4150 fall 2009 review
I. Complex differentiability and the Cauchy Riemann equations
We are studying the behavior of a special class of differentiable functions f:C-->C, where C is the complex number plane. The complex number plane C ≈ R^2 is just the real number plane equipped with an additional multiplication, and our complex differentiable functions are a subclass of the real differentiable functions from R^2 to R^2, familiar from vector calculus. When thought of as a complex number, the point (a,b) of R^2 is written as a+bi.

The usual definition of complex differentiability is this: f is holomorphic at p, if it is defined on some neighborhood of p, and the limit of the difference quotient [f(z)-f(p)]/(z-p) exists as z approaches p. If f'(p) is that limit, then the linear function L(h) = f'(p).h, is the best complex linear approximation to the function f(p+h) - f(p), both considered as functions of h.

Here is the relation to real differentiability.
Recall that a function f:R^2-->R^2 is differentiable at the point p = (a,b) in the real sense, if it has a good approximation near (a,b) by a real linear transformation L. ["Good approximation" means the graph of L(h) is tangent at (0,0) to the graph of f(p+h) - f(p).] A function f:C-->C is said to be complex differentiable, or holomorphic, at the point a+bi, if it is differentiable in the real sense, and if moreover the linear function L that approximates it near a+bi is not just real - linear but complex linear. That is, not only is L(v+w) = L(v)+L(w) for all v,w in C, and L(av) = aL(v) for all v in C and all a in R, but also L(cv) = cL(v) for all v in C and all c in C.

We know all real linear maps of R^2 are represented by real 2by2 matrices, so all complex linear maps also have such matrices, but because of the extra requirement for complex linearity, the real matrices representing complex linear maps have a special form. A real linear map L is also complex linear if L(iv) = iL(v) for all v in C, so we ask next what this says about the real matrix for L.

The real matrix for L has the coordinates of L(<1,0>) in its first column, and the coordinates of L(<0,1>) in its second column. Now multiplication by i is itself a real linear map, taking 1 = <1,0> to i = <0,1>, and taking <0,1> = i, to -1 = <-1,0>, so the matrix representing multiplication by i is:
$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. If L has matrix $\begin{bmatrix} a & c \\ b & d \end{bmatrix}$, then to say that L(iv) = iL(v) for all v, means that the matrices for L and for i commute, i.e. Li = iL, hence $\begin{bmatrix} a & c \\ b & d \end{bmatrix}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} =$

$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} a & c \\ b & d \end{bmatrix}$. Multiplying out, we see that $\begin{bmatrix} c & -a \\ d & -b \end{bmatrix} = \begin{bmatrix} -b & -d \\ a & c \end{bmatrix}$. Thus we must have c = -b, and d = a. So the real matrix of a complex linear map looks like $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$.

Another way to look at this is to notice that every complex linear map L:C-->C is determined by where it sends 1 = <1,0>, and if it sends 1 to a+ib, then the map is just multiplication by a+ib. I.e., if L sends 1 = <1,0> to a+ib, it sends i = <0,1> to i(a+ib) = -b+ia. Thus its columns are the coordinates of these two images, hence again we get $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$, exactly the matrix found above.

If we translate this into a statement about the linear approximation (i.e. the derivative) of a function, we get the famous Cauchy Riemann equations. Recall the real matrix for the linear map that approximates a differentiable function f:R^2-->R^2 has the partial derivatives of f as entries. Thus if z = x+iy and f(z) = U(z) + iV(z), and if f is real differentiable, the entries in the matrix of the derivative of f are the partials of U and V: f' = $\begin{bmatrix} \frac{\partial U}{\partial x} & \frac{\partial U}{\partial y} \\ \frac{\partial V}{\partial x} & \frac{\partial V}{\partial y} \end{bmatrix}$. To say this matrix represents a complex linear map then just says that ∂U/∂x = ∂V/∂y, and ∂U/∂y = -∂V/∂x, precisely the Cauchy Riemann equations.

Conversely, we know from advanced calculus that if all 4 partials ∂U/∂x, ∂V/∂y, ∂U/∂y, ∂V/∂x, of f:R^2-->R^2, exist at every point of an open set and are continuous there, then f is real differentiable on that set, i.e. the real linear map represented by the matrix of partials is a good approximation to f locally everywhere. In the same way, if in addition to the existence and continuity of these partials, we have the Cauchy Riemann equations also satisfied, then f:C-->C is holomorphic in the given open set.

The connection between the linear map point of view and the classical point of view where the derivative is a complex number is this: The matrix $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ represents the linear map defined by multiplication by a+ib. Hence if $\begin{bmatrix} \frac{\partial U}{\partial x} & \frac{\partial U}{\partial y} \\ \frac{\partial V}{\partial x} & \frac{\partial V}{\partial y} \end{bmatrix}$ is the matrix of partials of a holomorphic function satisfying the Cauchy Riemann equations, then the complex number $\partial U/\partial x$ (p) + i $\partial V/\partial x$ (p) equals the limit of [f(z)-f(p)]/(z-p) as z approaches p. I.e. $\partial U/\partial x$ (p) + i $\partial V/\partial x$ (p) = f'(p), in the classical sense, whereas in the modern linear map sense, f'(p) is considered to be the linear map obtained by multiplying by that number. To reduce confusion, in class I called the linear map f'(p).h the differential of f at p, and called the number f'(p) the derivative of f at p.

