Buck Chapter 2: Continuity [pp 55-96]

Buck Chapter 2: Continuity **1**
2.2 Basics. 1
2.3 Approximation and Uniform Continuity 3
2.4 Properties of Continuous Functions 8
2.5 Limits of functions 11
2.6 Discontinuities 12
2.7 Mean value theorems and L'Hospital. 14

2.1 is a 1 paragraph intro.

Note: "continuity" is a property that *functions* can have, it is not a property of sets or of sequences.

2.2 Basics. Definition p 56 of "f(p) continuous at a point p_0" : For any given $\varepsilon > 0$, you can find a δ small enough so any p in the δ-ball around p_0 meets the requirement $| f(p) - f(p_0) | < \varepsilon$. Shrinking the δ ball in the domain shrinks the image of the δ ball in the range. If you have continuity at all points p_o in some domain region D, then you have "f(p) continuous on D". Note that the required small δ will in general be a function of the location of point p_0, as shown in Fig 2.1 p 56.

Theorem 1: Continuity of f(p) implies that, if $p_n \rightarrow p$ in the domain, $f(p_n) \rightarrow f(p)$ in the range. One says then that "continuity preserves convergence".

Theorem 2 (converse of above). If for all sequences $p_n \rightarrow p$ you have $f(p_n) \rightarrow f(p)$, then f continuous.

Comment: Bucks are tying the notion of continuity of a function $f:D \rightarrow \mathcal{R}$ to the notion of a sequence of points p_n in domain D, and another sequence of points $f_n = f(p_n)$ in range $\mathcal{R}$. The sequence was a major topic of Chapter 1 and thus had to be treated first.

Example: Let D be E^2 and let f(p) be $f(x,y) = xy^2/(x+y)$ AND $f(0,0) = 0$. Bucks consider two sequences in the domain of f, both of which $\rightarrow (0,0)$. For the first called p_n you get $\lim f(p_n) = 0$ so things seem OK since $f(0,0) = 0$. For the second sequence q_n they find $\lim f(p_n) = 1/2$, but the number 1/2 is not $f(p_0) = 0$! Thus, according to Theorem 2, since not all sequences $t_n \rightarrow p$ have $f(t_n) \rightarrow f(0)$, it must be that f is not continuous at $p_0 = (0,0)$. Just looking at the functional form, you can expect trouble at this point.

In Definition 1' we get an alternate version of the definition of continuity in terms of Fig 2.2. The role of the δ ball is played by set U in the domain, and the image by set V in the range. The idea is this: for any V in the range, no matter how small, you can find a U in the domain whose entire image lies inside V.
They then think of set U as the **inverse image** or **pre-image** of V. You could write $f(U) = V$ and $f^{-1}(V) = U$.

Theorem 3. Consider sets D and f(D) under some f. Let V be an arbitrary open set in f(D), the range. If for *any* such V you find that the inverse image $f^{-1}(V)$ is an open set in the domain, then f is continuous on D. The converse is also true:

Let's try to restate this better.

Theorem 3. [Any V open in range $\Rightarrow$ $f^{-1}(V)$ open in domain] $\Leftrightarrow$ f is continuous on domain

Here is my little word proof. Suppose V were open in range, but $U = f^{-1}(V)$ were NOT open in domain. Then some point u in $U = f^{-1}(V)$ must lie on the boundary of U. If f were continuous, that boundary point in the domain would map into a boundary point v in V because Theorem 1 says convergence is preserved by a continuous function. But then V would not be open since if would contain a point on its boundary. Since we assumed it was open, it must therefore be true that U is also open.

Bucks then momentarily talk about a set of functions rather than points in E^n. Strangely, instead of referring to a "set of functions" they call it a "collection of functions". [set = collection]

Theorem 4. Consider the set of all the real-valued continuous functions f defined on D in the domain. The claim is that this set forms "**an algebra**". This is a term I never define in my Galois doc. It seems that the criteria for a set of functions being an algebra is this:

f,g in set $\Rightarrow$ f+g, fg and $\alpha f + \beta g$ are in the set

If you omit fg, then the set is a vector space, so it is the multiplication item that makes it an algebra. You can imagine that the proofs are easy, for example to show that fg is continuous if f and g are continuous. Division f/g is also included as long as $g \neq 0$.

Corollary. Since 1 and x and y are continuous functions, so are all multivariable polynomials like $f = 2 + 3x^3 + 2xy$.

They mention division of functions f/g on p 61, but this is not part of the algebra. The obvious claim is that f/g is continuous if f and g are, as long as you avoid places where g = 0. Well, maybe it is part of the algebra with this caveat.

From wiki:

> In other words, an algebra over a field is a set together with operations of multiplication, addition, and scalar multiplication by elements of the underlying field, that satisfy the axioms implied by "vector space" and "bilinear".[1]

You can also have an algebra over a ring instead of over a field.

Theorem 5: Continuity is preserved under functional concatenation, as in f(g(p)). That is to say, if f and g are continuous on their respective domains , then f*g is continuous as well. Bucks give no symbol for concatenation and do not use the word concatenation -- the refer to "a composite function". I think "chaining" of functions is also used by authors, leading to the "chain rule" of differentiation.

On page 63 Bucks ask: how do we know sin(x) is continuous? They claim that the usual circle picture argument is not solid, so you don't get "proof by picture" that sin(x) is continuous. It all depends on how you define sin(x), and perhaps the infinite series (a poly!) is better. They are just pointing out this little detail in passing. No doubt one of them worked on this issue at some time.

