

NOTES FOR 18.04, Spring 2017

Jeremy Orloff

0 18.04 course introduction

0.1 Brief course description

Complex analysis is a beautiful, tightly integrated subject. It revolves around complex analytic functions. These are functions that have a complex derivative. Unlike calculus using real variables, the mere existence of a complex derivative has strong implications for the properties of the function.

There are a small number of far-reaching theorems that we'll explore in the first part of the class. Along the way, we'll touch on some mathematical and engineering applications of these theorems. The last third of the class will be devoted to a deeper look at applications.

The main theorems are Cauchy's Theorem, Cauchy's integral formula, and the existence of Taylor and Laurent series. Among the applications will be harmonic functions, two dimensional fluid flow, easy methods for computing (seemingly) hard integrals, Laplace transforms, and Fourier transforms with applications to engineering and physics.

0.2 Topics needed from prerequisite math classes

We will review these topics as we need them:

- Limits
- Power series
- Vector fields
- Line integrals
- Green's theorem

0.3 Level of mathematical rigor

We will make careful arguments to justify our results. Though, in many places we will allow ourselves to skip some technical details if they get in the way of understanding the main point, but we will note what was left out.

1 Complex algebra and the complex plane

We will start with a review of the basic algebra and geometry of complex numbers. Most likely you have encountered this previously in 18.03 or elsewhere.

1.1 Motivation

The equation $x^2 = -1$ has no real solutions, yet we know that this equation arises naturally and we want to use its roots. So we make up a new symbol for the roots and call it a complex number.

Definition. The symbols $\pm i$ will stand for the solutions to the equation $x^2 = -1$. We will call these new numbers **complex numbers**. We will also write

$$\sqrt{-1} = \pm i$$

Notes: 1. i is also called an **imaginary number**. This is a historical term. These are perfectly valid numbers that don't happen to lie on the real number line.

2. Our motivation for using complex numbers is not the same as the historical motivation. Historically, mathematicians were willing to say $x^2 = -1$ had no solutions. The issue that pushed them to accept complex numbers had to do with the formula for the roots of cubics. Cubics always have at least one real root, and when square roots of negative numbers appeared in this formula, even for the real roots, mathematicians were forced to take a closer look at these (seemingly) exotic objects.

We're going to look at the algebra, geometry and, most important for us, the exponentiation of complex numbers.

Before starting a systematic exposition of complex numbers, we'll work a simple example.

Example 1.1. Solve the equation $z^2 + z + 1 = 0$.

answer: We can apply the quadratic formula to get

$$z = \frac{-1 \pm \sqrt{1-4}}{2} = \frac{-1 \pm \sqrt{-3}}{2} = \frac{-1 \pm \sqrt{3}\sqrt{-1}}{2} = \frac{-1 \pm \sqrt{3}i}{2}.$$

Think: Do you know how to solve quadratic equations by completing the square? This is how the quadratic formula is derived and is well worth knowing!

1.1.1 Fundamental theorem of algebra

One of the reasons for using complex numbers is because allowing complex roots means every polynomial has exactly the expected number of roots. This is the fundamental theorem of algebra:

Fundamental theorem of algebra. A polynomial of degree n has exactly n complex roots (repeated roots are counted with multiplicity).

In a few weeks, we will be able to prove this theorem as a remarkably simple consequence of one of our main theorems.

1.2 Terminology and basic arithmetic

Definitions.

- **Complex numbers** are defined as the set of all numbers

$$z = x + yi,$$

where x and y are real numbers.

- We denote the set of all complex numbers by **C**. (On the blackboard we will usually write \mathbb{C} –this font is called *blackboard bold*.)
- We call x the **real part** of z . This is denoted by $x = \operatorname{Re}(z)$.
- We call y the **imaginary part** of z . This is denoted by $y = \operatorname{Im}(z)$.

Note well: The imaginary part of z is a real number. It **DOES NOT** include the i .

Note. Engineers typically use j instead of i . We'll follow mathematical custom in 18.04.

The basic arithmetic operations follow the standard rules. All you have to remember is that $i^2 = -1$. We will go through these quickly using some simple examples. It almost goes without saying that in 18.04 it is essential that you become fluent with these manipulations.

- **Addition:** $(3 + 4i) + (7 + 11i) = 10 + 15i$
- **Subtraction:** $(3 + 4i) - (7 + 11i) = -4 - 7i$
- **Multiplication:** $(3 + 4i)(7 + 11i) = 21 + 28i + 33i + 44i^2 = -23 + 61i$. Here we have used the fact that $44i^2 = -44$.

Before talking about division and absolute value we introduce a new operation called conjugation. It will prove useful to have a name and symbol for this, since we will use it frequently.

Complex conjugation is denoted with a bar and defined by

$$\overline{x + iy} = x - iy.$$

If $z = x + iy$ then its conjugate is $\bar{z} = x - iy$ and we read this as “z-bar = $x - iy$ ”.

Example 1.2. $\overline{3 + 5i} = 3 - 5i$.

The following is a very useful property of conjugation. We will use it in the next example to help with division.

Useful property of conjugation: If $z = x + iy$ then $z\bar{z} = (x + iy)(x - iy) = x^2 + y^2$ (is real).

Example 1.3. (Division.) Write $\frac{3 + 4i}{1 + 2i}$ in the standard form $x + iy$.

answer: We use the useful property of conjugation to clear the denominator:

$$\frac{3 + 4i}{1 + 2i} = \frac{3 + 4i}{1 + 2i} \cdot \frac{1 - 2i}{1 - 2i} = \frac{11 - 2i}{5} = \frac{11}{5} - \frac{2}{5}i.$$

In the next section we will discuss the geometry of complex numbers, which give some insight into the meaning of the magnitude of a complex number. For now we just give the definition.

Definition. The **magnitude** of the complex number $x + iy$ is defined as

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

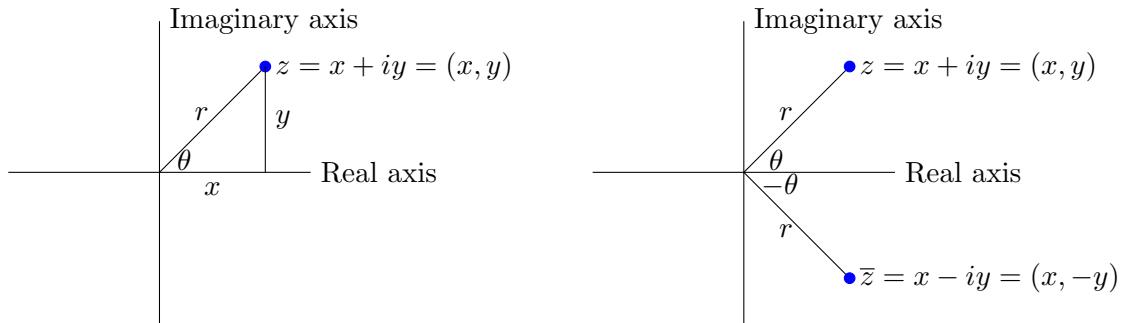
The magnitude is also called the **absolute value**, **norm** or **modulus**.

Example 1.4. The norm of $3 + 5i = \sqrt{9 + 25} = \sqrt{34}$.

Note this really well. The norm is the sum of x^2 and y^2 **it does not include the i !** **Therefore it is always positive.**

1.3 The complex plane and the geometry of complex numbers

Because it takes two numbers x and y to describe the complex number $z = x + iy$ we can visualize complex numbers as points in the xy -plane. When we do this we call it the **complex plane**. Since x is the real part of z we call the x -axis the **real axis**. Likewise, the y -axis is the **imaginary axis**.



1.4 Polar coordinates

In the figures above we have marked the length r and polar angle θ of the vector from the origin to the point $z = x + iy$. These are the same polar coordinates you saw in 18.02 and 18.03. There are a number of synonyms for both r and θ

$$r = |z| = \text{magnitude} = \text{length} = \text{norm} = \text{absolute value} = \text{modulus}$$

$$\theta = \arg(z) = \text{argument of } z = \text{polar angle of } z$$

As in 18.02 you should be able to visualize polar coordinates by thinking about the distance r from the origin and the angle θ with the x -axis.

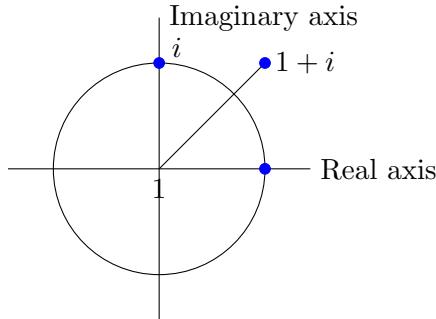
Example 1.5. In this example we make a table of z , r and θ for some complex numbers. Notice that θ is not uniquely defined since we can always add a multiple of 2π to θ and still be at the same point in the plane.

$$z = a + bi \quad r \quad \theta$$

$$1 \quad 1 \quad 0, 2\pi, 4\pi, \dots \quad \text{Argument} = 0, \text{ means } z \text{ is along the } x\text{-axis}$$

$$i \quad 1 \quad \pi/2, \pi/2 + 2\pi, \dots \quad \text{Argument} = \pi/2, \text{ means } z \text{ is along the } y\text{-axis}$$

$$1+i \quad \sqrt{2} \quad \pi/4, \pi/4 + 2\pi, \dots \quad \text{Argument} = \pi/4, \text{ means } z \text{ is along the ray at } 45^\circ \text{ to the } x\text{-axis}$$



When we want to be clear which value of θ is meant, we will specify a **branch** of \arg . For example, $0 \leq \theta < 2\pi$ or $-\pi < \theta \leq \pi$. This will be discussed in much more detail in the coming weeks. Keeping careful track of the branches of \arg will turn out to be one of the key requirements of complex analysis.

1.5 Euler's Formula

Euler's (pronounced 'oilers') formula connects complex exponentials, polar coordinates, and sines and cosines. It turns messy trig identities into tidy rules for exponentials. We will use it a lot.

The formula is the following:

$$e^{i\theta} = \cos(\theta) + i \sin(\theta). \quad (1)$$

There are many ways to approach Euler's formula. Our approach is to simply take Equation 1 as the definition of complex exponentials. This is legal, but does not show that it's a good definition. To do that we need to show the $e^{i\theta}$ obeys all the rules we expect of an exponential. To do that we go systematically through the properties of exponentials and check that they hold for complex exponentials.

1.5.1 $e^{i\theta}$ behaves like a true exponential

1. e^{it} differentiates as expected: $\frac{de^{it}}{dt} = ie^{it}$.

Proof. This follows directly from the definition:

$$\frac{de^{it}}{dt} = \frac{d}{dt}(\cos(t) + i \sin(t)) = -\sin(t) + i \cos(t) = i(\cos(t) + i \sin(t)) = ie^{it}. \quad \text{QED}$$

2. $e^{i \cdot 0} = 1$.

Proof. $e^{i \cdot 0} = \cos(0) + i \sin(0) = 1$. QED

3. The usual rules of exponents hold: $e^{ia} e^{ib} = e^{i(a+b)}$.

Proof. This relies on the cosine and sine addition formulas.

$$\begin{aligned} e^{ia} \cdot e^{ib} &= \cos(a) + i \sin(b) \cdot (\cos(a) + i \sin(b)) \\ &= \cos(a) \cos(b) - \sin(a) \sin(b) + i(\cos(a) \sin(b) + \sin(a) \cos(b)) \\ &= \cos(a+b) + i \sin(a+b) = e^{i(a+b)}. \end{aligned}$$

4. The definition of $e^{i\theta}$ is consistent with the power series for e^x .

Proof. To see this we have to recall the power series for e^x , $\cos(x)$ and $\sin(x)$. They are

$$\begin{aligned} e^x &= 1 + x + x^2/2! + x^3/3! + x^4/4! + \dots \\ \cos(x) &= 1 - x^2/2! + x^4/4 - x^6/6! + \dots \\ \sin(x) &= x - x^3/3! + x^5/5! + \dots \end{aligned}$$

Now we can write the power series for $e^{i\theta}$ and then split it into the power series for sine and cosine:

$$\begin{aligned} e^{i\theta} &= \sum_0^\infty \frac{(i\theta)^n}{n!} \\ &= \sum_0^\infty (-1)^k \frac{\theta^{2k}}{(2k)!} + i \sum_0^\infty (-1)^k \frac{\theta^{2k+1}}{(2k+1)!} \\ &= \cos(\theta) + i \sin(\theta). \end{aligned}$$

So the Euler formula definition is consistent with the usual power series for e^z .

1-4 should convince you that $e^{i\theta}$ behaves like an exponential.

1.5.2 Complex exponentials and polar form

Now let's turn to the relation between polar coordinates and complex exponentials.

Suppose $z = x + iy$ has polar coordinates r and θ . That is, we have $x = r \cos(\theta)$ and $y = r \sin(\theta)$. Thus, we get the important relationship

$$z = x + iy = r \cos(\theta) + ir \sin(\theta) = r(\cos(\theta) + i \sin(\theta)) = re^{i\theta}.$$

This is so important you shouldn't proceed without understanding. We also record it without the intermediate equation.

$$z = x + iy = re^{i\theta}. \quad (2)$$

Because r and θ are the polar coordinates of (x, y) we call $z = re^{i\theta}$ the **polar form of z** .

Magnitude, argument, conjugate, multiplication and division are easy in polar form.

Magnitude. $|e^{i\theta}| = 1$.

Proof. $|e^{i\theta}| = |\cos(\theta) + i \sin(\theta)| = \sqrt{\cos^2(\theta) + \sin^2(\theta)} = 1$.

In words, this says that $e^{i\theta}$ is always on the unit circle –this is useful to remember!

Likewise, if $z = re^{i\theta}$ then $|z| = r$. You can calculate this, but it should be clear from the definitions: $|z|$ is the distance from z to the origin, which is exactly the same definition as for r .

Argument. If $z = re^{i\theta}$ then $\arg(z) = \theta$.

Proof. This is again the definition: the argument is the polar angle θ .

Conjugate. $\overline{(re^{i\theta})} = re^{-i\theta}$.

Proof. $\overline{(re^{i\theta})} = \overline{(r(\cos(\theta) + i \sin(\theta)))} = r(\cos(\theta) - i \sin(\theta)) = re^{-i\theta}$.

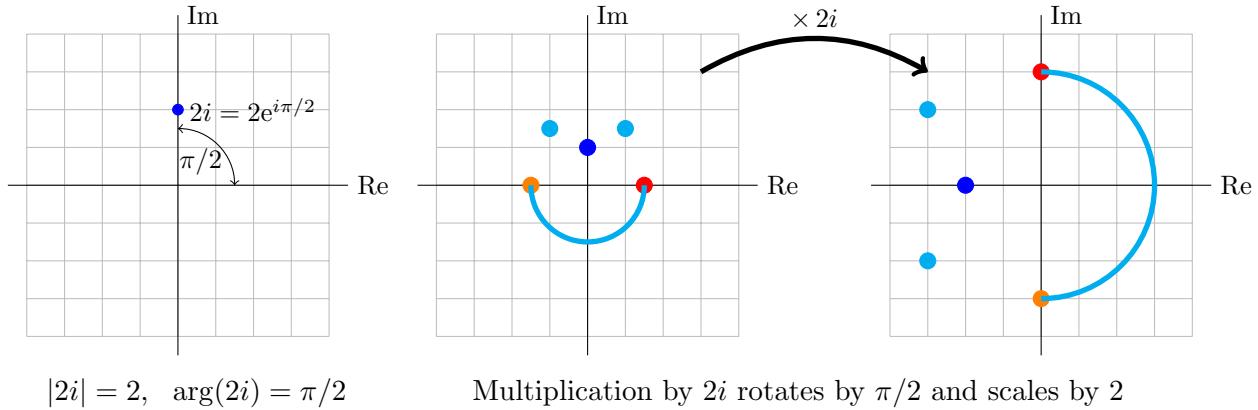
In words: complex conjugation changes the sign of the argument.

Multiplication. If $z_1 = r_1 e^{i\theta_1}$ and $z_2 = r_2 e^{i\theta_2}$ then $z_1 z_2 = r_1 r_2 e^{i(\theta_1 + \theta_2)}$.

This is what mathematicians call trivial to see, just write the multiplication down. In words, the formula says the for $z_1 z_2$ the magnitudes multiply and the arguments add.

Division. Again it's trivial that $\frac{r_1 e^{i\theta_1}}{r_2 e^{i\theta_2}} = \frac{r_1}{r_2} e^{i(\theta_1 - \theta_2)}$.

Example 1.6. Multiplication by $2i$. Here's a simple but important example. By looking at the graph we see that the number $2i$ has magnitude 2 and argument $\pi/2$. So in polar coordinates it equals $2e^{i\pi/2}$. This means that multiplication by $2i$ multiplies lengths by 2 and adds $\pi/2$ to arguments, i.e. rotates by 90° . The effect is shown in the figures below



$$|2i| = 2, \quad \arg(2i) = \pi/2$$

Multiplication by $2i$ rotates by $\pi/2$ and scales by 2

Example 1.7. Raising to a power. Compute (i) $(1+i)^6$; (ii) $\left(\frac{1+i\sqrt{3}}{2}\right)^3$

answer: (i) $1+i$ has magnitude $= \sqrt{2}$ and $\arg = \pi/4$, so $1+i = \sqrt{2}e^{i\pi/4}$. Raising to a power is now easy:

$$(1+i)^6 = \left(\sqrt{2}e^{i\pi/4}\right)^6 = 8e^{6i\pi/4} = 8e^{3i\pi/2} = -8i.$$

$$\text{(ii)} \quad \frac{1+i\sqrt{3}}{2} = e^{i\pi/3}, \text{ so } \left(\frac{1+i\sqrt{3}}{2}\right)^3 = (1 \cdot e^{i\pi/3})^3 = e^{i\pi} = -1$$

Example 1.8. Complexification or complex replacement. In this example we illustrate the technique of complex replacement to simplify a trigonometric integral. This technique will come in handy when we need to compute certain integrals.

Problem. Compute $I = \int e^x \cos(2x) dx$.

answer: We have Euler's formula $e^{2ix} = \cos(2x) + i \sin(2x)$, so $\cos(2x) = \operatorname{Re}(e^{2ix})$. The complex replacement trick is to replace $\cos(2x)$ by e^{2ix} . We get (justification below)

$$I_c = \int e^x \cos 2x + ie^x \sin 2x dx, \quad I = \operatorname{Re}(I_c).$$

Computing I_c is straightforward:

$$I_c = \int e^x e^{i2x} dx = \int e^{x(1+2i)} dx = \frac{e^{x(1+2i)}}{1+2i}.$$

Using rectangular coordinates. Here we will do the computation in rectangular coordinates. In applications, for example throughout 18.03, polar form is preferred because it is easier and gives the answer in a more useable form.

$$\begin{aligned} I_c &= \frac{e^{x(1+2i)}}{1+2i} \cdot \frac{1-2i}{1-2i} = \frac{e^x(\cos(2x) + i \sin(2x))(1-2i)}{5} \\ &= \frac{1}{5} e^x (\cos(2x) + 2 \sin(2x) + i(-2 \cos(2x) + \sin(2x))) \end{aligned}$$

So, $I = \operatorname{Re}(I_c) = \frac{1}{5} e^x (\cos(2x) + 2 \sin(2x))$.

Justification of complex replacement. The trick comes by cleverly adding a new integral to I as follows. Let $J = \int e^x \sin(2x) dx$. Then we let

$$I_c = I + iJ = \int e^x (\cos(2x) + i \sin(2x)) dx = \int e^x e^{2ix} dx.$$

Clearly, by construction, $\operatorname{Re}(I_c) = I$ as claimed above.

Alternative using polar coordinates to simplify the expression for I_c :

Write $1+2i = re^{i\phi}$, where $r = \sqrt{5}$ and $\phi = \arg(1+2i) = \tan^{-1}(2)$ in the first quadrant. Then:

$$I_c = \frac{e^{x(1+2i)}}{\sqrt{5}e^{i\phi}} = \frac{e^x}{\sqrt{5}} e^{i(2x-\phi)} = \frac{e^x}{\sqrt{5}} (\cos(2x-\phi) + i \sin(2x-\phi)).$$

Thus, $I = \operatorname{Re}(I_c) = \frac{e^x}{\sqrt{5}} \cos(2x-\phi)$.

1.5.3 Nth roots

We are going to need to be able to find the n th roots of complex numbers. The trick is to recall that a complex number has more than one argument, that is we can always add a multiple of 2π . For example,

$$2 = 2e^{0i} = 2e^{2\pi i} = 2e^{4\pi i} \dots = 2e^{2n\pi i}$$

Example 1.9. Find all 5 fifth roots of 2.

answer: In polar form: $(2e^{2n\pi i})^{1/5} = 2^{1/5}e^{2n\pi i/5}$. So the fifth roots of 2 are

$$2^{1/5} = 2^{1/5}e^{2n\pi i/5}, \text{ where } n = 0, 1, 2, \dots$$

The notation is a little strange, because the $2^{1/5}$ on the left side of the equation means the complex roots and the $2^{1/5}$ on the right hand side is a magnitude, so it is the positive real root.

Looking at the left hand side we see that for $n = 5$ we have $2^{1/5}e^{2\pi i}$ which is exactly the same as the root when $n = 0$, i.e. $2^{1/5}e^{0i}$. Likewise $n = 6$ gives exactly the same root as $n = 1$. So, we have 5 different roots corresponding to $n = 0, 1, 2, 3, 4$.

$$2^{1/5} = 2^{1/5}, 2^{1/5}e^{2\pi i/5}, 2^{1/5}e^{4\pi i/5}, 2^{1/5}e^{6\pi i/5}, 2^{1/5}e^{8\pi i/5}.$$

Similarly we can say that in general $z = re^{i\theta}$ has N different N th roots:

$$z^{1/N} = r^{1/N} e^{i\theta/N + i2\pi(n/N)} \text{ for } 0, 1, 2, \dots, N-1.$$

Example 1.10. Find the 4 fourth roots of 1.

answer: $1 = e^{i2\pi n}$, so $1^{1/4} = e^{i2\pi(n/4)}$. So the 4 different fourth roots are $1, e^{i\pi/2}, e^{i\pi}, e^{i3\pi/2}, e^{i2\pi}$.

