LECTURE 25

At the end of last lecture, we proved

Proposition 6.14: Suppose that $f\colon [a,b] \to \mathbb{R}$ is continuous and 1-1 with image $f([a,b]) = [c,d]$. Then $f^{-1}\colon [c,d] \to \mathbb{R}$ is continuous.

There is a version of this proposition which deals with open intervals: if I is an open interval of the form (a,b), (a,∞), $(-\infty,b)$ or $\mathbb{R}$ and $f\colon I \to f(I)$ is continuous and 1-1, then $f^{-1}\colon f(I) \to I$ is continuous.

Our next goal is to prove the following result - the *inverse function theorem.*

Proposition 6.15: Suppose $f\colon [a,b] \to [c,d]$ is continuous, 1-1 and onto. If f is differentiable at x_0 with $f'(x_0) \neq 0$, then $f^{-1}\colon [c,d] \to [a,b]$ is differentiable at $f(x_0)$ with

$$(f^{-1})'(f(x_0)) = \frac{1}{f'(x_0)}.$$

Put another way, if $y_0 = f(x_0)$, then

$$(f^{-1})'(y_0) = \frac{1}{f'(f^{-1}(y_0))}.$$

Proof: Since f is differentiable at x_0, there exists a function $R_f\colon [a,b] \to \mathbb{R}$ such that

$$f(x) - f(x_0) = (f'(x_0) + R_f(x))(x - x_0)$$

for all $x \in [a,b]$, with $\lim_{x\to x_0} R_f(x) = R_f(x_0) = 0$. Putting $y = f(x)$ and $y_0 = f(x_0)$, the equation above becomes

$$y - y_0 = (f'(x_0) + R_f(f^{-1}(y))(f^{-1}(y) - f^{-1}(y_0)).$$

Since R_f is continuous at x_0, and f^{-1} is continuous at y_0 (Proposition 6.14), it follows that the composite function $R_f(f^{-1}(y))$ is continuous at y_0 (Theorem 4.5), and hence that $f'(x_0) + R_f(f^{-1}(y))$ is continuous at y_0. Since $f'(x_0) + R_f(f^{-1}(y_0)) = f'(x_0) \neq 0$, it follows that there is an open interval $I \subset [c,d]$, containing y_0, such that $f'(x_0)R_f(f^{-1}(y)) \neq 0$ for all $y \in I$.

Therefore, if $y \in I$, and $y \neq y_0$, then

$$\frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \frac{1}{f'(x_0) + R_f(f^{-1}(y))}.$$

Since $f'(x_0) + R_f(f^{-1}(y)) \to f'(x_0)$ as $y \to y_0$, it follows from Proposition 6.6 that

$$\lim_{y\to y_0} \frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \lim_{y\to y_0} \frac{1}{f'(x_0) + R_f(f^{-1}(y))} = \frac{1}{f'(x_0)}.$$

Therefore f^{-1} is differentiable at y_0, with

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)}.$$

This completes the proof of the Proposition. ■

Note: more generally, suppose I is an open interval of the form (a, b), (a, ∞), $(-\infty, a)$ or $\mathbb{R}$, and $f\colon I \to f(I)$ is 1-1 and continuous. If f is differentiable at $x_0 \in I$ with $f'(x_0) \neq 0$, then $f^{-1}\colon f(I) \to I$ is differentiable at $f(x_0)$ with $(f^{-1})'(f(x_0)) = 1/f'(x_0)$. The proof is analogous to the proof above.

Suppose now that $f\colon S \to \mathbb{R}$ is a function. Recall that we say that f has a *maximum* (respectively *minimum*) at $x_0 \in S$ if $f(x) \leq f(x_0)$ for all $x \in S$ (respectively $f(x) \geq f(x_0)$ for all $x \in S$). We say that f has a *local maximum* at $x_0 \in S$ if there exists $\delta > 0$ such that $f(x) \leq f(x_0)$ for all $x \in S$ such that $|x - x_0| < \delta$. Similarly we say that f has a *local minimum* at $x_0 \in S$ if there exists $\delta > 0$ such that $f(x) \geq f(x_0)$ for all $x \in S$ such that $|x - x_0| < \delta$.

Lemma 6.16: Suppose that $f\colon (a, b) \to \mathbb{R}$ is differentiable on (a, b). If f has a local maximum at $x_0 \in (a, b)$ then $f'(x_0) = 0$.

