To Dad,
my best teacher and my favorite student

Chapter 1

The Complex Number System

1.1 Constructions of $\mathbb{C}$

The set of complex numbers $\mathbb{C}$ is defined by accepting $\sqrt{-1}$, usually written i, into the community of numbers. Once we accept $\sqrt{-1}$, we have to accept arithmetical combinations which involve it, such as $5+6\sqrt{-1}$, usually written $5+6i$. We have to accept division by i, which fortunately is the same as multiplication by $-i$. (Since $-i \cdot i = 1$, the reciprocal of i is $-i$.) It turns out that general combinations of the form $a+bi$ are "enough" in the sense that for any two of them we can identify a sum, difference, product, or quotient also of the form $a+bi$.

We therefore wish to define $\mathbb{C}$ to be "only" the set of numbers $a+bi$. But in order to do so, we need to give an account of what $a+bi$ represents. Some authors define $\mathbb{C}$ to be the set of "formal sums" $a+bi$. This seems to avoid, rather than answer, the basic question ("What does $a+bi$ mean?") by defining "$a+bi$" to be a form representing its own form. It leaves many questions unanswered: Is $3+1i$ identical with $3+i$? or $1+2+i$? or $i+3$? We feel these should be the same, but they differ in form. If $\mathbb{C}$ is defined to be a set of forms, these things should be unequal! Quite generally, to define a thing as a "formal" [form] is to intentionally confuse its nature with its method of expression, in order to avoid the real work of describing its nature. We can do much better:

Definition 1. $\mathbb{C}$ is defined to be the set of ordered pairs (a, b):

$$\mathbb{C} \equiv \{(a, b) \,|\, a, b \in \mathbb{R}\}$$

with associated addition and multiplication formulas: $(a, b) + (c, d) = ((a + c), (b + d))$, and $(a, b) * (c, d) = ((ac - bd), (ad + bc))$.

This definition is clear and rigorous, but it has two disadvantages: First, $\mathbb{R}$ is not a subset of this set, but we want $\mathbb{R} \subset \mathbb{C}$. We solve this problem by identifying (refusing to distinguish) $x \in \mathbb{R}$ with $(x, 0) \in \mathbb{C}$. (The alternative is to use $\mathbb{C} = (\mathbb{R} \times \mathbb{R}) - (\mathbb{R} \times \{0\}) \cup \mathbb{R}$ instead, but that forces the sum and product rules to break into cases.)

Next, the definition requires us to write complex numbers in the form "$(2, 5)$", which is strange and unnatural. We would rather write $2 + 5i$. A sleight of hand solves the problem: First, we *define* i to be the complex number $(0, 1)$. With the identifications $a = (a, 0)$ and $b = (b, 0)$ in mind, we compute $a + bi = (a, 0) + (b, 0)(0, 1) = (a, b)$, as desired. Thus $a + bi = (a, b)$ is not a "shorthand" but a legitimate computation.

There is another option for those who know a little ring theory:

Definition 2. $\mathbb{C}$ is that quotient ring of $\mathbb{R}[i]$ (here, i is used as an indeterminant variable, just like the variable usually written x) obtained by modding out by the principle ideal generated by $i^2 + 1$.

$$\mathbb{C} = \mathbb{R}[i]/(i^2 + 1)$$

This definition has many advantages: It explains the addition and multiplication formulas, it makes it trivial to see that $\mathbb{C}$ is a ring, and pretty easy to show that $\mathbb{C}$ is a field. But the elements of $\mathbb{R}[i]/(i^2 + 1)$

are cosets, and we really don't want complex numbers to be cosets, so it adds an unnecessary level of complication. Again we have the problem that $\mathbb{R}$ is not a subset of $\mathbb{R}[i]/(i^2+1)$, but it is standard practice in algebra to identify a ring R with its image in $R[x]$ or R/I, and these conventions give us $\mathbb{R} \subset \mathbb{R}[i]/(i^2+1)$.

Theorem 3. *The definitions are equivalent in the sense that both structures are fields and are isomorphic.*

Proof. Since $i^2+1 \in \mathbb{R}[i]$ is monic of degree 2, every element of $\mathbb{R}[i]/(i^2+1)$ can be written uniquely in the form $a+bi$ (The linear polynomial remaining after polynomial long division by i^2+1.) We may define $\phi : \mathbb{R}[i]/(i^2+1) \to \mathbb{R} \times \mathbb{R}$ by $\phi(a+bi) = (a,b)$. It's easy to check that addition and multiplication are preserved. It follows that $\mathbb{R} \times \mathbb{R}$ is a ring (in the first place! We hadn't yet checked it.) and is isomorphic to $\mathbb{R}[i]/(i^2+1)$. To show that $\mathbb{R}[i]/(i^2+1)$ is a field, note that $\mathbb{R}[i]$ is a PID and i^2+1 is irreducible, so (i^2+1) is a maximal ideal. Thus the quotient is a field. $\square$

The fact that it is a field indicates that it is a full number system − no other combinations of i are necessary. It's reassuring that quotients $(a+bi)/(c+di)$ exist, but we also need to calculate them. To find the reciprocal of a complex number $a+bi$, we apply the following trick:

$$\frac{1}{a+bi} = \frac{a-bi}{(a+bi)(a-bi)} = \frac{a-bi}{a^2+b^2} = \frac{a}{a^2+b^2} + \frac{-b}{a^2+b^2}i$$

Note that the trick fails exactly when $a+bi=0$. The number $a-bi$ is called the conjugate of $a+bi$. The conjugate function is usually written with an overline: $\overline{a+bi} = a-bi$.

Exercise 1. (Factoring 13) The Gaussian integers are numbers of the form $a+bi$ for $a,b \in \mathbb{Z}$. They form a ring, usually written $\mathbb{Z}[i]$. They play the role of integers for the complex numbers. Prove that the "prime" number 13 is the product of two smaller Gaussian integers. Here "smaller" means in norm: w is smaller than z if $|w| < |z|$.

Exercise 2. (Matrix representation) Show that the system of all matrices of the form $\begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}$ (with $\alpha, \beta \in \mathbb{R}$) with matrix addition and multiplication is isomorphic to the field of complex numbers. Furthermore, show that any such matrix can be written in the form $\begin{pmatrix} r & 0 \\ 0 & r \end{pmatrix} \cdot \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}$ for some $r \in [0,\infty)$ and $\theta \in [0,2\pi)$.

1.2 More operations

We have four more operations on complex numbers which help us understand them in terms of real numbers:

Definition 4. We define four special functions on $\mathbb{C}$: For a complex number $z=a+bi$, we define:

1. $\overline{z} = a-bi$. This is called the *conjugate* of z.
2. $|z| = \sqrt{a^2+b^2}$. This is called the *modulus* of z.
3. $\Re(z) = Re(z) = a$. This is called the real part of z.
4. $\Im(z) = Im(z) = b$. This is called the imaginary part of z. Notice that the "imaginary part" is a real number!

Proposition 5. *For any complex numbers w and z:*

1. $|z|^2 = z\overline{z}$
2. $Re(z) = \frac{1}{2}(z+\overline{z})$
3. $Im(z) = \frac{1}{2i}(z-\overline{z})$
4. $|wz| = |w||z|$.

Exercise 3. Prove the above proposition by writing z in the form $z=a+bi$. (This technique − proving things by writing z in the form $a+bi$ is one we wish to outgrow, and the statement of this proposition will help us do so. We do not wish to constantly regard z as two variables to be independently manipulated!)

So all four special functions can be expressed using the conjugate. What makes conjugation so special? To start, it's an automorphism of $\mathbb{C}$:

Exercise 4. (Conjugation is an automorphism) Verify by direct calculation, using the definitions of $+$ and $\cdot$, that $z \mapsto \overline{z}$ is a field homomorphism from $\mathbb{C}$ to $\mathbb{C}$. In order to show that it is bijective, provide its inverse function. Verify that conjugation fixes every element of $\mathbb{R}$.

Corollary 6. *For complex numbers w and z:*

1. $\overline{w \pm z} = \overline{w} \pm \overline{z}$
2. $\overline{wz} = \overline{w}\ \overline{z}$
3. $\overline{w/z} = \overline{w}/\overline{z}$

Of course the identity map $id : \mathbb{C} \to \mathbb{C}$ is also an automorphism fixing $\mathbb{R}$. But there are no other such maps. In the language of Galois Theory, conjugation is the single nontrivial element of the Galois group of $\mathbb{C}$ over $\mathbb{R}$.

Exercise 5. Prove that any field automorphism $\phi : \mathbb{C} \to \mathbb{C}$ which fixes $\mathbb{R}$ must be conjugation or the identity. (Hint: What can $\phi(i)$ be?)

1.3 Basic Inequalities and relations

Algebra is powerful, but we wish to study the analysis of $\mathbb{C}$. Analysis begins with inequalities.

Proposition 7. *For any complex number z:*

$$-|z| \leq Re(z) \leq |z| \qquad \text{and} \qquad -|z| \leq Im(z) \leq |z|$$

...with equality, respectively, if z is nonpositive real, nonnegative real, nonpositive imaginary (i.e., bi for $b \leq 0$), and nonnegative imaginary.

Exercise 6. Prove the previous proposition by writing $z = a + bi$ and using properties of real numbers a and b.

Metric topological reasoning requires a triangle inequality in $\mathbb{C}$:

Proposition 8 (Triangle Inequality)**.** *Let w and z be complex numbers. Then:*

1. $|w \pm z| \leq |w| + |z|$ *with equality if and only if $z = 0$ or $w/z \in \mathbb{R}^{\pm}$*
2. $|w \pm z| \geq |w| - |z|$ *with equality if and only if $z = 0$ or $w/z \in \mathbb{R}^{\mp}$*

Proof. The modulus of a sum $|w + z|$ involves an inconvenient square root of a sum. It can be avoided by reasoning about square modulus instead:

$|w + z|^2 = (w + z)\overline{(w + z)} = w\overline{w} + z\overline{z} + w\overline{z} + \overline{w}z = |w|^2 + |z|^2 + w\overline{z} + \overline{w\overline{z}} = |w|^2 + |z|^2 + 2Re(w\overline{z})$

Now Proposition 7 allows us to relate $Re(w\overline{z})$ to $|w\overline{z}| = |w||\overline{z}| = |w||z|$, obtaining:

$|w|^2 + |z|^2 - 2|w||z| \leq |w + z|^2 \leq |w|^2 + |z|^2 + 2|w||z|$, thus

$(|w| - |z|)^2 \leq |w + z|^2 \leq (|w| + |z|)^2$. All these values except possibly the far left are positive real numbers, so we may drop the squares:

$|w| - |z| \leq |w + z| \leq |w| + |z|$.

Again by Proposition 7, equality holds, respectively, when $w\overline{z}$ is nonpositive real, or nonnegative real, respectively. If $z \neq 0$, $w\overline{z} = |z|^2 w/z$, and $|z|^2$ is necessarily positive real.

This completes the proof for $|w + z|$. For the case $|w - z|$ simply apply the first case with $-z$ in place of z.

□

Exercise 7. Use the triangle inequality to prove that $|z_1 + z_2 + \ldots + z_n| \leq |z_1| + |z_2| + \ldots + |z_n|$

1.4 Polar representation

Let $w = a + bi$ be a complex number. Let r denote the distance from 0 and θ denote the counterclockwise angle from the positive real axis, in radians. Note that we have already discussed an operation on w that yields r, namely $r = |w|$. The operation which yields θ is called $\arg(w)$, but it is not well behaved because $\theta + n2\pi$ has the same geometric significance as θ. Usually we insist $\arg(w) \in [0, 2\pi)$(sometimes $(-\pi, \pi]$). This makes arg a definitive function on $\mathbb{C} - \{0\}$, but doesn't make it continuous.

Notice that $a = |w|\cos(\theta)$ and $b = |w|\sin(\theta)$, so $w = |w|(\cos(\theta) + i\sin(\theta))$. The parenthesized portion is called $\operatorname{cis}(\theta)$:

Definition 9.

$$\operatorname{cis}(\theta) \equiv_{\text{def}} \cos(\theta) + i\sin(\theta)$$

Thus $w = |w|\operatorname{cis}(\theta)$.

Proposition 10. *The function* $\operatorname{cis} : \mathbb{R} \to \mathbb{C}$ *has the following properties*

1. $\operatorname{cis}(0) = 1$
2. $\operatorname{cis}(\alpha + \beta) = \operatorname{cis}(\alpha)\operatorname{cis}(\beta)$
3. $\operatorname{cis}(\alpha - \beta) = \operatorname{cis}(\alpha)/\operatorname{cis}(\beta)$
4. $\operatorname{cis}(n\alpha) = \operatorname{cis}(\alpha)^n$.
5. $|\operatorname{cis}(\theta)| = 1$

Proof. Parts 1 and 5 are obvious. 2 is essentially the angle sum formulas for cos and sin. (Digression: Proof of angle sum formulas, using geometry: Represent the angle $\theta + \phi$ as on a unit circle, then drop a perpendicular to the angle θ, and from there another perpendicular to x-axis, and calculate the height.) 3 follows from 2. 4 follows from 2 with induction. $\square$

We would like to claim that $\frac{d}{d\theta}\operatorname{cis}(\theta) = -\sin(\theta) + i\cos(\theta) = i\operatorname{cis}(\theta)$. Compare to normal rules for differentiating exponential functions $\frac{d}{dt}e^{kt} = ke^{kt}$, which reproduce the same function times a constant. We can't explain the formula $\frac{d}{d\theta}\operatorname{cis}(\theta) = i\operatorname{cis}(\theta)$ without a theory of derivatives of functions $f : \mathbb{R} \to \mathbb{C}$. We are nevertheless tempted by this derivative calculation and by parts 1-4 above to regard cis as a kind of exponentiation. Later we will learn that $\operatorname{cis}(\theta) = e^{i\theta}$, but we are not yet ready to even define complex exponentiation. The students should understand, however, that cis is a "temporary" notation, to be replaced later. All of our theorems about cis will be regarded, later, as theorems about $e^{i\theta}$.

Exercise 8. Use the laws for $\operatorname{cis}(\theta)$ to derive "triple angle formulas" for sin and cos. *Do not* use the angle sum formulas directly (although of course the proof of proposition 10 depends on them).

Exercise 9. The function cis is a group homomorphism. What are its domain and range? (Tell each group by telling the set and its operation.) What is its kernel?

Exercise 10. Find the sixth roots of unity. Give their values arithmetically and include a picture of them in the complex plane. Together they form a group, by proposition 10 (do not prove), isomorphic to what familiar group? What is the operation of the group? What is the identity of the group?

1.5 Geometry of operations

If $\mathbb{C}$ can be identified with $\mathbb{R} \times \mathbb{R}$, then it is the domain of classical Euclidean geometry. Complex operations should have geometric interpretations. What is the geometric significance of addition? of multiplication?

Addition is vector addition, so it corresponds geometrically to the parallelogram rule. Using polar coordinates, we can now derive the geometric significance of multiplication: If $w = |w|\operatorname{cis}(\alpha)$ and $z = |z|\operatorname{cis}(\beta)$ are complex numbers, then $wz = |w||z|\operatorname{cis}(\alpha)\operatorname{cis}(\beta) = |w||z|\operatorname{cis}(\alpha + \beta)$, a complex number whose norm is the product, and whose argument is the sum. In short, magnitudes multiply and angles add. Note that for any fixed $w \in \mathbb{C}$, the map $z \mapsto wz$ is a fairly simple transformation of the plane – an expansion by a factor of $|w|$, and a rotation by $\arg(w)$, fixing the origin.

We now summarize the geometric significance of each of the fundamental operations on complex numbers:

1. Addition can be interpreted as a parallelogram rule, as for vectors.
2. Subtraction can be interpreted as a parallelogram rule with a reversal, as for vectors.
3. Multiplication requires multiplication of moduli and addition of angles.
4. Division requires division of moduli and subtraction of angles.
5. Modulus represents distance from 0.
6. Conjugation is a vertical flip over the real axis.
7. Real part and imaginary part are horizontal and vertical coordinate.

1.6 The Riemann Sphere

For real functions, we often attach ∞ and $-\infty$ to the system $\mathbb{R}$, but not algebraically: $8 + \infty$ is not defined. We do attach them topologically, though: $\lim_{a\to\infty} f(x)$ and $\lim_{a\to 0} f(x) = \infty$ are both defined. Topologically speaking, the extended real line is a compactification of $\mathbb{R}$, a topological extension space which is compact.

The extension of $\mathbb{C}$ is similar in that it is topological, and causes these limits to be defined. It is different in the following ways:

1. There is only one infinity. The alternative was not two, but infinitely many, one for each direction of escape! (This "alternative" comment is open-ended. It could suggest spaces isomorphic to $P^2_{\mathbb{R}}$, a closed unit disk, or even the Stone-Cech compactification.)
2. We often allow a substantial amount of arithmetic with infinity: $a * \infty = \infty$ (for nonzero a) , $a/\infty = 0$, and $a + \infty = a - \infty = \infty$. In fact the only arithmetical forms we don't allow are $\infty \pm \infty$, ∞/∞, $0 * \infty$, $0/0$. We even allow $a/0 = \infty$ and $b/\infty = 0$. However, we cannot do these things without spoiling the *field structure* of $\mathbb{C}$, so the extended system is not a good algebraic structure. The guiding principle is whether or not a potential operation $a \star b$ is continuous near infinity, zero, or wherever, as a function of both of its inputs. To explain this we need to have a topology in place which allows us to define continuity.
3. We use a sphere model to picture $\mathbb{C} \cup \{\infty\}$, and linear projection to explain the correspondence. Therefore we need to:
4. define a topological structure on the Riemann sphere and show it topologically equivalent to the sphere. We can do this via standard compactifications. (illustrated b/c you haven't seen compactifications!)
5. We use coordinatized analytic geometry to compare the sphere to the Riemann Sphere in a nontrivial way.

The projection

We define a map p from $S^2 - \{(0,0,1)\}$ to $\mathbb{C}$ by projection from the north pole, identifying $\mathbb{C}$ with the $x - y$ plane in $\mathbb{R}^3$ (That is, $a + bi$ is identified with $(a, b, 0) \in \mathbb{R}^3$). We wish to find a formula for this projection map, in both directions. Let us write (x_1, x_2, x_3) for an arbitrary point on S^2, subject of course to the condition $x_1^2 + x_2^2 + x_3^2 = 1$, and let us write $z = a + bi$ for an arbitrary point on $\mathbb{C}$, which we identify with the vector $(a, b, 0) \in \mathbb{R}^3$. Since (x_1, x_2, x_3) is on the line connecting $(0, 0, 1)$ to $(a, b, 0)$, we have some t for which $t(0, 0, 1) + (1-t)(a, b, 0) = (x_1, x_2, x_3)$. We see immediately that $t = x_3$, so $a = \frac{x_1}{1-x_3}$ and $b = \frac{x_2}{1-x_3}$. This gives the formula for $z = a + bi$ in terms of (x_1, x_2, x_3):

$$z = a + bi = p(x_1, x_2, x_3) = \frac{x_1 + ix_2}{1 - x_3}$$

We wish to also provide a formula for the inverse function. This is harder because we must solve for x_1, x_2, and x_3 in terms of $z = a + bi$. We begin by taking square moduli of the equation above. (Note the denominator $1 - x_3$ is a real constant.)

$$
\begin{aligned}
|z|^2 &= \tfrac{x_1^2+x_2^2}{(1-x_3)^2} && (1.1)\\
&= \tfrac{1-x_3^2}{(1-x_3)^2} \quad \text{(because } (x_1,x_2,x_3) \text{ is on the sphere.)} && (1.2)\\
&= \tfrac{1+x_3}{1-x_3} \quad \text{(cancel after factoring } 1-x_3^2\text{)} && (1.3)\\
& && (1.4)
\end{aligned}
$$

We multiply by $1 - x_3$ to solve for x_3:

$$|z|^2 - |z|^2 x_3 = 1 + x_3 \qquad \text{and} \qquad x_3 = \frac{|z|^2 - 1}{|z|^2 + 1}$$

If follows that the denominator $1-x_3$ is equal to $1-x_3 = \frac{2}{|z|^2+1}$. Since $x_1 = a(1-x_3)$ and $x_2 = b(1-x_3)$ we now have:

$$x_1 = \frac{2a}{|z|^2+1} \qquad x_2 = \frac{2b}{|z|^2+1} \qquad x_3 = \frac{|z|^2-1}{|z|^2+1}$$

$$p^{-1}(z) = \left(\frac{2a}{|z|^2+1}, \frac{2b}{|z|^2+1}, \frac{|z|^2-1}{|z|^2+1}\right)$$

The maps p and p^{-1} above are continuous and give a topological homeomorphism between $\mathbb{C}$ and $S^2 - \{(0,0,1)\}$. Notice that the algebra above fails if $1 - x_3 = 0$. There is only one such point on the sphere, namely the north pole itself $(0,0,1)$. Here the geometry of the "unique" line through the north pole and *itself* also fails, but this is our intention. The coordinate x_3 is an *increasing* function of $|z|$, taking values in $[-1,1)$, and that regions $B(0,M) \subseteq \mathbb{C}$ correspond to regions $\left\{(x_1,x_2,x_3) \in S^2 | x_3 < \frac{M^2-1}{M^2+1}\right\}$. As $|z|$ grows larger, x_3 approaches 1. The set $\mathbb{C} \cup \{\infty\}$ is not a topological space in any reasonable way, but the completed set S^2 is. By extending p and p^{-1} so identify $p(0,0,1) = \infty$ and $p^{-1}(\infty) = (0,0,1)$, we can assign a topology to $\mathbb{C} \cup \{\infty\}$ homeomorphic to that of S^2, making p a homeomorphism by fiat.

Exercise 11. For each of the functions from $\mathbb{C}$ to $\mathbb{C}$, find a formula for the corresponding function from S^2 to S^2 (in x_1, x_2, x_3 coordinates.) You can do this by creating the composition $p^{-1} \circ f \circ p$ of maps $S^2 \to \mathbb{C} \to \mathbb{C} \to S^2$ and simplifying algebraically.

For each function, describe how the resulting spherical function can be interpreted as a geometric motion on the sphere.

1. $f(z) = \overline{z}$
2. $g(z) = -z$
3. $h(z) = 1/z$

Exercise 12 (hard)**.** Prove that circles on S^2 correspond, by stereographic projection, to circles and lines on $\mathbb{C}$. Which circles on S^2 correspond to lines on $\mathbb{C}$? Explain with proof.

Exercise 13 (hard)**.** Let $w, z \in \mathbb{C}$ and let $d(w,z)$ denote the distance between the points corresponding to w and z on S^2. Prove:

$$d(w,z) = \frac{2|w-z|}{\sqrt{1+|w|^2}\sqrt{1+|z|^2}}$$

Chapter 2

Metric Spaces and the Topology of $\mathbb{C}$

2.1 Metric Spaces

Definition 11. A **metric space** is a set X together with a function $d : X \times X \to \mathbb{R}$ (called a **distance function** or *metric*) so that

1. (Zero) $d(x, y) = 0$ iff $x = y$
2. (Nonnegative) $d(x, y) \geq 0$
3. (Symmetry) $d(x, y) = d(y, z)$
4. (Triangle inequality) $d(x, z) \leq d(x, y) + d(y, z)$

The two main examples are $\mathbb{R}$ with distance function $d(x, y) = |y - x|$ (absolute value) and $\mathbb{C}$ with distance function $d(w, z) = |z - w|$. These two alone justify the abstraction above, but there are countless other important metric spaces. Note that we have proven the triangle inequality for $\mathbb{C}$.

Some discussion: Analysis is often a game of estimation, defined by proving that some absolute difference $|P - Q|$ is less than ε. Any proof that $|P - Q| < \varepsilon$ with *arbitrary* $\varepsilon > 0$ is a roundabout way to prove the equality $P = Q$. The role of the triangle inequality is to estimate by waypoints: If P is nearly P', which is close to Q', which is about Q, we find a way to prove $|P-P'| < \varepsilon/3$, $|P'-Q'| < \varepsilon/3$, and $|Q'-Q| < \varepsilon/3$. The combination of these facts to say $|P - Q| < \varepsilon$ is the triangle inequality. Such a proof is usually written this way: $|P-Q| = |P-P'+P'-Q'+Q'-Q| \leq_{\text{(triangle inequality)}} |P-P'|+|P'-Q'|+|Q'-Q| < \varepsilon/3+\varepsilon/3+\varepsilon/3 = \varepsilon$. Note that the first step seems "magical" until we become accustomed the trick.

