Math 2310H: Several versions of continuity

Recall the useful shorthand notations: "$\forall$" means "for all", and "$\exists$" means "for some", or "there exists...such that", "$\Leftrightarrow$" means "if and only if", and "$\Rightarrow$" means "implies" or "if...then".

Continuity on an interval:
Then we have the familiar definition of continuity of a function f at every point of an interval I.

Definition: A function f defined on an interval I, is continuous at every point of I $\Leftrightarrow$ ($\forall$ a in I)($\forall\ \varepsilon>0$)($\exists\ \delta>0$):($\forall$ x in I)($|x-a|<\delta \Rightarrow |f(x)-f(a)|<\varepsilon$).

In English this says: f is a continuous on I if and only if, for every point a of I, and every positive number ε, there is a corresponding positive number $\delta = \delta(a,\varepsilon)$, [i.e. δ may depend on both a and ε], such that, for every x in I, if x is within a distance δ of a, then f(x) is within a distance ε of f(a). Very, very informally, this means: for numbers x close to a, the values f(x) are close to f(a). But it is much more precise than this: i.e. you can make the values f(x) as close to f(a) as you want, just by making the number x close enough to a.

Example: The function f(x) = 1/x, is continuous on the interval
$(0, +\infty)$.
proof: Given a >0, and $\varepsilon>0$, we must find δ so that $|x-a|<\delta$ implies that $|1/x - 1/a| < \varepsilon$. To do this we set x = a+h, where $|h| = |x-a| < \delta$, and ask how small δ needs to be so that $|1/x - 1/a| < \varepsilon$. I.e. we set
$|1/(a+h) - 1/a| < \varepsilon$, and try to see how small we need h to be.

Since 1/x will be extremely large if x is near zero, we start by asking that $|h| < a/2$, in order to keep x away from 0. Then note that $|a+h| \leq |a|-|h|$, by the triangle inequality, so since $|h| < |a/2|$, we have $-|h| > -|a/2|$, hence after adding $|a|$ to both sides, we get $|a|-|h| > |a|-|a/2| = |a/2|$.

Now simplify the expression $|1/(a+h) - 1/a| = |h/[a(a+h)]|$, and we see that
$|1/(a+h) - 1/a| = |h/[a(a+h)]| < \varepsilon$, if and only if $|h| < \varepsilon\,|a||a+h|$. Since $|a+h| \geq |a| - |h| \geq |a/2|$, we claim it will suffice to have $|h| <$ both $|a/2|$ and $\varepsilon\,|a||a/2|$.

I.e. if $|h| = |x-a| <$ both $|a/2|$ and $\varepsilon\,|a||a/2|$, then the reasoning above shows that $|1/(a+h) - 1/a| = |h/[a(a+h)]| < |h/[a(a/2)]| < \varepsilon\,|a||a/2|/|a||a/2| = \varepsilon$.

Thus given ε, choose $\delta < |a/2|$ and also $\delta < \varepsilon\,|a||a/2|$. Then $|x-a| < \delta$, will imply that $|f(x)-f(a)| < \varepsilon$. **QED.**
OK, This was kind of a pain. In fact I remember after 40 years that this proof looked hard in my freshman class too, but we got it, anyway. If it makes you feel better, this is probably the hardest δ, ε proof.

Note: In this proof, $\delta = \min\{|a|/2, \varepsilon\,|a||a/2|\}$ depends on both a and ε.

Remember: to negate a quantified statement, you just change all the quantifiers to their opposites and then negate the assertion inside the quantifiers.

Hence to say f is not continuous on I, means:
$(\exists \text{ a in I})\ (\exists\ \varepsilon > 0)(\forall\ \delta > 0){:}(\exists \text{x in I})(|x\text{-}a| < \delta$ does not imply $|f(x)\text{-}f(a)| < \varepsilon)$.

Now rememeber, A imples B, menas that B is true whenever A is true, so A does not imply B means that A is true "but" (i.e. and) B is not.