When the angles are ones we know about, e.g. 30, 60, 90, 45, etc., we should simplify the complex exponentials. In this case, the roots are $1, i, -1, -i$.

Example 1.11. Find the 3 cube roots of -1.

answer: $-1 = e^{i\pi+i2\pi n}$. So, $(-1)^{1/3} = e^{i\pi/3+i2\pi(n/3)}$ and the 3 cube roots are $e^{i\pi/3}, e^{i\pi}, e^{i5\pi/3}$.

Since $\pi/3$ radians is 60° we can simplify:

$$e^{i\pi/3} = \cos(\pi/3) + i \sin(\pi/3) = \frac{1}{2} + i \frac{\sqrt{3}}{2} \Rightarrow (-1)^{1/3} = -1, \frac{1}{2} \pm \frac{\sqrt{3}}{2}.$$

Example 1.12. Find the 5 fifth roots of $1+i$.

answer: $1+i = \sqrt{2}e^{i(\pi/4+2n\pi)}$, for $n = 0, 1, 2, \dots$. So, the 5 fifth roots are $(1+i)^{1/5} = 2^{1/10}e^{i\pi/20}, 2^{1/10}e^{i9\pi/20}, 2^{1/10}e^{i17\pi/20}, 2^{1/10}e^{i25\pi/20}, 2^{1/10}e^{i33\pi/20}$.

Using a calculator we could write these numerically as $a+bi$, but there is no easy simplification.

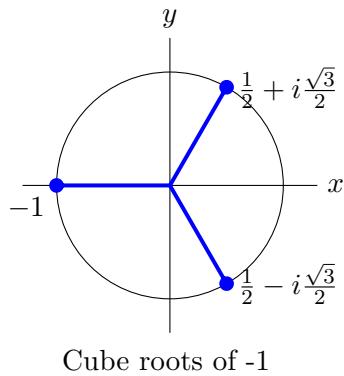
Example 1.13. We should check that our technique works as expected for a simple problem. Find the 2 square roots of 4.

answer: $4 = 4e^{i2\pi n}$. So, $4^{1/2} = 2e^{i\pi n}$. So the 2 square roots are $= 2e^0, 2e^{i\pi} = \pm 2$ as expected!

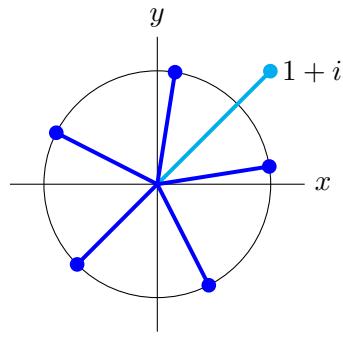
1.5.4 The geometry of n th roots

Looking at the examples above we see that roots are always spaced evenly around a circle centered at the origin. For example, the fifth roots of $1+i$ are spaced at increments of $2\pi/5$ radians around the circle of radius $2^{1/5}$.

Note also that the roots of real numbers always come in conjugate pairs.



Cube roots of -1



Fifth roots of 1+i

1.6 Inverse Euler Formula

Euler's formula gives a complex exponential in terms of sines and cosines. We can turn this around to get the [inverse Euler formulas](#).

Euler's formula says:

$$e^{it} = \cos(t) + i \sin(t) \quad \text{and} \quad e^{-it} = \cos(t) - i \sin(t).$$

By adding and subtracting we get:

$$\cos(t) = \frac{e^{it} + e^{-it}}{2} \quad \text{and} \quad \sin(t) = \frac{e^{it} - e^{-it}}{2i}.$$

Please take note of these formulas we will use them frequently!

1.7 Representing complex multiplication as matrix multiplication

Consider the product

$$w = (a + bi)(x + iy) = (ax - by) + i(bx + ay)$$

We can think of points in the complex plane $z = x + iy$ as two-dimensional vectors in \mathbf{R}^2 , e.g.

$$z = x + iy \equiv \begin{bmatrix} x \\ y \end{bmatrix}, \quad w = (ax - by) + i(bx + ay) \equiv \begin{bmatrix} ax - by \\ bx + ay \end{bmatrix}$$

It's easy to see that as vectors we have

$$\begin{bmatrix} ax - by \\ bx + ay \end{bmatrix} = \begin{bmatrix} a & -b \\ b & a \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

That is, in this sense, multiplication by $a + ib$ is the same as multiplication by the matrix $\begin{bmatrix} a & -b \\ b & a \end{bmatrix}$. We will not make a lot of use of this fact, but later it will help us remember certain formulas and facts.

Example 1.14. Multiplication by i is represented by the matrix $\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$.

1.8 de Moivre's formula

de Moivre's formula: For positive integers n :

$$(\cos(\theta) + i \sin(\theta))^n = \cos(n\theta) + i \sin(n\theta)$$

Proof. This is a simple consequence of Euler's formula:

$$(\cos(\theta) + i \sin(\theta))^n = (e^{i\theta})^n = e^{in\theta} = \cos(n\theta) + i \sin(n\theta) \quad \text{QED}$$

The reason this simple fact has a name is that historically de Moivre stated it before Euler's formula was known. Without Euler's formula there is not such a simple proof.

1.9 The exponential function

We have Euler's formula: $e^{i\theta} = \cos(\theta) + i\sin(\theta)$. We can extend this to the complex exponential function e^z .

Definition. For $z = x + iy$ the **complex exponential function** is defined as

$$e^z = e^{x+iy} = e^x e^{iy} = e^x (\cos(y) + i\sin(y)).$$

In this definition e^x is the usual exponential function for a real variable x .

It is easy to see that all the usual rules of exponents hold:

1. $e^0 = 1$
2. $e^{z_1+z_2} = e^{z_1}e^{z_2}$
3. $(e^z)^n = e^{nz}$ for positive integers z .
4. $(e^z)^{-1} = e^{-z}$
5. $e^z \neq 0$

It will turn out that the property $\frac{de^z}{dz} = e^z$ also holds, but we can't prove this yet because we haven't defined what we mean by the complex derivative $\frac{d}{dz}$.

Here are some more simple, but extremely important properties of e^z . You should become fluent in their use and know how to prove them.

6. $|e^{i\theta}| = 1$

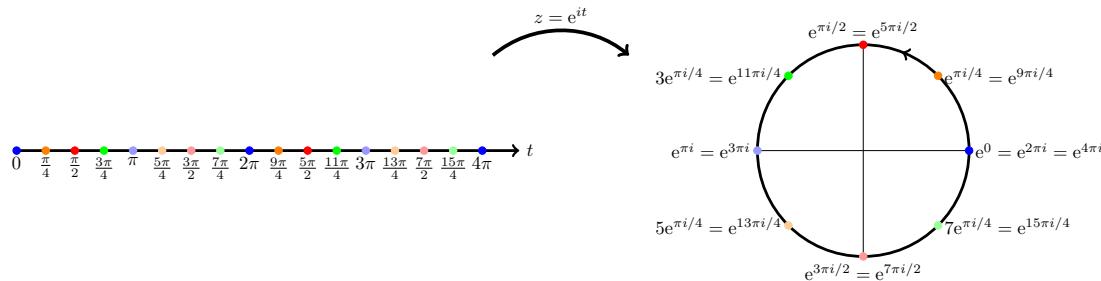
Proof.

$$|e^{i\theta}| = |\cos(\theta) + i\sin(\theta)| = \sqrt{\cos^2(\theta) + \sin^2(\theta)} = 1.$$

7. $|e^{x+iy}| = e^x$ (as usual $z = x + iy$ and x, y are real).

Proof. You should be able to supply this. **If not: ask a teacher or TA.**

8. The path e^{it} for $0 < t < \infty$ wraps counterclockwise around the unit circle. It does so infinitely many times. This is illustrated in the following picture.



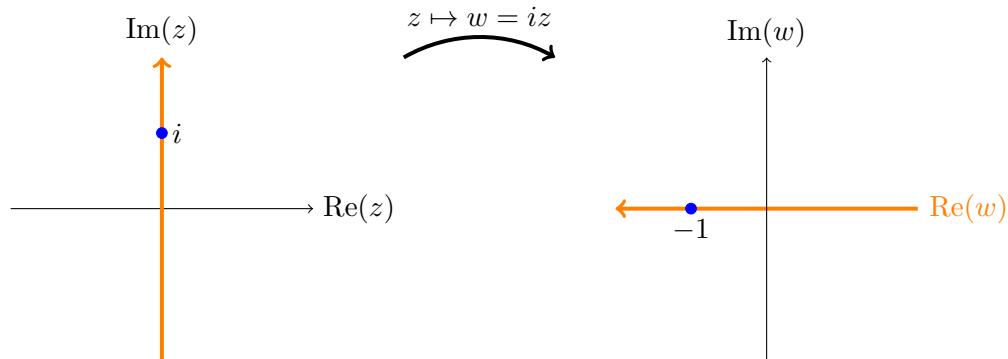
The map $t \rightarrow e^{it}$ wraps around the unit circle.

1.10 Complex functions as mappings

A complex function $w = f(z)$ is hard to graph because it takes 4 dimensions: 2 for z and 2 for w . So, to visualize them we will think of [complex functions as mappings](#). That is we will think of $w = f(z)$ as taking a number in the complex z -plane and mapping (sending) it to a number in the complex w -plane.

We will use the following terms and symbols to discuss mappings.

- A function $w = f(z)$ will also be called a [mapping](#) of z to w .
- Alternatively we will write $z \mapsto w$ or $z \mapsto f(z)$. This is read as “ z maps to w ”.
- We will say that “ w is the [image](#) of z under the mapping” or more simply “ w is the image of z ”.
- If we have a set of points in the z -plane we will talk of the image of that set under the mapping. For example, under the mapping $z \mapsto iz$ the image of the imaginary z -axis is the real w -axis.



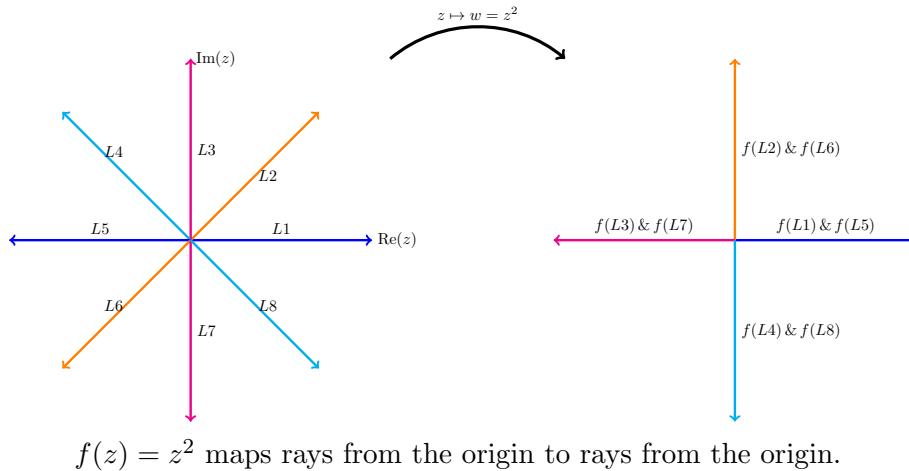
The image of the imaginary axis under $z \mapsto iz$.

Now we'll illustrate visualizing mappings with some examples:

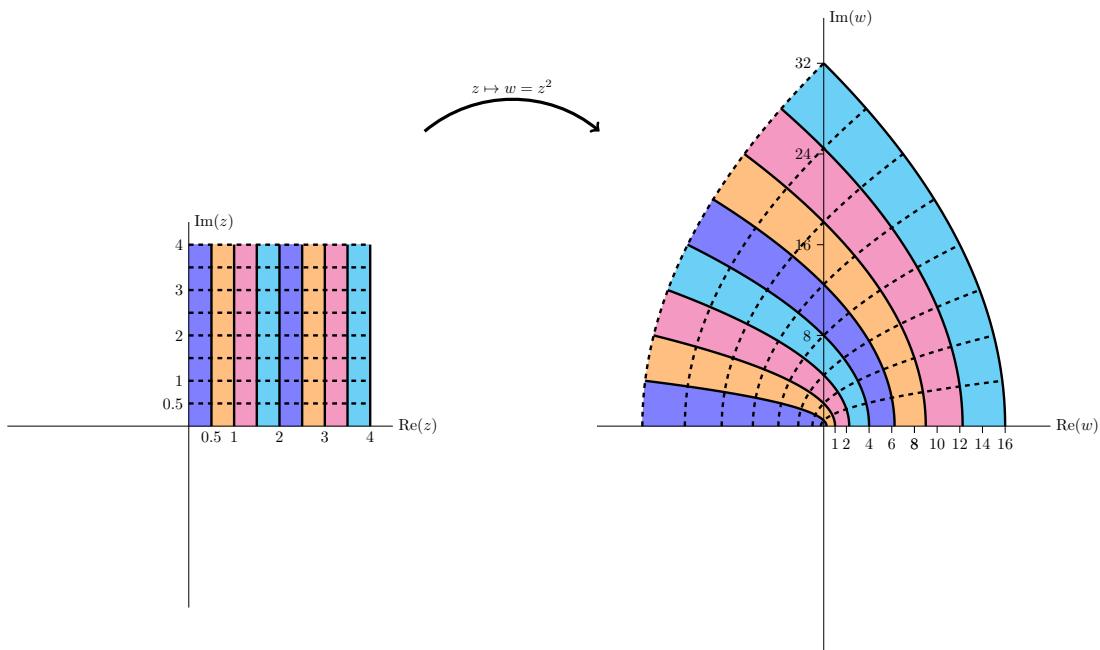
Example 1.15. The mapping $w = z^2$. We visualize this by putting the z -plane on the left and the w -plane on the right. We then draw various curves and regions in the z -plane and the corresponding image under z^2 in the w -plane.

In the first figure we show that rays from the origin are mapped by z^2 to rays from the origin. We see that

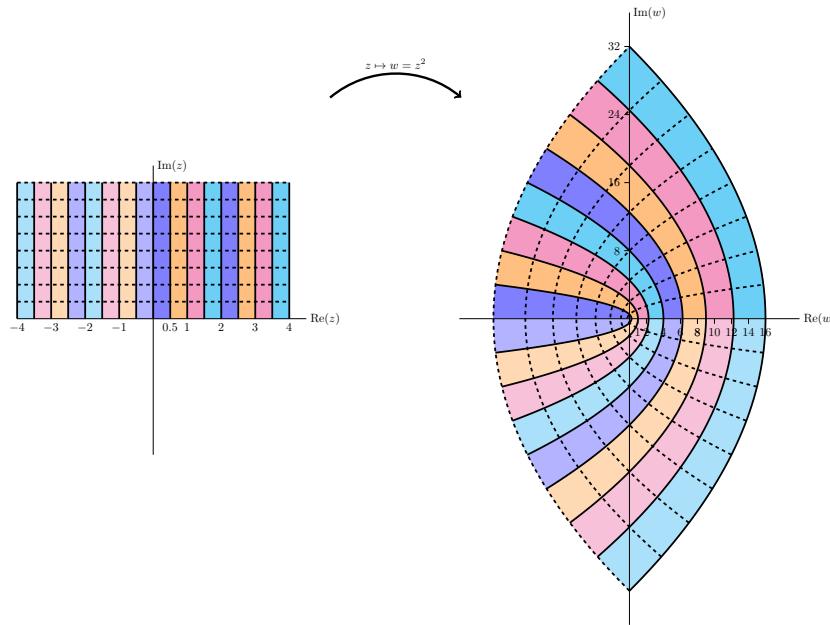
1. The ray L_2 at $\pi/4$ radians is mapped to the ray $f(L_2)$ at $\pi/2$ radians.
2. The rays L_2 and L_6 are both mapped to the same ray. This is true for each pair of diametrically opposed rays.
3. A ray at angle θ is mapped to the ray at angle 2θ .



The next figure gives another view of the mapping. Here we see vertical stripes in the first quadrant are mapped to parabolic stripes that live in the first and second quadrants.

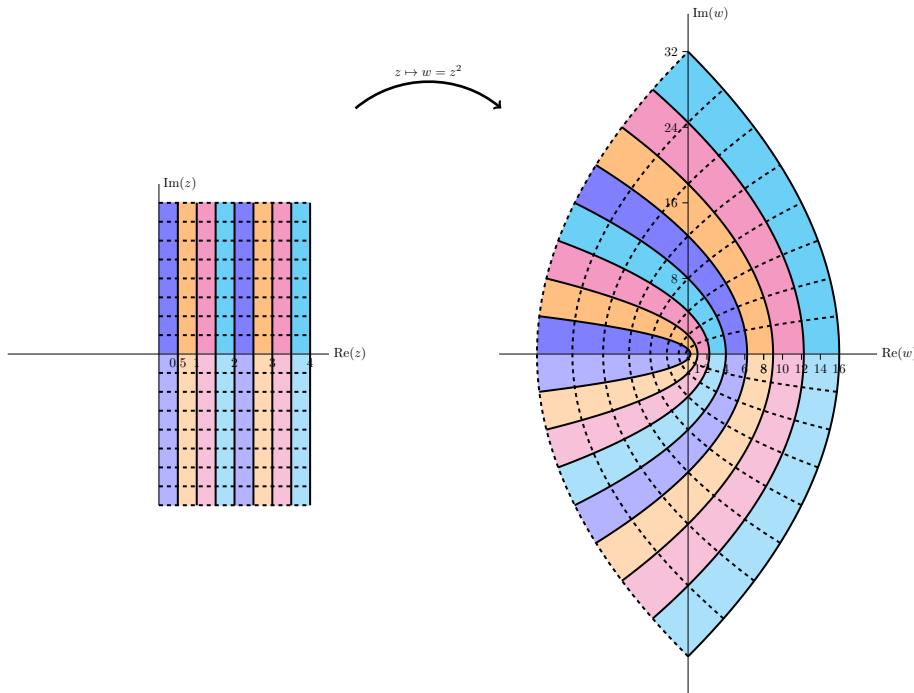


The next figure is similar to the previous one, except in this figure we look at vertical stripes in both the first and second quadrants. We see that they map to parabolic stripes that live in all four quadrants.

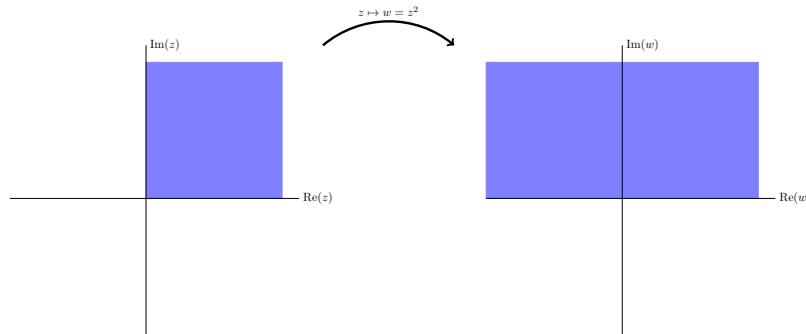


$f(z) = z^2$ maps the first two quadrants to the entire plane.

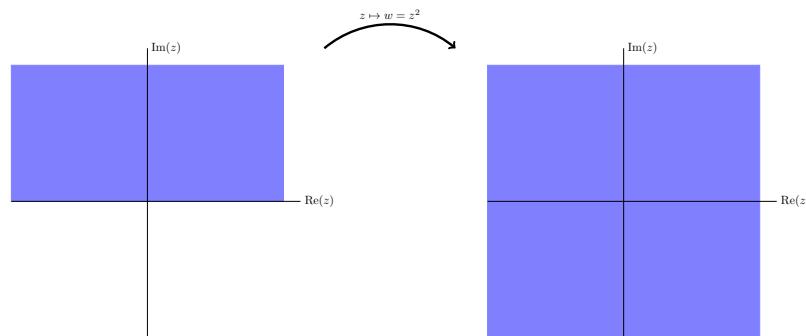
The next figure shows the mapping of stripes in the first and fourth quadrants. The image map is identical to the previous figure. This is because the fourth quadrant is minus the second quadrant, but $z^2 = (-z)^2$.



Vertical stripes in quadrant 4 are mapped identically to vertical stripes in quadrant 2.

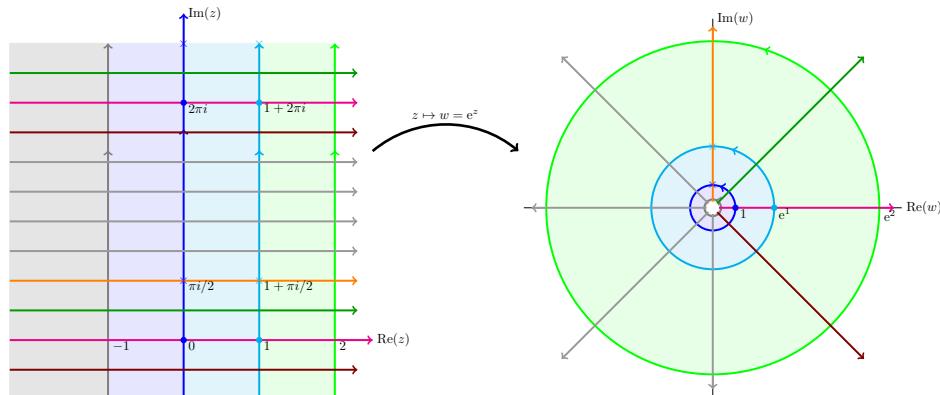


Simplified view of the first quadrant being mapped to the first two quadrants.



Simplified view of the first two quadrants being mapped to the entire plane.

Example 1.16. The mapping $w = e^z$. Here we present a series of plots showing how the exponential function maps z to w .