Example: If U(x,y) = e^x cos(y) , V(x,y) = e^x sin(y), then $\partial U/\partial x$ = e^x cos(y) = $\partial V/\partial y$, and $\partial U/\partial y$ = e^x sin(y) = -$\partial V/\partial x$, and these partials are all continuous. Thus if z = x+iy, then f(z) = U(x,y)+iV(x,y) = e^x cos(y) + i e^x sin(y), is a holomorphic function everywhere in C, which restricts where y=0 to e^x.

Exercise: Let U(x,y) = ln(sqrt(x^2+y^2), and V(x,y) = Arctan(y/x) for x>0. Then check that U+iV is a holomorphic function which restricts where y = 0 to ln(x), compute the matrix of its derivative, and check that matrix is inverse to the derivative matrix for the function in the previous example. [Multiply the matrices together.]

II. Geometry, of complex linear and holomorphic maps

Complex linear maps are closely related to rotations about the origin of R^2, in the following sense. Recall that a rotation through t radians about (0,0), is a real linear map that takes <1,0> to <cos(t), sin(t)>, and <0,1> to <-sin(t), cos(t)>. You can see the first fact by recalling the polar coordinates of the point on the unit circle and angle t are <cos(t), sin(t)>. You can check the second one by recalling that rotation preserves angles, so the perpendicular vectors <1,0> and <0,1> should go to perpendicular vectors. Then note <-sin(t), cos(t)> is indeed perpendicular to <cos(t), sin(t)> since their dot product is zero.

This says the matrix of a rotation is of form $\begin{bmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{bmatrix}$, which looks just like the matrix $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ for multiplication by the complex number a+ib, where a+ib = cos(t)+isin(t) is a complex number of absolute value 1. In particular, if $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$ is the matrix for multiplication by any complex number a+ib, and if r = sqrt(a^2+b^2) is its absolute value, then we can divide all entries by r, to get c = a/r, s = b/r. Then the complex number c+is has absolute value 1, a+ib = r(c+is), and our matrix $\begin{bmatrix} a & -b \\ b & a \end{bmatrix} = r \begin{bmatrix} c & -s \\ s & c \end{bmatrix}$, where $\begin{bmatrix} c & -s \\ s & c \end{bmatrix}$ is the matrix of a rotation. Thus multiplication by any $\neq 0$ complex number is a composition of a rotation and scaling by some positive r. In particular, since rotations and scalings both preserves angles, so does multiplication by any complex number.

Since multiplying by the product of two complex numbers gives a linear map which is composed of first multiplying by one and then by the other, it follows that the geometric result of multiplying by the product of two complex numbers is the same as scaling consecutively by the length of both of them, and rotating consecutively by both of them. Consequently, the product of two complex numbers has length equal to the product of their lengths, and its angle is the sum of their angles.

Finally, since any differentiable map is locally approximated linearly by its derivative matrix, any complex differentiable map with non zero derivative also preserves the angle between curves. I.e. the angle between two curves is measured by the angle between their tangent vectors, and the tangent vectors of the two curves are mapped by the linear approximation, and this preserves angles. Since holomorphic functions with non zero derivative have this property of preserving angles, we may call them "conformal", but it is common to use the term "conformal maps" for angle preserving maps that are also bijections and even orientation preserving. I.e. holomorphic bijections are conformal maps, and it is common to use the term "conformal maps" only for holomorphic bijections. [It can be proved that all orientation preserving conformal maps are holomorphic. So when we say "conformal maps" for holomorphic bijections, we are taking for granted the properties orientation - preserving, and bijection.]

III. Examples of holomorphic maps and their inverses.
The exponential map.
Define e^z = e^x cos(y) + i e^x sin(y), which we have already checked is a holomorphic map C-->C, in our example above. Notice the length of e^z equals e^x, and the angle equals y, so for x = 0, only the angle of e^z changes as y changes, the length remaining constantly 1. So this function winds the vertical line x = 0 around the unit circle counterclockwise, again and again, as y runs from -infinity to infinity. If we let x range between -infinity and 0, and y range from 0 to 2π inclusive, this half infinite horizontal strip covers the punctured unit disc, missing only 0, and covering the half open interval (0,1] on the x axis twice, once when y = 0 and once when y = 2π.

