Notes for Math 121A

Harrison Bachrach

Spring 2015

Notes for Math 121A based on lectures given by Dr. Nikhil Srivastava in Spring 2015 at UC Berkeley in Berkeley, California. The following notes assume basic knowledge of series, multivariable calculus and linear algebra; however, the first few sections are dedicated to the brief review of those topics.

For those of you scientists and engineers who have taken up to linear algebra & multivariable calculus *but no further*, I entreat you to continue reading. After the review, Sections 5 and 6 (on Complex Analysis and Harmonic Analysis respectively) demonstrate the humbling power and beauty of higher-level mathematics. While the topics are perhaps not expressed here with the rigor that would satisfy a mathematician, I hope that they can offer a stimulating preview of "upper division" mathematics.

Sections suffixed with an asterix, i.e. the "*" symbol, are generally "optional" in that they don't *necessarily* have applicability to the physical sciences. However, they are often quite beautiful and/or significant within pure mathematics.

Contents

1 Notation

1.1 Screen & Print Notation

Many symbols of set theory and boolean logic are useful beyond their original applications. In this section, I will be explaining the meaming of those symbols which are used here which you may not have encountered previously.

- A *set* of things is denoted by curly braces. E.g. The set of natural numbers $\mathbb{N} = \{0, 1, 2, 3, \ldots\}$.

- The symbol $\in$ means "in" or "is an element of". For example, the statement $2 \in \mathbb{N}$ is true, while $-3 \in \mathbb{N}$ is not.

- The symbol $\forall$ means "for all". For example, $\forall x \in \mathbb{N}\ x > -1$ is a true statement.

- The symbol $\Rightarrow$ means "implies", that is, "$p \Rightarrow r$" is to be read as "if p is true, then q is also true". Note that it states nothing about the status of p if we are given the status of q, nor does it tell us the status of q if p is false.

- The symbol $\iff$ means "if and only if"[1], that is, "$p \iff r$" is to be read as "if and only if p is true, then q is also true" *and* "if and only if q is true, then p is also true." Intuitively, this logically links p and q.

- The symbol $\wedge$ means "and". A statement "$P(x) \wedge Q(x)$" is only true if *both* $P(x)$ and $Q(x)$ are true statements. E.g., $(5 > 3 \wedge 2 \neq 0)$ is true, while $(3 > 6 \wedge 4 = 4)$ is false.

- We can build sets by using *set builder notation*. The notation works as follows: a set S of elements that satisfy a condition is denoted

  $$S = \{\text{Element} \mid \text{Condition on element}\}.$$

  For example, the set of natural numbers can be expressed as

  $$\mathbb{N} = \{x \mid x \in \mathbb{Z} \wedge x \geq 0\}$$

  where $\mathbb{Z}$ is the set of integers.

- A very useful notation for talking about intervals on the real line is *interval notation*. It works as follows:

  $$\begin{aligned}(a, b) &= \{x \mid x \in \mathbb{R} \wedge a < x < b\} \\ [a, b) &= \{x \mid x \in \mathbb{R} \wedge a \leq x < b\} \\ (a, b] &= \{x \mid x \in \mathbb{R} \wedge a < x \leq b\} \\ [a, b] &= \{x \mid x \in \mathbb{R} \wedge a \leq x \leq b\}\end{aligned}$$

[1] "if and only if" is often written as "iff"

1.2 Handwritten Notation

2 Series

Definition 2.1. A *series* is the sum of the items in a sequence.

Finite sums are always well defined, while *infinite* sums must meet certain criteria to be well defined. A finite sum up to term a_N would be denoted

$$\sum_{n=0}^{N} a_n = S_N$$

While an infinite sum, that is the limit of a finite sum (assuming the limit exists) is denoted as

$$\lim_{N\to\infty} S_N = \lim_{N\to\infty} \sum_{n=0}^{N} a_n = \sum_{n=0}^{\infty} a_n = S$$

If the above limit exists, the series is *convergent*, while if it doesn't, the series is *divergent*.

Definition 2.2. A *geometric series* is one where terms are related by a common ratio r. The general geometric series up to order N could be expressed as

$$S_N = a + ar + ar^2 + \cdots + ar^N$$

The formula for a finite geometric series is

$$S_N = \frac{a(1-r^N)}{1-r} \tag{2.0.1}$$

and the formula for an infinite geometric series (assuming the series exists, i.e. assuming $|r| < 1$) is simply (by taking the limit $N \to \infty$),

$$S = \frac{a}{1-r} \tag{2.0.2}$$

where in both cases, a and r represent the first term and common ratio, respectively.

Definition 2.3. The *remainder* R_N is the difference between the infinite series S

and the "partial sum" S_N, that is,

$$R_N = \sum_{k=N+1}^{\infty} a_k = S - S_N$$

Series are a way of breaking something difficult into a sum of easy to handle terms.

Expand this in relation to rest of course

2.1 Tests for Convergence

Definition 2.4 (Comparing Magnitudes)**.**

$$f \ll g \iff \lim_{n\to\infty} \frac{f(n)}{g(n)} = 0$$

Do note some general relative magnitudes of growth (assume $n, k \in \mathbb{Z}$):

$$\log n \ll n \ll n^k \lll 2^n \ll n! \tag{2.1.1}$$

2.1.1 Preliminary Test

If $\lim_{n\to\infty} a_n \neq 0$, then $\sum_n^\infty a_n$ diverges.

2.1.2 Integral Test

If a_n is non-negative and non-increasing, then

$$\sum_{n=1}^{\infty} a_n \text{ converging} \iff \int_b^\infty a(x)\,dx \text{ converging}$$

where $a(x)$ is the continuation of a_n for the real variable x; e.g. if $a_n = 1/n$, then $a(x) = 1/x$, where $x \in \mathbb{R}$.

2.1.3 Comparison Principle

Suppose a_n and b_n are sequences with $0 \leq a_n \leq b_n$ for all n.

$$\sum b_n \text{ converges} \Rightarrow \sum a_n \text{ converges.}$$
$$\sum a_n \text{ diverges} \Rightarrow \sum b_n \text{ diverges.}$$

2.1.4 Special Comparison Principle

If a_n, b_n are non-negative sequences and $\lim_{n\to\infty} \frac{a_n}{b_n} < \infty$ (that is to say, the limit is finite), then

$$\sum b_n \text{ converges} \Rightarrow \sum a_n \text{ converges.}$$

2.1.5 Ratio Test

Defining

$$\rho_n = \left|\frac{a_{n+1}}{a_n}\right|, \quad \rho = \lim_{n\to\infty} \rho_n = \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|$$

then

$$\begin{aligned}
&\rho < 1 \Rightarrow \text{the series converges.} \\
&\rho > 1 \Rightarrow \text{the series diverges.} \\
&\rho = 1 \Rightarrow \text{the test is inconclusive.}
\end{aligned}$$

2.1.6 Alternating Series Test

If a_n is an alternating series (i.e., $\operatorname{sign}(a_{n+1}) = -\operatorname{sign}(a_n)$), $|a_{n+1}| \leq |a_n|$, *and* $\lim_{n\to\infty} = 0$, then the series converges.