Definition 12. An **open ball** $B(x, \varepsilon)$ in a metric space X is a set of the form $\{y \in X | d(x, y) < \varepsilon\}$. A **closed ball** is similar: $\overline{B}(x, \varepsilon) = \{y \in X | d(x, y) \leq \varepsilon\}$.

Definition 13. A set A is **open** if every point $a \in A$ is contained in some open ball $a \in B(a, \varepsilon) \subseteq A$. A set is **closed** if its complement is open.

The terms *open* and *closed* are oddly chosen, since a set can be both, or neither. Perhaps whoever chose these terms was unfamiliar with doors.

Proposition 14. *The set of open sets in a metric space (X, d) satisfies the axioms for open sets in an abstract topology. That is,*

1. *$\emptyset$ and X are open sets.*
2. *Unions of arbitrary families of open sets are open.*
3. *Intersections of finite families of open sets are open.*

Proposition 15. *The set of closed sets in a metric space (X, d) satisfies the axioms for closed sets in an abstract topology. That is,*

1. *$\emptyset$ and X are closed sets.*

2. *Intersections of arbitrary families of closed sets are closed.*

3. *Unions of finite families of closed sets are closed.*

Notation 16. We write the symbol "$\imath$" (an upside-down Greek iota), pronounced "that" or "the unique," to refer to some definite thing by means of a property that it alone has. For example, "$\imath x\ \ x^3 = 64$" means "the unique x whose cube is 64," which of course is 4: $(\imath x\ \ x^3 = 64) = 4$. In some cases no such thing exists, or many exist, so the construction cannot be interpreted to refer to an object. For example, "$\imath x\ \ x > 0$" fails to refer to anything. This should remind the reader of standard limit notation, which can also fail to refer to an answer if a limit "is not defined" or "does not exist". In fact, limits can be described as a special case of $\imath$-notation, and we will often define limits below using $\imath$.

The notation "$\imath$" is Bertrand Russell's notation for definite description, but our policy for interpretation differs from Russell's. The symbol $\imath x$ acts like a quantifier in that it anticipates a statement involving x, but a statement $\imath x P(x)$ is grammatically a noun clause, not a complete statement.

Definition 17. (Sequence convergence) Let $(x_n)_{n\in\mathbb{N}}$ be a sequence in a metric space X. The limit of the sequence is defined to be

$$\lim_{n\to\infty} x_n \equiv \imath L\ \ \forall\varepsilon > 0\ \ \exists N\ \ \forall n > N\ \ |x_n - L| < \varepsilon$$

equivalently,

$$\lim_{n\to\infty} x_n \equiv \imath L\ \ \forall \text{ open } U \ni L\ \ \exists N\ \ \forall n > N\ \ x_n \in U$$

Of course the limit of a sequence may or may not exist. If it exists we say that x_n is convergent and that it converges to L. We write $x_n \to L$. (If a limit is *bad*, like $\lim_{n\to\infty}(-1)^n$, we contend that it lacks *existence*, not *definition.* Any sequential limit in a metric space is "defined" in the literal sense that a definition is provided above and an alternate definition would conflict. This author's position – that a notation can be "defined" but not refer to an existent thing – is not standard.)

Exercise 14. Prove the implied equivalence in the previous definition

Proposition 18. *(Sequentially closed and open) Let X be a metric space. A set $A \subseteq X$ is closed if and only if for any convergent sequence $(x_n) \subseteq A$, the limit $\lim_n x_n \in A$. On the other hand, a set A is open if and only if no sequence outside A converges to a point of A.*

The reader should beware that in nonmetric topologies, convergence of sequences does not characterize open and closed sets.

Definition 19. Let $A \subseteq X$ be a subset of a metric space. The **interior** of A, written A^0 is, equivalently,

1. $\bigcup\{U \,|\, U \subseteq A, U \text{ open}\}$

2. $\{x \in X \,|\, x \in U \subseteq A \text{ for some open } U\}$

3. $\{x \in X \,|\, B(x,\varepsilon) \subseteq A \text{ for some } \epsilon > 0\}$

Exercise 15. Prove the implied equivalences in the previous definition

Definition 20. Let $A \subseteq X$ be a subset of a metric space. The **closure** of A, written $\overline{A}$ is, equivalently,

1. $\bigcap\{C \,|\, C \supseteq A, C \text{ closed}\}$

2. $\{x \in X \,|\, U \cap A \neq \emptyset \text{ for every open } U \ni x\}$

3. $\{x \in X \,|\, B(x,\varepsilon) \cap A \neq \emptyset \text{ for every } \epsilon > 0\}$

4. $\{x \in X \,|\, x \text{ is the limit of some sequence } (x_n) \text{ in } A\}$

Exercise 16. Prove the implied equivalences in the previous definition

In general $A^0 \subseteq A \subseteq \overline{A}$. Any open set equals its interior, and any closed set equals its closure.

2.2 Continuity

Definition 21. A function $f : (X, d) \to (Y, \rho)$ from one metric space to another is **continuous** if, equivalently,

1. $\forall x_0 \in X \;\; \forall \varepsilon > 0 \;\; \exists \delta > 0 \;\; \forall x \in X \;\; d(x, x_0) < \delta \to \rho(f(x), f(x_0)) < \varepsilon$
2. For any open set $U \subseteq Y$, $f^{-1}(U)$ is open.
3. For any closed set $C \subseteq Y$, $f^{-1}(C)$ is closed.
4. For any convergent sequence (x_n), $\lim_n f(x_n) = f(\lim_n x_n)$

Exercise 17. Prove the implied equivalences in the previous definition

Sequential limits are identified by some integer input (the index) which goes to infinity. A sequence is just a function of such an input. We also have limits in which some input x, varying in a metric space X, approaches a fixed point x_0. Thus we define the limit of a (not necessarily continuous!) function defined on X.

Definition 22. Let $f : (X, d) \to (Y, \rho)$ be a function from a metric space to another. Let $x_0 \in X$. We define the limit:

$$\lim_{x \to x_0} f(x) \equiv_{def} \imath y_0 \in Y \;\; \forall \varepsilon > 0 \;\; \exists \delta > 0 \;\; \forall x \in X \;\; 0 < d(x, x_0) < \delta \to \rho(f(x), y_0) < \varepsilon$$

As in basic Calculus, the limit $\lim_{x \to x_0} f(x)$ may or may not exist. Note that this empowers us to discuss $\lim_{x \to x_0} f(x)$ when $f : \mathbb{R} \to \mathbb{R}$, $f : \mathbb{C} \to \mathbb{C}$, $f : \mathbb{R} \to \mathbb{C}$, or $f : \mathbb{C} \to \mathbb{R}$. All four cases will be important.

Definition 23. A function f is **continuous at** x_0 if $\lim_{x \to x_0} f(x) = f(x_0)$.

Note that continuity at x_0 does not require continuity, not even in a neighborhood of x_0. But if f is continuous at every point, it is continuous.

If we have a sequence $\{f_n\}$ of functions and a sequence (x_m) of inputs, we may take limits in two ways: $\lim_{n \to \infty} \lim_{m \to \infty} f_n(x_m)$ and $\lim_{m \to \infty} \lim_{n \to \infty} f_n(x_m)$. Limits of this type do not always commute, and we must use caution to distinguish them carefully.

2.3 Connectedness

Definition 24. A set U in a metric space (X, d) is **clopen** if it is both open and closed.

Definition 25. A metric space X is **disconnected** if it is the union of two nonempty disjoint open sets. Equivalently, if it has a clopen set other than $\emptyset$ and X. A subset $A \subset X$ is **disconnected** if the metric space (A, d) is disconnected. A set is **connected** if it is not disconnected.

Proposition 26. *Intervals in $\mathbb{R}$ (i.e., sets of the form (a, b), $[a, b]$, $[a, b)$, or $(a, b]$, where a might be $-\infty$ and b might be $+\infty$) are connected, and every connected set is an interval.*

Proposition 27. *(Continuous image of a connected set is connected.) Let $f : X \to Y$ be a continuous map of metric spaces, with X connected. Then $f(X)$ is connected.*

2.4 Compactness

Definition 28. A topological space A is called **compact** if, equivalently,

1. Every open cover $A = \bigcup_{\alpha \in I} U_\alpha$ has a finite subcover $A = \bigcup_{i=1}^{n} U_{\alpha_i}$.
2. Every family $\mathcal{F}$ of closed subsets of A with the finite intersection property (i.e., any finite subfamily has nonempty intersection) has nonempty intersection (i.e., $\bigcap \mathcal{F} \neq \emptyset$).

If $A \subset X$ is a subset of a larger topological space X, we call A compact if it is compact in the subspace topology.

Exercise 18. Prove the implied equivalence in the previous definition

This equivalence is easier than it looks, little more than DeMorgan's laws.

Proposition 29. *Let $K \subseteq X$ be compact in a metric space X. Then K is closed.*

Proof. If $x \notin K$, then K has an open cover of the form $\{X - \overline{B(x,\varepsilon)} | \varepsilon > 0\}$. Any finite subcover of that yields a ball $B(x, \varepsilon)$ disjoint from K. $\square$

Proposition 30. *If $C \subseteq K \subseteq X$, with C closed and K compact, then C is compact.*

Proposition 31. *Let $f : X \to Y$ be a continuous function, and $K \subseteq X$ compact. Then $f(K)$ is compact.*

Proposition 32. *The product $X \times Y$ (in the product topology) of two compact spaces X and Y is compact.*

Theorem 33 (Heine-Borel Theorem, 1D)**.** *Let $K \subseteq \mathbb{R}$. Then K is compact if and only if it is closed and bounded.*

Proof. Let K be compact. Then it is closed, by proposition 29. Boundedness is trivial. Conversely, it suffices to show that $[a, b]$ is compact, since K is a closed subset of some closed interval. Let $[a, b] \subseteq \bigcup \mathcal{U}$, where $\mathcal{U}$ is a family of open sets. We use a connectedness argument. Let

$$C = \{c \in [a, b] \,|\, [a, c] \text{ can be covered by a finite subcollection of } \mathcal{U}\}$$

If $[a, c]$ is finitely covered, then we have $c \in B(c, \varepsilon) \subseteq U_c \in \mathcal{U}$, and for any $c' \in B(c, \varepsilon)$, $[a, c']$ is finitely covered by the same cover. So C is open. On the other hand, if $[a, c]$ is not finitely covered, then $c \in B(c, \varepsilon) \subseteq U_c \in \mathcal{U}$. If any $c' \in B(c, \varepsilon)$ were such that $[a, c']$ were finitely coverable, then $[a, c]$ could be finitely covered by adding the cover element U_c. So $[a, b] - C$ is also open. Since $a \in C$, connectedness yields $[a, b] = C$ as desired. $\square$

Theorem 34 (Heine-Borel Theorem, general)**.** *Let $K \subseteq \mathbb{R}^n$. Then K is compact if and only if it is closed and bounded.*

Proof. If K is closed and bounded, it's a subset of some interval product $[a_1, b_1] \times \ldots \times [a_n, b_n]$, which is compact by the previous two statements. $\square$

Caution: For many students, "closed and bounded" is conceptually easier than "compact" and the Heine-Borel theorem sounds like permission not to grapple with the subtleties of open covers. Please note that "compact" is a question of general topology, but "bounded" is meaningless in nonmetric spaces. Moreover, simple metric spaces other than $\mathbb{R}^n$ often have closed bounded noncompact sets. The Heine-Borel property is a specialized result for a specialized space and shouldn't underpin one's concept of compactness!

2.5 Completeness

Consider the sequence $x_n = 1/n$ in the space $\mathbb{R}$ and also in the space $\mathbb{R} - \{0\}$. In the former, it converges to zero, but in the latter it is a divergent sequence. In $\mathbb{R} - \{0\}$, there is no way to measure how close x_n gets to 0, but we can say that the sequence terms get close to one another. The Cauchy criterion expresses that the terms of a sequence get close to each other. Of course it is strongly related to convergence. (In fact, Cauchy called this condition "convergence"!)

Definition 35. Let x_n be a sequence in a metric space. We call x_n **Cauchy** if

$$\forall \varepsilon > 0 \ \exists N \ \forall m > N \ \forall n > N \ d(x_m, x_n) < \varepsilon$$

Note that the definition of convergence depends on the limit point, which may not be known, whereas the Cauchy condition depends only on the terms of the sequence.

Proposition 36. *Every convergent sequence is Cauchy.*

Exercise 19. Prove the previous proposition.

Returning to the sequence $x_n = 1/n$, we see that it converges in $\mathbb{R}$, so it is Cauchy. Despite being Cauchy, it does not converge in $\mathbb{R} - \{0\}$, and this can be considered a "flaw" with the metric space $\mathbb{R} - \{0\}$. Spaces without such problems are called complete:

Definition 37. Let X be a metric spaces. We call X **complete** if every Cauchy sequence in X converges.

Proposition 38. $\mathbb{R}$ *is complete.*

Proof. Let x_n be a Cauchy sequence in $\mathbb{R}$. Since x_n is Cauchy, it is bounded in some interval $[a, b]$. For each N, let C_N be the closure of the tail sequence $\{x_N, x_{N+1}, \ldots\}$ in $\mathbb{R}$. The sequence (C_N) is a decreasing sequence of nonempty closed sets in a compact set, so there is some $L \in \bigcap C_N$. Now for any $\varepsilon > 0$ we may choose N so that for all larger n, m we have $|x_m - x_n| < \varepsilon/2$. But the tail sequence $\{x_N, x_{N+1}, \ldots\}$ has within it some $x_n \in B(L, \varepsilon/2)$. By the triangle inequality, $\{x_N, x_{N+1}, \ldots\} \subseteq B(L, \varepsilon)$. □

Corollary 39. $\mathbb{R}^n$ *and* $\mathbb{C}$ *are complete.*

Proof. A Cauchy sequence (v_n) in $\mathbb{R}^n$ has as its coordinates n Cauchy sequences on $\mathbb{R}$. They converge individually, producing a vector to which (v_n) converges. □

Exercise 20. Prove that a closed subset of a complete metric space is complete.

2.6 Uniform convergence and continuity

A sequence of functions f_n on a space X produces, for each point $x \in X$, a sequence $f_n(x)$. Pointwise convergence refers to the individual convergence of such sequences:

Definition 40. Let $f_n : X \to Y$ be a sequence of functions. We say f_n **converges pointwise** to f if, for each point $x \in X$, $\lim_{n\to\infty} f_n(x) = f(x)$. That is,

$$\forall x \in X \ \ \forall \varepsilon > 0 \ \ \exists N \ \ \forall n > N \ \ \rho(f(x), f_n(x)) < \varepsilon$$

Now if each function f_n is continuous, it's natural to expect that the limit function f is also continuous, but this need not be the case:

Example 41. Let $f_n : [0, 1] \to \mathbb{R}$ by $f_n(x) = x^n$. Then the sequence of functions converges pointwise to the (discontinuous) function $f : [0, 1] \to \mathbb{R}$ which is identically zero except $f(1) = 1$.

There is a stronger form of convergence, however, which makes things work out better:

Definition 42. Let (f_n) be a sequence of functions. We say that (f_n) **converges uniformly** to f if

$$\forall \varepsilon > 0 \ \ \exists N \ \ \forall n > N \ \ \forall x \in X \ \ \rho(f(x), f_n(x)) < \varepsilon$$

The only difference between this definition and the previous is the relative ordering of "$\forall x \in X$" and "$\exists N$". Thus the difference between pointwise convergence and uniform convergence is that uniform convergence allows us to choose N, depending on $\varepsilon > 0$, but *not* depending on x. Abstractly, "$\exists N \forall x \, P(N, x)$" implies "$\forall x \exists N \, P(N, x)$", so uniform convergence implies convergence. The example above, restricted to the interval $[0, 1/2]$ converges uniformly because $|f_n(x)| \le 2^{-n}$. The point of this refined notion of convergence is that the uniform limit of continuous functions is continuous:

Theorem 43. *Let (f_n) be a sequence of continuous functions converging uniformly to f. Then f is continuous.*

Proof. In spirit, continuity boils down to the estimation $f(x) \approx f(x_0)$. Our strategy is to estimate $f(x) \approx f_n(x) \approx f_n(x_0) \approx f(x_0)$. Let $\varepsilon > 0$. By uniform convergence, choose N so that

$$\forall n > N \ \ \forall x \in X \ \ \rho(f(x), f_n(x)) < \varepsilon/3$$

Fix some $n > N$. Let x_0 be fixed. By the continuity assumption on f_n, choose $\delta > 0$ so that

$$\text{for given } x_0 \in X \ \ \forall x \in X \quad \text{if } d(x_0, x) < \delta \text{ then } \rho(f_n(x), f_n(x_0)) < \varepsilon/3$$

Now for given $x_0 \in X$ and for any $x \in X$ we have by the triangle inequality for ρ:

$$\text{if } d(x_0, x) < \delta \text{ then } \rho(f(x), f(x_0)) \leq \rho(f(x), f_n(x)) + \rho(f_n(x), f_n(x_0)) + \rho(f_n(x_0), f(x_0)) = \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$$

as desired. ☐

While we're interchanging quantifiers, we would like to point out a stronger form of *continuity* in which we convert $\forall x_0 \ \exists \delta$ to $\exists \delta \ \forall x_0$:

Definition 44. Let $f : X \to Y$. We say that f is **uniformly continuous** if

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall x_0 \in X \ \forall x \in X \ \ d(x, x_0) < \delta \to \rho(f(x), f(x_0)) < \varepsilon$$

Given the terminology, the following theorem is obligatory:

Theorem 45. *Let (f_n) be a sequence of uniformly continuous functions converging uniformly to f. Then f is uniformly continuous.*

Proof. In the proof of theorem 43, strike out "Let x_0 be fixed" and replace "given x_0" with "any x_0". ☐

Theorem 46. *If $f : X \to Y$ is continuous on a* **compact** *domain X, then f is uniformly continuous.*

Proof. Let $\varepsilon > 0$. For each $z \in X$, using continuity, choose $B(z, \delta_z)$ so that for any $y \in B(z, 2\delta_z)$ we have $\rho(f(y), f(z)) < \varepsilon/2$. Find a finite subcover of the open cover $\{B(z, \delta_z)\}$, and let δ_0 be the minimum of all δ_z in the subcover. Now let $x, y \in X$ be arbitrary with $d(y, x) < \delta_0$. There is some z (from the finite subcover) with $d(y, z) \leq \delta_z$, so also $d(x, z) \leq \delta_0 + \delta_z \leq 2\delta_z$ so $\rho(f(y), f(x)) \leq \rho(f(y), f(z)) + \rho(f(z), f(x)) \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$ as desired. ☐

Exercise 21. By example, prove that a sequence of uniformly continuous functions on a compact domain can converge, *not uniformly*, to a uniformly continuous function.

Next we introduce absolute convergence:

Definition 47. We call the series $\sum_{n=0}^{\infty} a_n$ of complex numbers **absolutely convergent** if the series $\sum_{n=0}^{\infty} |a_n|$ converges.

Here "absolute" refers to the absolute value function, but in $\mathbb{C}$ the function $|a_n|$ is usually called modulus. Of course absolute convergence implies convergence. It's possible to deduce this from the M-test using constant functions $f_n(z) = a_n$, but a straightforward proof is more informative:

Proof. For $\varepsilon > 0$ choose N so that for all $n > m > N$ we have $\sum_{k=m+1}^{n} |a_k| < \varepsilon$. This is possible because $\sum_{n=0}^{\infty} |a_n|$ converges, so its partial sum sequence is Cauchy. Then by the triangle inequality, $\left|\sum_{k=m+1}^{n} a_k\right| < \varepsilon$. This shows that the sequence of partial sums $\left(\sum_{k=0}^{N} a_k\right)_{N \in \mathbb{N}}$ is Cauchy, so converges. ☐

The proof of the M-test uses an infinitary triangle inequality, so we start with that:

Proposition 48. *Let (z_n) be a sequence of complex numbers. Then*

$$\left|\sum_{n=1}^{\infty} z_n\right| \leq \sum_{n=1}^{\infty} |z_n|$$

Proof.

$$\left|\sum_{n=1}^{\infty} z_n\right| = \left|\lim_N \sum_{n=1}^{N} z_n\right| = \lim_N \left|\sum_{n=1}^{N} z_n\right| \leq \lim_N \sum_{n=1}^{N} |z_n| = \sum_{n=1}^{\infty} |z_n|$$

...because $| - |$ is continuous and $a_n \leq b_n \to \lim a_n \leq \lim b_n$. ☐

The next theorem is fundamental to our use of power series.

Theorem 49 (Weierstrass M-test)**.** *Let $u_n : X \to \mathbb{C}$ be a sequence of functions having $\forall x \in X \ \ |u_n(x)| \le M_n$ for a sequence of bounds M_n. If $\sum_{n=1}^{\infty} M_n < \infty$, then $\sum_{n=1}^{\infty} u_n$ converges uniformly and absolutely on X.*

Proof of the M-test. First, fix x. Absolute convergence of $\sum_{n=1}^{\infty} u_n(x)$ follows from the comparison test since $|u_n(x)| \le M_n$. This shows that $\sum_{n=1}^{\infty} u_n(x)$ converges pointwise to a function we call $f(x)$.

For uniform convergence, let $\varepsilon > 0$. Choose N so that $\sum_{n=N+1}^{\infty} M_n < \varepsilon$ and let $n > N$. Then for any x we have

$$\left| f(x) - \sum_{i=1}^{n} u_i(x) \right| = \left| \sum_{i=n+1}^{\infty} u_i(x) \right| \le \sum_{i=n+1}^{\infty} |u_i(x)| \le \sum_{i=n+1}^{\infty} M_i < \varepsilon$$

as desired. □

Exercise 22. Let $f(z) = \sum_{n=0}^{\infty} a_n z^n$ be a power series whose (complex) coefficients a_n are bounded above in the sense that $|a_n| \le B \in \mathbb{R}$ for some B. Prove that the sum $\sum_{n=0}^{\infty} a_n z^n$ is uniformly convergent on any disk $B(0, r)$ with $0 < r < 1$.

Chapter 3

Analytic Functions

3.1 Derivatives and Limits

We define the derivative of a complex-valued function exactly as one would expect:

Definition 50. Let $f : \mathbb{C} \to \mathbb{C}$. The **derivative** of f is defined to be:

$$f'(z_0) = \lim_{h\to 0} \frac{f(z_0+h) - f(z_0)}{h}$$

Equivalently (with $z = z_0 + h$):

$$f'(z_0) = \lim_{z\to z_0} \frac{f(z) - f(z_0)}{z - z_0}$$

When the derivative exists, we say that f is **differentiable** at z_0.

If we unpack this definition we have:

$$f'(z_0) = \imath L \in \mathbb{C}\ \ \forall \varepsilon > 0\ \ \exists \delta > 0\ \ \forall z \in \mathbb{C}\ \ 0 < |z - z_0| < \delta \to \left| \frac{f(z) - f(z_0)}{z - z_0} - L \right| < \varepsilon$$

The key difference here is "$\forall z \in \mathbb{C}$" instead of "$\forall z \in \mathbb{R}$" (which, given the implication, amounts to $\forall z \in B(z_0, \delta)$ instead of $\forall z \in (z_0 - \delta, z_0 + \delta)$). This is a much larger, two-dimensional set of z values which must satisfy the condition $\left| \frac{f(z)-f(z_0)}{z-z_0} - L \right| < \varepsilon$. We will see, thanks to this small change, that differentiability is a much stronger condition on complex functions than it is on real functions. For example, if f is differentiable, then its derivative f' is likewise differentiable, and so on.

For now, we distinguish between three conditions later proven equivalent:

Definition 51. Let $U \subseteq \mathbb{C}$ be an open set. We say that $f : U \to \mathbb{C}$ is **differentiable** if $f'(z)$ exists for all $z \in U$. We say that f is **continuously differentiable** if f is differentiable and f' is a continuous function. We say that f is **infinitely differentiable** if all higher derivatives $f^{(n)}(z)$ exist.

Much later, we will prove that these conditions are all equivalent, and replace these terms with the single term **analytic**. For now we use the more direct language.

We begin with an easy example from which we will later generate more.