Indeed since e^x is never 0, and cos(y) and sin(y) cannot both be zero, e^z is never zero for any z in C, but every other complex value occurs infinitely many times. The values of e^z on the doubly infinite horizontal strip $0 \leq y < 2\pi$ cover the entire plane except for 0. Each strip $2\pi \leq y < 4\pi$, $4\pi \leq y < 6\pi$,....also covers the entire punctured plane when mapped by e^z in the same way. Thus the exponential function sends horizontal lines $y = b$ bijectively to half - infinite rays from 0 to infinity along a fixed angle of b radians, and sends vertical lines $x = a$ onto circles of radius e^x, wrapping each vertical line around the corresponding circle infinitely many times.

Logarithms

A logarithm is an inverse to the exponential map. Since the exponential map is not injective, it does not have a global inverse, but it does have an inverse on any set which is the image of a subset where e^z is injective. Logs are never uniquely defined, since we could always add $2\pi i$ to any value of the log and get another one, since $e^{(z+2\pi i)} = e^z$. Hence we sometimes call any choice of a logarithm a "branch" of the logarithm. We have to be a bit more restrictive if we want a continuous, holomorphic inverse. For example there is a "branch" of the log on the whole, punctured plane with values in the strip $0 \leq y < 2\pi$, but it is not continuous along the x axis. So to get a continuous log we restrict further to the complement of the closed set $\{x \geq 0, y=0\}$, we stay off the closure of the positive real axis. Then the logarithm defined on that set, with values in the strip $0 < y < 2\pi$ is continuous and holomorphic. We will se later that a logarithm can be defined by a path integral, $\mathrm{Log}(z) = \int_1^z \frac{dz}{z}$ for any path from 1 to z. For this to be single valued we need to restrict our path, hence our domain of the log function to some region not containing zero and not containing any paths that wind around zero. E.g. removing the non negative real axis works fine, but there are many other open regions that work as well. It is a consequence of the open mapping principle for holomorphic functions that to find a holomorphic branch of the inverse of a holomorphic function, it suffices to find a continuous branch, and that a holomorphic function has a holomorphic inverse on any open set where it is injective. (Lang, page 82, thm.6.4)

nth power functions

The function z^n for an integer $n \geq 1$, is absolutely fundamental, and it expands each pie wedge with vertex at the origin, by widening its angle n times. Thus a wedge with width $\pi/2$, when squared becomes a wedge of width π, and if cubed becomes a wedge of width $3\pi/2$. Hence if the wedge has width more than π, when squared it covers some points of the plane more than once. Similarly, the whole plane itself thought of as a wedge of "width" 2π and centered at 0, is wound three times around on itself by the map z^3.

nth root functions

These are inverses of power functions, and since nth power functions are not always injective, nth roots exist only on restricted regions. Just as power functions z^n with $n \geq 1$ expand angles at the origin, rooi functions z^(1/n) shrink these angles. For example, a holomorphic branch of the square root function is defined on the upper half plane (width π), and maps it onto the open upper right quadrant (width $\pi/2$). Conveniently, nth root functions are defined on any set where a branch of the logarithm can be defined, i.e. on any open set not containing a path which winds around the origin, since we may define an nth root of z by z^(1/n) = e^(1/n log(z)). This also shows that nth root functions are holomorphic since log(z) and e^z are holomorphic (where defined).

IV. Topology of complex sets, open maps

A fundamental concept is that of an open set. A subset U of C is open in C if for every point p in U, there is a disc D centered at p which is also completely contained in U. Thus no points on the "edge" of U are included in U. E.g the interior of the unit disc, the set U = {z:, |z| < 1} is an open set, since for any point p of length less than 1, all points close enough to p also have length less than 1. The "closed unit disc" E = {z: $|z| \leq 1$} is not open since it includes some points of length equal to 1, e.g. z = 1, and no matter how small a disc F we choose centered at such a point, there will be some points in that disc F which have length > 1, so they will not be in the closed disc E.

Not every set is open, so it is useful to consider the largest open subset of a given set. This is called then "interior" of the set. If S is any subset of C, p is an interior point of S if there is some disc centered at p and completely contained in S. So a set is open if all its points are interior points. Some subsets of C have no interior points, such as the set Q^2 of points whose coordinates are both rational, or more simply just the x axis, since no disc is contained entirely in the x axis.

A map f:C-->C is called continuous if the inverse image f^(-1)(V) of every open set V is open, and is called an open map if the (forward) image of every open set U is also an open set f(U). Notice that an open map can never have even a local maximum modulus at any point of the open set U. I.e. if p is any point of U, then f(p) is never on the edge of the set f(U), so there are always points q near p in U, such that |f(q)| > |f(p)|. Ironically, this is very useful in helping us find maximum moduli for functions defined on sets that are not open. For example if f:C-->C is continuous and E is a closed disc, then f does have a maximum modulus somewhere on E. However if f is also open, then that maximum modulus cannot occur in the interior of E, thus it occurs on the "frontier" or "boundary" of E, what we were calling intuitively the "edge" of E.

Recall that a set E in C is by definition "compact" if every sequence in E has an accumulation point in E, or equivalently has a convergent subsequence with limit in E. By a theorem this holds if and only if E is both “closed” (i.e. its complement is open) and bounded. Then the image of a compact set under a continuous function f is also compact, and thus on a compact set the modulus of a continuous complex valued function always has a maximum and a minimum.