So the previous statement becomes:

To say f is not continuous on I, means:
$(\exists \text{ a in I})\ (\exists\ \varepsilon > 0)(\forall\ \delta > 0){:}(\exists \text{x in I})(|x\text{-}a| < \delta$ and $|f(x)\text{-}f(a)| \geq \varepsilon)$.

In English this says, there is some point a, and some distance ε, such that f(x) cannot be guaranteed to be as close to f(a) as ε, no matter how close we take x to a. I.e. for some ε, there are x's with $|f(x)\text{-}f(a)| \geq \varepsilon$ in every interval around a.

Here is an example of a discontinuous function:

Let f(t) = sec(t) for $t \neq \pi/2$ on the interval $[0,\pi]$, and $f(\pi/2) = 0$. Then f is not continuous on $[0,\pi]$ because it is not continuous at $\pi/2$. I.e. let $\varepsilon = 1$. Then I claim, no matter how close we get to $\pi/2$, we can always find an x that close, but with f(x) further from $f(\pi/2) = 0$ than 1. By looking at the circle, note that on every interval around $\pi/2$, there are points t where $\cos(t) < 1$, hence where $\sec(t) = 1/\cos(t) > 1$. That proves it.

Note that this is an "infinite discontinuity" like the one on page 37 of our book.

Uniform continuity on an interval:
If we can choose δ to depend only on ε, and not on a, then we say f is "uniformly continuous" on the interval I. In symbols this says:

Definition: A function f defined on an interval I, is uniformly continuous on I $\Leftrightarrow$
$(\forall\ \varepsilon > 0)(\exists\ \delta > 0){:}(\forall \text{ a in I})(\forall \text{x in I})(|x\text{-}a| < \delta \Rightarrow |f(x)\text{-}f(a)| < \varepsilon)$.

The following big theorem is proved in analysis courses and also in my handout from math 2300H:

Theorem: If the interval I is closed and bounded, i.e. if I = [a,b] where a and b are finite numbers, then every function f which is continuous at every point of I, is also uniformly continuous on I.

Example: Let f(x) = sqrt(x), on the interval $[0,+\infty)$. Then we claim f is uniformly continuous on $[0,+\infty)$.

To prove it , given ε, we must find a δ that works in the definition of continuity at every point a of the interval $[0,+\infty)$. Using the insight we got from working this problem in class, we start by finding a δ that works at a = 0. I.e. we want $|x-0| = x < \delta$ to imply that $|\text{sqrt}(x) - \text{sqrt}(0)| = \text{sqrt}(x) < \varepsilon$. But squaring both sides gives that if $x < \varepsilon^2$, then $\text{sqrt}(x) < \varepsilon$. So just take $\delta = \varepsilon^2$. Now we claim this same δ works at every positive a. I.e. we claim that if a , x > 0, and $|x-a| < \varepsilon^2$, then $|\text{sqrt}(x) - \text{sqrt}(a)| < \varepsilon$.

OK, I'm not claiming I expect you to be able to do this, but I did sit down and knock this out on scratch paper. You tell me how it seems to you:
First of all, I will assume that a < x, without loss of generality, i.e. since we are talking about any two points a, x >0, we can name the larger one x.

Then I want to show that if $x-a < \varepsilon^2$, then $\text{sqrt}(x) - \text{sqrt}(a) < \varepsilon$. But
$x-a < \varepsilon^2$, implies that $x < a+\varepsilon^2$, so taking square roots of both sides this says that $\text{sqrt}(x) < \text{sqrt}(a+\varepsilon^2)$. Now I claim that $\text{sqrt}(a+\varepsilon^2) \leq \text{sqrt}(a) + \varepsilon$. I.e. in general, I claim that for positive numbers C,D, that
$\text{sqrt}(C+D) \leq \text{sqrt}(C) + \text{sqrt}(D)$. to see this note that by squaring both sides it suffices to show that
$C+D \leq [\text{sqrt}(C)+\text{sqrt}(D)]^2 = C + D + 2\text{sqrt}(C)\text{sqrt}(D)$, but this is obvious since $2\text{sqrt}(C)\text{sqrt}(D) \geq 0$. So indeed we have $\text{sqrt}(a+\varepsilon^2) \leq \text{sqrt}(a) + \varepsilon$. Combining this with what we had before, we see that:

$\text{sqrt}(x) < \text{sqrt}(a+\varepsilon^2) \leq \text{sqrt}(a) + \varepsilon$. So $\text{sqrt}(x) - \text{sqrt}(a) < \varepsilon$, as desired.
QED.