Example 52. [Derivative of a linear function] If $f(z) = mz + b$ (where m, b are complex constants) then $f'(z) = m$

Proof. The expression $\frac{f(z)-f(z_0)}{z-z_0}$ is always m, independent of both z and z_0, so the limit exists and equals m. □

Before we can prove the various derivative rules, we need some limit laws:

Theorem 53 (Limit Rules)**.** *Let X be a metric space and let $f, g : X \to \mathbb{C}$ or $\mathbb{C}^n$. (Note this includes the case $f, g : X \to \mathbb{R}$.) The following equations are true provided their right hand sides exist.*

1. $\lim_{x\to x_0} f(x) \pm g(x) = \lim_{x\to x_0} f(x) \pm \lim_{x\to x_0} g(x)$
2. $\lim_{x\to x_0} f(x)g(x) = \lim_{x\to x_0} f(x) \cdot \lim_{x\to x_0} g(x)$
3. $\displaystyle\lim_{x\to x_0} f(x)/g(x) = \frac{\lim_{x\to x_0} f(x)}{\lim_{x\to x_0} g(x)}$
4. *If* $\square : \mathbb{C}^2 \to \mathbb{C}$ *is a continuous binary operation then* $\lim_{x\to x_0} f_1(x)\square f_2(x) = \lim_{x\to x_0} f_1(x)\square \lim_{x\to x_0} f_2(x)$
5. *If* $\lim_{x\to x_0} g(x) = y_0$ *and f is continuous at y_0, then* $\lim_{x\to x_0} f(g(x)) = f(y_0)$

Proof. First we prove 5. Let $\varepsilon > 0$. Choose $\rho > 0$ so that $|y - y_0| < \rho \to |f(y) - f(y_0)| < \varepsilon$. Choose $\delta > 0$ so that $0 < |x - x_0| < \delta \to |g(x) - y_0| < \rho$. We may substitute $y = g(x)$ and chain together the implications: $0 < |x - x_0| < \delta \to |f(g(x)) - f(y_0)| < \varepsilon$ as desired.

Applying 5 with $g(x) = (f_1, f_2)$ and $f = \square$ yields 4. It remains to prove the continuity of arithmetical operations. In each case we assume $a\square b = c$ and $\varepsilon > 0$, and give a neighborhood $U = B(a, \delta_1) \times B(b, \delta_2) \subseteq \mathbb{C} \times \mathbb{C}$ so that $\square(U) \subseteq B(c, \varepsilon)$. For addition and subtraction we may take $B(a, \varepsilon/2) \times B(b, \varepsilon/2)$, by the triangle inequality. For multiplication suppose hypothetically $\delta_1 > 0$ and $\delta_2 > 0$ are already chosen. If $(x, y) \in B(a, \delta_1) \times B(b, \delta_2)$ then $|xy - ab| \le |xy - bx + bx - ab| \le |x||y - b| + |b||x - a| < |x|\delta_2 + |b|\delta_1 \le |x - a + a|\delta_2 + |b|\delta_1 \le \delta_1\delta_2 + |a|\delta_2 + |b|\delta_1$. We use this computation to guide our choices of δ_1 and δ_2. We may make $|b|\delta_1 \le \varepsilon/3$ by choosing $\delta_1 = \frac{\varepsilon}{3|b|}$ (or 1 if $b = 0$). We may then assure that the other two terms are each less than or equal to $\varepsilon/3$ by choosing $\delta_2 = \min(\frac{\varepsilon}{3\delta_1}, \frac{\varepsilon}{3|a|})$. (Or simply $\delta_2 = \frac{\varepsilon}{3\delta_1}$ if $a = 0$.)

We can deduce the division law from the multiplication law if we can show that $r(x) = 1/x$ is a continuous function at any $x_0 \ne 0$. Let $\varepsilon > 0$ and suppose δ somehow already chosen, with $|x - x_0| < \delta$. Then

$$\left|\frac{1}{x} - \frac{1}{x_0}\right| = \left|\frac{x_0 - x}{xx_0}\right| < \left|\frac{\delta}{(x - x_0 + x_0)x_0}\right| = \frac{\delta}{|x - x_0 + x_0||x_0|} \le \frac{\delta}{||x_0| - \delta||x_0|} = \frac{\delta}{(|x_0| - \delta)|x_0|}$$

provided $|x_0| - \delta$ is positive. If $\delta < \frac{|x_0|}{2}$ it is, and we may furthermore continue

$$\le \frac{\delta}{\frac{|x_0|}{2}|x_0|} = \frac{2\delta}{|x_0|^2} \le^? \varepsilon$$

Therefore if $\delta = \min(\frac{x_0}{2}, \frac{\varepsilon|x_0|^2}{2})$ we have our result. $\square$

Exercise 23. Let f be a function and let g be a continuous function. Consider the limits $\lim_{x\to x_0} f(g(x))$ and $\lim_{y\to g(x_0)} f(y)$. Does either limit's existence imply the other's? Explain with proof. If they both exist are they necessarily equal? Explain with proof.

Exercise 24. Assess whether each function is differentiable. Answer with proof.

1. $f(z) = \overline{z}$, at any point $z \in \mathbb{C}$.
2. $f(z) = |z|^2$, at any point $z \in \mathbb{C}$. (Hint: Consider $z = 0$ and $z \ne 0$ as separate cases.)
3. $f(z) = \frac{1}{z}$.

We may now begin the theory of differentiable functions in earnest.

Proposition 54. *If f is differentiable, then f is continuous.*

Proof.

$$\lim_{w\to z} f(w) - f(z) = \lim_{w\to z} \frac{f(w) - f(z)}{w - z} \cdot (w - z) = f'(z) \lim_{w\to z}(w - z) = 0$$

$\square$

Theorem 55 (Linear Approximation)**.** *Let f be a complex function defined in a neighborhood of a point z. Then f is differentiable at z if and only if there is a linear function L and an "error" function g with $f = L + g$, $g(z) = 0$, and $\lim_{w\to z} \frac{g(w)}{w-z} = 0$.*

Proof. A function f is differentiable at z exactly when there is a value m so that $\lim_{w\to z} \frac{f(w)-f(z)}{w-z} - m = 0$ Equivalently $\lim_{w\to z} \frac{f(w)-f(z)-m(w-z)}{w-z} = 0$. Notice that $L(w) = f(z)+m(w-z)$ is linear and the numerator suffices for $g(w)$. Conversely if $f = L + g$ as above, then $f'(z) = L'(z) + g'(z) = m + 0 = m$, so $f'(z)$ exists. □

Exercise 25. Compare the following two statements about limits:

$$\begin{aligned} \lim_{w\to z} g(w) =&0 \quad \text{versus} \\ \lim_{w\to z} \frac{g(w)}{w-z} =&0 \end{aligned}$$

One of these statements is logically *stronger*, meaning that it implies the other. Let us call the stronger statement "S" and the weaker "W".

1. Which is which?
2. Prove $S \to W$.
3. Illustrate $\neg(W \to S)$ by counterexample (and prove it's a counterexample).
4. Let f be a complex function defined in a neighborhood of a point z. Now f is differentiable at z if and only if there is a linear function L and an "error" function g with $f = L + g$ and [S] and $g(z) = 0$. What familiar condition on f is equivalent to the above, if we replace S with W? Answer without proof.

Theorem 56. *[Derivative Rules] Let $f, g : U \to \mathbb{C}$. The following equations are true provided the right hand sides exist:*

1. *(sum/difference rule)* $\frac{d}{dz}(f(z) \pm g(z)) = f'(z) \pm g'(z)$
2. *(constant multiple rule)* $\frac{d}{dz}(kf(z)) = kf'(z)$
3. *(product rule)* $\frac{d}{dz}(f(z) \cdot g(z)) = f'(z)g(z) + f(z)g'(z)$
4. *(quotient rule)* $\frac{d}{dz}(f(z)/g(z)) = \frac{f'(z)g(z)-f(z)g'(z)}{(g(z))^2}$
5. *(chain rule)* $\frac{d}{dz}(f(g(z))) = f'(g(z))g'(z)$

Proof. The proofs are the same as for real functions $f : \mathbb{R} \to \mathbb{R}$, but since students may not have seen proof-based calculus, we provide them.

1. $\frac{d}{dz}(f(z) + g(z)) = \lim_{w\to z} \frac{f(w)+g(w)-f(z)-g(z)}{w-z} = \lim_{w\to z} \frac{f(w)-f(z)}{w-z} + \lim_{w\to z} \frac{g(w)-g(z)}{w-z} = f'(z) + g'(z)$
2. Exercise (below).
3. $$\begin{aligned} \frac{d}{dz}(f(z)g(z)) &= \lim_{w\to z} \frac{f(w)g(w) - f(z)g(z)}{w-z} \\ &= \lim_{w\to z} \frac{f(w)g(w) - f(z)g(w) + f(z)g(w) - f(z)g(z)}{w-z} \\ &= \left(\lim_{w\to z} \frac{f(w)-f(z)}{w-z} g(w)\right) + \left(\lim_{x\to z} f(z) \frac{g(w)-g(z)}{w-z}\right) \\ &= f'(z)g(z) + f(z)g'(z) \end{aligned}$$

4. Exercise (below).

5. The most attractive demonstration of the chain rule is sadly wrong. It goes like this:

$$\begin{aligned}\frac{d}{dz}(f(g(z))) &= \lim_{w\to z}\frac{f(g(w))-f(g(z))}{w-z}\\ &= \lim_{w\to z}\frac{f(g(w))-f(g(z))}{g(w)-g(z)}\cdot\lim_{w\to z}\frac{g(w)-g(z)}{w-z} \text{ but see below}\\ &= f'(g(z))\cdot g'(z)\end{aligned}$$

The problem is that the introduced denominator $g(w)-g(z)$ may be zero, even though $w-z$ is not. We therefore replace the fraction $\frac{f(g(w))-f(g(z))}{g(w)-g(z)}$ with a piecewise defined function designed to deal with that possibility:

$$\frac{d}{dz}(f(g(z))) = \lim_{w\to z}\frac{f(g(w))-f(g(z))}{w-z} \tag{3.1}$$

$$= \lim_{w\to z}\left[\left\{\begin{array}{ll}\frac{f(g(w))-f(g(z))}{g(w)-g(z)} & \text{if } g(w)\neq g(z)\\ f'(g(z)) & \text{if } g(w)=g(z)\end{array}\right\}\cdot\frac{g(w)-g(z)}{w-z}\right] \tag{3.2}$$

$$= f'(g(z))\cdot g'(z) \tag{3.3}$$

Notice that 3.1 and 3.2 remain equal when $g(w)=g(z)$, because both sides are zero. To evaluate the limit 3.2 we used the product law and a continuity argument: The piecewise part may be regarded as the composition of $w\mapsto g(w)$ (continuous at $w=z$) with the function $x\mapsto\left\{\begin{array}{ll}\frac{f(x)-f(g(z))}{x-g(z)} & \text{if } x\neq g(z)\\ f'(g(z)) & \text{if } x=g(z)\end{array}\right.$, which is continuous at $x=g(z)$ by the definition of the derivative $f'(g(z))$.

□

Exercise 26. The chain rule makes implicit use of the important function $q(w,z)$, which we will revisit in section 3.4.

$$q(w,z)=\left\{\begin{array}{ll}\frac{f(w)-f(z)}{w-z} & \text{if } w\neq z\\ f'(z) & \text{if } w=z\end{array}\right.$$

(However, in the chain rule prove we have the substitution $q(w,g(z))$.)

Prove:

1. The function q is symmetric: $q(w,z)=q(z,w)$.
2. If $f(z)$ is differentiable on $U\subseteq C$, then $q(w,z)$ is a continuous function of w for each fixed $z\in U$ and vice versa.

Exercise 27. Prove parts 2 and 4 of the previous theorem. For the quotient rule, use common denominator and take inspiration from the product rule.

Our last derivative rule is for inverse functions:

Theorem 57. *Let $g(f(z))=z$ for all $z\in G$ (open), f continuous at z. Then*

$$f'(z)=\frac{1}{g'(f(z))}$$

provided the right hand side exists.

Proof. The chain rule suffices if we assume $f'(z)$ exists, but we wish to prove it exists. We need to partly mimic that proof:

$$1=\frac{d}{dz}g(f(z))=\lim_{w\to z}\frac{g(f(w))-g(f(z))}{w-z}=\lim_{w\to z}\frac{g(f(w))-g(f(z))}{f(w)-f(z)}\frac{f(w)-f(z)}{w-z}$$

Note: Since $g(f(z)) = z$ in G, we know f is injective on G (open), so the introduced denominator $f(w) - f(z)$ is nonzero. Now $\lim_{w \to z} \frac{g(f(w)) - g(f(z))}{f(w) - f(z)}$ is $g'(f(z))$, as in the chain rule proof, and is presumed nonzero. We may divide both sides by this limit and apply the quotient law for limits to conclude. $\square$

Corollary 58. *If $g : X \to Y$ ($X \subseteq \mathbb{C}$, $Y \subseteq \mathbb{C}$) is a differentiable bijection with nonzero derivative and continuous right inverse function f (so $g \circ f = \mathrm{id}$). Then f is also differentiable, and $f'(z) = \frac{1}{g'(f(z))}$ for all $z \in Y$.*

We would like to apply the previous theorem to square roots, but first we must discuss the square root of a complex number. Notice that any complex number z can be written in the form $r \operatorname{cis}(\theta)$, with $r \geq 0$ and $\theta \in (-\pi, \pi]$ (or alternatively $[0, 2\pi)$). Then let $w = \sqrt{r} \operatorname{cis}(\theta/2)$. Explicitly as a function of z, we have $w = \sqrt{|z|} \operatorname{cis}(\arg(z)/2)$, where $\arg(z)$ is an **argument function**, or angle, of z. From the properties of the function cis (Proposition 10), we see that $w^2 = z$ and it's legitimate to write $w = \sqrt{z}$. Since $(\sqrt{z})^2 = z$, we appear to be in the situation of theorem 57, but only if we can say $\sqrt{z}$ is continuous.

Notice that $\operatorname{cis}(\theta) = \cos(\theta) + i \sin(\theta)$ is continuous by sum and multiple rules, since cos and sin are continuous. Furthermore $r = |z|$ is a continuous function of z, and $\sqrt{-}$ is continuous on $[0, \infty)$. It follows that $f(z) = \sqrt{z}$ is continuous whenever $\arg(z)$ is continuous.

Proposition 59. *Let $\arg : \mathbb{C} - \{0\} \to (-\pi, \pi]$ with $\arg(r \operatorname{cis}(\theta)) = \theta$ whenever $\theta \in (-\pi, \pi]$. Then $\arg$ is continuous at z when $\arg(z)$ is in the interior of the range and discontinuous at z when $arg(z)$ is on the boundary. The statement remains true if the range $(-\pi, \pi]$ is replaced with any other halfopen interval of length 2π.*

Proof. Let $\theta_0 = \arg(z)$. If $\theta_0 \in (-\pi, \pi)$, let $\arg(z) \in B(\theta_0, \varepsilon) \subseteq (-\pi, \pi)$. Then $\arg^{-1}(B(\theta_0, \varepsilon))$ is the wedge $\{r \operatorname{cis}(\theta) | \theta \in B(\theta_0, \varepsilon)\}$, which is an open set containing z. On the other hand if $\arg(z) = \theta_0 = \pi$, then z is negative real and every neighborhood $B(z, \delta)$ contains points y with negative argument, so $\arg^{-1}(B(\theta_0, 1))$ is not open.

To elaborate, notice that by restriction $\arg(z) : S^1 \to (-\pi, \pi]$. This map has compact domain but noncompact image, so it *must* be discontinuous somewhere. $\square$

Returning to the square root, we see that $w = \sqrt{z} = \sqrt{|z|} \operatorname{cis}(\arg(z)/2)$ is discontinuous exactly when z is negative real. A technical argument shows that it is continuous at $z = 0$ as well. The derivative formula $\frac{d}{dz}\sqrt{z} = \frac{1}{2\sqrt{x}}$ is valid for all z except negative reals, where $\sqrt{z}$ is discontinuous, and zero, where z^2 has zero derivative. A different choice of range for $\arg(z)$ gives a different square root function, but always one discontinuous on an open ray from the origin.

Exercise 28. Recall that the kernel of the cis map is known. Let $w, z \in \mathbb{C}$. By using polar coordinates for w and z, prove that $\arg(wz) = \arg(w) + \arg(z) + 2\pi n$ for some $n \in \mathbb{Z}$. Here the $2\pi n$ term records the inherent ambiguity of angle measure. Except for this term, the law should remind you of a logarithmic function.

3.2 Cauchy Riemann Equations and Harmonic Functions

Because a single complex number determines two real numbers, its real and imaginary part, a function $f : \mathbb{C} \to \mathbb{C}$ can be thought of as a function with two real inputs and two real outputs, or even two distinct functions $u(x, y) = Re(f(x + iy))$ and $v(x, y) = Im(f(x + iy))$, so that $f(x + iy) = u(x, y) + iv(x, y)$, where $u : \mathbb{R}^2 \to \mathbb{R}$ and $v : \mathbb{R}^2 \to \mathbb{R}$. This perspective is arguably more complicated, but allows us to relate properties of complex functions to the multivariable calculus of real functions. The first application of this line of thinking is called the Cauchy Riemann Equations

Proposition 60 (Cauchy-Riemann Equations)**.** *Let $f(z) = f(x + iy) = u(x, y) + iv(x, y)$ be a complex function differentiable at $z \in \mathbb{C}$ (where $u : \mathbb{R}^2 \to \mathbb{R}$ and $v : \mathbb{R}^2 \to \mathbb{R}$). Then at z:*

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \textit{and} \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$$

Proof. If the limit $f'(z) = \lim_{h\to 0} \frac{f(z+h)-f(z)}{h}$ exists, then it must agree with limits in which the variable h remains purely real or purely imaginary. So if $f'(z)$ exists, then

$$f'(z) = \lim_{\substack{h\to 0 \\ h\in\mathbb{R}}} \frac{f(z+h)-f(z)}{h} = \lim_{\substack{h\to 0 \\ h\in\mathbb{R}}} \frac{f(z+ih)-f(z)}{ih}$$

If $f'(x+iy) = u(x,y) + iv(x,y)$ then each limit is a combination of partial derivatives of u and v:

$$f'(z) = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} = \frac{1}{i}\left(\frac{\partial u}{\partial y} + i\frac{\partial v}{\partial y}\right)$$

Setting the real parts equal yields the first Cauchy Riemann equation, and setting imaginary parts equal yields the second.

□

If the function $f : \mathbb{C} \to \mathbb{C}$ is a *twice* differentiable, then the Cauchy Riemann equations apply to its derivative $f'(z) = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x}$. This means

$$\frac{\partial}{\partial x}\frac{\partial u}{\partial x} = \frac{\partial}{\partial y}\frac{\partial v}{\partial x} \quad \text{and} \quad \frac{\partial}{\partial y}\frac{\partial u}{\partial x} = -\frac{\partial}{\partial x}\frac{\partial v}{\partial x}$$

Applying the same to the other formula $if'(z) = \frac{\partial u}{\partial y} + i\frac{\partial v}{\partial y}$ produces

$$\frac{\partial}{\partial x}\frac{\partial u}{\partial y} = \frac{\partial}{\partial y}\frac{\partial v}{\partial y} \quad \text{and} \quad \frac{\partial}{\partial y}\frac{\partial u}{\partial y} = -\frac{\partial}{\partial x}\frac{\partial v}{\partial y}$$

Equating mixed partial derivatives and adding these equations yields:

$$\frac{\partial^2}{\partial x^2}u + \frac{\partial^2}{\partial y^2}u = 0 \quad \text{and} \quad \frac{\partial^2}{\partial x^2}v + \frac{\partial^2}{\partial y^2}v = 0$$

That is, the real and imaginary parts of a twice differentiable complex function are harmonic, in the following sense.

Definition 61. The second order differential operator above is called the **Laplace operator**, typically written

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$$

A function u for which $\nabla^2 u = 0$ is called a **harmonic function**.

The operator ∇^2 calculates an overall concavity in two dimensions. Harmonic functions lack local minima and maxima and have an overall curvature which is balanced at every point between positive and negative. A harmonic function is a solution to the constant-time version of the heat equation. That is, a harmonic function can describe the temperature function of a body whose temperature is not changing with time.

3.3 Power Series

Using Example 52 and the combinations of Theorem 56 we can quickly prove that all polynomials are differentiable, and that their derivatives are calculated exactly as one would expect:

Example 62. The derivative of $\displaystyle\sum_{n=0}^{N} a_n z^n$ is $\displaystyle\sum_{n=0}^{N} a_n n z^{n-1}$

We would like to extend this example to include infinite sums, but first we must consider matters of convergence.

Definition 63. A **power series** is a function of z of the form

$$f(z) = \sum_{n=0}^{\infty} a_n z^n, \text{ or more generally } f(z) = \sum_{n=0}^{\infty} a_n (z-a)^n$$

When $\sum_{n=0}^{\infty} a_n z^n$ converges we say f converges at z. The set of z at which f converges is called the domain of convergence.

The generalization to sums of the form $\sum a_n (z-a)^n$ is unnecesary in the case of polynomials because powers $(z-a)^n$ may be simply expanded. In the case of series, such expansion introduces convergence problems. For example, what is the constant term of $\sum_{n=0}^{\infty} 1(z+1)^n$ after expansion?

Example 64. The function $f(z) = \sum_{n=0}^{\infty} z^n$ is a **geometric series**. It converges uniformly on any region $B(0,r)$ with $0 < r < 1$, but if $|z| \geq 1$ it diverges. (More generally, $\sum_{n=0}^{\infty}(z-a)^n$ converges uniformly on $B(a,r)$ and diverges when $|z-a| \geq 1$.)

Proof. Let $0 < r < 1$. On $B(0,r)$ we have $|z| < r$ so $|z^n| < r^n$. The values r^n satisfy the requirements of the Weierstrass M-test (Theorem 49) because $\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}$, so $\sum_{n=0}^{\infty} z^n$ is uniformly convergent on $B(0,r)$. If on the other hand $|z| \geq 1$, then the terms of $\sum_{n=0}^{\infty} z^n$ do not even approach zero, so the series diverges. $\square$

We sometimes prove convergence or divergence of series by comparison to a geometric series. This technique is at the heart of the root test:

Theorem 65 (Cauchy's Root Test)**.** *Let $\sum_{n=0}^{\infty} a_n$ be a series, and let $C = \limsup |a_n|^{1/n}$.*

1. *If $C < 1$ then the series absolutely converges.*
2. *If $C > 1$ then the series diverges.*

Proof. In the first case, we have $\limsup |a_n|^{1/n} < 1$, so we may choose s with $\limsup |a_n|^{1/n} < s < 1$. For all but finitely many n we have $|a_n|^{1/n} < s$, so $|a_n| < s^n$. Thus $\sum_{n=0}^{\infty} |a_n| < \sum_{n=0}^{\infty} s^n +$ [something finite] and $\sum a_n$ is absolutely convergent.

In the second case, we have $1 < \limsup |a_n|^{1/n}$. This time we choose s with $1 < s < \limsup |a_n|^{1/n}$, and as before $|a_n| > s^n$ for infinitely many n. Then the terms a_n do not even approach zero, so $\sum_{n=0}^{\infty} a_n$ diverges. $\square$

Theorem 66 (Cauchy-Hadamard test)**.** *Let $f(z) = \sum_{n=0}^{\infty} a_n (z-a)^n$ be a power series. Let*

$$R = 1/\limsup |a_n|^{1/n} \qquad \textit{(with } R = \infty \textit{ if the } \limsup \textit{ is } 0.)$$

Then $f(x)$ converges uniformly and absolutely on any region $B(a,r)$ with $0 < r < R$, but if $|z-a| > R$ it diverges.

Proof. As in the proof of the root test, we may choose s so that $1/R = \limsup |a_n|^{1/n} < s < 1/r$, and for all but finitely many n we have $|a_n|^{1/n} < s$ and $|a_n| < s^n$. For these n, and for all $z \in B(a,r)$, $|a_n(z-a)^n| < s^n r^n = (sr)^n$, so the M-test gives uniform and absolute convergence.