In real calculus, a function whose derivative is never zero is open, and it follows that the maximum of a real valued differentiable function on an open interval, if it exists, occurs at some point where the derivative is zero. Complex differentiable functions are different, since they define open maps even at isolated zeroes of the derivative. Thus as long as a complex differentiable (“holomorphic”) function on a connected open set U is not constant, then it defines an open map on U. Hence if E is any closed disc in U, the maximum modulus of f on E must occur on the boundary circle of E.

It is not trivial to prove that a non constant holomorphic function on a connected open set is an open map. Our proof was given in stages, and used the deep theorem from real advanced calculus that a real C^1 function $f:R^2 \to R^2$ whose partial derivative matrix is invertible, i.e. has non zero determinant at p, defines an invertible map and hence an open map, on some neighborhood of p. For a holomorphic function f, the invertibility of the matrix of partials follows whenever the derivative is non zero. I.e. from the Cauchy Riemann equations, the determinant of the matrix of partials equal $a^2 + b^2$ when the complex derivative is $a+ib$, and hence that determinant is non zero if either a or b is non zero, i.e. whenever $a+ib \neq 0$.

Thus the inverse function theorem on R^2 guarantees that a holomorphic function defines an open map on any set where the derivative is never zero.
This called the “open mapping theorem”, Thm. 2.21 in Greenleaf, page 103, but this is not what I call the open mapping theorem for holomorphic (i.e. analytic) functions. For some reason unknown to me, Greenleaf completely ignores the much more powerful fact that a holomorphic function defines an open mapping on any connected open set where the derivative is not identically zero. I.e. even if the function does have zero derivative some places, as long as the function is not constant, then the map is still open. This is what gives complex functions their special geometric properties, and this is the true open mapping theorem for holomorphic functions, and it is given in Lang, pages 80-81.

I do not know why Greenleaf does not point this out, since the hard part of the proof is the real variable theorem that he is using and taking for granted. Well, maybe he did not know the easy proof, since he does give a lengthy proof of the maximum modulus principle for analytic functions on pages 165-167? No, more likely he wanted to be self contained, and give a proof that did not assume the real variable proof from advanced calculus. Also notice he does not get out quite as strong a result in his thm.2.21 as the open mapping theorem would give. I.e. he gets the corollary of the maximum modulus principle but does not get the local structure theorem.

Lang's and our proof of the open mapping principle for complex functions starts from the fact that the basic power functions z^n are open mappings, for n ≥ 1. This is not true in real calculus, since x^2 is not an open mapping near x=0. This fact lets us prove the open mapping principle for complex functions as follows. Basically we just use the inverse function theorem to show that every analytic function looks locally near an isolated zero of the derivative, essentially the same as a power function. [Of course to get the result for holomorphic functions, we have to wait until we know all holomorphic functions are analytic.]

Note that the product z.g(z), where g is holomorphic, has derivative equal to g(0) at z=0. Hence if g(0) ≠ 0, this is a locally invertible function, by the inverse function theorem. If f is analytic and has derivative zero at z = 0, by factoring out a suitable power of z, it can be written as z^n.h(z) for some analytic h with h(0) ≠ 0. Thus by the theory of logarithms we can find an analytic nth root of h, so that f = [z.g(z)]^n, near z=0. Then zg(z) has ≠0 derivative at z=0 hence is invertible. And f is the composition of the invertible open map z.g(z) with the nth power function. Since a composition of open maps is open, f is open near z=0. This explains why all analytic functions are open near points where they are not constant, since the power series for f has some non zero coefficient a(n), and we can thus factor out z^n and give this argument with g(z) being the other factor, i.e. with f(z) = z^n.g(z), where g(0) = a(n) ≠ 0.

V. Applications of the open mapping theorem for holomorphic functions

The open mapping principle implies the maximum and minimum modulus principles: i.e. a non constant holomorphic function cannot have a maximum modulus at an interior point of a connected domain, and cannot have a minimum modulus at an interior point unless the value there is zero. However if we have a holomorphic function f defined on a closed disc E, then it is also continuous, and by compactness of the closed disc, the modulus |f| has both a maximum and a minimum on E. Hence either f is constant on E, or if not, then the maximum of |f| occurs on the boundary circle. Moreover, if f is non constant and the minimum modulus occurs inside the disc, it must be zero.

It follows that if f is a non constant "entire" function, i.e. analyic on the whole plane, and if M(r) = the maximum of |f(z)| for |z| = r, then M(r) is a strictly increasing function of r. This holds by the maximu modulus principle, since for r1 < r2, any point with |z| = r1 is interior to the disc |z| ≤ r2, so the maximum value M(r2) cannot occur on |z| = r1. Thus if f is entire and |f|--≥0 as r-->infinity, then f is identically zero, since M(r) cannot strictly increase and also converge to zero, hence f is constant, and the only constant that converges to zero is zero.