Proof: Without loss of generality $f(x) \leq f(x_0)$ for all $x \in (a, b)$. Let (x_n) be a sequence in (a, b) such that $x_0 < x_n < b$ for all n and such that $x_n \to x_0$. Since f is differentiable at x_0 we have

$$\lim_{n\to\infty} \frac{f(x_n) - f(x_0)}{x_n - x_0} = f'(x_0).$$

Since f has a local maximum at x_0 we have $f(x_n) - f(x_0) \leq 0$. Therefore $(f(x_n) - f(x_0))/(x_n - x_0) \leq 0$ since $x_0 < x_n$. Therefore $f'(x_0) \leq 0$ by Preservation of Inequalities (Theorem 2.7). Now suppose that (x_n) is a sequence in (a, b) such that $x_n \to x_0$ and $a < x_n < x_0$ for all n. Since f is differentiable at x_0 we have

$$\lim_{n\to\infty} \frac{f(x_n) - f(x_0)}{x_n - x_0} = f'(x_0).$$

Since f has a local maximum at x_0 we have $f(x_n) - f(x_0) \leq 0$. Therefore $(f(x_n) - f(x_0))/(x_n - x_0) \geq 0$ since $x_n < x_0$. Therefore $f'(x_0) \geq 0$ by Preservation of Inequalities (Theorem 2.7). Therefore $f'(x_0) = 0$. ■

Note: Similarly, if $f\colon (a, b) \to \mathbb{R}$ is differentiable on (a, b) and f has a local minimum at x_0 then $f'(x_0) = 0$.

Example: Suppose that $f\colon [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b). Since f is continuous on $[a, b]$, f attains its maximum at some point $x_0 \in [a, b]$. Either $x_0 = a$, $x_0 = b$ or $x_0 \in (a, b)$ and $f'(x_0) = 0$.

Theorem 6.17 (Rolle's Theorem): Suppose that $f\colon [a, b] \to \mathbb{R}$ is continuous on $[a, b]$ and differentiable on (a, b). If $f(a) = f(b)$ then $f'(c) = 0$ for some $c \in (a, b)$.

Proof: Since f is continuous on $[a, b]$ it attains its maximum at $x_0 \in [a, b]$ and its minimum at $x_1 \in [a, b]$. If $x_0 \in (a, b)$ then $f'(x_0) = 0$. If $x_1 \in (a, b)$ then $f'(x_1) = 0$. Otherwise $x_0, x_1 \in \{\, a, b \,\}$. But then $f(x_0) \leq f(x) \leq f(x_1)$ for all $x \in [a, b]$. Since $f(a) = f(b)$ the function f must be a constant. Therefore $f'(x) = 0$ for all $x \in [a, b]$. ■

Example: the following example shows that it is essential to have continuity at the end-points of the interval $[a, b]$. Let $f\colon [0, 1] \to \mathbb{R}$ be defined by

$$f(x) = \begin{cases} x & \text{if } 0 \leq x < 1, \\ 0 & \text{if } x = 1. \end{cases}$$

Then f is differentiable on $(0, 1)$ (and hence continuous on $(0, 1)$). We have $f(0) = f(1)$. But $f'(x) = 1 \neq 0$ for all $x \in (0, 1)$. Note that f is not continuous at the right end-point of $[0, 1]$.

Proposition 6.18 (Mean Value Theorem): Let $f\colon [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b). Then there exists $c \in (a,b)$ such that

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Proof: Let $g\colon [a,b] \to \mathbb{R}$ be the function defined by

$$g(x) = f(x) - \left(f(a) + \frac{f(b) - f(a)}{b - a}(x - a)\right).$$

Then g is continuous on $[a,b]$ and differentiable on (a,b) by the hypothesis on f. We have $g(a) = f(a) - f(a) = 0$ and $g(b) = f(b) - f(a) - (f(b) - f(a))(b - a)/(b - a) = 0$. Therefore, by Rolle's Theorem, there exists $c \in (a,b)$ such that $g'(c) = 0$. Therefore

$$0 = f'(c) - \frac{f(b) - f(a)}{b - a}$$

which implies the result. ■

Note: It is easy to see that Rolle's Theorem and the Mean Value Theorem are equivalent to one another. We've just seen that Rolle's Theorem implies the Mean Value Theorem. Conversely, assuming the Mean Value Theorem, we see that if $f\colon [a,b] \to \mathbb{R}$ satisfies the hypotheses of Rolle's Theorem, then by the Mean Value Theorem there exists $c \in (a,b)$ such that $f'(c) = (f(b) - f(a))/(b - a)$. But $f(b) = f(a)$ from which Rolle's Theorem follows. Less obviously, both Rolle's Theorem and the Mean Value Theorem are equivalent to the Least Upper Bound Property of $\mathbb{R}$, Axiom III. We will see this shortly, once we have proven some corollaries to the Mean Value Theorem.