If on the other hand $|z-a| > R$, then $C = \limsup |a_n(z-a)^n|^{1/n} = |z-a| \limsup |a_n|^{1/n} = |z-a|/R > 1$, so the root test guarantees divergence. $\square$

Although the Hadamard test is powerful, the ratio test is often easier to use. Before we prove it we need a technical lemma:

Lemma 67. *Let a_n be a sequence of positive real numbers. If $\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = L$ then $\lim_{n\to\infty} a_n^{1/n} = L$.*

Proof. Let $b_n = \ln_{\mathbb{R}}(a_n)$, where $\ln_{\mathbb{R}}$ is the ordinary real natural logarithm function. We must prove that if $\lim_{n\to\infty} b_{n+1} - b_n = K$ then $\lim_{n\to\infty} \frac{b_n}{n} = K$. (These are equivalent since e^x and $\ln_{\mathbb{R}}$ are continuous.)

First, assume $K = 0$, and let $\varepsilon > 0$. Choose N beyond which $b_{n+1} - b_n < \varepsilon/2$, and let $n > N$. We compute: $\frac{b_n}{n} = \frac{b_n - b_N}{n} + \frac{b_N}{n} \leq \frac{(n-N)\varepsilon/2}{n} + \frac{b_N}{n} \leq \varepsilon/2 + \frac{b_N}{n}$. Thus if $n > \max(N, 2b_N/\varepsilon)$ we have $\frac{b_n}{n} < \varepsilon$ as desired.

If $K \neq 0$, we apply the previous case to the modified sequence $b_n - nK$.

$\square$

Theorem 68. *[Ratio Test for Power Series] Let $f(z) = \sum_{n=0}^{\infty} a_n(z-a)^n$ be a power series. If $\lim_{n\to\infty} \left|\frac{a_n}{a_{n+1}}\right|$ exists then it equals the radius of convergence of f.*

Proof. Apply the lemma and the Hadamard test to the sequence $|a_n|$. □

Exercise 29. Give examples, with proof, of a power series . . .

1. . . . with infinite radius of convergence.
2. . . . with radius of convergence equal to 11.
3. . . . with zero radius of convergence.

Lemma 69. *Let $A = \sum_{n=0}^{\infty} a_n$ and $B = \sum_{n=0}^{\infty} b_n$ be convergent series in $\mathbb{C}$.*

1. $\sum_{n=0}^{\infty} ka_n = kA$.
2. $\sum_{n=0}^{\infty} a_n + b_n = A + B$.
3. $\sum_{n=0}^{\infty} c_n = AB$ *where* $c_n = \sum_{k=0}^{n} a_k b_{n-k}$, *provided at least one series converges absolutely.*

Proof. The first two are straightforward, but the third is delicate. Let $A_N = \sum_{n=0}^{N} a_n$ and $B_N = \sum_{n=0}^{N} b_n$.

1. An easy exercise.
2. $\sum_{n=0}^{\infty}(a_n + b_n) = \lim_N \sum_{n=0}^{N}(a_n + b_n) = \lim_N (A_N + B_N) = \lim_N A_N + \lim_N B_N = A + B$ as desired.
3. Link to Figure

   Let $N \in \mathbb{N}$, to be specifically chosen later.

$$\begin{aligned}
\left|AB - \sum_{n<N} c_n\right| &= \left|\sum_{i<N} a_i B + \sum_{i\geq N} a_i B - \sum_{i+j<N} a_i b_j\right| \\
&= \left|\sum_{i<N} a_i \left(B - \sum_{j<N-i} b_j\right) + \sum_{i\geq N} a_i B\right| \\
&= \left|\sum_{i<N} \left(a_i \sum_{j\geq N-i} b_j\right) + \sum_{i\geq N} a_i B\right| \\
&\leq \sum_{i\leq N} |a_i| \left|\sum_{j\geq N-i} b_j\right| + \sum_{i\geq N} |a_i||B|
\end{aligned}$$

In the right term, $|B|$ is a fixed value and the tail $\sum_{i\geq N} |a_i|$ is small for large N by absolute convergence. In the left term, we can control $|a_i|$ if i is large and $|\sum_{j\geq N-i} b_j|$ if i is small (so that j is large). Therefore we split the sum in half. Let $0 \leq M < N$. Here M will also be specifically chosen later. We continue...

$$= \overbrace{\sum_{0\leq i<M} |a_i| \left|\sum_{j>N-i} b_j\right|}^{\alpha} + \overbrace{\sum_{M\leq i<N} |a_i| \left|\sum_{j>N-i} b_j\right|}^{\beta} + \overbrace{\sum_{i\geq N} |a_i||B|}^{\gamma}$$

Without further delay let $\varepsilon > 0$.

We control term β first. The sequence $\left|\sum_{j>N-i} b_j\right|$ (as a function of N) converges to 0, so it has an upper bound U_B , independent of N and M. As a function of N, the convergent sequence $\sum_{i\geq N} |a_i|$ is Cauchy, so its differences $\sum_{M\leq i<N} |a_i|$ can be made small by choosing M sufficiently large. Choose

M so that (for all $N > M$) we have $\sum_{M \leq i < N} |a_i| < \frac{\varepsilon}{3U_B}$. For this M and for any $N > M$ we have $\beta < \frac{\varepsilon}{3}$

Next we control term α. Having chosen M we may treat $\sum_{0 \leq i < M} |a_i|$ as a constant V_A. Again because the sequence $\left|\sum_{j > N-i} b_j\right|$ converges to 0, we may choose K so that if $N - i > K$ then $\left|\sum_{j > N-i} b_j\right| < \frac{\varepsilon}{3V_A}$. Now for any $N > K + M$, we have $N - i > K$ and so $\alpha < \frac{\varepsilon}{3}$.

Finally we control term γ, in which $|B|$ is a fixed constant. We assume $|B| \neq 0$, the contrary case being trivial. Since the series $\sum a_i$ is absolutely convergent there is some G so that if $N > G$ then $\sum_{i \geq N} |a_i| < \frac{\varepsilon}{3|B|}$, and $\gamma < \frac{\varepsilon}{3}$.

If $N > \max(K + M, G)$ then $\left|AB - \sum_{n<N} c_n\right| < \varepsilon$ as required.

□

Theorem 70. *Let $f(z) = \sum_{n=0}^{\infty} a_n(z-a)^n$ and $g(z) = \sum_{n=0}^{\infty} b_n(z-a)^n$ be power series with radius of convergence at least R. Then (the series below converge on $B(a, R)$ and)*

1. $\sum_{n=0}^{\infty}(ka_n)(z-a)^n = kf(z)$.
2. $\sum_{n=0}^{\infty}(a_n + b_n)(z-a)^n = f(z) + g(z)$.
3. $\sum_{n=0}^{\infty} c_n(z-a)^n = f(z)g(z)$ *where* $c_n = \sum_{k=0}^{n}(a_k b_{n-k})$.

Proof. All three parts follow from Lemma 69. □

Exercise 30. Find series $A = \sum_{n=0}^{\infty} a_n$ and $B = \sum_{n=0}^{\infty} b_n$, both convergent but neither absolutely convergent, for which $\sum_{n=0}^{\infty} c_n$ does not converge to AB. (Hints: Choose $\sum a_n$ and $\sum b_n$ real and alternating but decreasing in absolute value. This will guarantee their convergence but leave many options to violate absolute convergence (a condition you must avoid in light of Lemma 69). Argue that $c_n = \sum_{k=0}^{n} a_k b_{n-k}$ is a sum with positive terms. Bound them below, deduce a lower bound for c_n, and force $\sum c_n$ to diverge.)

Exercise 31. Calculate a power series for $\frac{1}{(1-z)^2}$ by the series multiplication rule, and find its radius of convergence.

Definition 71. Let $f(z) = \sum_{n=0}^{\infty} a_n(z-a)^n$ be a power series. The **formal derivative** is the power series

$$\sum_{n=0}^{\infty} a_n n(z-a)^{n-1}$$

We wish to show that the formal derivative equals the derivative, but first we need to check its convergence:

Theorem 72. *Let $\sum_n a_n(z-a)^n$ be a power series. Then the formal derivative $\sum_n na_n(z-a)^{n-1}$ has the same radius of convergence.*

Proof. First notice that $\sum_n na_n(z-a)^{n-1}$ converges iff $\sum_n na_n(z-z_0)^n$ converges, because $z - z_0$ is a constant multiple in the sum on n. It's more convenient to use the Hadamard test on the latter:

$$\limsup \sqrt[n]{|na_n|} = \limsup \sqrt[n]{n}\sqrt[n]{|a_n|} = \lim_n \sqrt[n]{n} \limsup \sqrt[n]{|a_n|} = \limsup \sqrt[n]{|a_n|}$$

(This is a valid limit law when the limit is defined. The limit $\lim_n \sqrt[n]{n}$ can be evaluated by Lemma 67.) □

Corollary 73. *As above, the formal antiderivative has the same radius of convergence.*

Proof. The formal antiderivative of f has some radius of convergence, which matches that of *its* formal derivative (by Theorem 72), which of course is f. □

Theorem 74. *The formally derived series converges to the derivative (inside the radius of convergence $B(a, R)$).*

Proof. For simplicity we assume $a = 0$. Let $f(z) = \sum_{n=0}^{\infty} a_n z^n$. We write f_1 for the formally derived series and f' for the derivative.

To show $f_1(z) = f'(z)$ need to show that the difference $\frac{f(w)-f(z)}{w-z} - f_1(z)$ approaches 0 as $w \to z$. Assume $w, z \in B(0, s)$ where $s < R$. Let $f(z) = s_n(z) + R_n(z)$, where $s_n(z) = \sum_{k=0}^{n-1} a_n z^n$. Note $f'(z) = \frac{d}{dz} \lim_n s_n(z)$ whereas $f_1 = \lim_n \frac{d}{dz} s_n(z)$, so our obligation here is a case of interchanging limits.

$$\begin{aligned}
\frac{f(w)-f(z)}{w-z} - f_1(z) &= \frac{s_n(w) + R_n(w) - s_n(z) - R_n(z)}{w-z} - f_1(w) \\
&= \frac{s_n(w) - s_n(z)}{w-z} - f_1(z) + \frac{R_n(w) - R_n(z)}{w-z} \\
&= \left(\frac{s_n(w) - s_n(z)}{w-z} - s_n'(z) \right) + (s_n'(z) - f_1(z)) + \frac{R_n(w) - R_n(z)}{w-z}
\end{aligned}$$

Let $\varepsilon > 0$ and fix $z \in B(a, s)$. We will bound each of the three terms by $\varepsilon/3$. We need only find one n for which this is possible, since the left side of the equation is independent of n. The last term is hardest:

$$\begin{aligned}
\left| \frac{R_n(w) - R_n(z)}{w-z} \right| &= \left| \frac{\sum_{k=n}^{\infty} a_k (w^k - z^k)}{w-z} \right| \\
&= \left| \sum_{k=n}^{\infty} a_k (w^{k-1} + z w^{k-2} + \ldots + z^{k-1}) \right| \\
&= \sum_{k=n}^{\infty} \left| a_k (w^{k-1} + z w^{k-2} + \ldots + z^{k-1}) \right| \\
&\leq \sum_{k=n}^{\infty} |a_k| k s^k
\end{aligned}$$

The last expression is the tail of a series (the formal derivative power series at $z = s$, times s) known convergent by 72. Thus for n sufficiently large the third term can be made less than $\varepsilon/3$, for *all* $w \in B(a, s)$.

As for the second term $(s_n'(z) - f_1(z))$, $f_1(z)$ is by definition $\lim_{n\to\infty} s_n'(z)$, so for n large enough it can be made less than $\varepsilon/3$, independent of w.

Now fix n for which the last two terms are small as described. The first term $\frac{s_n(w)-s_n(z)}{w-z} - s_n'(z)$ is the difference between a function and its limit as w approaches z, so for w in some neighborhood $B(z, \delta)$ it can be bounded (in absolute value) by $\varepsilon/3$.

We've shown that for $\varepsilon > 0$, we can find some n and $\delta > 0$ so that if $0 < |w - z| < \delta$ then $\left| \frac{f(w)-f(z)}{w-z} - f_1(z) \right| < \varepsilon$. We disregard n since the expression doesn't depend on n, and conclude that $\lim_{w\to z} \frac{f(w)-f(z)}{w-z} = f_1(z)$ as desired. □

Corollary 75. *Let $f(z) = \sum_{n=0}^{\infty} a_n (z-a)^n$ be a power series with radius of convergence R. Then the formal antiderivative is an antiderivative on $B(a, R)$.*

Proof. Compare the derivative and the formal derivative of the formal antiderivative of f. □

Exercise 32. Theorem 74 proves that $\lim_{n\to\infty} \frac{d}{dz} s_n(z)$ equals $\frac{d}{dz} \lim_{n\to\infty} s_n(z)$ for certain functions s_n. In this exercise you show that the equality is not a generic law. Find functions $s_n : \mathbb{R} \to \mathbb{R}$ or $s_n : (a, b) \to \mathbb{R}$ and z in their domain for which the limits:

$$\lim_{n\to\infty} \frac{d}{dz} s_n(z) \text{ and } \frac{d}{dz} \lim_{n\to\infty} s_n(z)$$

exist and are unequal.

Exercise 33. Calculate a power series for $\frac{1}{(1-z)^2}$ by the series differentiation rule, and find its radius of convergence. Compare to the previous problem.

Power series give us a versatile example of analytic functions. We will show (much) later that, conversely, all analytic functions are locally equal to power series.

At this point our general theory of analytic functions is quite strong, but our variety of examples is not. Next we use power series to introduce complex versions of three important real-valued functions.

Definition 76. We define:

$$e^z \equiv \sum_{n=0}^{\infty} \frac{z^n}{n!} = 1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \frac{z^4}{24} + \dots$$
$$\cos(z) \equiv \sum_{\substack{n=0 \\ n \text{ even}}}^{\infty} \frac{(-1)^{\frac{n}{2}} z^n}{n!} = 1 - \frac{z^2}{2} + \frac{z^4}{24} - \frac{z^6}{720} + \dots$$
$$\sin(z) \equiv \sum_{\substack{n=1 \\ n \text{ odd}}}^{\infty} \frac{(-1)^{\lfloor \frac{n}{2} \rfloor} z^n}{n!} = x - \frac{z^3}{6} + \frac{z^5}{120} - \frac{z^7}{5040} + \dots$$

Proposition 77. *Each function e^z, $\cos(z)$, and $\sin(z)$ has infinite radius of convergence.*

Proof. For e^x, the ratio test (68) gives $\lim_n \left|\frac{a_n}{a_{n+1}}\right| = \lim_n (n+1) = \infty$. The ratio test doesn't apply for cos and sin because their zero terms, but its use for e^x allows us to conclude that $\limsup \left|\frac{1}{n!}\right|^{1/n} = 0$, and the lim sup's for cos and sin must agree with this result, so by Cauchy-Hadamard they also have infinite radius of convergence. □

Theorem 78 (Properties of e^z, $\cos(z)$, and $\sin(z)$)**.** *The functions e^z, $\cos(z)$, and $\sin(z)$ have the following properties:*

1. *Exponent laws: For any $w, z \in \mathbb{C}$:*
   (a) $e^0 = 1$
   (b) $e^{w+z} = e^w e^z$
   (c) $e^{w-z} = e^w / e^z$
   (d) $e^{nz} = (e^z)^n$ *for all* $n \in \mathbb{N}$.
2. *Interrelationships:*
   (a) *Euler's Formula:* $e^{iz} = \cos(z) + i\sin(z)$ *for all* $z \in \mathbb{C}$.
   (b) $e^{a+ib} = e^a \operatorname{cis}(b) = e^a(\cos(b) + i\sin(b))$ *for all* $a, b \in \mathbb{R}$
   (c) $\cos(z) = \frac{1}{2}(e^{iz} + e^{-iz})$
   (d) $\sin(z) = \frac{1}{2i}(e^{iz} - e^{-iz})$
3. *Symmetries: For all $z \in \mathbb{C}$:*
   (a) $e^z = e^{z+2\pi i}$
   (b) $\cos(z) = \cos(z + 2\pi)$
   (c) $\sin(z) = \sin(z + 2\pi)$
   (d) $\cos(-z) = \cos(z)$
   (e) $\sin(-z) = -\sin(z)$
4. *Derivatives:*
   (a) $\frac{d}{dz} e^z = e^z$
   (b) $\frac{d}{dz} \cos(z) = -\sin(z)$

(c) $\frac{d}{dz}\sin(z) = \cos(z)$

Proof. The derivative formulas are straightforward consequences of Theorem 74, and the value of each function at 0 is trivial. The alert reader may worry that we are equivocating by using the symbols $\sin(x)$, $\cos(x)$ and e^x with two different meanings, especially in 2b where $\text{cis}(t)$ is defined in terms of prior knowledge sin and cos. Writing the "new" functions in bold, we calculate the derivatives of the functions:

$$\frac{\boldsymbol{e^t}}{e^t} \quad \text{and} \quad \frac{\mathbf{cos}(\boldsymbol{t}) + i\,\mathbf{sin}(\boldsymbol{t})}{\cos(t) + i\sin(t)}$$

Derivative laws give zero in each case, so each ratio is constant. Evaluating at $t = 0$ shows the constant is 1, so the functions agree. This shows that the newly defined complex exponential and trigonometric functions, restricted to $\mathbb{R}$, agree with their original versions.

Parts 1b and 2a are the most significant, and most of the other properties follow from these.

First we prove 1b:

$$e^{a+b} = \sum_{n=0}^{\infty} \frac{(a+b)^n}{n!} = \sum_{n=0}^{\infty}\sum_{k=0}^{n} \frac{a^k b^{n-k}}{k!(n-k)!} = \left(\sum_{n=0}^{\infty} \frac{a^n}{n!}\right)\left(\sum_{n=0}^{\infty} \frac{b^n}{n!}\right) = e^a e^b$$

Note the use of the binomial formula and Theorem 70.

To prove 2a, simply insert iz into the series for e^z and substitute $i^2 = -1$, $i^3 = -i$, $i^4 = 1$, etc. The even terms make $\cos(z)$, and the odd terms make $i\sin(z)$. Notice that this requires Theorem 70 for the validity of term-by-term addition of series.

The equation $e^0 = 1$ is obvious, and the other exponent laws follow easily from 1b. Equation 2b, sometimes used as a definition of e^{a+ib}, follows from 1b and 2a.

The symmetries 3d and 3e follow by direct substitution into the series. Combined with Euler's formula 2a for e^{iz} and e^{-iz} we obtain 2c and 2d. Next 3a follows from 2b and the periodicities of the *real* functions cos and sin. (I would prefer to give a self-contained argument, not dependent on prior knowledge of cos and sin, but this would require a rigorous definition of π and would take us far afield! See Spivak's Calculus for this argument.) Using 3a we deduce 3b from 2c, and 3c from 2d

$\square$

Exercise 34. Prove the angle sum formulas for arbitrary complex a and b

$$\cos(a+b) = \cos(a)\cos(b) - \sin(a)\sin(b)$$

$$\sin(a+b) = \sin(a)\cos(b) + \cos(a)\sin(b)$$

Note: This is a simple consequence of Euler's identity. Can you prove it directly from the power series definitions?

We're ready now to face the challenge of the logarithm. Of course the job of $\log(z)$ is to invert the function e^z, but 3a above reveals that e^z is far from injective. We must restrict its domain. In fact, for any $w \in \mathbb{C}$, we have $e^w = e^z$ for some z in the region $D = \{a + bi | b \in (-\pi, \pi]\}$, where e^z is injective (though we haven't proven it). Thus D seems a natural choice for the restricted domain of e^z.

Let us set $w = e^z$, where $z = a + bi \in D$ and try to solve for z: Using 2b we have $w = e^a \,\text{cis}(b)$. Taking moduli, $|w| = |e^a| \cdot 1$, and $a = \ln_{\mathbb{R}}(|w|)$, where of course $\ln_{\mathbb{R}}$ is the real natural logarithm. Returning to $w = e^a \,\text{cis}(b)$ and taking arguments, we have $\arg(w) = \arg \text{cis}(b) = b$ (since e^a is a positive real, and since $b \in (-\pi, \pi]$). Thus $z = a + bi = \ln_{\mathbb{R}}(|w|) + i\arg(w)$. This leads us to the following definition and proposition:

Definition 79. Let $z \in \mathbb{C}$. We define the **logarithm** $\log : \mathbb{C} \to \{a + bi | b \in (-\pi, \pi]\}$ of z by the formula:

$$\log(z) \equiv \ln_{\mathbb{R}}(|z|) + i\arg(z)$$

where arg is an argument function.

Notice that the ambiguity of arg is built into log. We can define log using any argument function, making appropriate changes below.

Proposition 80. *The function* $\log(z)$ *is undefined at* 0 *and discontinuous exactly where* $\arg(z)$ *is (i.e., the negative reals). Where continuous, the derivative is* $\log'(z) = 1/z$*. For all* z *we have* $e^{\log(z)} = z$*, and whenever* $\Im(z) \in (-\pi, \pi]$ *we have* $\log(e^z) = z$*.*

Proof. At $z = 0$, $\ln_{\mathbb{R}}(|z|)$ is undefined. Everywhere else, $\ln_{\mathbb{R}}(|z|)$ is continuous, so all discontinuities come from the term $i\arg(z)$. For any z, $e^{\log(z)} = e^{\ln_{\mathbb{R}}(|z|)+i\arg(z)} = e^{\ln_{\mathbb{R}}(|z|)} \operatorname{cis}(\arg(z)) = |z| \operatorname{cis}(\arg(z)) = z$. For $b \in (-\pi, \pi]$, we have $\log(e^{a+bi}) = \log(e^a \operatorname{cis}(b)) = \ln_{\mathbb{R}}(|e^a|) + i\arg(\operatorname{cis}(b)) = a + ib$ as desired. □

It's useful to look closely at the discontinuities of log. As z crosses the negative real line, its argument $\arg(z)$ jumps from nearly $-\pi$ to π, and the imaginary part of $\log(z)$ does the same. By choosing a different argument function, say $\arg_2$ with range $[0, 2\pi]$, we obtain a modified logarithm $\log_2(z)$, now discontinuous along the positive reals. On the upper half plane ($Im(z) > 0$) we have $\arg_2(z) = \arg(z)$, so $\log_2(z) = \log(z)$. But on the lower half plane ($Im(z) < 0$) $\arg_2(z) = \arg(z) + 2\pi$, so $\log_2(z) = \log(z) + 2\pi i$. The problem of the discontinuities of the logarithm is not to be blamed on the negative real axis, but is more essential: For any logarithm function, as z moves counterclockwise around the origin, $Im(\log(z))$ increases continuously with θ until, after one full rotation, $\log(z)$ has "gained" $2\pi i$ and must snap back to its starting value.

Notice that the function $1/z$ is analytic on $\mathbb{C} - \{0\}$, and $\log(z)$ is nearly but not quite an antiderivative on this region, because $\log(z)$ is discontinuous on $(-\infty, 0)$. This is the best we can do: Later we'll prove $1/z$ has no antiderivative on its entire domain.

Exercise 35. Let $f(x)$ be a differentiable function. For which x is it true that $\frac{d}{dx}\log(f(x)) = \frac{f'(x)}{f(x)}$? Answer with proof.

Exercise 36. The rules of logarithms are usually stated for positive real numbers: For $a, b \in (0, \infty)$:

1. $\ln_{\mathbb{R}}(1) = 0$
2. $\ln_{\mathbb{R}}(ab) = \ln_{\mathbb{R}}(a) + \ln_{\mathbb{R}}(b)$
3. $\ln_{\mathbb{R}}(a/b) = \ln_{\mathbb{R}}(a) - \ln_{\mathbb{R}}(b)$
4. $\ln_{\mathbb{R}}(a^n) = n\ln_{\mathbb{R}}(a)$

Everything hinges on the product law $\ln_{\mathbb{R}}(ab) = \ln_{\mathbb{R}}(a) + \ln_{\mathbb{R}}(b)$.

1. Give a complex counterexample (using the function log) to the product law for logarithms (Hint: It's the fault of the range of the argument function!)
2. As functions of z for fixed b, prove that $\log(zb)$ and $\log(z) + \log(b)$ agree at $z = 1$.
3. As functions of z for fixed b, compare the derivatives of $\log(zb)$ and $\log(z) + \log(b)$. On what set does each function have undefined derivative?
4. Assuming $\Re(b) > 0$, prove that $\log(zb)$ and $\log(z) + \log(b)$ agree for all z with $\Re(z) > 0$.