Example: Consider the polynomial f(z) = 6z^3 -2z^2 + 2z -1. On the boundary |z| = 1 of the unit disc, |f(z)| = |6z^3 -2z^2 + z -1| ≥ |6z^3| - |2z^2| - |2z| -|1| = 1. But at z=0, |f(0)| = |-1| = 1 also, so the minimum modulus does occur somewhere in the interior of the disc. [Since f is continuous and |z| ≤ 1 is compact, there is a minimum modulus somewhere.] Since f is a non constant holomorphic function and the open disc is connected, that minimum must be zero, i.e. this polynomial does have a root inside the unit disc, (but I do not know where. Somebody may make me look foolish again by finding an obvious root, but I hope not.)

Proving zeroes exist: If f is analytic and non constant on a neighborhood of a compact connected set K, and if there is a point p in the interior of K where |f(p)| is as small as the minimum value of |f| on the boundary of K, then the minimum modulus of f does occur in the interior of K and hence must be zero.

Example: If f(z) = z^5 + 10z^4 - 13z^3 + 12z + 31, prove f has a zero in the disc |z| < 15. (Does that work?)

Fundamental theorem of algebra: If f(z) is any polynomial of degree n ≥ 1, by dividing by its lead coefficient we may assume it has lead coefficient 1 without changing whether it has roots, so let f(z) = z^n + a(1)z^(n-1) +...+a(n-1)z + a(n). We claim f has a root in some disc |z| ≤ r. Factoring out z^n leaves
f(z) = z^n[1 + a(1)/z +....+a(n-1)/z^(n-1) + a(n)/z^n]. As z-->infinity, the factor in the bracket approaches 1 and z^n approaches infinity, so f(z)-->infinity. In particular for some r, |z| ≥ r implies that |f(z)| > |a(n)|. Since f(0) = a(n), the minimum modulus on the closed disc |z| ≤ r does not occur on the boundary circle |z| = r. Since it does occur somewhere on this compact disc, it occurs in the interior and is thus zero by the minimum modulus principle. Hence f has a root.

VI. Principle of isolated zeroes, and "analytic continuation".

An analytic function has a fundamental property that it cannot be zero on a very large set. This is called the "principle of isolated zeroes". A point p is isolated in a set S, if there is some disc D centered at p that contains no oither points of S except p. So p is a certain finite distance away from all other points of S. A set S of points in C is called discrete if all its points are isolated, i.e. if for every point p in

S, there is a disc D centered at p such that no other point of D except p is in the set S. For instance the integers on the x axis form a discrete set. An open set is "connected" if it is not the union of two disjoint, non empty open subsets, or equivalently if any two points in it can be joined by a polygonal path with edges parallel to the axes. If f is an analytic function in a connected open set, then either f is constantly zero, or else its set of zeroies is discrete, i.e. all the zeroes of a non constant analytic function in a connected open set are isolated: near one zero there are no other zeroes.

This is locally easy, is true globally using connectedness, and has important consequences. To prove it, assume f(p) = 0, but some derivative of f is not zero at p, so the Taylor series of f at p looks like (z-p)^n h(z), where h(p) ≠ 0. Then by continuity h is not zero on some disc centered at p, and z^n is only zero at p. This proves that if f(p) = 0, but some derivative of f is not zero at p, then p is an isolated zero. Thus if at every zero of f, some derivative is non zero, then all zeroes of f are isolated. That is the easy local part.

Now we claim there cannot be a point p where all derivatives of f are zero, unless f is a constant function. This global statement uses connectedness of the domain. For if p is a point where all derivatives of f are zero, then since f is analytic, it will have zero Taylor series near p, hence f will be constant in some disc centered at p, and hence at all points near p f will also have all derivatives equal to zero. Thus the set where all derivatives are zero is open.

But since each derivative is a continuous function, the set where some derivative is non zero is also open, being the inverse image under the continuous derivative of the open set of non zero complex numbers. Thus the subset of the domain of f where all derivatives are zero is open, and also the complementary subset where some derivative is non zero is open. If we assume the domain of f is connected then by definition of connected sets, one of these two sets is empty. I.e. either all the derivatives of f are zero at all points, or at every point some derivative is non zero.

Thus if all derivatives are zero at some point then f is constant, while if at every point some derivative is not zero, then all zeroes are isolated. Hence for an analytic function on a connected open set, either all zeroes are isolated or else f is constant.

The contrapositive statement is the one we want: if an analytic function has a non isolated zero, then f is identically zero on any connected subset of its domain containing that zero. I.e. <u>among all analytic functions on a connected domain, only the identically zero function has a non isolated zero.</u>

Thus if we have two analytic functions f,g defined on C, and they agree at the points 1/n on the real axis for n>0 a positive integer, then f = g everywhere. Why? well, f-g = 0 at every point of form 1/n, hence also at 0 by continuity, so f-g has z=0 as a non isolated zero, hence f-g is constant on the connected set C, hence is always equal to 0. Thus f(z) = g(z) for all z.

