ contains (x, t) and is disjoint from $f[U]$. Finally, suppose that $t > 0$. Then by continuity there is an open set $V \subset \overline{U}$ such that $x \in V$ and

$$\mathbf{d}(y, X - U) \quad < \quad \frac{2}{3t}$$

for all $y \in V$. It follows that $V \times (t/2, 3t/2)$ contains (x, t) and is disjoint from $f[U]$, and this completes the proof of the final case.■

Additional exercises

1. Follow the hint, and try to see what a function in the intersection would look like. In the first place it has to satisfy $f(0) = 1$, but for each $n > 0$ it must be zero for $t \geq 1/n$. The latter means that the $f(t) = 0$ for all $t > 0$. Thus we have determined the values of f everywhere, but the function we obtained is not continuous at zero. Therefore the intersection is empty. Since every function in the set A_n takes values in the closed unit interval, it follows that if f and g belong to A_n then $\|f-g\| \leq 1$ and thus the diameter of A_n is at most 1 for all n. In fact, the diameter is exactly 1 because $f(0) = 1$.

For the sake of completeness, we should note that each set A_n is nonempty. One can construct a "piecewise linear" function in the set that is zero for $t \geq 1/n$ and decreases linearly from the 1 to 0 as t increases from 0 to $1/n$. (Try to draw a picture of the graph of this function!)∎

2. For each positive integer k let $H_k(y)$ be the vector whose first k coordinates are the same as those of y and whose remaining coordinates are zero, and let $T_k = I - H_k$ (informally, these are "head" and "tail" functions). Then H_k and T_k are linear transformations and $|H_k(y)|,\ |T_k(y)|\ \leq |y|$ for all y.

Since Cauchy sequences are bounded there is some $B > 0$ such that $|x_n| \leq B$ for all n.

Let x be given as in the hint. We claim that $x \in \ell^2$; by construction we know that $H_k(x) \in \ell^2$ for all k. By the completeness of $\mathbb{R}^M$ we know that $\lim_{n\to\infty} H_k(x_n) = H_k(x)$ and hence there is an integer P such that $n \geq P$ implies $|H_k(x) - H_k(x_n)| < \varepsilon$. If $n \geq P$ we then have that

$$|H_k(x)| \ \leq \ |H_k(x_n)| + |H_k(x) - H_M(k_n)| \ \leq \ |x_n| + |H_M(x) - H_M(x_n)| \ < \ B + \varepsilon \ .$$

By construction $|x|$ is the least upper bound of the numbers $|H_k(x)|$ if the latter are bounded, and we have just shown the latter are bounded. Therefore $x \in \ell^2$; in fact, the argument can be pushed further to show that $|x| \leq B$, but we shall not need this.

We must now show that x is the limit of the Cauchy sequence. Let $\varepsilon > 0$ and choose M this time so that $n,\, m \geq M$ implies $|x_m - x_n| < \varepsilon/6$. Now choose N so that $k \geq N$ implies $|T_k(x_M)| < \varepsilon/6$ and $|T_k(x)| < \varepsilon/3$; this can be done because the sums of the squares of the coordinates for x_M and x are convergent. If $n \geq M$ then it follows that

$$|T_k(x_n)| \ \leq \ |T_k(x_M)| + |T_k(x_n) - T_k(x_M)| = \ |T_k(x_n)| \ \leq \ |T_k(x_M)| + |T_k(x_n - x_M)| \ \leq$$

$$|T_k(x_M)| + |x_n - x_M| \ \leq \ \frac{\varepsilon}{6} + \frac{\varepsilon}{6} \ = \ \frac{\varepsilon}{3} \ .$$

Choose P so that $P \geq M + N$ and $n \geq P$ implies $|H_N(x) - H_N(x_n)| < \frac{\varepsilon}{3}$. If $n \geq P$ we then have

$$|x - x_n| \ \leq \ |H_N(x) - H_N(x_n)| + |T_N(x) - T_N(x_n)| \ \leq$$

$$|H_N(x) - H_N(x_n)| \ + \ |T_N(x)| \ + \ |T_N(x_n)| \ < \ \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \ = \ \varepsilon \ .$$

This completes the proof.∎

III.3 : Implications of completeness

Problems from Munkres, § 48, pp. 298 – 300

1. If the interior of the closure of B_n is empty, then B_n is nowhere dense. Thus if the interior of the closure of each B_n is empty then X is of the first category.∎

2. View the real numbers as a vector space over the rationals. The previous argument on the existence of bases implies that the set $\{1\}$ is contained in a basis (work out the details!). Let B be such a basis (over the rationals), and let W be the rational subspace spanned by all the remaining vectors in the basis. Then $\mathbb{R}$ is the union of the cosets $c + W$ where c runs over all rational numbers, and this is a countable union. We claim that each of these cosets has a nonempty interior. To see this, note that each coset contains exactly one rational number, while if there interior were nonempty it would contain an open interval and every open interval contains infinitely many rational numbers (*e.g.*, if $c \in (a, b)$ there is a strictly decreasing sequence of rational numbers $r_n \in (c, b)$ whose limit is c).■

4. According to the hint, we need to show that if $\{ V_n \}$ is a sequence of open dense subsets in X, then $(\cap_n V_n)$ is dense.

Let $x \in X$, and let W be an open neighborhood of x that is a Baire space. We claim that the open sets $V_n \cap W$ are all dense in W. Suppose that W_0 is a nonempty open subset of W; then W_0 is also open in X, and since $\cap_n V_n$ is dense in X it follows that

$$W_0 \cap (W \cap V_n) = (W_0 \cap W) \cap V_n = W_0 \cap V_n \neq \emptyset$$

for all n. Therefore $V_n \cap W$ is dense in W. Since W is a Baire space and, it follows that

$$\bigcap_n (V_n \cap W) = W \cap \left(\bigcap_n V_n \right)$$

is dense in W.

To show that $(\cap_n V_n)$ is dense in X, let U be a nonempty open subset of X, let $a \in U$, let W_a be an open neighborhood of a that is a Baire space, and let $U_0 = U \cap W$ (hence $a \in U_0$). By the previous paragraph the intersection

$$U_0 \cap \left(\bigcap_n (W \cap V_n) \right)$$

is nonempty, and since this intersection is contained in

$$U \quad \cap \quad \left(\bigcap_n V_n \right)$$

it follows that the latter is also nonempty, which implies that the original intersection $(\cap_n V_n)$ is dense.■

Problem from Munkres, § 27, pp. 178 – 179

6. PRELIMINARY OBSERVATION. Each of the intervals that is removed from A_{n-1} to construct A_n is entirely contained in the former. One way of doing this is to partition $[0, 1]$ into the 3^n intervals

$$\left[\frac{b}{3^n}, \frac{b+1}{3^n} \right]$$

where b is an integer between 0 and $3^n - 1$. If we write down the unique 3-adic expansion of b as a sum

$$\sum_{i=0}^{n-1} a_i \, 3^i$$

where $b_i \in \{0, 1, 2\}$, then A_k consists of the intervals associated to numbers b such that none of the coefficients b_i is equal to 1. Note that these intervals are pairwise disjoint; if the interval corresponding to b lies in A_k then either the interval corresponding to $b+1$ or $b-1$ is not one of the intervals that are used to construct A_k (the base 3 expansion must end with a 0 or a 2, and thus one of the adjacent numbers has a base 3 expansion ending in a 1). The inductive construction of A_n reflects the fact that for each k the middle third is removed from the closed interval

$$\left[\frac{k}{3^{n-1}}, \frac{k+1}{3^{n-1}}\right] .$$

(a) By induction A_k is a union of 2^k pairwise disjoint closed intervals, each of which has length 3^{-k}; at each step one removes the middle third from each of the intervals at the previous step. Suppose that $K \subset C$ is nonempty and connected. Then for each n the set K must lie in one of the 2^n disjoint intervals of length 3^{-n} in A_n. Hence the diameter of K is $\leq 3^{-n}$ for all $n \geq 0$, and this means that the diameter is zero; *i.e.*, K consists of a single point.■

(b) By construction C is an intersection of closed subsets o the compact space $[0,1]$, and therefore it is closed and hence compact.■

(c) If we know that the left and right endpoints for the intervals comprising A_n are also (respectively) left and right hand endpoints for intervals comprising A_{n+1}, then by induction it will follow that they are similar endpoints for intervals comprising A_{n+k} for all $k \geq 0$ and therefore they will all be points of C. By the descriptions given before, the left hand endpoints for A_n are all numbers of the form $b/3^n$ where b is a nonnegative integer of the form

$$\sum_{i=0}^{n-1} b_i \, 3^i$$

with $b_i = 0$ or 2 for each i, and the right hand endpoints have the form $(b+1)/3^n$ where b has the same form. If b has the indicated form then

$$\frac{b}{3^n} \quad = \quad \frac{3b}{3^{n+1}} \ , \quad \text{where} \quad 3b = \quad \sum_{i=1}^{n} b_{i-1} \, 3^i$$

shows that $b/3^n$ is also a left hand endpoint for one of the intervals comprising A_{n+1}. Similarly, if $(b+1)/3^n$ is a right hand endpoint for an interval in A_n and b is expanded as before, then the equations

$$\frac{b+1}{3^n} \quad = \quad \frac{3b+3}{3^{n+1}}$$

and

$$(3b+3) - 1 \quad = \quad 3b+2 \quad = \quad 2 + \sum_{i=1}^{n} b_{i-1} \, 3^i$$

show that $(b+1)/3^n$ is also a right hand endpoint for one of the intervals comprising A_{n+1}.■

(d) If $x \in C$ then for each n one has a unique closed interval J_n of length 2^{-n} in A_n such that $x \in J_n$. Let $\lambda_n(x)$ denote the left hand endpoint of that interval unless x is that point, and let $\lambda_n(x)$ be the right hand endpoint in that case; we then have $\lambda_n(x) \neq x$. By construction $\left| \lambda_n(x) - x \right| < 2^{-n}$ for all n, and therefore $x = \lim_{n \to \infty} \lambda_n(x)$. On the other hand, by the preceding portion of this problem we know that $\lambda_n(x) \in C$, and therefore we have shown that x is a limit point of C, which means that x is not an isolated point.■

(e) One way to do this is by using (d) and the Baire Category Theorem. By construction C is a compact, hence complete, metric space. It is infinite by (c), and every point is a limit point by (d). Since a countable complete metric space has isolated points, it follows that C cannot be countable.

In fact, one can show that $|C| = 2^{\aleph_0}$. The first step is to note that *if* $\{ a_k \}$ *is an infinite sequence such that* $a_k \in \{0, 2\}$ *for each* k*, then the series*

$$\sum_{k=1}^{\infty} \frac{a_k}{3^k}$$

converges and its sum lies in C*.*

This assertion may be verified as follows: The infinite series converges by a comparison test with the convergent series such that $a_k = 2$ for all k. Given a point as above, the partial sum

$$\sum_{k=1}^{n} \frac{a_k}{3^k}$$

is a left hand endpoint for one of the intervals comprising A_n. The original point will lie in A^n if the sum of the rest of the terms is $\leq 1/3$. But

$$\sum_{k=n+1}^{\infty} \frac{a_k}{3^k} \;\leq\; \sum_{k=n}^{\infty} \frac{2}{3^k} \;=\; \frac{2}{3^{n+1}} \cdot \frac{1}{(1 - \frac{1}{3})} \;=\; \frac{1}{3}$$

so the point does lie in A_n. Since n was arbitrary, this means that the sum lies in $\cap_n \, A_n = C$.

Returning to the original problem of determining $|C|$, we note that the set $\mathcal{A}$ of all sequences described in the assertion is in a natural 1–1 correspondence with the set of all functions from the positive integers to $\{0, 1\}$. Let $\mathcal{A}_0$ be the set of all functions whose values are nonzero for infinitely many values of n, and let $\mathcal{A}_1$ be the functions that are equal to zero for all but finitely many values of n. We then have that $|\mathcal{A}_1| = \aleph_0$ and $\mathcal{A}_0$ is infinite (why?). The map sending a function in $\mathcal{A}_0$ to the associated sum of an infinite series is 1–1 (this is just a standard property of base N expansions — work out the details), and therefore we have

$$|\mathcal{A}_0| \;\leq\; |C| \;\leq\; |\mathbb{R}| \;=\; 2^{\aleph_0}$$

and

$$|\mathcal{A}_0| \;=\; |\mathcal{A}_0| + \aleph_0 \;=\; |\mathcal{A}_0| + |\mathcal{A}_1| \;=\; |\mathcal{A}| \;=\; 2^{\aleph_0}$$

which combine to imply $|C| = 2^{\aleph_0}$.■

Additional exercises

1. Suppose that the conclusion is false: *i.e.*, A is not nowhere dense in X and B is not nowhere dense in Y. Then there are nonempty open sets U and V contained in the closures of A and B respectively, and thus we have

$$\emptyset \;\neq\; U \times V \;\subset\; \overline{A} \times \overline{B} \;=\; \overline{A \times B}$$

and therefore $A \times B$ is not nowhere dense in $X \times Y$.

To see that we cannot replace "or" with "and" take $X = Y = \mathbb{R}$ and let A and B be equal to $[0, 1]$ and $\{0\}$ respectively. Then A is not nowhere dense in X but $A \times B$ is nowhere dense in $X \times Y$.■

2. The space $\mathbb{R}^\infty$ has this property because it is the union of the closed nowhere dense subspaces A_n that are defined by the condition $x_i = 0$ for $i > n$.■

3. %vfil

First of all, the map f is 1–1 onto; we are given that it is onto, and it is 1–1 because $u \neq v$ implies $\mathbf{d}(\, f(u),\, f(v)\,) > \mathbf{d}(u,v) > 0$. Therefore f has an inverse, at least set-theoretically, and we denote f^{-1} by T.

We claim that T satisfies the hypotheses of the Contraction Lemma. The proof of this begins with the relations

$$\mathbf{d}(\, T(u),\, T(v)\,) = \mathbf{d}(\, f^{-1}(u),\, f^{-1}(v)\,) = \frac{1}{C1}\,\mathbf{d}\left(\, f\left(f^{-1}(u)\right),\, f\left(f^{-1}(v)\right)\,\right) = \mathbf{d}(u,v)\ .$$

Since $C > 1$ it follows that $0 < 1/C < 1$ and consequently the hypotheses of the Contraction Lemma apply to our example.

Therefore T has a unique fixed point p; we claim it is also a fixed point for f. We shall follow the hint. Since T is 1–1 and onto, it follows that $x = T(\,T^{-1}(x)\,)$ and that $T(x) = x \implies x = T^{-1}(x)$; the converse is even easier to establish, for if $x = T^{-1}(x)$ the application of T yields $T(x) = x$. Since there is a unique fixed point p such that $T(p) = p$, it follows that there is a unique point, in fact the same one as before, such that $p = T^{-1}(p)$, which is equal to $f(p)$ by definition.■

THE CLASSICAL EUCLIDEAN CASE.

This has two parts. The first is that every expanding similarity of $\mathbb{R}^n$ is expressible as a so-called affine transformation $T(v) = cAv + b$ where A is given by an orthogonal matrix. The second part is to verify that each transformation of the type described has a unique fixed point. By the formula, the equation $T(x) = x$ is equivalent to the equation $x = cAx + b$, which in turn is equivalent to $(I - cA)x = b$. The assertion that T has a unique fixed point is equivalent to the assertion that this linear equation has a unique solution. The latter will happen if $I - cA$ is invertible, or equivalently if $\det(I - cA) \neq 0$, and this is equivalent to saying that c^{-1} is not an eigenvalue of A. But if A is orthogonal this means that $|Av| = |v|$ for all v and hence the only possible eigenvalues are ± 1; on the other hand, by construction we have $0 < c^{-1} < 1$ and therefore all of the desired conclusions follow. The same argument works if $0 < c < 1$, the only change being that one must substitute $c^{-1} > 1$ for $0 < c^{-1} < 1$ in the preceding sentence.■

4. We need to show that φ maps $[\sqrt{a}, x_0]$ into itself and that the absolute value of its derivative takes a maximum value that is less than 1.

In this example, the best starting point is the computation of the derivative, which is simply an exercise in first year calculus:

$$\varphi'(x) \quad = \quad \frac{1}{2}\left(1 - \frac{a}{x^2}\right)$$

This expression is an increasing function of x over the set $[\sqrt{a}, +\infty)$; its value at $\sqrt{a}$ is 0 and the limit at $+\infty$ is $1/2$. In particular, the absolute value of the derivative on $[\sqrt{a}, x_0]$ is less than $1/2$, and by the Mean Value Theorem the latter in turn implies that φ maps the interval in question to

$$\left[\sqrt{a}, \frac{\sqrt{a} + x_0}{2}\right]$$

which is contained in $[\sqrt{a}, x_0]$.■

5. (i) We shall first prove that the interior of H in $\mathbb{R}^n$ is empty. Suppose to the contrary that there is some $p \in H$ and some $\varepsilon > 0$ such that $N_\varepsilon(p) \subset H$. Then by the definition of H we know that $F = 0$ on $N_\varepsilon(p)$. However, we have

$$F(p + ta) \quad = \quad F(p) \ + \ t|a|^2 \quad = \quad t|a|^2$$

so $p + ta \notin H$ for $t \neq 0$; since $t < \varepsilon/|a|$ implies that $p + ta \in N_\varepsilon(p)$, we cannot have $N_\varepsilon(p) \subset H$. This contradiction implies that the interior of H is empty. Note also that H is closed because it is the zero set of a continuous real valued function.

We shall now prove that $\mathbb{R}^n - H$ is dense in $\mathbb{R}^n$. If this were not the case, then the complement of the closure contains some open subset of the form $N_r(q)$, and this open subset would have to be contained in H. Since H has an empty interior, this cannot happen, and therefore $\mathbb{R}^n - H$ must be dense in $\mathbb{R}^n$.■

(*ii*) We have $\mathbb{R}^n - \cup_i H_i = \cap_i (\mathbb{R}^n - H_i)$, and by (*i*) we know that each factor in the intersection is an open dense subset. Since a finite intersection of open and dense subsets is dense, it follows that the set in question is dense.■

6. (*i*) The function in question can be rewritten in the vector form $F(x) = \langle a, x\rangle - b$, where $a = (a_1, \cdots, a_n)$ and $\langle\ ,\ \rangle$ denotes the usual dot product. Suppose that x_0 lies in H, and consider the curve $x_0 + ta$ for $t \in R$. Then $F(x + ta) = \langle a, x + ta\rangle + b = \langle a, x_0\rangle + t|a|^2 - b$, and since x_0 lies on the hyperplane defined by $F(x) = 0$ we have $F(x + ta) = t|a|^2$, where $a \neq 0$. This quantity is positive if $t > 0$ and negative if $t < 0$, so both H_+ and H_- are nonempty. These sets are open by continuity of F, so the only thing left to prove is that these subsets are convex. Following the hint, we shall first prove that $F(ty + (1-t)x) = t\,F(y) + (1-t)\,F(x)$ for all $x, y \in \mathbb{R}^n$ and $t \in \mathbb{R}$:

$$F(ty \;+\; (1-t)x) \;=\; \langle a, ty + (1-t)x\rangle \;-\; b \;=\; t\langle a, y\rangle \;+\; (1-t)\,\langle a, x\rangle \;-\; b \;=$$

$$t\langle a, y\rangle \;+\; (1-t)\,\langle a, x\rangle \;-\; \big(t\,b + (1-t)\,b\big) \;=\; t\,(\langle a, y\rangle \;-\; b) \;+\; (1-t)\,(\langle a, x\rangle \;-\; b) \;=$$

$$t\,F(y) \;+\; (1-t)\,F(x)$$

If $0 < t < 1$ and the values $F(x), F(y)$ are both positive or both negative, the identity and convexity of $(0, \infty) \subset \mathbb{R}$ imply that $F(ty + (1-t)x)$ is also positive or negative. Therefore both H_+ and H_- are convex.■

(*ii*) Suppose that $F(x) > 0 > F(y)$. We want to find some t such that $0 < t < 1$ and $0 = F(tx + (1-t)y) = t\,F(x) + (1-t)\,F(y)$. This equation has the solution

$$t \;=\; \frac{-F(y)}{F(x) - F(y)}$$

so we need to show that the right hand side lies between 0 and 1. Since $F(y) < 0$, the numerator is positive, and likewise for the denominator because $F(x) > 0$ and hence $F(x) - F(y) > F(x) > 0$. Furthermore, the latter chain of inequalities implies that $t < 1$.■

(*iii*) Let $\mathbf{H}$ denote the topology on $\mathbb{R}^n$ generated by half-spaces, and let $\mathbf{M}$ be the usual metric topology. Since half-spaces are open in the metric topology, it follows that $\mathbf{H} \subset \mathbf{M}$. To prove the reverse implication, note that it suffices to prove that the base for $\mathbf{M}$ consisting of the $\mathbf{d}^{\langle\infty\rangle}$ neighborhoods

$$\prod_{j=1}^{n} (a_i - \varepsilon, a_i + \varepsilon)$$

is contained in $\mathbf{H}$. This will follow if we can show that each of the displayed sets is a finite intersection of half-spaces, and the latter in turn follows because the subspace in the display is the intersection of the half-spaces defined by the strict inequalities $x_j - a_j > -\varepsilon$ and $x_j - a_j < \varepsilon$ for $1 \leq j \leq n$.■

Note. Many systems of synthetic axioms for classical geometries include some form of *Space Separation Axiom* which states that the complement of a hyperplane (a line in two dimensions, a plane in three dimensions, *etc.*) satisfies the conclusions of this exercise, and for such a system there is a natural synthetic topology generated by the half-spaces. One implication of the preceding exercise is that this synthetic topology coincides with the usual one if we are working with axioms for Euclidean geometry or closely related systems.

III.4 : Connected spaces

Problems from Munkres, § 23, p. 152

2. Let C be the connected component of $\cup_n A_n$ containing A_1. We shall prove by induction that $A_k \subset C$ for each positive integer k. It will follow that $C = \cup_n A_n$ and that the latter is connected.

Suppose that A_k is contained in C; we want to show that $A_{k+1} \subset C$. We are given that there is at least one point $p \in A_k \cup A_{k+1}$, and therefore we know that this union of the connected subsets A_k and A_{k+1} is connected. Since C is a connected component of $\cup_n A_n$ containing A_k, it follows that C must also contain the connected subset $A_k \cup A_{k+1}$, and hence it follows that $A_{k+1} \subset C$, completing the proof of the inductive step.■

3. Suppose that C is a nonempty open and closed subset of $Y = A \cup B$. Then by connectedness either $A \subset C$ or $A \cap C = \emptyset$; without loss of generality we may assume the first holds, for the argument in the second case will follow by interchanging the roles of C and $Y - C$. We need to prove that $C = Y$.

Since each A_α for each α we either have $A_\alpha \subset C$ or $A_\alpha \cap C = \emptyset$. In each case the latter cannot hold because $A \cap A_\alpha$ is a nonempty subset of C, and therefore A_α is contained in C for all α. Therefore $C = Y$ and hence Y is connected.■

4. Suppose that we can write $X = C \cup D$ where C and D are disjoint nonempty open and closed subsets. Since C is open it follows that either D is finite or all of X; the latter cannot happen because $X - D = C$ is nonempty. On the other hand, since C is closed it follows that either C is finite or $C = X$. Once again, the latter cannot happen because $X - C = D$ is nonempty. Thus X is a union of two finite sets and must be finite, which contradicts our assumption that X is infinite. This forces X to be connected.■

5. Suppose that C is a maximal connected subset; then C also has the discrete topology, and the only discrete spaces that are connected are those with at most one point.

There are many examples of totally disconnected spaces that are not discrete. The set of rational numbers and all of its subspaces are fundamental examples; here is the proof: Let x and y be distinct rational numbers with $x < y$. Then there is an irrational number r between them, and the identity

$$\mathbb{Q} = \Big(\mathbb{Q} \cap (-\infty, r)\Big) \quad \cup \quad \Big(\mathbb{Q} \cap (r, +\infty, r)\Big)$$

gives a separation of $\mathbb{Q}$, thus showing that x and y lie in different connected components of $\mathbb{Q}$. But x and y were arbitrary so this means that no pair of distinct points can lie in the same connected component of Q. The argument for subsets of $\mathbb{Q}$ proceeds similarly.■

9. A good way to approach this problem is to begin by drawing a picture in which $X \times Y$ is a square and $A \times B$ is a smaller concentric square. It might be helpful to work with this picture while reading the argument given here.

Let $x_0 \in X - A$ and $y_0 \in Y - B$, and let C be the connected component of (x_0, y_0) in $X \times Y - A \times B$. We need to show that C is the entire space, and in order to do this it is enough to show that given any other point (x, y) in the space there is a connected subset of $X \times Y - A \times B$ containing it and (x_0, y_0). There are three cases depending upon whether or not $x \in A$ or $y \in B$ (there are three options rather than four because we know that both cannot be true).

If $x \notin A$ and $y \notin B$ then the sets $X \times \{y_0\}$ and $\{x\} \times Y$ are connected subsets such that (x_0, y_0) and (x, y_0) lie in the first subset while (x, y_0) and (x, y) lie in the second. Therefore there is a connected subset containing (x, y) and (x_0, y_0) by Exercise 3.— Now suppose that $x \in A$ but $y \notin B$. Then the two points in question are both contained in the connected subset $X \times \{y\} \cup times\{x_0\} \times Y$. Finally, if $x \notin A$ but $y \in B$, then the two points in question are both contained in the connected subset $X \times \{y_0\} \cup \{x\} \times Y$. Therefore the set $X \times Y - A \times B$ is connected.■

12. [*Assuming Y is a closed subset of X. Munkres does not explicitly assume $Y \neq \emptyset$, but without this assumption the conclusion is false.*] Since $X - Y$ is open in X and A and B are disjoint open subsets of $X - Y$, it follows that A and B are open in X. The latter in turn implies that $X \cup A$ and $Y \cup B$ are both closed in X.

We shall only give the argument for $X \cup A$; the proof for $X \cup B$ is the similar, the only change being that the roles of A and B are interchanged. Once again, it might be helpful to draw a picture.