Exercise 37. Use the definition of e^z to estimate the value of e.

Exercise 38. The general exponential function a^z is often defined by the formula

$$a^z \equiv_{def} e^{z\log(a)}$$

(To remember the formula, just apply rules of logarithms to $a^z = e^{\log(a^z)}$.) Now let a be an arbitrary complex number. Prove that the exponential function a^z satisfies the sum law for exponential functions $a^{w+z} = a^w a^z$

3.4 Derivatives and Difference quotients

We now pause to analyze the difference quotient function. For this section, we let $f : X \to \mathbb{C}$ be a function defined and differentiable on $X \subseteq \mathbb{C}$ (or indeed $f : X \to \mathbb{R}$ with $X \subseteq \mathbb{R}$). We define $q(w, z)$ by the formula:

$$q(w,z) = \begin{cases} \frac{f(w)-f(z)}{w-z} & \text{if } w \neq z \\ f'(z) & \text{if } w = z \end{cases}$$

Notice that $q(w, z)$ is the difference quotient function, completed along the diagonal with its natural limit. Thus, $q(w, z)$ (above) is a continuous function of w for each fixed $z \in X$ and vice versa. However, this does not automatically mean that $q : X \times X \to \mathbb{C}$ is continuous. Furthermore note that q is symmetric in the sense that $q(w, z) = q(z, w)$.

Theorem 81. *Let X be an open convex subset of $\mathbb{C}$ or $\mathbb{R}$, and let $f : X \to \mathbb{C}$ be differentiable with $|f'| < \varepsilon$. Then $\forall w \neq z \in X$ $\left|\frac{f(w)-f(z)}{w-z}\right| < \varepsilon$.*

Proof. Notice that $q(w, z)$ (above) is continuous in w for fixed z and vice versa, and $|q(w, z)| = |q(w, z)|$.

We prove a sort of transitivity along straight lines: Assume $|q(w, m)| < \varepsilon$, $|q(m, z)| < \varepsilon$, and m is between w and z on a straight line segment in $\mathbb{C}$. Then

$$|q(w, z)| = \left|\frac{f(w) - f(z)}{w - z}\right| \leq \frac{|f(w) - f(m)| + |f(m) - f(z)|}{|w - m| + |m - z|} < \varepsilon$$

by the triangle inequality in the numerator and "triangle equality" for $|w - z|$. By basic properties of fractions, this is between $|q(w, m)|$ and $|q(m, z)|$, so it too is less than ε.

From here we need only topology. Let $w, z \in X$ and let I be the line segment connecting them. Since $q(w, z)$ is continuous in w, $S = \{w|q(w, z) < \varepsilon\}$ is open. Since $|f'(z)| < \varepsilon$, S contains z. Since I is compact, $I - S$ is empty or has a minimal (i.e. closest to z) element, $w_0 \neq z$. But then $S' = \{w : |q(w_0, w)| < \varepsilon\}$ is likewise open and contains w_0, so there is some m between z and w_0 for which $|q(z, m)| < \varepsilon$ and $|q(m, w_0)| < \varepsilon$, contradicting the transitivity property. This shows $I - S$ is empty, so $q(w, z) < \varepsilon$ as desired. □

Corollary 82. *If $f : U \to \mathbb{C}$ is differentiable with zero derivative then f is locally constant in the sense that $\forall z \in U$ $\exists \delta > 0$ f is constant on $B(z, \delta)$.*

Proof. Theorem 81 applies for all $\varepsilon > 0$. □

Corollary 83. *If $f : U \to \mathbb{C}$ is differentiable with zero derivative and U is connected, then f is constant.*

Proof. Choose $z_0 \in U$ and let $c = f(z_0)$. Then $U = f^{-1}(c) \cup f^{-1}(\mathbb{C} - \{c\})$. The first set is open by Corollary 82, and the second is open by continuity. By connectedness, the latter set must be empty, so f is constant. □

Corollary 84. *Let $f : U \to \mathbb{C}$, with U open and connected. If f has an antiderivative, then it is unique up to constant.*

Proof. If F_1 and F_2 are antiderivatives of f, then $F_1 - F_2$ has zero derivative, so is constant. □

From Theorem 81 we also deduce:

Corollary 85. *Let X be an open convex subset of $\mathbb{C}$ or $\mathbb{R}$, and let $f : X \to \mathbb{C}$ be differentiable with $f' \in B(m, \varepsilon)$, with $m \in \mathbb{C}$. Then $\forall w \neq z \in X$ $\frac{f(w)-f(z)}{w-z} \in B(m, \varepsilon)$.*

Proof. Apply Theorem 81 to the function $f(x) - mx$. □

If furthermore f' is a continuous function, then so is the function q above:

Theorem 86. *Let $f : X \to \mathbb{C}$, where X is an open subset of $\mathbb{C}$ or $\mathbb{R}$. If f' is defined and continuous on X, then the function*

$$q(w, z) = \begin{cases} \frac{f(w)-f(z)}{w-z} & \text{if } w \neq z \\ f'(z) & \text{if } w = z \end{cases}$$

is a continuous function from $X \times X$ to $\mathbb{C}$.

Proof. We prove continuity at an arbitrary point (z_0, w_0). If $w_0 \neq z_0$ then in a neighborhood of (z_0, w_0) we have $q(w, z) = \frac{f(w)-f(z)}{w-z}$, which is continuous since both $f(z)$ and $\frac{1}{w-z}$ are. If $w_0 = z_0$, let $\varepsilon > 0$. Let $m = f'(z_0) = q(w_0, z_0)$. Choose $\delta > 0$ so that $|w - z_0| < \delta \to f'(w) \in B(m, \varepsilon)$ by continuity of f'. Since $B(z_0, \delta)$ is convex, Corollary 85 applies, so for all $(w, z) \in B(z_0, \delta) \times B(z_0, \delta)$ we have $q(w, z) \in B(m, \varepsilon)$. Therefore q is continuous. □

Any continuous function will be uniformly continuous on compact sets. If X is furthermore compact we get a very useful tool which allows us to uniformly estimate $\frac{f(z)-f(w)}{z-w}$ using the derivative $f'(z_0)$ at a different, but nearby point z_0:

Theorem 87. *Let $f : K \to \mathbb{C}$, where K is a compact subset of $\mathbb{C}$ or $\mathbb{R}$. If f' is defined and continuous on some open superset U of K, then*

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall z_0 \in K \ \forall w \neq z \in B(z_0, \delta) \cap K \to \left| \frac{f(z) - f(w)}{z - w} - f'(z_0) \right| < \varepsilon$$

Proof. We apply the uniform continuity of $q(w, z) = \left\{ \begin{array}{ll} \frac{f(w)-f(z)}{w-z} & \text{if } w \neq z \\ f'(z) & \text{if } w = z \end{array} \right\}$ on the domain $K \times K$ to learn:

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall (w_0, z_0) \in K \times K \ \forall (w, z) \in K \times K \ d((w_0, z_0), (w, z)) < \delta \to |q(w, z) - q(w_0, z_0)| < \varepsilon$$

Here $d((w_0, z_0), (w, z)) = \max(|w_0 - w|, |z_0 - z|)$. The desired formula is obtained by taking the special case $w_0 = z_0$ and noting that when $w_0 = z_0$ the condition $w \neq z \in B(z_0, \delta)$ implies the condition $d((w_0, z_0), (w, z)) < \delta$. □

Exercise 39. A standard example of a function $f : \mathbb{R} \to \mathbb{R}$ with discontinuous derivative is:

$$f(x) = \begin{cases} x^2 \sin(1/x) & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

1. Prove that f is differentiable (everywhere) but f' is discontinuous.
2. Justify: $q(w, z) : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ is continuous in w for fixed z and continuous in z for fixed w.
3. Prove: $q(w, z)$ is discontinuous at $(0, 0)$.

Chapter 4

Integration

4.1 Motivation

4.1.1 The Riemann Integral on $\mathbb{R}$

We wish to define a theory of complex integration which directly generalizes the real integral $\int_a^b f(x)dx$. It isn't obvious how the theory should change to accomodate the complex numbers, and the basic definitions are a bit mysterious to the beginner. So we begin with a conversation designed to review real integration and motivate new definitions.

The real (definite) integral $\int_a^b f(x)dx$ is meant of course to measure the signed area under the curve $f(x)$ between $x = a$ and $x = b$. It requires remarkably sophisticated geometry to directly define "area" in a way that applies to the great variety of graphs of everyday functions, so instead we approximate the area using rectangles, and define the integral to be the limit of rectangular approximations as rectangles narrow toward zero width:

Definition 88 (Real integral – incomplete definition)**.** Let $f : \mathbb{R} \to \mathbb{R}$ be a function. The **definite integral of** f **from** a **to** b is defined to be

$$\int_a^b f(x)dx \equiv_{\text{def}} \lim \sum_{i=1}^{m} f(\tau_i)(t_i - t_{i-1})$$

We have divided the interval $[a, b]$ into m parts $a = t_0 < t_1 < \ldots < t_m = b$ and placed sample points τ_i into these parts. Geometrically, $t_i - t_{i-1}$ measures the width of the i'th rectangle and $f(\tau_i)$ its height. Often it is presumed that the rectangle boundaries t_i are evenly spaced, so $(t_i - t_{i-1})$ is a constant width, which then must be $\frac{b-a}{m}$. Note that there is an arbitrariness in the argument of f: We could take $f(t_i)$ ("right hand rule"), $f(t_{i-1})$ ("left hand rule"), $f(\frac{t_i+t_{i-1}}{2})$ ("midpoint rule") or more generally any $f(\tau_i)$ with $t_{i-1} \le \tau_i \le t_i$.

What sort of limit is meant in the definition of the real integral? Superficially, this is easy: To make the rectangles ever narrower, we mean the limit as $m \to \infty$ with equal width rectangles.

What could be wrong with that? For practical purposes, some integrals are much easier to calculate using strategically unequal widths. For example, to calculate $\int_1^b \frac{1}{x}dx$, a subdivision in which $t_i = b^{i/m}$ gives each rectangle equal area. On a more theoretical note, assuming that a and b are rational, setting $\tau_i = a + i\frac{b-a}{m}$ commits us to only evaluating the function $f : [a, b] \to \mathbb{R}$ at rational inputs. It seems improper that "most" of the values $f(x)$ of f, that is the values at irrational x, should be disregarded in the definition of $\int_a^b f(x)dx$.

These are reasons to consider more general arrangements of points $a = t_0 < t_1 < \ldots < t_m = b$, and more general sample points τ_i: We call this data a *partition.*

Definition 89. A **partition** P of a closed interval $[a, b]$ is a finite set of points $a < t_0 < t_1 < \ldots < t_m = b$, called the **endpoints** of P. A **tagged partition** is a partition as above, together with a list of **tags** $\tau_1, \ldots \tau_m$, so that $t_{i-1} \le \tau_i \le t_i$. We say τ_i **tags** the interval $[t_{i-1}, t_i]$ for P.

The intervals $[t_{i-1}, t_i]$ divide $[a, b]$ into regions on which we'll draw rectangles. The tags τ_i are sample points: We evaluate f at τ_i to assess the height of the rectangle based on $[t_{i-1}, t_i]$. Incorporating this useful generalization into our definition of the integral gives us an immediate problem interpreting the limit, because we can no longer regard t_i and τ_i as functions of m. To describe the limit "as the rectangles get narrower," we measure the **norm** of a partition to be the maximum width of its intervals.

Definition 90. Let P be a partition of $[a, b]$ with endpoints $a < t_0 < t_1 < \ldots < t_m = b$. The **norm** of P is the maximum width:

$$||P|| \equiv_{\text{def}} \max\{t_i - t_{i-1} : 1 \leq i \leq m\}$$

Definition 91 (Real integral – partitions definition)**.** Let $f : \mathbb{R} \to \mathbb{R}$ be a function. The **definite integral of f from a to b** is defined to be

$$\int_a^b f(x)dx \equiv_{\text{def}} \lim_{\|P\| \to 0} \sum_{i=1}^{m} f(\tau_i)(t_i - t_{i-1})$$

If the limit exists, we say that f is **Riemann integrable.**

...but this is a new kind of limit entirely, demanding a new definition.

Definition 92. Let ϕ be a function from tagged partitions of $[a, b]$ to $\mathbb{R}$ (or $\mathbb{C}$). We define:

$$\lim_{\|P\| \to 0} \phi(P) \equiv_{\text{def}} \imath L \ \forall \varepsilon > 0 \ \exists \delta > 0 \ \forall \text{ tagged partitions } P \text{ on } [a, b] \ \big(\|P\| < \delta \to |\phi(P) - L| < \varepsilon\big)$$

Note that the codomain of ϕ can be changed to any metric space by replacing $|\phi(P) - L|$ with $d(\phi(P), L)$.

4.1.2 Complex integrals

Let's consider an integral $\int_a^b f(z)dz$, where $f : \mathbb{C} \to \mathbb{C}$ is a complex function and $a, b \in \mathbb{C}$ are complex numbers. To mimic the work above we would introduce waypoints $a = z_0, z_1, \ldots, z_n = b$, and sample points τ_i. Our waypoints z_n can no longer be ordered because $\mathbb{C}$ has no order structure. We can no longer insist that $z_{i-1} \leq \tau_i \leq z_i$ for the same reason, but let us ignore this problem for now. We could, however, generalize the idea of partition norm to be $\max_i |z_i - z_{i-1}|$. But after all this we would still discover an intractable difficulty: For many very reasonable functions such as $f(z) = 1/z$, the limit of sums $\sum_{i=1}^m f(\tau_i)(z_i - z_{i-1})$ depends on how the points z_i meander in two dimensions from a to b, and straight line paths are not always possible. We must therefore separate the question of subdivision from the question of how these points wander around in the complex plane. We define an object called a *path* (written γ) which exactly prescribes these wanderings. Then the integral depends not only on endpoints a and b but on the path γ.

Definition 93. A **path** is a continuous function $\gamma : [a, b] \to \mathbb{C}$ from a closed interval to $\mathbb{C}$. A path is **closed** if $\gamma(a) = \gamma(b)$.

This frees us to return to ordinary, real tagged partitions: If P is a partition of the domain $[a, b]$ of a path γ, then $\{\gamma(t_i)\}$ will be the needed finite sequence of waypoints in $\mathbb{C}$, and $\gamma(\tau_i)$ will be the required sample points. We now define the path integral of a complex function.

Definition 94 (Path integral)**.** Let $f : \mathbb{C} \to \mathbb{C}$ be a function, and $\gamma : [a, b] \to \mathbb{C}$ a path. The **path integral of f along γ** is defined to be

$$\int_\gamma f(z)dz \equiv_{\text{def}} \lim_{\|P\| \to 0} \sum_{i=1}^{m} f(\gamma(\tau_i))(\gamma(t_i) - \gamma(t_{i-1}))$$

where the limit is over tagged partitions P of $[a, b]$.

Exercise 40. Prove directly from the previous definition that when f is the constant function $f(z) = k$, then $\int_\gamma f(z)dz = k(\gamma(b) - \gamma(a))$. (You may take for granted that the limit in this new sense of a constant is that constant.)

Exercise 41. Let $\chi_{\mathbb{Q}} : \mathbb{R} \to \mathbb{R}$ by $\chi_{\mathbb{Q}}(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{if } x \notin \mathbb{Q} \end{cases}$

1. Show that $\chi_{\mathbb{Q}}$ is integrable (i.e., the integral limit exists) on $[0, 1]$, and calculate its integral, using the definition
$$\int_a^b f(x)dx \equiv_{\text{def}} \lim_{m\to\infty} \sum_{i=1}^{m} f(\tau_i)(t_i - t_{i-1})$$
where $t_i = \frac{i(b-a)}{m}$ and $\tau_i = t_i$.
2. Show that $\chi_{\mathbb{Q}}$ is not integrable on $[0, 1]$ using the (more sophisticated) partition-based definition
$$\int_a^b f(x)dx \equiv_{\text{def}} \lim_{\|P\|\to 0} \sum_{i=1}^{m} f(\tau_i)(t_i - t_{i-1})$$
where t_i and τ_i come from the partition P.

Exercise 42. Prove that if $f : [a, b] \to \mathbb{R}$ is unbounded, then $\int_a^b f(x)dx$ does not exist. (This should bother you when you remember your improper integrals!)

Since the function $f : \mathbb{C} \to \mathbb{C}$ enters into the right hand side only to form the composition $f\circ\gamma : [a, b] \to \mathbb{C}$, we have an obvious generalization in which $f \circ \gamma$ is replaced with an arbitrary function on $[a, b]$:

Definition 95 (Riemann-Stieltjes integral)**.** Let $f : [a, b] \to \mathbb{C}$ be a function, and $\gamma : [a, b] \to \mathbb{C}$ a function (usually a path). The **Riemann-Stieltjes integral of** f **with respect to** γ is defined to be
$$\int_a^b f d\gamma = \int_a^b f(t)d\gamma(t) \equiv_{\text{def}} \lim_{\|P\|\to 0} \sum_{i=1}^{m} f(\tau_i)(\gamma(t_i) - \gamma(t_{i-1}))$$
where the limit is over tagged partitions P of $[a, b]$.

This is a rich and interesting generalization of the Riemann integral, even in case $\gamma : [a, b] \to \mathbb{R}$ is a nonincreasing (not necessarily continuous) real valued function. In that case, the Riemann-Stieltjes integral acts like a weighted integral of f, in which portions of $[a, b]$ are given more weight in proportion to γ' (if indeed γ is differentiable). Jump discontinuities of γ give nonzero weight, in the sense of integration, to single points of $[a, b]$.

Note 96. *If $\gamma(t) = t$ in the Riemann-Stieltjes integral, then*
$$\int_a^b f(t)dt = \int_a^b f d\gamma$$
If γ is a path, then
$$\int_\gamma f(z)dz = \int_a^b (f \circ \gamma)d\gamma$$
By studying the Riemann-Stieltjes integral, we are developing a common foundation of both complex path integrals and ordinary Riemann integration.

Exercise 43. Let $\gamma : [0, 1] \to \mathbb{C}$ be a straight line path from 1 to i. Give an explicit equation for γ.

1. Calculate $\int_\gamma |z|^2 dz$, by converting the integrand to a real-valued function of a real variable $t \in [0, 1]$ and using real methods.
2. Calculate $\int_\gamma z^2 dz$. In this case the integrand is complex. Expand and separate it into real and imaginary part, and reduce the problem to two real integrals.

4.2 Limit properties of functions on partitions

We wish to prove a criterion for existence of the Riemann-Stieltjes integral. The integral is a limit as $\|P\| \to 0$, and it's convenient to develop the properties of such limits:

Definition 97. Let P and Q be partitions of $[a, b]$ (tagged or not). We say that Q **refines** P, and write $P \preccurlyeq Q$, if the endpoints $\{t_i\}$ of P form a subset of the endpoints $\{s_j\}$ of Q. (Note: For $P \preccurlyeq Q$, we require no relationship between tags. Some authors do!)

Note that when $P \preccurlyeq Q$, Q is greater in terms of *more division points*, but P usually has wider intervals.

Proposition 98. *Let $\phi(P)$ be a function from tagged partitions P of $[a, b]$ to $\mathbb{R}$ with ϕ increasing in the sense that if $P \preccurlyeq Q$ then $\phi(P) \leq \phi(Q)$. If $\lim\limits_{\|P\| \to 0} \phi(P)$ exists then it equals $\sup_P(\phi(P))$*

Proof. Let $L = \lim_{\|P\| \to 0} \phi(P)$. Let Q be any partition. Let $\varepsilon > 0$. Choose $\delta > 0$ for ε in the definition of the limit. Choose some refinement $Q \preccurlyeq R$ with $\|R\| < \delta$. Then $\phi(Q) \leq \phi(R) < L + \epsilon$. Since ε was arbitary, $\phi(Q) \leq L$ for all Q and $\sup_P(\phi(P)) \leq L$. Since $\phi(P)$ can be made arbitarily close to L, $\sup_P(\phi(P)) \geq L$. □

Definition 99. Let $\phi(P)$ be a function from tagged partitions P of $[a, b]$ to a metric space X (usually $\mathbb{C}$). We say that ϕ is **Cauchy** if:

$$\forall \varepsilon > 0 \ \exists \delta > 0 \ \forall \text{ tagged partitions } P, Q \text{ with } \|P\| < \delta \text{ and } P \preccurlyeq Q \text{ we have } d(\phi(P), \phi(Q)) < \varepsilon$$

Theorem 100. *Let $\phi(P)$ be a function from tagged partitions P of $[a, b]$ to a complete metric space X. Then ϕ is Cauchy if and only if $\lim\limits_{\|P\| \to 0} \phi(P)$ exists.*

Proof. If $\lim_{\|P\| \to 0} \phi(P) = L$, let $\varepsilon > 0$ and choose $\delta > 0$ so that $\|P\| < \delta$ implies $d(\phi(P), L) < \varepsilon/2$. Now if $P \preccurlyeq Q$ then $\|Q\| < \delta$ and $d(\phi(P), \phi(Q)) \leq d(\phi(P), L) + d(\phi(Q), L) < \varepsilon/2 + \varepsilon/2 = \varepsilon$.

Conversely let ϕ be Cauchy. Construct a sequence of partitions $P_0 \preccurlyeq P_1 \preccurlyeq P_2 \preccurlyeq \ldots$ with $\lim_n \|P_n\| = 0$. For example, subdivide $[a, b]$ into 2^n equal intervals, tagged at midpoints. Then the sequence $(\phi(P_n))$ is Cauchy in the usual sense, so has a limit L. To see that $\lim_{\|P\| \to 0} \phi(P) = L$, let $\varepsilon > 0$. Choose δ for $\varepsilon/3$ in the Cauchy criterion for ϕ, and choose N so that for $n > N$ we have $\|P_n\| < \delta$ and $d(\phi(P_n), L) < \varepsilon/3$. To show partition convergence, let $\|Q\| < \delta$. Choose some $n > N$ and a joint refinement filter R so that $P_n \preccurlyeq R$ and $Q \preccurlyeq R$. Then the Cauchy criterion applies to both $d(\phi(Q), \phi(R))$ and $d(\phi(R), \phi(P_n))$, so:

$$d(\phi(Q), L) \leq d(\phi(Q), \phi(R)) + d(\phi(R), \phi(P_n)) + d(\phi(P_n), L) \leq \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon$$

□

4.3 Existence of Integrals

Now we have some headache: There are some *very* strange continuous functions $\gamma : [a, b] \to \mathbb{C}$: Maps whose images have infinite length, whose images are fractals (e.g., the Koch curve), or even two-dimensional images (e.g., the Peano space-filling curve). None of these sets are convenient domains of integration, so we must identify a subclass of "reasonable" curves, in particular those of finite length in a sense we call "bounded variation."

Definition 101. Let $\gamma : [a, b] \to \mathbb{C}$ be a function. The **variation** of γ on a partition P with endpoints $\{a = t_0 < \ldots < t_m = b\}$ is the total distance along the m straight line segments connecting $\gamma(t_0), \ldots, \gamma(t_m)$:

$$v(\gamma; P) \equiv_{\text{def}} \sum_{i=1}^{m} |\gamma(t_i) - \gamma(t_{i-1})|$$

The **total variation** of γ is the supremum over all partitions P of $[a, b]$:

$$v(\gamma) \equiv_{\text{def}} \sup_P v(\gamma; P)$$

If $V(\gamma)$ is finite, we say γ has **bounded variation**.

We will show that if f is continuous and γ has bounded variation, then the integral $\int_a^b f(t)d\gamma(t)$ exists. Integrals over paths of unbounded variation sometimes do not exist.