Note: if $\sum |a_n|$ converges, the series is *absolutely convergent*.

2.2 Power Series

Definition 2.5. A *power series* is a function expressed as an infinite sum of monomials. That is, a function $f(x)$ can be expanded in a power series if it can be expressed as

$$f(x) = \sum_{n=0}^{\infty} a_n x^n$$

for some $\{a_n\}$.

Theorem 2.1. A power series expansion of $f(x)$ converges on only 3 types of intervals.

1. It converges everywhere (that is, $\forall x$)
2. It converges for $x = 0$ only
3. It converges when $|x| < R$ and diverges when $|x| > R$, where R is the radius of convergence. (The points $x = \pm R$ must be checked explicitly and are not, in general, symmetric.)

2.2.1 Taylor Series

Definition 2.6. The power series expansion of an analytic function about a number a is known as a *Taylor series*. When $a = 0$, this is known as a *Maclaurin series*. If $f(x)$ is analytic,

$$f(x-a) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(n)}}{n!}(x-a)^n + \cdots$$

Note: the radius of convergence for Taylor series depends upon a.

Here are some common Maclaurin series, with their interval of convergence in parentheses:

$\sin(x)$	$=$	$\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!}$	$=$	$x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$	$(x \in \mathbb{R})$
$\cos(x)$	$=$	$\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}$	$=$	$1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots$	$(x \in \mathbb{R})$
e^x	$=$	$\sum_{n=0}^{\infty} \frac{x^n}{n!}$	$=$	$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$	$(x \in \mathbb{R})$
$\log(1+x)$	$=$	$\sum_{n=1}^{\infty} \frac{(-1)^{n+1} x^n}{n}$	$=$	$x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} \cdots$	$(x \in (-1, 1]\,)$
$(1+x)^p$	$=$	$\sum_{n=0}^{\infty} \binom{p}{n} x^n$	$=$	$1 + px + \frac{p(p-1)}{2!}x^2 + \frac{p(p-1)(p-2)}{3!}x^3 + \cdots$	$(x \in (-1, 1)\,)$

2.2.2 Asymptotic Notation

How do we write higher order terms we are not concerned with so that we can keep track of them? With "Little-oh" and "Big-Oh" notation!

Definition 2.7 (Little-oh notation)**.** Given continuous functions $f(x)$ and $g(x)$, we say that $f(x) = o(g(x))$ as $x \to a$ if

$$\lim_{x \to a} \frac{f(x)}{g(x)} = 0$$

e.g. $x^5 = o(x)$ as $x \to 0$, $x^4 = o(x^5)$ as $x \to \infty$, etc.

Definition 2.8 (Big-Oh notation)**.** Given continuous functions $f(x)$ and $g(x)$, we say that $f(x) = (g(x))$ as $x \to a$ if

$$\lim_{x \to a} \left| \frac{f(x)}{g(x)} \right| < \infty$$

e.g. $x^2 = O(x^2)$ as $x \to 0$, $2 \sin x = O(1)$ as $x \to \infty$, etc.

The rules for manipulating the notation is as follows:

Condsider changing examples

1. If $c \in \mathbb{R}$ and $f(x) = o(g(x))$, then $cf(x) = o(g(x))$.
2. If $f_1(x) = o(g_1(x))$ and $f_2(x) = o(g_2(x))$, then $f_1(x)f_2(x) = o(g_1(x)g_2(x))$.
3. If $f(x) = o(g(x))$, then $x \cdot f(x) = o(x \cdot g(x))$.
4. If $\lim_{x \to 0} g(x) = 0$, then $\dfrac{1}{1 + g(x)} = 1 - g(x) + o(g(x))$.
5. $o(f(x) + g(x)) = o(f(x)) + o(g(x))$
6. $o(o(f(x))) = o(f(x))$

All of the above apply to "Big-Oh" notation as well.

Both notations are commonly used in expressing series so that one can keep track of the order of an approximation. For example, if we expand e^x about $x = 0$ but are only concerned with second-order terms and below, we can write it as

$$e^x = 1 + x + \frac{x^2}{2!} + o(x^2) = 1 + x + \frac{x^2}{2!} + O(x^3)$$

2.3 Error of Series Approximations

In general, the error for a finite *power* series can be expressed as

$$R_N(x) = f(x) - \left(f(x) + (x - a)f'(a) + \cdots + (x - a)^N \frac{f^{(N)}(a)}{N!} \right) \tag{2.3.1}$$

$$= \frac{(x - a)^{N+1} f^{(N+1)}(c)}{(N + 1)!} \tag{2.3.2}$$

for some $c \in [a, x]$. This is known as *Taylor's Theorem*. However, it is not of much practical use, as we do not have a formula for c.

However, useful for error bounds

In the case where the power series coefficients are *decreasing* and *non-negative*, on the interval $x \in (-1, 1)$ we can use the formula

$$R_N(x) = \frac{a_{N+1} x^N}{1 - |x|} \tag{2.3.3}$$

In the case where the (not necessarily power) series is *alternating*, and absolutely decreasing (that is, $|a_{n+1}| < |a_n|$ for all n) and $\lim_{n\to\infty} a_n = 0$, then

$$|R_N| \leq |a_{N+1}|. \tag{2.3.4}$$

3 Linear Algebra

3.1 Notation

- A *vector* in a vector space is denoted as $\mathbf{x}$.
- The *coordinate vector* of $\mathbf{x}$ with respect to a basis $\mathcal{B}$ is denoted as $[\mathbf{x}]_{\mathcal{B}}$. If there is no subscript, e.g. $[\mathbf{x}]$, the basis is understood to be the *standard basis*.
- The *standard basis elements* are denoted $\mathbf{e}_1, \mathbf{e}_2, \ldots \mathbf{e}_n$.

