2250: Elementary proofs of big theorems

The first theoretical result is the Intermediate Value Theorem (IVT) for continuous functions on an interval.

Theorem: If f is continuous on then interval I, then the set of values f assumes on I is also an interval. I.e. if a,b are points in I, then any number between f(a) and f(b) is also a value of f, taken at some point between a and b.

proof: We assume $f(a) < 0$ and $f(b) > 0$, and try to find c with $f(c) = 0$. We assume that every infinite decimal represents a real number.

Lemma: If f is continuous at c and f changes sign on every interval containing c, then $f(c) = 0$.

proof: This is the contrapositive of the fact that if f is positive (or negative) at c and continuous at c, then f is positive (or negative) on some interval containing c. This is immediate from the definition of continuity, and is an important exercise in understanding that definition. **QED.**

Thus it suffices to find a real number c such that f changes sign on every interval containing c. Assume $[a,b] = [0,1]$.

Since $f(0) < 0$ and $f(1) > 0$, then f changes sign on some interval of form $[r/10, (r+1)/10]$. let c start out as the decimal .r.

Then since $f(r/10) < 0$ and $f((r+1)/10) > 0$, f changes sign on some interval of form $[(10r+s)/100, (10r+s+1)/100]$. Then c continues as the decimal .rs.

Continuing in this way forever, we obtain an infinite decimal, i.e. a real number c = .rs....., in the interval [0,1], such that f changes sign on every interval containing c. Hence $f(c) = 0$. **QED.**

We know the continuous image of an open bounded interval may be neither open nor bounded. But the next big theorem says that the continuous image of a closed bounded interval is also closed and bounded. We do it in two steps.

Theorem: If f is a function which is continuous everywhere on a closed bounded interval [a,b], then f is bounded there.

proof: We prove it by contradiction, i.e. assuming f is unbounded leads to finding a point where f is not continuous.

Lemma: A function which is continuous at c, is also bounded on some interval containing c.

proof: This is immediate from the definition of continuity. E.g. if $e = 1$, by continuity of f at c, there is an interval I containing c where the values of f lie between $f(c)-1$ and $f(c)+1$. Thus f is bounded on I. **QED.**

Hence it suffices to show that if f is unbounded on [a,b], then there is a point c of [a,b] such that f is unbounded on every interval containing c.

Assume $[a,b] = [0,1]$, and that f is unbounded on [0,1]. Then there is some interval of form $[r/10, (r+1)/10]$ where f is unbounded. Start out the decimal c as .r.

Then there is some interval of form $[(10r+s)/100, (10r+s+1)/100]$ where f is unbounded. Continue the expansion of c as the decimal .rs.

Continuing forever, we construct an infinite decimal c = .rs......, in the interval [0,1], such that f is unbounded on every interval containing c. Thus f is not continuous at c. **QED.**

Theorem: If f is continuous on the closed bounded interval [a,b], then f assumes a maximum value there.
proof: We know the set of values f takes on [a,b] is a bounded interval. If not closed it has form (c,d) or [c,d) or (c,d]. If of form [c,d) say, then the continuous function 1/(f(x)-d) is unbounded on [a,b], contradiction. **QED.**

Longer, more painful argument:
We know f has an upper bound on [a,b]. We want a smallest one.
Lemma: Every non empty bounded set S has a smallest upper bound c.
proof: The proof is trivial if S is a finite set, so assume I particular that S does not consist only of the point 0. Then S is a non empty bounded infinite subset say of [0,1], then 1 is an upper bound for S but 0 is not, so there is some number of form r/10 which is not an upper bound of S, but such that (r+1)/10 is an upper bound. Let the decimal c start out as .r.

Then there is some number of form 10r+s which is not an upper bound but such that (10r+s+1)/100 is an upper bound. Continue c as .rs.

Continuing forever, we construct a real number c = .rs...., such that no smaller number is an upper bound, but every larger number is. Hence c is the smallest upper bound for S. **QED.**

Now if K is the smallest upper bound of all the values of f on [a,b], it will suffice to show f assumes the value K on [a,b]. But if not, then g(x) = 1/(K-f(x)) would be both continuous and unbounded on [a,b], which contradicts the previous result. **QED.**

A similar criterion for open intervals is this:
Cor: If f is continuous on (a,b), and f(x) approaches plus infinity as x approaches a, and also as x approaches b, then f has a global minimum on (a,b). (Apply the previous result to a suitable closed interval of form [a-e, b-e].)

For more general max/min problems on open intervals, it helps to have a little more understanding of how the derivative of a function affects the behavior of the graph . For some reason the following simple principle seems not to be stated in standard books.

Theorem: A continuous function on an interval cannot change direction except at a critical point. I.e. f is strictly monotone on any interval not containing a critical point.
proof: If has no critical points on an interval, then f is not constant. If f is not monotone then the IVT implies that f takes the same value twice on that interval, say f(a) = f(b) for some points a < b in the interval. Then f has both a maximum and a minimum on the interval [a,b] and since f is not constant, one of these extrema occurs at some c with a < c < b. Then c must be a critical point, as we know. Since there are no critical points, in fact f is monotone. **QED.**

This gives us many easily verifiable criteria for finding maxima and minima on open intervals. Here is a typical one.
Cor: If f is differentiable on (a,b) and has only one critical point c in that interval, then f(c) is a global minimum for f on (a,b) provided there are points u, v, with a< u< c < v < b, i.e. u and v are on either side of the critical point c, and f(c) < f(u) and f(c) < f(v).
proof: f is monotone and greater somewhere on each side of c, hence greater everywhere. **QED.**

More generally if f has a finite number of critical points $c1,\ldots.cn$, and f is higher somewhere to the left of the first critical point c_1 (higher than at c_1), and also higher somewhere to the right of the last one cn, then f has a global minimum at some critical point.

Finally we have:

Cor: If f' = 0 zero everywhere on an interval I, then f is constant on I.

proof: By the argument in the proof of the previous theorem, if a differentiable function h takes the same value at two different points, then h' = 0 somewhere in between. Applying this to a difference h = f-g, shows two differentiable functions f,g taking the same value at two points must have the same derivative somewhere in between. Since for any a,b in I, f agrees at a and at b with the linear function passing through (a,f(a)), (b,f(b)), whose slope is everywhere equal to ([f(b)-f(a)]/[b-a]), the graph of f must have that same slope at some c between a and b. In particular if f(a) differs from f(b), there is a c with $a < c < b$ and f'(c) = ([f(b)-f(a)]/[b-a]), which is not zero. **QED.**

Cor: If f, g have the same derivative everywhere on an interval, then f differs from g by a constant on that interval.

proof: Since (f-g)' is zero everywhere, (f-g) must be constant. **QED.**

Cor: If f has no critical points on an interval (a,b) and if $f'(c) > 0$ at some point c of (a,b), then f is strictly increasing on (a,b).

proof: By definition of the derivative, f is increasing at c, and f is monotone on (a,b), so f is increasing everywhere on (a,b). **QED.**