2.3 Approximation and Uniform Continuity.

Suppose you want function F to "approximate" function f. One way to do this is to require that the abs value $| f - F | < \varepsilon$ at every point in the domain. Called "uniform ε approx". Recall that the L_1 norm of a function is the integral of $| f |$ over the domain interval, and then the distance between two functions by the L_1 norm would be $\int | f - F | dx$, but this is NOT what we are talking about here. This is more of a max norm, saying max $| f - F | < \varepsilon$

Weierstrass Approx Theorem. [66] On a closed interval [a,b] you can ε approximate any continuous f by a polynomial.

Bucks do not prove this an concentrate instead on doing a simple linear approximation with straight line segments as in Fig 2-5. Notice that the function f is allowed to have a sharp corner and still be continuous at that corner (not differentiable of course). So I think Bucks on pages 66-67 are proving that you can fit any continuous f with a polygonal line approximation in this ε sense, but I skip this text.

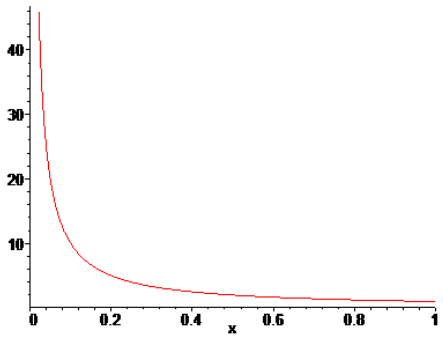
Uniform Continuity. [67] Idea is that you require $| f(p)-f(q) | < \varepsilon$ not just for one particular q in the domain (which would be regular continuity at point q), but you require $| f(p)-f(q) | < \varepsilon$ for all p and q in the domain. You then have continuity uniformly over the domain. The idea is that given any ε, the same δ has to work for all q's in D.

Temperature and prongs is a good example. For a fixed δ (prong separation on metal surface D), you get $\Delta T < \varepsilon$ no matter where you put *both* prongs in D. The temperature is then uniformly continuous on D.

Classic examples: consider $f(x) = 1/x$ on [0,1]. For a given ε, as you move toward x = 0 you require a smaller and smaller δ, so there is no δ>0 that works on the entire domain for a given ε, so not uniformly continuous. However, for example $f(x) = \sqrt{x}$ on [0,1] things ARE uniformly continuous. But this example also has an infinite slope at x = 0, so you would think that a single δ would not work. This shows that the slope argument is bogus. For $f(x) = \sqrt{x}$ they find a δ that works everywhere, though they have to consider two cases. The result is $\delta < \varepsilon^2/9$. So I learned something new here!

[68] Decoding text regarding non-uniform continuity of f(x) = 1/x.

As shown below, the Bucks have summarized a lot of work in a tiny amount of text, and I will here attempt to fill in the large gaps with description. First, here is a plot showing that infinite slope at x = 0,



Our interest is in finding δ such that, when $|x-x_0| < \delta$, we have $|1/x-1/x_0| < \varepsilon$ for any ε, and where δ does not depend on the point x_0. Can this be done on the entire closed interval [0,1] ?

Step 1: Pick point $x_0 > 0$ and fix it. In the continuity definition, our real interest is in very small ε and very small δ. We crank down on ε, and hope to always find some small $\delta>0$ that works. So assume that ε has been cranked down enough to force $\delta < x_0/2$ for our choice of x_0.

Step 2: Then $|x-x_0| < x_0/2$ and $(x-x_0)^2 < (x_0/2)^2$ which says $x^2+x_0^2- 2xx_0 < (1/4)x_0^2$ or $x^2 - 2xx_0 + (3/4)x_0^2 < 0$. $B^2 - 4AC = (2x_0)^2 - 3x_0^2 = x_0^2$ so the solution to $x^2 - 2xx_0 + (3/4)x_0^2 = 0$ is $x = (2x_0 \pm x_0)/2 = x_0/2$ and $3x_0/2$. These are the two intercepts of the up-curving parabola $x^2 - 2xx_0 + (3/4)x_0^2 = 0$ and we have $x^2 - 2xx_0 + (3/4)x_0^2 = 0 < 0$ between these two points, so we then find that our assumption $\delta < x_0/2$ and $|x-x_0| < \delta$ leads to this

$$x_0/2 < x < 3x_0/2 .$$

The left inequality is the main interest and it says $x > x_0/2$ and since $x_0 > 0$, also $xx_0 > x_0^2/2$.

Step 3: Now consider $|f(x)-f(x_0)| = |1/x-1/x_0| = |(x_0- x)/(xx_0)| = |(x_0-x)| / |xx_0|$. We know that in fact $|(x_0-x)| = |x-x_0| < \delta$ is required. Therefore

$$|f(x)-f(x_0)| = |(x_0-x)| / |xx_0| < \delta /(xx_0) .$$

But we showed in Step 2 that $xx_0 > x_0^2/2$ and therefore $1/(xx_0) < 2x_0^2$ so then

$$|f(x)-f(x_0)| = |(x_0-x)| / |xx_0| < \delta /(xx_0) < 2\,\delta\, x_0^2$$

We want $|f(x)-f(x_0)| < \varepsilon$, so to achieve this we can choose to make $2\,\delta\, x_0^2 < \varepsilon$ so then

$$|f(x)-f(x_0)| = |(x_0-x)| / |xx_0| < \delta /(xx_0) < 2\,\delta\, x_0^2 < \varepsilon$$

But $2\,\delta\, x_0^2 < \varepsilon$ says that $\delta < \varepsilon / (2x_0^2)$. This choice of a δ then meets our requirement at least for our particular $x_0 > 0$.

Conclusion at this point: $f(x) = 1/x$ is continuous at all points $0 < x_0 < 1$. So all the work above was to just show this fact, which seems totally obvious just looking at the picture.

