We noted last time that a space that is homeomorphic to a connected space is also connected. In fact,it is very easy to establish a stronger result: continuous images of connected sets are connected.

Theorem. *Let $f\colon X \to Y$ be continuous, and suppose that A is a connected subset of X. Then $f(A)$ is a connected subset of Y.*

Proof. If $f(A)$ is disconnected then there exist open subsets V_1, V_2 of Y such that

$$f(A) \subseteq V_1 \cup V_2, \tag{1}$$
$$f(A) \cap V_1 \neq \emptyset \quad \text{and} \quad f(A) \cap V_2 \neq \emptyset, \tag{2}$$
$$f(A) \cap V_1 \cap V_2 = \emptyset. \tag{3}$$

Since f is continuous, $f^{-1}(V_1)$ and $f^{-1}(V_1)$ are open subsets of X. By (1) above, for each $a \in A$ we have either $f(a) \in V_1$ or $f(a) \in V_2$; that is, either $a \in f^{-1}(V_1)$ or $a \in f^{-1}(V_2)$. So $A \subseteq f^{-1}(V_1) \cup f^{-1}(V_2)$. By (2) there is an $a \in A$ such that $f(a) \in V_1$, giving $a \in f^{-1}(V_1)$, and similarly there is an $a' \in A$ with $a' \in f^{-1}(V_2)$. So $A \cap f^{-1}(V_1) \neq \emptyset$ and $A \cap f^{-1}(V_2) \neq \emptyset$. Finally, $A \cap f^{-1}(V_1) \cap f^{-1}(V_2) = \emptyset$, since if there were some element a in this set it would follow that $f(a) \in f(A) \cap V_1 \cap V_2$, contradicting (3). We have shown that

$$A \subseteq f^{-1}(V_1) \cup f^{-1}(V_2), \tag{1'}$$
$$A \cap f^{-1}(V_1) \neq \emptyset \quad \text{and} \quad A \cap f^{-1}(V_2) \neq \emptyset, \tag{2'}$$
$$A \cap f^{-1}(V_1) \cap f^{-1}(V_2) = \emptyset, \tag{3'}$$

contradicting the fact that A is connected. So the assumption that $f(A)$ is disconnected has led to a contradiction; so $f(A)$ is connected. □

There is any even shorter proof using the fact that a set is disconnected if and only if there is a continuous surjective function from the set to the discrete space $\{0, 1\}$. If $f(A)$ is disconnected then there is a surjective continuous function $g\colon f(A) \to \{0, 1\}$, and then the function from A to $\{0, 1\}$ given by $a \mapsto g(f(a))$ is continuous (since composites of continuous functions are continuous) and surjective (since g is surjective). So A is disconnected.

As we shall see, it is not clear that the definition of connectedness that we have given really captures the everyday concept of connectedness, which is perhaps more to do with path-connectedness, a concept that we shall define in due course, and that is stonger than connectedness. However, if intuition suggests that a set is connected, then it ought to be true that the set is indeed connected in the technical sense. In particular, intervals in $\mathbb{R}$ are connected sets.

There are nine different kinds of intervals: (a, b), $(a, b]$, $[a, b)$, for any a, $b \in \mathbb{R}$ with $a < b$, $[a, b]$, for any a, $b \in \mathbb{R}$ with $a \leq b$, $(-\infty, a)$, $(\infty, a]$, (a, ∞), $[a, \infty)$, for any $a \in \mathbb{R}$, and $(-\infty, \infty)$ (the whole real line). † Intervals can be characterized as follows: a subset I of $\mathbb{R}$ is an interval if and only if I is nonempty, and for all a, $b \in I$ and $x \in \mathbb{R}$, if $a \leq x \leq b$

† We have deviated from the convention adopted in Choo's notes by permitting one-element subsets of $\mathbb{R}$ to be counted as closed intervals.

then $x \in I$. That is to say, if $I \neq \emptyset$ then I is an interval if and only if every point of $\mathbb{R}$ that lies between two points of I is also in I.

Lemma. *Let a, $b \in \mathbb{R}$ with $a < b$, and S a subset of $\mathbb{R}$ such that $a \in S$ and $b \notin S$, and let $p = \sup(S \cap [a, b])$.*
(i) If S is closed in $\mathbb{R}$ then $p \in S$.
(ii) If S is open in $\mathbb{R}$ then $p \notin S$.

Proof. Note that $S \cap [a, b]$ is nonempty (since $a \in S \cap [a, b]$) and bounded above (by b). So, by an axiom of the real number system, $S \cap [a, b]$ has a least upper bound. So the definition of p in the statement of the lemma is meaningful. Observe that $a \leq p$ (since $a \in S \cap [a, b]$ and p is an upper bound for $S \cap [a, b]$) and $p \leq b$ (since b is an upper bound for $S \cap [a, b]$ and p is the least upper bound for $S \cap [a, b]$).