Suppose that C is a nonempty proper subset of $Y \cup A$ that is both open and closed, and let $D = (Y \cup A) - C$. One of the subsets C, D must contain some point of Y, and without loss of generality we may assume it is C. Since Y is connected it follows that all of Y must be contained in C. Suppose that $D \neq \emptyset$. Since D is closed in $Y \cup A$ and the latter is closed in X, it follows that D is closed in X. On the other hand, since D is open in $Y \cup A$ and disjoint from Y it follows that D is open in A, which is the complement of Y in $Y \cup A$. But A is open in X and therefore D is open in X. By connectedness we must have $D = X$, contradicting our previous observation that $D \cap Y = \emptyset$. This forces the conclusion that D must be empty and hence $Y \cup A$ is connected.■

Problems from Munkres, § 24, pp. 157 – 159

1. (*b*) Take $X = (0, 1) \cup (2, 3)$ and $Y = (0, 3)$, let f be inclusion, and let g be multiplication by $1/3$. There are many other examples. In the spirit of part **(a)** of this exercise, one can also take $X = (0, 1)$, $Y = (0, 1]$, $f =$ inclusion and $g =$ multiplication by $1/2$, and similarly one can take $X = (0, 1]$, $Y = [0, 1]$, $f =$ inclusion and $g(t) = (t + 1)/2$.■

FOOTNOTE. Notwithstanding the sort of examples described in the exercises, a result of S. Banach provides a "resolution" of X and Y if there are continuous embeddings $f : X \to Y$ and $g : Y \to X$; namely, there are decompositions $X = X_1 \cup X_2$ and $Y = Y_1 \cup Y_2$ such that $X_1 \cap X_2 = Y_1 \cap Y_1 = \emptyset$ and there are homeomorphisms $X_1 \to X_2$ and $Y_1 \to Y_2$. The reference is S. Banach, *Un théorème sur les transformations biunivoques*, Fundamenta Mathematicæ **6** (1924), 236–239.

Additional exercises

1. Define a binary relation $\sim$ on X such that $u \sim v$ if and only if there are open subsets U and V in $\mathcal{U}$ such that $u \in U$, $v \in V$, and there is a sequence of open sets $\{U_0, U_1, \cdots U_n\}$ in $\mathcal{U}$ such that $U = U_0$, $V = U_n$, and $U_i \cap U_{i+1} \neq \emptyset$ for all i. This is an equivalence relation (verify this in detail!). Since every point lies in an open subset that belongs to $\mathcal{U}$ it follows that the equivalence classes are open. Therefore the union of all but one equivalence class is also open, and hence a single equivalence class is also closed in the space. If X is connected, this can only happen if there is exactly one equivalence class.■

2. This is similar to some arguments in the course notes and to the preceding exercise. The hypothesis that the relation is locally constant implies that equivalence classes are open. Therefore the union of all but one equivalence class is also open, and hence a single equivalence class is also closed in the space. If X is connected, this can only happen if there is exactly one equivalence class.$\blacksquare$

3. The important point is that $\mathbb{Q}^n$ is dense in $\mathbb{R}^n$; given a point $x \in \mathbb{R}^n$ with coordinates x_i and $\varepsilon > 0$ we can choose rational numbers q_i such that $|q_i - x_i| < \varepsilon/n$; if y has coordinates q_i then $|x - y| < \varepsilon$ follows immediately. Since $\mathbb{R}^n$ is locally connected, the components of open sets are open, and therefore we conclude that every component of a nonempty open set $U \subset \mathbb{R}^n$ contains some point of $\mathbb{Q}^n$. Picking one such point for each component we obtain a 1–1 map from the set of components into the countable set $\mathbb{Q}^n$.

The Cantor set is an example of a closed subset of $\mathbb{R}$ for which the components are the one point subsets and there are uncountably many points.$\blacksquare$

4. (i) FALSE. One simple counterexample is given by taking $X = Y = \mathbb{R}$ and $A = B = \{0\}$. Then $X \times Y - A \times B$ is merely $\mathbb{R}^2 - \{\mathbf{0}\}$, which is the image of the connected set $(0, \infty) \times [0, 2\pi]$ under the polar coordinate map sending (r, θ) to $(r \cos\theta, r \sin\theta)$.$\blacksquare$

(ii) FALSE. Take $A = (0, 1]$ and $B = \{0, 1\}$ so that $A \cap B = \{1\}$ and $A \cup B = [0, 1]$.

(iii) TRUE. Given that the hypotheses are nearly the same as in (ii) but slightly stronger, one might guess this answer, but of course a proof is still needed.

We shall only prove that A is connected; the connectedness of B will then follow by interchanging the roles of A and B in the argument given here. — Suppose that A is not connected, and write it as a union of two nonempty closed subsets $A_1 \cup A_2$. Since $A \cap B$ is connected, this intersection is contained in either A_1 or in A_2; renumbering these sets if necessary, we may as well assume that $A \cap B \subset A_1$, which means that $\emptyset = A \cap B \cap A_2 = B \cap A_2$. Since $B \neq \emptyset$ and $B \cap A_2$ is empty, the set $B \cap A_1$ must be nonempty. Consider the closed subspaces $B \cup A_1$ and A_2. They are both nonempty, they are disjoint, and their union is $A \cup B$. Therefore $A \cup B$ is disconnected. On the other hand, we assumed that $A \cup B$ was connected, so we have reached a contradiction. The source of the problem was our added assumption that A was disconnected, so this must be false and we are forced to conclude that A is connected.$\blacksquare$

5. Assume that the conclusion is false and that $B \cap \mathrm{Bdy}\,(A) = \emptyset$. Therefore no point of B is a boundary point of A; in other words, for every point $b \in B$ there is some open neighborhood U such that either U does not contain any points of X or else U does not contain any points of $X - A$. This can be rephrased to state that either $U \cap B$ is contained in $B \cap A$ or in $B \cap X - A$ depending upon which of these sets contains b. It follows that $B \cap A$ and $B \cap X - A$ are open subsets in B which are disjoint and whose union is B. By our hypotheses we also know that each intersection is nonempty, and therefore we conclude that B is disconnected. The latter contradicts another of our hypotheses; the source of the contradiction was the added condition that $B \cap \mathrm{Bdy}\,(A) = \emptyset$, and therefore this statement cannot be true. We can rephrase this to say that $B \cap \mathrm{Bdy}\,(A) \neq \emptyset$ must be true.$\blacksquare$

5. (i) If X is discrete and $x \in X$, then $\{x\}$ by itself forms a neighborhood base at x; since $\{x\}$ is open and closed in X, it follows that X is totally disconnected in the sense of the definition.

As noted in the hint, a discrete space has no limit points because the deleted neighborhood $\{x\} - \{x\}$ is empty. However, the set A of all points x in $\mathbb{R}$ such that $x = 0$ or $x = 1/n$ for some positive integer n does have a limit point at 0, and we claim it is totally disconnected. Since each one point subset $\{1/n\}$ is open and closed, there is a clopen (closed + open) neighborhood base at such points. There is also a clopen neighborhood base at 0, and it is given by the sets

$(-1/n, 1/n) \cap A$ because this set is equal to $[-2/(2n+1), 2/(2n+1)] \cap A$ (the point $1/(n+1)$ is the maximal point in this set). Therefore A is not discrete but A is totally disconnected.

(*ii*) The first set is open and the second set is closed, and the sets in the first family form a neighborhood base, so it is enough to show that the intersections of the open and closed intervals with $\mathbb{Q}$ are equal. Since the closed intervals are obtained from the open intervals by adjoining the two endpoints, it is enough to note that these endpoints are not rational numbers — a consequence of the fact that $\sqrt{2}$ is irrational. This proves that $\mathbb{Q}$ is totally disconnected.

To see that every point of $\mathbb{Q}$ is a limit point of $\mathbb{Q}$, note that if $q \in \mathbb{Q}$ then the sequence of points $q + \frac{1}{n}$ is rational and no point in the sequence is equal to q, but nevertheless the limit of the sequence is q.∎

(*iii*) Suppose that X and Y are totally disconnected and $(x, y) \in X \times Y$. Let $\{U_\alpha\}$ and $\{Y_\beta\}$ be clopen neighborhood bases at x and y respectively. Since the product of two clopen subsets is clopen, it follows that $\{U_\alpha \times Y_\beta\}$ forms a clopen neighborhood base at (x, y). Since the latter point was arbitrary, it follows that $X \times Y$ is totally disconnected.∎

7. Clearly there are two cases, one when $n = 1$ and the other when $n \geq 2$.

The case $n = 1$. We know that the connected subsets of $\mathbb{R}$ are precisely those sets A such that if $x < y$ and $x, y \in A$, then the closed interval $[x, y]$ is contained in A. Intuitively, we expect that the only sets with this property are open intervals (possibly with $\pm\infty$ as endpoints), half-open intervals (possibly with $\pm\infty$ as the open endpoint), and closed intervals. If this is true, then there are $2^{\aleph_0}$ choices for each endpoint, and for every pair of endpoints there are at most four intervals depending upon which of the ≤ 2 endpoints belong to the interval, so an upper bound for the cardinality of the family of connected subsets is $4 \cdot 2^{\aleph_0} \cdot 2^{\aleph_0} = 2^{\aleph_0}$, This number is also a lower bound because we have the family of intervals $[r, \infty)$ where r runs through the elements of $\mathbb{R}$, and therefore the cardinality of the family of connected subsets is exactly $2^{\aleph_0}$.

We now have to verify the assertion that connected subsets of $\mathbb{R}$ are intervals. Let C be a nonempty connected subset of $\mathbb{R}$, so that $u, v \in C$ and $u < v$ imply $[u, v] \subset C$. Let $a(C)$ be the greatest lower bound of C if C has a lower bound, and let $a(C) = -\infty$ otherwise. Similarly, let $b(C)$ be the least upper bound of C if C has an upper bound, and let $b(C) = +\infty$ otherwise. Let p be some point of C.

CLAIM: The set C is an interval whose lower endpoint is $a(C)$ and whose upper endpoint is $b(C)$. — Since $a(C)$ and $b(C)$ are lower and upper bounds for C, it follows that C is contained in the set of all points x satisfying $a(C) \leq x \leq b(C)$, with the convention that there is strict inequality if $a(C) = -\infty$ or $b(C) = +\infty$. If $a(C) < b(C)$ and $p \in C$ lies between these values, take sequences $\{u_n\}$ and $\{v_n\}$ in C such that $u_n \to a(C)$, $v_n \to b(C)$, and $u_n < p < v_n$ for all n. By the assertion in the first sentence of the first paragraph, then $[u_n; v_n] \subset C$ for all n, and therefore the union of these subsets, which is the set of all x such that $a(C) \leq x \leq b(C)$, is contained in C. Combining this with the previous paragraph, we have

$$\{x \in \mathbb{R} \mid a(C) \ < \ x \ < \ b(C)\} \quad \subset \quad C \quad \subset \quad \{x \in \mathbb{R} \mid a(C) \leq x \leq b(C)\}$$

with the previous conventions if $a(C) = -\infty$ or $b(C) = +\infty$. For each pair of values $a(C)$ and $b(C)$ there are up to four choices for C, depending upon whether or not $a(C)$ or $b(C)$ belong to C.∎

The case $n \geq 2$. In order to avoid awkward typographical problems, let $\mathbf{c} = 2^{\aleph_0} = |\mathbb{R}|$; recall that we then have $|\mathbb{R}^n| = \mathbf{c}$ for all positive integers n. Since the set of all subsets in $\mathbb{R}^n$ then has cardinality $2^{\mathbf{c}}$, it follows that $|C_n| \leq 2^{\mathbf{c}}$ for all n. Furthermore, since $\mathbb{R}^2$ is homeomorphic to a subspace of $\mathbb{R}^n$ if $n \geq 3$ it follows that $|C_2| \leq |C_n|$ for all $n \geq 3$. Thus if we can show that $|C_2| \geq 2^{\mathbf{c}}$, then it will follow that $|C_n| = 2^{\mathbf{c}} > \mathbf{c}$ for all $n \geq 2$.

We can construct a family of $2^{\mathbf{c}}$ connected subsets of $\mathbb{R}^2$ as follows: One of our results implies that if A is a connected subset of a topological space and $A \subset B \subset \overline{A}$, then B is connected. In particular, if we apply this to $U = (0,1)^2 \subset \mathbb{R}^2$, then each subset C of $\mathrm{Bdy}\,(U)$ determines a connected subset $U \cup C$ of $\mathbb{R}^2$, and $C \neq C'$ implies that $U \cup C \neq U \cup C'$ because $U \cap \mathrm{Bdy}\,(U) = \emptyset$. In particular, if we let C run through all the $2^{\mathbf{c}}$ subsets of $(0,1) \times \{0\}$, then this yields a family $\{U \cup C\}$ of $2^{\mathbf{c}}$ connected subsets of $\mathbb{R}^2$, and therefore we have proved the desired inequality $|C_2| \geq 2^{\mathbf{c}}$. As noted before, this completes the proof that $|C_n| = 2^{\mathbf{c}} > 2^{\aleph_0}$ if $n \geq 2$.■

8. (i) Each line px in $\mathrm{Lines}\,(p, \mathbb{R}^n - D)$ is an arcwise connected set containing p, so every point in the set lies in the same arc component as p, and therefore $\mathrm{Lines}\,(p, \mathbb{R}^n - D)$ is arcwise connected.

The complement of $\mathrm{Lines}\,(p, \mathbb{R}^n - D)$ consists of all points y on punctured lines $px - \{p\}$ such that px contains a point of D; this point in D cannot be p because $p \notin D$. Since there is a unique line joining two points (see the discussion below), every such line px is equal to pz for some $z \in D$. Since there are only countably many points in D, there are only countably many lines px such that $px - \{p\}$ is contained in the complement of $\mathrm{Lines}\,(p, \mathbb{R}^n - D)$.

(ii) The existence of a point $x \notin pq$ is discussed in the postscript to the discussion following the solution to this exercise, so the hint involves something which actually exists. If $y \in px$, then $q \notin px$ implies that the lines px and qy are distinct and hence can only have the point y in common (because there is a unique line joining two points). By hypotheses $q \notin D$, and therefore there are only countably many lines of the form qz where $z \in D$, and they are all distinct from px because $q \notin px$ (again by the uniqueness of a line joining two points). Therefore there are only countably many points $y \in px$ such that qy contains an element of D. Since there are uncountably many points on the line px, it follows that there is some point $w \in px - \{p\}$ such that $qw \cap D = \emptyset$. We then have $qw \subset \mathbb{R}^n - D$, and this implies that $pw \cup qw \subset \mathbb{R}^n - D$, which in turn means that p and q lie in the same arc component of $\mathbb{R}^n - D$. Since p and q are arbitrary points of $\mathbb{R}^n - D$, this shows that the latter is arcwise connected.■

Remark on lines in $\mathbb{R}^n$. The solution to this exercise used the following standard geometrical fact, both as intuition and as a logical step in the argument:

Two distinct lines in $\mathbb{R}^n$ have at most one point in common.

We also use a related statement: *Given two points, there is a unique line containing them.*

Since a rigorous proof of the first statement might not have been given in prerequisite undergraduate courses, for the sake of completeness we shall give a self-contained proof here using vector algebra. One can also use the concepts developed in Sections I.3 and II.2 of `pg-all.pdf` to view this result as part of a more general pattern. We should probably start by formally defining the $\mathbf{xy}$ joining two points $\mathbf{x}$, $\mathbf{y} in \mathbb{R}^n$ (or more generally vectors in any vector space V) to be the set of all points $\mathbf{z}$ such that $\mathbf{z} = t\,\mathbf{y} + (1-t)\,\mathbf{x} = \mathbf{x} + t\,(\mathbf{y} - \mathbf{x})$ for some scalar t. It is an elementary exercise in linear algebra to check that a point $\mathbf{w}$ belongs to this set if and only if $\mathbf{w} = \mathbf{y} + s\,(\mathbf{y} - \mathbf{x})$ for some scalar s (in fact, if $\mathbf{w} = \mathbf{z}$ where $\mathbf{z}$ is given as above then we can take $s = 1 - t$).

Assume now that we are given $\mathbf{a} \neq \mathbf{b}$ and $\mathbf{c} \neq \mathbf{d}$ in $\mathbb{R}^n$ such that the lines $\mathbf{ab}$ and $\mathbf{cd}$ have two distinct points in common. We need to prove that $\mathbf{ab}$ must be equal to $\mathbf{cd}$.

If the lines have two points in common, then there are scalars s, t, u, v such that $t\,\mathbf{b} + (1-t)\,\mathbf{a} = s\,\mathbf{d} + (1-s)\,\mathbf{c}$ and $v\,\mathbf{b} + (1-v)\,\mathbf{a} = u\,\mathbf{d} + (1-u)\,\mathbf{c}$. We may rewrite these in the forms

$$\mathbf{a} \;+\; t\,(\mathbf{b} - \mathbf{a}) \;\;=\;\; \mathbf{c} \;+\; s\,(\mathbf{d} - \mathbf{c})\;, \quad \mathbf{a} \;+\; v\,(\mathbf{b} - \mathbf{a}) \;\;=\;\; \mathbf{c} \;+ u\,(\mathbf{d} - \mathbf{c})$$

and if we take the difference of these equations we find that $(t-v)\,(\mathbf{b}-\mathbf{a}) = (s-u)\,(\mathbf{d}-\mathbf{c})$. Therefore there is a nonzero vector $\mathbf{w}$ such that $\mathbf{b}-\mathbf{a} = K\,\mathbf{w}$ and $\mathbf{d}-\mathbf{c} = L\,\mathbf{w}$; since the expressions on the left are nonzero, it follows that K and L are both nonzero. These yield the equation

$$\mathbf{c}-\mathbf{a} \quad = \quad K\,(t-v)\,\mathbf{w} \quad = \quad L\,(s-u)\,\mathbf{w}$$

so that $K\,(t-v) = L\,(s-u)$. Therefore for some scalars M and M' we have

$$\mathbf{c} \quad = \quad \mathbf{a} \,+\, M\,\mathbf{w} \quad = \quad \mathbf{a} \,+\, M'\,(\mathbf{b}-\mathbf{a})$$

so that $\mathbf{c} \in \mathbf{ab}$. If we now reverse the roles of $\mathbf{c}$ and $\mathbf{d}$ in the preceding argument, we also find that $\mathbf{d} \in \mathbf{ab}$.

We shall now prove that $\mathbf{cd} \subset \mathbf{ab}$; if we know this, then we can reverse the roles of the two lines in the argument and also obtain the conclusion $\mathbf{ab} \subset \mathbf{cd}$, and if we combine these we have $\mathbf{ab} = \mathbf{cd}$, which is what we wanted to prove. — To see this, let $\mathbf{x} in \mathbf{cd}$. Then $\mathbf{x} = r\,\mathbf{c} + (1-r)\,\mathbf{d}$ for some r. Now $\mathbf{c}, \mathbf{d} \in \mathbf{ab}$ implies that $\mathbf{c} = (1-M')\,\mathbf{a} + M'\,\mathbf{b}$ (as before) and $\mathbf{d} = (1-N')\,\mathbf{a} + N'\,\mathbf{b}$ for some scalars M' and N'. Direct substitution then yields the equation

$$\mathbf{x} \quad = \quad (1 - rM' + (r-1)N')\,\mathbf{a} + (rM' + (1-r)N')\,\mathbf{b}$$

which shows that $\mathbf{ab} \subset \mathbf{cd}$. As noted earlier in this paragraph, similar reasoning shows the reverse inclusion, and therefore the two lines must be equal.■

Postscript. In our solution we also use the fact that if $n \geq 2$ and $\mathbf{p} \neq \mathbf{q} \in \mathbb{R}^n$, then there is some point $\mathbf{x}$ which does not lie on the line $\mathbf{pq}$, so to be complete we should also verify this assertion. By definition, if $\mathbf{x}$ lies on $\mathbf{pq}$ then $\mathbf{x}-\mathbf{p}$ is a scalar multiple of $\mathbf{q}-\mathbf{p}$. Since $n \geq 2$ implies the existence of some vector $\mathbf{v}$ which is not a multiple of $\mathbf{q}-\mathbf{p}$, it follows that $\mathbf{x} = \mathbf{p}+\mathbf{v}$ does not lie on $\mathbf{pq}$.■

III.5 : Variants of connectedness

Problem from Munkres, § 24, pp. 157 – 159

8. (b) The closure $\overline{A}$ of an arc component A is NOT NECESSARILY arcwise connected; this contrasts with the fact that connected components are always closed subsets. Consider the graph of $\sin(1/x)$ for $x > 0$. It closure is obtained by adding the points in $\{0\} \times [-1,1]$ and we have shown that the space consisting of the graph and this closed segment is not path connected.■

Problem from Munkres, § 25, pp. 162 – 163

10. [Only (a), (b) and examples A and B from (c).]

(a) The proofs that $\sim$ is reflexive and symmetric are very elementary and left to the reader. Regarding transitivity, suppose that $x \sim y$ and $y \sim z$ and that we are given a separation of X as $A \cup B$. Without loss of generality we may assume that $x \in A$ (otherwise interchange the roles of A and B in the argument). Since $x \sim y$ it follows that $y \in A$, and the latter combines with $y \sim z$ to show that $z \in A$. Therefore if we are given a separation of X as above, then x and z lie in the same piece, and since this happens for all separations it follows that $x \sim z$.■

(b) Let C be a connected component of X. Then $A \cap C$ is an open and closed subset of C and therefore it is either empty or all of C. In the first case $C \subset B$ and in the second case $C \cap B$. In either case it follows that all points of C lie in the same equivalence class. – If X is locally connected then each connected component is both closed and open. Therefore, if x and y lie in different components, say C_x and C_y, then $X = C_x \cup X - C_x$ defines a separation such that x lies in the first subset and y lies in the second, so that $x \sim y$ is false if x and y do not lie in the same connected component. Combining this with the previous part of the exercise, if X is locally connected then $x \sim y$ if and only if x and y lie in the same connected component.■

(c) [*Note:* Example C in this part of the problem was not assigned; some comments appear below.]

> *Trying to solve this exercise without any pictures would probably be difficult at best and hopelessly impossible at worst.*

Drawings for Case A and Case C are in the file `math205Axolutions03a.pdf`.

Case A**.** We claim that the connected components are the closed segments $\{1/n\} \times [0,1]$ and the one point subsets $\{(0,0)\}$ and $\{(0,1)\}$. — Each segment is connected and compact (hence closed), and we claim each is also an open subset of A. This follows because the segment $\{1/n\} \times [0,1]$ is the intersection of A with the open subset

$$\left(\frac{1}{2}\left[\frac{1}{n} + \frac{1}{n+1}\right] \, , \, \frac{1}{2}\left[\frac{1}{n} + \frac{1}{n-1}\right] \right) \times \mathbb{R} \ .$$

Since the segment $\{1/n\} \times [0,1]$ is open, closed and connected, it follows that it must be both a component and a quasicomponent. It is also an arc component because it is arcwise connected. This leaves us with the two points $(0,0)$ and $(0,1)$. They cannot belong to the same component because they do not form a connected set. Therefore each belongs to a separate component and also to a separate path component. The only remaining question iw whether or not they determine the same quasicomponent. To show that they do lie in the same quasicomponent, it suffices to check that if $A = U \cup V$ is a separation of X into disjoint open subsets then both points lie in the same open set. Without loss of generality we may as well assume that the origin lies in U. It then follows that for all n sufficiently large the points $(1/n, 0)$ all lie in U, and the latter implies that all of the connected segments $\{1/n\} \times [0,1]$ is also lie in U for n sufficiently large. Since $(0,1)$ is a limit point of the union of all these segments (how?), it follows that $(0,1)$ also lies in U. This implies that $(0,1)$ lies in the same quasicomponent of A as $(0,0)$.■

Case B**.** This set turns out to be connected but not arcwise connected. We claim that the path components are given by $\{(0,1)\}$ and its complement. Here is the proof that the complement is path connected: Let P be the path component containing all points of $[0,1] \times \{0\}$. Since the latter has a nontrivial intersection with each vertical closed segment $\{1/n\} \times [0,1]$ it follows that all of these segments are also contained in P, and hence P consists of all points of A except perhaps $(0,1)$. Since $(0,1)$ is a limit point of P (as before) it follows that B is connected and thus there is only one component and one quasicomponent. We claim that there are two path components. Suppose the extra point $(0,1)$ also lies in P. Then, for example, there will be a continuous curve $\alpha : [0,1] \to A$ joining $(0,1)$ to $(1/2, 1)$. Let t_0 be the maximum point in the subset of $[0,1]$ where the first coordinate is zero. Since the first coordinate of $\alpha(a)$ is 1, we must have $t_0 < 1$. Since $A \cap \{0\} \times \mathbb{R}$ is equal to $\{(0,0)\} \cup \{(0,1)\}$, it follows that α is constant on $[0, t_0]$, and by continuity there is a $\delta > 0$ such that $|t - t_0| < \delta$ implies that the *second* coordinate of $\alpha(t)$ is greater than $1/2$. If $t \in (t_0, t_0 + (\delta/2))$, then the first coordinate of $\alpha(t)$ is positive and the second is greater than $1/2$. In fact we can choose some t_1 in the open interval such that the first coordinate is irrational. But there are no points in $B \cap (\frac{1}{2}, +\infty)$ whose first coordinates are irrational, so we have a contradiction. The latter arises because we assumed the existence of a continuous curve joining $(0,1)$ to another point in B. Therefore no such curve can exist and $(0,1)$ does not belong to the path component P. Hence B has two path components, and one of them contains only one point.$\blacksquare$

Case C. This was not assigned, but we shall include the solution as an example of an argument which is more challenging and should be understood passively. Before proceeding to the solution, we shall give away the answer: The space C is connected and each of the (infinitely many) closed segments given in the definition is a separate path component.

Once again, it would probably be extremely difficult to solve this part of the exercise without the sort of drawing we have inserted into `math205Asolutions03a.pdf`. This drawing shows that C is symmetric with respect to rotation through $90°$;*i.e.*, if $J(x,y) = (-y,x)$ is counterclockwise rotation through a right angle, then $J[C] = C$. As usual, one needs to give a written argument to verify this, but this is routine and the details are left to the reader (note that J maps the first summand of C to the fourth, the fourth summand to the second, the second summand to the third, and the third summand to the first).

The set C is a union of countably closed segments $E_{j,n}$ which are defined by $E_{0,n} = \{1/n\} \times [0,1]$ and $E_{k,n} = J^k[E_{0,n}] = J^{k-1}[E_{k-1,n}]$, where J is defined as in the preceding paragraph. Set $E_k = \cup_n \; E_{k,n}$. — Also, let $K = \{1, \frac{1}{2}, \frac{1}{3}, \; \cdots \}$ be defined as in Munkres' formulation of the exercise.