3.2 Inner Product Spaces

Definition 3.1. An *inner product* is a billinear map that takes two vectors of a vector space and returns a scalar of a field F (where in these notes, $F = \mathbb{R}$ or $F = \mathbb{C}$). In more symbolic notation, we have $\langle \cdot, \cdot \rangle : V \times V \to F$. A billinear map qualifies as an inner product if all of the following conditions are met[a]:

1. $\langle \mathbf{x} | \mathbf{x} \rangle \geq 0$ and $\langle \mathbf{x} | \mathbf{x} \rangle = 0 \iff \mathbf{x} = \mathbf{0}$
2. $\langle a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2 | \mathbf{y} \rangle = a_1^* \langle \mathbf{x}_1 | \mathbf{y} \rangle + a_2^* \langle \mathbf{x}_2 | \mathbf{y} \rangle$
3. $\langle \mathbf{x} | b_1 \mathbf{y}_1 + b_2 \mathbf{y}_2 \rangle = b_1 \langle \mathbf{x} | \mathbf{y}_1 \rangle + b_2 \langle \mathbf{x} | \mathbf{y}_2 \rangle$
4. $\langle \mathbf{x} | \mathbf{y} \rangle = \langle \mathbf{y} | \mathbf{x} \rangle^*$

[a] Where z^* is the complex conjugate of z, also written as $\bar{z}$. For more information, see Section 5.1.

Definition 3.2. The *norm* of a vector $\mathbf{x}$ is $\|\mathbf{x}\| = \sqrt{\langle \mathbf{x} | \mathbf{x} \rangle}$.

From these we get a few theorems, which we will state here without proof.

First we have the *Cauchy-Schwarz Inequality*:

$$\langle \mathbf{p}|\mathbf{q}\rangle \leq \|\mathbf{p}\| \cdot \|\mathbf{q}\|. \tag{3.2.1}$$

Next we have the *Triangle Inequality*:

$$\|\mathbf{p}+\mathbf{q}\| \leq \|\mathbf{p}\| + \|\mathbf{q}\|. \tag{3.2.2}$$

Lastly we have the *Pythagorean theorem*:

$$\langle \mathbf{p}|\mathbf{q}\rangle = 0 \iff \|\mathbf{p}\|^2 + \|\mathbf{q}\|^2 = \|\mathbf{p}+\mathbf{q}\|^2. \tag{3.2.3}$$

For our purposes, there are only four important inner products to concern ourselves with. When we're dealing with vectors in $\mathbb{R}^n$, we will use the familiar *dot product*; if $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, then

$$\langle \mathbf{x}|\mathbf{y}\rangle = \mathbf{x}^T\mathbf{y} = \sum_{i=1}^{n} x_i y_i \tag{3.2.4}$$

With vectors $\mathbf{x}, \mathbf{y} \in \mathbb{C}^n$ we have a similarly defined inner product defined thusly:

$$\langle \mathbf{x}|\mathbf{y}\rangle = \mathbf{x}^\dagger\mathbf{y} = \sum_{i=1}^{n} x_i^* y_i \tag{3.2.5}$$

For functions/vectors in $L^2(X)$ meaning[2] the set of functions $\{f_i\}$ that satisfy

$$\int_X |f_i(x)|^2\, dx < \infty$$

3.3 Coordinates and Change of Bases

The coordinate mapping of a transformed vector $T(\mathbf{x})$ is given simply as

$$[T(\mathbf{x})] = \begin{bmatrix} [T(\mathbf{e}_1)] & [T(\mathbf{e}_2)] & \cdots & [T(\mathbf{e}_n)] \end{bmatrix} [\mathbf{x}] = [T][\mathbf{x}]. \tag{3.3.1}$$

In a basis $\mathcal{B}$ with basis $\{\mathbf{b}_i\}$, the transformation can be written

$$[T(\mathbf{x})]_\mathcal{B} = \begin{bmatrix} [T(\mathbf{b}_1)]_\mathcal{B} & [T(\mathbf{b}_2)]_\mathcal{B} & \cdots & [T(\mathbf{b}_n)]_\mathcal{B} \end{bmatrix} [\mathbf{x}]_\mathcal{B} = [T]_\mathcal{B}[\mathbf{x}]_\mathcal{B}. \tag{3.3.2}$$

To switch between these two transformation matrices, we have

$$[T] = B[T]_\mathcal{B}B^{-1} \iff [T]_\mathcal{B} = B^{-1}[T]B \tag{3.3.3}$$

where

$$B = \begin{bmatrix} [\mathbf{b}_1] & [\mathbf{b}_2] & \cdots & [\mathbf{b}_n] \end{bmatrix}. \tag{3.3.4}$$

From the above, we can see

$$[\mathbf{x}] = B[\mathbf{x}]_\mathcal{B} \iff [\mathbf{x}]_\mathcal{B} = B^{-1}[\mathbf{x}]. \tag{3.3.5}$$

[2] where X is most commonly $[a, b]$ or $\mathbb{R}$

3.4 Diagonalization

Definition 3.3. To *diagonalize* a matrix A is to express it as $A = CDC^{-1}$, where C is an invertible matrix and D is a diagonal matrix.

TODO: How to diagonalize a matrix

Definition 3.4. An *orthogonal matrix* is a real matrix which has the following (mutually equivalent) properties:

1. All its columns (if treated as column vectors) are mutually orthogonal.
2. Its transpose equals its inverse, e.g. if Q is an orthogonal matrix, $Q^T = Q^{-1}$.

Theorem 3.1. If a matrix A (with real entries) is *symmetric* (i.e. $A = A^T$), then it also has the following properties:

1. It is diagonalizable as $A = QDQ^T$, where D is a diagonal matrix and Q is a orthogonal matrix.
2. Its basis is orthogonal.
3. All its eigenvalues are real.

Definition 3.5. The *adjoint* of a matrix A is $A^\dagger = (A^T)^*$. This is pronounced "A dagger." In physics, this is commonly known as the *Hermitian conjugate* of A.

Definition 3.6. An *unitary matrix* is a complex matrix which has the following (mutually equivalent) properties:

1. All its columns (if treated as column vectors) are mutually orthogonal.
2. Its adjoint equals its inverse, e.g. if Q is a unitary matrix, $Q^\dagger = Q^{-1}$.

Theorem 3.2. If A is *Hermitian* (i.e. $A = A^\dagger$), then it also has the following properties:

1. It is diagonalizable as $A = UDU^\dagger$, where D is a diagonal matrix and U is

a unitary matrix.

2. Its basis is orthogonal.
3. All its eigenvalues are real.

4 Partial Differentiation

Definition 4.1. A *partial derivative* is the multivariate analogue of a traditional derivative.

$$\frac{\partial f}{\partial x_i} = \lim_{\Delta x_i \to 0} \frac{f(x_1, \cdots, x_i + \Delta x_i, \cdots, x_n) - f(x_1, \cdots, x_i, \cdots, x_n)}{\Delta x_i}$$

What this means in a practical context is that when taking a partial derivative, we take all other variables to be constant. E.g. $f(x, y) = 3x^2 \cos(y)$

$$\frac{\partial f}{\partial x} = 6x \cos(y), \quad \frac{\partial f}{\partial y} = -3x^2 \sin(y)$$

A function is *differentiable* at a point if it is well approximated by a linear function in the neighborhood of that point.