Now it is crucial here that both f and g are defined and analytic also at 0 for this to be true. Otherwise we would not have z as a non isolated zero if it were not in the domain. E.g. f(z) = [e^(2πi/z) - 1], is zero whenever e^(2πi/z) = 1, and that happens at all z such that 2πi/z = 2nπi for some n. Hence it happens at all z = 1/n. So this function does equal zero at all z = 1/n, but it is not analytic at z = 0, so z = 0 is not a non isolated zero.

As a consequence, and assuming we know that all holomorphic functions are analytic, the holomorphic function above f(z) = e^x cos(y) + i e^x sin(y), is the only holomorphic function defined on C which equals e^x when y = 0. I.e. e^x has exactly one extension to a holomorphic function on all of C. We say that the "analytic continuation" of the real function e^x into the complex plane is unique.

In the same way, all familiar real functions like cos(x), sin(x), tan(x), ln(x), arctan(x), x^n, x^(1/n), etc.... have at most one analytic extension, or continuation, into the whole complex plane.

Beware: The uniqueness of analytic continuation depends on the domains of the continuations being connected and the same. I.e. if U is a connected open set in C such that Umeet{x axis} contains at least one non isolated point, and if f,g are two analytic functions in U, which agree on Umeet{x axis} = the part of the x axis contained in U, then f = g everywhere on U.

But, if U and V are two <u>different</u> connected open sets in C, such that Umeet{x axis} = Vmeet{x axis}, and if f and g are two analytic functions, f defined in U and g defined in V, such that f=g in Umeet{x axis} = Vmeet{x axis}, there is NO guarantee that f(z) = g(z) for points z that happen to be in both U and V.

Example of non uniqueness when the domains are different:
If U = C - {non positive y axis} = {x+iy, x≠0 or y>0}, and V = C - {non negative y axis} = { x+iy: x≠0 or y<0}, then there are two different branches of the log function, one = f defined in U, and another = g defined in V, such that both agree with ln(x) on the positive x axis, but they disagree on the negative x axis. I.e. f will have value iπ at z= -1, while g will have value -iπ at z = -1.
This is because Log(z) = ln(|z|) + iArg(z), and from the shape of these domains, we

must take Arg(-1) = π in U, but Arg(-1) = $-\pi$ in V. [I hope I got this right, but as usual I am doing this in my head, so I may have screwed up the minus signs, or worse. I have already corrected it once, so please check me and do not just believe it blindly.]

Remark: I personally think an even better reason to believe the uniqueness of analytic continuation, at least locally near a point p of the x axis, is the fact that the continuation has the same Taylor series. I.e. if there is an analytic continuation of f into the complex plane then the Taylor series should converge also in the complex disc centered at p, but the coefficients are determined by the derivatives of f at p, which can be calculated from just real values approaching p. So if f is going to have any analytic continuation into the complex plane, it will have to come from the same power series, but with complex values of z plugged in. Thus there is only one choice of this continuation, at least near p, namely the one given by the same Taylor series as for the real version of f. But we still need some connectedness arguments to make this argument work in other subsets. I guess in some sense this is really the same argument as before, since that isolated zeroes argument also used the coefficients of the Taylor series, but this seems more concrete.

Anyway, the surprizing result is that if you tell me the values of f at some convergent sequence in C, then there is at most one way to define f in the whole complex plane and having those values on that sequence. Thus the values of an analytic function in one part of the plane, completely determine the values in every other part of the plane. This is a very strong statement.

Example: If I want an analytic function defined in all of C, with value i/n at the point 1/n, the only such function is f(z) = iz. If I want an analytic function g with all values = 2 at the points 1/n, the only such function is f(z) = 2 for all z in C.

VII. Path integration

Cauchy's integral theorem

Cauchy's integral theorem says that a function has integral equal to zero over a closed path, provided f is analytic (or holomorphic, since that is equivalent) at every point on the path and "inside" the path.

But what does "inside the path" mean? If the path is a simple closed curve, like a circle, then the meaning is clear, it means points inside the circle. If the path has more than one component, like the border of a larger disc with some smaller discs removed, it means any point in the part of the larger disc that is not removed.

If the path is complicated and crosses itself, maybe many times, a point is inside according to how many times, and in what direction, the path winds around the point. Basically, a point is inside a path if and only if the path winds around the point some non zero number of times. And a point is outside the path if the winding number for that point is zero.

You have to count the winding number including the direction of winding to get the right answer, since going back the other way can cause canceling in the number, i.e. winding numbers can be positive and negative and can cancel each other. If we take the counterclockwise direction as positive, then the winding number equals the number of times the path winds counterclockwise around the point, minus the number of times it goes clockwise around the point.

If a region U is convex, then for every closed path in U, all points inside the path are also in U. I.e. a path in a convex region cannot wind around any points outside that region. Since a disc is convex, if f is analytic in a disc, then the integral of f around any closed path in that disc is zero.m Similarly the uppwer half plane is convex, so any analytic function integrates to zero around any closed path in the upper half plane.