Suppose that D is a clopen subset of C which contains $(1,0)$. Since D is open, it also contains points of the form $(1/n, 1)$ for all but finitely many values of n, and hence it contains points from all but finitely many of the intervals $E_{3,n}$. By the connectedness of these intervals and the clopenness of D, it follows that $E_{3,n} \subset D$ for $n \geq N_0$ for some N_0. Each of the points $(1/m, 0)$ is a limit point of $\cup_{n \geq N_0} \; E_{3,n}$ (consider the sequence $(1/m, 1/n)$ for $n \geq N_0$), and since D is closed it also follows that $(1/m, 0) \in D$. Since the latter point lies in $E_{0,m}$ and the latter is connected, it follows that the clopen set D also contains E_0. Furthermore, since $(0,1)$ is a limit point of E_0 (look at the sequence $(1/m, 1)$ in E_0), it follows that $J(1,0) = (0,1) \in D$.

By the same reasoning, if $k \geq 1$ we know that if $J^k(0,1) \in D$ then $E_j \subset D$ and $J^{k+1}(0,1) \in D$. Therefore we can put together an inductive argument to conclude that $E_k \subset D$ for all k (since J^4 is the identity, there are only finitely many steps where a new conclusion is obtained). Finally, since $C = \cup_j \; E_j$ it follows that $C = D$, which means that C is connected. Therefore there is only one connected component, and there is also only one quasicomponent because each quasicomponent is a union of components.

We must now describe the arc components of C; as noted earlier, we want to show that these are the subspaces $E_{k,n}$. Most of the work involves proving the following:

> CLAIM . *If $x \in E_{k,n}$ and U is an open neighborhood of x in C, then there is a subneighborhood V of U such that the arc component of x in V is equal to $E_{k,n} \cap V$.*

Because of the symmetry of C under right angle rotations, it will suffice to prove this claim when $k = 0$.

To simplify the notation, let $Sq_\varepsilon(x) \subset \mathbb{R}^2$ denote the open square of all points $y \in \mathbb{R}^2$ such that $\mathbf{d}^{\langle\infty\rangle}(x,y) < \varepsilon$. We shall show that if $x \in C$ as above then we can take V to have the form $Sq_\varepsilon(x) \cap C$ for some $\varepsilon > 0$ (which suffices since the latter subsets form a neighborhood base at x).

As above, we need only consider the case where $x \in E_{0,n_0}$ for some positive integer n_0, so we have $x = (1/n_0, t_0)$ for some $t_0 \in [0,1]$. If $t_0 > 0$ then $V = Sq_\varepsilon(x) \cap C$ has the desired properties if we take ε to be smaller than both t_0 and $1/(2n_0)$; in fact, for this case the intersection lies on the vertical line through x and it is arcwise connected. We now turn to the more complicated case where $t_0 = 0$. In this case, if $\varepsilon < 1/(2n_0)$ then $V = Sq_\varepsilon(x) \cap C$ is a disjoint union of the vertical segment $\{1/n_0\} \times [0, \varepsilon)$ and the horizontal segments $(1/n_0 - \varepsilon, 1/n_0 + \varepsilon) \times \{1/m\}$ for all integers m such that $1/m < \varepsilon$. From this description it is clear that the arc component of x in V is the intersection of the latter with the vertical segment E_{0,n_0}. This completes the proof of the claim (for the sets $E_{0,n}$; as noted before, symmetry considerations yield the same conclusion for $E_{k,n}$ when $k = 1, 2, 3$).

Using the claim we have just verified, we can prove that each subset $E_{k,n}$ is an arc component as follows: Let $\gamma : [0,1] \to C$ be a continuous curve. If $\gamma(t) \in E_{k,n}$ then the claim and the continuity of γ imply that γ maps some interval $(t-h, t+h) \cap [0,1]$ into $E_{0,n}$, and this implies that $\gamma^{-1}[E_{k,n}]$ is an open subset of $[0,1]$ for each k and n. Since the sets $E_{k,n}$ are pairwise disjoint and their union is C, it follows that the open subsets $\gamma^{-1}[E_{k,n}] \subset [0,1]$ are also pairwise disjoint and their union is $[0,1]$. Since the unit interval is connected, it follows that exactly one of these subsets is nonempty, and therefore it must be the entire unit interval. In other words, we have shown that the image of a continuous mapping from $[0,1]$ to C must be contained in some $E_{k,n}$, which means that every arc component of C must be contained in some $E_{k,n}$. Finally, since the latter are arcwise connected, it follows that each $E_{k,n}$ is an arc component of C.■

Additional exercises

1. In a locally connected space the connected components are open (and pairwise disjoint). These sets form an open covering and by compactness there is a finite subcovering. Since no proper subcollection of the set of components is an open covering, this implies that the set of components must be finite.■

2. Let $Y = \mathbb{R}^2$ and let $X \subset \mathbb{R}^2$ be the union of the horizontal half-line $(0, \infty) \times \{0\}$ and the vertical closed segment $\{-1\} \times [-1, 1]$. These subsets of X are closed in X and pairwise disjoint. Let $f : X \to \mathbb{R}^2$ be the continuous map defined on $(0, \infty) \times \{0\}$ by the formula $f(t, 0) = (t, \sin(1/t))$ and on $\{-1\} \times [-1, 1]$ by the formula $f(-1, s) = (0, s)$. The image $f[X]$ is then the example of a non-locally connected space that is described in the course notes.■

3. Consider the polar coordinate map $[1, \sqrt{2}] \times [0, 2\pi]$ which sends (r, θ) to $(r\cos\theta, r\sin\theta)$. This is a continuous onto mapping, and the domain is arcwise connected (in fact, it is a convex subset of $\mathbb{R}^2$. Therfore the image, which is the annulus in the exercise, is also arcwise connected.■

4. The result for $\mathbb{R}^n - \{\mathbf{0}\}$ is implicit in one of the earlier starred exercises, but we shall give a proof here because it can be done simply and the result for S^{n-1} depends upon this fact.

If $x, y \in \mathbb{R}^n - \{\mathbf{0}\}$ are linearly independent, then the line segment $ty + (1-t)x$ $(0 \le t \le 1)$ does not pass through $\mathbf{0}$ (if it did, then x and y would be linearly dependent), and hence x and y lie in the same arc component of $\mathbb{R}^n - \{\mathbf{0}\}$. Suppose now that x and y are linearly dependent, so that each is a nonzero multiple of the other. Since $n \ge 2$ there is some vector z such that x and z are linearly independent, and it follows immediately that y and z are also linearly independent. Two applications of the previous argument then show that x, y, z all lie in the same arc component. Thus in all cases we have shown that two arbitrary points $x, y \in \mathbb{R}^n - \{\mathbf{0}\}$ always lie in the same arc component, which means that the space under consideration is arcwise connected.

To prove the result for S^{n-1}, consider the map

$$\sigma : \mathbb{R}^n - \{\mathbf{0}\} \longrightarrow S^{n-1}$$

which sends a nonzero vector v to the unit vector $|v|^{-1} \cdot v$ pointing in the same direction. This map is continuous since $|v| \neq 0$ on the domain, and it is clearly onto because $\sigma(v) = v$ if $v \in S^{n-1}$ (so that $|v| = 1$). Since $\mathbb{R}^n - \{\mathbf{0}\}$ is arcwise connected, its image under σ — which is S^{n-1} — must also be arcwise connected.■

5. (*i*) Let Q denote the quasicomponent of p in X. By definition $q \in Q$ if and only if for each separation of X into clopen subsets $A \cup B$ (with $A \cap B = \emptyset$) both p and q lie in the same subset.

By definition, if $q \in Q$ and $C \subset X$ is a clopen subset containing p, then $q \in C$. Therefore Q is contained in the intersection of all clopen subsets C such that $p \in C$. Conversely, if q lies in this intersection, then if C is a clopen set containing p we have $q \in C$, which means that q lies in the same quasicomponent as p.■

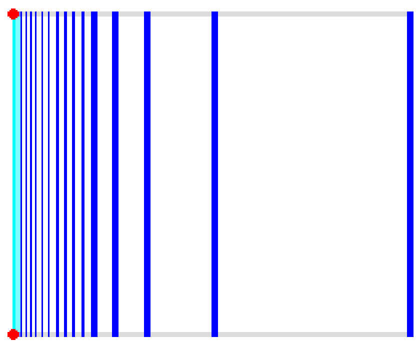
(*ii*) We are assuming there are only finitely many components, so list them as $C_1 \ \cdots \ , C_r$. Each of these subsets is closed since it is a component, and we claim that each is also open. This is true because we have

$$X \ - \ C_j \quad = \quad \bigcup_{i \neq j} C_i$$

for all j, so that the right hand side is a closed subset and hence its complement — which is C_j — must be open. This means that every component of X is clopen, and by (*i*) it follows that the quasicomponent of a point is contained in the component of a point. On the other hand, by Exercise 25.10 in Munkres we know that the reverse inclusion is true, and therefore the quasicomponent of a point is equal to the connected component of that point.■

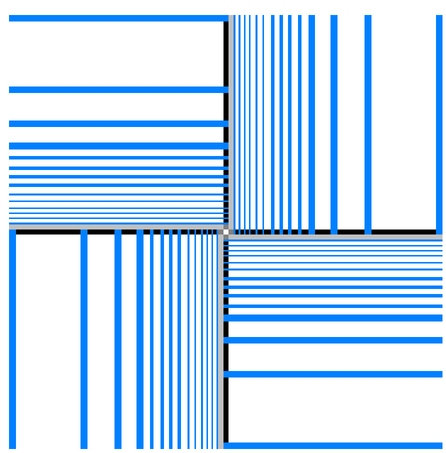
Drawing to accompany Munkres, Exercise $25.10(c)$

The following drawing of the set A should clarify many of the issues arising in the solution to this exercise. By definition, this set is given by a countable union of vertical line segments, which are depicted in blue, and two points which are depicted in red. The region shaded in light blue contains infinitely many additional closed segments which are too closely spaced to be shown due to the resolution of the drawing, but A does ***not*** contain the closed segment on the $y-$ axis which joins the two red points.



The crucial observation is that if U is a clopen subset which contains one red point, it must also contain the other. This is true because every neighborhood of a red point contains the endpoints of all but finitely many vertical segments. Since these segments are connected, the set U must contain each of these vertical segments, and since the other red point is in the closure of the union of these segments, the other red point must also be in U.

Also, here is a drawing of the set C in this exercise; the vertical and horizontal segments are in light blue, and the regions in light gray contain infinitely many segments which are too closely spaced to show. One advantage of the picture is how it shows that this set is symmetric with respect to rotation through one or more right angles. This can also be verified analytically (some additional information is given in the file **math205Asolutions03.pdf**).



SOLUTIONS TO EXERCISES FOR

MATHEMATICS 205A — Part 4

Fall 2014

IV. Function spaces

IV.1 : General properties

Additional exercises

1. The mapping q is 1–1 because $q(f) = q(g)$ implies that for all x we have $f(x) = p_x \circ q(f) = p_x \circ q(g) = g(x)$, which means that $f = g$.

To prove continuity, we need to show that the inverse images of subbasic open sets in Y^X are open in $\mathbf{C}(X, Y)$. The standard subbasic open subsets have the form $\mathcal{W}(\{x\}, U) = p_x^{-1}(U)$ where $x \in X$ and U is open in Y. In fact, there is a smaller subbasis consisting of all such sets $\mathcal{W}(\{x\}, U)$ such that $U = N_\varepsilon(y)$ for some $y \in Y$ and $\varepsilon > 0$. Suppose that f is a continuous function such that $q(f)$ lies in $\mathcal{W}(\{x\}, U)$. By definition the later condition means that $f(x) \in U$. The latter in turn implies that $\delta = \varepsilon - \mathbf{d}(f(x), y) > 0$, and if $\mathbf{d}(f, g) < \delta$ then the Triangle Inequality implies that $\mathbf{d}(g(x), y) < \varepsilon$, which in turn means that $g(x) \in U$. Therefore q is continuous at f, and since f is arbitrary this shows q is a continuous mapping.■

2. It suffices to show that the map in question is onto and distance-preserving. The map is onto because if u and v are continuous functions into Y and Z respectively, then we can retrieve f by the formula $f(x) = \big(u(x), v(x)\big)$. Suppose now that f and g are continuous functions from X to $Y \times Z$. Then the distance from f to g is the maximum of $\mathbf{d}\big(f(x), g(x)\big)$. The latter is less than or equal to the greater of $\mathbf{d}\big(p_Y f(s), p_Y g(s)\big)$ and $\mathbf{d}\big(p_Y f(t), p_Y g(t)\big)$. Thus if

$$\Phi : \mathbf{C}(X, Y \times Z) \to \mathbf{C}(X, Y) \times \mathbf{C}(X, Z)$$

then the distance from $\Phi(f)$ to $\Phi(g)$ is greater than or equal to the distance from f to g. This means that the map Φ^{-1} is uniformly continuous. Conversely, we claim that the distance from f to g is greater than or equal to the distance between $\Phi(f)$ and $\Phi(g)$. The latter is equal to the larger of the maximum values of $\mathbf{d}\big(p_Y \circ f(s), p_Y \circ g(s)\big)$ and $\mathbf{d}\big(p_Z \circ f(t), p_Z \circ g(t)\big)$. If $w \in X$ is where $\mathbf{d}(f, g)$ takes its maximum, it follows that

$$\mathbf{d}\big(f(w), g(t)\big) \quad =$$

$$\max\Big\{\ \mathbf{d}\big(p_Y \circ f(w), p_Y \circ g(w)\big),\ \mathbf{d}\big(p_Z \circ f(w), p_Z \circ g(w)\big)\ \Big\}$$

which is less than or equal to the distance between $\Phi(f)$ and $\Phi(g)$ as described above.■

3. The distance between f and g is the maximum value of the distance between $f(x)$ and $g(x)$ as x runs through the elements of x, which is the greater of the maximum distances between $f(x)$ and $g(x)$ as x runs through the elements of C, where C runs through the set $\{A, B\}$. But the second expression is equal to the larger of the distances between $\mathbf{d}(f|A, g|A)$ and $\mathbf{d}(f|B, g|B)$.

Therefore the map described in the problem is distance preserving. As in the previous exercise, to complete the proof it will suffice to verify that the map is onto. The surjectivity is equivalent to saying that a function is continuous if its restrictions to the closed subsets A and B are continuous. But we know the latter is true.■

4. Let $\varepsilon > 0$ be given. We claim that the distance between $f \times g$ and $f' \times g'$ is less than ε if the distance between f and f' is less than ε and the distance between g and g' is less than ε. Choose $u_0 \in X$ and $v_0 \in Z$ so that

$$\mathbf{d}\big(\, f' \times g'(u_0, v_0),\ f \times g(u_0, v_0),\,\big)$$

is maximal and hence equal to the distance between $f \times g$ and $f' \times g'$. The displayed quantity is equal to the greater of $\mathbf{d}\big(\, f'(u_0),\, f(u_0)\,\big)$ and $\mathbf{d}\big(\, g'(v_0),\, g(v_0)\,\big)$. These quantities in turn are less than or equal to $\mathbf{d}(f', f)$ and $\mathbf{d}(g', g)$ respectively. Therefore if both of the latter are less than ε it follows that the distance between $f' \times g'$ and $f \times g$ is less than ε.■

5. (i) Follow the hint. We then have $V(h) \circ V(f) = V(h \circ f) = V(\mathrm{id})$, which is the identity. Likewise, we also have $V(f) \circ V(h) = V(f \circ h) = V(\mathrm{id})$, which is the identity.■

(ii) Again follow the hint. We then have $U(f) \circ U(h) = U(h \circ f) = V(\mathrm{id})$, which is the identity. Likewise, we also have $U(h) \circ V(f) = V(f \circ h) = V(\mathrm{id})$, which is the identity.■

(iii) Let $f : A \to A'$ and $g : B \to B'$ be the homeomorphisms. Let $U(f^{-1}) \circ V(g) = V(g) \circ U(f^{-1})$ — equality holds by associativity of composition. By the first two parts of the exercise this map is the required homeomorphism from $\mathbf{C}(A, B)$ to $\mathbf{C}(A', B')$.■

6. It will be convenient to denote $\alpha(\gamma, \gamma')$ generically by $\gamma + \gamma'$. The construction then implies that the distance between $\xi + \xi'$ and $\eta + \eta'$ is the larger $\mathbf{d}(\xi, \eta)$ and $\mathbf{d}(\xi', \eta')$. Therefore the concatenation map is distance preserving.

7. Follow the hint. If $y \neq y'$, then the values of $k(y)$ at every point is X, and hence it is not equal to y', the value of $k(y')$ at every point of X. Therefore k is 1–1.

Next, we shall verify the set-theoretic identities described above. If $y \in U$ then since $k(y)$ is the function whose value is y at every point we clearly have $k(y) \in \mathcal{W}(K, U)$, and hence $k(U) \subset \mathcal{W}(K, U) \cap \mathrm{Image}(k)$. Conversely, any constant function in the image of the latter is equal to $k(y)$ for some $y \in U$. The identity $k^{-1}(\, \mathcal{W}(K, U)\,) = U$ follows similarly.■

8. As indicated in the hint, without loss of generality we may assume that $s < t$. Given a function f in $\mathbf{Diff}(A, B)$, the Mean Value Theorem implies that

$$f(t) \ - \ f(s) \quad = \quad f'(\xi) \cdot (t - s)$$

for some $\xi \in (s, t)$. By hypothesis $|f'(x)| \leq B$ for all x, and therefore $|f(t) - f(s)| = |f'(\xi)||t - s| \leq B \cdot |t - s|$. Therefore, if $\varepsilon > 0$ and $\delta = \varepsilon / B$, then $|s - t| < \delta$ implies $|f(s) - f(t)| < \varepsilon$.■

IV.2 : Adjoint equivalences

Additional exercises

1. The map $A : \mathbf{F}(X \times Y, Z) \to \mathbf{F}(X, \mathbf{F}(Y, Z)\,)$ sends $h : X \times Y \to Z$ to the function $h^\flat$ such that $[h^\flat(x)](y) = h(x, y)$. The argument proving the adjoint formula for spaces of continuous functions modifies easily to cover these examples, and in fact in this case the proof is a bit easier because it is not necessary to consider metrics or topologies.

2. Let $y_0 \in Y$, and let $L : Y \times [0,1] \to Y$ be the map sending (y,t) to the point $(1-t)y + ty_0$ on the line segment joining y to y_0. If we set $H(x,t) = L(f(x),t)$, then H satisfies the conditions in the hint and defines a continuous map in $\mathbf{C}(X,Y)$ joining f to the constant function whose values is always y_0. Thus for each f we know that f and the constant function with value y_0 lie in the same arc component of $\mathbf{C}(X,Y)$. Therefore there must be only one arc component.■

3. By the adjoint formula there are homeomorphisms

$$\mathbf{C}\big(X,\,\mathbf{C}(Y,Z)\big) \;\cong\; \mathbf{C}(X \times Y, Z) \;\cong\; \mathbf{C}(Y \times X, Z) \;\cong\; \mathbf{C}\big(Y,\,\mathbf{C}(X,Z)\big)$$

and this yields the desired 1–1 correspondence of sets.■

V. Constructions on spaces

V.1 : Quotient spaces

Problem from Munkres, § 22, pp. 144 – 145

4. (*a*) The hint describes a well-defined continuous map from the quotient space W to the real numbers. The equivalence classes are simply the curves $g(x, y) = C$ for various values of C, and they are parabolas that open to the left and whose axes of symmetry are the x-axis. It follows that there is a 1–1 onto continuous map from W to $\mathbb{R}$. How do we show it has a continuous inverse? The trick is to find a continuous map in the other direction. Specifically, this can be done by composing the inclusion of $\mathbb{R}$ in $\mathbb{R}^2$ as the x-axis with the quotient projection from $\mathbb{R}^2$ to W. This gives the set-theoretic inverse to $\mathbb{R}^2 \to W$ and by construction it is continuous. Therefore the quotient space is homeomorphic to $\mathbb{R}$ with the usual topology.$\blacksquare$

(*b*) Here we define $g(x, y) = x^2 + y^2$ and the equivalence classes are the circles $g(x, y) = C$ for $C > 0$ along with the origin. In this case we have a continuous 1–1 onto map from the quotient space V to the nonnegative real numbers, which we denote by $[0, \infty)$ as usual. To verify that this map is a homeomorphism, consider the map from $[0, \infty)$ to V given by composing the standard inclusion of the former as part of the x-axis with the quotient map $\mathbb{R}^2 \to V$. This is a set-theoretic inverse to the map from V to $[0, \infty)$ and by construction it is continuous.$\blacksquare$

Additional exercises

0. We claim that every subset of $X/\mathcal{R}$ is both open and closed. But a subset of the quotient is open and closed if and only if the inverse image has these properties, and every subset of a discrete space has these properties.$\blacksquare$

1. We need to show that $U \subset A$ is open if and only if $r^{-1}[U]$ is open in X. The ($\Longrightarrow$) implication is immediate from the continuity of r. To prove the other direction, note that $r \circ i = \mathrm{id}_A$ implies that

$$U \quad = \quad i^{-1}\left[r^{-1}[U]\right] \quad = \quad A \cap r^{-1}[U]$$

and thus if the inverse image of U is open in X then U must be open in A.$\blacksquare$

2. Let p_X and p_A denote the quotient space projections for X and A respectively. By construction, j is the unique function such that $j \circ p_A = p_X|A$ and therefore j is continuous. We shall define an explicit continuous inverse $k : X/\mathcal{R} \to A/\mathcal{R}_0$. To define the latter, consider the continuous map $p_A \circ r : X \to A/\mathcal{R}_0$. If $y\mathcal{R}z$ holds then $r(y)\mathcal{R}_0\, r(z)$, and therefore the images of y and z in $A/\mathcal{R}_0$ are equal. Therefore there is a unique continuous map k of quotient spaces such that $k \circ p_X = p_A \circ r$. This map is a set-theoretic inverse to j and therefore j is a homeomorphism.$\blacksquare$

3. (*a*) The relation is reflexive because $x = 1 \cdot x$, and it is reflexive because $y = \alpha x$ for some $\alpha \neq 0$ implies $x = \alpha^{-1} y$. The relation is transitive because $y = \alpha x$ for $\alpha \neq 0$ and $z = \beta y$ for $y \neq 0$ implies $z = \beta\alpha x$, and $\beta\alpha \neq 0$ because the product of nonzero real numbers is nonzero.$\blacksquare$

(*b*) Use the hint to define r; we may apply the preceding exercise if we can show that for each $a \in S^2$ the set $r^{-1}(\{a\})$ is contained in an $\mathcal{R}$-equivalence class. By construction $r(v) = |v|^{-1}v$, so $r(x) = a$ if and only if x is a positive multiple of a (if $x = \rho\, a$ then $|x| = \rho$ and $r(x) = a$, while if

$a = r(x)$ then by definition a and x are positive multiples of each other). Therefore if $x\mathcal{R}y$ then $r(x) = \pm\, r(y)$, so that $r(x)\mathcal{R}_0\, r(y)$ and the map

$$S^2/[x \equiv \pm x] \quad \longrightarrow \quad \mathbb{RP}^2$$

is a homeomorphism.■

4. Needless to say we shall follow the hints in a step by step manner.

Let $h : D^2 \to S^2$ be defined by

$$h(x,y) = (x, y, \sqrt{1 - x^2 - y^2}\,).$$

Verify that h preserves equivalence classes and therefore induces a continuous map $\overline{h}$ on quotient spaces.

To show that $\overline{h}$ is well-defined it is only necessary to show that its values on the $\mathcal{R}'$-equivalence classes with two elements are the same for both representatives. If $\pi : S^2 \to \mathbb{RP}^2$ is the quotient projection, this means that we need $\pi\circ h(u) = \pi\circ h(v)$ if $|u| = |v| = 1$ and $u = -v$. This is immediate from the definition of the equivalence relation on S^2 and the fact that $h(w) = w$ if $|w| = 1$.

Why is $\overline{h}$ a $1-1$ and onto mapping?

By construction h maps the equivalence classes of points on the unit circle onto the points of S^2 with $z = 0$ in a 1–1 onto fashion. On the other hand, if u and v are distinct points that are not on the unit circle, then $h(u)$ cannot be equal to $\pm\, h(v)$. The inequality $h(u) \neq -h(v)$ follows because the first point has a positive z-coordinate while the second has a negative z-coordinate. The other inequality $h(u) \neq h(v)$ follows because the projections of these points onto the first two coordinates are u and v respectively. This shows that $\overline{h}$ is 1–1. To see that it is onto, recall that we already know this if the third coordinate is zero. But every point on S^2 with nonzero third coordinate is equivalent to one with positive third coordinate, and if $(x,y,z) \in S^2$ with $z > 0$ then simple algebra shows that the point is equal to $h(x,y)$.

Finally, prove that $\mathbb{RP}^2$ is Hausdorff and $\overline{h}$ is a closed mapping.

If the first statement is true, then the second one follows because the domain of $\overline{h}$ is a quotient space of a compact space and continuous maps from compact spaces to Hausdorff spaces are always closed. Since $\overline{h}$ is already known to be continuous, 1–1 and onto, this will prove that it is a homeomorphism.