Proposition 102. *$v(\gamma, P)$ is an increasing function of P in the sense that if $P \preccurlyeq Q$ then $v(\gamma, P) \leq v(\gamma, Q)$*

Proof. Let P have endpoints $\{t_i\}_{i=0}^m$ and let Q have endpoints $\{s_j\}_{j=0}^n$. Apply the triangle inequality to $|\gamma(t_i) - \gamma(t_{i-1})| = \left|\gamma(s_{j_i}) - \gamma(s_{j_i-1}) + \gamma(s_{j_i-1}) \ldots - \gamma(s_{j_{i-1}})\right|$. $\square$

Corollary 103. *Let $\gamma : [a, b] \to \mathbb{C}$ be a function. If $\lim_{\|P\|\to 0} v(\gamma, P)$ exists then it equals $v(\gamma)$.*

We are now ready to prove the existence of the integral:

Theorem 104. *Let $\gamma : [a, b] \to \mathbb{C}$ have bounded variation, and let $f : [a, b] \to \mathbb{C}$ be continuous. Then $\int_a^b f(t)d\gamma(t)$ exists.*

Proof. The integral is a limit as $\|P\| \to 0$, and of course we plan to use the Cauchy criterion. Let $\varepsilon > 0$. Since f is continuous on a compact set, it is uniformly continuous and we may choose $\delta > 0$ so that whenever $|x - y| < \delta$ then $|f(x) - f(y)| < \frac{\varepsilon}{v(\gamma)}$. Here $v(\gamma)$ is the variation of γ, presumed finite. Let $P \preccurlyeq Q$ be partitions of $[a, b]$, where P has endpoints t_i and tags τ_i and Q has endpoints s_j and tags σ_j, and assume $\|P\| < \delta$. Since Q refines P, each t_i is equal to some s_j. Say $t_i = s_{k_i}$. The difference $\left|\sum_{i=1}^m f(\tau_i)(\gamma(t_i) - \gamma(t_{i-1})) - \sum_{j=1}^{m'} f(\sigma_j)(\gamma(s_j) - \gamma(s_{j-1}))\right|$ can be reorganized by expanding each single term $f(\tau_i)(\gamma(t_i) - \gamma(t_{i-1}))$ to a subsum $\sum_{j=k_{i-1}+1}^{k_i} f(\tau_i)(\gamma(s_j) - \gamma(s_{j-1}))$ We let $\tau_j' = \tau_i$ for all j in this range, so that the term $f(\tau_i)$ can be expressed as a function of j. Thus:

$$\begin{aligned}
\sum_{i=1}^m f(\tau_i)(\gamma(t_i) - \gamma(t_{i-1})) &= \sum_{i=1}^m \sum_{j=k_{i-1}+1}^{k_i} f(\tau_j')(\gamma(s_j) - \gamma(s_{j-1})) \\
&= \sum_{j=1}^{m'} f(\tau_j')(\gamma(s_j) - \gamma(s_{j-1}))
\end{aligned}$$

The watchful reader will notice that the sum above is not a normal partition estimate to an integral because τ_j' is not necessarily in the range $[s_{j-1}, s_j]$. Nevertheless we may combine the two sums:

$$\begin{aligned}
&\left|\sum_{i=1}^m f(\tau_i)(\gamma(t_i) - \gamma(t_{i-1})) - \sum_{j=1}^{m'} f(\sigma_j)(\gamma(s_j) - \gamma(s_{j-1}))\right| = \\
&\quad \left|\sum_{j=1}^{m'} f(\tau_j')(\gamma(s_j) - \gamma(s_{j-1})) - f(\sigma_j)(\gamma(s_j) - \gamma(s_{j-1}))\right| = \left|\sum_{j=1}^{m'} (f(\tau_j') - f(\sigma_j))(\gamma(s_j) - \gamma(s_{j-1}))\right| \\
&\qquad \leq \sum_{j=1}^{m'} \left|f(\tau_j') - f(\sigma_j)\right| \left|\gamma(s_j) - \gamma(s_{j-1})\right| \\
&\qquad < \sum_{j=1}^{m'} \frac{\varepsilon}{v(\gamma)} |\gamma(s_j) - \gamma(s_{j-1})| \leq \varepsilon
\end{aligned}$$

We conclude that the function $\sum_{i=1}^m f(\tau_i)(\gamma(t_i) - \gamma(t_{i-1}))$ is a Cauchy function of the partition P, so it converges. That is, the integral $\int_a^b f(t)d\gamma(t)$ exists. $\square$

Corollary 105. *Let $f : [a, b] \to \mathbb{C}$ be continuous. Then $\int_a^b f(t)dt$ exists.*

Proof. $\gamma(t) = t$ has bounded variation $v(\gamma) = b - a$. $\square$

Exercise 44. A function $h : X \to \mathbb{C}$ (with $X \subseteq \mathbb{C}$) is called **Lipschitz** if it has bounded difference quotients: That is, there is a constant M so that for all $x \neq y \in X$ $\left|\frac{h(y)-h(x)}{y-x}\right| < M$. Prove that if $f : [a,b] \to \mathbb{C}$ and $\gamma : [a,b] \to \mathbb{C}$ are both Lipschitz, then $\int_a^b f d\gamma$ exists. (Hint: Not from the definitions!)

The next theorem is easy but extremely useful:

Theorem 106 (ML inequality)**.** *Let $f : [a,b] \to \mathbb{C}$ be continuous. Let $|f| \leq M$ on $[a,b]$, and let $L = v(\gamma)$. Then $\left|\int_a^b f(t) d\gamma\right| \leq ML$.*

Proof. Notice the integral exists by Theorem 104 For any partition P we have $\left|\sum_{i=1}^{m} f(\tau_i)(\gamma(t_i) - \gamma(t_{i-1}))\right| \leq \sum_{i=1}^{m} M|\gamma(t_i) - \gamma(t_{i-1})| \leq Mv(\gamma) \leq ML$. □

Corollary 107. *Let $g(z,w)$ be a continuous function of two variables on $U \times im(\gamma)$, where γ is continuous and bounded variation. Let $z_0 \in U$. Then*

$$\lim_{z \to z_0} \int_\gamma g(z,w)dw = \int_\gamma \lim_{z \to z_0} g(z,w)dw = \int_\gamma g(z_0,w)dw$$

Proof. Choose a ball $\overline{B(z_0,\rho)} \subseteq U$ and choose $\delta < \rho$ for $\frac{\varepsilon}{v(\gamma)}$ in the uniform continuity of g on the compact set $\overline{B(z_0,\rho)} \times im(\gamma)$. Then when $|z - z_0| < \delta$ the ML inequality gives:

$$\left|\int_\gamma g(z,w)dw - \int_\gamma g(z_0,w)dw\right| = \left|\int_\gamma g(z,w) - g(z_0,w)dw\right| \leq \frac{\varepsilon}{v(\gamma)} v(\gamma) = \varepsilon$$

□

Theorem 108 (Additivity properties of Riemann-Stieltjes integral)**.** *Let $f, g : [a,b] \to \mathbb{C}$ be continuous, let $\gamma, \sigma : [a,b] \to \mathbb{C}$ be bounded variation, and let $c_1, c_2 \in \mathbb{C}$. Then*

1. $\int_a^b (c_1 f + c_2 g) d\gamma = c_1 \int_a^b f d\gamma + c_2 \int_a^b g d\gamma$
2. $\int_a^b f d(c_1\gamma + c_2\sigma) = c_1 \int_a^b f d\gamma + c_2 \int_a^b f d\sigma$
3. $\int_a^{b'} f d\gamma = \int_a^b f d\gamma + \int_b^{b'} f d\gamma$

Proof. (Parts 1 and 2 are routine. Part 3 is subtle because an arbitrary partition P of the interval $[a,b']$ may not have b as an endpoint.) □

Example 109. (Differentiability does not guarantee bounded variation.) Let

$$\gamma(t) = \begin{cases} t^2 \cos(\frac{1}{t^2}) & \text{if } t \neq 0 \\ 0 & \text{if } t = 0 \end{cases}$$

Then γ is everywhere differentiable, γ' is discontinuous at 0, and γ has unbounded variation on any neighborhood of 0.

Proof. By the usual differentiation rules, if $t \neq 0$ then $\gamma'(t) = 2t\cos(\frac{1}{t^2}) + \frac{2}{t}\sin(\frac{1}{t^2})$, which is unbounded in a neighborhood of 0. Furthermore $\gamma'(0) = 0$ because $-t^2 \leq \gamma(t) \leq t^2$. Thus γ' is everywhere defined but discontinuous at 0.

Next consider $v(\gamma, P)$ where P is a partition of $[a,b]$, $a < 0 < b$ having endpoints rigged to make the cosine ± 1:

$$a < 0 < \frac{1}{\sqrt{\pi M}} < \frac{1}{\sqrt{\pi(M-1)}} < \ldots < \frac{1}{\sqrt{\pi N}} < b$$

(Note if $b > 0$ we may choose $0 << N < M$ so that $\frac{1}{\sqrt{\pi N}} < b$. For simplicity take both M and N even.) The corresponding values of γ are:

$$\gamma(a), 0, \frac{1}{\pi M}, \frac{-1}{\pi(M-1)}, \frac{1}{\pi(M-2)}, \dots, \frac{-1}{\pi(N+1)}, \frac{1}{\pi N}, \gamma(b)$$

Because the harmonic series diverges, $v(\gamma, P)$ can be made arbitrarily large by choosing M sufficient large. $\square$

The previous example indicates that we need a stronger condition than differentiability to guarantee bounded variation:

Definition 110. Let $\gamma : [a, b] \to \mathbb{C}$ be a function. We call γ **smooth** if γ is differentiable and γ' is continuous.

Proposition 111. *If $\gamma : [a, b] \to \mathbb{C}$ is smooth, then γ has bounded variation and moreover $v(\gamma) = \int_a^b |\gamma'(t)| dt$.*

Proof. We claim that the integral exists and for any $\varepsilon > 0$ there is a $\delta > 0$ so that if $\|P\| < \delta$ then both

$$\left|\sum_{i=1}^{m} |\gamma(t_i) - \gamma(t_{i-1})| - \sum_{i=1}^{m} |\gamma'(\tau_i)|(t_i - t_{i-1})\right| < \varepsilon/2 \qquad \text{and} \qquad \left|\sum_{i=1}^{m} |\gamma'(\tau_i)|(t_i - t_{i-1}) - \int_a^b |\gamma'(t)| dt\right| < \varepsilon/2$$

Together these will establish that $\lim_{\|P\| \to 0} \sum_{i=1}^{m} |\gamma(t_i) - \gamma(t_{i-1})| = \int_a^b |\gamma'(t)| dt$, after which proposition 103 applies.

By Corollary 105, we know that $\int_a^b |\gamma'(t)| dt$ exists. The righthand inequality then follows by the definition of the integral.

For the lefthand inequality, let P be a partition with endpoints t_i and tags τ_i.

$$\begin{aligned}\left|\sum_{i=1}^{m} |\gamma(t_i) - \gamma(t_{i-1})| - \sum_{i=1}^{m} |\gamma'(\tau_i)|(t_i - t_{i-1})\right| &= \left|\sum_{i=1}^{m} \left(\left|\frac{\gamma(t_i) - \gamma(t_{i-1})}{t_i - t_{i-1}}\right| - |\gamma'(\tau_i)|\right)(t_i - t_{i-1})\right| \\ &\leq \sum_{i=1}^{m} \left|\frac{\gamma(t_i) - \gamma(t_{i-1})}{t_i - t_{i-1}} - \gamma'(\tau_i)\right|(t_i - t_{i-1})\end{aligned}$$

Now let $\varepsilon > 0$. Using Theorem 87, we can choose δ so that for any partition P with $\|P\| < \delta$ we have $\left|\frac{\gamma(t_i) - \gamma(t_{i-1})}{t_i - t_{i-1}} - \gamma'(\tau_i)\right| < \frac{\varepsilon}{2(b-a)}$. The difference above is then less than $\varepsilon/2$.

$\square$

Exercise 45. Compute the total variation $v(\gamma)$ for each of the paths below.

1. $\gamma(t) = Re^{it}$ on $[0, 2\pi]$, where $R > 0$.
2. $\gamma(t) = mt + b$ on $[a, b]$, where $m, b \in \mathbb{C}$
3. $\gamma(t) = \lfloor t/2 \rfloor$ on $[0, 5]$.

If γ is smooth, then $d\gamma$ may be replaced with $\gamma'(t)dt$ (as in traditional u-substitution), converting the Riemann-Stieltjes integral to a traditional Riemann integral:

Theorem 112. *Let $\gamma : [a, b] \to \mathbb{C}$ be smooth, and let $f : [a, b] \to \mathbb{C}$ be continuous. Then*

$$\int_a^b f d\gamma = \int_a^b f(t)\gamma'(t) dt$$

Proof. Since γ is smooth, it has bounded variation by Proposition 111. Therefore both integrals exist by Theorem 104 and its Corollary 105. Let P be a partition with endpoints t_i and tags τ_i and consider the difference between the sums:

$$\sum_{i=1}^{m} f(\tau_i)(\gamma(t_i) - \gamma(t_{i-1})) - \sum_{i=1}^{m} f(\tau_i)\gamma'(\tau_i)(t_i - t_{i-1})$$

Regrouping gives:

$$= \sum_{i=1}^{m} f(\tau_i) \left(\frac{\gamma(t_i) - \gamma(t_{i-1})}{t_i - t_{i-1}} - \gamma'(\tau_i) \right) (t_i - t_{i-1})$$

Let $\varepsilon > 0$. Since f is continuous on a compact set we may bound it: $|f| < B$. Again by Theorem 87, we choose δ so that whenever $\|P\| < \delta$ we have $\left| \frac{\gamma(t_i) - \gamma(t_{i-1})}{t_i - t_{i-1}} - \gamma'(\tau_i) \right| < \frac{\varepsilon}{B(b-a)}$, which makes the modulus of the sum less than ε as required. $\square$

Next we prove a substitution theorem:

Theorem 113. *Let $\gamma : [a,b] \to \mathbb{C}$ and $f : [a,b] \to \mathbb{C}$ be continuous. Let $\phi : [c,d] \to [a,b]$ be continuous, nondecreasing, and onto. Then*

$$\int_c^d (f \circ \phi)(t) d(\gamma \circ \phi)(t) = \int_a^b f(t) d\gamma(t)$$

provided the right hand side exists.

Proof. Let $\varepsilon > 0$. Since $I = \int_a^b f(t)d\gamma(t)$ exists, we may choose $\rho > 0$ according to the definition of the integral, so that for any partition P of $[a,b]$ with $\|P\| < \rho$ we have $|I - \sum_{i=1}^{m} f(\tau_i)(\gamma(t)_i - \gamma(t_{i-1}))| < \varepsilon$. Next choose δ for ρ in the definition of absolute continuity of ϕ, so that $|s - s'| < \delta \to |\phi(s) - \phi(s')| < \rho$. Now if Q partitions $[c,d]$ with endpoints s_i, tags σ_i, and $\|Q\| < \delta$, consider the integral estimate using Q:

$$\sum_{i=1}^{m} f(\phi(\sigma_i))(\gamma(\phi(s_i)) - \gamma(\phi(s_{i-1})))$$

This *is* an integral estimate using a partition "$\phi(Q)$" of $[a,b]$ having endpoints $\phi(s_i)$ and tags $\phi(\sigma_i)$, and $\|\phi(Q)\| < \rho$ by construction, so it is in $B(I, \varepsilon)$ as required. (If ϕ is not injective, we may need to drop multiple copies of $\phi(s_i)$ to get a partition, but this does not affect the sum.) $\square$

Recall that our interest in Riemann-Stieltjes integrals originated with path integrals $\int_\gamma f \equiv_{def} \int_a^b f(\gamma(t))d\gamma(t)$. Let us collect together our knowledge about this special case:

Theorem 114. *Let $\gamma : [a,b] \to \mathbb{C}$ be a function and let f be defined on $\gamma([a,b])$.*

1. *If f is continuous and γ is smooth then $\int_\gamma f$ exists.*
2. *If $\phi : [c,d] \to [a,b]$ is continuous, nondecreasing, and onto, then $\int_{\gamma \circ \phi} f = \int_\gamma f$ provided the right hand side exists.*
3. *If f is continuous and γ is smooth then the path integral can be computed by an ordinary Riemann integral: $\int_\gamma f = \int_a^b f(\gamma(t))\gamma'(t)dt$.*

Proof. The first is 104 and 111. The second is 113 with $f \circ \gamma$ in place of f. The third is 112. $\square$

Finally we prove a Fundamental Theorem of Calculus for path integrals:

Theorem 115. *Let $\gamma : [a,b] \to \mathbb{C}$ be continuous with bounded variation. Let $f : U \to \mathbb{C}$ be continuous on an open set U containing $im(\gamma)$ with primitive F (i.e. $F' = f$). Then*

$$\int_\gamma f(z)dz = F(\gamma(b)) - F(\gamma(a))$$

Proof. Notice that F is continuously differentiable on the compact set $\gamma([a,b])$, so Theorem 87 applies. Consider the absolute difference between the integral estimate with partition P of $[a,b]$ (with endpoints t_i and tags τ_i) and the right hand side of the above equation:

$$
\begin{aligned}
&\left|\sum_{i=1}^{m} f(\gamma(\tau_i))(\gamma(t_i)-\gamma(t_{i-1})) - (F(\gamma(b)) - F(\gamma(a)))\right| \\
=&\left|\sum_{i=1}^{m} f(\gamma(\tau_i))(\gamma(t_i)-\gamma(t_{i-1})) - \sum_{i=1}^{m} F(\gamma(t_i)) - F(\gamma(t_{i-1}))\right| \\
=&\left|\sum_{i=1}^{m} \left(F'(\gamma(\tau_i)) - \frac{F(\gamma(t_i)) - F(\gamma(t_{i-1}))}{\gamma(t_i)-\gamma(t_{i-1})}\right)(\gamma(t_i)-\gamma(t_{i-1}))\right| \\
\leq&\sum_{i=1}^{m}\left|F'(\gamma(\tau_i)) - \frac{F(\gamma(t_i)) - F(\gamma(t_{i-1}))}{\gamma(t_i)-\gamma(t_{i-1})}\right||\gamma(t_i)-\gamma(t_{i-1})|
\end{aligned}
$$

Let $\varepsilon > 0$ and choose $\rho > 0$ according to Theorem 87 so that $w \neq z \in B(z_0,\rho) \cap \gamma([a,b]) \to \left|F'(z_0) - \frac{F(z)-F(w)}{z-w}\right| < \frac{\varepsilon}{v(\gamma)}$. By uniform continuity of γ choose $\delta > 0$ so that $|x-y| < \delta \to |\gamma(x)-\gamma(y)| < \rho$. Then for any partition P with $\|P\| < \delta$ we have $\left|F'(\gamma(\tau_i)) - \frac{F(\gamma(t_i))-F(\gamma(t_{i-1}))}{\gamma(t_i)-\gamma(t_{i-1})}\right| < \frac{\varepsilon}{v(\gamma)}$. Then

$$
\sum_{i=1}^{m}\left|F'(\gamma(\tau_i)) - \frac{F(\gamma(t_i)) - F(\gamma(t_{i-1}))}{\gamma(t_i)-\gamma(t_{i-1})}\right||\gamma(t_i)-\gamma(t_{i-1})| < \sum_{i=1}^{m}\frac{\varepsilon}{v(\gamma)}|\gamma(t_i)-\gamma(t_{i-1})| \leq \frac{\varepsilon}{v(\gamma)}v(\gamma) = \varepsilon
$$

It follows that $\int_\gamma f = F(\gamma(b)) - F(\gamma(a))$ as desired. □

Corollary 116. *Let $f = F'$ on γ, a closed path. Then $\int_\gamma f = 0$.*

Our proof does not rely upon the Fundamental Theorem of Calculus for real integrals, so we may safely deduce it as a corollary.

Corollary 117 (Fundamental Theorem of Calculus, Real Case)**.** *Let $f : [a,b] \to \mathbb{R}$ be a continuous with antiderivative F (i.e. $F' = f$). Then*

$$
\int_a^b f(x)dx = F(b) - F(a)
$$

Proof. Use the identity function for $\gamma : [a,b] \to [a,b]$ in the previous theorem. □

Exercise 46. Deduce the real Fundamental Theorem of Calculus directly from Theorem 87 – without using a path γ – as directly and simply as possible, by using Theorem 115 as a guide.

Exercise 47. In this exercise we compute $\int_\gamma \frac{1}{z}dz$ for a circle γ. Specifically, $\gamma(t) = e^{it}$ on the domain $[0,2\pi]$.

1. Explain why $\ln(z)$ is not an antiderivative of $\frac{1}{z}$ suitable for use with the fundamental theorem of calculus for this problem.
2. Breaking the integral into two parts using $\gamma_1 = \gamma|_{[0,\pi]}$ and $\gamma_2 = \gamma_{[\pi,2\pi]}$, calculate it using the fundamental theorem of calculus for path integrals (twice). Be very specific about the choices of antiderivatives. Provide pictures illustrating the path decomposition.
3. Starting over, calculate the integral using Theorem 112 without breaking it into two parts. Your answer should agree with the previous part.
4. Using the previous parts of this problem and the fundamental theorem, prove that $\frac{1}{z}$ does not have an antiderivative on (all of) the unit circle.

Chapter 5

Complex Integrals

When it comes to calculating integrals, the Fundamental Theorem of Calculus is key. It allows us to calculate any integral over any smooth path, provided we can find an antiderivative.

Example 118. Let $\gamma : [0,1]$ be any smooth path starting from $\gamma(0) = \pi$ and ending at $\gamma(1) = i$. Then $\int_\gamma (\cos(z) + z^2)dz = (\sin(i) + \frac{i^3}{3}) - (\sin(\pi) + \frac{\pi^3}{3})$

We can give hundreds of similar examples. Most math majors have ample practice calculating integrals via antiderivatives in Calculus II, and all those methods – substitution, integration by parts, partial fractions, double angle formulas, etc – apply so long as the antiderivatives in question are continuously differentiable on $\mathbb{C}$, or at least on $\gamma([a,b])$. However, we must use caution that certain real functions with a "reputation" for continuity, in particular $\ln(x)$ and $\sqrt[n]{x}$ have complex versions with discontinuities along the negative real axis. They cannot be used as antiderivatives if $\gamma([a,b])$ contains a negative real number.

Example 119. Let $\gamma : [0, 2\pi] \to \mathbb{C}$ be a "loop around a of radius r": $\gamma(t) = a + re^{it}$ $(r > 0)$, and let $f(z) = \frac{1}{z-a}$. Then $\int_\gamma f(z)dz = 2\pi i$.

Proof. Since γ is smooth we use Theorem 114, Part 3.

$$\int_\gamma f = \int_0^{2\pi} f(\gamma(t))\gamma'(t)dt = \int_0^{2\pi} \frac{1}{a + re^{it} - a} ire^{it}dt = \int_0^{2\pi} idt = 2\pi i$$

□

There is much to learn from this simple example. First, notice that the Fundamental Theorem would suggest the antiderivative $\ln(x-a)$ and the answer $\ln(\gamma(2\pi)-a) - \ln(\gamma(0)-a) = \ln((a+r)-a) - \ln((a+r)-a) = 0$. Why is it wrong? Because $\ln(z-a)$ is discontinuous (and therefore not an antiderivative!) when $z-a$ is a negative real number, which occurs on the image of γ: In particular $\gamma(\pi) = a - r$, so $\ln(\gamma(\pi)) = \ln(-r)$. It's important to remember the locus of discontinuity of $\ln(z)$!

Next, notice that the answer is independent of the radius r. A fast path around a big circle and a slow path around a small circle yield the same path integral. In a sense to be made precise later, the only thing that matters here is that γ winds once counterclockwise around the point $z = a$. It could travel in a square or a trace a picture of Olaf the Snowman and still yield $2\pi i$.

Finally, notice that the value $2\pi i$ of the integral is exactly the size of the jump discontinuity of $\ln(z)$ at any negative real point. If $\ln(z)$ could somehow take multiple values according to the context of the integral, it could remain continuous and would yield $2\pi i$ by the fundamental theorem, but of course a function can't do that. (Can it?)

Exercise 48. In this exercise you'll calculate $\int_\gamma \frac{1}{z}dz$ in two parts using the Fundamental Theorem and both logarithms. Here for convenience $\gamma : [-\pi/2, 3\pi/2] \to \mathbb{C}$ by $\gamma(t) = e^{it}$. The altered domain of γ doesn't change the answer.