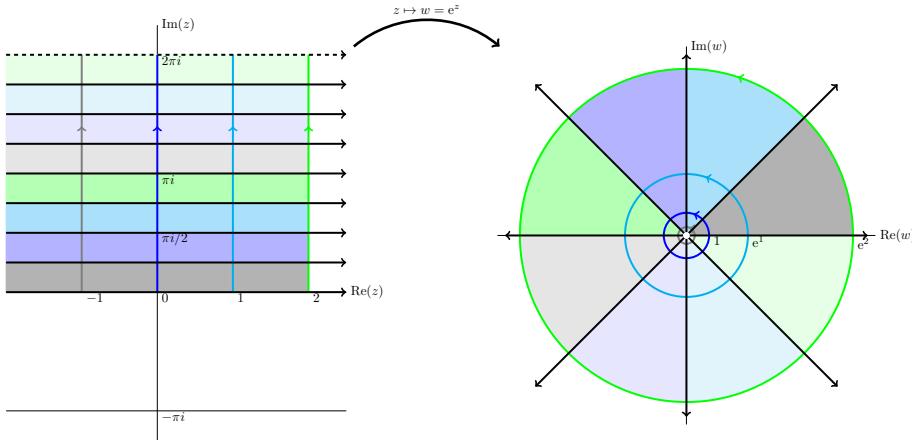


Notice that vertical lines are mapped to circles and horizontal lines to rays from the origin.

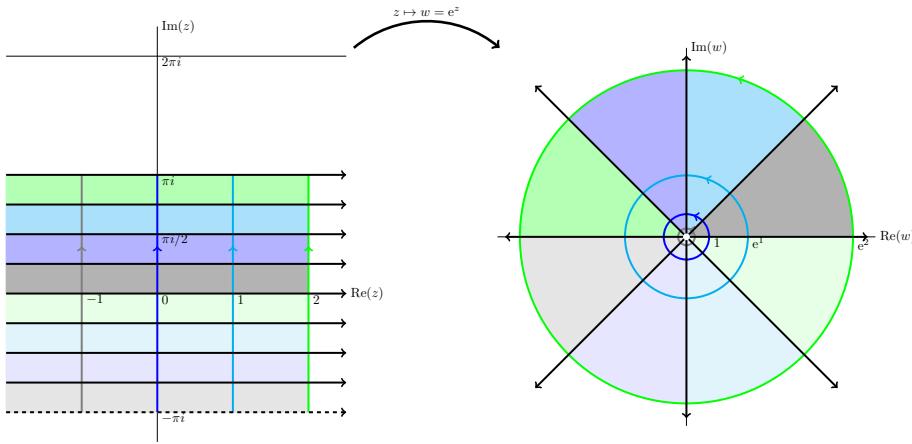
The next four figures all show essentially the same thing: the exponential function maps horizontal stripes to circular sectors. Any horizontal stripe of width 2π gets mapped to the entire plane minus the origin,

Because the plane minus the origin comes up frequently we give it a name:

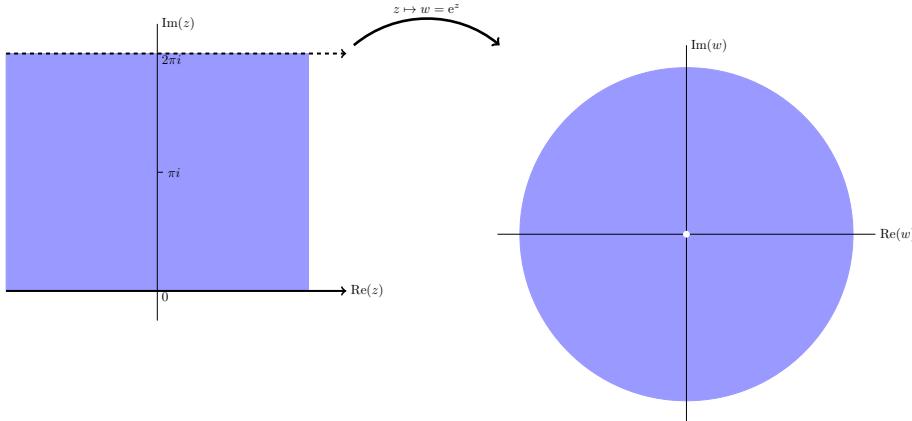
Definition. The punctured plane is the complex plane minus the origin. In symbols we can write it as $\mathbf{C} - \{0\}$ or $\mathbf{C}/\{0\}$.



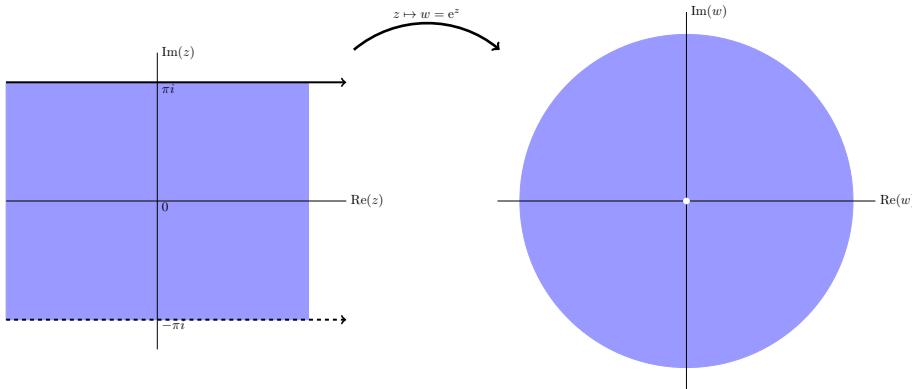
The horizontal $0 \leq y < 2\pi$ is mapped to the punctured plane



The horizontal $-\pi < y \leq \pi$ is mapped to the punctured plane



Simplified view showing e^z maps the horizontal stripe $0 \leq y < 2\pi$ to the punctured plane.



Simplified view showing e^z maps the horizontal stripe $-\pi < y \leq \pi$ to the punctured plane.

1.11 $\log(z)$ and $\arg(z)$

1.11.1 Many-to-one functions

The function $f(z) = z^2$ maps $\pm z$ to the same value, e.g. $f(2) = f(-2) = 4$. We say that $f(z)$ is a 2-to-1 function. That is, it maps 2 different points to each value. (Technically, it only maps one point to 0, but we will gloss over that for now.) Here are some other examples of many-to-one functions.

Example 1.17. $w = z^3$ is a 3-to-1 function. For example, 3 different z values get mapped to $w = 1$:

$$1^3 = \left(\frac{-1 + \sqrt{3}i}{2} \right)^3 = \left(\frac{-1 - \sqrt{3}i}{2} \right)^3 = 1$$

Example 1.18. The function $w = e^z$ maps infinitely many points to each value. For example

$$\begin{aligned} e^0 &= e^{2\pi i} = e^{4\pi i} = \dots e^{2n\pi i} = \dots = 1 \\ e^{i\pi/2} &= e^{i\pi/2+2\pi i} = e^{i\pi/2+4\pi i} = \dots e^{i\pi/2+2n\pi i} = \dots = i \end{aligned}$$

In general, $e^{z+2n\pi i}$ has the same value for every integer n .

1.11.2 Branches of $\arg(z)$

Important note. You should master this section. Branches of $\arg(z)$ are the key that really underlies all our other examples. Fortunately it is reasonably straightforward.

The key point is that the argument is only defined up to multiples of $2\pi i$ so every z produces infinitely many values for $\arg(z)$. Because of this we will say that $\arg(z)$ is a **multiple-valued function**.

Note. You might complain that by definition a function must take just one value. And you would be right. But we will call functions like $\arg(z)$ or \sqrt{z} (which has 2 possible values) multiple-valued. What that means in practice is that whenever we use such a function

will have to be careful to specify which of the possible values we mean. This is known as specifying a branch of the function.

Definition. By a [branch of the argument function](#) we mean any interval of length 2π . By specifying a branch we are saying that we will take the single value of $\arg(z)$ that lies in the branch.

Let's look at several different branches to understand how they work:

(i) If we specify the branch as $0 \leq \arg(z) < 2\pi$ then we have the following arguments.

$$\arg(1) = 0; \quad \arg(i) = \pi/2; \quad \arg(-1) = \pi; \quad \arg(-i) = 3\pi/2$$

This branch and these points are shown graphically in Figure (i) below.

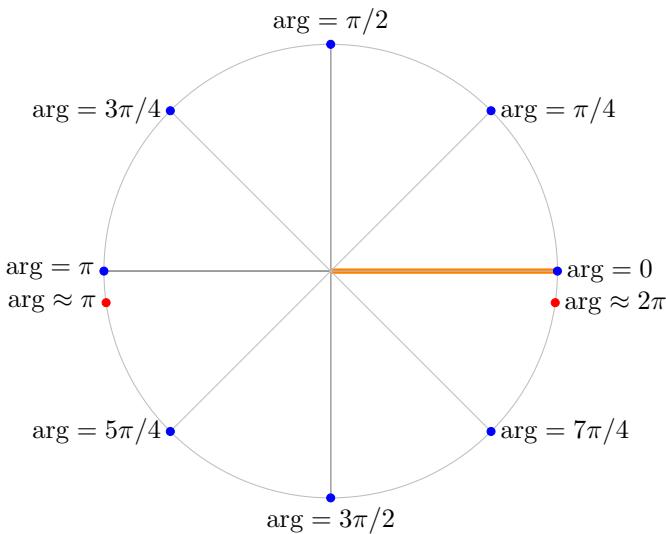


Figure (i): The branch $0 \leq \arg(z) < 2\pi$ of $\arg(z)$.

Notice that if we start at $z = 1$ on the positive real axis we have $\arg(z) = 0$. Then $\arg(z)$ increases as we move counterclockwise around the circle. The argument is continuous until we get back to the positive real axis. There it jumps from almost 2π back to 0.

There is no getting around (no pun intended) this discontinuity. We will call the line where the discontinuity happens the [branch cut](#) for this branch of $\arg(z)$. It is shown as a thick orange line in the figure.

For future reference you should note that near the negative real axis $\arg(z)$ is continuous and the arguments of nearby points are close to each other.

(ii) If we specify the branch as $-\pi < \arg(z) \leq \pi$ then we have the following arguments:

$$\arg(1) = 0; \quad \arg(i) = \pi/2; \quad \arg(-1) = \pi; \quad \arg(-i) = -\pi/2$$

This branch and these points are shown graphically in Figure (ii) below.

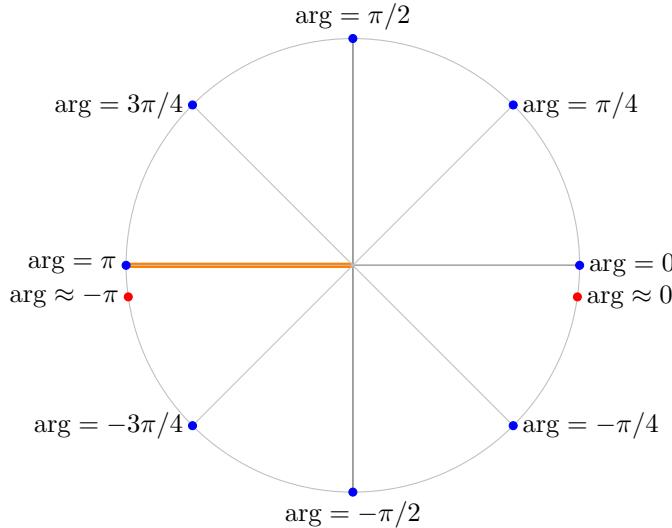


Figure (ii): The branch $-\pi < \arg(z) \leq \pi$ of $\arg(z)$.

Compare this figure with Figure (i). The values of $\arg(z)$ are the same in the upper half plane, but in the lower half plane they differ by 2π .

For this branch the branch cut is along the negative real axis. As we cross the branch cut the value of $\arg(z)$ jumps from π to something close to $-\pi$.

(iii) Figure (iii) shows the branch of $\arg(z)$ with $\pi/4 \leq \arg(z) < 9\pi/4$.

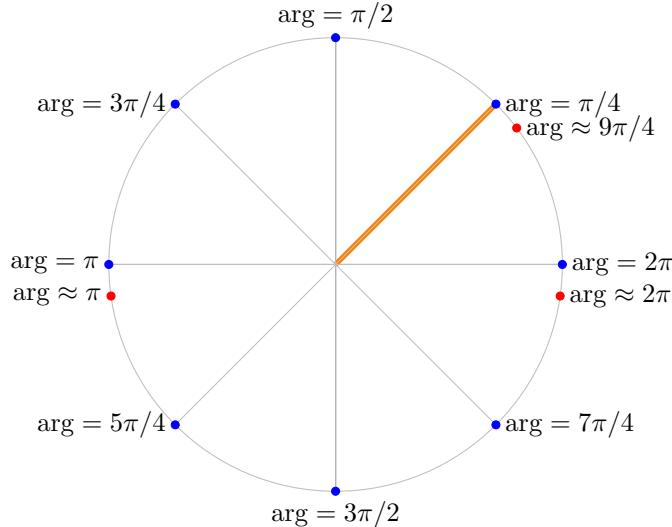


Figure (iii): The branch $\pi/4 \leq \arg(z) < 9\pi/4$ of $\arg(z)$.

Notice that on this branch $\arg(z)$ is continuous at both the positive and negative real axes. The jump of 2π occurs along the ray at angle $\pi/4$.

(iv) Obviously, there are many many possible branches. For example,

$$42 < \arg(z) \leq 42 + 2\pi.$$

(v) We won't make use of this in 18.04, but, in fact, the branch cut doesn't have to be a

straight line. Any curve that goes from the origin to infinity will do. The argument will be continuous except for a jump by 2π when z crosses the branch cut.

1.12 The principal branch of $\arg(z)$

Branch (ii) in the previous section is singled out and given a name:

Definition. The branch $-\pi < \arg(z) \leq \pi$ is called [the principal branch of \$\arg\(z\)\$](#) . We will use the notation $\text{Arg}(z)$ (capital A) to indicate that we are using the principal branch. (Of course, in cases where we don't want there to be any doubt we will say explicitly that we are using the principal branch.)

1.13 $\log(z)$

Our goal in this section is to define the log function. We want $\log(z)$ to be the inverse of e^z . That is, we want $e^{\log(z)} = z$. We will see that $\log(z)$ is multiple-valued, so when we use it we will have to specify a branch.

We start by looking at the simplest example which illustrates that $\log(z)$ is multiple-valued.

Example 1.19. Find $\log(1)$.

answer: We know that $e^0 = 1$, so $\log(1) = 0$ is one answer.

We also know that $e^{2\pi i} = 1$, so $\log(1) = 2\pi i$ is another possible answer. In fact, we can choose any multiple of $2\pi i$:

$$\log(1) = 2n\pi i, \text{ where } n \text{ is any integer}$$

This example leads us to consider the polar form for z as we try to define $\log(z)$. If $z = re^{i\theta}$ then one possible value for $\log(z)$ is

$$\log(z) = \log(re^{i\theta}) = \log(r) + i\theta,$$

here $\log(r)$ is the usual logarithm of a real positive number. For completeness we show explicitly that with this definition $e^{\log(z)} = z$:

$$e^{\log(z)} = e^{\log(r)+i\theta} = e^{\log(r)}e^{i\theta} = re^{i\theta} = z \quad \text{QED.}$$

Since $r = |z|$ and $\theta = \arg(z)$ we have arrived at our definition.

Definition. The function $\log(z)$ is defined as

$$\log(z) = \log(|z|) + i\arg(z),$$

where $\log(|z|)$ is the usual natural logarithm of a positive real number.

- Notes.**
1. Since $\arg(z)$ has infinitely many possible values, so does $\log(z)$.
 2. $\log(0)$ is not defined. (Both because $\arg(0)$ is not defined and $\log(|0|)$ is not defined.)
 3. Choosing a branch for $\arg(z)$ makes $\log(z)$ single valued. The usual terminology is to say we have chosen a [branch of the log function](#).

Example 1.20. Compute all the values of $\log(i)$. Specify which one comes from the principal branch.

answer: We have that $|i| = 1$ and $\arg(i) = \pi/2 + 2n\pi$, so

$$\log(i) = \log(1) + \frac{\pi}{2}i + 2n\pi i = \frac{\pi}{2}i + 2n\pi i, \text{ where } n \text{ is any integer.}$$

The principal branch of $\arg(z)$ is between $-\pi$ and π , so $\operatorname{Arg}(i) = \pi/2$. The value of $\log(i)$ from the principal branch is $i\pi/2$.

Example 1.21. Compute all the values of $\log(-1 - \sqrt{3}i)$. Specify which one comes from the principal branch.

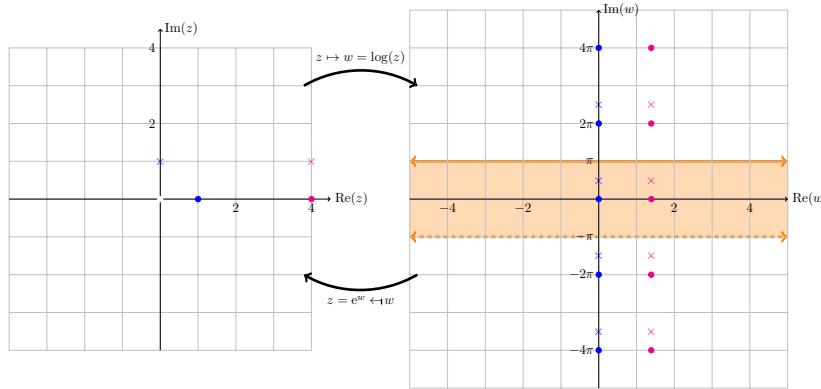
answer: Let $z = -1 - \sqrt{3}i$. Then $|z| = 2$ and in the principal branch $\operatorname{Arg}(z) = -2\pi/3$. So all the values of $\log(z)$ are

$$\log(z) = \log(2) - \frac{2\pi}{3}i + 2n\pi i.$$

The value from the principal branch is $\log(z) = \log(2) - 2\pi i/3$.

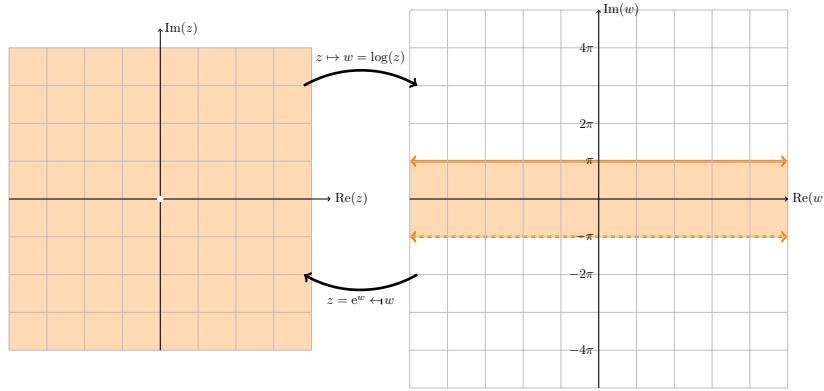
1.14 Figures showing $w = \log(z)$ as a mapping

The figures below show different aspects of the mapping given by $\log(z)$. In the first figure we see that a point z is mapped to (infinitely) many values of w . In this case we show $\log(1)$ (blue dots), $\log(4)$ (red dots), $\log(i)$ (blue cross), $\log(4i)$ (red cross). The values in the principal branch are inside the shaded region in the w -plane. Note that the values of $\log(z)$ for a given z are placed at intervals of $2\pi i$ in the w -plane.

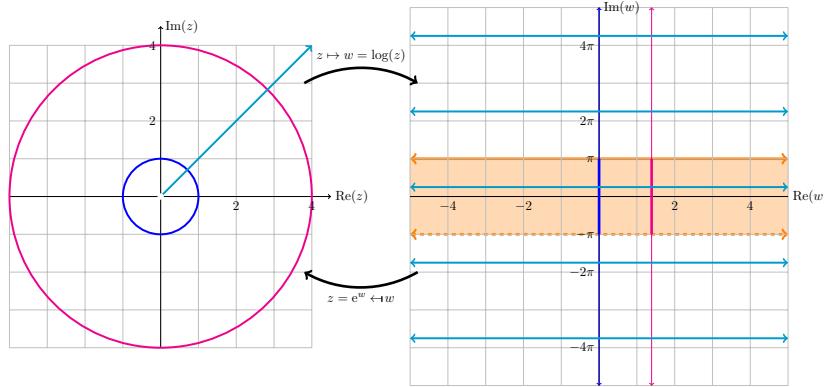


Mapping $\log(z)$: $\log(1)$, $\log(4)$, $\log(i)$, $\log(4i)$

The second figure illustrates that the principal branch of \log maps the punctured plane to the horizontal strip $-\pi < \operatorname{Im}(w) \leq \pi$.

Mapping $\log(z)$: the principal branch and the punctured plane

The third figure shows how circles centered on 0 are mapped to vertical lines, and rays from the origin are mapped to horizontal lines. If we restrict ourselves to the principal branch the circles are mapped to vertical line segments and rays to a single horizontal line in the principal (shaded) region of the w -plane.

Mapping $\log(z)$: mapping circles and rays

1.15 Complex powers

We can use the log function to define complex powers.

Definition. Let z and a be complex numbers then the power z^a is defined as

$$z^a = e^{a \log(z)}.$$

This is generally multiple-valued, so to specify a single value requires choosing a branch of $\log(z)$.

Example 1.22. Compute all the values of $\sqrt{2i}$. Give the value associated to the principal branch of $\log(z)$.

answer: $\log(2i) = \log(2e^{\frac{i\pi}{2}}) = \log(2) + i\frac{\pi}{2} + i2n\pi$. So,

$$\sqrt{2i} = (2i)^{1/2} = e^{\left(\frac{\log(2i)}{2}\right)} = e^{\left(\frac{\log(2)}{2} + \frac{i\pi}{4} + in\pi\right)} = \sqrt{2}e^{\left(\frac{i\pi}{4} + in\pi\right)}.$$

(As usual n is an integer.) As we saw earlier, this only gives two distinct values. The principal branch has $\text{Arg}(2i) = \pi/2$, so

$$\sqrt{2i} = \sqrt{2}e^{(\frac{i\pi}{4})} = \sqrt{2}\frac{(1+i)}{\sqrt{2}} = 1+i.$$

The other distinct value is when $n = 1$ and gives minus the value just above.

Example 1.23. Cube roots: To be added later –you should try adding it yourself!

Example 1.24. Compute all the values of 1^i . What is the value from the principal branch?

answer: This is similar to the problems above. $\log(1) = 2n\pi i$, so

$$1^i = e^{i\log(1)} = e^{i2n\pi i} = e^{-2n\pi}, \text{ where } n \text{ is an integer.}$$

The principal branch has $\log(1) = 0$ so $1^i = 1$.