Let's say we have a function $z = f(x, y)$. We can then approximate small changes in z like so:

$$\Delta z = f(x_0 + \Delta x, y_0 + \Delta y) - f(x_0, y_0) = \frac{\partial f}{\partial x}\Delta x + \frac{\partial f}{\partial y}\Delta y + o(\Delta x) + o(\Delta y)$$

Taking the limit of the above, we get the total differential

$$dz = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy \tag{4.0.1}$$

If we have more than 2 variables of concern in a problem, e.g. $z = f(x, y)$ and $z = g(x, \theta)$, then we can use a subscript to denote what variable(s) we are holding constant,

$$\left(\frac{\partial z}{\partial x}\right)_\theta = \frac{\partial g}{\partial x}, \quad \left(\frac{\partial z}{\partial x}\right)_y = \frac{\partial f}{\partial x}$$

4.1 Implicit Partial Differentiation

Careful manipulation of total differentials can yield partial derivatives for complicated equations. For example, given $x^2 + y^2 + z^2 = 1$, what is $\left(\frac{\partial z}{\partial x}\right)_y$? We can manipulate differentials as if they were algebraic quantities to generate this derivative, but we must understand that "under the hood," we are really manipulating these "Δx" type quantities (which are *not* symbolic in the same manner as differentials) and then subsequently taking the *limit* as described in the definition of the partial derivative.

If we consider this equation $x^2 + y^2 + z^2 = 1$ a constant function, that is, $F(x, y, z) = x^2 + y^2 + z^2 = 1$, then

$$dF = 2x\,dx + 2y\,dy + 2z\,dz$$

Now, we are holding y constant, so $\Delta y = 0$, thus we are interested in

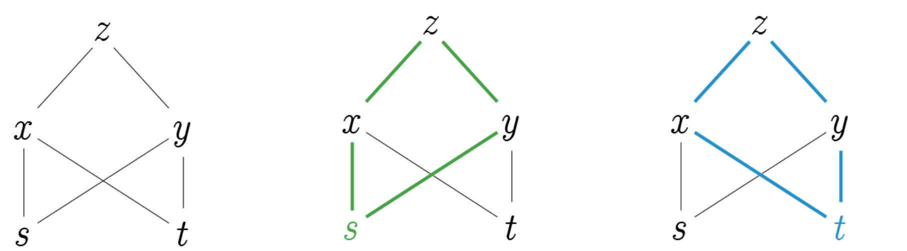
$$dF_y = 2x\,dx + 2z\,dz$$

where the subscript y denotes that we are taking y as constant. Since F is a constant function, its differential must be zero. Thus,

$$\begin{aligned} 2x\,dx + 2z\,dz &= 0 \\ 2x\,dx &= -2z\,dz \\ \left(\frac{\partial z}{\partial x}\right)_y &= -\frac{x}{z} \end{aligned}$$

4.2 The Chain Rule

Figure 1: The chain rule, visualized



(a) Interdependancies of variables in example

(b) "Connections" used when finding $\left(\frac{\partial z}{\partial s}\right)_t$

(c) "Connections" used when finding $\left(\frac{\partial z}{\partial t}\right)_s$

Just as in single variable calculus, when we differentiate compositions of two or more functions, we have to use the chain rule. However, the process is slightly more complicated than in the single-variable case. Again, we are going to utilize the *total differential*. While we *can* use a formula like we do in the single-variable case, it is fairly cumbersome, difficult to remember, and not as illuminating as the alternative.

Suppose we have $z = z(x, y)$, $x = x(s, t)$, and $y = y(s, t)$. We know from earlier that

$$dz = \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy \tag{4.2.1}$$

$$dx = \frac{\partial x}{\partial s} ds + \frac{\partial x}{\partial t} dt \tag{4.2.2}$$

$$dy = \frac{\partial y}{\partial s} ds + \frac{\partial y}{\partial t} dt \tag{4.2.3}$$

Now, suppose we are interested in $\left(\frac{\partial z}{\partial s}\right)_t$. In this case, we are holding t constant, and as a result, we may remove any term with dt. Thus,

$$dz = \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy \tag{4.2.4}$$

$$dx = \frac{\partial x}{\partial s} ds \tag{4.2.5}$$

$$dy = \frac{\partial y}{\partial s} ds \tag{4.2.6}$$

Recalling that, in truth, we are manipulating these "Δx" type quantites, we may use the principle of substitution. Thus,

$$dz_t = \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} ds + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s} ds = \left(\frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s}\right) ds \tag{4.2.7}$$

If we bring the ds to the other side (by utilizing the fact we are manipulating "Δx" type quantities), we get

$$\left(\frac{\partial z}{\partial s}\right)_t = \frac{\partial z}{\partial x}\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\frac{\partial y}{\partial s} \tag{4.2.8}$$

As you can see, the chain rule can be understood as the "chains" of dependence between variables. If you look at Figure 2b on page 13, you can see the tree-like structure of the system of variables. Each term represents a pathway down from the function of interest to the independent variable that is not fixed. Each line connecting two variables represents a derivative of the higher w.r.t. the lower.

4.3 Optimization

4.3.1 Gradient

4.3.2 Lagrange Multipliers

5 Complex Analysis

5.1 Complex Numbers

Definition 5.1. The *imaginary unit*, denoted as i, is the solution[a] to the following equation:

$$i^2 = -1$$

[a]Well, not exactly. The fact is, there are two solutions to this equation, and we are picking one. But we could have just as easily picked the other! If we replaced i with $-i$ in every math book, nothing would be incorrect (except maybe the spacing).

From this, we can generate a definition of the Complex Numbers.

Definition 5.2. The *set of complex numbers*, denoted $\mathbb{C}$, can be defined as

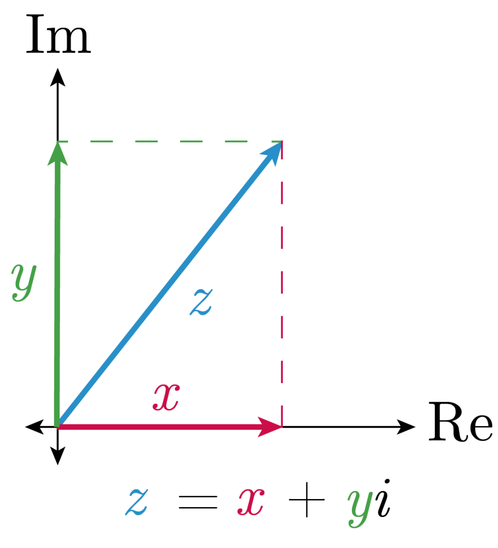
$$\mathbb{C} = \{x + iy \mid x, y \in \mathbb{R}\}$$