So how do we prove that $\mathbb{RP}^2$ is Hausdorff? Let v and w be points of S^2 whose images in $\mathbb{RP}^2$ are distinct, and let P_v and P_w be their orthogonal complements in $\mathbb{R}^3$ (hence each is a 2-dimensional vector subspace and a closed subset). Since Euclidean spaces are Hausdorff, we can find an $\varepsilon > 0$ such that $N_\varepsilon(v) \cap P_v = \emptyset$, $N_\varepsilon(w) \cap P_w = \emptyset$, $N_\varepsilon(v) \cap N_\varepsilon(w) = \emptyset$, and $N_\varepsilon(-v) \cap N_\varepsilon(w) = \emptyset$. If T denotes multiplication by -1 on $\mathbb{R}^3$, then these conditions imply that the four open sets

$$N_\varepsilon(v), \quad N_\varepsilon(w), \quad N_\varepsilon(-v) = T\left(N_\varepsilon(v)\right), \quad N_\varepsilon(-w) = T\left(N_\varepsilon(w)\right)$$

are pairwise disjoint. This implies that the images of the distinct points $\pi(v)$ and $\pi(w)$ in $\mathbb{RP}^2$ lie in the disjoint subsets $\pi\left[\,N_\varepsilon(v)\,\right]$ and $\pi\left[\,N_\varepsilon(w)\,\right]$ respectively. These are open subsets in $\mathbb{RP}^2$ because their inverse images are given by the open sets $N_\varepsilon(v) \cup N_\varepsilon(-v)$ and $N_\varepsilon(w) \cup N_\varepsilon(-w)$ respectively.■

5. A set W belongs to $(g\circ f)_*\mathbf{T}$ if and only if $(g\circ f)^{-1}[W]$ is open in X. But

$$(g\circ f)^{-1}[W] = f^{-1}\left[\,g^{-1}[W]\,\right]$$

so the condition on W holds if and only if $g^{-1}[W]$ belongs to $f_*\mathbf{T}$. The latter in turn holds if and only if w belongs to $g_*\,(f_*\mathbf{T})$.■

6. The object on the left hand side is the family of all sets having the form $(f\circ h)^{-1}[V]$ where V belongs to $\mathbf{T}$. As in the preceding exercise we have

$$(f\circ h)^{-1}[V] = h^{-1}\left[\,f^{-1}[V]\,\right]$$

so the family in question is just $h^*\,(f^*\mathbf{T})$.■

7. Let $p : X \times Y \to X$ be projection onto the first coordinate. Then $u\mathcal{R}v$ implies $p(u) = p(v)$ and therefore there is a unique continuous map $X \times Y/\mathcal{R} \to X$ sending the equivalence class of (x, y) to x. Set-theoretic considerations imply this map is 1–1 and onto, and it is a homeomorphism because p is an open mapping.■

8. (a) If $X/\mathcal{R}$ is Hausdorff then the diagonal $\Delta(X/\mathcal{R})$ is a closed subset of $(X/\mathcal{R}) \times (X/\mathcal{R})$. But $\pi \times \pi$ is continuous, and therefore the inverse image of $\Delta(X/\mathcal{R})$ must be a closed subset of $X \times X$. But this set is simply the graph of $\mathcal{R}$.■

(b) If π is open then so is $pi \times \pi$, for the openness of π implies that $\pi \times \pi$ takes basic open subsets of $X \times X$ into open subsets of $(X/\mathcal{R}) \times (X/\mathcal{R})$. By hypothesis the complementary set $X \times X - \Gamma_{\mathcal{R}}$ is open in $X \times X$, and therefore its image, which is

$$(X/\mathcal{R}) \times (X/\mathcal{R}) \;-\; \Delta(X/\mathcal{R})$$

must be open in $(X/\mathcal{R}) \times (X/\mathcal{R})$. But this means that the diagonal $\Delta(X/\mathcal{R})$ must be a closed subset of $(X/\mathcal{R})\times(X/\mathcal{R})$ and therefore that $X/\mathcal{R}$ must satisfy the Hausdorff Separation Property.■

(c) The condition on $\Gamma_{\mathcal{R}}$ implies that each equivalence class is open. But this means that each point in $X/\mathcal{R}$ must be open and hence the latter must be discrete.■

9. (i) The binary relation $\mathcal{R}$ is symmetric and transitive but not symmetric. Therefore the equivalence relation $\mathcal{E}$ generated by $\mathcal{R}$ consists of the union $\mathcal{R}$ with the diagonal of D^2; in other words, $u\;\mathcal{E}\;v$ if and only if $u = v$ or $u\;\mathcal{R}\;v$, and if $u \neq v$ then $u\;\mathcal{E}\;v$ if and only if $u = \alpha v$, where $|u| = |v| = 1$ and $\alpha^d = 1$.■

(ii) In order to prove the existence of the continuous mapping h^* we need to show that $h(u) = h(v)$ if $u\;\mathcal{R}\;v$ and $u \neq v$; *i.e.*, $u = \alpha v$, where $|u| = |v| = 1$ and $\alpha^d = 1$. Under these conditions we have $h(z) = (0, z^d)$, so if u and v satisfy the given conditions then $h(u) = (0, u^d) = (0, \alpha^d v^d)$ because $\alpha^d = 1$. Therefore h is constant on equivalence classes, which implies the existence of h^*.

Since D^2 is compact and $\mathbb{C}^2$ is Hausdorff, by Theorem III.1.9, the mapping h^* is a homeomorphism onto its image if and only if it is 1–1. This is equivalent to showing that $h(u) = h(v)$ implies $u\;\mathcal{R}\;v$.

Suppose that $h(u) = h(v)$; taking coordinates, we must have $(1-|u|)u = (1-|v|)v$ and $u^d = v^d$. The latter implies that $|u|^d = |v|^d$, which in turn implies that $|u| = |v|$. There are now two cases. CASE 1: Suppose that $|u| = |v| < 1$. Then $|u| = |v|$ implies $1 - |u| = 1 - |v|$, and if we combine this with the equation for first coordinates we see that $u = v$. CASE 2: Suppose that $|u| = |v| = 1$. In this case $h(u) = (0, u^d)$ and $h(v) = (0, v^d)$, so if the two image points are equal then $1 = (u/v)^d$; therefore if $\alpha = u/v$, then $\alpha^d = 1$ and $u = \alpha\, v$, so that $u\;\mathcal{R}\;v$, which is what we wanted to show.■

(iii) Follow the hint. We know that $z \to z^d$ maps D^2 to itself, and if $z \neq 0$ then the equivalence class of z consists of all numbers of the form $\alpha\, z$, where $\alpha^d = 1$; in the exceptional case where $z = 0$, the equivalence class of z is merely $\{0\}$. In all cases we know that $u \sim v$ implies $u^d = v^d$, so if $h(z) = z^d$ then $u \sim v$ implies $h(u) = h(v)$, which means that $h = h^* \circ p$, where $p : D^2 \to D^2/\mathcal{F}$ is the quotient projection. The mapping h^* is onto because h is onto, so by Theorem III.1.9 it is only necessary to verify that h^* is 1–1, or equivalently that $h(u) = h(v)$ implies $u \sim v$. If $u \neq 0$, then $v \neq 0$ too and we have $u^p = v^p$; as in (ii), this implies that $u = \alpha\, v$ for some α such that $\alpha^d = 1$. On the other hand if $u = 0$ and $h(v) = h(0) = 0$, then $v = 0$ so that $u \sim v$ in this case too.■

V.2 : Sums and cutting and pasting

Additional exercises

1. ($\Longrightarrow$) If X is locally connected then so is every open subset. But each A_α is an open subset, so each is locally connected.■

($\Longleftarrow$) We need to show that for each $x \in X$ and each open set U containing x there is an open subset $V \subset U$ such that $x \in V$ and V is connected. There is a unique α such that $x = i_\alpha(a)$ for some $a \in A_\alpha$. Let $U_0 = i_\alpha^{-1}(U)$. Then by the local connectedness of A_α and the openness of U_0 there is an open connected set V_0 such that $x \in V_0 \subset U_0$. If $V = i_\alpha(V_0)$, then V has the required properties.■

2. X is compact if and only if each A_α is compact and there are only finitely many (nonempty) subsets in the collection.

The ($\Longrightarrow$) implication follows because each A_α is an open and closed subspace of the compact space X and hence compact, and the only way that the open covering $\{ A_\alpha \}$ of X, which consists of pairwise disjoint subsets, can have a finite subcovering is if it contains only finitely many subsets. To prove the reverse implication, one need only use a previous exercise which shows that a finite union of compact subspaces is compact.■

3. Since the exercise asks for details in a sketch to be filled in, we shall begin by reprinting this sketch: — *Let $A \subset S^2$ be the set of all points $(x, y, z) \in S^2$ such that $|z| \leq \frac{1}{2}$, and let B be the set of all points where $|z| \geq \frac{1}{2}$. If $T(x) = -x$, then $T[A] = A$ and $T[B] = B$ so that each of A and B (as well as their intersection) can be viewed as a union of equivalence classes for the equivalence relation that produces $\mathbb{RP}^2$. By construction B is a disjoint union of two pieces $B_\pm$ consisting of all points where sign$(z) = \pm 1$, and thus it follows that the image of B in the quotient space is homeomorphic to $B_+ \cong D^2$. Now consider A. There is a homeomorphism h from $S^1 \times [-1, 1]$ to A sending (x, y, t) to $(\alpha(t)x,\ \alpha(t)y,\ \frac{1}{2}t)$ where*

$$\alpha(t) \quad = \quad \sqrt{1 - \frac{t^2}{4}}$$

and by construction $h(-v) = -h(v)$. The image of A in the quotient space is thus the quotient of $S^1 \times [-1, 1]$ modulo the equivalence relation $u \sim v \iff u = \pm v$. This quotient space is in turn homeomorphic to the quotient space of the upper semicircular arc S^1_+ (all points with nonnegative y-coordinate) modulo the equivalence relation generated by setting $(-1, 0, t)$ equivalent to $(1, 0, -t)$, which yields the Möbius strip. The intersection of this subset in the quotient with the image of B is just the image of the closed curve on the edge of B_+, which also represents the edge curve on the Möbius strip.

We shall go through the insertions needed at various steps in this argument.

Let $A \subset S^2$ be the set of all points $(x, y, z) \in S^2$ such that $|z| \leq \frac{1}{2}$, and let B be the set of all points where $|z| \geq \frac{1}{2}$. If $T(x) = -x$, then $T[A] = A$ and $T(B) = B$ [etc.]

This is true because if $T(v) = w$, then the third coordinates of both points have the same absolute values and of course they satisfy the same inequality relation with respect to $\frac{1}{2}$.

By construction B is a disjoint union of two pieces $B_\pm$ consisting of all points where sign$(z) = \pm 1$,

This is true the third coordinates of all points in B are nonzero.

There is a homeomorphism h from $S^1 \times [-1, 1]$ to A sending (x, y, t) to $(\alpha(t)x, \alpha(t)y, \frac{1}{2}t)$ where

$$\alpha(t)s \quad = \quad \sqrt{1 - \frac{t^2}{4}}$$

One needs to verify that h is 1–1 onto; this is essentially an exercise in algebra. Since we are dealing with compact Hausdorff spaces, continuous mappings that are 1–1 onto are automatically homeomorphisms.

This quotient space $[S^1 \times [-1, 1]$ modulo the equivalence relation $u \sim v \iff u = \pm v]$ is in turn homeomorphic to the quotient space of the upper semicircular arc S^1_+ (all points with nonnegative y-coordinate) modulo the equivalence relation generated by setting $(-1, 0, t)$ equivalent to $(1, 0, -t)$, which yields the Möbius strip.

Let $\mathcal{A}$ and $\mathcal{B}$ be the respective equivalence relations on $S^1_+ \times [-1, 1]$ and $S^1 \times [-1, 1]$, and let **A** and **B** be the respective quotient spaces. By construction the inclusion $S^1_+ \times [-1, 1] \subset S^1 \times [-1, 1]$ passes to a continuous map of quotients, and it is necessary and sufficient to check that this map is 1–1 and onto. This is similar to a previous exercise. Points in $S^1 - S^1_+$ all have negative second coordinates and are equivalent to unique points with positive second coordinates. This implies that the mapping from **A** to **B** is 1–1 and onto at all points except perhaps those whose second coordinates are zero. For such points the equivalence relations given by $\mathcal{A}$ and $\mathcal{B}$ are identical, and therefore the mapping from **A** to **B** is also 1–1 and onto at all remaining points.■

4. We can and shall view X as $A \cup_{\mathrm{id}} B$.

Consider the map $F_0 : A \sqcup B \to A \sqcup B$ defined by H^{-1} on A and the identity on B. We claim that this passes to a unique continuous map of quotients from X to $A \cup_h B$; *i.e.*, the map F_0 sends each nonatomic equivalence classes $\{ (c, 1), (c, 2) \}$ for $X = A \cup_{\mathrm{id}} B$ to a nonatomic equivalence class of the form $\{ (u, 1), (h(u), 2) \}$ for $A \cup_h B$. Since F_0 sends $(c, 1)$ to $(h^{-1}(c), 1)$ and $(c, 2)$ to itself, we can verify the compatibility of F_0 with the equivalence relations by taking $u = h^{-1}(c)$. Passage to the quotients then yields the desired map $F : X \to A \cup_h B$.

To show this map is a homeomorphism, it suffices to define Specifically, start with $G_0 = F_0^{-1}$, so that $G_0 = H$ on A and the identity on B. In this case it is necessary to show that a nonatomic equivalence class of the form $\{ (u, 1), (h(u), 2) \}$ for $A \cup_h B$ gets sent to a nonatomic equivalence class of the form $\{ (c, 1), (c, 2) \}$ for $X = A \cup_{\mathrm{id}} B$. Since G_0 maps the first set to $\{ (h(u), 1), (h(u), 2) \}$ this is indeed the case, and therefore G_0 also passes to a map of quotients which we shall call G.

Finally we need to verify that F and G are inverses to each other. By construction the maps F_0 and G_0 satisfy $F([y]) = [F_0(y)]$ and $G([z]) = [G_0(z)]$, where square brackets denote equivalence classes. Therefore we have

$$G \circ F([y]) = G\left([F_0(y)]\right) = [G_0\left(F_0(y)\right)]$$

which is equal to $[y]$ because F_0 and G_0 are inverse to each other. Therefore $G \circ F$ is the identity on X. A similar argument shows that $F \circ G$ is the identity on $A \cup_h B$.

To construct the example where X is **not** homeomorphic to $A \cup_h B$, we follow the hint and try to find a homeomorphism of the four point space $\{\pm 1\} \times \{1, 2\}$ to itself such that X is **not** homeomorphic to $A \cup_h B$ is connected; this suffices because we know that X is not connected. Sketches on paper or physical experimentation with wires or string are helpful in finding the right formula.

Specifically, the homeomorphism we want is given as follows:

$$\begin{array}{rcl} (-1,1) \in A_+ & \longrightarrow & (1,2) \in A_- \\ (1,1) \in A_+ & \longrightarrow & (1,1) \in A_- \\ (1,2) \in A_+ & \longrightarrow & (-1,1) \in A_- \\ (-1,2) \in A_+ & \longrightarrow & (-1,2) \in A_- \end{array}$$

The first of these implies that the images of $S^1_+ \times \{2\}$ and $S^1_- \times \{1\}$ lie in the same component of the quotient space, the second of these implies that the images of $S^1_- \times \{1\}$ and $S^1_+ \times \{1\}$ both lie in the same component, and the third of these implies that the images of $S^1_+ \times \{2\}$ and $S^1_- \times \{2\}$ also lie in the same component. Since the entire space is the union of the images of the connected subsets $S^1_\pm \times \{1\}$ and $S^1_\pm \times \{2\}$ it follows that $A \cup_h B$ is connected.■

FOOTNOTE. The argument in the first part of the exercise remains valid if A and B are open rather than closed subsets.■

5. (*a*) For each j let $\mathbf{in}_j : X_j \to \coprod_k X_k$ be the standard injection into the disjoint union, and let

$$P : \coprod_k X_k \longrightarrow \bigvee_k (X_k, x_k)$$

be the quotient map defining the wedge. Define Y_j to be $P \circ \mathbf{in}_j[\, X_j\,]$. By construction the map $P \circ \mathbf{in}_j$ is continuous and 1–1; we claim it also sends closed subsets of X_j to closed subsets of the wedge. Suppose that $F \subset X_j$ is closed; then $P \circ \mathbf{in}_j[F]$ is closed in the wedge if and only if its inverse image under P is closed. But this inverse image is the union of the closed subsets $\mathbf{in}_j[F]$ and $\coprod_k \{x_k\}$ (which is a finite union of one point subsets that are assumed to be closed). It follows that Y_j is homeomorphic to $X + j$. The condition on $Y_k \cap Y_\ell$ for $k \neq \ell$ is an immediate consequence of the construction.

The assertion that the wedge is Hausdorff if and only if each summand is follows because a subspace of a Hausdorff space is Hausdorff, and a finite union of closed Hausdorff subspaces is always Hausdorff (by a previous exercise).

To verify the assertions about compactness, note first that for each j there is a continuous collapsing map q_j from $\vee_k\ (X_k, x_k)$ to (X_j, x_j), defined by the identity on the image of (X_j, x_j) and by sending everything to the base point on every other summand. If the whole wedge is compact, then its continuous under q_j, which is the image of X_j, must also be compact. Conversely if the sets X_j are compact for all j, then the (finite!) union of their images, which is the entire wedge, must be compact.

To verify the assertions about connectedness, note first that for each j there is a continuous collapsing map q_j from $\vee_k\ (X_k, x_k)$ to (X_j, x_j), defined by the identity on the image of (X_j, x_j) and by sending everything to the base point on every other summand. If the whole wedge is connected, then its continuous under q_j, which is the image of X_j, must also be connected. Conversely if the sets X_j are connected for all j, then the union of their images, which is the entire wedge, must be connected because all these images contain the base point. Similar statements hold for arcwise connectedness and follow by inserting "arcwise" in front of "connected" at every step of the argument.■

(b) To prove existence, first observe that there is a unique continuous map $\widetilde{F} : \coprod_k X_k \to Y$ such that $\mathbf{in}_j \circ \widetilde{F} = F_j$ for all j. This passes to a unique continuous map F on the quotient space $\vee_k \, (X_k, x_k)$ because $\widetilde{F}$ is constant on the equivalence classes associated to the quotient projection P. This constructs the map we want; uniqueness follows because the conditions prescribe the definition at every point of the wedge.■

(c) Strictly speaking, one should verify that the so-called weak topology is indeed a topology on the wedge. We shall leave this to the reader.

To prove [1], note that ($\Longrightarrow$) is trivial. For the reverse direction, we need to show that if E is closed in Y then $h^{-1}[E]$ is closed with respect to the so-called weak topology we have defined. The subset in question is closed with respect to this topology if and only if $h^{-1}[E] \cap \varphi[X_j]$ is closed in $\varphi[X_j]$ for all j, and since φ_j maps its domain homeomorphically onto its image, the latter is true if and only if $\varphi^{-1} \circ h^{-1}[E]$ is closed in X_j for all j. But these conditions hold because each of the maps $\varphi_j \circ h$ is continuous. To prove [2], note first that there is a unique set-theoretic map, and then use [1] to conclude that it is continuous.■

(d) For each j let $y_j \in X_j$ be a point other than x_j, and consider the set E of all points y_j. This is a closed subset of the wedge because its intersection with each set $\varphi[X_j]$ is a one point subset and hence closed. In fact, every subset of E is also closed by a similar argument (the intersections with the summands are either empty or contain only one point), so E is a discrete closed subset of the wedge. Compact spaces do not have infinite discrete closed subspaces, and therefore it follows that the infinite wedge with the weak topology is not compact.■

We shall conclude this document by filling in some details in the final remark in the exercises for Section V.2. This remark is reprinted here for the sake of convenience:

Remark. If each of the summands in (d) is compact Hausdorff, then there is a natural candidate for a *strong topology* on a countably infinite wedge which makes the latter into a compact Hausdorff space. In some cases this topology can be viewed more geometrically; for example, if each (X_j, x_j) is equal to $(S^1, 1)$ and there are countably infinitely many of them, then the space one obtains is the *Hawaiian earring* in $\mathbb{R}^2$ given by the union of the circles defined by the equations

$$\left(x - \frac{1}{2^k}\right)^2 + + y^2 \quad = \quad \frac{1}{2^{2k}} \ .$$

As usual, drawing a picture may be helpful. The k^{th} circle has center $(1/2^k,\, 0)$ and passes through the origin; the y-axis is the tangent line to each circle at the origin.

SKETCHES OF VERIFICATIONS OF ASSERTIONS IN THIS REMARK.

If we are given an infinite sequence of compact Hausdorff pointed spaces $\{\,(X_n, x_n)\,\}$ we can put a compact Hausdorff topology on their wedge as follows. Let W_k be the wedge of the first k spaces; then for each k there is a continuous map

$$q_k : \bigvee_n (X_n, x_n) \quad \longrightarrow \quad W_k$$

(with the so-called weak topology on the wedge) that is the identity on the first k summands and collapses the remaining ones to the base point. These maps are in turn define a continuous function

$$\mathbf{q} : \bigvee_n (X_n, x_n) \quad \longrightarrow \quad \prod_k W_k$$

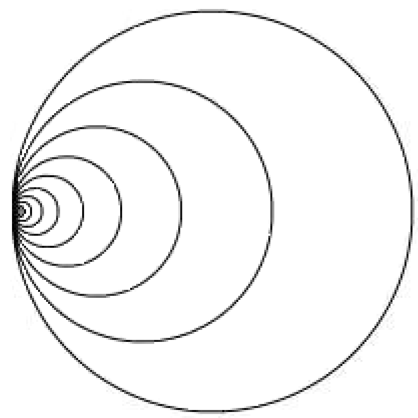
whose projection onto W_k is q_k. This mapping is continuous and 1–1; if its image is closed in the (compact!) product topology, then this defines a compact Hausdorff topology on the infinite wedge $\vee_n(X_n, x_n)$.

Here is one way of verifying that the image is closed. For each k let $c_k : W_k \to W_{k-1}$ be the map that is the identity on the first $(k-1)$ summands and collapses the last one to a point. Then we may define a continuous map C on $\prod_{k\geq 1} W_k$ by first projecting onto the product $\prod_{k\geq 2} W_k$ (forget the first factor) and then forming the map $\prod_{k\geq 2} W_k$. The image of $\mathbf{q}$ turns out to be the set of all points $\mathbf{x}$ in the product such that $C(\mathbf{x}) = \mathbf{x}$. Since the product is Hausdorff the image set is closed in the product and thus compact.

A comment about the compactness of the Hawaiian earring E might be useful. Let F_k be the union of the circles of radius 2^{-j} that are contained in E, where $j \leq k$, together with the closed disk bounded by the circle of radius $2^{-(k+1)}$ in E. Then F_k is certainly closed and compact. Since E is the intersection of all the sets F_k it follows that E is also closed and compact.■

Drawing to accompany Additional Exercise $\mathrm{V}.2.8(b)$

The Hawaiian earring is a union of circles which are tangent to the $y-$ axis at the origin, lie in the right hand half – plane defined by that line, and have decreasing radii which go to 0 in the limit. One way of distinguishing the weak and strong topologies is the following: Suppose that we pick one point on each circle, never choosing the common point of all the circles. In the weak topology these points form a closed subset, but in the strong topology they have a limit point (namely, the point which is common to all the circles).



(***Source:*** **http://en.wikipedia.org/wiki/Semi-locally_simply_connected#mediaviewer/File:Hawaiian-Earring-250.jpg**)

SOLUTIONS TO EXERCISES FOR

MATHEMATICS 205A — Part 5

Fall 2014

VI. Spaces with additional properties

VI.1 : Second countable spaces

Problems from Munkres, § 30, *pp.* 194 – 195

9. [*First part only*] The statement and proof are parallel to a result about compact spaces in the course notes, the only change being that "compact" is replaced by "Lindelöf."■

10. Suppose first that X is a finite product of spaces Y_i such that each Y_i has a countable dense subset D_i. Then $\prod_i D_i$ is countable and

$$X \;=\; \prod_i Y_i \;=\; \prod_i \overline{D_i} \;=\; \overline{\prod_\alpha D_i}\;.$$

Suppose now that X is countably infinite. The same formula holds, but the product of the D_i's is not necessarily countable. To adjust for this, pick some point $\delta_j \in D_j$ for each j and consider the set E of all points $(a_0, a_1, \cdots)$ in $\prod_j D_j$ such that $a_j = \delta_j$ for all but at most finitely many values of j. This set is countable, and we claim it is dense. It suffices to show that every basic open subset contains at least one point of E. But suppose we are given such a set $V = \prod_j V_j$ where V_j is open in X_j and $V_j = X_j$ for all but finitely many j; for the sake of definiteness, suppose this happens for $j > M$. For $j \leq M$, let $b_j \in D_j \cap V_j$; such a point can be found since D_j is dense in X_j. Set $b_j = \delta_j$ for $j > M$. If we let $b = (b_0, b_1, \cdots)$, then it then follows that $b \in E \cap V$, and this implies E is dense in the product.■

13. This is similar to the proof that an open subset of $\mathbb{R}^n$ has only countably many components.■

14. This is essentially the same argument as the one showing that a product of two compact spaces is compact. The only difference is that after one constructs an open covering of Y at one step in the proof, then one only has a countable subcovering and this leads to the existence of a countable subcovering of the product. Details are left to the reader.■

Additional exercises

1. List the basic open subsets in $\mathcal{B}$ as a sequence $U_0, U_1, U_2, \cdots$ If W is an open subset of U define a function $\psi_W : \mathbb{N} \to \{0,1\}$ by $\psi_W(i) = 1$ if $U_i \subset W$ and 0 otherwise. Since $\mathcal{B}$ is a basis, every open set W is the union of the sets U_i for which $\psi_W(i) = 1$. In particular, the latter implies that if $\psi_V = \psi_W$ then $V = W$ and hence we have a 1–1 map from the set of all open subsets to the set of functions from $\mathbb{N}$ to $\{0,1\}$. Therefore the cardinality of the family of open subsets is at most the cardinality of the set of functions, which is $2^{\aleph_0}$.

If X is Hausdorff, or even if we only know that every one point subset of X is closed in X, then we may associate to each $x \in X$ the open subset $X - \{x\}$. If $x \neq y$ then $X - \{x\} \neq X - \{y\}$

and therefore the map $C : X \to \mathbf{T}$ defined in this fashion is 1–1. Therefore we have $|X| \leq |\mathbf{T}|$, and therefore by the preceding paragraph we know that $|X| \leq 2^{\aleph_0}$.■

A somewhat more complicated argument yields similar conclusions for spaces satisfying the weaker $\mathbf{T_0}$ condition stated in Section VI.3.■

2. (*a*) The x-axis is closed because it is closed in the ordinary Euclidean topology and the "new" topology contains the Euclidean topology; therefore the x-axis is closed in the "new" topology. The subspace topology on the x-axis is the discrete topology intersection of the open set $T_\varepsilon(x)$ with the real axis is $\{\, x \,\}$.■

(*b*) In general if $(X, \mathbf{T})$ is Hausdorff and $\mathbf{T} \subset \mathbf{T}^*$ then $(X, \mathbf{T}^*)$ is also Hausdorff, for a pair of disjoint $\mathbf{T}$-open subsets containing distinct point u and v will also be a pair of $\mathbf{T}$-open subsets with the same properties.■

(*c*) To show that a subset D of a topological space X is dense, it suffices to show that the intersection of D with every nonempty open subset in some base $\mathcal{B}$ is nonempty (why?). Thus we need to show that the generating set is a base for the topology and that every basic open subset contains a point with rational coordinates.

The crucial point to showing that we have a base for the topology is to check that if U and V are basic open subsets containing some point $z = (x, y)$ then there is some open subset in the generating collection that contains z and is contained in $U \cap V$. If U and V are both basic open subsets we already know this, while if $z \in T_\varepsilon(a) \cap T_\delta(b)$ there are two cases depending upon whether or not z lies on the x-axis; denote the latter by A. If $z \notin A$ then it lies in the metrically open subsets $T_\varepsilon(a) - A$ and $T_\varepsilon(b) - A$, and one can find a metrically open subset that contains z and is contained in the intersection. On the other hand, if $z \in A$, then $T_\varepsilon(a) \cap A = \{a\}$ and $T_\delta(b) \cap A = \{b\}$ imply $a = b$, and the condition on intersections is immediate because the intersection of the subsets is either $T_\varepsilon(a)$ or $T_\delta(a)$ depending upon whether $\delta \leq \varepsilon$ or vice versa.■

3. For each property $\mathcal{P}$ given in the exercise, the space X has property $\mathcal{P}$ if and only if each A_α does and there are only finitely many α for which A_α is nonempty. The verifications for the separate cases are different and will be given in reverse sequence.