2 Analytic functions

2.1 Introduction

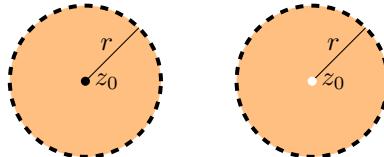
The main goal of this topic is to define and give some of the important properties of complex analytic functions. A function $f(z)$ is analytic if it has a complex derivative $f'(z)$. In general, the rules for computing derivatives will be familiar to you from single variable calculus. However, a much richer set of conclusions can be drawn about a complex analytic function than is generally true about real differentiable functions.

Since the derivative is defined as a limit and we will want to define our functions on open regions, the first few sections will give definitions of these concepts which will suffice for 18.04. For anyone who is interested, we put a more formal discussion of limits in an appendix at the end of this topic's notes.

2.2 Open disks, open deleted disks, open regions

Definition. The **open disk of radius r around z_0** is the set of points z with $|z - z_0| < r$, i.e. all points within distance r of z_0 .

The **open deleted disk of radius r around z_0** is the set of points z with $0 < |z - z_0| < r$. That is, we remove the center z_0 from the open disk.



Left: an open disk around z_0 ; right: a deleted open disk around z_0

A deleted disk is also called a **punctured disk**.

Definition. An [open region](#) in the complex plane is a set A with the property that every point in A can be surrounded by an open disk that lies entirely in A . We will often drop the word open and simply call A a region.

In the figure below, the set A on the left is an open region because for every point in A we can draw a little circle around the point that is completely in A . The set B is not an open region because it contains its boundary, so we can't draw a circle around the point z that is completely inside B .



Left: an open region A ; right: B is not an open region

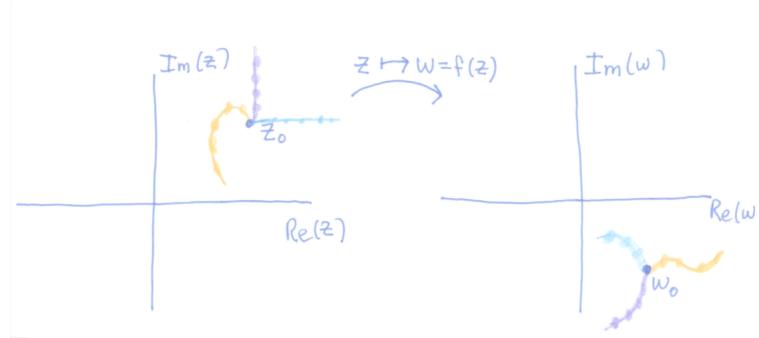
2.3 Limits and continuous functions

Definition. If $f(z)$ is defined on a deleted disk around z_0 then we say

$$\lim_{z \rightarrow z_0} f(z) = w_0$$

if $f(z)$ goes to w_0 no matter what direction z approaches z_0 .

The figure below shows several sequences of points that approach z_0 . If $\lim_{z \rightarrow z_0} f(z) = w_0$ then $f(z)$ must go to w_0 along each of these sequences.



Sequences going to z_0 are mapped to sequences going to w_0 .

Example 2.1. Many functions have obvious limits. For example: $\lim_{z \rightarrow 2} z^2 = 4$ and $\lim_{z \rightarrow 2} (z^2 + 2)/(z^3 + 1) = 6/9$.

Here is an example where the limit doesn't exist because different sequences give different limits.

Example 2.2. (No limit) Show that

$$\lim_{z \rightarrow 0} \frac{z}{\bar{z}} = \lim_{z \rightarrow 0} \frac{x+iy}{x-iy}$$

does not exist.

answer: On the real axis we have $\frac{z}{\bar{z}} = \frac{x}{x} = 1$, so the limit as $z \rightarrow 0$ along the real axis is 1.

On the imaginary axis we have $\frac{z}{\bar{z}} = \frac{iy}{-iy} = -1$, so the limit as $z \rightarrow 0$ along the imaginary axis is -1.

Since the two limits do not agree the limit as $z \rightarrow 0$ does not exist!

2.3.1 Properties of limits

We have the usual properties of limits. Suppose

$$\lim_{z \rightarrow z_0} f(z) = w_1 \text{ and } \lim_{z \rightarrow z_0} g(z) = w_2$$

then

- $\lim_{z \rightarrow z_0} f(z) + g(z) = w_1 + w_2$.
- $\lim_{z \rightarrow z_0} f(z)g(z) = w_1 \cdot w_2$.
- If $w_2 \neq 0$ then $\lim_{z \rightarrow z_0} f(z)/g(z) = w_1/w_2$
- If $h(z)$ is continuous and defined on a neighborhood of w_1 then $\lim_{z \rightarrow z_0} h(f(z)) = h(w_1)$
(Note: we will give the official definition of continuity in the next section.)

We won't give a proof of these properties. As a challenge, you can try to supply it using the formal definition of limits given in the appendix.

We can restate the definition of limit in terms of functions of (x, y) ,

Write $f(z) = f(x + iy) = u(x, y) + iv(x, y)$. Then

$$\lim_{z \rightarrow z_0} f(z) = w_0 \text{ iff } \begin{cases} \lim_{P \rightarrow P_0} u(x, y) = u_0 \\ \lim_{P \rightarrow P_0} v(x, y) = v_0, \end{cases}$$

where $P = (x, y)$, $P_0 = (x_0, y_0)$, $w_0 = u_0 + iv_0$.

Notes. 1. The term 'iff' stands for 'if and only if' which is another way of saying 'is equivalent to'.

2.3.2 Continuous functions

A function is continuous if it doesn't have any sudden jumps. This is the gist of the following definition.

Definition. If the function $f(z)$ is defined on an open disk around z_0 and $\lim_{z \rightarrow z_0} f(z) = f(z_0)$ then we say f is continuous at z_0 .

If f is defined on an open region A then the phrase ' f is continuous on A ' means that f is continuous at every point in A .

As usual, we can rephrase this in terms of functions of (x, y) :

Fact. $f(z) = u(x, y) + iv(x, y)$ is continuous iff $u(x, y)$ and $v(x, y)$ are continuous as functions of two variables.

Example 2.3. (Some continuous functions)

(i) A polynomial $P(z) = a_0 + a_1z + a_2z^2 + \dots + a_nz^n$ is continuous on the entire plane.

Reason: it is clear that each power $(x + iy)^k$ is continuous as a function of (x, y) .

(ii) The exponential function is continuous on the entire plane. Reason: $e^z = e^{x+iy} = e^x \cos(y) + ie^x \sin(y)$. So the both the real and imaginary parts are clearly continuous as a function of (x, y) .

(iii) The function $\text{Arg}(z)$ is continuous on the plane minus the non-positive real axis. Reason: this is clear and is the reason we defined branch cuts for \arg . We have to remove the negative real axis because $\text{Arg}(z)$ jumps by 2π when you cross it. We also have to remove $z = 0$ because $\text{Arg}(z)$ is not even defined at 0.

(iv) The principal branch of the function $\log(z)$ is continuous on the plane minus the non-positive real axis. Reason: the principal branch of \log has $\log(z) = \log(r) + i \text{Arg}(z)$. So the continuity of $\log(z)$ follows from the continuity of $\text{Arg}(z)$.

2.3.3 Properties of continuous functions

Since continuity is defined in terms of limits, we have the following properties of continuous functions.

Suppose $f(z)$ and $g(z)$ are continuous on a region A . Then

- $f(z) + g(z)$ is continuous on A .
- $f(z)g(z)$ is continuous on A .
- $f(z)/g(z)$ is continuous on A except at points where $g(z) = 0$.
- If h is continuous on $f(A)$ then $h(f(z))$ is continuous on A .

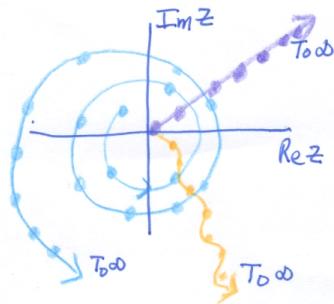
Using these properties we can claim continuity for each of the following functions:

- e^{z^2}
- $\cos(z) = (\mathrm{e}^{iz} + \mathrm{e}^{-iz})/2$
- If $P(z)$ and $Q(z)$ are polynomials then $P(z)/Q(z)$ is continuous except at roots of $Q(z)$.

2.4 The point at infinity

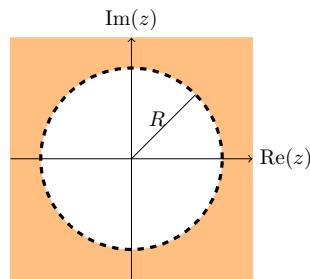
By definition the **extended complex plane** = $\mathbf{C} \cup \{\infty\}$. That is, we have **one point at infinity** to be thought of in a limiting sense described as follows.

A sequence of points $\{z_n\}$ goes to infinity if $|z_n|$ goes to infinity. This “point at infinity” is approached in any direction we go. All of the sequences shown in the figure below are growing, so they all go to the (same) “point at infinity”.



Various sequences all going to infinity.

If we draw a large circle around 0 in the plane, then we call the region **outside** this circle a **neighborhood of infinity**.



The shaded region outside the circle of radius R is a neighborhood of infinity.

2.4.1 Limits involving infinity

The key idea is $1/\infty = 0$. By this we mean

$$\lim_{z \rightarrow \infty} \frac{1}{z} = 0$$

We then have the following facts:

- $\lim_{z \rightarrow z_0} f(z) = \infty \Leftrightarrow \lim_{z \rightarrow z_0} 1/f(z) = 0$
- $\lim_{z \rightarrow \infty} f(z) = w_0 \Leftrightarrow \lim_{z \rightarrow 0} f(1/z) = w_0$
- $\lim_{z \rightarrow \infty} f(z) = \infty \Leftrightarrow \lim_{z \rightarrow 0} \frac{1}{f(1/z)} = 0$

Example 2.4. $\lim_{z \rightarrow \infty} e^z$ is not defined because it has different values if we go to infinity in different directions, e.g. we have $e^z = e^x e^{iy}$ and

$$\lim_{x \rightarrow -\infty} e^x e^{iy} = 0$$

$$\lim_{x \rightarrow +\infty} e^x e^{iy} = \infty$$

$\lim_{y \rightarrow +\infty} e^x e^{iy}$ is not defined, i.e. x is constant, so $e^x e^{iy}$ loops in a circle indefinitely.)

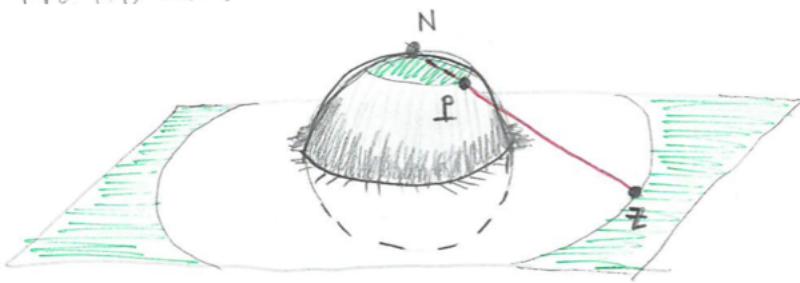
Example 2.5. Show $\lim_{z \rightarrow \infty} z^n = \infty$ (for n a positive integer).

answer: We need to show that $|z^n|$ gets large as $|z|$ gets large. This is obvious because if $|z| = R$ is large then $|z|^n = R^n$ is even larger.

2.4.2 Stereographic projection from the Riemann sphere

This is a lovely section and we suggest you read it. However it will be a while before we use it in 18.04.

One way to visualize the point at ∞ is by using a (unit) Riemann sphere and the associated stereographic projection. The figure below shows a sphere whose equator is the unit circle in the complex plane.



Stereographic projection from the sphere to the plane.

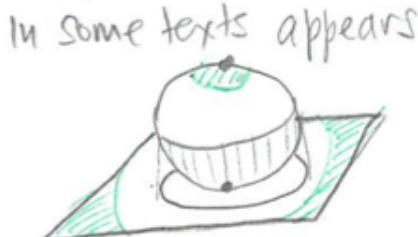
Stereographic projection from the sphere to the plane is accomplished by drawing the secant line from the north pole N through a point on the sphere and seeing where it intersects the plane. This gives a 1-1 correspondence between a point on the sphere P and a point in the complex plane z . It is easy to see show that the formula for stereographic projection is

$$P = (a, b, c) \mapsto z = \frac{a}{1 - c} + i \frac{b}{1 - c}.$$

The point $N = (0, 0, 1)$ is special, the secant lines from N through P become tangent lines to the sphere at N which never intersect the plane. We consider N the point at infinity.

In the figure above, the region outside the large circle through the point z is a neighborhood of infinity. It corresponds to the small circular cap around N on the sphere. That is, the small cap around N is a neighborhood of the point at infinity on the sphere!

The figure below shows another common version of stereographic projection. In this figure the sphere sits with its south pole at the origin. We still project using secant lines from the north pole.



2.5 Derivatives

The definition of the complex derivative of a complex function is similar to that of a real derivative of a real function: For a function $f(z)$ the derivative f at z_0 is defined as

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

Provided, of course, that the limit exists. If the limit exists we say f is analytic at z_0 or f is differentiable at z_0 .

Remember! The limit has to exist and be the same no matter how you approach z_0 .

If f is analytic at all the points in an open region A then we say f is analytic on A .

As usual with derivatives there are several alternative notations. For example, if $w = f(z)$ we can write

$$f'(z) = \frac{dw}{dz} = \lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0} = \lim_{\Delta z \rightarrow 0} \frac{\Delta w}{\Delta z}$$

Example 2.6. Find the derivative of $f(z) = z^2$.

answer: We compute using the definition of the derivative as a limit.

$$\lim_{\Delta z \rightarrow 0} \frac{(z + \Delta z)^2 - z^2}{\Delta z} = \lim_{\Delta z \rightarrow 0} \frac{z^2 + 2z\Delta z + (\Delta z)^2 - z^2}{\Delta z} = 2z.$$

Example 2.7. Show $f(z) = \bar{z}$ is not differentiable at any point z .

answer: Again we look at the limit definition for $f'(z_0)$. In this case we will see that the limit changes as we approach z_0 from different directions. Let $z = x + iy$ and $z_0 = x_0 + iy_0$. Then the limit we want is

$$\lim_{z \rightarrow z_0} \frac{\bar{z} - \bar{z}_0}{z - z_0} = \lim_{z \rightarrow z_0} \frac{(x - x_0) - i(y - y_0)}{(x - x_0) + i(y - y_0)}$$

If we approach z_0 along the horizontal line through z_0 , i.e. $z = x + iy_0$ and $x \rightarrow x_0$ the limit becomes

$$\lim_{x \rightarrow x_0} \frac{x - x_0}{x - x_0} = 1.$$

Likewise, if we approach z_0 along the vertical line through z_0 the limit becomes

$$\lim_{y \rightarrow y_0} \frac{-i(y - y_0)}{i(y - y_0)} = -1.$$

Since these limits are different, the derivative does not exist.

Challenge. Use polar coordinates to show the limit in the previous example can be any value with modulus 1 depending on the angle at which z approaches z_0 .

2.5.1 Derivative rules

It wouldn't be much fun to compute every derivative using limits. Fortunately, we have the same differentiation formulas as for real-valued functions. That is, assuming f and g are differentiable we have:

- Sum rule: $\frac{d}{dz}(f(z) + g(z)) = f' + g'$
- Product rule: $\frac{d}{dz}(f(z)g(z)) = f'g + fg'$
- Quotient rule: $\frac{d}{dz}(f(z)/g(z)) = \frac{f'g - fg'}{g^2}$
- Chain rule: $\frac{d}{dz}g(f(z)) = g'(f(z))f'(z)$

To give you the flavor of these arguments we'll prove the product rule.

$$\begin{aligned}\frac{d}{dz}(f(z)g(z)) &= \lim_{z \rightarrow z_0} \frac{f(z)g(z) - f(z_0)g(z_0)}{z - z_0} \\ &= \lim_{z \rightarrow z_0} \frac{(f(z) - f(z_0))g(z) + f(z_0)(g(z) - g(z_0))}{z - z_0} \\ &= \lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0}g(z) + f(z_0)\frac{(g(z) - g(z_0))}{z - z_0} \\ &= f'(z_0)g(z_0) + f(z_0)g'(z_0)\end{aligned}$$

Here is an important fact that you would have guessed. We will prove it in the next section.

Theorem. If $f(z)$ is defined and differentiable on an open disk and $f'(z) = 0$ on the disk then $f(z)$ is constant.

2.6 Cauchy-Riemann equations

The Cauchy-Riemann equations are our first consequence of the fact that the limit defining $f(z)$ must be the same no matter which direction you approach z from. They will be one of the most important tools in our toolbox.

2.6.1 Partial derivatives as limits

Before getting to the Cauchy-Riemann equations we remind you about partial derivatives. If $u(x, y)$ is a function of two variables then the partial derivatives of u are defined as

$$\frac{\partial u}{\partial x}(x, y) = \lim_{\Delta x \rightarrow 0} \frac{u(x + \Delta x, y) - u(x, y)}{\Delta x},$$

i.e. the derivative of u holding y constant.

$$\frac{\partial u}{\partial y}(x, y) = \lim_{\Delta y \rightarrow 0} \frac{u(x, y + \Delta y) - u(x, y)}{\Delta y},$$

i.e. the derivative of u holding x constant.

2.6.2 The Cauchy-Riemann equations

The following shows how to derive the Cauchy-Riemann equations. The full statement of these equations is stated at the end as a theorem.

Let's suppose that $f(z)$ is differentiable in some region A and

$$f(z) = f(x + iy) = u(x, y) + iv(x, y).$$

We'll compute $f'(z)$ by approaching z first from the horizontal direction and then from the vertical direction. We'll use the formula

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

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

Horizontal direction: $\Delta y = 0$, $\Delta z = \Delta x$

$$\begin{aligned} f'(z) &= \lim_{\Delta z \rightarrow 0} \frac{f(z + \Delta z) - f(z)}{\Delta z} \\ &= \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x + iy) - f(x + iy)}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} \frac{(u(x + \Delta x, y) + iv(x + \Delta x, y)) - (u(x, y) + iv(x, y))}{\Delta x} \\ &= \lim_{\Delta x \rightarrow 0} \frac{u(x + \Delta x, y) - u(x, y)}{\Delta x} + i \frac{v(x + \Delta x, y) - v(x, y)}{\Delta x} \\ &= \frac{\partial u}{\partial x}(x, y) + i \frac{\partial v}{\partial x}(x, y) \end{aligned}$$

Vertical direction: $\Delta x = 0$, $\Delta z = i\Delta y$ (We'll do this one a little faster.)

$$\begin{aligned} f'(z) &= \lim_{\Delta z \rightarrow 0} \frac{f(z + \Delta z) - f(z)}{\Delta z} \\ &= \lim_{\Delta y \rightarrow 0} \frac{(u(x, y + \Delta y) + iv(x, y + \Delta y)) - (u(x, y) + iv(x, y))}{i\Delta y} \\ &= \lim_{\Delta y \rightarrow 0} \frac{u(x, y + \Delta y) - u(x, y)}{i\Delta y} + i \frac{v(x, y + \Delta y) - v(x, y)}{i\Delta y} \\ &= \frac{1}{i} \frac{\partial u}{\partial y}(x, y) + \frac{\partial v}{\partial y}(x, y) \\ &= \frac{\partial v}{\partial y}(x, y) - i \frac{\partial u}{\partial y}(x, y) \end{aligned}$$

We have found two different representations of $f'(z)$ in terms of the partials of u and v . We put them together and state our big theorem.

Theorem 2.8. (Cauchy-Riemann equations) If $f(z) = u(x, y) + iv(x, y)$ is differentiable then

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

In particular,

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

This last set of partial differential equations is what is usually meant by the Cauchy-Riemann equations.

Here is the short form of the Cauchy-Riemann equations:

$$\begin{aligned} u_x &= v_y \\ u_y &= -v_x \end{aligned}$$

It turns out that the converse is true and will be very useful to us.

Theorem. Consider the function $f(z) = u(x, y) + iv(x, y)$ defined on a region A . If u and v satisfy the Cauchy-Riemann equations and have continuous partials then $f(z)$ is differentiable on A .

The proof of this would require a careful consideration of all the limits involved.

2.6.3 Using the Cauchy-Riemann equations

The Cauchy-Riemann equations provide us with a direct way of checking that a function is differentiable and computing its derivative.

Example 2.9. Show e^z is differentiable and compute its derivative is e^z .

answer: We write $e^z = e^{x+iy} = e^x \cos(y) + ie^x \sin(y)$. So

$$u(x, y) = e^x \cos(y) \text{ and } v(x, y) = e^x \sin(y).$$

Computing partial derivatives we have

$$\begin{aligned} u_x &= e^x \cos(y), u_y = -e^x \sin(y) \\ v_x &= e^x \sin(y), v_y = e^x \cos(y) \end{aligned}$$

We see that $u_x = v_y$ and $u_y = -v_x$, so the Cauchy-Riemann equations are satisfied. Thus, e^z is differentiable and

$$\frac{d}{dz} e^z = u_x + iv_x = e^x \cos(y) + ie^x \sin(y) = e^z.$$

Example 2.10. Show that $f(z) = \bar{z}$ is not differentiable.

answer: $f(x + iy) = x - iy$, so $u(x, y) = x$, $v(x, y) = -y$. Taking partial derivatives

$$u_x = 1, u_y = 0, v_x = 0, v_y = -1$$

Since $u_x \neq v_y$ the Cauchy-Riemann equations are not satisfied and therefore f is not differentiable.

Theorem. If $f(z)$ is differentiable on a disk and $f'(z) = 0$ on the disk then $f(z)$ is constant.

Proof. Since f is differentiable and $f'(z) \equiv 0$, the Cauchy-Riemann equations show that

$$u_x(x, y) = u_y(x, y) = v_x(x, y) = v_y(x, y) = 0$$

We know from multivariable calculus that a function of (x, y) with both partials identically zero is constant. Thus u and v are constant, and therefore so is f .