We also define the functions $\mathrm{Re}(x + iy) = x$ and $\mathrm{Im}(x + iy) = y$ to work with the real and imaginary parts more easily. Two complex numbers are equal if and only if their corresponding real and imaginary parts are both equal. That is,

$$z_1 = z_2 \iff \mathrm{Re}(z_1) = \mathrm{Re}(z_2) \wedge \mathrm{Im}(z_1) = \mathrm{Im}(z_2).$$

We can think of complex numbers as analogous to Euclidean vectors with the real

Figure 3: The complex number z as a Eulclidean vector.



and imaginary lines as the x and y axes. See Figure 3. Addition and subtraction are defined relatively simply enough:

$$z_1 \pm z_2 = (x_1 + iy_1) \pm (x_2 + iy_2) = (x_1 \pm x_2) + i(y_1 \pm y_2)$$

Multiplication follows what you'd expect:

$$\begin{aligned} z_1 \cdot z_2 &= (x_1 + iy_1)(x_2 + iy_2) \\ &= x_1x_2 + ix_2y_1 + ix_1y_2 + i^2y_1y_2 \\ &= (x_1x_2 - y_1y_2) + i(x_2y_1 + x_1y_2) \end{aligned}$$

For division, what we are really asking is, "given that $z \neq 0$, does z^{-1} exist $\forall z$ such that $zz^{-1} = 1$, and if so, what is it?" Reformulating the question we get, "given $z = x + iy \neq 0$, what is $z^{-1} = a + ib$ such that $(x + iy)(a + ib) = 1$?" Expanding this out gives us

$$(xa - yb) + i(ya + xb) = 1 + 0i$$

From the definition of equality for complex numbers that we saw earlier, the real and imaginary parts of two complex numbers must be equal if the the complex numbers are to be equal. Thus,

$$\begin{aligned} xa - yb &= 1 \\ ya + xb &= 0 \end{aligned}$$

Noting that this is simply a system of linear equations, we can put it into matrix form,

$$\begin{pmatrix} x & -y \\ y & x \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

Since

$$\det \begin{pmatrix} x & -y \\ y & x \end{pmatrix} = x^2 + y^2 \neq 0$$

the matrix is invertable, thus a and b exist and are unique, given by

$$\begin{aligned} \begin{pmatrix} a \\ b \end{pmatrix} &= \begin{pmatrix} x & -y \\ y & x \end{pmatrix}^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \frac{1}{x^2 + y^2} \begin{pmatrix} x & y \\ -y & x \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} \frac{x}{x^2+y^2} \\ \frac{-y}{x^2+y^2} \end{pmatrix} \end{aligned}$$

Thus, if $z \neq 0$, then

$$z^{-1} = \frac{x}{x^2 + y^2} + i\frac{-y}{x^2 + y^2}$$

Definition 5.3. The *complex conjugate* of a complex number z (or often times, simply the *conjugate* of z), denoted[a] with either a star or overbar (e.g. z^* or $\bar{z}$), is the negation of z's imaginary part. That is, if $z = x + iy$, then

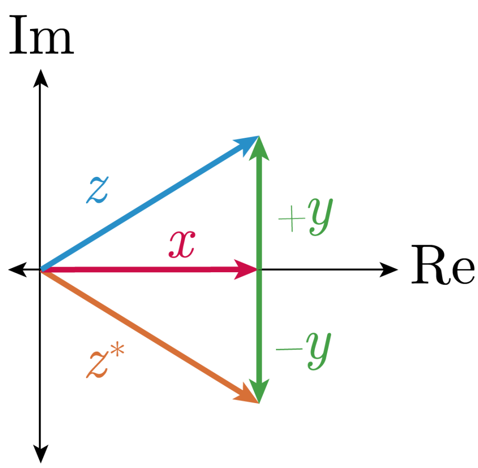
$$z^* = x - iy.$$

[a]Physicists tend to use the z^* notation, while mathematicians use the $\bar{z}$ notation.

Some important properties of the complex conjugate are as follows (assuming $z_1, z_2 \in \mathbb{C}$):

- $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$
- $\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$
- $\overline{z_1^{-1}} = (\overline{z_1})^{-1}$

Figure 4: The complex conjugate of z, visualized



One can think of complex conjugation of a reflection a complex coordinate about the real axis; see Figure 4.

Definition 5.4. The *absolute value* or *magnitude* of a complex number z, denoted $|z|$, is the "length" of the Euclidean vector from the origin to the point z in the

$\mathbb{C}$ plane. If $z = x + iy$ where $x, y \in \mathbb{R}$,

$$|z| = \sqrt{x^2 + y^2}$$

The absolute value of a complex number is a type of norm, specifically the *Euclidean norm* or the L^2 *norm*. Some useful properties of the absolute value include:

- $|z_1 z_2| = |z_1||z_2|$
- $|z^{-1}| = \frac{1}{|z|}$

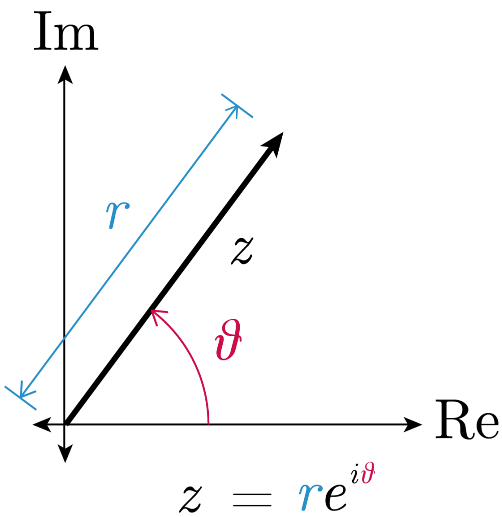
With these definitions, we can express the inverse of a complex number as

$$z^{-1} = \frac{z^*}{|z|^2} \tag{5.1.1}$$

Or, equivalently

$$zz^* = |z|^2 \tag{5.1.2}$$

Figure 5: The complex number z in polar coordinates



We can also express complex numbers with polar coordinates; see Figure 5. If we define $r = |z|$ and θ as the angle from the positive real axis to the Euclidean vector representation of z, then

$$x = r\cos\theta \tag{5.1.3}$$

$$y = r\sin\theta \tag{5.1.4}$$

Using Euler's formula, which I will simply state here as[3]

$$e^{i\theta} = \cos\theta + i\sin\theta,$$

we can also express the complex number z as $z = re^{i\theta}$. Euler's formula has many powerful results that are quite simple to derive. One of the most trivial is De Moive's Theorem.

