MATH 4331/6312: HANDOUT 1

B. G. BODMANN

Connected Sets and the Intermediate Value Theorem

We will see in this section that the intermediate value theorem from calculus is really a consequence of the fact that an interval of real numbers is a connected set. First we need to define this concept.

1. **Definition.** *A subset A of $\mathbb{R}^n$ is* disconnected *if there are two open sets U, V that are disjoint, $U \cap V = \emptyset$, each of them has a non-empty intersection with A, and $A \subset U \cup V$. Otherwise, we say that A is* connected.

2. **Example.** *If A is a non-empty set containing a finite number of points $\{p_1, p_2, \ldots, p_k\}$ in $\mathbb{R}^n$, and $k \geq 2$, then A is disconnected.*

The most important example of a connected space is an *interval* in $\mathbb{R}$, which means either an open interval, closed interval, or half-open interval. The limits can be numbers or $+\infty$ or $-\infty$.

3. **Theorem.** *Let $I \subseteq \mathbb{R}$ be an interval or all of $\mathbb{R}$, then I is a connected set. Conversely, if A is a connected set in $\mathbb{R}$, then A is an interval.*

Proof. Suppose I is an interval but not connected, then we would have $I \subset U \cup V$ where U and V are both non-empty and open, $U \cap I \neq \emptyset$, $V \cap I \neq \emptyset$, and $U \cap V = \emptyset$. Let $a \in U$ and $b \in V$. Without loss of generality, we can assume that $a < b$ (otherwise just change the names of U and V). Let $U_1 = U \cap I \cap [a, b]$, and $V_1 = V \cap I \cap [a, b]$. Then U_1 and V_1 are disjoint, non-empty sets because $a \in U_1$ and $b \in V_1$. Since U_1 is bounded, $c = \sup\{x : x \in U_1\} < \infty$. From $b \in V$, $a < b$, and V being open, we know there is $\delta > 0$ such that $(v - \delta, v) \subset V_1$. Let δ be the maximal choice that satisfies this inclusion in V_1 (this exists because we know $\delta \leq b - a$). We then know $a \leq c \leq b - \delta$. Moreover, by the assumption on δ, $c \notin V$, otherwise the openness of V would allow us to increase δ.
Since c is not the right limit of I, we also know that $c \notin U$, otherwise by the openness of U and $c < b$, it would not be the supremum of U_1.
Hence $c \notin U \cup V$, but this union has I as its subset, so $c \notin I$. Thus, I is not an interval.
Conversely, assume A is a connected set in $\mathbb{R}$. Take $a, b \in A$ with $a < b$. To show A is an interval, we prove that each $x \in \mathbb{R}$ with $a < x < b$ satisfies $x \in A$. If this is not the case for some x, we can define $U = (-\infty, x)$ and $V = (x, \infty)$, then $a \in U \cap A$, $b \in V \cap A$ but U and V are open and disjoint, hence A is disconnected. □

Now we come to a general version of the Intermediate Value Theorem.

4. **Theorem** (Intermediate Value Theorem in $\mathbb{R}^n$). *Let $f : S \subset \mathbb{R}^n \to \mathbb{R}$ be a continuous function and S a connected set. Given $x, y \in S$ and $L \in \mathbb{R}$ with $f(x) < L < f(y)$, then there is $z \in S$ with $f(z) = L$.*

Proof. Suppose there is no such z. Let $X = f^{-1}((-\infty, L))$ and $Y = f^{-1}((L, \infty))$, then X and Y are disjoint, $S = X \cup Y$ and by the continuity both of these sets are open in S. Consequently, from the definition of relative openness, $X = S \setminus \overline{Y}$ and $Y = S \setminus \overline{X}$. Let U and V be open sets in $\mathbb{R}^n$ such that $X = U \cap S$ and $Y = V \cap S$, then $U_1 = U \setminus \overline{V}$ and $V_1 = V \setminus \overline{U}$ are open and disjoint and we retain the intersection property $X = U \cap S \setminus \overline{V} = U_1 \cap S$ and $Y = V \cap S \setminus \overline{U} = V_1 \cap S$.
Thus, S is a subset of two open disjoint sets U_1 and V_1 in $\mathbb{R}^n$, with a non-empty intersection with each of them, which means it is disconnected, contradicting our assumption. Hence, L is in the range of f. □

We conclude the usual Intermediate Value Theorem by specializing to $n = 1$ and $S = I$ with I an interval.

5. **Corollary** (Intermediate Value Theorem). *Let $I \subseteq \mathbb{R}$ be an interval or the whole real line and let $f : I \to \mathbb{R}$ be continuous. If $x_0, x_1 \in I$, $L \in \mathbb{R}$ and $f(x_0) < L < f(x_1)$, then there is x_2 between x_0 and x_1 with $f(x_2) = L$.*

The same type of proof gives another insight that relates continuity and connectedness.

6. **Theorem.** *Let $f : S \subset \mathbb{R}^m \to \mathbb{R}^n$ be continuous. If S is connected, then $f(S)$ is also connected.*

Proof. Let $A = f(S)$. Assume A is disconnected, then there exist U, V open disjoint, each having a non-empty intersection with A and $A \subset U \cup V$. By the continuity of f, $f^{-1}(U)$ and $f^{-1}(V)$ are open in S, and satisfy that they are disjoint, have a non-empty intersection with S, and give $S = f^{-1}(U) \cup f^{-1}(V)$. Using the same argument as in the preceding theorem, S is disconnected, contradicting our assumption. □

Again specializing to $\mathbb{R}$, we obtain the following consequence.

7. **Corollary.** *Let $I \subseteq \mathbb{R}$ be an interval. If $f : I \to \mathbb{R}$ is continuous, then $f(I)$ is an interval.*

We recall the informal statement that "a real-valued function is continuous if we can draw its graph without lifting the pen." To make this statement precise, we need a stronger form of connectedness.

8. **Definition.** *A set $A \subset \mathbb{R}^n$ is called* path connected *provided that for any two points, $a, b \in A$ there exists a continuous function $f : [0, 1] \to A$ such that $f(0) = a$ and $f(1) = b$.*

Intuitively, a space is pathwise connected if and only if you can draw a curve between any two points with no gaps in the curve.

9. **Theorem.** *If $A \subset \mathbb{R}^n$ is path connected, then it is connected.*

Proof. Suppose A is path connected, but not connected. Then there are open sets U, V with $A \subset U \cup V$, $U \cap V = \emptyset$ and there are $a \in A \cap U$, $b \in A \cap V$. By path connectedness, there is a continuous function $f : [0, 1] \to A$ with $f(0) = a$ and $f(1) = b$. Extending f to all of $\mathbb{R}$ by $f(x) = a$ if $x < 0$ and $f(x) = b$ if $x > 1$ retains the continuity. This then gives open disjoint sets $f^{-1}(U)$ and $f^{-1}(V)$ whose union contains $[0, 1]$ as a subset and $0 \in f^{-1}(U)$, $1 \in f^{-1}(V)$, so $[0, 1]$ is disconnected, which is a contradiction. Hence, A is connected. □

10. **Definition.** *Let $g : [a, b] \to \mathbb{R}$. By the* graph of g *we mean the set $G = \{(x, g(x)) : a \leq x \leq b\} \subseteq \mathbb{R}^2$.*

11. **Theorem.** *Let $g : [a, b] \to \mathbb{R}$. The function g is continuous if and only if the graph of g is a path connected subset of $\mathbb{R}^2$.*