2.6.4 $f'(z)$ as a 2×2 matrix

Recall that we could represent a complex number $a + ib$ as a 2×2 matrix

$$a + ib \leftrightarrow \begin{bmatrix} a & -b \\ b & a \end{bmatrix}. \quad (3)$$

Now if we write $f(z)$ in terms of (x, y) we have

$$f(z) = f(x + iy) = u(x, y) + iv(x, y) \leftrightarrow f(x, y) = (u(x, y), v(x, y))$$

Recall that the Jacobian of $f(x, y)$ is $\begin{bmatrix} u_x & u_y \\ v_x & v_y \end{bmatrix}$.

Finally, we have $f'(z) = u_x + iv_x$, so we can represent $f'(z)$ as $\begin{bmatrix} u_x & -v_x \\ v_x & u_x \end{bmatrix}$. Using the Cauchy-Riemann equations we can replace $-v_x$ by v_y and u_x by v_y which gives us the representation

$$f'(z) \leftrightarrow \begin{bmatrix} u_x & u_y \\ v_x & v_y \end{bmatrix},$$

i.e. $f'(z)$ is just the Jacobian of $f(x, y)$.

For me, it is easier to remember the Jacobian than the Cauchy-Riemann equations. Since $f'(z)$ is a complex number I can use the matrix representation in Equation 3 to remember the Cauchy-Riemann equations!

2.7 Gallery of functions

In this section we'll look at many of the functions you know and love as functions of z . For each one we'll have to do three things.

1. Define how to compute it.
2. Specify a domain where it is analytic.
3. Specify a branch and a branch cut (if needed) giving its range.
4. Compute its derivative.

Most often, we can compute the derivatives of a function using the algebraic rules like the quotient rule. If necessary we can use the Cauchy-Riemann equations or, as a last resort, even the definition of the derivative as a limit.

Before we start on the gallery we define the term “entire function”.

Definition. A function that is analytic at every point in the complex plane is called an **entire function**. We will see that e^z , z^n , $\sin(z)$ are all entire functions.

2.7.1 List of functions, derivatives and properties

The following is a concise list of a number of functions and their derivatives. None of the derivatives will surprise you. We also give important properties for some of the functions. The proofs for each follow below.

1. $f(z) = e^z = e^x \cos(y) + ie^x \sin(y).$

Domain = all of \mathbf{C} (f is entire).

$$f'(z) = e^z.$$

2. $f(z) \equiv c$ (constant)

Domain = all of \mathbf{C} (f is entire).

$$f'(z) = 0.$$

3. $f(z) = z^n$ (n an integer ≥ 0)

Domain = all of \mathbf{C} (f is entire).

$$f'(z) = nz^{n-1}.$$

4. $P(z)$ (polynomial)

A polynomial has the form $P(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_0.$

Domain = all of \mathbf{C} ($P(z)$ is entire).

$$P'(z) = na_n z^{n-1} + (n-1)a_{n-1} z^{n-2} + \dots + 2a_2 z + a_1.$$

5. $f(z) = 1/z$ Domain = $\mathbf{C} - \{0\}$ (the punctured plane).

$$f'(z) = -1/z^2.$$

6. $f(z) = P(z)/Q(z)$ (rational function). When P and Q are polynomials $P(z)/Q(z)$ is called a rational function.

We'll assume that they have no common factors, then

$$\text{Domain} = \mathbf{C} - \{\text{roots of } Q\} \quad f'(z) = \frac{P'Q - PQ'}{Q^2}.$$

7. $\sin(z), \cos(z)$

Definition. $\cos(z) = \frac{e^{iz} + e^{-iz}}{2}, \quad \sin(z) = \frac{e^{iz} - e^{-iz}}{2i}$

(By Euler's formula we know this is consistent with $\cos(x)$ and $\sin(x)$ when $z = x$ is real.)

Domain: these functions are entire.

$$\frac{d \cos(z)}{dz} = -\sin(z), \quad \frac{d \sin(z)}{dz} = \cos(z)$$

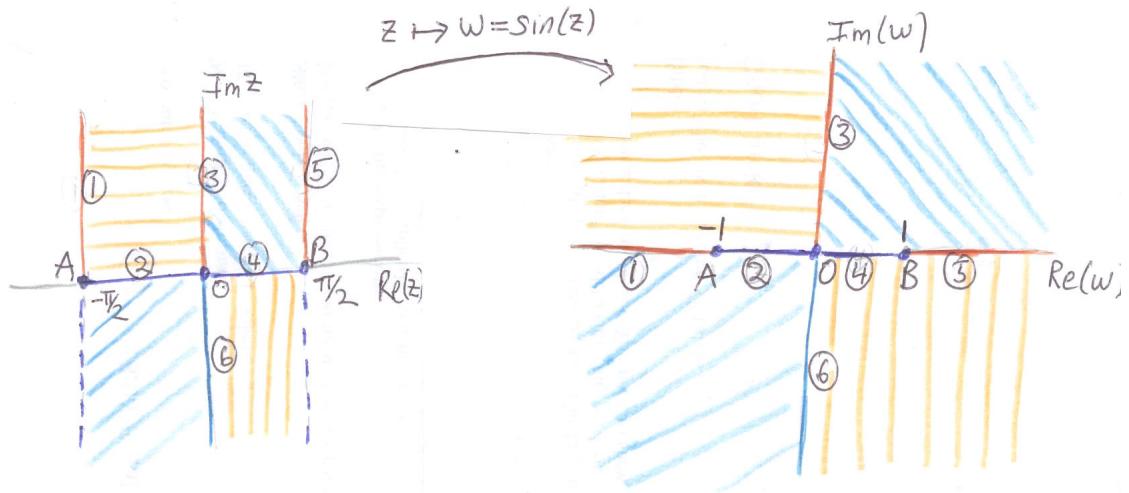
Other key properties of cos and sin:

- $\cos^2(z) + \sin^2(z) = 1$
- $e^z = \cos(z) + i \sin(z)$
- Periodic in x with period 2π , e.g. $\sin(x + 2\pi + iy) = \sin(x + iy)$.
- They are not bounded!
- In the form $f(z) = u(x, y) + iv(x, y)$ we have

$$\cos(z) = \cosh(y) \cos(x) + i \sinh(y) \sin(x)$$

$$\sin(z) = \cosh(y) \sin(x) + i \sinh(y) \cos(x)$$

- They are not 1-1 functions so we will have to choose a branch (i.e. a range for the image) in order to define inverses. The figure below shows a domain where $\sin(z)$ is 1-1. The domain consists of the vertical strip $z = x + iy$ with $-\pi/2 < x < \pi/2$ together with the two rays on boundary where $y \geq 0$. The figure indicates how the regions making up the domain in the z -plane are mapped to the quadrants in the w -plane.



A domain where $z \mapsto w = \sin(z)$ is one-to-one

- The zeros of $\sin(z)$ are $z = n\pi$ for n any integer.
- The zeros of $\cos(z)$ are $z = \pi/2 + n\pi$ for n any integer.
- (I.e., they have only real zeros that you learned about in your trig. class.)

8. Other trig functions $\cot(z)$, $\sec(z)$ etc.

Definition. The same as for the real versions of these functions, e.g. $\cot(z) = \cos(z)/\sin(z)$, $\sec(z) = 1/\cos(z)$.

Domain: The entire plane minus the zeros of the denominator.

Derivative: Compute using the quotient rule, e.g.

$$\frac{d \tan(z)}{dz} = \frac{d}{dz} \left(\frac{\sin(z)}{\cos(z)} \right) = \frac{\cos(z)\cos(z) - \sin(z)(-\sin(z))}{\cos^2(z)} = \frac{1}{\cos^2(z)} = \sec^2 z$$

(No surprises there!)

9. $\sinh(z)$, $\cosh(z)$ (hyperbolic sine and cosine)

Definition.

$$\cosh(z) = \frac{e^z + e^{-z}}{2}, \quad \sinh(z) = \frac{e^z - e^{-z}}{2}$$

Domain: these functions are entire.

$$\frac{d \cosh(z)}{dz} = \sinh(z), \quad \frac{d \sinh(z)}{dz} = \cosh(z)$$

Other key properties of \cosh and \sinh :

- $\cosh^2(z) - \sinh^2(z) = 1$
- For real x , $\cosh(x)$ is real and positive, $\sinh(x)$ is real.

- $\cosh(iz) = \cos(z)$, $\sinh(z) = -i \sin(iz)$.

10. $\log(z)$ (See Topic 1.)

Definition. $\log(z) = \log(|z|) + i \arg(z)$, where a branch of $\arg(z)$ is specified.

Domain: \mathbf{C} minus a branch cut from 0 to ∞ .

$$\frac{d}{dz} \log(z) = \frac{1}{z}$$

11. z^a (any complex a) (See Topic 1.)

Definition. $z^a = e^{a \log(z)}$, where a branch of $\log(z)$ is specified.

Domain: Generally the domain is \mathbf{C} minus a branch cut of \log . If a is an integer ≥ 0 then z^a is entire. If a is a negative integer then z^a is defined and analytic on $\mathbf{C} = \{0\}$.

$$\frac{dz^a}{dz} = az^{a-1}.$$

2.7.2 Proofs

Here we prove at least some of the facts stated in the list just above.

1. $f(z) = e^z$. This was done in Example 2.9 using the Cauchy-Riemann equations.

2. $f(z) \equiv c$ (constant). This case is trivial.

3. $f(z) = z^n$ (n an integer ≥ 0): show $f'(z) = nz^{n-1}$

It's probably easiest to use the definition of derivative directly. Before doing that we note the factorization

$$z^n - z_0^n = (z - z_0)(z^{n-1} + z^{n-2}z_0 + z^{n-3}z_0^2 + \dots + z^2z_0^{n-3} + zz_0^{n-2} + z_0^{n-1})$$

Now

$$\begin{aligned} f'(z_0) &= \lim_{z \rightarrow z_0} \frac{f(z) - f(z_0)}{z - z_0} = \lim_{z \rightarrow z_0} \frac{z^n - z_0^n}{z - z_0} \\ &= \lim_{z \rightarrow z_0} (z^{n-1} + z^{n-2}z_0 + z^{n-3}z_0^2 + \dots + z^2z_0^{n-3} + zz_0^{n-2} + z_0^{n-1}) \\ &= nz_0^{n-1} \quad QED. \end{aligned}$$

Since we showed directly that the derivative exists for all z , the function is must be entire.

4. $P(z)$ (polynomial). Since a polynomial is a sum of monomials, the formula for the derivative follows from the derivative rule for sums and the case $f(z) = z^n$. Likewise the fact the $P(z)$ is entire.

5. $f(z) = 1/z$. This follows from the quotient rule.

6. $f(z) = P(z)/Q(z)$. This also follows from the quotient rule.

7. $\sin(z)$, $\cos(z)$. All the facts about $\sin(z)$ and $\cos(z)$ follow from their definition in terms of exponentials.

8. Other trig functions $\cot(z)$, $\sec(z)$ etc. Since these are all defined in terms of \cos and \sin , all the facts about these functions follow from the derivative rules.
9. $\sinh(z)$, $\cosh(z)$. All the facts about $\sinh(z)$ and $\cosh(z)$ follow from their definition in terms of exponentials.
10. $\log(z)$. The derivative of $\log(z)$ can be found by differentiating the relation $e^{\log(z)} = z$ using the chain rule. Let $w = \log(z)$, so $e^w = z$ and

$$\begin{aligned} \frac{d}{dz} e^w &= \frac{dz}{dz} = 1. \\ \text{So, } \frac{de^w}{dw} \frac{dw}{dz} &= 1. \\ \text{So, } e^w \frac{dw}{dz} &= 1 \Rightarrow \frac{dw}{dz} = \frac{1}{e^w} \end{aligned}$$

Using $w = \log(z)$ we get

$$\frac{d \log(z)}{dz} = \frac{1}{z}. \quad QED$$

11. z^a (any complex a). The derivative for this follows from the formula

$$\begin{aligned} z^a &= e^{a \log(z)} \\ \Rightarrow \frac{dz^a}{dz} &= e^{a \log(z)} \cdot \frac{a}{z} = \frac{az^a}{z} = az^{a-1} \end{aligned}$$

2.8 Branch cuts and function composition

We often compose functions, i.e. $f(g(z))$. In general in this case we have the chain rule to compute the derivative. However we need to specify the domain for z where the function is analytic. And when branches and branch cuts are involved we need to take care.

Example 2.11. Let $f(z) = e^{z^2}$. Since e^z and z^2 are both entire functions, so is $f(z) = e^{z^2}$. The chain rule gives us

$$f'(z) = e^{z^2}(2z).$$

Example 2.12. Let $f(z) = e^z$ and $g(z) = 1/z$. $f(z)$ is entire and $g(z)$ is analytic everywhere but 0. So $f(g(z))$ is analytic except at 0 and

$$\frac{df(g(z))}{dz} = f'(g(z))g'(z) = e^{1/z} \cdot \frac{-1}{z^2}.$$

Example 2.13. Let $h(z) = 1/(e^z - 1)$. Clearly h is entire except where the denominator is 0. The denominator is 0 when $e^z - 1 = 0$. That is, when $z = 2\pi ni$ for any integer n . Thus, $h(z)$ is analytic on the set

$$\mathbf{C} - \{2\pi ni, \text{ where } n \text{ is any integer}\}$$

The quotient rule gives $h'(z) = -e^z/(e^z - 1)^2$. A little more formally: $h(z) = f(g(z))$, where $f(w) = 1/w$ and $w = g(z) = e^z - 1$. We know that $g(z)$ is entire and $f(w)$ is analytic everywhere except $w = 0$. Therefore, $f(g(z))$ is analytic everywhere except where $g(z) = 0$.

Example 2.14. It can happen that the derivative has a larger domain where it is analytic than the original function. The main example is $f(z) = \log(z)$. This is analytic on C minus a branch cut. However $\frac{d}{dz} \log(z) = \frac{1}{z}$ is analytic on $\mathbf{C} - \{0\}$. [The converse can't happen.](#)

Example 2.15. Define a region where $\sqrt{1-z}$ is analytic.

answer: Choosing the principal branch of argument, we have \sqrt{w} is analytic on

$$\mathbf{C} - \{x \leq 0, y = 0\}, \text{ (see figure below.)}$$

So $\sqrt{1-z}$ is analytic except where $w = 1-z$ is on the branch cut, i.e. where $w = 1-z$ is real and ≤ 0 . It's easy to see that

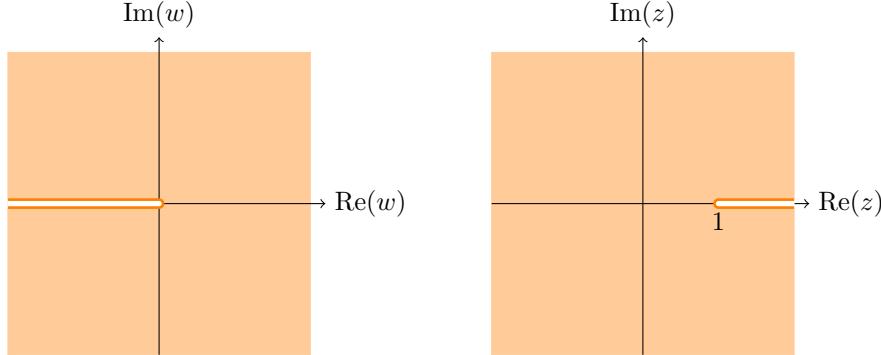
$$w = 1-z \text{ is real and } \leq 0 \Leftrightarrow z \text{ is real and } \geq 1.$$

So $\sqrt{1-z}$ is analytic on the region (see figure below)

$$\mathbf{C} - \{x \geq 1, y = 0\}$$

Note. A different branch choice for \sqrt{z} would lead to a different region where $\sqrt{1-z}$ is analytic.

The figure below shows the domains with branch cuts for this example.



Left: domain for \sqrt{w} , right: domain for $\sqrt{1-z}$

Example 2.16. Define a region where $f(z) = \sqrt{1+e^z}$ is analytic.

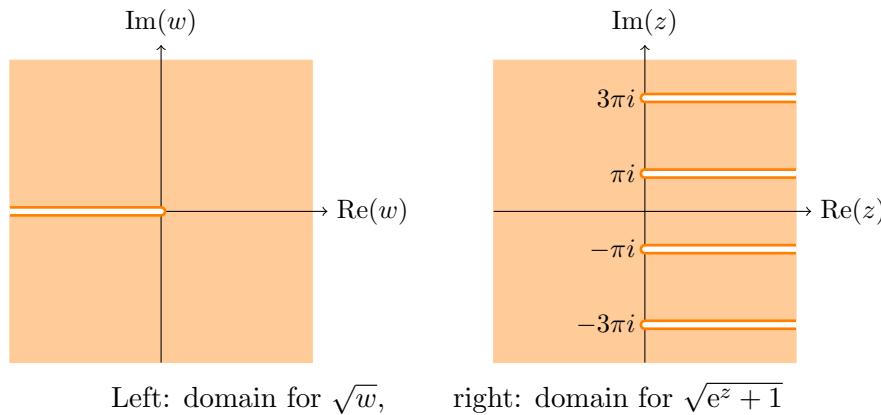
answer: Again, let's take \sqrt{w} to be analytic on the region

$$\mathbf{C} - \{x \leq 0, y = 0\}$$

So, $f(z)$ is analytic except where $1+e^z$ is real and ≤ 0 . That is, except where e^z is real and ≤ -1 . Now, $e^z = e^x e^{iy}$ is real only when y is a multiple of π . It is negative only when y is an odd multiple of π . It has magnitude greater than 1 only when $x > 0$. Therefore $f(z)$ is analytic on the region

$$\mathbf{C} - \{x \geq 0, y = \text{odd multiple of } \pi\}$$

The figure below shows the domains with branch cuts for this example.



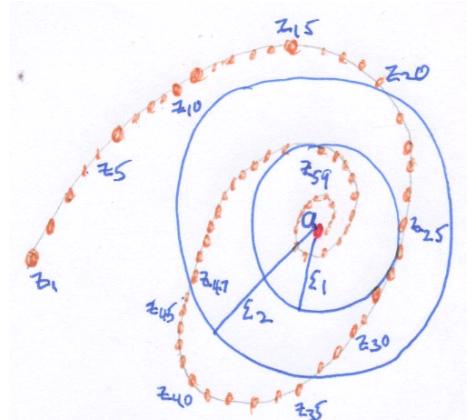
2.9 Appendix Limits

The intuitive idea behind limits is relatively simple. Still, in the 19th century mathematicians were troubled by the lack of rigor, so they set about putting limits and analysis on a firm footing with careful definitions and proofs. In this appendix we give you the formal definition and connect it to the intuitive idea. In 18.04 we will not need this level of formality. Still, it's nice to know the foundations are solid, and some students may find this interesting.

2.9.1 Limits of sequences

Intuitively, we say a sequence of complex numbers z_1, z_2, \dots converges to a if for large n , z_n is really close to a . To be a little more precise, if we put a small circle of radius ϵ around a then eventually the sequence should stay inside the circle. Let's refer to this as [the sequence being captured by the circle](#). This has to be true for any circle no matter how small, though it may take longer for the sequence to be 'captured' by a smaller circle.

This is illustrated in the figure below. The sequence is strung along the curve shown heading towards a . The bigger circle of radius ϵ_2 captures the sequence by the time $n = 47$, the smaller circle doesn't capture it till $n = 59$. Note that z_{25} is inside the larger circle, but since later points are outside the circle we don't say the sequence is captured at $n = 25$.

A sequence of points converging to a

Definition. The sequence z_1, z_2, z_3, \dots converges to the value a if for every $\epsilon > 0$ there is a number N_ϵ such that $|z_n - a| < \epsilon$ for all $n > N_\epsilon$. We write this as

$$\lim_{n \rightarrow \infty} z_n = a.$$

Again, the definition just says that eventually the sequence is within ϵ of a , no matter how small you choose ϵ .

Example 2.17. Show that the sequence $z_n = (1/n + i)^2$ has limit -1.

answer: This is clear because $1/n \rightarrow 0$. For practice, let's phrase it in terms of epsilons: given $\epsilon > 0$ we have to choose N_ϵ such that

$$|z_n - (-1)| < \epsilon \text{ for all } n > N_\epsilon$$

One strategy is to look at $|z_n + 1|$ and see what N_ϵ should be. We have

$$|z_n - (-1)| = \left| \left(\frac{1}{n} + i \right)^2 + 1 \right| = \left| \frac{1}{n^2} + \frac{2i}{n} \right| < \frac{1}{n^2} + \frac{2}{n}$$

So all we have to do is pick N_ϵ large enough that

$$\frac{1}{N_\epsilon^2} + \frac{2}{N_\epsilon} < \epsilon$$

Since this can clearly be done we have proved that $z_n \rightarrow i$.

This was clearly more work than we want to do for every limit. Fortunately, most of the time we can apply general rules to determine a limit without resorting to epsilons!

- Notes.**
1. In 18.04 we will be able to spot the limit of most concrete examples of sequences. The formal definition is needed when dealing abstractly with sequences.
 2. To mathematicians ϵ is one of the go-to symbols for a small number. The prominent and rather eccentric mathematician Paul Erdos used to refer to children as epsilons, as in ‘How are the epsilons doing?’
 3. The term ‘captured by the circle’ is not in common usage, but it does capture what is happening.

2.9.2 Limits of functions

Sometimes we need limits of the form $\lim_{z \rightarrow z_0} f(z) = a$. Again, the intuitive meaning is clear: as z gets close to z_0 we should see $f(z)$ get close to a . Here is the technical definition

Definition. Suppose $f(z)$ is defined on a punctured disk $0 < |z - z_0| < r$ around z_0 . We say $\lim_{z \rightarrow z_0} f(z) = a$ if for every $\epsilon > 0$ there is a δ such that

$$|f(z) - a| < \epsilon \text{ whenever } 0 < |z - z_0| < \delta$$

This says exactly that as z gets closer (within δ) to z_0 we have $f(z)$ is close (within ϵ) to a . Since ϵ can be made as small as we want, $f(z)$ must go to a .

Notes. 1. Using the punctured disk (also called a [deleted neighborhood](#)) means that $f(z)$ does not have to be defined at z_0 and, if it is then $f(z_0)$ does not necessarily equal a . If $f(z_0) = a$ then we say the f is continuous at z_0 .