The Lindelöf property.

The proof in this case is the same as the proof we gave for compactness in an earlier exercises with "countable" replacing "finite" throughout.■

Separability.

($\Longrightarrow$) Let D be the countable dense subset. Each A_α must contain some point of D, and by construction this point is not contained in any of the remaining sets A_β. Thus we have a 1–1 function from $\mathcal{A}$ to D sending α to a point $d(\alpha) \in A_\alpha \cap D$. This implies that the cardinality of $\mathcal{A}$ is at most $|D| \leq \aleph_0$.■

($\Longleftarrow$) If D_α is a dense subset of A_α and $\mathcal{A}$ is countable, then $\cup_\alpha\, D_\alpha$ is a countable dense subset of $\mathcal{A}$.■

Second countability.

($\Longrightarrow$) Since a subspace of a second countable space is second countable, each A_α must be second countable. Since the latter condition implies both separability and the Lindelöf property, the preceding arguments show that only countably many summands can be nontrivial.

($\Longleftarrow$) If $\mathcal{A}$ is countable and $\mathcal{B}_\alpha$ is a countable base for A_α then $\cup_\alpha\, \mathcal{B}_\alpha$ determines a countable base for X (work out the details!).■

4. More generally, if $\mathcal{B}_X$ is a base for the topology on X and $\mathcal{B}_Y$ is a base for the topology on Y, then a base for the topology on $X \times Y$ is given by all sets of the form $U \times V$ where $U \in \mathcal{B}_X$ and

$V \in \mathcal{B}_Y$. — To prove this, it is enough to show it for basic open subsets $M \times N$ in $X \times Y$. We then have $M = \cup_i U_i$ and $N = \cup_j V_j$ for $U_i \in \mathcal{B}_X$ and $V_j \in \mathcal{B}_Y$, which means that $M \times N = \cup_{i,j} U_i \times V_j$.

If both bases are countable, then the unions in the preceding sentence can be taken to be countable if we remove duplicate summands (there are only countably many possibilities for U_i or V_j), and this implies that the indexing set for the decomposition $M \times N = \cup_{i,j} U_i \times V_j$ is also countable.■

VI.2 : Compact spaces – II

Problems from Munkres, § 28, pp. 181 – 182

6. Follow the hint to define a and the sequence. The existence of ε is guaranteed because $a \notin f[X]$ and the compactness of $f[X]$ imply that the continuous function $g(x) = \mathbf{d}(a, f(x))$ is positive valued, and it is bounded away from 0 because it attains a minimum value. Since $x_k \in f[X]$ for all k it follows that $\mathbf{d}(a, x_k) > \varepsilon$ for all k. Given $n \neq m$ write $m = n + k$; reversing the roles of m and n if necessary we can assume that $k > 0$. If f is distance preserving, then so is every n-fold iterated composite $\circ^n f$ of f with itself. Therefore we have that

$$\mathbf{d}(x_n, x_m) \quad = \quad \mathbf{d}\,(\circ^n f(a), \circ^n f(x_k)) \quad = \quad \mathbf{d}(a, x_k) \quad > \quad \varepsilon$$

for all distinct nonnegative integers m and n. But this cannot happen if X is compact, because the latter implies that $\{\, x_n \,\}$ has a convergent subsequence. This contradiction implies that our original assumption about the existence of a point $a \notin f[X]$ is false, so if is onto. On the other hand $x \neq y$ implies

$$0 \quad < \quad \mathbf{d}(x, y) \quad = \quad \mathbf{d}\,(\, f(x), f(y)\,)$$

and thus that f is also 1–1. Previous results now also imply that f is a homeomorphism onto its image.■

FOOTNOTE. To see the need for compactness rather than (say) completeness, consider the map $f(x) = x + 1$ on the set $[0, +\infty)$ of nonnegative real numbers.■

Additional exercises

1. Follow the steps in the hint.

> *A map g is $1 - 1$ on a subset B is and only if $B \times B \cap (g \times g)^{-1}[\,\Delta_Y\,] = \Delta_B$ where Δ_S denotes the diagonal in $S \times S$.*

This is true because the set on the left hand side of the set-theoretic equation is the set of all (b, b') such that $g(b) = g(b')$. If there is a nondiagonal point in this set then the function is not 1–1, and conversely if the function is 1–1 then there cannot be any off-diagonal terms in the set.■

> *Show that if $D' = (f \times f)^{-1}[\,\Delta_Y\,] - \Delta_X$, then D' is closed and disjoint from the diagonal.*

Since f is locally 1–1, for each $x \in X$ there is an open set U_x such that $f|U_x$ is 1–1; by the first step we have

$$U_x \times U_x \cap (f \times f)^{-1}[\,\Delta_Y\,] \quad = \quad \Delta_{U_x} \ .$$

If we set $W = \cap_x U_x \times U_x$ then W is an open neighborhood of Δ_X and by construction we have

$$(f \times f)^{-1}[\,\Delta_Y\,] \;\cap\; W \;\;=\;\; \Delta_X$$

which shows that Δ_X is open in $(f \times f)^{-1}[\,\Delta_Y\,]$ and thus its relative complement in the latter — which is D' — must be closed.■

Show that the subsets D' and $A \times A$ are disjoint, and find a square neighborhood of $A \times A$ disjoint from D'.

Since $f|A$ is 1–1 we have $A \times A \cap (f \times f)^{-1}[\,\Delta_Y\,] = \Delta_A$, which lies in $\Delta_X = (f \times f)^{-1}[\,\Delta_Y\,] - D'$. Since A is a compact subset of the open set $X \times X - D'$, by Wallace's Theorem there is an open set U such that

$$A \times A \quad \subset \quad U \times U \quad \subset \quad X \times X - D' \; .$$

The first part of the proof now implies that $f|U$ is 1–1.■

2. By the Inverse Function Theorem we know that f is locally 1–1, and therefore by the previous exercise we know that f is 1–1 on an open set V such that $A \subset V \subset U$. But the Inverse Function Theorem also implies that f locally has a $\mathbf{C}^1$ inverse on U. Since f has a global set-theoretic inverse from $f[V]$ back to V, it follows that this global inverse is also $\mathbf{C}^1$.■

3. As usual, we follow the steps in the hint.

For each integer $r > 0$ show that every integer is within p^{-r} of one of the first p^{r+1} nonnegative integers.

If n is an integer, use the long division property to write $n = p^{r+1}a$ where $0 \leq a < p^{r+1}$. We then have $\mathbf{d}_p(n, a) \leq p^{-(r+1)} < p^r$.■

Furthermore, each open neighborhood of radius p^{-r} centered at one of these integers a is a union of p neighborhoods of radius $p^{-(r+1)}$ over all of the first p^{r+2} integers b such that $b \equiv a \bmod p^{r+1}$.

Suppose we do long division by p^{r+2} rather than p^{r+1} and get a new remainder a'. How is it related to a? Very simply, a' must be one of the p nonnegative integers b such that $0 \leq b < p^{r+2}$ and $b \equiv a \bmod p^{r+1}$. Since $\mathbf{d}(u, v) < p^{-r}$ if and only if $\mathbf{d}_p(u, v,) \leq p^{-(r+1)}$, it follows that $N_{p^{-r}}(a)$ is a union of p open subsets of the form $N_{p^{-(r+1)}}(b)$ as claimed (with the numbers b satisfying the asserted conditions).

Find a sequence of nonnegative integers $\{\, b_r \,\}$ such that the open neighborhood of radius p^{-r} centered at b_r contains infinitely many terms in the sequence and $b_{r+1} \equiv b_r \bmod p^{r+1}$.

This is very similar to a standard proof of the Bolzano-Weierstrass Theorem in real variables. Infinitely many terms of the original sequence must lie in one of the sets $N_1(b_0)$ where $1 \leq b < p$. Using the second step we know there is some b_1 such that $b_1 \equiv b_0 \bmod p$ and infinitely many terms of the sequence lie in $N_{p^{-1}}(b_1)$, and one can continue by induction (fill in the details!).■

Form a subsequence of $\{\, a_n \,\}$ by choosing distinct points $a_{n(k)}$ recursively such that $n(k) > n(k-1)$ and $a_{n(k)} \in N_{p^{-k}}(b_k)$. Prove that this subsequence is a Cauchy sequence and hence converges.]

We know that the neighborhoods in question contain infinitely many values of the sequence, and this allows us to find $n(k)$ recursively. It remains to show that the construction yields a Cauchy sequence. The key to this is to observe that $a_{n(k)} \equiv b_k \bmod p^{k+1}$ and thus we also have

$$\mathbf{d}_p\left(a_{n(k+1)}, a_{n(k)}\right) \quad < \quad \frac{1}{p^k} \; .$$

Similarly, if $\ell > k$ then we have

$$\mathbf{d}_p \left(a_{n(\ell)}, a_{n(k)}\right) \quad < \quad \frac{1}{p^{\ell-1}} \; + \; \cdots \; + \; \frac{1}{p^k} \; .$$

Since the geometric series $\sum_k p^{-k}$ converges, for every $\varepsilon > 0$ there is an M such that $\ell, k \geq M$ implies the right had side of the displayed inequality is less than ε, and therefore it follows that the constructed subsequence is indeed a Cauchy sequence. By completeness (of the completion) this sequence converges. Therefore $\widehat{\mathbb{Z}_p}$ is (sequentially) compact.■

VI.3 : Separation axioms

Problem from Munkres, § 26*, pp.* 170 – 172

11. Follow the suggestion of the hint to define C, D and to find U, V. There are disjoint open subset sets $U, V \subset X$ containing C and D respectively because a compact Hausdorff space is $\mathbf{T_4}$.

For each α the set $B_\alpha = A_\alpha - (U \cup V)$ is a closed and hence compact subset of X. If each of these subsets is nonempty, then the linear ordering condition implies that the family of closed compact subsets B_α has the finite intersection property; specifically, the intersection

$$B_{\alpha(1)} \cap \; \dots \; \cap B_{\alpha(k)}$$

is equal to $B_{\alpha(j)}$ where $A_{\alpha(j)}$ is the smallest subset in the linearly ordered collection

$$\{ \, A_{\alpha(1)}, \; \dots \; , A_{\alpha(k)} \, \} \; .$$

Therefore by compactness it will follow that the intersection

$$\bigcap_\alpha B_\alpha \quad = \quad \left(\bigcap_\alpha A_\alpha \right) - (U \cup V)$$

is nonempty. But this contradicts the conditions $\cap_\alpha A_\alpha \quad = \quad C \cup D \quad \subset \quad U \cup V$. Therefore it follows that the $\cap_\alpha A_\alpha$ must be connected.

Therefore we only need to answer the following question: Why should the sets $A_\alpha - (U \cup V)$ be nonempty? If the intersection is empty then $A_\alpha \subset U \cup V$. By construction we have $C \subset A_\alpha \cap U$ and $D \subset A_\alpha \cap V$, and therefore we can write A_α as a union of two nonempty disjoint open subsets. However, this contradicts our assumption that A_α is connected and therefore we must have $A_\alpha - (U \cup V) \neq \emptyset$.■

Problems from Munkres, § 33*, pp.* 212 – 214

2. (a) [*For metric spaces.*]

The proof is based upon Urysohn's Lemma and therefore is valid in arbitrary $\mathbf{T_4}$ spaces; we have stated it only for metric spaces because we have only established Urysohn's Lemma in that case.

Let X be the ambient topological space and suppose that u and v are distinct points of X. Then $\{u\}$ and $\{v\}$ are disjoint closed subsets and therefore there is a continuous function $f : X \to \mathbb{R}$

such that $f(u) = 0$ and $f(v) = 1$. Since $f[X]$ is connected, for each $t \in [0,1]$ there is a point $x_t \in X$ such that $f(x_t) = t$. Since $s \neq t$ implies $f(x_s) \neq f(x_t)$, it follows that the map $x : [0,1] \to X$ sending t to x_t is 1–1. Therefore we have $2^{\aleph_0} \leq |X|$, and hence X is uncountable.■

6. (*a*) Let $(X, \mathbf{d})$ be a metric space, let $A \subset X$ be closed and let $f(x) = \mathbf{d}(x, A)$. Then

$$A = f^{-1}(\{0\}) = \bigcap_n f^{-1}\Big[\, \big(-\tfrac{1}{n}, \tfrac{1}{n}\big) \,\Big]$$

presents A as a countable intersection of open subsets.■

8. For each $a \in A$ there is a continuous function $f_a : X \to [-1,1]$ that is -1 at a and 0 on B. Let $U_a = f_a^{-1}[[-1,0)]$. Then the sets U_a define an open covering of A and hence there is a finite subcovering corresponding to $U_{a(1)}, \cdots, U_{a(k)}$. Let f_i be the function associated to $a(i)$, let g_i be the maximum of f_i and 0 (so g_i is continuous by a previous exercise), and define

$$f = \prod_{i=1}^{k} g_k .$$

By construction the value of f is 1 on B because each factor is 1 on B, and $f = 0$ on $\cup_i U_{a(i)}$ because $g_j = 0$ on $U_{a(j)}$; since the union contains A, it follows that $f = 0$ on A.■

Additional exercises

1. The strict containment condition implies that the identity map from $(X, \mathbf{T})$ to $(X, \mathbf{T}^*)$ is continuous but not a homeomorphism. Since the image of a compact set is compact, it follows that $(X, \mathbf{T}^*)$ is compact. If it were Hausdorff, the identity map would be closed and thus a homeomorphism. Therefore $(X, \mathbf{T}^*)$ is not Hausdorff.■

2. (*a*) ($\Longrightarrow$) Suppose that X is $\mathbf{T_3}$, let $x \in X$ and let V be a basic open subset containing x. Then there is an open set U in X such that $x \in U \subset \overline{U} \subset V$. Since $\mathcal{B}$ is a basis for the topology, one can find a basic open subset W in $\mathcal{B}$ such that $x \in W \subset U$, and thus we have $x \in W \subset \overline{W} \subset \overline{U} \subset V$.■

($\Longleftarrow$) Suppose that X is $\mathbf{T_1}$ and satisfies the condition in the exercise. If U is an open subset and $x \in U$, let V be a basic open set from $\mathcal{B}$ such that $x \in V \subset U$, and let W be the basic open set that exists by the hypothesis in the exercise. Then we have $x \in W \subset \overline{W} \subset V \subset U$ and therefore X is regular.■

(*b*) By the first part of the exercise we only need to show this for points in basic open subsets. If the basic open subset comes from the metric topology, this follows because the metric topology is $\mathbf{T_3}$; note that the closure in the new topology might be smaller than the closure in the metric topology, but if a metrically open set contains the metric closure it also contains the "new" closure. If the basic open subset has the form $T_\varepsilon(a)$ for some a and z belongs to this set, there are two cases depending upon whether or not z lies on the x-axis, which we again call A. If $z \notin A$, then z lies in the metrically open subset $T_\varepsilon(a) - A$, and one gets a subneighborhood whose closure lies in the latter exactly as before. On the other hand, if $z \in A$ then $z = (a, 0)$ and the closure of the set $T_\delta(a)$ in either topology is contained in the set $T_\varepsilon(a)$.■

3. (*a*) If X is $\mathbf{T_3}$, then for each point $x \notin A$ there are disjoint open subsets U and V such that $x \in U$ and $A \subset V$. Let $\pi : X \to X/A$ be the quotient projection; we claim that $\pi[U]$ and $\pi[V]$ are open disjoint subsets of X/A. Disjointness follows immediately from the definition of the equivalence relation, and the sets are open because their inverse images are the open sets $U = \pi^{-1}[\,\pi[U]\,]$ and $V = \pi^{-1}[\,\pi[V]\,)$ respectively.■

(b) Suppose that $F \subset X$ is closed; we need to show that $\pi^{-1}[\,\pi(F)\,]$ is also closed. There are two cases depending upon whether or not $A \cap F = \emptyset$. If the two sets are disjoint, then $\pi^{-1}[\,\pi(F)\,] = F$, and if the intersection is nonempty then $\pi^{-1}[\,\pi(F)\,] = F \cup A$. In either case the inverse image is closed, and therefore the image of F is always closed in the quotient space.

We claim that π is not necessarily open if A has a nonempty interior. Suppose that both of the statements in the previous sentence are true for a specific example, and let v be a nonempty open subset of X that is contained in A. If π is open then $\pi[V] = \pi[A] = \{A\} \in X/A$ is an open set. Therefore its inverse image, which is A, must be open in X. But it is also closed in X. Therefore we have the following conclusion: *If A is a nonempty proper closed subset of the connected space X with a nonempty interior, then $\pi : X \to X/A$ is not an open mapping.*■

(c) Let $p : X \to X/A$ and $q : Y \to Y/B$ be the projection maps, and consider the composite $q \circ f$. Then the condition $f[A] \subset B$ implies that $q \circ f$ sends each equivalence class for the relation defining X/A to a point in Y/B, and thus by the basic properties of quotient maps it follows that there is a unique continuous map $F : X/A \to Y/B$ such that $q \circ f = F \circ p$; this is equivalent to the conclusion stated in this part of the problem.■

4. The map $f : \mathbb{R} \to \mathbb{R}$ given by $f(x) = -x$ defines a homeomorphism from $(\mathbb{R}, \mathbf{U}) \to (\mathbb{R}, \mathbf{L})$, so it suffices to consider the case of the lower semicontinuity topology $\mathbf{U}$.

We shall first prove that $(\mathbb{R}, \mathbf{U})$ is $\mathbf{T}_0$. Given two points in $\mathbb{R}$ we can label them as b_1, b_2 such that $b_1 < b_2$. If $a = \frac{1}{2}(b_a + b_2)$, then (a, ∞) is an open subset in $\mathbf{U}$ which contains b_2 but not b_1. — In contrast, note that there is no open subset which contains b_1 but not b_2.■

To see that $(\mathbb{R}, \mathbf{U})$ is not $\mathbf{T}_1$, we need to show that for each $x \in \mathbb{R}$ the set $\mathbb{R} - \{x\}$ is not in $\mathbf{U}$. By definition, every subset in $\mathbf{U}$ is connected in the metric topology, and since $\mathbb{R} - \{x\}$ is not connected in the metric topology it follows that $\mathbb{R} - \{x\}$ is not in $\mathbf{U}$.■

VI.4 : Local compactness and compactifications

Problems from Munkres, § 38, *pp.* 241 – 242

2. We need to begin by describing the compactification mentioned in the exercise. It is given by taking the closure of the embedding (homeomorphism onto its image) $g : (0, 1) \to [-1, 1]^2$ that is inclusion on the first coordinate and $\sin(1/x)$ on the second.

Why is it impossible to extend $\cos(1/x)$ to the closure of the image? Look at the points in the image with coordinates $(1/k\pi, \sin k\pi)$ where k is a positive integer. The second coordinates of these points are always 0, so this sequence converges to the origin. If there is a continuous extension F, it will follow that

$$F(0,0) \quad = \quad \lim_{k\to\infty} \cos\left(\frac{1}{1/k\pi}\right) \quad = \quad \cos k\pi \ .$$

But the terms on the right hand side are equal to $(-1)^k$ and therefore do not have a limit as $k \to \infty$. Therefore no continuous extension to the compactification exists.

One can construct a compactification on which x, $\sin(1/x)$ and $\cos(1/x)$ extend by taking the closure of the image of the embedding $h : (0, 1) \to [-1, 1]^3$ defined by

$$h(x) \quad = \quad (x, \sin(1/x), \cos(1/x)\,) \ .$$

The continuous extensions are given by restricting the projections onto the first, second and third coordinates.■

3. [*Just give a necessary condition on the topology of the space.*]

If A is a dense subset of a compact metric space, then A must be second countable because a compact metric space is second countable and a subspace of a second countable space is also second countable.

This condition is also sufficient, but the sufficiency part was not assigned because it requires the Urysohn Metrization Theorem. The latter says that a $\mathbf{T_3}$ and second countable topological space is homeomorphic to a product of a countably infinite product of copies of $[0,1]$; this space is compact by Tychonoff's Theorem, and another basic result not covered in the course states that a countable product of metrizable spaces is metrizable in the product topology (see Munkres, Exercise 3 on pages 133–134). So if X is metrizable and second countable, the Urysohn Theorem maps it homeomorphically to a subspace of a compact metrizable space, and the closure of its image will be a metrizable compactification of X.■

Additional exercises

Definition. If $f : X \to Y$ is continuous, then f is *proper* (or *perfect*) if for each compact subset $K \subset Y$ the inverse image $f^{-1}[K]$ is a compact subset of X.

1. ($\Longrightarrow$) Suppose that f is proper and U is open in $Y^\bullet$. There are two cases depending upon whether $\infty_Y \in U$. If not, then $U \subset Y$ and thus $[f^\bullet]^{-1}[U] = f^{-1}[U]$ is an open subset of X; since X is open in $X^\bullet$ it follows that $f^{-1}[U]$ is open in $X^\bullet$. On the other hand, if $\infty_Y \in U$ then $Y - U$ is compact, and since f is proper it follows that

$$C \quad = \quad f^{-1}[Y - U] \quad = \quad X - f^{-1}\left[U - \{\infty_Y\}\right]$$

is a compact, hence closed, subset of X and $X^\bullet$. Therefore

$$[f^\bullet]^{-1}[U] \quad = \quad f^{-1}[U] \cup \{\infty_X\} \quad = \quad X^\bullet - C$$

is an open subset of $X^\bullet$.■

($\Longleftarrow$) Suppose that $f^\bullet$ is continuous and $A \subset Y$ is compact. Then

$$f^{-1}[A] = (f^\bullet)^{-1}[A]$$

is a closed, hence compact subset of $X^\bullet$ and likewise it is a compact subset of X.■

2. Let $f : X \to Y$ be a proper map of noncompact locally compact Hausdorff spaces, and let $f^\bullet$ be its continuous extension to a map of one point compactifications. Since the latter are compact Hausdorff it follows that $f^\bullet$ is closed. Suppose now that $F \subset X$ is closed. If F is compact, then so is $f[F]$ and hence the latter is closed in Y. Suppose now that F is not compact, and consider the closure E of F in $X^\bullet$. This set is either F itself or $F \cup \{\infty_X\}$ (since F is its own closure in X it follows that $E \cap X = F$). Since the closed subset $E \subset X^\bullet$ is compact, clearly $E \neq F$, so this implies the second alternative. Once again we can use the fact that $f^\bullet$ is closed to show that $f^\bullet[E] = f[F] \cup \{\infty_Y\}$ is closed in $Y^\bullet$. But the latter equation implies that $f[F] = f^\bullet[F] \cap Y$ is closed in Y.■

3. Write the polynomial as

$$p(z) \quad = \quad \sum_{k=0}^{n} a_k \, z^k$$

where $a_n \neq 0$ and $n > 0$, and rewrite it in the following form:

$$a_n z^n \cdot \left(1 \quad + \quad \sum_{k=1}^{n-1} \frac{a_k}{a_n} \frac{1}{z^{n-k}} \right)$$

The expression inside the parentheses goes to 1 as $n \to \infty$, so we can find $N_0 > 0$ such that $|z| \geq N_0$ implies that the absolute value (or modulus) of this expression is at least $\frac{1}{2}$.

Let $M > 0$ be arbitrary, and define

$$N_1 \quad = \quad \left(\frac{2M_0}{|a_n|} \right)^{1/n} .$$

Then $|z| > \max(N_0, \, N_1)$ implies $|p(z)| > M$. This proves the limit formula.

To see that p is proper, suppose that K is a compact subset of $\mathbf{F}$, and choose $M > 0$ such that $w \in K$ implies $|w| \leq M$. Let N be the maximum of N_1 and N_2, where these are defined as in the preceding paragraph. We then have that the closed set $p^{-1}[K]$ lies in the bounded set of points satisfying $|z| \leq N$, and therefore $p^{-1}[K]$ is a compact subset of $\mathbf{F}$.■

4. As usual, we follow the hints.

> [*Show that*] *the family* $\{ A_n \}$ *is a locally finite family of closed locally compact subspaces in* A.

If $(x, t) \in \ell^2 \times (0, +\infty)$ and U is the open set $\ell^2 \times (t/2, \, +\infty)$, then $x \in U$ and $A_n \cap U = \emptyset$ unless $2^{-n} \geq t/2$, and therefore the family is locally finite. Furthermore, each set is closed in $\ell^2 \times (0, +\infty)$ and therefore also closed in $A = \bigcup_n A_n$.

> *Use this to show that the union is locally compact.*

We shall show that *if A is a $\mathbf{T_3}$ space that is a union of a locally finite family of closed locally compact subsets A_α, then A is locally compact.* Let $x \in A$, and let U be an open subset of A containing x such that $U \cap A_\alpha = \emptyset$ unless $\alpha = \alpha_1 \ \cdots \ , \alpha_k$. Let V_0 be an open subset of A such that $x \in V_0 \subset \overline{V_0} \subset U$, for each i choose an open set $W_i \in A_{\alpha_i}$ such that the closure of W_i is compact, express W_i as an intersection $V_i \cap A_{\alpha_i}$ where V_i is open in A, and finally let $V = \cap_i \, V_i$. Then we have

$$\overline{V} \quad = \quad \bigcup_{i=1}^{k} (\overline{V} \cap A_{\alpha_i}) \quad \subset \quad \bigcup_{i=1}^{k} (\overline{V_i} \cap A_{\alpha_i}) \quad = \quad \bigcup_{i=1}^{k} \textbf{Closure} \, (W_i \ \text{ in } \ A_{\alpha_i}) \; .$$

Since the set on the right hand side is compact, the same is true for $\overline{V}$. Therefore we have shown that A is locally compact.