Step 4: Set $\varepsilon = 1$ and try to find a δ that works in all of $[0,1000]$. We need $|x-x'| \le \delta$ for every x and x' in the interval, for our solution δ, because this is the definition of *uniform* continuity on page 67. Assume some solution δ exists. Then try these two particular points in the interval: $x = 1/n$ and $x' = \delta + 1/n$ where we think of n as being some large positive integer, so that x and x' are then always in the interval. Notice that for this δ and this x and this x' we have $|x-x'| = \delta$. I set the right end to 1000 just to make sure x' lies in the interval without a lot of extra chatter. Then

$$| f(x)-f(x') | = |1/x-1/x' | = | n - 1/(\delta+1/n)| = | n(\delta+1/n) - 1 |/(\delta+1/n) = n\delta /(\delta+1/n)$$

Can we always find n large enough to that RHS > 1 ? If so, then we fail the uniform continuity test. That would require

$n\delta /(\delta+1/n) > 1$?

$n\delta > (\delta+1/n)$?

$n^2\delta > (\delta n+1)$?

$\delta n^2 - \delta n - 1 > 0$? $\qquad\qquad$ $B^2 - 4AC = \delta^2 + 4\delta$

$n = (\delta \pm \sqrt{\delta^2 + 4\delta}\,)/(2\delta) = (1 \pm \sqrt{1 + 4/\delta}\,)/(2)$

So any n larger than $(1 + \sqrt{1 + 4/\delta}\,)/(2)$ makes $| f(x)-f(x') | < 1$. So no matter how small you make $\delta>0$, you can always find an n large enough that $| f(x)-f(x') | < \varepsilon = 1$.

So all is well, but what remains is the mysterious Buck "for example" comment at the end which seems wrong to me and to which I return below.

Question: Is our failure here associated with the left end of the interval? Yes, because we are choosing $x = 1/n$ and as n gets larger, x is impinging on this left endpoint, and really so is x'.

Buck mystery comment: Suppose we go ahead and pick $n > 1/\delta$ and $n>3$. Why does this result in $n\delta /(\delta+1/n) > 1$?

Lemma: Conditions can be written $n\delta > 1$ and $n\delta > 3\delta$.

$n\delta /(\delta+1/n) > 1$?

$n\delta > (\delta+1/n)$?

$n\delta - (\delta+1/n) > 0$?

Well, consider

$$n\delta > 3\delta$$

=>

$$n\delta - (\delta+1/n) > 3\delta - (\delta+1/n) = 2\delta - 1/n$$

So with their assumptions, we know that

$$n\delta - (\delta+1/n) > 2\delta - 1/n$$

But $\delta > 1/n$ is one of the conditions, so we know $2\delta > 2/n$ and then

$$n\delta - (\delta+1/n) > 2\delta - 1/n > 2/n - 1/n = 1/n > 0$$

and so indeed I have shown that, with their two assumptions, $n\delta - (\delta+1/n) > 0$ which is the condition that says $| f(x)-f(x') | > 1$. So I have verified their comment, but it does not seem very enlightening.

Comments on the Above. We first showed that $f(x) = 1/x$ is continuous on $0 < x < 1$. Next, we showed really that $f(x)$ is NOT uniformly continuous on $0 < x < 1$. The reason is that we were able to find two points x and x' in the range such that we could find no $\delta > 0$ which caused $| f(x)-f(x') | < \varepsilon = 1$. Our special points sort of tweeze together and approach the left (but open) end of the interval. Our negative proof was this: pick certain x and x' as functions of large n, and show that for any tiny $\delta>0$, we could find n large enough to violate the requirement $| f(x)-f(x') | < \varepsilon = 1$. We did NOT make the argument that because the slope at the left end approaches infinity, there is no δ what works.

[69] Decoding parallel text regarding uniform continuity of $f(x) = \sqrt{x}$.

This case really seems mysterious.

First, they show that this $f(x)$ is continuous on $0 < x < 1$. They first show (and I have verified) that

$$| f(x)-f(x') | \leq |x-x_0|/\sqrt{x_0}$$

Now if we require that $f(x)-f(x') | \leq \varepsilon$, then we can do this by choosing $|x-x_0|/\sqrt{x_0} < \varepsilon$ which means we do this by choosing $|x-x_0| < \varepsilon \sqrt{x_0}$. So our solution is $\delta = \varepsilon \sqrt{x_0}$ and then when $|x-x_0| < \delta$ we have

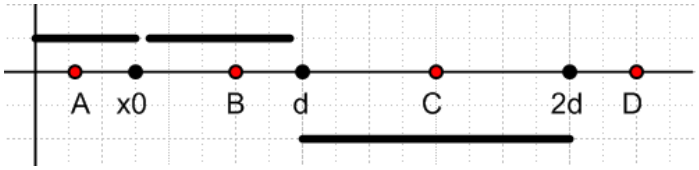
$$| f(x)-f(x') | \leq |x-x_0|/\sqrt{x_0} < \delta / \sqrt{x_0} = \varepsilon \sqrt{x_0} / \sqrt{x_0} = \varepsilon$$

So choosing any $\delta < \varepsilon \sqrt{x_0}$ will guarantee that $|f(x)-f(x_0)| \leq \varepsilon$. Thus, we have proven that $\sqrt{x}$ is continuous in the open interval $0 < x < 1$.

Now we are going to show that, in fact, $f(x) = \sqrt{x}$ is uniformly continuous on $0 < x < 1$ despite the fact that the slope is infinite at $x = 0$. This does indeed seem strange.

Step 1: First pick x_0 and nail it down. Bifurcate into two cases: $x_0 < \delta$ and $x_0 > \delta$.