2. Ask any mathematician to complete the phrase “For every ϵ ” and the odds are that they will respond “there is a $\delta \dots$ ”

2.9.3 Connection between limits of sequences and limits of functions

Here's an equivalent way to define limits of functions: the limit $\lim_{z \rightarrow z_0} f(z) = a$ if for every sequence of points $\{z_n\}$ with limit z_0 the sequence $\{f(z_n)\}$ has limit a .

3 Line integrals and Cauchy's theorem

3.1 Introduction

The basic theme here is that complex line integrals will mirror much of what we've seen for multivariable calculus line integrals. But, just like working with $e^{i\theta}$ is easier than working with sine and cosine, complex line integrals are easier to work with than their multivariable analogs. At the same time they will give deep insight into the workings of these integrals.

3.2 Ingredients

- The complex plane: $z = x + iy$
- The complex differential $dz = dx + idy$
- A curve in the complex plane: $\gamma(t) = x(t) + iy(t)$, defined for $a \leq t \leq b$.
- A complex function: $f(z) = u(x, y) + iv(x, y)$

3.3 Complex line integrals

Note. Line integrals are also called [path or contour](#) integrals.

Given the ingredients we define the [complex line integral](#) $\int_{\gamma} f(z) dz$ by

$$\int_{\gamma} f(z) dz = \int_a^b f(\gamma(t)) \gamma'(t) dt. \quad (4)$$

You should note that this notation looks just like integrals of a real variable. We don't need the vectors and dot products of line integrals in \mathbf{R}^2 . Also, make sure you understand that the product $f(\gamma(t))\gamma'(t)$ is just a product of complex numbers.

An alternative notation uses $dz = dx + idy$ to write

$$\int_{\gamma} f(z) dz = \int_{\gamma} (u + iv)(dx + idy) \quad (5)$$

Let's check that Equations 4 and 5 are the same. Equation 5 is really a multivariable calculus expression, so thinking of $\gamma(t)$ as $(x(t), y(t))$ it becomes

$$\int_{\gamma} f(z) dz = \int_a^b (u(x(t), y(t)) + iv(x(t), y(t))(x'(t) + iy'(t))dt$$

But, $u(x(t), y(t)) + iv(x(t), y(t)) = f(\gamma(t))$ and $x'(t) + iy'(t) = \gamma'(t)$ so the right hand side of this equation is

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

That is, it is exactly the same as the expression in Equation 4.

Example 3.1. Compute $\int_{\gamma} z^2 dz$ along the straight line from 0 to $1+i$.

answer: We parametrize the curve as $\gamma(t) = t(1+i)$ with $0 \leq t \leq 1$. So $\gamma'(t) = 1+i$. The line integral is

$$\int z^2 dz = \int_0^1 t^2 (1+i)^2 (1+i) dt = \frac{2i(1+i)}{3}.$$

Example 3.2. Compute $\int_{\gamma} \bar{z} dz$ along the straight line from 0 to $1+i$.

answer: We can use the same parametrization as in the previous example. So,

$$\int_{\gamma} \bar{z} dz = \int_0^1 t(1-i)(1+i) dt = 1.$$

Example 3.3. Compute $\int_{\gamma} z^2 dz$ along the unit circle.

answer: We parametrize the unit circle by $\gamma(\theta) = e^{i\theta}$, where $0 \leq \theta \leq 2\pi$. We have $\gamma'(\theta) = ie^{i\theta}$. So, the integral becomes

$$\int_{\gamma} z^2 dz = \int_0^{2\pi} e^{2i\theta} ie^{i\theta} d\theta = \int_0^{2\pi} ie^{3i\theta} d\theta = \frac{e^{i3\theta}}{3} \Big|_0^{2\pi} = 0.$$

3.4 Fundamental theorem for complex line integrals

This is exactly analogous to the fundamental theorem of calculus.

Theorem. (Fundamental theorem of complex line integrals)

If $f(z)$ is a complex analytic function and γ is a curve from z_0 to z_1 then

$$\int_{\gamma} f'(z) dz = f(z_1) - f(z_0).$$

Proof. This is an application of the chain rule. We have

$$\frac{df(\gamma(t))}{dt} = f'(\gamma(t)) \gamma'(t).$$

So

$$\int_{\gamma} f'(z) dz = \int_a^b f'(\gamma(t)) \gamma'(t) dt = \int_a^b \frac{df(\gamma(t))}{dt} dt = f(\gamma(t))|_a^b = f(z_1) - f(z_0).$$

Another way to state the fundamental theorem is: if f has an antiderivative F , i.e. $F' = f$ then

$$\int_{\gamma} f(z) dz = F(z_1) - F(z_0).$$

Example 3.4. Redo $\int_{\gamma} z^2 dz$, with γ the straight line from 0 to $1+i$.

answer: We can check by inspection that z^2 has an antiderivative $F(z) = z^3/3$. Therefore the fundamental theorem implies

$$\int_{\gamma} z^2 dz = \frac{z^3}{3} \Big|_0^{1+i} = \frac{(1+i)^3}{3} = \frac{2i(1+i)}{3}.$$

Example 3.5. Redo $\int_{\gamma} z^2 dz$, with γ the unit circle.

answer: Again, since z^2 had antiderivative $z^3/3$ we can evaluate the integral by plugging the endpoints of γ into the $z^3/3$. Since the endpoints are the same the resulting difference will be 0!

3.5 Cauchy-Riemann all the way down

We've defined an analytic function as one having a complex derivative. The following theorem shows that if f is analytic then so is f' . Thus, there are derivatives all the way down!

Theorem. If $f(z)$ is analytic then so is $f'(z)$.

Proof. To show this we have to prove that $f'(z)$ satisfies the Cauchy-Riemann equations. If $f = u + iv$ we know

$$u_x = v_y; \quad u_y = -v_x; \quad f' = u_x + iv_x.$$

Let's write $f' = U + iV$, so $U = u_x = v_y$ and $V = v_x = -u_y$. We want to show that $U_x = V_y$ and $U_y = -V_x$. We do them one at a time.

$U_x = V_y$: $U_x = v_{yx}$ and $V_y = v_{xy}$. Since $v_{xy} = v_{yx}$, we have $U_x = V_y$.

$U_y = -V_x$: $U_y = u_{xy}$ and $V_x = -u_{yx}$. So, $U_y = -V_x$. QED.

Technical point. We've assumed as many partials as we need. So far we can't guarantee that all the partials exist. Soon we will have a theorem which says that an analytic function has derivatives of all order. We'll just assume that for now. In any case, in most examples this will be obvious.

3.6 Path independence

We say the integral $\int_{\gamma} f(z) dz$ is path independent if it has the same value for any two paths with the same endpoints.

More precisely, if $f(z)$ is defined on a region A then $\int_{\gamma} f(z) dz$ is path independent in D , if it has the same value for any two paths in D with the same endpoints.

Theorem. If $f(z)$ has an antiderivative then the path integral $\int_{\gamma} f(z) dz$ is path independent.

Proof. This is the same argument as in Example 3.5. Since $f(z)$ has an antiderivative of $f(z)$, the fundamental theorem tells us that the integral only depends on the endpoint of γ , i.e.

$$\int_{\gamma} f(z) dz = F(z_1) - F(z_0)$$

where z_0 and z_1 are the beginning and end point of γ .

An alternative way to express path independence uses closed paths.

Theorem 3.6. The following two things are equivalent.

- (1) The integral $\int_{\gamma} f(z) dz$ is path independent.
- (2) The integral $\int_{\gamma} f(z) dz$ around any closed path is 0.

Proof. This is essentially identical to the equivalent multivariable proof. We have to show two things:

- (i) Path independence implies the line integral around any closed path is 0.
- (ii) If the line integral around all closed paths is 0 then we have path independence.

To see (i), assume path independence and consider the closed path C shown in figure (i) below. Since the starting point z_0 is the same as the endpoint z_1 the line integral $\int_C f(z) dz$ must have the same value as the line integral over the curve consisting of the single point z_0 . Since that is clearly 0 we must have the integral over C is 0.

To see (ii), assume $\int_C f(z) dz = 0$ for any closed curve. Consider the two curves C_1 and

C_2 shown in figure (ii). Both start at z_0 and end at z_1 . By the assumption that integrals over closed paths are 0 we have $\int_{C_1-C_2} f(z) dz = 0$. So,

$$\int_{C_1} f(z) dz = \int_{C_2} f(z) dz.$$

That is, any two paths from z_0 to z_1 have the same line integral. This shows that the line integrals are path independent.

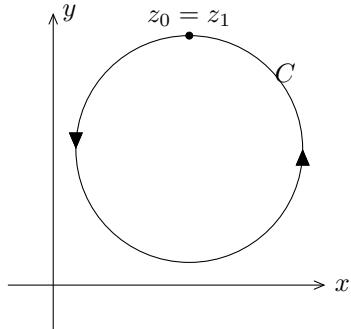


Figure (i)

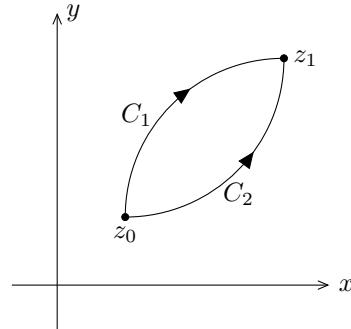


Figure (ii)

3.7 Examples

Example 3.7. Why can't we compute $\int_{\gamma} \bar{z} dz$ using the fundamental theorem.

answer: Because \bar{z} doesn't have an antiderivative.

Example 3.8. Compute $\int_{\gamma} 1/z dz$ over several contours

- (i) The line from 1 to $1 + i$.
- (ii) The circle of radius 1 around $z = 3$.
- (iii) The unit circle.

answer: For parts (i) and (ii) there is no problem using the antiderivative $\log(z)$ because these curves are contained in a simply connected region that doesn't contain origin.

(i) **answer:** $\log(1 + i) - \log(1) = \log(\sqrt{2}) + i\pi/4$.

(ii) **answer:** 0 (since the beginning and end points are the same).

(iii) **answer:** We parametrize the unit circle by $\gamma(\theta) = e^{i\theta}$ with $0 \leq \theta \leq 2\pi$. We compute $\gamma'(\theta) = ie^{i\theta}$. So the integral becomes

$$\int_{\gamma} \frac{1}{z} dz = \int_0^{2\pi} \frac{1}{e^{i\theta}} ie^{i\theta} dt = \int_0^{2\pi} i dt = 2\pi i.$$

Notice that we could use $\log(z)$ if we were careful to let it the argument increase by 2π as it went around the origin once.

Example 3.9. Compute $\int_{\gamma} \frac{1}{z^2} dz$ around unit circle in two ways.

- (i) Using the fundamental theorem
- (ii) Directly from the definition.

answer: 0. (You should supply the details for this.)

3.8 Cauchy's theorem

Cauchy's theorem is analogous to Green's theorem for curl free vector fields.

Theorem. (Cauchy's theorem) Suppose A is a simply connected region, $f(z)$ is analytic on A and C is a simple closed curve in A . Then the following three things hold:

$$(i) \int_C f(z) dz = 0$$

(i') We can drop the requirement that C is simple in part (i).

(ii) Integrals of f on paths within A are path independent. That is, two paths with same endpoints integrate to the same value.

(iii) f has an antiderivative in A .

Proof. We will prove (i) using Green's theorem, We could give a proof that didn't rely on Green's, but it would be quite similar in flavor to the proof of Green's theorem.

Let R be the region inside the curve. And write $f = u + iv$. Now we write out the integral as follows

$$\int_C f(z) dz = \int_C (u + iv) (dx + idy) = \int_C (u dx - v dy) + i(v dx + u dy).$$

Let's apply Green's theorem to the real and imaginary pieces separately. First the real piece.

$$\int_C u dx - v dy = \int_R -v_x - u_y dx dy = 0.$$

We get 0 because the Cauchy-Riemann equations say $u_y = -v_x$, so $-v_x - u_y = 0$.

Likewise for the imaginary piece:

$$\int_C v dx + u dy = \int_R u_x - v_y dx dy = 0.$$

We get 0 because the Cauchy-Riemann equations say $u_x = v_y$, so $u_x - v_y = 0$.

To see part (i') you should draw a few curves that intersect themselves and convince yourself that they can be broken into a sum of simple closed curves. Thus, (i') follows from (i).

Note: In order to truly prove part (i') we would need a more technically precise definition of simply connected so we could say that all closed curves within A can be continuously deformed to each other.

Part (ii) follows from (i) and Theorem 3.6

To see (iii), pick a base point $z_0 \in A$ and let

$$F(z) = \int_{z_0}^z f(w) dw.$$

Here the integral is over any path in A connecting z_0 to z . By part (ii), $F(z)$ is well defined. If we can show that $F'(z) = f(z)$ then we'll be done. Doing this amounts to managing the notation to apply the fundamental theorem of calculus and the Cauchy-Riemann equations.

Let's write $f(z) = u(x, y) + iv(x, y)$ and $F(z) = U(x, y) + iV(x, y)$. Then we can write

$$\frac{\partial f}{\partial x} = u_x + iv_x, \text{ etc.}$$

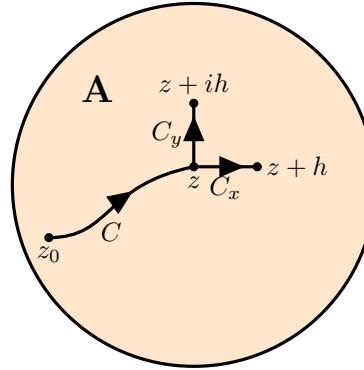
We can formulate the Cauchy-Riemann equations for $F(z)$ as

$$F'(z) = \frac{\partial F}{\partial x} = \frac{1}{i} \frac{\partial F}{\partial y}, \text{ i.e. } F'(z) = U_x + iV_x = \frac{1}{i}(U_y + iV_y) = V_y - iU_y. \quad (6)$$

For reference we note that using the path $\gamma(t) = x(t) + iy(t)$, with $\gamma(0) = z_0$ and $\gamma(b) = z$ we have

$$F(z) = \int_{z_0}^z f(z) dz = \int_{z_0}^z (u(x, y) + iv(x, y))(dx + idy) = \int_0^b (u(x(t), y(t)) + iv(x(t), y(t))(x'(t) + iy'(t)) dt. \quad (7)$$

Our goal now is to prove that the Cauchy-Riemann equations given in Equation 7 hold for $F(z)$. The figure below shows an arbitrary path from z_0 to z , which can be used to compute $F(z)$. To compute the partials of F we'll need the straight lines that continue C to $z + h$ or $z + ih$.



Paths for proof of Cauchy's theorem

First we'll look at $\frac{\partial F}{\partial x}$. So, fix $z = x + iy$. Looking at the paths in the figure above we have

$$F(z + h) - F(z) = \int_{C+C_x} f(w) dw - \int_C f(w) dw = \int_{C_x} f(w) dw.$$

The curve C_x is parametrized by $\gamma(t) = x + t + iy$, with $0 \leq t \leq h$. So,

$$\begin{aligned} \frac{\partial F}{\partial x} &= \lim_{h \rightarrow 0} \frac{F(z + h) - F(z)}{h} = \lim_{h \rightarrow 0} \frac{\int_{C_x} f(w) dw}{h} \\ &= \lim_{h \rightarrow 0} \frac{\int_0^h u(x + t, y) + iv(x + t, y) dt}{h} = u(x, y) + iv(x, y) = f(z). \end{aligned} \quad (8)$$

(The second to last equality uses the fundamental theorem of calculus.)

Similarly, we get

$$\begin{aligned} \frac{1}{i} \frac{\partial F}{\partial y} &= \lim_{h \rightarrow 0} \frac{F(z + ih) - F(z)}{ih} = \lim_{h \rightarrow 0} \frac{\int_{C_y} f(w) dw}{ih} \\ &= \lim_{h \rightarrow 0} \frac{\int_0^h u(x, y + t) + iv(x, y + t) dt}{h} = u(x, y) + iv(x, y) = f(z). \end{aligned} \quad (9)$$

Together Equations 8 and 9 show

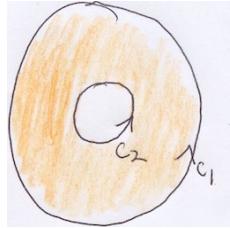
$$f(z) = \frac{\partial F}{\partial x} = \frac{1}{i} \frac{\partial F}{\partial y}$$

By Equation 6 we have shown that F is analytic and $F' = f$. QED

3.9 Extensions of Cauchy's theorem

Cauchy's theorem requires that the function $f(z)$ be analytic on a simply connected region. In cases where it is not, we can extend it in a useful way.

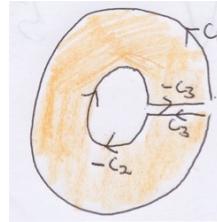
Suppose R is the region between the two simple closed curves C_1 and C_2 .



Theorem. (Extended Cauchy's theorem) If $f(z)$ is analytic on R then

$$\int_{C_1} f(z) dz = \int_{C_2} f(z) dz$$

Proof. The proof is based on the following figure. We 'cut' both C_1 and C_2 and connect them by two copies of C_3 , one in each direction. (In the figure we have drawn the two copies of C_3 as separate curves, in reality they are the same curve traversed in opposite directions.)



With C_3 acting as a cut, the region enclosed by $C_1 + C_3 - C_2 - C_3$ is simply connected, so Cauchy's theorem applies. We get

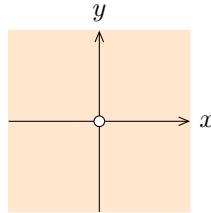
$$\int_{C_1+C_3-C_2-C_3} f(z) dz = 0$$

Breaking the integral up we have

$$\int_{C_1} f(z) dz + \int_{C_3} f(z) dz - \int_{C_2} f(z) dz - \int_{C_3} f(z) dz = 0.$$

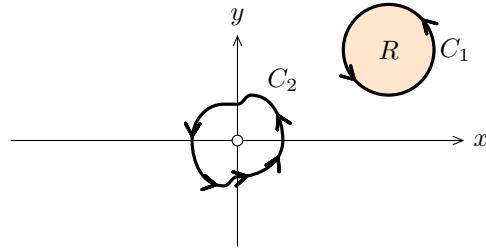
This clearly implies $\int_{C_1} f(z) dz = \int_{C_2} f(z) dz$ as claimed.

Example 3.10. Let $f(z) = 1/z$. $f(z)$ is clearly defined and analytic on the punctured plane.

Punctured plane: $\mathbf{C} - \{0\}$

Question: What values can $\int_C f(z) dz$ take for C a simple closed curve (positively oriented) in the plane?

answer: We have two cases (i) C_1 not around 0 (ii) C_2 around 0



In case (i) Cauchy's theorem applies directly because the interior does not contain the problem point at the origin. Thus,

$$\int_{C_1} f(z) dz = 0.$$

For case (ii) we will show that $\int_{C_2} f(z) dz = 2\pi i$.

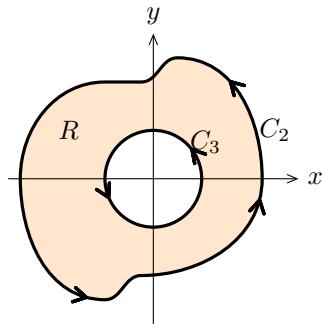


Figure for part (ii)

Let C_3 be a small circle of radius a centered at 0 and entirely inside C_2 . By the extended Cauchy theorem we have

$$\int_{C_2} f(z) dz = \int_{C_3} f(z) dz.$$

Thus, $\oint_{C_2} f(z) dz = \oint_{C_3} f(z) dz$.

Using the usual parametrization of a circle we can easily compute that the line integral is

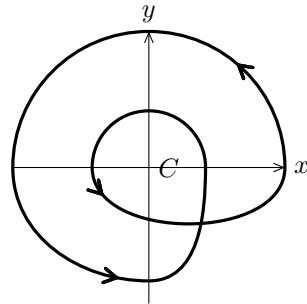
$$\int_{C_3} f(z) dz = \int_0^{2\pi} i dt = 2\pi i. \quad QED.$$

Answer to the question: The only possible values are 0 and $2\pi i$.

We can extend this answer in the following way:

If C is not simple, then the possible values of $\int_C f(z) dz$ are $2\pi n i$, where n is the number of times C goes (counterclockwise) around the origin 0.

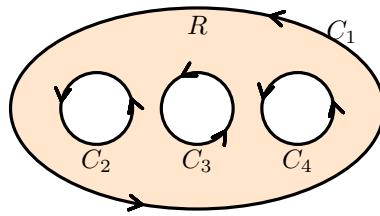
Definition. n is called the *winding number* of C around 0. (n also equals the number of times C crosses the positive x -axis, counting +1 for crossing from below and -1 for crossing from above.)



A curve with winding number 2 around the origin.

Example 3.11. A further extension: using the same trick of cutting the region by curves to make it simply connected we can show that if f is analytic in the region R shown below then

$$\int_{C_1-C_2-C_3-C_4} f(z) dz = 0.$$



4 Cauchy's integral formula

4.1 Introduction

Cauchy's theorem is a big theorem which we will use almost daily from here on out. Right away it will reveal a number of interesting and useful properties of analytic functions. More will follow as the course progresses.

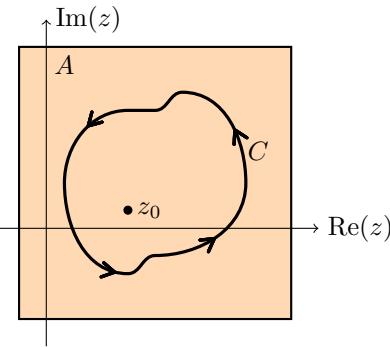
If you learn just one theorem this week it should be Cauchy's integral formula!

We start with a statement of the theorem for functions and some examples. Then we give the generalization to all of its derivatives. After some more examples we will prove the theorems. After that we will see some remarkable consequences that follow fairly directly from the theorem.

4.2 Cauchy's integral for functions

Theorem 4.1. (Cauchy's integral formula) Suppose C is a simple closed curve and the function $f(z)$ is analytic on a region containing C and its interior. We assume C is oriented counterclockwise. Then for any z_0 inside C :

$$f(z_0) = \frac{1}{2\pi i} \int_C \frac{f(z)}{z - z_0} dz \quad (10)$$



Cauchy's integral formula: simple closed curve C , $f(z)$ analytic on and inside C .