> *Show that the closure of A contains all of $\ell^2 \times \{0\}$. Explain why ℓ^2 is not locally compact.*

Let $x \in \ell^2$, and for each positive integer k let $P_k(x) \in A_k$ be the point $(H_k(x), 2^{-(k+1)})$, where $H_k(x)$ is the point whose first k coordinates are those of x and whose remaining coordinates are 0. It is an elementary exercise to verify that $(x, 0) = \lim_{k \to \infty} P_k(x)$. To conclude we need to show that ℓ^2 is not locally compact. If it were, then there would be some $\varepsilon > 0$ such that the set of all $y \in \ell^2$ satisfying every $|y| \leq \varepsilon$ would be compact, and consequently infinite sequence $\{ y_n \}$ in ℓ^2

with $|y_n| \leq \varepsilon$ (for all n) would have a convergent subsequence. To see this does not happen, let $y_k = \frac{1}{2}\varepsilon\, \mathbf{e}_k$, where $\mathbf{e}_k$ is the k^{th} standard unit vector in ℓ^2. This sequence satisfies the boundedness condition but does not have a convergent subsequence. Therefore ℓ^2 is not locally compact.■

FOOTNOTE. Basic theorems from functional analysis imply that a normed vector space is locally compact if and only if it is finite-dimensional.■

5. Suppose that V is open in $U^\bullet$. If $\infty_U \notin V$ then $V \subset U$ and $c^{-1}[V] = V$, so the set on the left hand side of the equation is open. Suppose now that $\infty_U \in V$; then $A = U^\bullet - V$ is a compact subset of U and $c^{-1}[V] = X - c^{-1}[A] = X - A$, which is open because the compact set A is also closed in X.

There are many examples for which c is not open. For example, let $X = [0, 5]$ and $A = [1, 3]$; in this example the image J of the open set $(2, 4)$ is not open because the inverse image of J is $[1, 4)$, which is not open. More generally, if X is connected and $X - U$ has a nonempty interior, then $X \to U^\bullet$ is not open (try to prove this!).■

FOOTNOTE. In fact, if $F = X - U$ then $U^\bullet$ is homeomorphic to the space X/F described in a previous exercise. This is true because the collapse map passes to a continuous map from X/F to $U^\bullet$ that is 1–1 onto, and this map is a homeomorphism because X/F is compact and $U^\bullet$ is Hausdorff.■

6. (a) If X is compact Hausdorff and $f : X \to Y$ is a continuous map into a Hausdorff space, then $f[X]$ is closed. Therefore $f[X] = Y$ if the image of f is dense, and in fact f is a homeomorphism.■

(b) This exercise shows that a formal analog of an important concept (the Stone-Čech compactification) is not necessarily as useful as the original concept itself; of course, there are also many situations in mathematics where the exact opposite happens. In any case, given an abstract closure (Y, f) we must have $(X, \mathrm{id}_X) \geq (Y, f)$ because $f : X \to Y$ trivially satisfies the condition $f = f \circ \mathrm{id}_X$.■

7. See the footnote to Exercise 5 above.■

8. Suppose that X is uniformly locally compact as above and $\{x_n\}$ is a Cauchy sequence in X. Choose M such that $m, n \geq M$ implies $\mathbf{d}(x_m, x_n) < \delta$, where δ is as in the problem. If we define a new sequence by $y_k = x_{k-M}$, then $\{y_k\}$ is a Cauchy sequence whose values lie in the (compact) closure of $N_\delta(x_M)$, and since compact metric spaces are complete the sequence $\{y_k\}$ has some limit z. Since $\{y_k\}$ is just $\{x_{k-M}\}$, it follows that the sequence $\{x_{k-M}\}$, and hence also the sequence $\{x_n\}$, must converge to the same limit z.

One of the simplest examples of a noncomplete metric space with the weaker property is the half-open interval $(0, \infty)$. In this case, if we are given x we can take $\delta_x = x/2$.■

9. Let $B = A \cup \{\infty_X\}$, and consider the function $g : A^\bullet \to X$ such that g is the inclusion on A and g sends ∞_A to ∞_X. By construction this map is continuous at all ordinary points of A, and if g is also continuous at ∞_A then g will define a 1–1 continuous map from $A^\bullet$ onto B, and this map will be a homeomorphism.

Suppose now that V is an open neighborhood of ∞_X in $X^\bullet$. By definition we know that $K = X - V$ is compact. But now we have

$$g^{-1}[V] \;=\; A^\bullet - g^{-1}[X - V] \;=\; A^\bullet - A \cap K\ .$$

Now $A \cap K$ is compact because both A and $X - V$ are closed X and K is a compact subset of the Hausdorff space X, and therefore it follows that the set $g^{-1}[V] = A^\bullet - A \cap K$ is open in A, so that g is continuous everywhere.■

10. Follow the hint. The final sentence is equivalent to the desired conclusion because an arbitrary open neighborhood of ∞ has the form $\{\infty\} \cup (\mathbb{C} - K)$, where K is a compact subset of $\mathbb{C}$. Since the compact set K is bounded in $\mathbb{C}$ we can find some $M > 0$ such that all points of K have distance strictly less than M from the origin. Let $\varepsilon = 1/M$, and using the limit assumption for $1/f$ at a let δ be chosen so that $0 < |z - a| < \delta$ implies $|1/f(z)| < \varepsilon$. Then the given conditions on z also imply $|f(z)| > M$, so that $f(z) \notin K$, which is what we needed to show.■

VI.5 : Metrization theorems

Problems from Munkres, § 40, p. 252

2. If W is an $\mathbf{F}_\sigma$ set then $W = \cup_n F_n$ where n ranges over the nonnegative integers and F_n is closed in X. Therefore

$$X - W \quad = \quad \bigcap_n X - F_n$$

is a countable intersection of the open subsets $X - F_n$ and accordingly is a $\mathbf{G}_\delta$ set.

Conversely, if $V = X - W$ is a $\mathbf{G}_\delta$ set, then $V = \cap_n U_n$ where n ranges over the nonnegative integers and U_n is open in X. Therefore

$$W \quad = \quad X - V \quad = \quad \bigcup_n X - U_n$$

is a countable union of the closed subsets $X - U_n$ and accordingly is an $\mathbf{F}_\sigma$ set.■

FOOTNOTE. We have already shown that a closed subset of a metrizable space is a $\mathbf{G}_\delta$ set, and it follows that every open subset of a metrizable space is an $\mathbf{F}_\sigma$ set.■

3. Suppose that X is $\mathbf{T_1}$ and has a locally finite base $\mathcal{B}$. Then for each $x \in X$ there is an open set W_x containing x such that $W_x \cap V_\beta = \emptyset$ for all β except $\beta(1), \cdots, \beta(k)$. Let V^* be the intersection of all sets in the finite subcollection that contain x. Since $\mathcal{B}$ is a base for this topology it follows that some $V_{\beta(J)}$ contains x and is contained in this intersection, But this means that $V_{\beta(J)}$ must be contained in all the other open subsets in $\mathcal{B}$ that contain x and is therefore a *minimal* open subset containing x. Suppose this minimal open set has more than one point; let y be another point in the set. Since X is $\mathbf{T_1}$ it will follow that $V_{\beta(J)} - \{y\}$ is also an open subset containing x. However, this contradicts the minimality of $V_{\beta(J)}$ and shows that the latter consists only of the point $\{x\}$. Since x was arbitrary, this shows that every one point subset of X is open and thus that X is discrete.■

Additional exercises

1. (a) None of the arguments verifying the axioms for open sets in metric spaces rely on the assumption $\mathbf{d}(x, y) = 0 \Longrightarrow x = y$.■

(b) The relation is clearly reflexive and symmetric. To see that it is transitive, note that $\mathbf{d}(x, y) = \mathbf{d}(y, z) = 0$ and the Triangle Inequality imply

$$\mathbf{d}(x, z) \quad \le \quad \mathbf{d}(x, y) \ + \ \mathbf{d}(y, z) \quad = \quad 0 \ + \ 0 \quad = \quad 0\,.$$

Likewise, if $x \sim x'$ and $y \sim y'$ then

$$\mathbf{d}(x', y') \quad \le \quad \mathbf{d}(x', x) \ + \ \mathbf{d}(x, y) \ + \ \mathbf{d}(y, y') \quad = \quad 0 \ + \ \mathbf{d}(x, y) \ + \ 0 \quad = \quad \mathbf{d}(x, y)$$

and therefore the distance between two points only depends upon their equivalence classes with respect to the given relation.■

(c) First of all we verify that $\mathbf{d}_\infty$ defines a metric. In order to do this we must use some basic properties of the function

$$\varphi(x) \quad = \quad \frac{x}{1+x}\ .$$

This is a continuous and strictly increasing function defined on $[0, +\infty)$ and taking values in $[0,1)$ and it has the additional property $\varphi(x+y) \leq \varphi(x)+\varphi(y)$. The continuity and monotonicity properties of φ follow immediately from a computation of its derivative, the statement about its image follows because $x \geq 0$ implies $0 \leq \varphi(x) < 1$ and $\lim_{x\to+\infty}\varphi(x) = 1$ (both calculations are elementary exercises that are left to the reader).

The inequality $\varphi(x+y) \leq \varphi(x)+\varphi(y)$ is established by direct computation of the difference $\varphi(x)+\varphi(y)-\varphi(x+y)$:

$$\frac{x}{1+x} + \frac{y}{1+y} - \frac{x+y}{1+x+y} \quad = \quad \frac{x^2y \ + \ 2\,xy \ + \ xy^2}{(1+x)\,(1+y)\,(1+x+y)}$$

This expression is nonnegative if x and y are nonnegative, and therefore one has the desired inequality for φ. Another elementary but useful inequality is $\varphi(x) \leq x$ if $x \geq 0$ (this is true because $1 \leq 1+x$). Finally, we note that the inverse to the continuous strictly monotonic function φ is given by

$$\varphi^{-1}(y) \quad = \quad \frac{y}{1-y}\ .$$

It follows that if $\mathbf{d}$ is a pseudometric then so is $\varphi\circ\mathbf{d}$ with the additional property that $\varphi\circ\mathbf{d} \leq 1$. More generally, if $\{\,a_n\,\}$ is a convergent sequence of nonnegative real numbers and $\{\,\mathbf{d}_n\,\}$ is a sequence of pseudometrics on a set X, then

$$\mathbf{d}_\infty \quad = \quad \sum_{n=1}^{\infty}\ a_n\cdot\varphi\circ\mathbf{d}_n < \sum_{n=1}^{\infty}\ a_n < \infty$$

also defines a pseudometric on X (write out the details of this!). In our situation $a_n = 2^{-n}$. Therefore the only thing left to prove about $\mathbf{d}_\infty$ is that it is positive when $x \neq y$. But in our situation if $x \neq y$ then there is some n such that $\mathbf{d}_n(x,y) > 0$, and the latter in turn implies that

$$2^{-n}\varphi\big(\,\mathbf{d}_n(x,y)\,\big) \quad > \quad 0$$

and since the latter is one summand in the infinite sum of nonnegative real numbers given by $\mathbf{d}_\infty(x,y)$ it follows that the latter is also positive. *Therefore $\mathbf{d}_\infty$ defines a metric on X.*

To prove that the topology $\mathcal{M}$ defined by this metric and the topology $\mathbf{T}_\infty$ determined by the sequence of pseudometrics are the same. Let N_α denote an α-neighborhood with respect to the $\mathbf{d}_\infty$ metric, and for each n let $N_\beta^{\langle n\rangle}$ denote a β-neighborhood with respect to the pseudometric $\mathbf{d}_n$. Suppose that N_ε is a basic open subset for $\mathcal{M}$ where $\varepsilon > 0$ and $x \in X$. Choose A so large that $n \geq A$ implies

$$\sum_{k=A}^{\infty} 2^{-k} \quad < \quad \frac{\varepsilon}{2}$$

Let W_x be the set of all z such that $\mathbf{d}_k(x,z) < \varepsilon/2$ for $1 \leq k < A$. Then W_x is the finite intersection of the $\mathbf{T}_\infty$-open subsets

$$W^{\langle k\rangle}(x) \quad = \quad \{\ z \in X \ \mid\ \mathbf{d}_k(x,z) < \varepsilon/2\ \}$$

and therefore W_k is also $\mathbf{T}_\infty$-open. Direct computation shows that if $y \in W_k$ then

$$\mathbf{d}_\infty(x,y) \quad = \quad \sum_{n=1}^{\infty} 2^{-n}\,\varphi\,(\,\mathbf{d}_n(x,y)\,) \quad = \quad \sum_{n=1}^{A-1} 2^{-n}\,\varphi\,(\,\mathbf{d}_n(x,y)\,) \;+\; \sum_{n=A}^{\infty} 2^{-n}\,\varphi\,(\,\mathbf{d}_n(x,y)\,) \quad <$$

$$\left(\sum_{n=1}^{A-1} 2^{-n}\,\varphi\,(\,\mathbf{d}_n(x,y)\,)\right) \;+\; \frac{\varepsilon}{2} \quad < \quad \left(\sum_{n=1}^{A-1} \frac{2^{-n}\,\varepsilon}{2}\right) \;+\; \frac{\varepsilon}{2} \quad < \quad \frac{\varepsilon}{2} \;+\; \frac{\varepsilon}{2} \quad < \quad \varepsilon$$

so that $W_x \subset N_\varepsilon(x)$. Therefore, if U is open with respect to $\mathbf{d}_\infty$ we have

$$U \quad = \quad \bigcup_{x\in U} \{x\} \quad \subset \quad \bigcup_{x\in U} W_x \quad \subset \quad \bigcup_{x\in U} N_\varepsilon(x) \quad \subset \quad U$$

which shows that U is a union of $\mathbf{T}_\infty$-open subsets and therefore is $\mathbf{T}_\infty$-open. *Thus $\mathcal{M}$ is contained in $\mathbf{T}_\infty$.*

To show the reverse inclusion, consider an subbasic $\mathbf{T}_\infty$-open subset of the form $N_\varepsilon^{\langle n\rangle}(x)$. If y belongs to the latter then there is a $\delta > 0$ such that $N_\delta^{\langle n\rangle}(y) \subset N_\varepsilon^{\langle n\rangle}(x)$; without loss of generality we may as well assume that $\delta < 2^{-k}$. If we set

$$\eta(y) \quad = \quad 2^{-k}\,\varphi\big(\,\delta(y)\,\big)$$

then $\mathbf{d}_\infty(z,y) < \eta(y)$ and

$$2^{-k}\,\varphi\circ\mathbf{d}_k \quad \leq \quad \mathbf{d}_\infty$$

imply that $2^{-k}\,\varphi\circ\mathbf{d}_k(z,y) < \eta(y)$ so that $\varphi\circ\mathbf{d}_k(z,y) < 2^k\,\eta(y)$ and

$$\mathbf{d}_k(z,y) \quad < \quad \varphi^{-1}\big(\,2^k\,\delta(y)\,\big)$$

and by the definition of $\eta(y)$ the right hand side of this equation is equal to $\delta(y)$. Therefore if we set $W = N_\varepsilon^{\langle k\rangle}(x)$ then we have

$$W \quad = \quad N_\varepsilon^{\langle k\rangle}(x) \quad = \quad \bigcup_{y\in W} \{y\} \quad \subset \quad \bigcup_{y\in W} N_{\eta(y)}(y) \quad \subset \quad \bigcup_{y\in W} N_{\delta(y)}^{\langle k\rangle}(y) \quad \subset \quad N_\varepsilon^{\langle k\rangle}(x)$$

which shows that $N_\varepsilon^{\langle k\rangle}(x)$ belongs to $\mathcal{M}$. *Therefore the topologies $\mathbf{T}_\infty$ and $\mathcal{M}$ are equal.*■

(d) Take the pseudometrics $\mathbf{d}_n$ as in the hint, and given $h \in X$ let $J^{\langle n\rangle}(h)$ be its restriction to $[-n,n]$. Furthermore, let $\|...\|_n$ be the uniform metric on $\mathbf{BC}\,(\,[-n,n]\,)$, so that

$$\mathbf{d}_n(f,g) \quad = \quad \|\,J^{\langle n\rangle}(f) \;-\; J^{\langle n\rangle}(g)\,\|_n\ .$$

Given $\varepsilon > 0$ such that $\varepsilon < 1$, let $\delta = 2^{-n}\,\varphi(\varepsilon)$. Then $\mathbf{d}_\infty(f,g) < \delta$ implies $2^{-n}\,\varphi\circ\mathbf{d}_n(f,g) < \delta$, which in turn implies $\mathbf{d}_n(f,g) < \varphi^{-1}(2^n\,\delta) = \varepsilon$.■

(e) We first claim that $\lim_{n\to\infty} \mathbf{d}_\infty(f_n,f) = 0$ if and only if $\lim_{n\to\infty} \mathbf{d}_k(f_n,f) = 0$ for all k.

$(\implies)$ Let $\varepsilon > 0$ and fix k. Since $\lim_{n\to\infty} \mathbf{d}_\infty(f_n,f) = 0$ there is a positive integer M such that $n \geq M$ implies $\mathbf{d}_\infty(f_n,f) < 2^k\,\varphi(\varepsilon)$. Since

$$2^{-k}\varphi\circ\mathbf{d}_k \quad \leq \quad \mathbf{d}_\infty$$

it follows that $\mathbf{d}_k \leq \varphi^{-1}(2^{-k}\,\mathbf{d}_\infty)$ and hence that $\mathbf{d}_k(f_n, f) < \varepsilon$ if $n \geq M$.■

($\Longleftarrow$) Let $\varepsilon > 0$ and choose A such that

$$\sum_{k=n}^{\infty} 2^{-k} \quad < \quad \frac{\varepsilon}{2}\,.$$

Now choose B so that $n \geq B$ implies that

$$\mathbf{d}_k(f_n, f) \quad < \quad \frac{\varepsilon}{2A}$$

for all $k < A$. Then if $n \geq A + B$ we have

$$\mathbf{d}_\infty(f_n, f) \quad = \quad \sum_{k=1}^{\infty} 2^{-k}\,\varphi^\circ\mathbf{d}_k(f_n, f) \quad = \quad \sum_{k=1}^{A} 2^{-k}\,\varphi^\circ\mathbf{d}_k(f_n, f) \;+\; \sum_{k=A}^{\infty} 2^{-k}\,\varphi^\circ\mathbf{d}_k(f_n, f) \quad <$$

$$\sum_{k=1}^{A} 2^{-k}\,\varphi^\circ\mathbf{d}_k(f_n, f) \;+\; \frac{\varepsilon}{2} \quad < \quad \left(\sum_{k=1}^{A} 2^{-k}\frac{\varepsilon}{2A}\right) \;+\; \frac{\varepsilon}{2} \quad = \quad \varepsilon$$

so that $\lim_{n\to\infty} \mathbf{d}_\infty(f_n, f) = 0$.■

Now suppose that K is a compact subset of $\mathbb{R}$, and let $\|...\|_K$ be the uniform norm on $\mathbf{BC}(K)$. Then $K \subset [-n, n]$ for some n and thus for all $g \in X$ we have $\|g\|_K \leq \mathbf{d}_n(g, 0)$. Therefore if $\{ f_n \}$ converges to f then the sequence of restricted functions $\{ f_n|K \}$ converges uniformly to $f|K$. Conversely, if for each compact subset $K \subset \mathbb{R}$ the sequence of restricted functions $\{ f_n|K \}$ converges to $f|K$, then this is true in particular for $K = [-L, L]$ and accordingly $\lim_{n\to\infty} \mathbf{d}_L(f_n, f) = 0$ for all L. However, as noted above this implies that $\lim_{n\to\infty} \mathbf{d}_\infty(f_n, f) = 0$ and hence that $\{ f_n \}$ converges to f.■

(f) The answer is **YES**, and here is a proof: Let $\{ f_n \}$ be a Cauchy sequence in X. Then if $K_m = [-m, m]$ the sequence of restricted functions $\{ f_n|K_m \}$ is a Cauchy sequence in $\mathbf{BC}(K_m)$ and therefore converges to a limit function $g_m \in \mathbf{BC}(K_m)$. Since $\lim_{n\to\infty} f_n|K_m(x) = g_m(x)$ for all $x \in K_m$ it follows that $p \leq m$ implies $g_m|K_p = g_p$ for all such m and p. Therefore if we define $g(x) = g_m(x)$ if $|x| < m$ then the definition does not depend upon the choice of m and the continuity of g_m for each m implies the continuity of g. Furthermore, by construction it follows that

$$\lim_{n\to\infty} \mathbf{d}_m(f_n, g) \quad = \quad \lim_{n\to\infty} \|\,(f_n|K_m) \;-\; g_m\,\| \quad = \quad 0$$

for all m and hence that $\lim_{n\to\infty} \mathbf{d}_\infty(f_n, f) = 0$ by part **(e)** above.■

(g) The key idea is to express U as an increasing union of bounded open subsets V_n such that $\overline{V_n} \subset V_{n+1}$ for all n. If U is a proper open subset of $\mathbb{R}^n$ let $F = \mathbb{R}^n - U$ (hence F is closed), and let V_m be the set of all points x such that $|x| < m$ and $\mathbf{d}_{[\mathbf{2}]}(x, F_m) > 1/m$, where $\mathbf{d}_{[\mathbf{2}]}$ denotes the usual Euclidean metric; if $U = \mathbb{R}^n$ let V_m be the set of all points x such that $|x| < m$. Since $y \in U$ if and only if $\mathbf{d}_{[\mathbf{2}]}(y, F) = 0$, it follows that $U = \cup_m V_m$. Furthermore, since $y \in \overline{V_m}$ implies $|x| \leq m$ and $\mathbf{d}_{[\mathbf{2}]}(x, F_n) \geq 1/n$ (why?), we have $\overline{V_m} \subset V_{m+1}$. Since $\overline{V_m}$ is bounded it is compact.

If f and g are continuous real valued functions on U, define $\mathbf{d}_n(f, g)$ to be the maximum value of $|f(x) - g(x)|$ on $\overline{V_m}$. In this setting the conclusions of parts **(d)** through **(f)** go through with only one significant modification; namely, one needs to check that every compact subset of U is contained in some V_m. To see this, note that K is a compact subset that is disjoint from the closed subset F, and therefore the continuous function $\mathbf{d}_{[\mathbf{2}]}(y, F)$ assumes a positive minimum value c_1 on K and that there is a positive constant c_2 such that $y \in K$ implies $|y| \le c_2$. If we choose m such that $m > 1/c_1$ and $m > c_2$, then K will be contained in V_m as required.■

FOOTNOTE. Here is another situation where one encounters metrics defined by an infinite sequence of pseudometrics. Let Y be the set of all infinitely differentiable functions on $[0, 1]$, let D^k denote the operation of taking the k^{th} derivative, and let $\mathbf{d}_k(f, g)$ be the maximum value of $|D^k f \ - \ D^k g|$, where $0 \le k < \infty$. One can also mix this sort of example with the one studied in the exercise; for instance, one can consider the topology on the set of infinitely differentiable functions on $\mathbb{R}$ defined by the countable family of pseudometrics $\mathbf{d}_{k,n}$ where $\mathbf{d}_{k,n}(f, g)$ is the maximum of $|D^k f \ - \ D^k g|$ on the closed interval $[-n, n]$.■

2. (a) We shall follow the steps indicated in the hint(s).

Note that

$$(\mathbb{R}^n)^\bullet \ - \ \{0\} \quad \cong \quad \mathbb{R}^n$$

This is true because the left hand side is homeomorphic to the complement of a point in S^n, and such a complement is homeomorphic to $\mathbb{R}^n$ via stereographic projection (which may be taken with respect to an arbitrary unit vector on the sphere).■

Consider the continuous function on

$$\left(\overline{V} - V\right) \sqcup \{\infty\} \quad \subset \quad (\mathbb{R}^n)^\bullet \ - \ \{0\}$$

defined on the respective pieces by the restriction of f and ∞. Why can this be extended to a continuous function on $(\mathbb{R}^n)^\bullet \ - \ V$ with the same codomain?

In the first step we noted that the codomain was homeomorphic to $\mathbb{R}^n$, and the Tietze Extension Theorem implies that a continuous function from a closed subset A of a metric space X into $\mathbb{R}^n$ extends to all of X.■

What happens if we try to piece this together with the original function f defined on U?

If h is the function defined above, then we can piece h and $f|\overline{V}$ together and obtain a continuous function on all of $(\mathbb{R}^n)^\bullet$ if and only if the given functions agree on the intersection of the two closed subsets. This intersection is equal to $\overline{V} - V$, and by construction the restriction of h to this subset is equal to the restriction of f to this subset.■

(b) Let h and h' denote the associated maps from $(\mathbb{R}^n)^\bullet \ - \ V$ to $(\mathbb{R}^n)^\bullet \ - \ \{0\} \cong \mathbb{R}^n$ that are essentially given by the restrictions of g and g'. Each of these maps has the same restriction to $\left(\overline{V} - V\right) \sqcup \{\infty\}$. Define a continuous mapping $H : ((\mathbb{R}^n)^\bullet \ - \ V) \times [0, 1] \longrightarrow (\mathbb{R}^n)^\bullet \ - \ \{0\} \cong \mathbb{R}^n$ by $H(x, t) = t\, h'(x) + (1 - t)\, h(x)$, and define F' on $\overline{V} \times [0, 1]$ by $F(x, t) = f(x)$. As in the first part of this exercise, the mappings H and F' agree on the intersection of their domains and therefore they define a continuous map G on all of $S^n \times [0, 1]$. Verification that $G(x, 0) = g(x)$ and $G(x, 1) = g'(x)$ is an elementary exercise.■

3. (a) Let U be open in X, and let $F = X - U$. Then by the examples cited above we know that F is a closed set which is also a G_δ set. Apply Exercise 2 on page 252 to show that $U = X - F$ must be an open F_σ set.■

(b) If A is the zero set of a continuous function as above, then

$$A \quad = \quad \cap_n \ge 1 \ f^{-1}\left[\,[0, 1/n)\,\right]$$

where each half open interval $[0, n)$ is open in $[0, 1]$, so that their inverse images are open in X. Conversely, if A is a closed G_δ of the form $\cap_n U_n$, then if we set $V_n = U_1 \cap \cdots \cap U_n$, we have $A = \cap_n V_n$ and $V_1 \supset V_2 \supset \cdots$. Define functgions $g_n : X \to [0, 1]$ such that $g_n = 0$ on A and $g_n = 1$ on $F_n = X - V_n$. If we take $f = \sum_n 2^{-n} \cdot g_n$, then the infinite series converges absolutely and uniformly to a continuous function, and the zero set of this function is precisely A.■

(c) Let $A = X - W$, and let f be as in the previous part of this exercise. Then the set of zero points is A, so the set of points where f is nonzero (and in fact positive by construction) will be the set $W = X - A$.■

4. For each x let φ_x be a continuous real valued function in $\mathcal{F}$ such that $\varphi_x = 0$ on some open set U_x containing x. The family of open sets $\{ U_x \}$ is an open covering of X and therefore has a finite subcovering by sets $U_{x(i)}$ for $1 \leq i \leq k$. The product of the functions $\prod_i f_{x(i)}$ belongs to $\mathcal{F}$ and is zero on each $U_{x(i)}$; since the latter sets cover X it follows that the product is the zero function and therefore the latter belongs to $\mathcal{F}$.■

5. REMINDER. If X is a topological space and $\mathcal{U} = \{U_1, \cdots, U_n\}$ is a finite indexed open covering of X, then a partition of unity subordinate to $\mathcal{U}$ is an indexed family of continuous functions $\varphi_i : X \to [0, 1]$ for $1 \leq i \leq n$ such that for each i the zero set of the function φ_i contains contains $X - U_i$ in its interior and

$$\sum_{i=1}^{n} \varphi_i \quad = \quad 1 \ .$$

Theorem 36.1 on pages 225–226 of Munkres states that for each finite indexed open covering $\mathcal{U}$ of a $\mathbf{T_4}$ space (hence for each such covering of a compact Hausdorff space), there is a partition of unity subordinate to $\mathcal{U}$. The proof of this is based upon Urysohn's Lemma, so the methods in Munkres can be combined with our proof of the result for metric space to prove the existence of partitions of unity for indexed finite open coverings of compact metric spaces.