Step 2: Assume the first case that $x_0 < \delta$, and we need $|x-x_0| < \delta$ as well. What do these two inequalities tell us? I had trouble showing the result analytically, but graphically consider:



where the red circles show possible values of x and d means δ. For which positions of x do we have $|x-x_0| < \delta$? The left bar shows the possible values of $|x-x_0|$ as point A is moved within its range. And similarly point B and point C. For any of these three positions, we have $|x-x_0| \leq \delta$. But position D is not allowed. Therefore our two inequalities imply

$$0 < x < 2\delta$$

Using this along with $x_0 < \delta$ Bucks show that $|f(x)-f(x_0)| \leq 3\sqrt{\delta}$. We hold on this point.

Step 3: Assume the second case that $x_0 > \delta$, and we need $|x-x_0| < \delta$ as well. Bucks then use earlier inequality to show again that in this case we also have $|f(x)-f(x_0)| \leq 3\sqrt{\delta}$.

Step 4: We therefore know that $|f(x)-f(x_0)| \leq 3\sqrt{\delta}$ no matter what the values x_0 and δ end up being. Therefore, in order to get $|f(x)-f(x_0)| \leq \varepsilon$, we select δ such that $3\sqrt{\delta} < \varepsilon$ which says $\delta < (\varepsilon/3)^2$. So, for any such δ, and for *any* x and x_0 in the interval, we get $|f(x)-f(x_0)| \leq \varepsilon$.

Comments on the Above. Obviously this proof must fail for f(x) = 1/x. The various inequalities won't be valid. But it is a little difficult to see exactly what makes one be non-uniform and the other be uniform. These are just examples. A completely general answer is provided by the next theorem which then let's you avoid all the work done in these example!!

Theorem 6. [69] If f is continuous on D, and if D is closed and bounded ("compact"), then f is uniformly continuous on D.

First, how does this apply to our two examples? In each case let D = [0,1] which is obviously closed and bounded.

f(x) = 1/x : This function is NOT continuous at x = 0, so the theorem says nothing -- I guess the function could or could not be uniformly continuous on D. We showed it is NOT uniformly continuous.

$f(x) = \sqrt{x}$: This function IS continuous at x = 0 and thus on D = [0,1]. The theorem then applies and we conclude that $\sqrt{x}$ is uniformly continuous on the interval.

$f(x) = x^p$: As long as $p \geq 0$, f(x) will be continuous at x = 0, and x^p will be uniformly continuous. I suspect that for any $p < 0$ it will not be uniformly continuous, but we have only shown that for p = -1, and the Theorem 6 has nothing to say on this question.

Comment: note that "bounded" refers to domain D, not to values of the function f(x)

They give a long proof of Theorem 6 which I skipped.

Reminder: if you define "compact1" as meaning the open subcover stuff, and "compact2" as set being bounded and closed, then compact1 = compact2 on E^n but not necessarily on other Hilbert Spaces.

They give a long proof which runs bottom p 69 down 2/3 of page 70. Within this discussion is a comment on how you form the contrapositive of a complicated statement, see top p 70. The proof makes use of the BW theorem regarding a cluster point existing in any bounded infinite set. An alternate proof is outlined in two Exercises on page 72 and that one uses the Heine-Borel.

Theorem 7. Consider f continuous on closed and bounded [a, b]. The domain is then compact2. Then it is possible to do the uniform ε approx by a polygonal line (a piecewise linear function).

They comment on how this theorem helps compute tables in books like A&S (1964). This general subject of **approximating things** was perhaps of more interest in 1965 than now. Of course all programs that compute functions now use such methods behind the black curtain, for example, Maple.

Consider 1/x on [0,1] I guess right away you have to go to (0,1] so you violate the theorem premise and maybe then it is NOT possible to do a linear ε fit to f = 1/x. I would have to try to do such a fit to see whether or not it would be possible, but I can imagine it might not be possible. Bucks do not give examples of violations of this theorem.

2.4 Properties of Continuous Functions

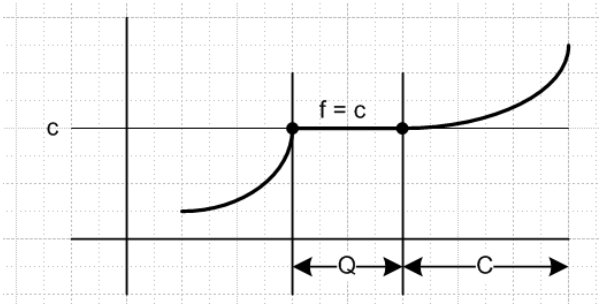
My own general comments: You can see, when $D \subset E^1 = R$, that it is probably wise to restrict your general study of functions to those that are "continuous". Otherwise you have points where the function is undefined or has 2 different values, and then you don't even have a function. So we have generalized this notion to $D \subset E^n$ and suspect that for these more general functions on the fancier vector space, we once again want to avoid functions which are poorly defined at certain points. Maybe in 2D as you approach some point, you get a different value for f(p) depending on the angle of approach. An analytic function I think has the same limit for all approach angles in x + iy space and so analytic functions are really then just continuous functions (have to check on this later). So when the historical developers of calculus like Weierstrass (as opposed to the founding fathers like Newton) started trying to "nail down" proofs and such, the notion of f(p) being continuous became important, and they came up with the ε and δ tools to define continuity in E^n instead of just staring at a continuous f(x) on E^1.

This section is basically a statement and proof of 6 theorems (8 thru 13), I will try to restate each one my way:

Theorem 8. Consider $f:E^n \rightarrow R$ and assume f is continuous on some D in E^n and p_0 is some point in D. Suppose you find that $f(p_0) > 0$. Then there must be a range ball B around p_0 such that $f(B) > 0$ for the entire ball (ball = neighborhood in Phil language).