Now do not be discouraged by the complexity of this. These arguments with inequalities, and δ, ε are considered the hardest arguments in elementary math, and I am not expecting you to master them. That will come in a much later class, at the 3000 or 4000 level, but it does not hurt to see a few now, done for you.

Lipschitz continuity:
Even if a function is uniformly continuous on I, it can still be very hard to actually find the δ that goes with a given ε, as perhaps is believable from the argument above. There is a special case that comes up very often however where this is easy. For this purpose we introduce the idea of "Lipschitz continuity".

A function is Lipschitz continuous on an interval I if not only can the
δ be chosen to depend only on ε, but if in fact δ can be chosen to be a fixed constant multiple of ε. More precisely:

Definition: A function f is "Lipschitz continuous" on the interval I, if and only if: $(\exists K >0)(\forall \varepsilon >0):(\forall$ a in I$)(\forall$ x in I$)(|x-a|<\varepsilon/K \Rightarrow |f(x)-f(a)|<\varepsilon)$.

This implies that f is uniformly continuous in I by taking $\delta = \varepsilon/K$ in the definition of uniform continuity.

Although Lipschitz continuity is stronger than uniform continuity and hence also stronger than continuity, still many functions are Lipschitz continuous, and it is often easier to recognize, and to prove, Lipschitz continuity than the other kinds. To see this, we restate the definition as follows.

Lemma: f is Lipschitz continuous on I if and only if the set of absolute values of secant slopes: $\{ |f(x)-f(a)|/|x-a|$, for all a,x, in I$\}$ is bounded. i.e. if and only if there is a number K>0 such that: for all a,x in I,
$|f(x)-f(a)|/|x-a| \leq K$, i.e. if and only if for all a,x in I, $|f(x)-f(a)| \leq K|x-a|$.
proof: If such a K exists, then whenever $0 < |x-a| < \varepsilon/K$, we conclude that
$|f(x)-f(a)| = (|x-a|)(|f(x)-f(a)|/|x-a|) \leq |x-a|\, K < (\varepsilon/K)K = \varepsilon$. So K is a Lipschitz constant. It is obvious that also when $0 = |x-a|$, then $|f(x)-f(a)| = 0 < \varepsilon$. On the other hand if K>0 is a Lipschitz constant for f on I, we claim that $|f(x)-f(a)|/|x-a| \leq K$ for all a,x in I. For if there did exist a,x in I such that $|f(x)-f(a)|/|x-a| > K$, then $|f(x)-f(a)|/K|x-a| > 1$. Then there is some number $c > 0$ such that $|f(x)-f(a)|/K|x-a| > 1+c$. Then if we take
$\varepsilon = K|x-a|(1+c)$, we have $|x-a| < |x-a|(1+c) = \varepsilon/K$, and yet
$|f(x)-f(a)| > \varepsilon = K|x-a|(1+c)$. This contradicts the assumption that K was a Lipschitz constant for f on I. **QED.**

Thus any number K larger than the absolute values $|\Delta y/\Delta x|$ of all secant slopes for f on I, serves as a Lipschitz constant for f on I. This is often taken as the definition of Lipschitz continuity. Such numbers are sometimes easy to find for differentiable functions f, as we will see next.