1. Explain how we know that $f(z) = \frac{1}{z}$ has an antiderivative defined on $\mathbb{C} - (-\infty, 0]$.

2. Find an antiderivative for $f(z) = \frac{1}{z}$ defined on $\mathbb{C} - [0, \infty)$. (No need to prove correctness.)
3. Let $\gamma : [-\pi/2, 3\pi/2] \to \mathbb{C}$ by $\gamma(t) = e^{it}$. Describe the geometry of the path.
4. Calculate $\int_\gamma \frac{1}{z} dz$ by dividing the integral into $\int_{-\pi/2}^{\pi/2} + \int_{\pi/2}^{3\pi/2}$ and using the fundamental theorem of calculus, twice.
5. Deduce *from your previous answer* that $\frac{1}{z}$ cannot have an antiderivative on its entire domain.

What if the point at which f is undefined is not the center of the circle? Everything works the same:

Example 120. Let $\gamma : [0, 2\pi] \to \mathbb{C}$ be a "loop around a of radius r": $\gamma(t) = a + re^{it}$ $(r > 0)$, and let $f(z) = \frac{1}{z-b}$, with $b \in B(a, r)$. Then $\int_\gamma f(z)dz = 2\pi i$.

Proof. By the previous example, it suffices to show $\int_\gamma \frac{1}{z-a} dz - \int_\gamma \frac{1}{z-b} dz = 0$. It suffices to find an antiderivative for $\frac{1}{z-a} - \frac{1}{z-b}$. Now $\ln(z - a) - \ln(z - b)$ is not a good answer because each part may be discontinuous. We have better luck with $\ln(\frac{z-a}{z-b})$, which has different discontinuities. The value $\frac{z-a}{z-b}$ is negative exactly when $z - a$ is a negative real multiple of $z - b$. As vectors, $z - a$ and $z - b$ point in opposite directions, so z is on the line segment from a to b. Therefore $\ln(\frac{z-a}{z-b})$ is differentiable on the complement in $\mathbb{C}$ of this segment, and the reader may check its derivative is indeed $\frac{1}{z-a} - \frac{1}{z-b}$. This set includes γ, so the fundamental theorem applies. □

Notice that this proof does more than advertised: It works when γ is any path at all disjoint from the line segment from a to b, telling us $\int_\gamma \frac{1}{z-b} dz = \int_\gamma \frac{1}{z-a} dz$. This will be useful later, but for now we can't calculate $\int_\gamma \frac{1}{z-a} dz$ except for circular paths!

5.0.1 Cauchy's Theorem

Early complex integration theory centers around two types of related theorems. The first type asserts that $\int_\gamma f(z)dz = 0$ for a differentiable function f and a certain kind of path γ. We'll prove several such theorems. Students should note that they differ not in the strength of their conclusion, but in the generality of the path γ to which they apply. In a way, Corollary 116 begins this story, but applies only when f has a primitive. The second type asserts that if f is analytic, its values within a region may be predicted from its values on the boundary of a region, if only one can compute the integral of a related function, differentiable except at one point. Again we have several versions of this theorem, differing chiefly in the paths and regions used. We begin with our most specific zero-integral theorem.

Theorem 121. *Let f be differentiable on a triangle $T = \triangle ABC$. Then $\int_{dT} f(z)dz = 0$*

Proof. Our strategy is Goursat's. Suppose the given integral were nonzero. We describe a decomposition of the triangle into four pieces, so that on at least one the integral is "still large." Continuing gives a shrinking sequence of triangles, converging to some point z_0. Since f is differentiable at z_0, we use the tangent line approximation to reduce to the error function g. The limit property of g, coupled with the lower bound on $\int g dz$ will give a contradiction.

Let $T_0 = T = \triangle ABC$. Let $C = \left|\int_{\partial T} f(z)dz\right|$. Let p and d be the perimeter and diameter of T. Let D, E, and F be the midpoints of BC, CA, and AB respectively. Then $\partial\triangle ABC = \partial\triangle AFE + \partial\triangle FBD + \partial\triangle DEF + \partial\triangle EDC$. For at least one of these, which we call T_1, we have $|\int_{\partial T_1} f(z)dz| \geq C/4$. Continuing in this way gives triangles T_n so that $|\int_{\partial T_n} f(z)dz| \geq C/4^n$. Note that T_n has perimeter $P/2^n$ and diameter $d/2^n$. Let z_0 be the point in $\bigcap T_n$ (by compactness).

Let $f = L + g$ be the linear approximation at z_0, with $\lim_{w\to z_0} \frac{g(w)}{w-z_0} = 0$. Since L has a primitive, $\int_{\partial T_n} f(z)dz = \int_{\partial T_n} g(z)dz$. For any $\varepsilon > 0$, for n sufficiently large we have for all $w \in T_n$: $\left|\frac{g(w)}{w-z_0}\right| < \varepsilon$, so $|g(w)| \leq \varepsilon|w - z_0| < \varepsilon d/2^n$. By the ML-inequality,

$$\left|\int_{\partial T_n} f(z)dz\right| = \left|\int_{\partial T_n} g(z)dz\right| \leq \varepsilon d/2^n P/2^n = \frac{\varepsilon dP}{4^n}$$

For ε sufficiently small, this is inconsistent with the lower bound $C/4^n$ above. □

Since the property $\int_{dT} f(z)dz = 0$ applies only to triangle paths, it may seem to specialized to be useful. But it forces f to have a primitive, as the next theorem shows.

Notation 122. For any $z_1, z_2 \in \mathbb{C}$ we write $[z_1, z_2]$ for the path $\gamma : [0,1] \to \mathbb{C}$ by $\gamma(t) = z_1 + t(z_2 - z_1)$.

Theorem 123. *Let f be continuous on a convex set D. If for any triangle $T = \triangle ABC$ in D we have $\int_{dT} f(z)dz = 0$, then f has a primitive $f = F'$.*

Proof. Fix $A \in D$, and let $F(B) = \int_{[A,B]} f(z)dz$. Now fix $B \in D$. Using triangle $\triangle ABC$ we calculate $\dfrac{F(C) - F(B)}{C - B} = \dfrac{\int_{[A,C]} f(z)dz - \int_{[A,B]} f(z)dz}{C - B} = \dfrac{\int_{[B,C]} f(z)dz}{C - B}$. To show $F'(B) = f(B)$, we must prove $\lim_{C \to B} \dfrac{\int_{[B,C]} f(z)dz}{C - B} = f(B)$. Since obviously $\lim_{C \to B} \dfrac{\int_{[B,C]} f(B)dz}{C - B} = f(B)$ it suffices to show $\lim_{C \to B} \dfrac{\int_{[B,C]} (f(z) - f(B))dz}{C - B} = 0$. The last is a simple application of the continuity of f at B and the ML inequality. □

This brings us to our first version of Cauchy's theorem:

Theorem 124. *[Cauchy's Theorem on a Convex Domain] Let f be differentiable on a convex set D. Then f has a primitive on D and $\int_\gamma f = 0$ for any closed path in D.*

Proof. Assemble 121, 123, and 116: □

The integral $\int_\gamma f(z)dz$ of a differentiable function around a loop in a convex domain is zero. Under the right conditions, it follows that $\int_\gamma f(z)dz = \int_\sigma f(z)dz$ if f is differentiable on the region "between" the loops:

Theorem 125. *Let $z \in B(z,r) \subseteq B(z_0, R)$, with boundary paths $\gamma(t) = z_0 + Re^{it}$ and $\sigma(t) = z + re^{it}$. If f is differentiable on $\overline{B(z_0, R)} - z$, then*

$$\int_\gamma f(z)dz = \int_\sigma f(z)dz$$

Proof. Disassemble the irregular annulus into three sections, producing three paths whose sum is $\gamma - \sigma$, each of which is contained in a convex set where f is differentiable. □

According to the previous theorem, if f is differentiable on a convex domain *except* at z, then a wide variety of loops γ around z all produce the same integral value $\int_\gamma f(w)dw$. The next theorem explains how to use the behavior of f on a neighborhood of z to predict the common value of these integrals.

This suggests that $\int_\gamma f(w)dw$ can be calculated by replacing γ with a very small loop σ with the same center z_0. One might hope to approximate:

$$\int_\gamma f(w)dw = \int_\sigma f(w)dw = \int_\sigma \frac{f(w)(w - z_0)}{w - z_0} dz \approx \int_\sigma \frac{L}{w - z_0} dz = 2\pi i L$$

where L is the limit of the term it replaces and the approximation is driven by the smallness of σ. The next theorem makes this idea precise.

Theorem 126. *Let f be differentiable on $\overline{B(z_0, R)} - z$. Let $\gamma(t) = z_0 + Re^{it}$ be the boundary path. Then*

$$\int_\gamma f(w)dw = 2\pi i \lim_{w \to z_0} f(w)(w - z_0)$$

provided the limit exists.

Proof. Let $L = \lim_{w \to z_0} f(w)(w - z_0)$. Let $\varepsilon > 0$ and choose r so that for all $w \in \overline{B(z_0, r)}$ we have $|f(w)(w - z_0) - L| < \varepsilon$. By the previous theorem, we may replace γ with any path $\sigma(t) = z_0 + re^{it}$ of small radius around z_0. We know that $\int_\sigma \frac{L}{w-z_0}dw = 2\pi i L$, so we calculate the difference:

$$\begin{aligned}\left|\int_\sigma f(w)dw - 2\pi i L\right| =& \left|\int_\sigma f(w)dw - \int_\sigma \frac{L}{w-z_0}dw\right| \\ =& \left|\int_\sigma \left(f(w) - \frac{L}{w-z_0}\right)dw\right| \\ =& \left|\int_\sigma \left(\frac{f(w)(w-z_0)-L}{w-z_0}\right)dw\right| \\ \leq& \frac{\varepsilon}{r}2\pi r = 2\pi\varepsilon && \text{By the ML inequality 106}\end{aligned}$$

Since this can be made arbitrarily small by choice of σ, we have equality. □

Exercise 49. Let $\gamma : [0, 2\pi] \to \mathbb{C}$ by $\gamma(t) = e^{it}$. Calculate

$$\int_\gamma z^n \sin(z)dz$$

for every integer n. Hints: The answers are not all equal. There is not need to integrate by parts.

Theorem 127 (Cauchy's Integral Formula for the Disk)**.** *Let f be differentiable on $\overline{B(z_0, R)}$. Let $\gamma(t) = z_0 + Re^{it}$ be the boundary path. Then*

$$f(z) = \frac{1}{2\pi i}\int_\gamma \frac{f(w)}{w-z}dw$$

Proof. The work is done in the previous theorem. We need only calculate

$$\frac{1}{2\pi i}\int_\gamma \frac{f(w)}{w-z}dw = \frac{1}{2\pi i}2\pi i \lim_{w\to z}\frac{f(w)}{w-z}(w-z) = f(z)$$

□

At first glance, the Cauchy Integral Formula looks like a method of evaluating a rather unlikely integral, but it is much more. It gives a formula for $f(z)$ for any $z \in B(z_0, R)$ in terms of the values of f on the circle. It not only demonstrates that any differentiable function on D is entirely determined by its boundary values, but gives an explicit method for performing such a calculation.

To illustrate its power, we will next show how it can be used to produce a power series for any differentiable function. We first need a lemma helping us pass an integral across a limit or sum:

Lemma 128. *Let $X \subseteq \mathbb{C}$ and let f_n be a sequence of functions converging uniformly to f on X, and let γ be a path of bounded variation $v(\gamma) < \infty$ in X. Then*

$$\lim_{n\to\infty}\int_\gamma f_n(z)dz = \int_\gamma f(z)dz$$

If instead $f(z) = \sum_{n=0}^\infty f_n(z)$, with uniform convergence on X, then

$$\sum_{n=0}^\infty \int_\gamma f_n(z)dz = \int_\gamma f(z)dz$$

Proof. For the first claim, let $\varepsilon > 0$ and choose N so that for $n > N$ we have $|f_n(z) - f(z)| < \varepsilon/v(\gamma)$ *uniformly* on X. Then

$$\left|\int_\gamma f_n(z)dz - \int_\gamma f(z)dz\right| = \left|\int_\gamma f_n(z) - f(z)dz\right| \leq \frac{\varepsilon}{v(\gamma)}v(\gamma) = \varepsilon$$

The second claim follows from the first since integrals commute with finite sums. □

Theorem 129. *Let f be differentiable on $\overline{B(z_0,R)}$. Then f has a power series expansion*

$$f(z) = \sum_{n=0}^{\infty} a_n (z-z_0)^n$$

with coefficients

$$a_n = \frac{1}{2\pi i} \int_\gamma \frac{f(w)}{(w-z_0)^{n+1}} dw$$

and radius of convergence at least R.

Proof. We begin with a geometric series for $\frac{1}{w-z}$, centered at z_0:

$$\frac{1}{w-z} = \frac{1}{(w-z_0)-(z-z_0)} = \frac{1}{w-z_0} \cdot \frac{1}{1-\frac{z-z_0}{w-z_0}} = \frac{1}{w-z_0} \cdot \sum_{n=0}^{\infty} \left(\frac{z-z_0}{w-z_0}\right)^n$$

For fixed $z \in B(z_0,R)$ and all w on the boundary, $\left|\frac{z-z_0}{w-z_0}\right| < 1$ is constant in w, so we have uniformly convergence

$$\begin{aligned} f(z) = \frac{1}{2\pi i} \int_\gamma \frac{f(w)}{w-z} dw =& \frac{1}{2\pi i} \int_\gamma f(w) \frac{1}{w-z_0} \cdot \sum_{n=0}^{\infty} \left(\frac{z-z_0}{w-z_0}\right)^n dw \\ =& \sum_{n=0}^{\infty} (z-z_0)^n \frac{1}{2\pi i} \int_\gamma \frac{f(w)}{(w-z_0)^{n+1}} dw \end{aligned}$$

□

This theorem sets complex analysis apart from real analysis more clearly than any other. For real functions the conditions of differentiability, continuous differentiability, twice differentiability, infinite differentiability, etc. are all distinguishable, but this theorem tells us that differentiability on an open set gives us infinite differentiability. It is of course a powerful tool, and we will now see major theorems begin to fall like dominoes.

Theorem 130. *Let f be differentiable on an open set U, and let $a \in U$. If $\forall n \ f^{(n)}(a) = 0$, then f is constantly zero on a neighborhood of a. Moreover the set S of points a for which $\forall n \ f^{(n)}(a) = 0$ is clopen. If U is connected and any such point exists, then $f = 0$ on U.*

Proof. Since f is differentiable on a neighborhood of a, it has a convergent power series with nonzero radius of convergence. Its coefficients $f^{(n)}(a)/n!$ are all zero, so $f = 0$ on an open neighborhood $B(a,r)$ of a. Furthermore the power series for f at any point $b \in B(a,r)$ is clearly zero as well. This proves that S is open. Since S is the intersection of preimages $(f^{(n)})^{-1}(\{0\})$ of closed sets under continuous maps, S is closed. □

Theorem 131 (Morera's Theorem)**.** *Let $V \subset \mathbb{C}$ be open, and let f be continuous on V. If $\int_{\partial T} f(z)dz = 0$ for any triangle $T \subset V$ then f is infinitely differentiable.*

Proof. On any (convex!) ball $B(z_0,R) \subseteq V$, f has a primitive F, and $F' = f$. Since F has a power series, it is twice (in fact infinitely) differentiable, so f is differentiable (infinitely). □

Theorem 132 (Cauchy's Estimates)**.** *Let f be differentiable on $\overline{B(z_0,R)}$, and assume $|f(z)| \leq M$ on the boundary. Then $\left|f^{(n)}(z_0)\right| \leq \frac{Mn!}{R^n}$.*

Proof. From the power series expansion 129, we know $\left|f^{(n)}(z_0)\right| = \left|n!\frac{1}{2\pi i}\int_\gamma \frac{f(w)}{(w-z_0)^{n+1}}dw\right|$. The ML-inequality bounds the integral by $n!\frac{1}{2\pi}\frac{M}{R^{(n+1)}}2\pi R = \frac{Mn!}{R^n}$. □

Definition 133. A function $f : \mathbb{C} \to \mathbb{C}$ is called **entire** if it is defined and differentiable on all of $\mathbb{C}$.

Theorem 134 (Liouville's Theorem)**.** *Any bounded entire function is constant.*

Proof. Assume $|f(z)| \leq M$ for all z. Fix $z_0 \in \mathbb{C}$. Then $|f'(z_0)| \leq M/R$, for all R, so $f'(z_0) = 0$. Since f has constant zero derivative (and $\mathbb{C}$ is connected), f is constant. $\square$

Theorem 135 (Fundamental Theorem of Algebra)**.** *Let $P(x) \in \mathbb{C}[x]$ be a nonconstant polynomial with complex coefficients. Then P has a root in $\mathbb{C}$.*

Proof. Certainly $P(z)$ is entire. If $P(z)$ has no root, then $1/P(z)$ is also entire. Let $a_n z^n$ be the leading coefficient. Now $|P(z)| = |z^n| \left| a_n + a_{n-1}z^{-1} + \ldots + a_0 z^{-n} \right| \geq |z^n|(|a_n| - \left|a_{n-1}z^{-1}\right| - \ldots - |a_0 z^{-n}|)$. Whenever $|z| \geq R = 2n \max(a_i/a_n, 1)$, we continue $P(x) \geq |z^n| \left(|a_n| - \frac{|a_{n-1}a_n|}{2na_{n-1}} - \ldots - \frac{|a_0 a_n|}{2na_0} \right) = |z^n| \left| \frac{a_n}{2} \right| \geq \left| \frac{a_n}{2} \right|$. Thus when $z \geq R$ we have $|\frac{1}{P(z)}| \leq \left| \frac{2}{a_n} \right|$. On the other hand, $\frac{1}{P(z)}$ is bounded on $\overline{B(z,R)}$ by continuity (if $P(z)$ has no roots) and compactness. Thus $\frac{1}{P(z)}$ is constant and so is $P(z)$. $\square$

Our next formula can be interpreted as a generalized Pythagorean theorem. It is also significant in Fourier analysis, where it says that the energy in a waveform f is the sum of the energies represented in each fourier component.

Theorem 136 (Parseval's Formula)**.** *Let $f(z) = \sum_{n=0}^{\infty} a_n(z-a)^n$ be a power series with radius of convergence R, and let $r < R$. Then*

$$\frac{1}{2\pi}\int_0^{2\pi} \left| f(a + re^{it}) \right|^2 dt = \sum_{n=0}^{\infty} |a_n r^n|^2$$

Notice the connection with Fourier series: $f(a+re^{it}) = \sum_{n=0}^{\infty} a_n (re^{it})^n = \sum_{n=0}^{\infty} a_n r^n e^{int} = \sum_{n=0}^{\infty} a_n r^n (\cos(nt) + i\sin(nt))$, a Fourier series with coefficients $a_n r^n$.

Proof. Consider the inner product $\langle f, g \rangle \equiv_{\text{def}} \frac{1}{2\pi}\int_0^{2\pi} f(t)\overline{g(t)}dt$ and functions $b_n(t) = e^{int}$ for $n \in \mathbb{N}$. Now $\int_0^{2\pi} e^{int}dt = \frac{1}{in}e^{int}\Big|_0^{2\pi} = 0$ unless $n = 0$, when $\int_0^{2\pi} e^{int}dt = 2\pi$. It follows that $\langle b_m, b_n \rangle = 1$ if $m = n$ and 0 otherwise. The power series expresses $f(a + re^{it})$ in terms of this basis: $f(a + re^{it}) = \sum_{n=0}^{\infty} a_n (re^{it})^n = \sum_{n=0}^{\infty} a_n r^n b_n(t) = \lim s_N$, where $s_N(t) = \sum_{n=0}^{N} a_n r^n b_n(t)$. Now the convergence $s_N \to f$ is uniform on $\overline{B(a,r)}$ since $r < R$, and the same is true of $|s_N|^2 \to |f|^2$. Therefore

$$\begin{aligned}
\frac{1}{2\pi}\int_0^{2\pi} \left| f(a + re^{it}) \right|^2 dt &= \lim_{N\to\infty} \frac{1}{2\pi}\int_0^{2\pi} |s_N(t)|^2 dt \\
&= \lim_{N\to\infty} \frac{1}{2\pi}\int_0^{2\pi} s_N(t)\overline{s_N(t)}dt \\
&= \lim_{N\to\infty} \frac{1}{2\pi}\int_0^{2\pi} \sum_{m=0}^{N} a_m r^m b_m(t) \overline{\sum_{n=0}^{N} a_n r^n b_n(t)}dt \\
&= \lim_{N\to\infty} \frac{1}{2\pi} \sum_{m=0}^{N}\sum_{n=0}^{N}\int_0^{2\pi} a_m r^m b_m(t)\overline{a_n r^n b_n(t)}dt \\
&= \lim_{N\to\infty} \frac{1}{2\pi} \sum_{n=0}^{N} |a_n r^n|^2 = \frac{1}{2\pi}\sum_{n=0}^{\infty} |a_n r^n|^2 \text{ as desired.}
\end{aligned}$$

$\square$

Theorem 137 (Maximum Modulus Theorem)**.** *Let f be differentiable and nonconstant on a connected open set U. Then $|f|$ does not achieve a local maximum on U.*

Proof. Let $a \in U$. On some ball $B(a, R)$ we have a convergent power series $f(z) = \sum_{n=0}^{\infty} a_n(z-a)^n$, and for some smaller ball $B(a, r)$, Parceval's formula holds. For some $n > 0$ we have $a_n \neq 0$ by 130. Suppose $\left|f(a + re^{it})\right| \leq |f(a)|$ for all t. Then

$$|f(a)|^2 = |a_0|^2 < \sum_{n=0}^{\infty} |a_n r^n|^2 = \frac{1}{2\pi} \int_0^{2\pi} \left|f(a + re^{it})\right|^2 dt \leq \frac{1}{2\pi} 2\pi |f(a)|^2 = |f(a)|^2$$

by Parceval's Formula (136) and the ML Inequality (106). This contradiction shows a cannot be a local maximum of $|f(z)|$. □

If we consider a compact domain, such as $\overline{B(0, R)}$, then of course $|f|$ attains a maximum value. The Maximum Modulus Theorem guarantees that it occurs on the boundary.

5.1 General Cauchy Integral Theorem

The theorems of the previous section testify to the strength of Theorem 127, the Cauchy Integral Formula for the Disk. In this section we will develop a version for more general paths γ. We follow Dixon's approach.

Throughout this section, let $\gamma : [\alpha, \beta] \to \mathbb{C}$ be a path. We assume γ is continuous, and bounded variation (for example, piecewise smooth). Fix a and consider $\int_\gamma \frac{1}{w-a} dw$. Naively we may wish to calculate $\int_\gamma \frac{1}{w-a} dw = \log(\gamma(\beta)) - \log(\gamma(\alpha)) = 0$, but of course this is wrong even for the circular path $\gamma = z + e^{it}$, because $\log(w - a)$ is not an antiderivative of $1/(w - a)$ when $w - a \in (-\infty, 0]$. The next proposition assures us that this wrong calculation is, at worst, off by a multiple of $2\pi i$.

Proposition 138. *Let $a \in \mathbb{C}$. Let $\gamma : [\alpha, \beta] \to \mathbb{C} - \{a\}$. Then*

$$\int_\gamma \frac{1}{w-a} dw = \log(\gamma(\beta)) - \log(\gamma(\alpha)) + 2\pi i n$$

for some $n \in \mathbb{Z}$.

Proof. We consider $x \in [\alpha, \beta]$. Let X be the set of all $x \in [\alpha, \beta]$ for which the theorem is true for $\gamma|_{[\alpha,x]}$. Now fix x_1. Let $\gamma(x_1) = z_1$. If $z_1 - a$ is not a negative real, choose $\varepsilon > 0$ so that $\log(w - a)$ is continuous on $B(z_1, \varepsilon)$ and $\delta > 0$ so that $\gamma(B(x_1, \delta)) \subseteq B(z_1, \varepsilon)$. Let $x_2 \in B(x_1, \delta)$. Let $\gamma_1 = \gamma|_{[\alpha,x_1]}$, $\gamma_2 = \gamma|_{[\alpha,x_2]}$, and $\gamma_\Delta = \gamma|_{[x_1,x_2]}$, so $\int_{\gamma_2} = \int_{\gamma_1} + \int_{\gamma_\Delta}$ Now log is an antiderivative for the integral $\int_{\gamma_\Delta} \frac{1}{w-a} dz = \log(\gamma(x_2)) - \log(\gamma(x_1))$. Thus the theorem statement is true for x_1 if and only if it is true for x_2.