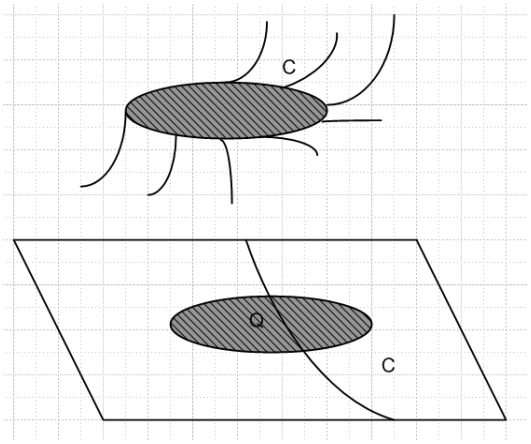
Theorem 9: Consider $f:E^n \rightarrow R$ and assume f is continuous on some D in E^n. Let C be that part of the domain D for which $f(C) > c$ and let Q by that part of D for which $f(C) = c$. The claim is that C is an open set and Q is a closed set, both "relative to D".

The "relative to D" part really just means you have to mentally truncate things to D within E^n in the case that D is finite. Here is a supporting picture



The dots indicate the very last value where $f(x) = c$. Set Q is closed since it includes the boundary (consisting of the two dots). Set C is open because it does not include the left boundary and it would be open if you go as far as you want off to the right. Intersecting this with some finite D would yield a boundary point, but such points don't count when you say "open relative to D".

Here is the same idea but with D in E^2



The oval is like a flat patch of roadway on a mountain slope, where the mountain goes up on the right side and down on the left side. The function $f > c$ on the right side which has support C in E^2 . The function $f = c$ on the patch, and on its edges. So now Q is again closed since it includes its boundary. And set C is open.

So far Bucks have talked about S being a bounded region in the domain. A **function f is bounded** if |f(S)| is bounded. That is to say, the image of S is bounded, so f is then "bounded on S". No big deal.

Comment: I think Bucks are always talking $f:E^n \to R$ with these theorems, so | | really does mean absolute value, not some multivariable norm.

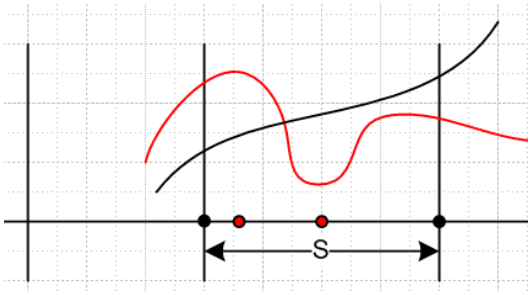
Theorem 10: If the domain set S is compact2 in E^n (closed and bounded) and if f is continuous on S, then f(S) is bounded in the range.

Go back to 1/x on [0,1]. The domain S is compact, but f is not continuous at the left end, and for that reason f(S) is not bounded in the range. Yes, $f(0) = \infty$, unbounded.

Theorem 11. If the domain set S is compact2 in E^n (closed and bounded) and if f is continuous on S, then f takes its min and max values in f(S). [For analytic function, min and max are on the boundary.]

Since f(S) is bounded, the min and max values must be finite, and the only question is whether these values are actually inside f(S). But since S is closed and f is continuous, we know f(S) is closed (a theorem Buck have not stated). Since continuous preserves convergence, if $p_n \to$ point on boundary in S, then $f(p_n) \to$ point on boundary in range, and that is why the image of a closed set must be closed. Finally, since f(S) is closed, it must actually contain those min and max points. Their proof is similar. [maybe my theorem requires uniform convergence?]

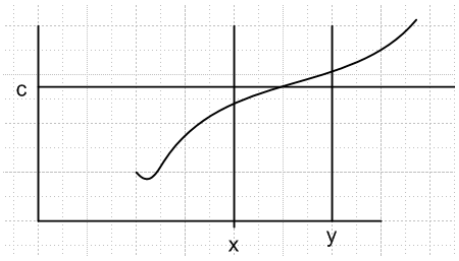
Well, here are two functions and for each we have the min and max corresponding to a point in S.



Theorem 12. f continuous on S which is compact2 in $E^n \Leftrightarrow$ **graph** of f is compact2.

Well, the graph has its "two" dimensions. The domain dimension S is compact2 as stated, and the range dimension we know is compact2 from Theorems 10 and 11. So the "graph" entity is also compact2. That is my intuitive proof, I could imagine doing out all the details. They do only the harder direction and leave easy direction as exercise.

Theorem 13. Assume f is continuous on S which is ***connected***. Assume $f(x) < c < f(y)$ for some points x and y in S. Then there must be some point z in S such that $f(z) = c$.



In the E^1 case you can see that z must exist as long as the interval [x,y] is not missing any points! If the domain is connected, no points can be missing. If even one point were missing, domain would be disconnected. For E^2 you can think of f(**x**) and f(**y**) being two points on the side of a hill. In this case as you go downhill from f(**y**), you will surely encounter some location z where f(**z**) = c. In this case there will generally be many paths along the hill surface and then **z** is not unique. If the "hill" were a gaussian of revolution, say, then the locus f(**z**) = c would be a full circle. Probably the solution locus for f(**z**) = c is a surface of dimension n-1. If the hill were disconnected by being burned along the curve f(**z**) = c, then none of the values of **z** works and in that case there is no viable **z**. You can see that more is going on for more dimensions and intuitive proofs are not going to work.

2.5 Limits of functions

Def 4 gives the usual ε δ definition of $\lim_{x\to b} f(x) = L$.