Theorem 5.1 (De Moive's Theorem)**.**

$$\begin{aligned} e^{in\theta} &= (\cos\theta + i\sin\theta)^n \\ \cos n\theta + i\sin n\theta &= (\cos\theta + i\sin\theta)^n \end{aligned}$$

This can be used to easily generate useful identites such as

$$\cos 3\theta = \cos^3\theta - 3\cos\theta\sin^2\theta$$

This $re^{i\theta}$ notation helps to elucidate the geometry of mathematical operations in $\mathbb{C}$:

$$z_1 = r_1 e^{i\theta_1}, z_2 = r_2 e^{i\theta_2} \Rightarrow z_1 z_2 = r_1 r_2 e^{i(\theta_1+\theta_2)},$$

or in other words, when multiplying two complex numbers, we multiply their magnitudes and add their angles from Re to get the complete product. Thus, multiplying a complex number by $e^{i\theta}$ will simply rotate it (while preserving its length).[4] Additionally,

$$z = re^{i\theta} \Rightarrow z^{-1} = re^{-i\theta},$$

or in other words, when inverting complex numbers, we simply negate the angle from Re, which makes perfect sense considering that z^{-1} is just a scaled z^*.

You might note that θ (known as the *argument* of z) is not unique for a given z. For example, the points (expressed as (r, θ) coordinate pairs) $(2, \frac{\pi}{4})$ and $(2, \frac{9\pi}{4})$ refer to the same point in $\mathbb{C}$. Thus, if we tried to create a function $\arg z$ such that if $z = re^{i\theta}$, $\arg z = \theta$, we would fail, as $\arg z = \theta + 2\pi$ is also perfectly valid. However, we *can* define $\arg z$ as a "*multivalued* function."Thus, for a given $z = re^{i\theta}$, $\arg z = \theta + 2n\pi$ (where $n \in \mathbb{Z}$). If we restrict the range of $\arg z$ to $[0, 2\pi)$, it becomes an honest-to-god *function*, which we denote as $\operatorname{Arg} z$.

add notes about Riemann surfaces

The functions $\arg z$ and $\operatorname{Arg} z$ have the following properties:

- $\arg z_1 z_2 = \arg z_1 + \arg z_2$
- $\arg z^{-1} = -\arg z$

[3] If you haven't seen this before, go derive it by Taylor expanding $e^{i\theta}, \sin\theta, \cos\theta$.

[4] Note that if $in\theta = i\theta \mod 2\pi$, repeated rotations can loop back upon themselves (e.g. $\left(e^{i\frac{\pi}{2}}\right)^4 z = z$), while if $in\theta \neq i\theta \mod 2\pi$, $\left(e^{i\theta}\right)^n z \neq z$ for $n \neq 0$. (The previous assumes $n \in \mathbb{Z}$).

which fits with our geometric understanding of $\mathbb{C}$ that we developed earlier.

Roots and fractional powers are not too much more compicated when dealing with complex numbers. Just as the root of a real number may have more than one answer (e.g.$\sqrt{64} = 8\text{or} - 8$), so too may the root of a complex number be multivalued. If $z = re^{i\theta}$,

$$\begin{aligned} z^{\frac{n}{m}} =& r^{\frac{n}{m}} e^{i\frac{n\theta}{m}}, \\ & r^{\frac{n}{m}} e^{i\frac{n\theta}{m} + \frac{2\pi}{m}}, \\ & r^{\frac{n}{m}} e^{i\frac{n\theta}{m} + \frac{2\pi(2)}{m}}, \\ & \dots, \\ & r^{\frac{n}{m}} e^{i\frac{n\theta}{m} + \frac{2\pi(m-1)}{m}} \end{aligned}$$

Thus, there are m distinct values for the term $z^{\frac{n}{m}}$.[5]

5.2 Complex Series, e^z, and the Logarithm

Overall, complex series behave similarly to their real counterparts. Let's define the infinite complex series S as

$$\sum_{n=1}^{\infty} a_n = S \qquad (a_n \in \mathbb{C}) \tag{5.2.1}$$

If $a_n = x_n + iy_n$ (where $x_n, y_n \in \mathbb{R}$), consider

$$X_N = \sum_{n=1}^{N} x_n, \qquad Y_N = \sum_{n=1}^{N} y_n.$$

Let's then define

$$X = \lim_{N\to\infty} X_N, \qquad Y = \lim_{N\to\infty} Y_N.$$

Thus, S exists (i.e. the series converges) iff X and Y exist.

The concept of absolute convergence also applies to complex series. That is, if $\sum |a_n|$ converges, then $\sum a_n$ converges absolutely. One can derive this from the fact that $|x_n| \leq |a_n| \wedge |y_n| \leq |a_n|$.

Complex power series also behave similarly to real power series. The generic complex power series about a point a can be defined as

$$S(z) = \sum_{n=1}^{\infty} a_n (z - a) \qquad (a, a_n \in \mathbb{C}) \tag{5.2.2}$$

[5]Assuming rational powers, i.e. $m, n \in \mathbb{N}$.

There is some R such that for

$$\begin{aligned} |z-a| < R, &\quad S(z) \text{ converges} \\ |z-a| > R, &\quad S(z) \text{ diverges} \end{aligned}$$

This power series will converge within this *disk of convergence* and diverge outside of it.

Note that we can still use some tests to determine convergence. The most common test to use in the complex case is the ratio test, which looks identical to the ratio test described earlier.

For example, let us examine the *complex exponential function*, i.e. e^z. Now, from Taylor series, we know that we can express e^x as

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots.$$

This also holds true for complex arguments. That is, for $z \in \mathbb{C}$,

$$e^z = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \dots. \tag{5.2.3}$$

But does this converge everywhere (as in the real case)? Let us use the ratio test.

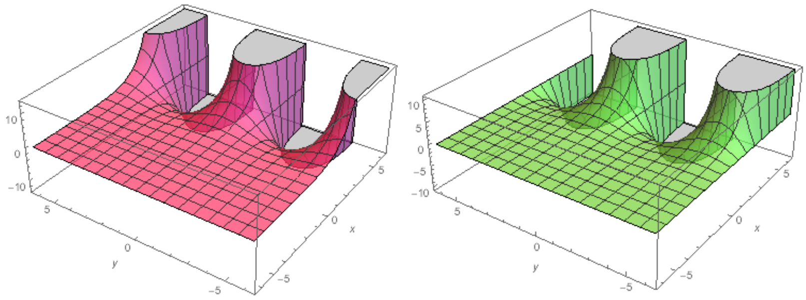
$$\begin{aligned} \rho &= \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| \\ &= \lim_{n\to\infty} \left| \frac{z^{n+1}}{(n+1)!} \cdot \frac{n!}{z^n} \right| \\ &= \lim_{n\to\infty} \left| \frac{z}{n+1} \right| \\ &= 0 \end{aligned}$$

Thus, the series converges regardless of the value of z, and therefore, converges everywhere on $\mathbb{C}$.