Now if $z_1 - a$ is a negative real, consider the function $\log'(z) = \ln(|z|) + i \arg'(z)$, where $\arg'(z) \in [0, 2\pi)$. Then $\log'(z) - \log(z) \in \{0, 2\pi i\}$. Moreover, $\log'(w - a)$ is an antiderivative of $\frac{1}{w-a}$ unless $w - a$ is a *positive* real. Using $\log'$ in place of log, the above argument goes through again.

We conclude that $x_1 \in X$ iff $x_2 \in X$. This shows that X is both open and closed. Since $\alpha \in X$, we deduce $\beta \in X$ as desired. □

In the case of a closed curve γ, $\log(\gamma(\beta)) - \log(\gamma(\alpha)) = 0$. The number n simply counts the number of times γ winds around z in the counterclockwise direction. We cannot assert this as a theorem without the algebraic topology to define "winds", so instead we take it as a definition:

Definition 139. Let $a \in \mathbb{C}$. Let $\gamma : [\alpha, \beta] \to \mathbb{C} - \{a\}$, $v(\gamma) < \infty$. The **Winding number** of γ around a is:

$$n(\gamma, a) \equiv_{def} \frac{1}{2\pi i} \int_\gamma \frac{1}{w-a} dw$$

Proposition 140. *For fixed γ with $v(\gamma) < \infty$, the winding number $W(\gamma, z)$ is a locally constant function of $z \in \mathbb{C} - im(\gamma)$.*

Proof. This is essentially a continuity argument. Fix $z \notin im(\gamma)$. Since γ is continuous, $im(\gamma)$ is compact, so closed. Let $B(z, \varepsilon_1)$ be disjoint from γ, and consider $\varepsilon_2 < \varepsilon_1$ to be decided. Let $z' \in B(z, \varepsilon_1)$. Then

$$W(\gamma, z) - W(\gamma, z') = \frac{1}{2\pi i} \int_\gamma \frac{1}{w-z} - \frac{1}{w-z'} dw = \frac{1}{2\pi i} \int_\gamma \frac{z - z'}{(w-z)(w-z')} dw$$

This integral can be bounded above (ML inequality) by $\frac{v(\gamma)\varepsilon_2}{2\pi\varepsilon_1^2}$. For sufficiently small $\varepsilon_2 < \varepsilon_1$, this is less than 1, so the $W(\gamma, z) - W(\gamma, z') = 0$ as desired. □

Here's another proof, arguably a better proof:

Proof. It suffices to show that $\int_\gamma \frac{1}{w-z} dw$ is a locally constant function of z, so let $z, z' \in B(z, \varepsilon)$, where $B(z, \varepsilon)$ lies in one component of $\mathbb{C} - \{\gamma\}$. Then

$$\int_\gamma \frac{1}{w-z} dw - \int_\gamma \frac{1}{w-z'} dw = \int_\gamma \frac{1}{w-z} - \frac{1}{w-z'} dw$$

As we've proved earlier, this integrand has an antiderivative $\log\left(\frac{w-z}{w-z'}\right)$ everywhere except the line segment connecting z to z'. By the fundamental theorem of calculus, the integral is zero. □

We take for granted that the complement of a path $\mathbb{C} - im(\gamma)$ has exactly one unbounded component. A path γ can't wind around a point in that component. This is obvious geometrically, but not from our integral definition.

Proposition 141. *If $\gamma : [\alpha, \beta] \to \mathbb{C}$, with $v(\gamma) < \infty$. Let z be in the unbounded component of $\mathbb{C} - im(\gamma)$. Then $W(\gamma, z) = 0$.*

Proof. By local constancy, $W(\gamma, z)$ is constant on the component of interest. When z is sufficiently far from $im(\gamma)$, (distance $\geq M$, say) we have $\left|\frac{1}{w-z}\right| < 1/M$, so $|W(\gamma, z)| \leq \frac{v(\gamma)}{2\pi M}$. For M sufficiently large, this is close to 0, so equal to 0. Since $W(\gamma, z)$ is constant for such z, $W(\gamma, z) = 0$ for all z in the component. □

5.2 Singularities

We turn our attention to **singularities** of a function f – points where f fails to be defined. Properties of singularities z_0 depend not on $f(z_0)$, which does not exist, but on $f(z)$ for z nearby z_0.

Definition 142. Let $z_0 \in \mathbb{C}$.

1. A **deleted neighborhood** of z_0 is a set of the form $U - \{z_0\}$, where $z_0 \in U$ and U is open.
2. A **singularity** of a function f is a point z_0 so that $f(z_0)$ is analytic on a deleted neighborhood of z_0 but undefined or not analytic at z_0.
3. If $z_0 \in \mathbb{C}$, the phrase **near** z_0 means "on some deleted neighborhood of z_0. For example, "f is constant near z_0" means there is some open set $U \ni z$ for which $f|_{U-\{z_0\}}$ is constant.

Notice that by definition a singularity a is isolated in the sense that some neighborhood of a has no singularities.

Example 143. Some common singularities.

1. The function
$$f(z) = \frac{e^z}{(z-3)^2}$$
is undefined at $z_0 = 3$. As z approaches 3, the numerator approaches $e^3 \approx 20.1$ and the denominator approaches 0, so the fraction is unbounded. Thus, there is no value which can be assigned to $f(3)$ which makes it continuous at 3. We call this singularity a **pole of order 2**.

2. The function
$$g(z) = \frac{z^2 - 9}{(z-3)}$$
is undefined at $z_0 = 3$. For all $z \neq z_0$, we have $g(z) = z + 3$. We call this singularity **removable** because the function can be assigned a value at z_0 which repairs the problem.

3. The function
$$h(z) = \frac{(z-3)^2(z-4)^2}{(z-3)}$$
is undefined at $z_0 = 3$. Again we can cancel, removing the singularity. In this case the resulting function has a root at 3.

4. The function
$$k(z) = e^{1/z}$$
is undefined at $z_0 = 0$. As z approaches $z_0 = 0$, the exponent goes to infinity. Because *the exponent* goes to infinity, the function $k(z)$ goes to infinity very quickly and has some strange properties. This is an example of an **essential** singularity.

Definition 144. Let f have a singularity at z_0. If there is some value w such that the function
$$\hat{f}(z) = \begin{cases} f(z) & z \neq z_0 \\ w & z = z_0 \end{cases}$$
is analytic at z_0, we call z_0 a **removable** singularity of f.

Proposition 145 (Riemann Continuation Theorem)**.** *If z_0 is a singularity of f and f is bounded near z_0, then z_0 is removable.*

Proof.
$$\text{Let } g(z) = \begin{cases} f(z)(z-a)^2 & z \neq a \\ 0 & z = 0 \end{cases}$$
If $z \neq z_0$ then $g'(z)$ can be calculated by the product rule. $g'(z_0) = \lim_{z\to z_0} f(z)(z - z_0) = 0$ since $z - z_0 \to 0$ and $f(z)$ is bounded near z_0. It follows that g has a power series expansion $g(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$, and $a_0 = a_1 = 0$. We may divide by $(z - z_0)^2$ termwise to find a power series expansion for $f(z)$ centered at a. ☐

The Riemann Continuation Theorem means that complex function theory has nothing like the "jump discontinuities" from ordinary real Calculus.

Roots and Poles should be understood as opposites, but first we need to understand functions which have neither:

Proposition 146. *Let g be analytic on a neighborhood of z_0. (So z_0 is not a singularity of g.) Then the following are equivalent:*

1. $g(z_0) \neq 0$
2. $1/g$ *is analytic on some (possibly smaller) neighborhood of* z_0
3. *The power series expansion of g at z_0 is of the form $g(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$ with $a_0 \neq 0$.*

Proof. From prior work, f has a power series expansion with positive radius of convergence, and $f(z_0) = a_0$. Thus $1 \leftrightarrow 3$ follows. Moreover $2 \to 1$ is obvious. For $1 \to 2$, note that f is nonzero on some neighborhood z_0 by continuity, and on this neighborhood $1/f$ is analytic by the quotient rule. ☐

Definition 147. Let f be analytic on a neighborhood of z_0. If $f(z) = (z - z_0)^k g(z)$ (for positive integer k) and $g(z_0) \neq 0$, we call z_0 a **root of order** k of f.

Definition 148. Let f be analytic near z_0. If $f(z) = (z - z_0)^{-k} g(z)$ (for positive integer k) and $g(z_0) \neq 0$, we call z_0 a **pole of order** k of f.

Definition 149. Let f be analytic near z_0. The **valuation** $v_{z_0}(f)$ of f at z_0 is defined to be

$$v_{z_0}(f) = \begin{cases} 0 & f \text{ if analytic at } z_0 \text{ and } f(z_0) \neq 0 \\ k & f \text{ has a root of order } k \text{ at } z_0. \\ -k & f \text{ has a pole of order } k \text{ at } z_0. \\ \infty & f = 0 \text{ on a neighborhood of } z_0. \\ \text{(undefined)} & \text{otherwise} \end{cases}$$

Proposition 150. *Let f be analytic near z_0.*

1. *If f has a root of order k at z_0 then f has power series expansion at z_0:*

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n$$

and $a_k \neq 0$ is the first nonzero coefficient.

2. *If f has a pole of order k at z_0 then near z_0, f can be written*

$$f(z) = \sum_{n=-k}^{\infty} a_n (z - z_0)^n$$

and $a_{-k} \neq 0$. This is called the **Laurent Series Expansion** *of f at z_0.*

3. *If f has a root (pole) of order k at z_0 then $1/f$ has a pole (root) of order k at z_0*
4. *$v_{z_0}(f)$ is the index of the first nonzero term in the Laurent (or power) series expansion of f.*
5. *$v_{z_0}(fg) = v_{z_0}(f) v_{z_0}(g)$ if the right hand side is defined.*

Proof. Directly from the definitions, using properties of g proved above. □

Just as locally bounded functions must have removable singularities, functions which become locally bounded when multiplied by $(z - z_0)^k$ must have poles:

Proposition 151. *If f is not bounded near a, but $f(z)(z - z_0)^k$ is bounded near z_0 for some k, and if k is minimal with this property, then f has a pole of order k at z_0.*

Proof. After multiplying by $(z - z_0)^k$ we may invoke the Riemann Continuation theorem. Subsequently dividing by $(z - z_0)^k$ we get a Laurent series whose first nonzero term is $a_{-k}(z - z_0)^{-k}$, which indicates a pole of order k. The coefficient a_{-k} is nonzero by the minimality of k. □

Definition 152. Let $U \subset \mathbb{C}$ be open. A **meromorphic function on** U is a function f analytic on U except on a set of points where f has poles. (Of course f may also have zeroes.) The set of all meromorphic functions on U is written $\mathcal{M}(U)$.

Proposition 153. *Let $U \subset \mathbb{C}$ be open. Then $\mathcal{M}(U)$ is a field under function addition and multiplication.*

Proof. We know that function addition and multiplication obey the field axioms. The point of the proposition is to identify additive and multiplicative identities and prove that if f and g are meromorphic, then so are $-f$, $f + g$, fg. Also, if $g \neq 0$ then $1/g$ must be meromorphic. Each part makes a nice exercise. Work locally at each singularity z_0. □

We have described functions bounded near z_0, which have removable singularities. We have described functions that become bounded only when multiplied by $(z - z_0)^k$, which have poles. The next theorem addresses the only other case. It describes the most exotic type of singularity.

Theorem 154 (Casorati-Weierstrass)**.** *If $f(z)(z - z_0)^n$ is not bounded near z_0 for any n, then $f(B(z_0, \varepsilon))$ is dense in $\mathbb{C}$ for every $\varepsilon > 0$. In this case we say that f has an* **essential singularity.**

Proof. If $f(B(z_0,\varepsilon)-\{z_0\})$ is not dense, choose an open ball $B(b,\delta)$ disjoint from its image. Then the function $g(z)=\frac{1}{f(z)-b}$ is analytic and nonzero on $B(z_0,\varepsilon)-\{z_0\}$ and bounded above by $1/\delta$. Since g is bounded it may be continued to z_0 (so we assume $g(z_0)$ exists). Solving for f gives:

$$f(z)=\frac{1}{g(z)}+b$$

Thus f is meromorphic on some neighborhood of $B(z_0,\varepsilon)$. It follows that $f(z)(z-z_0)^n$ is bounded near z_0 for some n. This proves (the contrapositive of) the claim. □

5.2.1 Generalized Cauchy Integral Formula

Lemma 155. *Let $g(z,w)$ be a continuous function of z and w on a domain $U\times V$, analytic in z for each w. Assume U is open and V is locally compact. Then the partial derivative function $\frac{\partial}{\partial z}g(z,w)=g_z(z,w)$ is likewise a continuous function of two variables. (and of course analytic in z)*

Proof. Fix z_0 and w_0.

Case 1: Assume $g(z_0,w_0)=g_z(z_0,w_0)=0$. By the definition of g_z we have $\lim_{z\to z_0}\frac{g(z,w_0)}{z-z_0}=0$. Let $\varepsilon>0$ and choose r so that $\overline{B(z_0,2r)}\subseteq U$ and whenever $|z-z_0|\le 2r$ then $\left|\frac{g(z,w_0)}{z-z_0}\right|<\varepsilon/2$. Thus on $B(z_0,2r)\times\{w_0\}$ we have $|g(z,w)|<r\varepsilon/2$ Notice that Cauchy's bound for $f'(z)$ says $|g_z(z_0,w_0)|\le\frac{\frac{r\varepsilon}{2}1!}{r}=\varepsilon/2$, which of course we already knew. We wish to extend this bound to a neighborhood of (z_0,w_0). On the compact region $\overline{B(z_0,2r)}\times\overline{B(w_0,r)}$ choose $\delta>0$ (with also $\delta<r$) for $r\varepsilon/2$ in the uniform continuity of g. We imagine an annulus in U with radii $r-\delta$ and $r+\delta$, given a thickness 2δ in the V-direction. It follows by the triangle inequality that if $r-\delta<|z-z_0|<r+\delta$ and $|w-w_0|<\delta$ then $|g(z,w)|=|g(z,w)-g(z,w_0)+g(z,w_0)|\le|g(z,w)-g(z,w_0)|+|g(z,w_0)|<r\varepsilon$. This bounds g along the boundary of any ball (in the U-direction) of radius r around any (z_1,w_1) in the neighborhood $B(z_0,\delta)\times B(w_0,\delta)$. Then for any such (z_1,w_1), Cauchy's bound gives $|g_z(z_1,w_1)|<\varepsilon$ as desired. This shows that g is a continuous function of two variables at z_0,w_0 in this case.

Case 2: Apply Case 1 to the function $g(z,w)-g_z(z_0,w_0)(z-z_0)-g(z_0,w_0)$, a function of z and w whose partial with respect to z is $g_z(z,w)-g_z(z_0,w_0)$. This shows that $g_z(w,z)$ plus a constant is continuous, thus so is $g_z(w,z)$. □

Theorem 156 (Leibniz Integral Formula)**.** *Let $g(z,w)$ a continuous function of two variables, and analytic in z for each w. Let γ be a continuous, bounded variation path. Then*

$$\frac{d}{dz}\int_\gamma g(z,w)dw=\int_\gamma g_z(z,w)dw$$

Proof. Fix z_0. For each w consider the tangent line approximation $g(z,w)=g_z(z_0,w)(z-z_0)+g(z_0,w)+e(z,w)$ where $\lim_{z\to z_0}\frac{e(z,w)}{z-z_0}=0$. By the previous lemma, $g_z(z_0,w)$ is a continuous function of w, so $e(z,w)$ is a continuous function of z and w.

$$\left(\frac{d}{dz}\int_\gamma g(z,w)dw\right)-\int_\gamma g_z(z,w)dw=\lim_{z\to z_0}\left[\frac{\int_\gamma g(z,w)dw-\int_\gamma g(z_0,w)dw}{z-z_0}-\int_\gamma g_z(z_0,w)dw\right]=\lim_{z\to z_0}\int_\gamma\frac{e(z,w)}{z-z_0}dw$$

On a compact region $\overline{B(z_0,\varepsilon)}\times im(\gamma)$, the function e is uniformly continuous, and by 107 the limit commutes with the integral:

$$\lim_{z\to z_0}\int_\gamma\frac{e(z,w)}{z-z_0}dw=\int_\gamma\lim_{z\to z_0}\frac{e(z,w)}{z-z_0}dw=0$$

□

Definition 157. Let $U\subseteq\mathbb{C}$ be open. Let γ be a closed path in U. We call γ **homologous to zero in** U if $W(\gamma,a)=0$ for all $a\notin U$.

Theorem 158 (Cauchy Integral Formula, general case)**.** *Let $U \subseteq \mathbb{C}$ be open, and let γ be homologous to zero in U. Let f be analytic on U. Then for all $z \in U$*

$$\int_\gamma \frac{f(w)}{w-z} dw = 2\pi i W(\gamma, z) f(z)$$

Proof. Now consider a path γ and a function f. Write E for the union of the components of $\mathbb{C} - \{\gamma\}$ on which $W(\gamma, z) = 0$. Notice that E includes the unbounded component of $\mathbb{C} - \{\gamma\}$, by 141. Let us assume f is differentiable on U, and $U \cup E = \mathbb{C}$. As before,

$$\text{let } q(z, w) = \begin{cases} \frac{f(w)-f(z)}{w-z} & \text{if } w \neq z \\ f'(z) & \text{if } w = z \end{cases}$$

Since f is differentiable, it is continuously so, and q is uniformly continuous on any compact set. For fixed z, $q(z, w)$ is a differentiable function of w and continuous at $w = z$, where it therefore has a removable singularity. Thus q is analytic in each variable separately.

$$\text{Let } h(z) = \begin{cases} \int_\gamma q(z, w) dw & \text{if } z \in U \\ \int_\gamma \frac{f(w)}{w-z} dw & \text{if } z \in E \end{cases}$$

On the (open!) overlap $U \cap E$, these agree because $\int_\gamma \frac{f(z)}{w-z} = 0$ by assumption. We claim $h(z)$ is entire. On U, the Leibniz rule applies and $h'(z) = \frac{d}{dz} \int_\gamma q(z, w) dw = \int_\gamma q_z(z, w) dw$. Notice q_z is a continuous function w, in fact of two variables, so the integral exists, so h' exists.

On $E - U$, Leibniz rule gives $\frac{d}{dz} \int_\gamma \frac{f(w)}{w-z} dw = \int_\gamma \frac{f(w)}{(w-z)^2} dw$. Again this is continuous since $w \neq z$ for $w \in \{\gamma\}$.

However, for z sufficiently far from the compact set $\{\gamma\}$ (say when $d(z, \{\gamma\}) > R$) we bound $h(z) = \int_\gamma \frac{f(w)}{w-z} dw$ by ML:

$$\left| \int_\gamma \frac{f(w)}{w-z} dw \right| \leq \max_{w \in \{\gamma\}} |f(w)| \frac{1}{R} |v(\gamma)|$$

That is, h is bounded for $\{z \in \mathbb{C} | d(z, \{\gamma\}) > R\}$, and bounded on $\{z \in \mathbb{C} | d(z, \{\gamma\}) \leq R\}$ by compactness. It follows that h is constant. Since R is arbitrary in the bound $|f(w)| \frac{1}{R} |v(\gamma)|$, h is the zero function.

Therefore, for all $z \in U$, we have

$$\int_\gamma \frac{f(w) - f(z)}{w-z} dw = 0$$

That is,

$$\int_\gamma \frac{f(w)}{w-z} dw = \int_\gamma \frac{f(z)}{w-z} dw = 2\pi i f(z) W(\gamma, z)$$

□

Corollary 159. *Let $U \subseteq \mathbb{C}$ be open, and let f be analytic on U, γ closed and homologous to zero in U. Then $\int_\gamma f(w) dw = 0$.*

Proof. $f(w) = f(w)(w-z)/(w-z)$. □

Corollary 160. *Let $U \subseteq \mathbb{C}$ be open, and let f be meromorphic on U, γ closed and homologous to zero in U. Then $\int_\gamma f(w) dw = 2\pi i \sum_i W(\gamma, z_i) Res(f, z_i)$.*

Proof. Subtract off higher negative power terms b/c they have antiderivatives. Subtract $\prod R_i/(w - z_i)$ and remove singularities to produce analytic function. Apply previous corollary. □

5.3 Using the method of residues to calculate real integrals

The residue theorem is so powerful that it can solve many ordinary real integrals, particularly improper ones:

Example 161. We calculate $\displaystyle\int_0^\infty \frac{1}{1+x^2}dx$.

With some familiarity with inverse trigonometric functions, one calculates the integral as $\lim_{b\to\infty} \arctan(b) - \arctan(0) = \pi/2$. In contrast, complex analysis solves the problem without leaving the domain of rational functions.

First we notice that the function is even and has residues at i and $-i$. We double the integral to produce the more convenient problem $\displaystyle\int_{-\infty}^\infty \frac{1}{1+x^2}dx$, replace it with the finite version $\displaystyle\int_{-R}^R \frac{1}{1+x^2}dx$, and regard this straight line real integral as the lower part of a path integral on γ, a path which proceeds from $-R$ to R along the real line, and then from R to $-R$ along a complex upper semicircle centered at 0. If we can calculate both the residue at i (interior to this path) and the integral along the semicircle, we can deduce the straight line integral from Cauchy's integral formula. The residue at i is $\lim_{z\to i}(z-i)\frac{1}{1+z^2} = \lim_{z\to i}\frac{1}{z+i} = \frac{1}{2i}$. The semicircular integral is harder to calculate, but the ML inequality bounds it at $\frac{2\pi R}{R^2-1}$, which goes to zero at R goes to infinity. In the limit, then, we have

$$\lim_{R\to\infty}\int_{-R}^R \frac{1}{1+z^2}dz = 2\pi i\frac{1}{2i} = \pi$$

Remember we have doubled the original problem, so the answer is $\pi/2$ as expected from trigonometry.

Example 162. We calculate $\displaystyle\int_{-\infty}^\infty \frac{1}{1+x^4}dx$. It's possible to factor this into two irreducible quadratics, complete squares, and apply trigonometric substitutions, but it's not fun. Instead we integrate as above. Again the semicircular integral is approaches zero, so we have

$$\lim_{R\to\infty}\int_{-R}^R \frac{1}{1+z^4}dz = 2\pi i(\text{sum of residues})$$

So we need the poles and residues. The roots are z^4+1 are $e^{\frac{1}{4}\pi i}$, $e^{\frac{3}{4}\pi i}$, $e^{\frac{5}{4}\pi i}$, and $e^{\frac{7}{4}\pi i}$. Setting $\alpha = e^{\frac{1}{4}\pi i}$, we write α^3, α^5, and α^7. Only α and α^3 are in the upper half plane, so we calculate the residues there. Note of course $\alpha^4 = -1$

$\lim_{z\to\alpha}(z-\alpha)\frac{1}{z^4+1} = \lim_{z\to\alpha}(z-\alpha)\frac{1}{(z-\alpha)(z-\alpha^3)(z-\alpha^5)(z-\alpha^7)} = \lim_{z\to\alpha}\frac{1}{(z-\alpha^3)(z-\alpha^5)(z-\alpha^7)} = \frac{1}{(\alpha-\alpha^3)(\alpha-\alpha^5)(\alpha-\alpha^7)} = \frac{1}{\alpha^3(1-\alpha^2)(1-\alpha^4)(1-\alpha^6)} = \frac{\alpha}{(-1)(1-i)(2)(1+i)} = \frac{-\alpha}{4}$

A sneakier calculation is to recognize the limit $\lim_{z\to\alpha}(z-\alpha)\frac{1}{z^4+1}$ as the reciprocal of the derivative of $f(z) = z^4+1$ at $z=\alpha$, which is $4\alpha^3$. The reciprocal is $-\alpha/4$. SImilarly, the residue at α^3 is the reciprocal of $4\alpha^9 = 4\alpha$, or $-\alpha^3/4$. The sum of the residues is $\frac{-1}{4}(\alpha+\alpha^3) = \frac{-\alpha}{4}(1+\alpha^2) = \frac{-\alpha}{4}(1+i) = \frac{-\alpha\sqrt{2}\alpha}{4} = \frac{-\sqrt{2}i}{4}$. We conclude:

$$\lim_{R\to\infty}\int_{-R}^R \frac{1}{1+4^2}dz = 2\pi i\frac{-\sqrt{2}i}{4} = \frac{\pi}{\sqrt{2}}$$