Suppose that S is closed, and suppose that $p \notin S$. Then $p \in \mathbb{R} \setminus S$, which is an open set since S is closed, and so there exists an $\varepsilon > 0$ such that $B(p, \varepsilon) \subseteq \mathbb{R} \setminus S$. Of course, since we are discussing $\mathbb{R}$ with its usual metric, $B(p, \varepsilon) = (p - \varepsilon, p + \varepsilon)$. Now let $x \in S \cap [a, b]$ be arbitrary. Since p is an upper bound for $S \cap [a, b]$ we have $x \leq p$, and so either $x \leq p - \varepsilon$ or $p - \varepsilon < x \leq p$. The latter alternative gives $x \in (p - \varepsilon, p] \subseteq (p - \varepsilon, p + \varepsilon) \subseteq \mathbb{R} \setminus S$, contradicting $x \in S$, and so we must have $x \leq p - \varepsilon$. Since this holds for all $x \in S \cap [a, b]$ it follows that $p - \varepsilon$ is an upper bound for $S \cap [a, b]$. But $p - \varepsilon < p$, and so this contradicts the fact that p is the least upper bound for S.

For the second part, suppose that S is open and $p \in S$. Since $b \notin S$ and $p \leq b$ it follows that $p < b$. Thus $p \in (-\infty, b) \cap S$, an open set since both $(-\infty, b)$ and S are open, and so there exists an $\varepsilon > 0$ such that $(p - \varepsilon, p + \varepsilon) \subseteq (-\infty, b) \cap S$. In particular, $p + (\varepsilon/2) \in S$ and $p + (\varepsilon/2) < b$, and since $a \leq p < p + (\varepsilon/2)$ it follows that $p + (\varepsilon/2) \in S \cap [a, b]$. But since $p + (\varepsilon/2) > p$, this contradicts the fact that p is an upper bound for $S \cap [a, b]$. □

Proposition. *Let $I \subseteq \mathbb{R}$ be an interval. Then I is connected.*

Proof. Suppose that I is not connected. Then there exist open subsets U_1, U_2 of $\mathbb{R}$ with $I \cap U_1$ and $I \cap U_2$ nonempty, $I \cap U_1 \cap U_2 = \emptyset$, and $I \subseteq U_1 \cup U_2$. We can choose $a \in I \cap U_1$ and $b \in I \cap U_2$ (since these sets are nonempty), and then $a \neq b$ (since $I \cap U_1 \cap U_2 = \emptyset$). Swapping the names of U_1 and U_2 if necessary, we may assume that $a < b$.

Since a, $b \in I$, and I is an interval, it follows from our characterization of intervals that $[a, b] \subseteq I$. Now put $A = [a, b] \cap U_1$ and $B = [a, b] \cap U_2$. Then

$$A \cup B = [a, b] \cap (U_1 \cup U_2) = [a, b],$$

since $[a, b] \subseteq I \subseteq U_1 \cup U_2$, and

$$A \cap B \subseteq I \cap U_1 \cap U_2 = \emptyset;$$

so $A = [a, b] \setminus B = [a, b] \setminus U_2$. Now if we define $p = \sup A$ then it follows from the first part of the lemma that $p \in A$, since $A = [a, b] \cap (\mathbb{R} \setminus U_2)$ and $\mathbb{R} \setminus U_2$ is closed. However, $A = [a, b] \cap U_1$ and U_1 is open; so it follows from the second part of the lemma that $p \notin A$. Thus we have obtained the desired contradiction. □

Our characterization of intervals also yields the following converse to the above result.

Proposition. *If $A \subseteq \mathbb{R}$ is connected and nonempty then A is an interval.*

Proof. Suppose that A is connected and $A \neq \emptyset$, and suppose that A is not an interval. By the characterization, of intervals there exist a, $b \in A$ and $x \notin A$ with $a \leq x \leq b$. Put $U_1 = (-\infty, x)$ and $U_2 = (x, \infty)$. Then U_1 and U_2 are open subsets of $\mathbb{R}$ with $a \in A \cap U_1$ and $b \in A \cap U_2$ (showing that $A \cap U_1$ and $A \cap U_2$ are both nonempty), $A \subseteq \mathbb{R} \setminus \{x\} = U_1 \cup U_2$, and $A \cap U_1 \cap U_2 = \emptyset$ (since $U_1 \cap U_2 = \emptyset$). This shows that A is not connected. □