For the task of visualizing e^z, we are stuck with a problem. We are mapping a 2D space to another 2D space, and unfortunately, we humans cannot visualize the 4D space.[6] This issue will arise with any complex to complex functions and thus, we must be clever with how we plot them. First, we may to plot the real and imaginary parts, like in Figure 6. However, this does not yield a great deal of new insight.

Another approach we may take is restrict ourselves to important 1D subsets of $\mathbb{C}$.

[6] Elon Musk, I'm looking at you to come up with a solution to this.



(a) $\mathrm{Re}\left(e^{x+iy}\right)$ (b) $\mathrm{Im}\left(e^{x+iy}\right)$

Figure 6: The function e^z.

5.3 Complex Differentiation

For the purposes of this section, let us assume $f : \mathbb{C} \to \mathbb{C}$. Recall that the *differentiation* is the approximation of a function by a linear function at a point.

Definition 5.5. The *derivative* of f at z is

$$f'(z) = \lim_{\Delta z \to 0} \frac{f(z + \Delta z) - f(z)}{\Delta z}$$

where $\Delta z \equiv \Delta x + i \Delta y$.

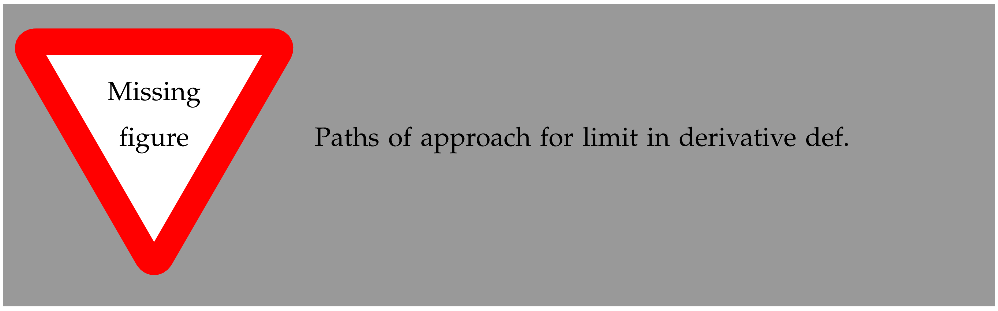
The above definition implies that

$$f(z + \Delta z) = f(z) + f'(z)\Delta z + o(\Delta z) \tag{5.3.1}$$

The usual rules apply to complex differentiation:

- $(f+g)' = f' + g'$
- $(fg)' = f'g + fg'$
- $f(g(z))' = f'(g(z))g'(z)$ (assuming derivatives exist)

Many derivatives take forms similar to their real counterparts, e.g. $(z^2)' = 2z$, $(e^z)' = e^z$.



However, not all functions are differentiable.

When we defined the limit included in the definition of the complex derivative, we appeared to treat the difference Δz identically to the difference Δx of the limit in the definition of the real derivative. However, while with real limits, we are only navigating the real number line, i.e. there are only two directions: positive or negative. With Δz, we are working with the complex plane: there are many ways for Δz to diminish to zero, as there are many ways one can approach the point z at which we are differentiating. If following different paths approach different values, the limit doesn't exist—just like how real limits don't exist if the left hand and right hand limits don't agree.

For example, consider the function $f(z) = \bar{z}$, equivalent to the map $x + iy \mapsto x - iy$:

$$\lim_{\Delta z \to 0} \frac{\overline{z + \Delta z} - \bar{z}}{\Delta z} = \lim_{\Delta z \to 0} \frac{\overline{\Delta z}}{\Delta z}$$

Let us now go through two different approaches towards the point z:

- $\Delta z \to 0$ horizontally: $\Delta z = \Delta x$, $\Delta x \to 0$

$$\lim_{\Delta x \to 0} \frac{\overline{\Delta x}}{\Delta x} = \lim_{\Delta x \to 0} \frac{\Delta x}{\Delta x} = \lim_{\Delta x \to 0} 1 = 1$$

- $\Delta z \to 0$ vertically: $\Delta z = i\Delta y$, $\Delta y \to 0$

$$\lim_{\Delta y \to 0} \frac{\overline{i\Delta y}}{i\Delta y} = \lim_{\Delta y \to 0} \frac{-i\Delta y}{i\Delta y} = \lim_{\Delta y \to 0} -1 = -1$$

These results are not equal, therefore the limit does not exist. Thus, $f(z) = \bar{z}$ is not differentiable. But why is that the case? What is it about the function $f(z) = \bar{z}$ that makes it different?

To look at this from another angle[7], consider the function F, equivalent to the map

$$\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} x \\ -y \end{bmatrix}$$

[7]Thanks are owed to the responses here.

This transformation can be written as the matrix

$$F = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

as

$$F\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ -y \end{bmatrix}.$$

Multiplication by a complex number $a + bi$ is equivalent to multiplication by a matrix

$$\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$$

as

$$(a + bi)(x + iy) = (ax - by) + i(bx + ay) \longleftrightarrow \begin{bmatrix} ax - by \\ bx + ay \end{bmatrix} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

But as we showed above, the transformation F doesn't take this form. Therefore, it cannot be expressed as a complex number. This means that we cannot express a first order approximation of f by expanding the function in a series about a complex number z_0 like

$$\begin{aligned} f(z_0 + \Delta z) &= f(z_0) + f'(z_0)\Delta z + \cdots \\ \overline{z_0 + \Delta z} &= \overline{z_0} + f'(z_0)\Delta z + \cdots \end{aligned}$$

5.3.1 Cauchy-Riemann Equations

Suppose $f : \mathbb{C} \to \mathbb{C}$, $f(x + iy) = u(x, y) + iv(x, y)$, and f is differentiable at $z = x + iy$.

By these assumptions, the limit

$$\lim_{\Delta z \to 0} \frac{f(z + \Delta z) - f(z)}{\Delta z}$$

exists for any approach towards z and its value is independent of that approach. If we examine the horizontal and vertical approaches,we derive the Cauchy-Riemann Include proof
equations:

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \tag{5.3.2}$$

$$\frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}. \tag{5.3.3}$$

Thus, satisfying the equations above is a *necessary* condition for differentiability. As it turns out, satisfying the Cauchy-Riemann equations is also a *sufficient* condition for differentiability.