Page 77 example is a bit confusing. If we start with $f(x) = (x^3-1)/(x-1)$ for $x \neq 1$, we can do simple algebra to rewrite f as $f(x) = x^2 + x + 1$. This function is continuous at x = 1 and has the value f(1) = 3. But in the function definition they define f by saying by fiat that $f(x) = (x^3-1)/(x-1)$ on the "deleted neighborhood of x = 1" and then by fiat that f(1) = 2 (they could have said f(1) = 25.63). I think the point is that, despite the claim in the definition of the function at x = 1, the *limit* is in fact $\lim_{x\to 1} f(x) = 3$.

Theorem 14 says that limits commute with addition, multiplication and division: For example,

$$\lim_{x\to b} [\, f(x)\, g(x)\,] = [\, \lim_{x\to b} f(x)\,]\, [\lim_{x\to b} g(x)\,]$$

Of course we assume that f and g have limits at b.

Theorem 15 is a bit strange. It involves $|f(x') - f(x'')| < \varepsilon$ as $x' \to b$ and $x'' \to b$ at the same time. In E^1 you might have one approach from above and the other from below. You might have sequences which define the approach, and the sequences could be different. In E^n the approaches could be from different directions and so on. Clearly this is a Cauchy type limit, whereas $|f(x') - f(b)| < \varepsilon$ as $x' \to b$ is the normal type limit, same idea as in Cauchy sequence convergence versus normal sequence convergence. We are not talking uniform here because we are only dealing with one point b.

The rest of the section is just notation as best I can tell.

$\lim_{x\to b} f(x) = +\infty$ // OK where applies, does not apply to $\lim_{x\to 0}(1/x) = +\infty$ due to 2 sides

$\lim_{x \downarrow 0} (1/x) = \infty$

$\lim_{x \uparrow 0} (1/x) = -\infty$

$\lim_{x \uparrow 0} (x\sin x) =$ undefined and $\neq \infty$

$\lim_{x \uparrow 0} (x\sin x + 3x) = \infty$ OK

So Bucks are ∞ the represent $+\infty$, not some general complex concept. It is a real number greater than all positive real numbers and they do give it a formal definition.

Finally they talk about a function being **monotonic** on some interval, no problem.

Theorem 16: If f(x) is bounded and monotonic on (a,b) then both end limits exist with appropriate arrow.

Example consider f(x) = 1/x on (0,1). The right limit is obviously 1, the left limit is $+\infty$ and does not exist therefore.

2.6 Discontinuities

Bucks now extend things from E^1 to E^n. Limit definition is obvious. Key idea is that the limit has to be the *same* regardless of direction of approach $p \rightarrow p_0$ in E^n, otherwise the limit does not exist. They mention limit taken toward the end of an arc as basically the same as the 1D situation.

Theorem 17. I don't see what this adds, maybe limit just be same for all approaches.

Example 1. $f(x,y) = xy/\sqrt{x^2+y^2}$. Example some approach directions. Along axes get f = 0, and along 45 degree line also get 0. But to really prove the limit exists, you have to do the ε δ stuff and they do it and this example really does have a limit.

Example 2. $f(x,y) = xy / (x^2+y^2)$. Axis approach gives 0, but 45 degree approach gives 1/2 ! So the limit does not exist! Since the limit does not exist, we see below this is an "essential" discontinuity.

Example 3. $f(x,y) = xy^2/(x^2+y^4)$. A very interesting example. You get f = 0 for approach at any angle, but when you approach along a certain curve, you get 1/2! So having all linear approaches work is not enough!! In the end, you have to do ε and δ.

Bucks make the distinction between the 2D limit in x,y space, and an iterated limit of x first they y or vice versa. The 2D limit is a tougher animal. If it exists, then I think the two iterated limits will give the same result regardless of order, since each limit is an axis approach. I wonder if Bucks are going to do the Moore's theorem thing at some point, I hope so. It concerns multiple ordered limits like this.

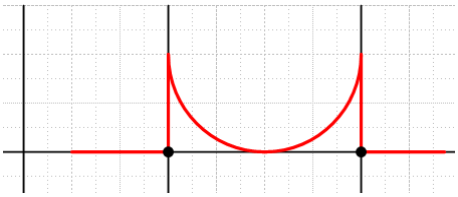
Def 9 extends definition to cases involving ∞, I don't have problems with this right now.

There are two ways to have a discontinuity at a point:

(1) function is defined at the point, but function is not continuous at the point

(2) function is not defined at the point.

Example: quadratic circle thing surrounded by 0. Within the closed circle, you are continuous at all boundary points, but the outer set is discontinuous as the boundary. Here are simpler examples. For case (2) can just consider our $f(x) = 1/x$ at $x = 0$ and the function is really undefined at the point. Another case would be the $\sin(1/x)$ thing at $x = 0$. For case (1) consider this 1D version of their disk thing:



The function is defined to be 0 at the two black points, so it is defined, but the function is obviously not continuous at the two points.

Suppose a function is given some strange value at a point which does not agree with the limit going to that point, which limit we assume exists. You can just alter the function value at that point to agree with the limit, and this is called a **removable discontinuity**. However, if the limit does not exist, then it is an **essential discontinuity**. In the drawing above, you could just raise the two black dots up to the curve and then you have "removed" those discontinuities with respect to the [a,b] interval shown. That is to say, you replace the bad values with the limits from the inside.