Corollary 1: Suppose $f\colon I \to \mathbb{R}$ is differentiable on I, where I is an interval. If $f'(x) = 0$ for all $x \in I$, then the function f is constant, i.e. there exists $C \in \mathbb{R}$ such that $f(x) = C$ for all $x \in I$.

Proof: Recall that I an interval means that if $x < y$ belong to I, then $z \in I$ if $x < z < y$. Let $x, y \in I$. We will prove that $f(x) = f(y)$. Since this is true for all $x, y \in I$ it follows that the function f is constant. Without loss of generality $x < y$. Since f is differentiable on I it follows that $f\colon [x,y] \to \mathbb{R}$ is continuous on $[x,y]$ and differentiable on (x,y). Therefore, the Mean Value Theorem applies and we see that there exists $c \in (x,y)$ such that

$$\frac{f(y) - f(x)}{y - x} = f'(c) = 0.$$

Therefore $f(x) - f(y) = 0$, i.e. $f(x) = f(y)$. ■

Note: It is important to note that Corollary 1 is only true if I is an interval. For example, if $A = (0,1) \cup (2,3)$ then the function $f\colon A \to \mathbb{R}$ defined by $f(x) = 1$ if $x \in (0,1)$ and $f(x) = 2$ if $x \in (2,3)$ satisfies $f'(x) = 0$ for all $x \in A$ but f is not constant. But A is not an interval.

Corollary 2: Suppose that $f\colon I \to \mathbb{R}$ is differentiable on I, where I is an interval. If $f'(x) > 0$ for all $x \in I$ then f is strictly increasing on I, i.e. $x < y \implies f(x) < f(y)$.

Proof: Suppose $x, y \in I$ and $x < y$. Then $f\colon [x,y] \to \mathbb{R}$ is continuous on $[x,y]$ and differentiable on (x,y). Therefore, by the Mean Value Theorem, there exists $c \in (x,y)$ such that

$$\frac{f(y) - f(x)}{y - x} = f'(c) > 0.$$

Since $y - x > 0$ it follows that $f(y) - f(x) > 0$, i.e. $f(x) < f(y)$. Hence f is strictly increasing on I. ■

Note: There are several variations on this corollary which are all proven in exactly the same way.

- if $f'(x) \geq 0$ for all $x \in I$ then f is increasing on I;
- if $f'(x) < 0$ for all $x \in I$ then f is strictly decreasing on I;
- if $f'(x) \leq 0$ for all $x \in I$ then f is decreasing on I.

Recall that f is *increasing* on I if $x < y \implies f(x) \leq f(y)$; f is *strictly decreasing* on I if $x < y \implies f(x) > f(y)$; and f is *decreasing* on I if $x < y \implies f(x) \geq f(y)$.

Note: With Corollary 1 in hand we can prove the equivalence of the Mean Value Theorem with Axiom III. Suppose that F is an ordered field, and the Mean Value Theorem holds on F. Let $A \subset F$ be a non-empty subset of F which is bounded above, but does not have a least upper bound. Define a function $f\colon F \to F$ by $f(x) = 1$ if x is an upper bound for A and $f(x) = 0$ otherwise. Let $x \in F$. If x is an upper bound for A, then there exists $y < x$ such that y is an upper bound for A since A does not have a least upper bound. Let $I = (y, \infty)$. If $z \in I$ then z is an upper bound for A. Hence $f(z) = 1$ for all $z \in I$. Suppose that x is not an upper bound for A. Then there exists $y > x$ such that $y \in A$. Let $I = (-\infty, y)$. If $z \in I$ then z is not an upper bound for A since $z < y$ and $y \in A$. Therefore $f(z) = 0$ for all $z \in I$. Therefore, for all $x \in F$, there exists an interval I containing x on which f is constant. Therefore f is differentiable on F with $f'(x) = 0$ for all x, but f is not equal to a constant. This contradicts Corollary 1 above.