This is remarkable: it says that knowing the values of f on the boundary curve C means we know **everything** about f inside C !! This is probably unlike anything you've encountered with functions of real variables.

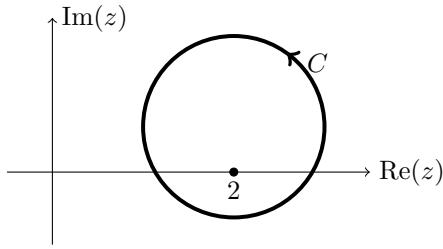
Aside 1. With a slight change of notation (z becomes w and z_0 becomes z) we often write the formula as

$$f(z) = \frac{1}{2\pi i} \int_C \frac{f(w)}{w - z} dw \quad (11)$$

Aside 2. We're not being entirely fair to functions of real variables. We will see that for $f = u + iv$ the parts u and v have many similar remarkable properties. u and v are called conjugate harmonic functions.

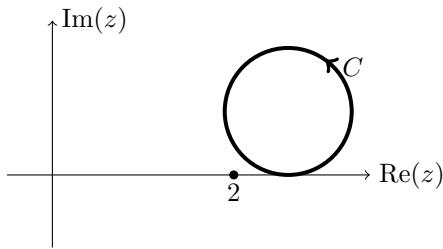
4.2.1 Examples

Example 4.2. Compute $\int_C \frac{e^{z^2}}{z - 2} dz$, where C is the curve shown.



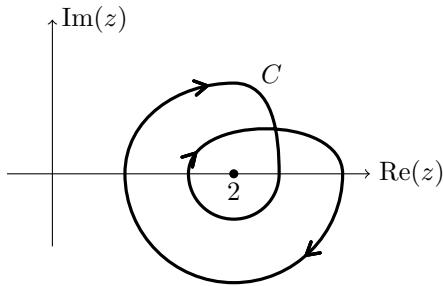
answer: Let $f(z) = e^{z^2}$. $f(z)$ is entire. Since C is a simple closed curve (counterclockwise) and $z = 2$ is inside C , Cauchy's integral formula says the the integral is $2\pi i f(2) = 2\pi i e^4$.

Example 4.3. Do the same integral as the previous example with C the curve shown.



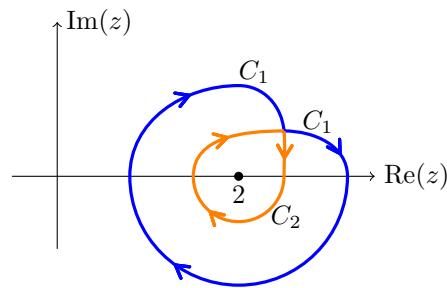
answer: Since $f(z) = e^{z^2}/(z - 2)$ is analytic on and inside C , Cauchy's theorem says the integral is 0.

Example 4.4. Do the same integral as the previous examples with C the curve shown.



answer: This one is trickier. Let $f(z) = e^{z^2}$. The curve C goes around $z = 2$ twice in the *clockwise* direction. Here's the answer, detailed reasoning will follow: the value of the integral is $-4\pi i f(2) = -4\pi i e^4$.

Reasoning: C loops around $z = 2$ twice, so we break C into $C_1 + C_2$ as shown. C_1 is the outer loop and C_2 is the inner loop.



These are both simple closed curves, so we can apply the Cauchy integral formula to each separately. (The negative signs are because they go clockwise around $z = 2$.)

$$\int_C \frac{f(z)}{z-2} dz = \int_{C_1} \frac{f(z)}{z-2} dz + \int_{C_2} \frac{f(z)}{z-2} dz = -2\pi i f(2) - 2\pi i f(2) = -4\pi i f(2). \text{ QED}$$

4.3 Cauchy's integral formula for derivatives

Cauchy's integral formula is worth repeating several times. So, now we give it for all derivatives $f^{(n)}(z)$ of f . This will include the formula for functions as a special case.

Theorem 4.5. *Cauchy's integral formula for derivatives.* If $f(z)$ and C satisfy the same hypotheses as for Cauchy's integral formula then, for all z inside C we have

$$f^{(n)}(z) = \frac{n!}{2\pi i} \int_C \frac{f(w)}{(w-z)^{n+1}} dw, \quad n = 0, 1, 2, \dots \quad (12)$$

where, C is a simple closed curve, oriented counterclockwise, z is inside C and $f(w)$ is analytic on and inside C .

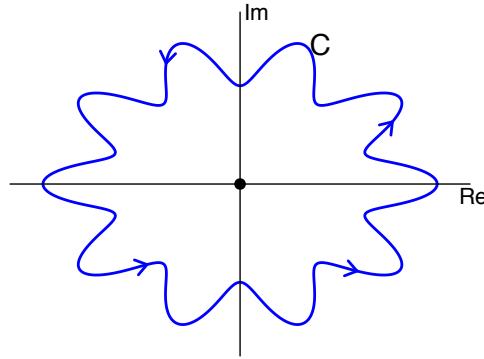
4.3.1 Examples

Example 4.6. Evaluate $I = \int_C \frac{e^{2z}}{z^4} dz$ where $C : |z| = 1$.

answer: With Cauchy's formula for derivatives this is easy. Let $f(z) = e^{2z}$. Then,

$$I = \int_C \frac{f(z)}{z^4} dz = \frac{2\pi i}{3!} f'''(0) = \frac{8}{3}\pi i.$$

Example 4.7. Now Let C be the contour shown below and evaluate the same integral as in the previous example.



answer: Again this is easy: the integral is the same as the previous example, i.e. $I = \frac{8}{3}\pi i$.

4.3.2 Another approach to some standard integrals

Suppose C is a simple closed curve around 0. We have seen that $\int_C \frac{1}{z} dz = 2\pi i$. The Cauchy integral formula gives the same result. That is, let $f(z) = 1$, then the formula says

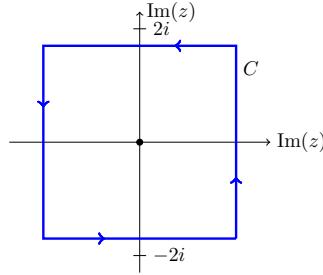
$$\frac{1}{2\pi i} \int_C \frac{f(z)}{z-0} dz = f(0) = 1.$$

Likewise Cauchy's formula for derivatives shows

$$\int_C \frac{1}{(z)^n} dz = \int_C \frac{f(z)}{z^{n+1}} dz = f^{(n)}(0) = 0, \text{ for integers } n > 1.$$

4.3.3 More examples

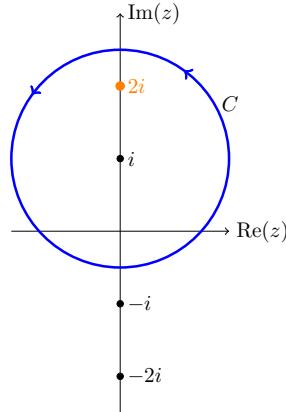
Example 4.8. Compute $\int_C \frac{\cos(z)}{z(z^2 + 8)} dz$ over the contour shown.



answer: Let $f(z) = \cos(z)/(z^2 + 8)$. $f(z)$ is analytic on and inside the curve C . That is, the roots of $z^2 + 8$ are outside the curve. So, we rewrite the integral as

$$\int_C \frac{\cos(z)/(z^2 + 8)}{z} dz = \int_C \frac{f(z)}{z} dz = 2\pi i f(0) = 2\pi i \frac{1}{8} = \frac{\pi i}{4}.$$

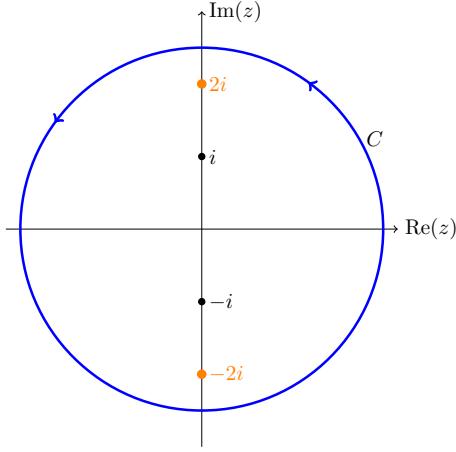
Example 4.9. Compute $\int_C \frac{1}{(z^2 + 4)^2} dz$ over the contour shown.



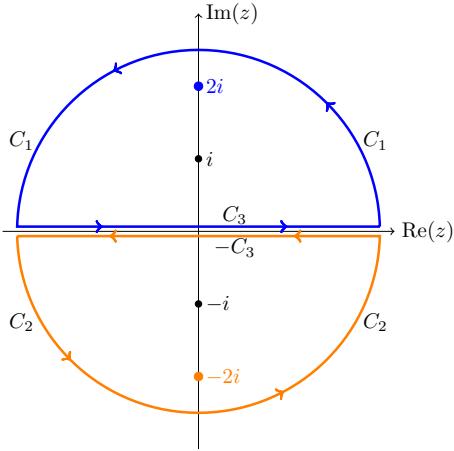
answer: We factor the denominator as $\frac{1}{(z^2 + 4)^2} = \frac{1}{(z - 2i)^2(z + 2i)^2}$. Let $f(z) = 1/(z + 2i)^2$. Clearly $f(z)$ is analytic inside C . So, by Cauchy's formula for derivatives:

$$\int_C \frac{1}{(z^2 + 4)^2} dz = \int_C \frac{f(z)}{(z - 2i)^2} dz = 2\pi i f'(2i) = 2\pi i \left[\frac{-2}{(z + 2i)^3} \right]_{z=2i} = \frac{4\pi i}{64i} = \frac{\pi}{16}$$

Example 4.10. Compute $\int_C \frac{z}{z^2 + 4} dz$ over the curve C shown below.



answer: The integrand has singularities at $\pm 2i$ and the curve C encloses them both. The solution to the previous solution won't work because we can't find an appropriate $f(z)$ that is analytic on the whole interior of C . Our solution is to split the curve into two pieces. Notice that C_3 is traversed both forward and backward.



Split the original curve C into 2 pieces that each surround just one singularity.

We have $\frac{z}{z^2 + 4} = \frac{z}{(z - 2i)(z + 2i)}$. We let $f_1(z) = z/(z + 2i)$ and $f_2(z) = z/(z - 2i)$. So,

$$\frac{z}{z^2 + 4} = \frac{f_1(z)}{z - 2i} = \frac{f_2(z)}{z + 2i}.$$

The integral, can be written out as

$$\int_C \frac{z}{z^2 + 4} dz = \int_{C_1 + C_3 - C_3 + C_2} \frac{z}{z^2 + 4} dz = \int_{C_1 + C_3} \frac{f_1(z)}{z - 2i} dz + \int_{C_2 - C_3} \frac{f_2(z)}{z + 2i} dz$$

Since f_1 is analytic inside the simple closed curve $C_1 + C_3$ and f_2 is analytic inside the simple closed curve $C_2 - C_3$, Cauchy's formula applies to both integrals. The total integral equals

$$2\pi i(f_1(2i) + f_2(-2i)) = 2\pi i(1/2 + 1/2) = 2\pi i.$$

Remarks. 1. We could also have done this problem using partial fractions.

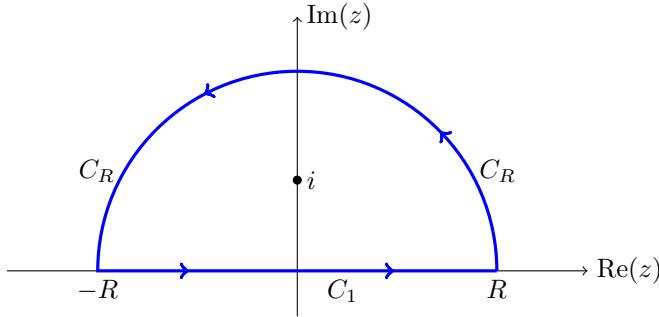
$$\frac{z}{(z - 2i)(z + 2i)} = \frac{A}{z - 2i} + \frac{B}{z + 2i}.$$

It will turn out that $A = f_1(2i)$ and $B = f_2(-2i)$. It is easy to apply the Cauchy integral formula to both terms.

2. In an upcoming topic we will formulate the residue theorem. This will encompass the integrals in Examples 4.8-4.10 and allow us to compute them with less effort.

Example 4.11. Compute the real integral $I = \int_{-\infty}^{\infty} \frac{1}{(x^2 + 1)^2} dx$

answer: The trick is to integrate $f(z) = 1/(z^2 + 1)^2$ over the closed contour $C_1 + C_R$ shown, and then show that the contribution of C_R to this integral vanishes as R goes to ∞ .



The only singularity of $f(z) = \frac{1}{(z+i)^2(z-i)^2}$ inside the contour is at $z = i$. Let $g(z) = 1/(z+i)^2$. Since g is analytic on and inside the contour, Cauchy's formula gives

$$\int_{C_1 + C_R} f(z) dz = \int_{C_1 + C_R} \frac{g(z)}{(z-i)^2} dz = 2\pi i g'(i) = 2\pi i \frac{-2}{(2i)^3} = \frac{\pi}{2}.$$

We parametrize C_1 by $\gamma(x) = x$, with $-R \leq x \leq R$. So

$$\int_{C_1} f(z) dz = \int_{-R}^R \frac{1}{(x^2 + 1)^2} dx.$$

This goes to the I as $R \rightarrow \infty$.

We parametrize C_R by $\gamma(\theta) = Re^{i\theta}$, with $0 \leq \theta \leq \pi$. So

$$\int_{C_R} f(z) dz = \int_0^\pi \frac{1}{(R^2 e^{2i\theta} + 1)^2} i R e^{i\theta} d\theta /$$

By the triangle inequality*, if $R > 1$

$$\left| \int_{C_R} f(z) dz \right| \leq \int_0^\pi \left| \frac{1}{(R^2 e^{2i\theta} + 1)^2} i R e^{i\theta} \right| d\theta \leq \int_0^\pi \frac{R}{(R^2 - 1)^2} d\theta = \frac{\pi R}{(R^2 - 1)^2}$$

Clearly this goes to 0 as R goes to infinity.

Thus, the integral over the contour $C_1 + C_R$ goes to I as R gets large. But this integral always has the same value $\pi/2$. We have shown that $I = \pi/2$.

As a sanity check, we note that our answer is real and positive as it needs to be.

*To expand on this statement: The triangle inequality says $|a + b| \leq |a| + |b|$. Using new notation we can write this as $|c| = |(c + d) - d| \leq |c + d| + |d|$. Rearranging terms we have

$$|c + d| \geq |c| - |d|$$

This is what we used above: $|R^2 e^{2i\theta} + 1| \geq |R^2 e^{2i\theta}| - 1 = R^2 - 1$ Thus $\frac{1}{|R^2 e^{2i\theta} + 1|} \leq \frac{1}{R^2 - 1}$

4.4 Proof of Cauchy's integral formula

4.4.1 Some useful theorems

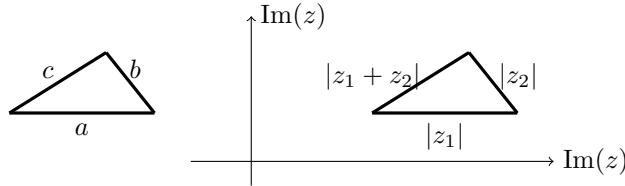
Before proving the theorem we'll need two theorems that will be useful in their own right.

Theorem 4.12. (Triangle inequality) For any function $f(z)$ and any curve γ , we have

$$\left| \int_\gamma f(z) dz \right| \leq \int_\gamma |f(z)| |dz|.$$

Here $dz = \gamma'(t) dt$ and $|dz| = |\gamma'(t)| dt$.

Proof. The name comes from the fact that if a, b, c are the three sides of a triangle then $c \leq a + b$. That is, the sum of the lengths of two sides is greater than the length of the third side.



The triangle inequality.

This translates to magnitudes of complex numbers as shown in the right hand part of the figure (think of complex addition as vector addition). That is,

$$|z_1 + z_2| \leq |z_1| + |z_2|$$

In this form, the triangle inequality says ‘the magnitude of a sum is less than the sum of the magnitudes.’

The lemma says that the magnitude of the integral is less than the integral of the magnitudes. Since an integral is essentially a sum (okay, a limit of sums) the triangle inequality for sums translates immediately to the version given in the theorem.

Corollary. If $|f(z)| < M$ on C then $\left| \int_C f(z) dz \right| \leq M (\text{length of } C)$.

Proof. Let $\gamma(t)$, with $a \leq t \leq b$, be a parametrization of C . Using the triangle inequality

$$\left| \int_C f(z) dz \right| \leq \int_C |f(z)| |dz| = \int_a^b |f(\gamma(t))| |\gamma'(t)| dt \leq \int_a^b M |\gamma'(t)| dt = M \text{length of } C.$$

Here we have used that $|\gamma'(t)| dt = \sqrt{(x')^2 + (y')^2} dt = ds$, the arclength element.

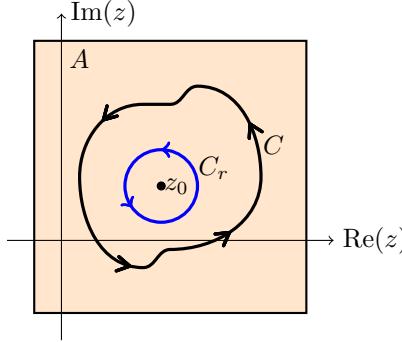
Theorem 4.13. (A second extension of Cauchy's theorem) Suppose that A is a simply connected region containing the point z_0 . Suppose g is a function which is

1. Analytic on $A - \{z_0\}$
2. Continuous on A . (In particular, g does not blow up at z_0 .)

Then, $\int_C g(z) dz = 0$ for all closed curves C in A .

Proof. The extended version of Cauchy's theorem in the topic 3 notes tells us that

$$\int_C g(z) dz = \int_{C_r} g(z) dz, \text{ where } C_r \text{ is a circle of radius } r \text{ around } z_0.$$



Since $g(z)$ is continuous we know that $|g(z)|$ is bounded inside C_r . Say, $|g(z)| < M$. The corollary to the triangle inequality says that

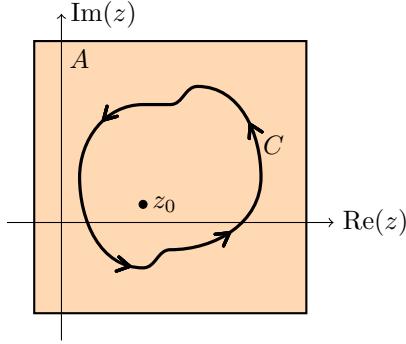
$$\left| \int_{C_r} g(z) dz \right| \leq M (\text{length of } C_r) = M 2\pi r.$$

Since r can be as small as we want, this implies that $\int_{C_r} g(z) dz = 0$. QED

Note. Using this, we can show that $g(z)$ is, in fact, analytic at z_0 . The proof will be the same as in our proof of Cauchy's theorem that $g(z)$ has an antiderivative.

4.4.2 Proof of Cauchy's integral formula

We reiterate Cauchy's integral formula from Equation 10: $f(z_0) = \frac{1}{2\pi i} \int_C \frac{f(z)}{z - z_0} dz$



Proof. (of Cauchy's integral formula) We use a trick that is useful enough to be worth remembering. Let

$$g(z) = \frac{f(z) - f(z_0)}{z - z_0}.$$

Since $f(z)$ is analytic on A , we know that $g(z)$ is analytic on $A - \{z_0\}$. Since the derivative of f exists at z_0 , we know that

$$\lim_{z \rightarrow z_0} g(z) = f'(z_0).$$

That is, if we define $g(z_0) = f'(z_0)$ then g is continuous at z_0 . From the extension of Cauchy's theorem just above, we have

$$\int_C g(z) dz = 0, \text{ i.e. } \int_C \frac{f(z) - f(z_0)}{z - z_0} dz = 0.$$

Thus

$$\int_C \frac{f(z)}{z - z_0} dz = \int_C \frac{f(z_0)}{z - z_0} dz = f(z_0) \int_C \frac{1}{z - z_0} dz = 2\pi i f(z_0).$$

(The last equality follows from our, by now, well known integral of $1/(z - z_0)$ on a loop around z_0 .)

4.5 Proof of Cauchy's integral formula for derivatives

Recall that Cauchy's integral formula in Equation 12 says

$$f^{(n)}(z) = \frac{n!}{2\pi i} \int_C \frac{f(w)}{(w - z)^{n+1}} dw, \quad n = 0, 1, 2, \dots$$

First we'll offer a quick proof which captures the reason behind the formula, and then a formal proof.

Quick proof.

We have an integral representation for $f(z)$, $z \in A$, we use that to find an integral representation for $f'(z)$, $z \in A$.

$$f'(z) = \frac{d}{dz} \left[\frac{1}{2\pi i} \int_C \frac{f(w)}{w - z} dw \right] = \frac{1}{2\pi i} \int_C \frac{d}{dz} \left(\frac{f(w)}{w - z} \right) dw = \frac{1}{2\pi i} \int_C \frac{f(w)}{(w - z)^2} dw$$

(Note, since $z \in A$ and $w \in C$, we know that $w - z \neq 0$) Thus,

$$f'(z) = \frac{1}{2\pi i} \int_C \frac{f(w)}{(w - z)^2} dw$$

Now, by iterating this process, i.e. by mathematical induction, we can show the formula for higher order derivatives.

Formal proof.

We do this by taking the limit of $\lim_{\Delta z \rightarrow 0} \frac{f(z + \Delta z) - f(z)}{\Delta z}$, using the integral representation of both terms.

$$\begin{aligned} f(z + \Delta z) &= \frac{1}{2\pi i} \int_C \frac{f(w)}{w - z - \Delta z} dw \\ f(z) &= \frac{1}{2\pi i} \int_C \frac{f(w)}{w - z} dw \end{aligned}$$

Now, using a little algebraic manipulation we get

$$\begin{aligned} \frac{f(z + \Delta z) - f(z)}{\Delta z} &= \frac{1}{2\pi i \Delta z} \int_C \frac{f(w)}{w - z - \Delta z} - \frac{f(w)}{w - z} dw \\ &= \frac{1}{2\pi i \Delta z} \int_C \frac{f(w)\Delta z}{(w - z - \Delta z)(w - z)} dw \\ &= \frac{1}{2\pi i} \int_C \frac{f(w)}{(w - z)^2 - \Delta z(w - z)} dw \end{aligned}$$

Letting Δz go to 0, we get Cauchy's formula for $f'(z)$.

$$f'(z) = \frac{1}{2\pi i} \int_C \frac{f(w)}{(w - z)^2} dw$$

There is no problem taking the limit under the integral sign because everything is continuous and the denominator is never 0.