Examples shown page 85. For $f(x) = x^x$ you can argue that f is not defined at $x = 0$. However, the limit approaching $x \to 0$ from above does exist and is 1. So you can just add this point to the definition of the function. In this example, the function itself is not defined at $x = 0$, and you extend the function by adding this point. So $x = 0$ in this case is a removable one, although in fact we didn't change the value at the point from some other value, we added a value. For $g(x)$ we have an oddball forced value at $x = 0$ and we just change this value, so clearly a removable singularity. For the $h(x) = \sin(1/x)$ case, the limit $x \to 0$ does not exist, so $x = 0$ is an essential singularity for this function. Some other examples are then given.

They claim it is not always obvious whether a discontinuity is removable or not, a research area in 1965.

Theorem 18. Let S = bounded in E^n and $E = \overline{S}$ is the closure. Assume all the limits exist so that the function can be extended from S to E so E exists. Then $\overline{S}$ is bounded and closed, and Theorem 6 tells us that if f is continuous on $\overline{S}$, then it is uniformly continuous on $\overline{S}$. But here I guess we only know that f is continuous on S. The upshot is that when you do this extension, f is in fact uniformly continuous on all of $E = \overline{S} = S \cup \partial S$. Somehow this theorem seems obvious to me, I am as usual missing something.

Tietze's Extension Theorem. If f is continuous and bounded on closed D in E^n, you can extend f to all of E^n and it will be bounded and continuous on all E^n. I think you just go around the "perimeter" of the function, at least in E^2, and extend the value at each point out to ∞. Very obvious in E^1, not very obvious in E^3. They will prove this theorem in Chapter 4.

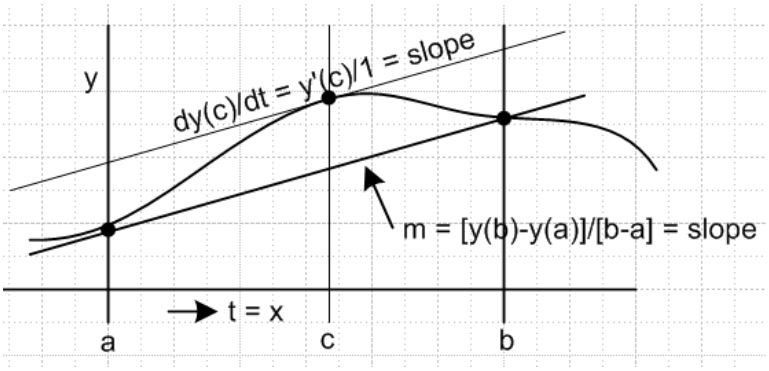
2.7 Mean value theorems and L'Hospital.

This spelling is acceptable and is in fact the original spelling of the guy's name!

Theorem 19. Consider f: R→R. If there is a local extremum at some point x_0 which lies inside the domain, then $f'(x_0) = 0$. If the extremum is at the end of a domain, then not true.

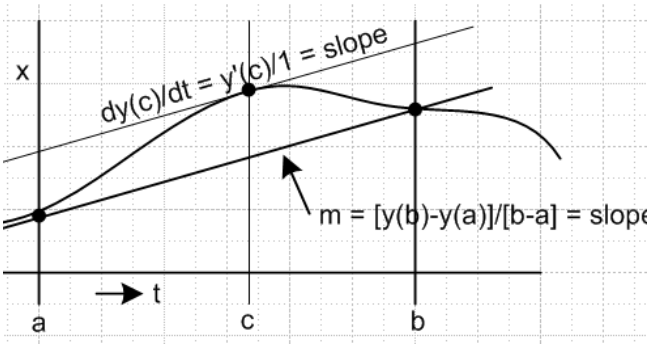
Theorem 20. (**Rolle's Theorem**) If f(a) = f(b) for an interval [a,b], and if f is *continuous* and *differentiable* on the interval, then there must be some point in that interval where the curve f(x) has zero slope.

Theorem 21. (**Mean Value Theorem**). Same context, but now function has different value at the two ends. The theorem says there is a point in the interval where the slope matches the secant slope shown.



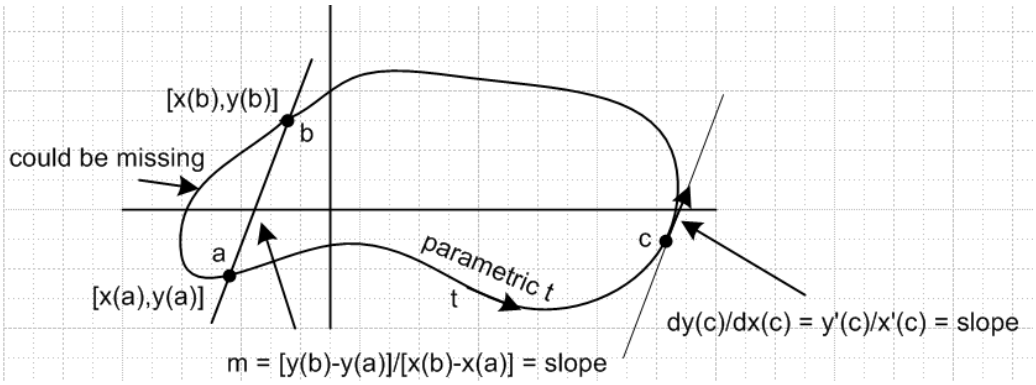
y'(c) = [y(b)-y(a)]/(b-a) c exists
y(b) = y(a) + y'(c)(b-a)
"Mean Value Theorem of Differential Calculus"

Why is this called "mean value theorem" ? Suppose the horizontal axis is time t and the vertical axis is position x for a trip along a straight track. Then redraw the picture



where now a and b are start and stop times for the trip. The slope of the curve at any point is the speed of the vehicle on the track. The secant line slope tells you the average (mean) speed for the trip, m. The theorem then says in this context that there must be at least one time instant where the instantaneous speed of the vehicle is the same as the **mean value** of the speed. I could not find any website which really explains this mean value name.