Lemma: If f is a differentiable function on the interval I, and if the absolute values $|f'(x)|$ of the derivative are bounded by K on I, then K is a Lipschitz constant for f on I. I.e. if there is a number K>0 such that for all x in I, we have $|f'(x)| \leq K$, then f is Lipschitz on I and K is a Lipschitz constant.
Proof: Since every derivative f'(x) is a limit of difference quotients $\Delta y/\Delta x$, the same is true for their absolute values. Hence if K is less than some number of form $|f'(x)|$, then K is also less than some number of form $|\Delta y/\Delta x|$. Thus, if K is <u>not</u> smaller than any difference quotient $|\Delta y/\Delta x|$, i.e. if $K \geq$ all numbers $|\Delta y/\Delta x|$, then also $K \geq |f'(x)|$ for all x in I.

Conversely, if $K \geq |f'(x)|$ for all x in I, then we claim that also $K \geq |\Delta y/\Delta x|$, for all difference quotients $\Delta y/\Delta x = (f(x_1)-f(x_0))/(x_1-x_0)$, with x_0, x_1 in I. This follows from the famous "Mean value theorem" of differential calculus. I.e. recall if $x_0 < x_1$ in I, and f is differentiable on $[x_0,x_1]$, theen there is some point x with $x_0 < x < x_1$, such that
$(f(x_1)-f(x_0))/(x_1-x_0) = f'(x)$. Hence the set of difference quotients $(f(x_1)-f(x_0))/(x_1-x_0)$, is actually a subset of the set of derivatives f'(x). So if K is $\geq$ than all derivatives, it must also be $\geq$ all difference quotients.
QED.

Example: If $f(x) = x^{1/3}$, on the interval $[1,+\infty)$, then f is Lipschitz there. Note that f is differentiable on this interval with derivative

$f'(x) = (1/3)x^{-(2/3)} = (1/3)/x^{(2/3)}$ And that for x in $[1,+\infty)$,
we have $x^{(2/3)} \geq 1$, hence $f'(x) \leq 1/3$. So K = 1/3 serves as a Lipschitz constant on this interval.

Recall another big theorem from first semester calculus, also proved in my handout notes.
MMV: (Max - min value theorem): If f is continuous on the clsoed bounded interval I = [a,b], then there is a number K>0, such that for all x in I, we have $|f(x) \leq K$.

Recall also that a function with a continuous derivative is called a "C^1" function, or "a function of class C^1". We may say for short simply that "f is C^1", using C^1 as an adjective modifying f.

Then we have the following corollary of our previous results:

Corollary: Any function f which is C^1 on a closed bounded interval I = [a,b], is also Lipschitz continuous on that interval, and any bound K for the absolute values $|f'(x)|$ of the derivative of f on I, serves as a Lipschitz constant.

Example: If f(x) = sin(x), on the interval $(-\infty,+\infty)$, then since the derivative sin'(x) = cos(x) is bounded between -1 and 1 everywhere, f(x) =sin(x) is Lipschitz on the whole real line, with Lipschitz constant 1. In particular for any pair of points a,b, we have $|\sin(a)-\sin(b)| \leq |b-a|$.

Example: If $f(x) = x^3$, then f is Lipschitz continuous on any closed bounded interval [a,b] but not on the whole real line. I.e. f is C^1, so it is Lipschitz on a closed bounded interval, but since $f'(x) = 3x^2$, the derivatives f'(x) are not bounded on the whole real line, so f is not Lipschitz there.

Example: $f(x) = x^{1/2}$ is uniformly continuous on the interval [0,1] but not Lipschitz there. Since f is continuous on this closed bounded interval it is uniformly continuous. But since f is diferentiable with derivative $f'(x) = (1/2)/x^{1/2}$ on the interval (0,1], with unbounded derivative there, f is not Lipschitz on the half open interval (0,1]. Hence it cannopt be Lipschitz on the entire interval [0,1]. I.e. if the set of secant slopes is already unbounded on the smaller interval (0,1], it is also unbounded on the larger interval [0,1].

Remark: The interest of the concept of Lipschitz continuity is further enhanced for us because if f is any Riemann integrable function on a closed bounded interval [a,b], then the indefinite integral of f is Lipschitz continuous on [a,b]. Hence Lipschitz continuity is a necessary condition for a function to be an indefinite integral of an integrable function on a closed bounded interval.