Convexity is not needed for this, and a weaker condition is called "simply connected". A region U is called simply connected if every closed path in U can be "shrunk" to a point continuous in U. More rigorously, a closed path C can be shrunk to a point in U if there is a continuous parameter map from the unit circle onto C, that can be extended to a continuous map of the unit disc into U. For instance, removing just the non positive x axis from the plane leaves a simply connected region, one such that no path can wind around any of the missing points. That is why the path integral of dz/z is well defined independent of path in this region and why the logarithm is defined there.

The general shrinking process for paths is called "homotopy" and is discussed with reference to Cauchy theorem in Lang pages 116-119, and 110-116. The winding number approach is also discussed in Lang, pages 138-145. Essentially a region is simply connected if no path in the region can wind around a point outside the region. This means that if q is a point outside the region, then dz/(z-q) always integrates to zero over any closed path lying inside the region. The famous Emil Artin, Lang's thesis advisor, is usually credited with developing this approach.

Applications of Cauchy's integral theorem, **integral formulas, Rouche'**
The integral formulas for values and derivatives of analytic functions are useful

corollaries of the Cauchy integral theorem. These allow one to apply the basic estimates on the size of a path integral in terms of the maximum of the integrand and the length of the path, to estimate the growth rate of the derivatives of analytic functions. In particular one obtains the Liouville theorem that the only bounded "entire" (analytic in the full plane) functions are constants, and similarly the only entire functions bounded by $B|z|^n$ for some constant B and some integer $n>0$, are polynomials of degree $\leq n$.

We also get residue formulas for path integrals. I.e. Cauchy says the integral around a closed path is zero if f is analytic at every point inside the path. More generally if f is analytic at all but a finite number of points inside the path, the integral equals $2\pi i$ times the sum of the residues at those points, each residue being counted as often as the path winds around the corresponding point. E.g. the integral of dz/z (clockwise) around the unit circle equals $2\pi i$, and for a path winding n times around the origin counterclockwise, the integral of dz/z is $2n\pi i$.

One can sometimes count the zeroes of an analytic function inside a closed contour (path) by reducing to an easier analytic function with the same number of zeroes. I.e. one can sometimes use Cauchy to prove two analytic functions have the same number of zeroes inside a contour, and then count them using the simpler function. This combines Cauchy's theorem with the so called "argument principle". The argument principle says that the number of zeroes of a function f analytic on and inside a closed contour, equals $(1/[2\pi i])$ times the integral of f'/f around that contour, provided f has no zeroes on the contour itself. By the change of variables formula, this integral also equals the integral of dw/w around the image of the contour under f. I.e. if the path is K, and $w = f(z)$, then dw/w integrated over w in f(K), equals df/f integrated over z in K.

Rouche' principle for counting zeroes

Rouche's principle says that if $|f| > |g|$ everywhere on K, and f,g are both analytic on and inside K, then f and f+g have the same number of zeroes (counted properly with multiplicities) inside K. This is because under the hypothesis that $|f| > |g|$ everywhere on K, it follows that f(K) and (f+g)(K) wind around $w = 0$ the same number of times. Hence dw/w has the same integral around f(K) and around (f+g)(K), hence also df/f and $d(f+g)/(f+g)$ have the same integral around K, i.e. f and f+g have the same number of zeroes inside K.

Example: If $f(z) = z^4$, $g = z^3 + z + 2$, and $K = \{z: |z| = 2\}$, then $|f| > |g|$ everywhere on K, so $f+g = z^4 + z^3 + z + 2$, has the same number of zeroes as $f = z^4$ inside K, namely 4 zeroes, counted with multiplicities. To see that $|f| > |g|$ on K, it suffices to note that on K, $|f| = 16$, and $|g| \leq |z^3| + |z| + |2| = 12$ on K.

Holomorphic versus analytic functions, the converse of Cauchy's theorem

1. A function f is called "holomorphic" in an open set U if f is complex differentiable everywhere in U.

2. f is called C^1 - holomorphic in U, if it is holomorphic and also the complex derivative is continuous everywhere in U.

3. f is called C^infinity - holomorphic in U, if f has infinitely many continuous complex derivatives at every point of U.

4. f is called analytic in U, if near every point of U f is represented by a power series.

The main theorem is that all these are equivalent. In particular all holomorphic functions in U are actually analytic in U.

We did not quite prove this in detail. We did prove that all analytic functions are C^infinity -holomorphic (on the take home test, problems IX and X and extra). In particular the only power series that can represent an analytic function is its Taylor series.

It is of course trivially true that all C^infinity -holomorphic functions are also C^1 - holomorphic. We also proved all C^1 - holomorphic functions are analytic in class, when we proved the converse of the Cauchy theorem.