Theorem 22. (**General Mean Value Theorem**). Here is my picture where the two functions are x(t) and y(t) and are parametrically describing some curve in 2D (need not be closed).



Think of this as a 2D trip in the plane. The secant line a b is direction from start to end of the trip. It would be the direction a bee would fly. The theorem says that there must be at least one instant in time during which the vehicle travelled in this direction (arrow and time c on the right). Geometrically, there must be some point c where the slope of the curve dy/dx = y'(t)/x'(t) is equal to the m shown. That is what Buck equation (2.7) says. Now the phrase "mean value" loses its meaning. Bucks refer to the parameter not as t but as x, and the two functions are then f(x) and g(x) and point c is called x_0.

In my notation, if x(t) = t, then Thm 22 becomes Thm 21.

The proof of 22 uses Rolle's theorem 20, no big surprise. Buck version of my picture above is 2-8. Their picture is nicer because it shows a case where there are two locations for c, not just one.

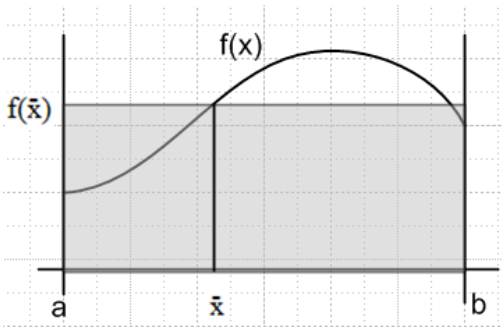
Theorem 23 seems not worth quoting.

Theorem 24 is **L'Hospital's Rule** for dealing with the limits of 0/0 or ∞/∞ ratios. It says that if the ratio of the derivatives of the functions has a limit, then that limit is also the limit of the ratio of functions. The proof uses Theorem 22 and I have not parsed it in detail.

Note Added. There is also a "**mean value theorem for integrals**" that they don't mention but which they use in the next chapter. This is different from the two mean value theorems stated above.

The claim is that $\bar{x}$ within the interval exists such that

$$\int_a^b f(x)\,dx = f(\bar{x}) \int_a^b dx = f(\bar{x})[b-a]$$



In my picture, the gray area is the same as the integral under the curve.

My Proof. f(x) must take every value between f(a) and f(b) since f is continuous ("**intermediate value theorem**"). Assume M and m are the max and min values that f(x) takes on the interval. Then

$$\int_a^b f(x)\,dx \le M\int_a^b dx \qquad\qquad \int_a^b f(x)\,dx \ge m\int_a^b dx$$

so

$$m\int_a^b dx \le \int_a^b f(x)\,dx \le M\int_a^b dx$$

or

$$m(b-a) \le \int_a^b f(x)\,dx \le M(b-a)$$

The integral must therefore equal x(b-a) where $m \le x \le M$. That is, such an x must exist. In other words, if you know that a point lies between A and B on the real axis, that point must exist.

A *different* Mean Value Theorem is given **page 106 Exercise 8 (Section 3.1)**. I will call this the:

Weighted mean value theorem for D in E^n

In nD the claim is shown on **page 106 Exercise 8 (Section 3.1)**. The claim again is that x̄ exists such that

$$\int_D f(\mathbf{x})g(\mathbf{x})d^n x = f(\bar{x})\int_D g(\mathbf{x})d^n x$$

D = open and connected in E^n

f,g = continuous and bounded

g = positive definite

I think I would just interpret g as a weight function and change to x' where $d^n x' = g(\mathbf{x})d^n x$ and then the theorem is saying

$$\int_{D'} f(\mathbf{x})\,d^n x = f(\bar{x})\int_{D'} d^n x = f(\bar{x}) * \text{Area}$$

and this says (for n=2) there is some point $\bar{x}$ in the plane such that the volume under a 2D surface equals the bottom area times the height of f evaluated at $\bar{x}$. Wiki gives a proof that looks marginal, but OK.

Well, let's try the same proof as above:

$$\int_D f(\mathbf{x})g(\mathbf{x})d^n x \leq M \int_D g(\mathbf{x})d^n x \qquad M = \max \text{ of } f \text{ on } D$$

$$\int_D f(\mathbf{x})g(\mathbf{x})d^n x \geq m \int_D g(\mathbf{x})d^n x \qquad m = \min \text{ of } f \text{ on } D$$

$$m \int_D g(\mathbf{x})d^n x \leq \int_D f(\mathbf{x})g(\mathbf{x})d^n x \leq M \int_D g(\mathbf{x})d^n x \ .$$

If g is positive definite, then the g integral is some positive number K so we have

$$mK \leq \int_D f(\mathbf{x})g(\mathbf{x})d^n x \leq MK \ .$$

There must then be some value k with $m \leq k \leq M$ such that (again, if point between A and B, it exists)

$$\int_D f(\mathbf{x})g(\mathbf{x})d^n x = kK = k \int_D g(\mathbf{x})d^n x$$

But since f must take all values between m and M, there is some $\bar{x}$ such that $k = f(\bar{x})$ since $m \leq k \leq M$. Then

$$\int_D f(\mathbf{x})g(\mathbf{x})d^n x = f(\bar{x}) \int_D g(\mathbf{x})d^n x$$

and this is the claim of the theorem. For f to take all values, f must be continuous. For various reasons, we must need D to be open and connected, and we must need f and g to be bounded and continuous. OK, enough on this for now.

We can write this in the special case of one dimension n = 1 to get

$$\int_a^b f(x)g(x)dx = f(\bar{x}) \int_a^b g(x)dx$$

Basically this is the "mean value theorem for 1D integrals with a weight added".