The solution to the exercise now proceeds as follows:

For each $x \in X$ we are assuming the existence of a continuous bounded function f_x such that $f_x(x) \neq 0$. Since J is closed under multiplication, we may replace this function by it square if necessary to obtain a function $g_x \in J$ such that $g_x(x) > 0$. Let U_x be the open set where g_x is nonzero, and choose an open subset V_x such that $\overline{V_x} \subset U_x$. The sets V_x determine an open covering of X; this open covering has a finite subcovering that we index and write as $\mathcal{V} = \{V_1, \cdots, V_n\}$. For each i let $g_i \in J$ be the previously chosen function that is positive on $\overline{V_i}$, and consider the function

$$g \quad = \quad \sum_{i=0}^{k} \varphi_i \cdot g_i \ .$$

This function belongs to J, and we claim that $g(y) > 0$ for all $y \in X$. Since $\sum_i \varphi_i = 1$ there is some index value m such that $\varphi_m(y) > 0$; by definition of a partition of unity this means that $y \in V_m$. But $y \in V_m$ implies $g_m(y) > 0$ too and therefore we have $g(y) \geq \varphi_m(y)\, g_m(y) > 0$. Since the reciprocal of a nowhere vanishing continuous real valued function on a compact space is continuous (and bounded!), we know that $1/g$ lies in $\mathbf{BC}(X)$. Since $g \in J$ it follows that $1 = g \cdot (1/g)$ also lies in J, and this in turn implies that $h = h \cdot 1$ lies in J for all $h \in \mathbf{BC}(X)$. Therefore $J = \mathbf{BC}(X)$ as claimed.■

6. Let $\mathcal{M}$ be the set of all maximal ideals. For each point $x \in X$ we need to show that $\mathbf{M}_x \in \mathcal{M}$. First of all, verification that $\mathbf{M}_x$ is an ideal is a sequence of elementary computations (which the reader should verify). To see that the ideal is maximal, consider the function $\widehat{x} : \mathbf{BC}(X) \to \mathbb{R}$ by the formula $\widehat{x}(f) = f(x)$. This mapping is a ring homomorphism, it is onto, and $\widehat{x}(f) = 0$ if and only if $f \in \mathbf{M}_x$. Suppose that the ideal is not maximal, and let $\mathbf{A}$ be an ideal such that $\mathbf{M}_x$ is properly contained in $\mathbf{A}$. Let $a \in A$ be an element that is not in $\mathbf{M}_x$. Then $a(x) = \alpha \neq 0$ and it follows that $a(x) - \alpha 1$ lies in $\mathbf{M}_x$. It follows that $\alpha 1 \in \mathbf{A}$, and since $\mathbf{A}$ is an ideal we also have that $1 = \alpha^{-1}(\alpha 1)$ lies in $\mathbf{A}$; the latter in turn implies that every element $f = f \cdot 1$ of $\mathbf{BC}(X)$ lies in $\mathbf{A}$.

We claim that the map from X to $\mathcal{M}$ sending x to $\mathbf{M}_x$ is 1–1 and onto. Given distinct points x and y there is a bounded continuous function f such that $f(x) = 0$ and $f(y) = 1$, and therefore it follows that $\mathbf{M}_x \neq \mathbf{M}_y$. To see that the map is onto, let $\mathbf{M}$ be a maximal ideal, and note that the preceding exercise implies the existence of some point $p \in X$ such that $f(p) = 0$ for all $f \in \mathbf{M}$. This immediately implies that $\mathbf{M} \subset \mathbf{M}_p$, and since both are maximal (proper) ideals it follows that they must be equal. Therefore the map from X to $\mathcal{M}$ is a 1–1 correspondence.■

FINAL FOOTNOTES. **(1)** One can use techniques from functional analysis and Tychonoff's Theorem to put a natural topology on $\mathcal{M}$ (depending only on the structure of $\mathbf{BC}(X)$ as a Banach space and an algebra over the reals) such that the correspondence above is a homeomorphism; see page 283 of Rudin, *Functional Analysis*, for more information about this.

(2) The preceding results are the first steps in the proof of an important result due to I. Gelfand and M. Naimark that give a complete set of abstract conditions under which a Banach algebra is isomorphic to the algebra of continuous complex valued functions on a compact Hausdorff space. A Banach algebra is a combination of Banach space and associative algebra (over the real or complex numbers) such that the multiplication and norm satisfy the compatibility relation $|xy| \leq |x| \cdot |y|$. The additional conditions required to prove that a Banach algebra over the complex numbers is isomorphic to the complex version of $\mathbf{BC}(X)$ are commutativity, the existence of a unit element, and the existence of an conjugation-like map (formally, an **involution**) $a \to a^*$ satisfying the additional condition $|aa^*| = |a|^2$. Details appear in Rudin's book on functional analysis, and a reference for the Gelfand-Naimark Theorem is Theorem 11.18 on page 289. A classic reference on Banach algebras (definitely NOT up-to-date in terms of current knowledge but an excellent source for the material it covers) is the book by Rickart in the bibliographic section of the course notes.

Solution to Munkres, Exercise 28.2
(for finite products only)

By induction it suffices to prove this for twofold products. Suppose X and Y are locally compact Hausdorff. Then previous results imply $X \times Y$ is Hausdorff.

Let $(x, y) \in X \times Y$, and let $W \subseteq X \times Y$ be open such that $(x, y) \in W$. Take open subsets $U_0 \subseteq X$, $V_0 \subseteq Y$ such that $(x, y) \in U_0 \times V_0 \subseteq W$ (we can do this because $X \times Y$ has the product topology). Since X and Y are locally compact there are open sets $U \subseteq X$, $V \subseteq Y$ such that $x \in U \subseteq \overline{U} \subseteq U_0$ and $y \in V \subseteq \overline{V} \subseteq V_0$

$\overline{U_0}$, $\overline{V_0}$ are compact. Since $\overline{U \times V} = \overline{U} \times \overline{V}$ and the product of compact spaces is compact, it follows that

$$(x, y) \in U \times V \subseteq \overline{U} \times \overline{V} = \overline{U \times V} \subseteq U_0 \times V_0 \subseteq W$$

where $\overline{U \times V}$ is compact. ∎

SOLUTIONS TO EXERCISES FOR

MATHEMATICS 205A — Part 6

Fall 2014

APPENDICES

Appendix A : Topological groups

These exercises are taken from various sections in Munkres. Some of these solutions are based upon material in the site `http://dbfin.com/topics/math/`.

Supplementary Exercises from Munkres, pp. 145 – 146

5. (a) Let $\pi : G \to G/H$ be the quotient projection. If $L_a : G \to G$ is the left multiplication homeomorphism for $a \in G$, then by associativity of multiplication we know that $\pi(x) = \pi(y)$ (equivalently, $xH = yH$) implies that $\pi \circ L_a(x) = axH = ayH = \pi \circ L_a(y)$, and therefore by the properties of quotient topologies there is a unique map $L'_a : G/H \to G/H$ such that $L'_a \circ \pi = \pi \circ \mathrm{L}_a$. If $a = 1$, so that L_a is the identity, then the unique mapping L'_a must also be the identity; furthermore, the uniqueness properties and $L_{ba} = L_b \circ L_a$ imply that $L'_{ba} = L'_b \circ L'_a$. If we now let $\{a, b\} = \{c, c^{-1}\}$ for some $c \in G$, the conclusions of the previous sentence imply that L'_c is a homeomorphism and $(L'_c)^{-1} = L'_{c^{-1}}$.

Suppose now that we have cosets xH and yH. If $a = yx^{-1}$, then by construction L'_a sends xH to yH, and therefore G/H is homogeneous.■

WARNING. In the overwhelming majority of mathematical writings over the past several decades, the phrase *homogeneous space* is reserved for the sorts of coset spaces we have described here. Normaly mathematicians would use the phrase “space which is homogeneous” for the spaces described as “homogeneous spaces” in Munkres.

(b) If H is a closed subgroup, then by the definition of the quotient topology the subset $\{H\} \subset G/H$ is closed (more generally, a one point subset in a quotient space is closed if and only if its equivalence class in the original space is closed).

More generally, if Y is a topological space such that $\{y\}$ is closed for some $y \in Y$, then for every homeomorphism $h : Y \to Y$ we know that $h[\{y\}] = \{h(y)\}$ is also a closed subset. Specializing to G/H, if $xH \in G/H$ then we know that there is a homeomorphism from G/H to itself sending H to xH by (a). If H is closed in G, then we know that $\{H\} \subset G/H$ is closed, and therefore it follows that $\{xH\}$ is also closed.■

(c) Suppose that U is open in G. Then $\pi[U]$ is the union of all cosets uH where u runs through the elements of U, and we also have $\pi^{-1}[\pi[U]] = \cup_{h \in H}\, U \cdot h$. Since right multiplication by h is a homeomorphism and U is open in G, it follows that each $U \cdot h$ is also open in G and hence their union is also open in G. By the definition of the quotient topology, this means that $\pi[U]$ is open in G/H.■

(d) We shall modify this part as follows: If H is a normal subgroup of G, then G/H is a topological group in the sense of the course notes (no assumption that $\{1\}$ is a closed subset), and if H is closed then this topological group is a $\mathbf{T}_1$ space (and hence a topological group in the sense of Munkres).

If G is a topological group and H is a normal subgroup then by standard results on groups we know that the group structure on G passes to a group structure on G/H. In terms of commutative diagrams, if m denotes the multiplication on G and χ denotes the inverse map on G, then these and the corresponding maps m' and χ' for G/H satisfy the following identities:

$$\pi \circ m \quad = \quad m' \circ (\pi \times \pi) \; , \qquad\qquad \pi \circ \chi \quad = \quad \chi' \circ \pi$$

We need to show that m' and χ' are continuous with respect to the quotient topology on G/H. The continuity of χ' follows immediately from the identity because π is a quotient map, but as noted in the next paragraph, the continuity of m' requires some further discussion.

If we impose the quotient topology on $G/H \times G/H$ associated to the continuous mapping $\pi \times \pi$, then general considerations imply that m' is continuous. However, the topology we really want on $G/H \times G/H$ is the self-product of the quotient topology for π. We need to verify that these two topologies are equal. If they are, then it will follow that the quotient group structure on G/H makes the latter (with the quotient topology) into a topological group.

We can formulate the issue more abstractly as follows: Suppose that we are given a topological space $(X, \mathbf{T}_X)$ and an onto map of sets $p : X \to Y$. Denote the quotient topology on Y by $p_*\mathbf{T}_X$. The product map $p \times p : X \times X \to Y \times Y$ is also onto, and thus we also have the quotient topology $(p \times p)_*\mathbf{T}_{X\times X}$ on $Y \times Y$. We need to show that, at least in some cases, the spaces $(Y \times Y, (p \times p)_*\mathbf{T}_{X\times X}$ and $(Y, p_*\mathbf{T}_X) \times (Y, p_*\mathbf{T}_X)$ are identical.

> CLAIM. *The statement in the preceding sentence is valid if there is a topology $\mathbf{U}$ on Y such that $p : (X, \mathbf{T}_X) \to (Y, \mathbf{U})$ is open.* — The openness hypothesis implies that $\mathbf{U} = p_*\mathbf{T}_X$, and the product map $p \times p : (X, \mathbf{T}_X) \times (X, \mathbf{T}_X) \to (Y, \mathbf{U}) \times (Y, \mathbf{U})$ is also open (since $p \times p$ sends basic open sets to basic open sets). Therefore $p \times p$ is a quotient map, and this implies that the spaces $(Y \times Y, (p \times p)_*\mathbf{T}_{X\times X}$ and $(Y, p_*\mathbf{T}_X) \times (Y, p_*\mathbf{T}_X)$ are identical.

Finally, we note that the preceding reasoning applies to the quotient group projection $\pi : G \to G/H$ because π is an open mapping.■

6. We claim that the quotient group is the multiplicative group $S^1 \subset \mathbb{C}$ of complex numbers such that $|z| = 1$. Consider the map $p : \mathbb{R} \to S^1$ defined by $p(t) = (\cos 2\pi\, t, \sin 2\pi\, t)$. Standard trigonometric identities imply that $p(t_1 + t_2) = p(t_1) \cdot p(t_2)$ for all $t_1, t_2 \in \mathbb{R}$. This map is continuous and onto, and the kernel is $\mathbb{Z}$. Therefore p passes to a continuous and 1–1 onto homomorphism $p^* : \mathbb{R}/\mathbb{Z} \to S^1$. Furthermore, the quotient group is compact because the restriction of p to the compact subset $[0, 1]$ is onto. Therefore Theorem III.1.9 implies that p^* is a homeomorphism, and this completes the proof that p^* is an isomorphism of topological groups.■

7. (a) If U is an open neighborhood of 1, then by continuity of multiplication there are open neighborhoods W_1 and W_2 of 1 such that $W_1 \cdot W_2 \subset U$. If

$$V \quad = \quad (W_1 \cap W_2) \;\cap\; (W_1 \cap W_2)^{-1}$$

then the identity $g = \left(g^{-1}\right)^{-1}$ implies that $V = V^{-1}$, and we also have $V \cdot V \subset W_1 \cdot W_2 \subset U$.■

(b) [Recall that Munkres' definition implies that G is $\mathbf{T}_1$.]

Let $U = G - \{y^{-1}x\}$, so that U is an open neighborhood of 1, and take V as in (a). We claim that $x \cdot V$ and $y \cdot V$ are disjoint. Assume that they do have some point z in common. Then we have $z = xv_1 = yv_2$ for some $v_1, v_2 \in V$, and thus we also have $y^{-1}x = v_2 \cdot v_1^{-1}$. Since V is symmetric and $V \cdot V \subset U$, it follows that $y^{-1}x \in U$, contradicting our choice of U. The source of the contradiction was our assumption regarding the existence of z, and therefore no such point can exist; in other words, we must have $(x \cdot V(\cap(y \cdot V) = \emptyset$.■

(c) Let $U = G - (x^{-1} \cdot A)$, so that $1 \in U$ and U is disjoint from the closed set $B = x^{-1} \cdot A$. Once again, take V as in (a). We claim that the open subsets $x \cdot V$ and $A \cdot V$ are disjoint. If not, then we have $xv_1 = av_2$ for some $V - 1, v_2 \in V$ and $a \in A$. We can rewrite the equation in the form $x^{-1}a = v_2 \cdot v_1^{-1}$, and as in (b) we have $x^{-1}a \in U$, which contradicts our choice of V. The source of the contradiction was the assumption that $x \cdot V$ and $A \cdot V$ had a point in common, and therefore these sets must be disjoint.■

(d) Follow the hint. Suppose that E is closed in G/H and $\pi(x) \notin E$. Taking inverse images, we see that $xH = \pi^{-1}[\pi[x]]$ is disjoint from the set $F = \pi^{-1}[E]$, which is closed and has the property that $F \times H = F$ (*i.e.*, F is right H-invariant). The reasoning of (c) implies that there is a symmetric neighborhood V of 1 such that $V \cdot x$ and $V \cdot F$ are disjoint, and the same also holds for $V \cdot xH$ and $V \cdot F$ because $F \cdot H = H$. Since V is open it follows that $V \cdot C$ is open for all $C \subset G$, and therefore $V \cdot xH$ and $V \cdot F = V \cdot F \cdot H$ are also open.

Consider the open sets $\pi[V \cdot xH]$ and $\pi[V \cdot F]$ in G/H. The first of these contains $\pi(x)$, and the second contains F. Since the respective inverse images of these open sets are the disjoint subsets $V \cdot xH$ and $V \cdot F = V \cdot F \cdot H$, it follows that $\pi[V \cdot xH]$ and $\pi[V \cdot F]$ are disjoint open subsets containing $\pi(x)$ and F respectively. Since each of the latter was arbitrary, this implies that G/H is regular.■

Problems from Munkres, § 26, *pp.* 170 – 172

12. See Theorem 1 in `proper.pdf`; the first half of the proof shows that, more generally, if a closed continuous (not necessarily onto) map satisfies the hypotheses, then inverse images of compact subsets are compact; although the statement of the theorem is restricted to Hausdorff spaces, the proof does not use this assumption on X and Y.■

13. [The definition of topological group in Munkres includes an assumption that one point subsets are closed. Although this additional hypothesis is not universally adopted, we need it in order to solve the exercise. By Exercise 22.7 in Munkres, the assumption implies that G is also Hausdorff and even satisfies the $\mathbf{T}_3$ separation axiom.]

(a) We shall follow the hint, so suppose that $c \notin A \cdot B$ where $A \subset G$ is closed and $B \subset G$ is compact. The for all $b \in B$ we have $c \notin A \cdot b$, and therefore there are two disjoint open sets U_b and V_b such that $c \in U_b$ and $A \cdot b \subset V_b$.

We claim there is an open neighborhood W_b of b such that $A \cdot b \subset W_b$. — By continuity of multiplication, for each $a \in A$ there is an open neighborhood N_a of b such that $ab \in a \cdot N_a \subset V_b$. If W_b is the union of all the neighborhoods N_a, then $A \cdot W_b \subset V_b$. The open sets W_b define an open covering of B, so by compactness there is a finite subcovering $W_{b_1}, \cdots, W_{b_k}$. If $U_c = U_{b_1} \cap \cdots \cap U_{b_k}$, then $c \in U$ and $U \cap A \cdot B = \emptyset$. Therefore $G - (A \cdot B)$ contains an open neighborhood of c for each $c \notin A \cdot B$, and consequently $G - (A \cdot B)$ is open, or equivalently $A \cdot B$ is closed in G.■

(b) Suppose that $F \subset G$ is closed. Then by (a) we know that $E \cdot H$ is also closed. Since p is onto, we have

$$p^{-1}[p[F]] \quad = \quad F \cdot H$$

and since p is a quotient map this implies that $p[F]$ is closed in G/H.■

(c) The projection $p : G \to G/H$ is closed by (b), and the inverse image of each point $[g] \in G/H$ is the coset gH, which is homeomorphic to the compact subgroup H. Therefore by Exercise 26.12 (proved above) we know that G is compact.■

Problems from Munkres, § 30, pp. 194 – 195

18. We shall follow the hint. Let $\mathcal{B} == \{B_n\}$ be a countable neighborhood base at $1 \in G$. Taking intersections $B_n \cap B_n^{-1}$ if necessary, we might as well assume that the open sets B_n are symmetric (with respect to taking inverses). Furthermore, by passing to a subsequence if necessary, we might as well assume that $B_{n+1} \cdot B_{n+1} \subset B_n$ (this uses the first part of Exercise 22.7 and the fact that we have a neighborhood base; taken together, these imply that there is some $k > n$ such that $B_k \cdot B_k \subset B_n$).

The basic idea is to imitate a proof that the additive groups $\mathbb{R}^n$ are second countable under the given hypotheses. The countable neighborhood base should serve as a weak substitute for the $\frac{1}{2^n}$-neighborhoods centered at points. — This underlying principle turns out to be very useful in working with topological groups; for example, one can use neighborhood bases at the identity to formulate a strong analog of uniform continuity for continuous maps of topological groups.

Suppose first that G has a countable dense subset D, and let U be an open subset of G. For each $x \in U$ the set $x^{-1} \cdot U$ is an open neighborhood of 1 and hence there is some open set $B_n \in \mathcal{B}$ such that $B_n \subset x^{-1} \cdot U$. The latter implies $x \cdot B_n \subset U$. Let $d \in D$ be such that $d \in x \cdot B_{n+1}$; then $x^{-1}d \in B_{n+1}$, and since B_n is symmetric we also have $d^{-1}x \in B_{n+1}$, which in turn implies that $x \in d \cdot B_{n+1}$. We now have

$$d \cdot B_{n+1} \quad = \quad x \cdot x^{-1} d \cdot B_{n+1} \quad \subset \quad x \cdot B_{n+1} \cdot B_{n+1} \subset x \cdot B_n$$

which means that U is a union of open sets in the countable family $\{\ d \cdot B_k \mid d \in D, k \in \mathbb{N}^*\ \}$, and therefore the latter is a countable base for G.

We shall now assume that G is Lindelöf. For each $B_k \in \mathcal{B}$, consider the open covering $\{\ g \cdot B_k \mid g \in G\ \}$ It has a finite subcovering, so choose $g_{k,j}$ such that the sets $g_{k,j} \cdot B_k$ form a countable subcovering, and let D be the set of points $g_{k,j}$ as k and j both run through all positive integers. This is a countable family, and we claim it is dense in G. As before, let U be an open subset of G and let $x \in G$, so that there is an open neighborhood of x having the form $x \cdot B_n \subset U$ for some n. By construction there is an open set $d \cdot B_{n+1}$ containing x, and as in the preceding discussion we have $d \in d \cdot B_{n+1} \subset x \cdot B_n$. Since the latter is contained in U, it follows that $d \in U$, which implies that D is dense in G. We can now apply the reasoning of the first part of the problem to conclude that G is second countable.■

Problems from Munkres, § 31, pp. 199 – 200

8. Let $p : X \to X/G$ denote the orbit space projection map (which is continuous since we are taking the quotient topology on X/G.)

There are several parts to this exercise, with one for each of the listed properties. It seems best to begin with some preliminary observations as in Section II.3 of the book by Bredon which is cited below.

CLAIM 1. *The orbit space projection $X \to X/G$ is an open mapping, with no compactness hypothesis on G.* — If U is open in X, then $G \cdot U = \cup_g \ g \cdot U$ is also open, and thus $p^{-1}[p[U]] = G \cdot U$ shows that $p[U]$ is open in X/G.

CLAIM 2. *If G is compact, then the orbit space projection $X \to X/G$ is closed.* — This follows by the same sort of reasoning used in Exercise 26.13(i). Let $F \subset X$ be closed. If we can show that $G \cdot F$ is closed when G is compact, then $p^{-1}[p[F]] = G \cdot F$ will show that $p[F]$ is closed in X/G. To show that $G \times F$ is closed, suppose that $x \notin G \cdot F$. Then for all $g \in G$ we have $x \notin G \cdot F$. For each $g \in G$ there is an open neighborhood V_g of $g \cdot x$ such that $V_g \cap A = \emptyset$ By continuity there is a neighborhood U_g of g such that $U_g \cdot V_g \subset X - A$. If $\mathcal{U}$ is the open covering of G by the sets U_g, then by compactness there is a finite subcovering $U_{g(1)}, \ \cdots \ , U_{g(k)}$. The corresponding open subsets $V_{g(j)}$ intersect in some open neighborhood V_x of x, and we have $G \cdot V_x \subset U_j \cdot V_j \subset X - A$. The latter implies that V_x is an open neighborhood of x which is contained in $X - G \cdot A$. Since $x \in X - G \cdot A$ was arbitrary, it follows that $X - G \cdot A$ is open in X and hence $G \cdot A$ is closed in X.■

CLAIM 3. *If G is compact, then the orbit space projection $X \to X/G$ is proper: Inverse images of compact subsets are compact.* — This follows from Exercise 1 in `proper.pdf` or equivalently from Exercise 26.12.

With these at our disposal, we can prove the statements in the exercise fairly directly.

Hausdorff. Suppose $[x] \neq [y]$ in X/G. If we lift everything back to X, this translates into a statement that the orbits $G \cdot x$ and $G \cdot y$ are disjoint. Each orbit is compact because it is a continuous image of G, so we have a pair of disjoint compact subsets. We can now proceed as in Section VI.3 to find disjoint open neighborhoods U_0 and V_0 of $G \cdot x$ and $G \cdot y$. By Wallace's Theorem there are open neighborhoods U and V of x and y such that $G \cdot U \subset \alpha^{-1}[U_0]$ and $G \cdot V \subset \alpha^{-1}[V_0]$. It follows that $p[G \cdot U]$ and $p[G \cdot V]$ are disjoint open neighborhoods of $[x]$ and $[y]$.■

Regularity. Let $[x] \in X/G$, and let $C \subset X/G$ be a closed subset such that $[x] \notin C$. Then $G \cdot x$ is a compact subset of X which is disjoint from the G-invariant closed subset $F = p^{-1}[C]$ (invariance means $G \cdot F = F$). In an arbitrary regular space Y, standard arguments show that if K is a compact subset which is disjoint from the closed subset E, then there are disjoint open subsets U and V containing K and E (this is left to the reader). If we specialize this to the case where $K = G \cdot x$ and $E = F$, then as in the preceding argument there is some open neighborhood W of x such that $G \times W \subset U$. It then follows that $[x]$ and C have disjoint open neighborhoods given by $p[G \cdot W]$ and $p[V]$.■

Normality. Suppose that A and B are disjoint closed subsets of X/G, and let E and F denote the inverse images, which are G-invariant disjoint closed subsets of X. Since X is normal, there are disjoint open subsets U and V containing E and F, and as before there are open neighborhoods M and N of U and V such that $G \times M \subset U$ and $G \times N \subset V$. It follows that $p[G \times M]$ and $p[G \times N]$ are disjoint open subsets containing A and B.■

Local compactness. Let $x \in X$, and suppose that x has an open neighborhood U such that $U \subset C$ where C is compact. Then $p[U]$ is an open neighborhood of $[x] = p(x)$ and $p[U]$ is contained in the compact subset $p[C]$.■

Second countability. Let $\mathcal{B}$ be a countable base for X, and define a family of open subsets in X/G by taking $V_k = p[B_k]$ for $B_k \in \mathcal{B}$. We claim that the sets V_k form a countable base for the topology on X/G. Suppose that W is open in X/G, and let $W' = p^{-1}[W]$. Then there is some subsequence of positive integers $k(n)$ such that $W' = \cup_n \ B_{k(n)}$, and since $p[W'] = W$ we have $W = \cup_n \ p[B_{k(n)}] = \cup_n \ V_{k(n)}$.■

10. This proof is very similar to the proof of Urysohn's Lemma. Since the course notes only give a reference for the proof of Urysohn's Lemma, we shall follow suit and note that the solution to this exercise is on pages 29–30 of the book by Montgomery and Zippin (which is listed in the bibliography at the end of the course notes).