I.e. we proved that any function satisfying the Cauchy integral theorem for some basic bordered regions, must be analytic. (This is proved in Greenleaf, pages 303-304.) It follows as a corollary that a uniform limit of analytic functions is also analytic, since the integral of the limit function must also satisfy the Cauchy integral theorem on the same bordered regions that the approximating functions do, because the integral of a uniform limit equals the limit of the integrals. So if all the {fn} have zero integral around the borders of simple regions then also the integral of the limit function f has zero integral around those borders.

Then we proved that all C^1 - holomorphic functions did satisfy Cauchy by using Green's theorem. This approach is mentioned in Greenleaf on pages 289-290, but not proven there. Thus all C^1 - holomorphic functions are analytic.

We skipped the proof of Goursat's theorem that all merely holomorphic functions satisfy Cauchy's theorem, but if we had done that then we could conclude they are also analytic. We skipped this proof because giving it is a little harder than just quoting Green's theorem.

The proof of Goursat's theorem is Thm. 3.1, p. 105 of Lang, and it is also the one called an "Assertion" on page 290 of Greenleaf, and proved on pages 292-295.

Since all these classes of functions are eventually proved the same, some books mix up the definitions, i.e. they use the word "analytic" for holomorphic functions with one complex derivative, simply because eventually such functions will be proved to actually have power series representations.

This is confusing in my opinion. Lang and Greenleaf both use the term holomorphic correctly, for functions with one complex derivative, on page 30 of Lang, and page 81 of Greenleaf.

Another proof we skipped was to show that a function which is analytic, or holomorphic, at all points except the center p of an open disc, i.e. a function which has an isolated singularity at p, has a “Laurent” expansion about the center, i.e. an expansion in powers of (z-p) which may involve both negative and non negative powers. We know this if the function arises as a quotient of two anaklytic functions, since we know how to divide power series with positive powers to obtain a power series with a finite number of negative powers. But an infinite number of negative powers is possible too at an isolated singularity.

This is proved exactly the same way as our proof that a function satisfying Cauchy’s formula has a Taylor series. I.e. Cauchy’s formula has integrand of form f(z)/[z-a], and if we assume p = 0, we rewrote the integrand as f(z)/{z(1-a/z)} and expanded the second factor in the bottom as a geometric series in (a/z), convergent when |a| < |z|, i.e. for a points inside a circle of points |z| = constant. But if we have a singularity at 0, we can give Cauchy’s formula also for a in the punctured disc, except we have to integrate around two circles, with a in between, and larger one |z| = r, and a smaller one |z| = e, where e can -->0. Then we can expand the integral for z in the larger circle in powers of (a/z), i.e. positive powers of a, since for z in the klarger circle we have |a| < |z|. But we can also expand for z in the smaller circle, where |z| < |a|, but then we have to write (z-a) = -a(1-z/a), and use the geomnetric series to expand in powers of (z/a), i.e. negative powers of a. Since the inner circle is oriented backwards, we also get a minus sign that cancels this one, and eventually we get a potentially doubly oinfinite power series, whose coefficient of the power a^n is the integral of f(z)dz/z^(n+1), whether n is negative or non negative. The only difference is the coefficients of the non negative powers are integrated over the larger circle, and those of the negative powers are integrated over the smaller circle.

Then we get some wonderful corollaries. The first is an analog of the Lioiuville theorem. Just as these integral expressions for the coeficients of give estiamtes that imply an entire function bounded near infinity, hence bounded, is constant, so also the similar estimates for the negative power coeficients imply that if f is bounded near p, then all negative powers have zero coefficient, hence p is a removable singularity. This is called Riemann's removable singularity theorem. An elaboration of this argument, shows that if f approaches infinity as a-->p, then there are only a finite number of negative powers, i.e. p is a "pole", whereas there are an infinite number if and only if the values of f in every neighborhood of p are actually dense in the complex plane. Here p is called an "essential" singularity. This last result is usually called the Casorati Weierstrass theorem. A stronger result due to Picard, shows that at an essential singularity p, f actually takes on all but possibly one complex value in every neighborhood of p. An example of such a function is f(z) = e^(1/z). It seems a good exercise to check Picard's theorem for this example.

Term by term differentiation of power series

On take home test 3 we also outlined the proof that one can differentiate power series term by term. We appealed to the fundamental theorem of calculus, parts 1 and 2. Recall by FTC1 we mean the statement that if a continuous differential has integral which is independent of path in a given region, then the function those integrals define is an antiderivative of the original differential, i.e. the derivative of the integral of a continuous f(z)dz, equals that f(z). Conversely, FTC2 says that if f(z)dz = dg is the differential of a smooth function g, then the path integral of f(z)dz is obtained by evaluating g at the endpoints and subtracting, i.e. the integral of the differential of a smooth g, is g again, up to a constant.

In particular FTC1 implies that in order to prove a power series has derivative equal to its derived series, it suffices to show it is the integral of that derived series. But the derived series does converge uniformly on compact sets, hence the integral of the derived series is the limit of the integrals of the individual terms. By definition each individual term of the derived series is the derivative of the corresponding term of the original series, so by FTC2, the original series, up to a constant term, is the integral of the derived series.