A side-effect of the Cauchy-Riemann equations is that both the functions u and v must satisfy Laplace's equation:

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \tag{5.3.4}$$

$$\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0 \tag{5.3.5}$$

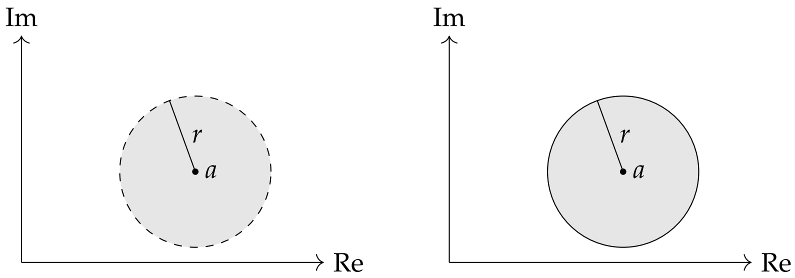
Differentiability generally depends upon the domain we are inspecting. For functions of a real variable, we generally discuss intervals. For functions of a *complex* variable, we instead use *disks*.

Definition 5.6. An *open disk* is a subset of the complex plane

$$\{z : (z - a) < r\} \subseteq \mathbb{C}$$

for some a, r. By contrast, a *closed disk* would be

$$\{z : (z - a) \leq r\} \subseteq \mathbb{C}.$$



(a) Open disk

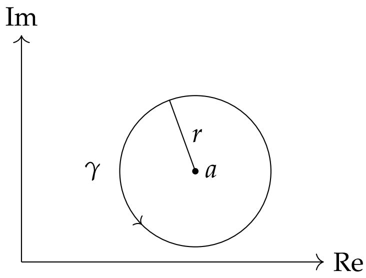
(b) Closed disk

Theorem 5.2. If f is differentiable in a disk containing z,

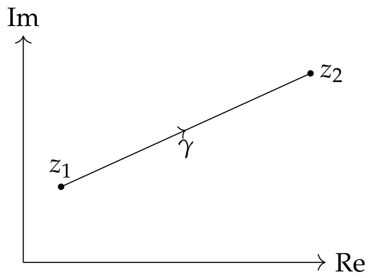
1. f is infinitely differentiable.

2. $f(z + \Delta z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z)}{n!}(\Delta z)^n$ converges absolutely in the disk, that is, the function is equal to its Taylor series, and thus, is "analytic".

The above theorem demonstrates that differentiable functions are incredibly well behaved, and that this "good behavior" is a local property.



(a) $\gamma(t) = a + re^{it}$, $(t \in [0, 2\pi])$



(b) $\gamma(t) = (1-t)z_1 + tz_2$, $(t \in [0, 1])$

Figure 8: Examples of curves

5.4 Contour Integration

For functions $f : \mathbb{R} \to \mathbb{R}$ and $f : \mathbb{R} \to \mathbb{C}$ we integrate over intervals (e.g. $[a, b]$). For functions $f : \mathbb{C} \to \mathbb{C}$, we integrate over *curves*.

Definition 5.7. A *curve* is a differentiable function $\gamma : [a, b] \to \mathbb{C}$ where $\gamma'(t) \neq 0$. If γ doesn't intersect itself[a] (except possibly at endpoints), then the curve is *simple*.

Integration along γ is simply

$$\int_\gamma f(z)\,\mathrm{d}z = \int_a^b f(\gamma(t))\gamma'(t)\,\mathrm{d}t\,,$$

noting that $\mathrm{d}z = \gamma'(t)\,\mathrm{d}t$.

[a] intersections would be cases where $\gamma(t_1) = \gamma(t_2)$ for $t_1 \neq t_2$

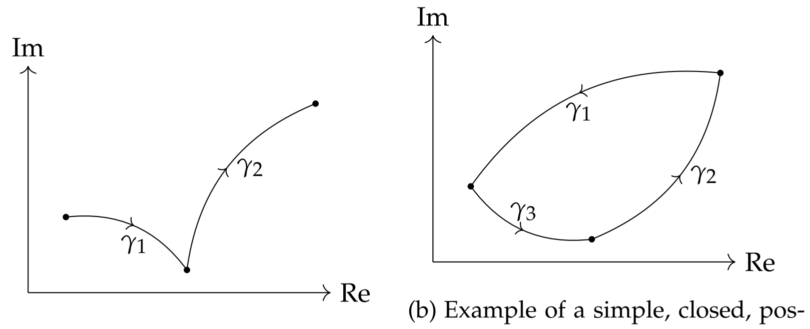
See Figure 8 for examples of curves.

Definition 5.8. A *contour* is a concatenation of a finite number of simple curves meeting at their endpoints. If it does not intersect itself, it is a *simple contour*.

Definition 5.9. A simple closed[a] contour is *positively oriented* if the interior[b] is to the *left* of the contour's direction (denoted by the arrow). Otherwise, the contour is *negatively oriented*.

[a] That is, it ends where it starts.

[b] The *interior* is the bounded region defined by the contour.



(a) Example of simple contour $C = \gamma_1 + \gamma_2$

(b) Example of a simple, closed, positively oriented contour $C = \gamma_1 + \gamma_2 + \gamma_3$

Figure 9: Examples of contours

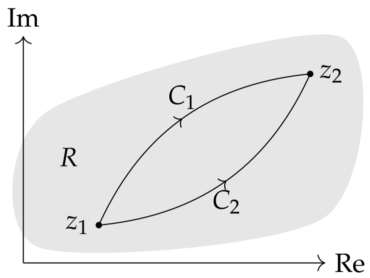


Figure 10: $\int_{z_1}^{z_2} f(z)\,\mathrm{d}z$ is the same, regardless of which path we take.

See Figure 9 for examples of contours.

Theorem 5.3 (Cauchy-Goursat Theorem)**.** If f is analytic *on* and *inside* a simple closed closed contour C, then

$$\oint_C f(z)\,\mathrm{d}z = 0.$$

In general, to prove Cauchy-Goursat, one may use

Proof. □

Proof of Cauchy-Goursat

One consequence of the Cauchy-Goursat theorem is that if f is analytic on a *simply connected* region R (that is, the region has no holes) and $z_1, z_2 \in R$, then

$$\int_{C_1} f(z)\,\mathrm{d}z = \int_{C_2} f(z)\,\mathrm{d}z$$

for any C_1, C_2 from z_1 to z_2. See Figure 10 for a visualization.

Proof. Consider $C = C_1 - C_2$. C is a closed simple contour; f is analytic on C and inside C.

$$\int_{C_1} f(z)\,\mathrm{d}z - \int_{C_2} f(z)\,\mathrm{d}z = \oint_C f(z)\,\mathrm{d}z = 0,$$

by Cauchy-Goursat. Thus,

$$\int_{C_1} f(z)\,\mathrm{d}z = \int_{C_2} f(z)\,\mathrm{d}z$$

□

5.5 Residue Calculus

6 Harmonic Analysis

6.1 Fourier Series

6.2 The Fourier Transform

6.3 The Laplace Transform