Note. Several other topological properties of orbit space projections (including some relevant to the second part of this course) are given in Chapter I and Section II.6 of the following book:

> **G. E. Bredon.** Introduction to Compact Transformation Groups. Pure and Applied Mathematics Series Vol. 46. *Academic Press, New York*, 1972.

We should note that the conjecture stated at the bottom of page 38 in that book has been shown to be true. One proof is given in Section IV.14 of another book by the same author which is cited below, and some of the result's implications for more general orbit spaces are summarized in Section 1 of the paper by Donnelly and Schultz.

> **G. E. Bredon.** Sheaf Theory (Second Edition). Graduate Texts in Mathematics Vol. 170. *Springer-Verlag, New York etc.*, 1997.

> **H. Donnelly and R. Schultz.** Compact group actions and maps into aspherical manifolds. *Topology* **21** (1982), 443–455.

Additional exercises

Notation. Let $\mathbb{F}$ be the real or complex numbers. Within the matrix group $\mathbf{GL}(n, \mathbb{F})$ there are certain subgroups of particular importance. One such subgroup is the *special linear group* $\mathbf{SL}(n, \mathbb{F})$ of all matrices of determinant 1.

0. We shall follow the steps indicated in the hint.

If N is a normal subgroup of $\mathbf{SL}(2, \mathbb{C})$ and $A \in N$ then N contains all matrices that are similar to A.

If B is similar to A then $B = P\,A\,P^{-1}$ for some $P \in \mathbf{GL}(2, \mathbb{C})$, so the point is to show that one can choose P so that $\det P = 1$. The easiest way to do this is to use a matrix of the form βP for some nonzero complex number β, and if we choose the latter so that $\beta^2 = (\det P)^{-1}$ then βP will be a matrix in in $\mathbf{SL}(2, \mathbb{C})$ and B will be equal to $(\beta P)\,A\,(\beta P)^{-1}$.■

Therefore the proof reduces to considering normal subgroups containing a Jordan form matrix of one of the following two types:

$$\begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix} \quad , \quad \begin{pmatrix} \varepsilon & 1 \\ 0 & \varepsilon \end{pmatrix}$$

Here α is a complex number not equal to 0 or ± 1 and $\varepsilon = \pm 1$.

The preceding step that if N contains a given matrix then it contains all matrices similar to that matrix. Therefore if N is a nontrivial normal subgroup of $\mathbf{SL}(2, \mathbb{C})$ then N is a union of similarity classes, and the latter in turn implies that N contains a given matrix A if and only if it contains a Jordan form for A. For 2×2 matrices there are only two basic Jordan forms; namely, diagonal matrices and elementary 2×2 Jordan matrices of the form

$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$$

where $\lambda \neq 0$. Such matrices have determinant 1 if and only if the product of the diagonal entries is the identity, and this means that the Jordan forms must satisfy conditions very close in the hint; the only difference involves the diagonal case where the diagonal entries may be equal to ± 1. If N is neither trivial nor equal to the subgroup $\{\pm I\}$, then it must contain a Jordan form given by a diagonal matrix whose nonzero entries are not ± 1 or else it must contain an elementary 2×2 Jordan matrix.

The idea is to show that if N contains one of these Jordan forms then it contains all such forms, and this is done by computing sufficiently many matrix products. Trial and error is a good way to approach this aspect of the problem.

The basic strategy is to show first that if N contains either of the matrices

$$\begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix} \quad , \quad \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix}$$

(where $\alpha \neq 0,\ \pm 1$), then N also contains

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

and next to show that if N contains the latter matrix then N contains every Jordan form. As noted before, the latter will imply that $N = \mathbf{SL}(2, \mathbb{C})$.

At this point one needs to do some explicit computations to find products of matrices in N with sufficiently many Jordan forms. Suppose first that N contains the matrix

$$A = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} .$$

Then N contains A^2, which is equal to

$$\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$$

and the latter is similar to

$$B = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

so N contains B. Suppose now that N contains some diagonal matrix A whose nonzero entries are α and α^{-1} where $\alpha \neq \pm 1$, and let B be given as above. Then N also contains the commutator $C = A\,B\,A^{-1}\,B^{-1}$. Direct computation shows that the latter matrix is given as follows:

$$C = \begin{pmatrix} 1 & \alpha^2 - 1 \\ 0 & 1 \end{pmatrix}$$

Since $\alpha \neq \pm 1$ it follows that C is similar to B and therefore $B \in N$.

We have now shown that if N is a nontrivial normal subgroup that is not equal to $\{\pm I\}$, then N must contain B. As noted before, the final step is to prove that if $B \in N$ then $N = \mathbf{SL}(2, \mathbb{C})$.

Since $B \in N$ implies that N contains all matrices similar to B it follows that all matrices of the forms

$$P = \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix} \quad , \qquad Q = \begin{pmatrix} 1 & 0 \\ z & 1 \end{pmatrix}$$

(where z is an arbitrary complex number)

also belong to N, which in turn shows that $P\,Q \in N$. But.

$$P\,Q = \begin{pmatrix} 1 & z \\ z & z^2+1 \end{pmatrix}$$

and its characteristic polynomial is $t^2 - (z^2+2)\,t - 1$. If $z \neq 0$ then this polynomial has distinct nonzero roots such that one is the reciprocal of the other. Furthermore, if $\alpha \neq 0$ then one can always solve the equation $\alpha + \alpha^{-1} = z^2 + 2$ for z over the complex numbers and therefore we see that for each $\alpha \neq 0$ there is a matrix $P\,Q$ similar to

$$\begin{pmatrix} \alpha & 0 \\ 0 & \alpha^{-1} \end{pmatrix}$$

such that $P\,Q \in N$. It follows that each of the displayed diagonal matrices also lies in N; in particular, this includes the case where $\alpha = -1$ so that the diagonal matrix is equal to $-I$. To complete the argument we need to show that N also contains a matrix similar to

$$A = \begin{pmatrix} -1 & 1 \\ 0 & -1 \end{pmatrix} \ .$$

But this is relatively easy because N must contain $-B = (-I)B$ and the latter is similar to A.■

FOOTNOTE. More generally, if $\mathbb{F}$ is an arbitrary field then every proper normal subgroup of $\mathbf{SL}(n, \mathbb{F})$ is contained in the central subgroup of diagonal matrices of determinant 1; of course, the latter is isomorphic to the group of n^{th} roots of unity in $\mathbb{F}$.

Definition. The *orthogonal group* $\mathbf{O}(n)$ consists of all transformations in $\mathbf{GL}(n, \mathbb{R})$ that take each orthonormal basis for $\mathbb{R}^n$ to another orthonormal basis, or equivalently, the subgroup of all matrices whose columns form an orthonormal basis. It is an easy exercise in linear algebra to show that the determinant of all matrices in $\mathbf{O}(n)$ is ± 1. The *special orthogonal group* $\mathbf{SO}(n)$ is the subgroup of $\mathbf{O}(n)$ consisting of all matrices whose determinants are equal to $+1$. Replacing If we replace the real numbers $\mathbb{R}$ by the complex numbers $\mathbb{C}$ we get the unitary groups $\mathbf{U}(n)$ and the *special unitary groups* $\mathbf{SU}(n)$, which are the subgroups of $\mathbf{U}(n)$ given by matrices with determinant 1. The determinant of every matrix in $\mathbf{U}(n)$ is of absolute value 1 just as before, but in the complex case this means that the determinant is a complex number on the unit circle. In Appendix A the orthogonal and unitary groups were shown to be compact.

1. The group $\mathbf{GL}(n, \mathbb{R})$ is isomorphic to the multiplicative group of nonzero real numbers. Therefore $\mathbf{O}(1)$ is isomorphic to set of nonzero real numbers that take 1 to an element of absolute value 1; but a nonzero real number has this property if and only if its absolute value is 1, or equivalently if it is equal to ± 1.

Consider now the group $\mathbf{SO}(2)$. Since it sends the standard basis into an orthonormal basis, it follows that its columns must be orthonormal. Therefore there is some θ such that the entries of the first column are $\cos\theta$ and $\sin\theta$, and there is some φ such that the entries of the second column are $\cos\varphi$ and $\sin\varphi$. Since the vectors in question are perpendicular, it follows that $|\theta - \varphi| = \pi/2$, and depending upon whether the sign of $\theta - \varphi$ is positive or negative there are two possibilities:

(1) If $\theta - \phi > 0$ then $\cos\varphi = -\sin\theta$ and $\sin\varphi = \cos\theta$.

(2) If $\theta - \phi < 0$ then $\cos\varphi = \sin\theta$ and $\sin\varphi = -\cos\theta$.

In the first case the determinant is 1 and in the second it is -1.

We may now construct the isomorphism from S^1 to $\mathbf{SO}(2)$ as follows: If $z = x + y\,i$ is a complex number such that $|z| = 1$, send z to the matrix

$$\begin{pmatrix} x & -y \\ y & x \end{pmatrix} .$$

This map is clearly 1–1 because one can retrieve z directly from the entries of the matrix. It is onto because the complex number $\cos\theta + i\,\sin\theta$ maps to the matrix associated to θ. Verification that the map takes complex products to matrix products is an exercise in bookkeeping.

A homeomorphism from $\mathbf{SO}(2) \times \mathbf{O}(1)$ to $\mathbf{O}(2)$ may be constructed as follows: Let $\alpha : \mathbf{O}(1) \to \mathbf{O}(2)$ be the map that sends ± 1 to the diagonal matrix whose entries are ± 1 and 1, and define $M(A, \varepsilon)$ to be the matrix product $A\,\alpha(\varepsilon)$. Since the image of -1 does not lie in $\mathbf{SO}(2)$, standard results on cosets in group theory imply that the map M is 1–1 and onto. But it is also continuous, and since it maps a compact space to a Hausdorff space it is a homeomorphism onto its image, which is $\mathbf{SO}(2)$.

To see that this map is not a group isomorphism, note that in the direct product group $\mathbf{SO}(2) \times \mathbf{O}(1)$ there are only finitely many elements of order 2 (specifically, the first coordinate must be $\pm I$). On the other hand, the results from Appendix D show that all elements of $\mathbf{O}(2)$ that are not in $\mathbf{SO}(2)$ have order 2 and hence there are infinitely many such element in $\mathbf{O}(2)$. Therefore the latter cannot be isomorphic to the direct product group.∎

2. This is similar to the first part of the preceding exercise. The group $\mathbf{GL}(n, \mathbb{C})$ is isomorphic to the multiplicative group of nonzero complex numbers. Therefore $\mathbf{U}(1)$ is isomorphic to set of nonzero complex numbers that take 1 to an element of absolute value 1; but a nonzero complex number has this property if and only if its absolute value is 1, or equivalently if it lies on the circle S^1.∎

3. The proof separates into three cases depending upon whether the group in question is $\mathbf{U}(n)$, $\mathbf{SU}(n)$ or $\mathbf{SO}(n)$,

The case $G = \mathbf{U}(n)$.

We start with the unitary group. The Spectral Theorem states that if A is an $n \times n$ unitary matrix then there is another invertible unitary matrix P such that $B = P\,A\,P^{-1}$ is diagonal. We claim that there is a continuous curve joining B to the identity. To see this, write the diagonal entries of B as $\exp it_j$ where t_j is real and $1 \leq j \leq n$. Let $C(s)$ be the continuous curve in the diagonal unitary matrices such that the diagonal entries of $C(s)$ are $\exp(i\,st_j)$ where $s \in [0, 1]$. It follows immediately that $C(0) = I$ and $C(1) = B$. Finally, if we let $\gamma(s) = P^{-1}\,C(s)\,P$ then γ is a continuous curve in $\mathbf{U}(n)$ such that $\gamma(0) = I$ and $\gamma(1) = A$. This shows that $\mathbf{U}(n)$ is in fact arcwise connected.∎

The case $G = \mathbf{SU}(n)$.

The preceding argument also shows that $\mathbf{SU}(n)$ is arcwise connected; it is only necessary to check that if A has determinant 1 then everything else in the construction also has this property. The determinants of B and A are equal because the determinants of similar matrices are equal. Furthermore, since the determinant of B is 1 it follows that $\sum_j t_j = 0$, and the latter in turn implies that the image of the curve $C(s)$ is contained in $\mathbf{SU}(n)$. Since the latter is a normal subgroup of $\mathbf{U}(n)$ it follows that the curve $\gamma(s)$ also lies in $\mathbf{SU}(n)$, and therefore we conclude that the latter group is also arcwise connected.

The product decomposition for $\mathbf{U}(n)$ is derived by an argument similar to the previous argument for $\mathbf{O}(2)$. More generally, suppose we have a group G with a normal subgroup K and a second subgroup H such that $G = H \cdot K$ and $H \cap K = \{1\}$. Then group theoretic considerations yield a 1–1 onto map $\varphi : K \times H \to G$ given by $\varphi(k, h) = k \cdot h$. If G is a topological group and the subgroups have the subspace topologies then the 1–1 onto map φ is continuous. Furthermore, if G is compact Hausdorff and H and K are closed subgroups of G then φ is a homeomorphism. In our particular situation we can take $G = \mathbf{U}(n)$, $K = \mathbf{SU}(n)$ and $H \cong \mathbf{U}(1)$ to be the subgroup of diagonal matrices with ones down the diagonal except perhaps for the $(1, 1)$ entry. These subgroups satisfy all the conditions we have imposed and therefore we have that G is homeomorphic to the product $K \times H$.∎

The case $G = \mathbf{SO}(n)$.

Finally, we need to verify that $\mathbf{SO}(n)$ is arcwise connected, and as indicated in the hint we use the normal form obtained in Appendix D. According to this result, for every orthogonal $n \times n$ matrix A there is another orthogonal matrix P such that $B = PAP^{-1}$ is is a block sum of orthogonal matrices that are either 2×2 or 1×1. These matrices may be sorted further using their determinants (recall that the determinant of an orthogonal matrix is ± 1).

In fact, one can choose the matrix P such that the block summands are sorted by size and determinant such that

(1) the 1×1 summands with determinant 1 come first,
(2) the 2×2 summands with determinant 1 come second,
(3) the 2×2 summands with determinant -1 come third,
(4) the 1×1 summands with determinant -1 come last.

Each matrix of the first type is a rotation matrix R_θ where the first row has entries $(\cos\theta \quad -\sin\theta)$, and each matrix of the second type is a matrix S_θ where the first row has entries $(\cos\theta \quad \sin\theta)$.

The first objective is to show that B lies in the same arc component as a matrix with no summands of the second type. Express the matrix B explicitly as a block sum as follows:

$$B = I_k \oplus \left(\bigoplus_{i=1}^{p} R_{\theta(i)} \right) \oplus \left(\bigoplus_{j=1}^{q} R_{\varphi(j)} \right) \oplus -I_\ell$$

Here I_m denotes an $m \times m$ identity matrix. Consider the continuous curve in $\mathbf{SO}(n)$ defined by the formula

$$C_1(t) = I_k \oplus \left(\bigoplus_{i=1}^{p} R_{t\theta(i)} \right) \oplus I_{2q+\ell} \ .$$

It follows that $C_1(1) = B$ and $C_1(0) = B_1$ is a block sum matrix with no summands of the second type.

The next objective is to show that B_1 lies in the same arc component as a matrix with no summands of the third type. This requires more work than the previous construction, and the initial step is to show that one can find a matrix in the same arc component with at most one summand of the third type. The crucial idea is to show that if the block sum has $q \geq 2$ summands of the third type, then it lies in the same arc component as a matrix with $q-2$ block summands of the third type. In fact, the continuous curve joining the two matrices will itself be a block sum, with one 4×4 summand corresponding to the two summands that are removed and identity matrices corresponding to the remaining summands. Therefore the argument reduces to looking at a 4×4 orthogonal matrix that is a block sum of two 2×2 matrices of the third type, and the objective is to show that such a matrix lies in the same arc component as the identity matrix.

Given real numbers α and β, let S_α and S_β be defined as above, and likewise for R_α and R_β. Then one has the multiplicative identity

$$S_\alpha \oplus S_\beta \quad = \quad \big(R_\alpha \oplus R_\beta\big) \cdot \big(1 \oplus (-1) \oplus 1 \oplus (-1)\big) \ .$$

We have already shown that there is a continuous curve $C(t)$ in the orthogonal group such that $C(0) = I$ and $C(1) = R_\alpha \oplus R_\beta$. If we can construct a continuous curve $Q(t)$ such that $Q(0) = I$ and $Q(1) = \big(1 \oplus (-1) \oplus 1 \oplus (-1)\big)$. Then the matrix product $C(t)\,Q(t)$ will be a continuous curve in the orthogonal group whose value at 0 is the identity and whose value at 1 is $S_\alpha \oplus S_\beta$. But $R_\alpha \oplus R_\beta$ is a rotation by 180 degrees in the second and fourth coordinates and the identity on the first and third coordinates, so it is natural to look for a curve $Q(t)$ that is rotation through $180\,t$ degrees in the even coordinates and the identity on the odd ones. One can write down such a curve explicitly as follows:

$$Q(t) \quad = \quad \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \pi t & 0 & -\sin \pi t \\ 0 & 0 & 1 & 0 \\ 0 & \sin \pi t & 0 & \cos \pi t \end{pmatrix}$$

Repeated use of this construction yields a matrix with no block summands of the first type and at most one summand of the second type. If there are no summands of the third type, we have reached the second objective, so assume that there is exactly one summand of the third type. Since our block sum has determinant 1, it follows that there is also at least one summand of the fourth type. We claim that the simplified matrix obtained thus far lies in the same arc component as a block sum matrix with no summands of the third type and one fewer summands of the fourth type. As in the previous step, everything reduces to showing that a 3×3 block sum $S_\alpha \oplus (-1)$ lies in the same arc component as the identity. In this case one has the multiplicative identity

$$S_\alpha \oplus (-1) \quad = \quad \big(R_\alpha \oplus 1\big) \oplus \Big(1 \oplus (-1) \oplus (-1)\Big)$$

and the required continuous curve is given by the formula

$$C(t) \quad = \quad \big(R_{t\alpha} \oplus 1\big) \oplus \Big((1 \oplus R_{t\pi}(-1)\Big) \ .$$

Thus far we have shown that every block sum matrix in $\mathbf{SO}(n)$ lies in the same arc component as a diagonal matrix (only summands of the first and last types). Let $I_s \oplus -I_\ell$ be this matrix. Since its determinant is 1, it follows that ℓ must be even, so write $\ell = 2k$. Then the diagonal matrix is a block sum of an identity matrix and k copies of R_π, and one can use the continuous curve $D(t)$

that is a block sum of an identity matrix with k copies of $R_{t\pi}$ to show that the diagonal matrix $I_s \oplus -I_\ell$ lies in the same arc component as the identity.

We have now shown that every block sum matrix with determinant 1 lies in the arc component of the identity. Suppose that $\Gamma(t)$ is a continuous curve whose value at 0 is the identity and whose value at 1 is the block sum. At the beginning of this proof we noted that for every orthogonal matrix A we can find another orthogonal matrix P such that $B = P\,A\,P^{-1}$ is is a block sum of the type discussed above; note that the determinant of B is equal to 1 if the same is true for the determinant of A. If $\Gamma(t)$ is the curve joining the identity to B, then $P^{-1}\,\Gamma(t)\,P$ will be a continuous curve in $\mathbf{SO}(n)$ joining the identity to A.■

4. One can use the same sort of arguments that established topological product decompositions for $\mathbf{O}(2)$ and $\mathbf{SU}(n)$ to show that $\mathbf{O}(n)$ is homeomorphic to $\mathbf{SO}(n) \times \mathbf{O}(1)$. Since $\mathbf{SO}(n)$ is connected, this proves that $\mathbf{O}(n)$ is homeomorphic to a disjoint union of two copies of $\mathbf{SO}(n)$.■

5. Let A be an invertible $n \times n$ matrix over the real or complex numbers, and consider the meaning of the Gram-Schmidt process for matrices. Let $\mathbf{a}_j$ represent the j^{th} column of A, and let B be the orthogonal or unitary matrix whose columns are the vectors $\mathbf{b}_j$ obtained from the columns of A by the Gram-Schmidt process. How are they related? The basic equations for defining the columns of B in terms of the columns of A have the form

$$\mathbf{b}_j = \sum_{k \leq j} c_{k,j}\mathbf{a}_k$$

where each $c_{j,j}$ is a positive real number, and if we set $c_{k,j} = 0$ for $k > j$ this means that $B = AC$ where C is lower triangular with diagonal entries that are positive and real. We claim that there is a continuous curve $\Gamma(t)$ such that $\Gamma(0) = I$ and $\Gamma(1) = C$. Specifically, define this curve by the following formula:

$$\Gamma(t) \quad = \quad I \ + \ t\,(C - I)$$

By construction $\Gamma(t)$ is a lower triangular matrix whose diagonal entries are positive real numbers, and therefore this matrix is invertible. If we let $\Phi(t)$ be the matrix product $A\,\Gamma(t)$ then we have a continuous curve in the group of invertible matrices such that $\Phi(1) = A$ and $\Phi(0)$ is orthogonal or unitary. Therefore it follows that every invertible matrix is in the same arc component as an orthogonal or unitary matrix.

In the unitary case this proves the result, for the arcwise connectedness of $\mathbf{U}(n)$ and the previous argument imply that $\mathbf{GL}(n,\mathbb{C})$ is also arcwise connected. However, in the orthogonal case a little more work is needed. The determinant of a matrix in $\mathbf{GL}(n,\mathbb{R})$ is either positive or negative, so the preceding argument shows that an invertible real matrix lies in the same arc component as the matrices of $\mathbf{SO}(n)$ if its determinant is positive and in the other arc component of $\mathbf{SO}(n)$ if its determinant is negative. Therefore there are at most two arc components of $\mathbf{GL}(n,\mathbb{R})$, and the assertion about arc components will be true if we can show that $\mathbf{GL}(n,\mathbb{R})$ is not connected. To see this, note that the determinant function defines a continuous onto map from $\mathbf{GL}(n,\mathbb{R})$ to the disconnected space $\mathbb{R} - \{0\}$.■

6. We need to show that the open neighborhood $\mathcal{N} = \mathcal{N}(\Phi)$ of $\{1\} \times X$ (the points where Φ is defined) is equal to all of $G \times X$. By Property (1) of a local group action, we know that $\{1\} \times X \subset \mathcal{N}$. Since X is compact, by Wallace's Theorem (see Section VI.2) we know that there is some open neighborhood V of 1 in G such that $V \times X \subset \mathcal{N}$. If we replace V by $U = V \cap V^{-1}$, then $U \times X \subset \mathcal{N}$ and $U = U^{-1}$ is an open neighborhood of the identity.

Since U is connected and $U = U^{-1}$, we know that every open neighborhood of U generates G. Given $g \in G$, write

$$g = \prod_{j=1}^{m} u_j^{\varepsilon_j}$$

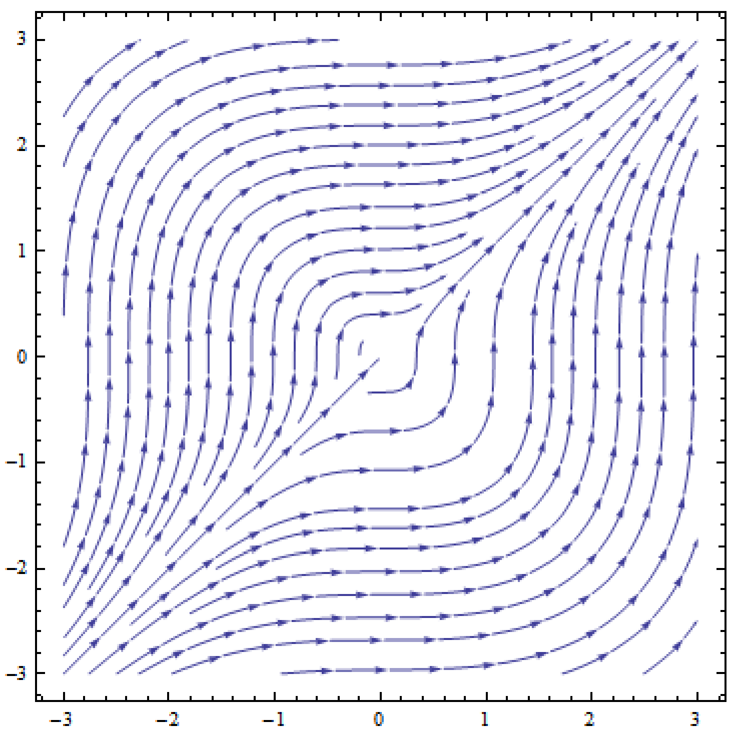
where $u_j \in U$ and $\varepsilon_j = \pm 1$. If p_j is the product of the last j factors (so $p_1 = u_m^{\varepsilon_m}$ and $p_j = u_{m+1-j}^{\varepsilon_{m+1-j}} \cdot p_{j-1}$ recursively), then induction on j and Property (2) imply that $(p_j, x) \in \mathcal{N}$ for $j = 1, \cdots, m$. Since $p_m = g$, it follows that for all x and all g we have $(g, x) \in \mathcal{N}$ and hence $\mathcal{N} = G \times X$.■

Further information on Additional Exercise A.6

We shall first say more about our assertion that the integral curves define a local action of $\mathbb{R}$ on an open set V. By definition, $\gamma(t) = \Phi(t, z)$ is the unique solution curve (or integral curve) for the system of differential equations $x' = M(x, y)$ and $y' = N(x, y)$ with initial condition $\gamma(t) = z$. Therefore the Chain Rule implies that $\beta(t) = \gamma(t + s)$ is the solution curve with initial condition $\beta(0) = \gamma(s)$. If we translate this back into a statement about the mapping Φ, we obtain the identity

$$\Phi(t, \Phi(s, z)) = \Phi(t + s, z).$$

For our example in which the local action does not extend to a global action, we have already noted that the solution curves satisfy the condition $y^3 = x^3 + C$, a fact which follows from the standard methods taught in a first course on differential equations like Mathematics **46**. A drawing which illustrates the solution curves is given below:



(***Source:*** http://math.stackexchange.com/questions/536370/how-to-show-that-a-given-vector-field-is-not-complete-in-mathbbr2)