4.6 Amazing consequence of Cauchy's integral formula

Theorem. Suppose $f(z)$ is analytic on a region A . Then, f has derivatives of all order.

Proof. This follows from Cauchy's integral formula for derivatives. That is, we have a formula for all the derivatives, so in particular the derivatives all exist.

A little more precisely: for any point z in A we can put a small disk around z_0 that is entirely contained in A . Let C be the boundary of the disk, then Cauchy's formula gives a formula for all the derivatives $f^{(n)}(z_0)$ in terms of integrals over C . In particular, those derivatives exist.

Remark. If you look at the proof of Cauchy's formula for derivatives you'll see that f having derivatives of all orders boils down to $1/(w - z)$ having derivatives of all orders for w on a curve not containing z .

Important remark. We have at times assumed that for $f = u + iv$ analytic, u and v have continuous higher order partial derivatives. This theorem confirms that fact. In particular, $u_{xy} = u_{yx}$, etc.

4.7 More consequences

4.7.1 Cauchy's inequality

Let C_R be the circle $|z - z_0| = R$. Assume that $f(z)$ is analytic on C_R and its interior, i.e. on the disk $|z - z_0| \leq R$. Finally let $M_R = \max |f(z)|$ over z on C_R .

Using Cauchy's integral formula for derivatives (Equation 12) we have

$$|f^{(n)}(z_0)| \leq \frac{n!}{2\pi} \int_{C_R} \frac{|f(w)|}{|w - z_0|^{n+1}} |dw| \leq \frac{n!}{2\pi} \frac{M_R}{R^{n+1}} \int_{C_R} |dw| = \frac{n!}{2\pi} \frac{M_R}{R^{n+1}} \cdot 2\pi R$$

Therefore, we have:

Theorem 4.14. (Cauchy's inequality)

$$|f^{(n)}(z_0)| \leq \frac{n! M_R}{R^n}, \quad n = 1, 2, 3, \dots \quad (13)$$

4.7.2 Liouville's theorem

Theorem 4.15. (Liouville's theorem) Assume $f(z)$ is entire and suppose it is bounded in the complex plane, namely $|f(z)| < M$ for all $z \in \mathbf{C}$ then $f(z)$ is constant.

Proof. For any circle of radius R around z_0 the Cauchy inequality says $|f'(z_0)| \leq \frac{M}{R}$.

But, R can be as large as we like so we conclude that $|f'(z_0)| = 0$ for every $z_0 \in \mathbf{C}$. Since the derivative is 0, the function itself is constant.

In short: If f is entire and bounded then f is constant.

Note. $P(z) = a_n z^n + \dots + a_0$, $\sin(z)$, e^z are all entire but **not bounded**.

Now, practically for free, we get the fundamental theorem of algebra.

Corollary. (Fundamental theorem of algebra)

Any polynomial P of degree $n \geq 1$, i.e.

$$P(z) = a_0 + a_1 z + \dots + a_n z^n, \quad a_n \neq 0,$$

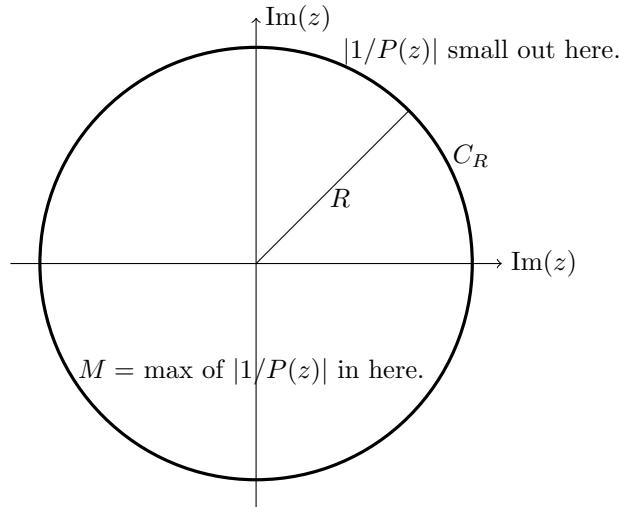
has at least exactly n roots.

Proof. There are two parts to the proof.

Hard part: Show that P has at least one root.

This is done by contradiction, together with Liouville's theorem. Suppose $P(z)$ **does not** have a zero. Then

1. $f(z) = 1/P(z)$ is entire. This is obvious because (by assumption) $P(z)$ has no zeros.
2. $f(z)$ is bounded. This follows because $1/P(z)$ goes to 0 as $|z|$ goes to ∞ .



(It is clear that $|1/P(z)|$ goes to 0 as z goes to infinity, i.e. $|1/P(z)|$ is small outside a large circle. So $|1/P(z)|$ is bounded by M .)

So, by Liouville's theorem $f(z)$ is constant. Thus $P(z)$ is constant, which is absurd and wrong! This contradicts the hypothesis of "No zeros", so P must have a zero.

Easy part. P has exactly n zeros. Let z_0 be one zero. We can factor $P(z) = (z - z_0)Q(z)$. $Q(z)$ has degree $n - 1$. If $n - 1 > 0$, then we can apply the result to $Q(z)$. We can continue this process until the degree of Q is 0.

4.7.3 Maximum modulus principle

Briefly, the maximum modulus principle states that if f is analytic and not constant in a domain A then $|f(z)|$ has no relative maximum in A and the absolute maximum of $|f|$ occurs on the boundary of A . In order to get this technically correct we need to state it carefully with the proper assumptions.

Theorem 4.16. (maximum modulus principle) Suppose $f(z)$ is analytic in a connected region A and z_0 is a point in A .

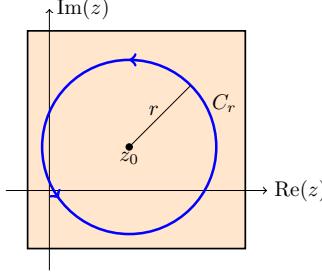
1. If $|f|$ has a relative maximum at z_0 then $f(z)$ is constant in a neighborhood of z_0 .
2. If A is bounded and connected, and f is continuous on A and its boundary, then either f is constant or the absolute maximum of $|f|$ occurs only on the boundary of A .

In order to prove the maximum modulus principle we will first prove the mean value property. This will give you a good feel for the maximum modulus principle. It is also important and interesting in its own right.

Theorem 4.17. (Mean value property) Suppose $f(z)$ is analytic on the closed disk of radius r centered at z_0 , i.e. the set $|z - z_0| \leq r$. Then,

$$f(z_0) = \frac{1}{2\pi} \int_0^{2\pi} f(z_0 + re^{i\theta}) d\theta \quad (14)$$

Proof. This is just Cauchy's integral formula on the disk $D_r = |z - z_0| \leq r$.



Since the disk is closed the region of analyticity must go beyond the circle.

We can parametrize C_r , the boundary of D_r , as

$$\gamma(t) = z_0 + re^{i\theta}, \text{ with } 0 \leq \theta \leq 2\pi, \text{ so } \gamma'(\theta) = ire^{i\theta}.$$

By Cauchy's formula we have

$$f(z_0) = \frac{1}{2\pi i} \int_{C_r} \frac{f(z)}{z - z_0} dz = \frac{1}{2\pi i} \int_0^{2\pi} \frac{f(z_0 + re^{i\theta})}{re^{i\theta}} ire^{i\theta} d\theta = \frac{1}{2\pi} \int_0^{2\pi} f(z_0 + re^{i\theta}) d\theta$$

This proves the property.

In words, the mean value property says $f(z_0)$ is the arithmetic mean of the values on the circle.

Proof of the maximum modulus principle. The argument for Theorem 4.16 part (1) is a little fussy. To make it we'll need to state the triangle inequality a little more carefully. For vectors the careful statement is

$$|\vec{v} + \vec{w}| \leq |\vec{v}| + |\vec{w}|$$

with equality only if \vec{v} and \vec{w} point in the same direction. This should be clear if you draw a picture.

The equivalent statement for complex numbers is

$$|z_1 + z_2| \leq |z_1| + |z_2|$$

with equality only if z_1 and z_2 lie along the same ray from the origin, i.e. $\arg(z_1) = \arg(z_2)$.

Finally for integrals the statement is

$$\left| \int_C f(z) dz \right| \leq \int_C |f(z)| |dz|$$

with equality only if, for all z on the curve C , $f(z)$ all lie along the same ray from the origin.

Now, to see part (1) we use the mean value property and the triangle inequality.

Since z_0 is a relative maximum of $|f|$, for every small enough circle $C : |z - z_0| = r$ around z_0 we have $|f(z)| \leq |f(z_0)|$ for z on C . Therefore, by the mean value property and the

triangle inequality

$$\begin{aligned}
 |f(z_0)| &= \left| \frac{1}{2\pi} \int_0^{2\pi} f(z_0 + re^{i\theta}) d\theta \right| && \text{(mean value property)} \\
 &\leq \frac{1}{2\pi} \int_0^{2\pi} |f(z_0 + re^{i\theta})| d\theta && \text{(triangle inequality)} \\
 &\leq \frac{1}{2\pi} \int_0^{2\pi} |f(z_0)| d\theta && (|f(z_0 + re^{i\theta})| \leq |f(z_0)|) \\
 &= |f(z_0)|
 \end{aligned}$$

Since the beginning and end of the above are both $|f(z_0)|$ all the inequalities in the chain must be equalities. The first inequality can only be an equality if for all θ , $f(z_0 + re^{i\theta})$ lie on the same ray from the origin. The second inequality can only be an equality if all $|f(z_0 + re^{i\theta})| = |f(z_0)|$. So we have all $f(z_0 + re^{i\theta})$ have the same magnitude and lie along the same ray. This implies they are all the same.

Finally, if $f(z)$ is constant along the circle and $f(z_0)$ is the average of $f(z)$ over the circle then $f(z) = f(z_0)$, i.e. f is constant on a small disk around z_0 . QED.

Proof of part (2): The assumptions that A is bounded and f is continuous on A and its boundary serve to guarantee that $|f|$ has an absolute maximum (on A combined with its boundary). Part (1) guarantees that the absolute maximum can not lie in the interior of the region A unless f is constant. (This requires a bit more argument. Do you see why?) If the absolute maximum is not in the interior it must be on the boundary.

Example 4.18. Find the maximum modulus of e^z on the unit square with $0 \leq x, y \leq 1$.

answer: $|e^{x+iy}| = e^x$, so the maximum is when $x = 1$, $0 \leq y \leq 1$ is arbitrary. This is indeed on the boundary of the unit square

Example 4.19. Find the maximum modulus for $\sin(z)$ on the square $[0, 2\pi] \times [0, 2\pi]$.

answer: We use the formula $\sin(z) = \sin x \cosh y + i \cos x \sinh y$. So

$$|\sin(z)|^2 = \sin^2 x \cosh^2 y + \cos^2 x \sinh^2 y = \sin^2 x \cosh^2 y + (1 - \sin^2 x) \sinh^2 y = \sin^2 x + \sinh^2 y.$$

Find the maximum over x of $\sin^2(x)$ is at $x = \pi/2$ and $x = 3\pi/2$. The maximum of $\sinh^2 y$ is at $y = 2\pi$. So maximum modulus is $\sqrt{1 + \sinh^2(2\pi)} = \sqrt{\cosh^2(2\pi)} = \cosh(2\pi)$. This occurs at the points $z = x + iy = \pi/2 + 2\pi i$ and $z = 3\pi/2 + 2\pi i$. Both these points are on the boundary of the region.

Example 4.20. Suppose $f(z)$ is entire. Show that if $\lim_{z \rightarrow \infty} f(z) = 0$ then $f(z) \equiv 0$.

answer: This is a standard use of the maximum modulus principle. The strategy is to show that the maximum of $|f(z)|$ is not on the boundary (of the appropriately chosen region), so $f(z)$ must be constant.

Fix z_0 . For $R > |z|$ let M_R be the maximum of $|f(z)|$ on the circle $|z| = R$. The maximum modulus theorem says that $|f(z_0)| < M_R$. Since $f(z)$ goes to 0, as R goes to infinity, we must have M_R also goes to 0. This means $|f(z_0)| = 0$. Since this is true for any z_0 , we have $f(z) \equiv 0$.

Example 4.21. Here is an example of why you need A to be bounded in the maximum

modulus theorem. Let A be the upper half-plane $\text{Im}(z) > 0$. So the boundary of A is the real axis.

Let $f(z) = e^{-iz}$. $|f(x)| = |e^{-ix}| = 1$ for x along the real axis. Since $|f(2i)| = |e^2| > 1$, we see $|f|$ cannot take its maximum along the boundary of A .

Of course, it can't take its maximum in the interior of A either. What happens here is that $f(z)$ doesn't have a maximum modulus. Indeed $|f(z)|$ goes to infinity along the positive imaginary axis.

5 Introduction to harmonic functions

5.1 Introduction

We will accept as a given that harmonic functions appear regularly and play a fundamental role in math, physics and engineering. In this topic we'll learn the definition, some key properties and their tight connection to complex analysis. The key connection to 18.04 is that both the real and imaginary parts of analytic functions are harmonic. We will see that this is a simple consequence of the Cauchy-Riemann equations. In the next topic we will look at some applications to hydrodynamics.

5.2 Harmonic functions

We start by defining harmonic functions and looking at some of their properties.

Definition 5.1. A function $u(x, y)$ is called **harmonic** if it is twice continuously differentiable and it satisfies the following partial differential equation:

$$\nabla^2 u = u_{xx} + u_{yy} = 0. \quad (15)$$

Equation 15 is called **Laplace's equation**. So a function is harmonic if it satisfies Laplace's equation. The operator ∇^2 is called the **Laplacian** and $\nabla^2 u$ is called the **Laplacian of u** .

5.3 Del notation

Here's a quick reminder on the use of the notation ∇ . For a function u and a vector field $\mathbf{F} = (u, v)$

- $\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right)$
- $\mathbf{grad} u = \nabla u = (u_x, u_y)$
- $\mathbf{curl} \mathbf{F} = \nabla \times \mathbf{F} = v_x - u_y$ (in two dimensions we take curl to be scalar valued).
- $\mathbf{div} \mathbf{F} = \nabla \cdot \mathbf{F} = u_x + v_y$
- $\mathbf{div grad} u = \nabla \cdot \nabla u = \nabla^2 u = u_{xx} + u_{yy}$.
- Also: $\mathbf{curl grad} u = \nabla \times \nabla u = 0$, and $\mathbf{div curl} \mathbf{F} = \nabla \cdot \nabla \times \mathbf{F} = 0$.

5.3.1 Analytic functions have harmonic pieces

The connection between analytic and harmonic functions is very strong. In many respects it mirrors the connection between e^z and sine and cosine.

Let $z = x + iy$ and write $f(z) = u(x, y) + iv(x, y)$.

Theorem 5.2. If $f(z) = u(x, y) + iv(x, y)$ is analytic on a region A then both u and v are harmonic functions on A .

Proof. This is a simple consequence of the Cauchy-Riemann equations. Since $u_x = v_y$ we have $u_{xx} = v_{yy}$. Likewise, $u_y = -v_x$ implies $u_{yy} = -v_{xy}$. Since $v_{xy} = v_{yx}$ we have $u_{xx} + v_{yy} = v_{yx} - v_{xy} = 0$. Therefore u is harmonic. We can handle v similarly.

Note. Since we know an analytic function is infinitely differentiable we know u and v have the required two continuous partial derivatives. This also insures that the mixed partials agree, i.e. $v_{xy} = v_{yx}$.

To complete the tight connection between analytic and harmonic functions we show that any harmonic function is the real part of an analytic function.

Theorem 5.3. If $u(x, y)$ is harmonic on a simply connected region A then u is the real part of an analytic function $f(z) = u(x, y) + iv(x, y)$.

Proof. This is similar to our proof that an analytic function has an antiderivative. First we come up with a candidate for $f(z)$ and then show it has the properties we need. Here are the details broken down into steps 1-4.

1. Find a candidate, call it $g(z)$, for $f'(z)$:

If we had an analytic f with $f = u + iv$, then Cauchy-Riemann says that $f' = u_x - iu_y$. So, let's define $g = u_x - iu_y$. This is our candidate for f' .

2. Show that $g(z)$ is analytic.

Write $g = \phi + i\psi$, where $\phi = u_x$ and $\psi = -u_y$. Checking the Cauchy-Riemann equations we have

$$\begin{bmatrix} \phi_x & \phi_y \\ \psi_x & \psi_y \end{bmatrix} = \begin{bmatrix} u_{xx} & u_{xy} \\ -u_{yx} & -u_{yy} \end{bmatrix}$$

Since u is harmonic we know $u_{xx} = -u_{yy}$, so $\phi_x = \psi_y$. It is clear that $\phi_y = -\psi_x$. Thus g satisfies the Cauchy-Riemann equations, so it is analytic.

3. Let f be an antiderivative of g .

Since A is simply connected our statement of Cauchy's theorem guarantees that $g(z)$ has an antiderivative in A . We'll need to fuss a little to get the constant of integration exactly right. So, pick a base point z_0 in A . Define the antiderivative of $g(z)$ by

$$f(z) = \int_{z_0}^z g(z) dz + u(x_0, y_0).$$

(Again, by Cauchy's theorem this integral can be along any path in A from z_0 to z .)

4. Show that the real part of f is u .

Let's write $f = U + iV$. So, $f' = U_x - iU_y$. By construction $f' = g = u_x - iu_y$. This means the first partials of U and u are the same, so U and u differ by at most a constant. However, also by construction, $f(z_0) = u(x_0, y_0) = U(x_0, y_0) + iV(x_0, y_0)$, So, $U(x_0, y_0) = u(x_0, y_0)$

(and $V(x_0, y_0) = 0$). Since they agree at one point we must have $U = u$, i.e. the real part of f is u as we wanted to prove.

Important corollary. u is infinitely differentiable.

Proof. By definition we only require a harmonic function u to have continuous second partials. Since the analytic f is infinitely differentiable, we have shown that so is u !

5.3.2 Harmonic conjugates

This leads to the following definition:

Definition. If u and v are the real and imaginary parts of an analytic function, then we say u and v are [harmonic conjugates](#).

Note. If $f(z) = u + iv$ is analytic then so is $if(z) = -v + iu$. So, u and v are harmonic conjugates and so are u and $-v$.

5.4 A second proof that u and v are harmonic

This fact is important enough that we will give a second proof using Cauchy's integral formula. One benefit of this proof is that it reminds us that Cauchy's integral formula can transfer a general question on analytic functions to a question about the function $1/z$. We start with an easy to derive fact.

Fact. The real and imaginary parts of $f(z) = 1/z$ are harmonic away from the origin. The same is true for the function $g(z) = f(z - a) = 1/(z - a)$.

Proof. $\frac{1}{z} = \frac{x}{x^2 + y^2} - i\frac{y}{x^2 + y^2}$. It is a simple matter to apply the Laplacian and see that you get 0. We'll leave the algebra to you! The statement about $g(z)$ follows in either exactly the same way, or by noting that the Laplacian is translation invariant.

Second proof that f analytic implies u and v are harmonic. We are proving that if $f = u + iv$ is analytic then u and v are harmonic. So, suppose f is analytic at the point z_0 . This means there is a disk of some radius, say r , around z_0 where f is analytic. Cauchy's formula says

$$f(z) = \frac{1}{2\pi i} \int_{C_r} \frac{f(w)}{w - z} dz,$$

where C_r is the circle $|w - z_0| = r$ and z is in the disk $|z - z_0| < r$.

Now, since the real and imaginary parts of $1/(w - z)$ are harmonic (as a function of z) The same must be true of the integral, which is limit of linear combinations of such functions. Since the circle is finite and f is continuous, interchanging the order of integration and differentiation is not a problem. We'll go through the details in the appendix.

5.5 Maximum principle, mean value property

These are similar to the corresponding properties of analytic functions. Indeed, we deduce them from those corresponding properties.

Theorem. (Mean value property) If u is a harmonic function then u satisfies the mean value property. That is, suppose u is harmonic on and inside a circle of radius r centered at $z_0 = x_0 + iy_0$ then

$$u(x_0, y_0) = \frac{1}{2\pi} \int_0^{2\pi} u(z_0 + r e^{i\theta}) d\theta$$

Proof. Let $f = u + iv$ be an analytic function with u as its real part. The mean value property for f says

$$u(x_0, y_0) + iv(x_0, y_0) = f(z_0) = \frac{1}{2\pi} \int_0^{2\pi} f(z_0 + r e^{i\theta}) d\theta = \frac{1}{2\pi} \int_0^{2\pi} u(z_0 + r e^{i\theta}) + iv(z_0 + r e^{i\theta}) d\theta$$

Looking at the real parts of this equation proves the theorem.

Theorem. (Maximum principle) Suppose $u(x, y)$ is harmonic on a region A .

(i) If u has a relative maximum or minimum at z_0 then u is constant on a disk centered at z_0 .

(ii) If A is bounded and connected and u is continuous on the boundary of A then the absolute maximum and absolute minimum of u occur on the boundary.

Proof. The proof for maxima is identical to the one for the maximum modulus principle. The proof for minima comes by looking at the maxima of $-u$.

Note. for analytic functions we only talked about maxima because we had to use the modulus in order to have real values. Since $|-f| = |f|$ we couldn't use the trick of turning minima into maxima by using a minus sign.

5.6 Orthogonality of curves

An important property of harmonic conjugates u and v is that their level curves are orthogonal. We start by showing their gradients are orthogonal.

Lemma 5.4. Let $z = x + iy$ and suppose that $f(z) = u(x, y) + iv(x, y)$ is analytic. Then the dot product of their gradients is 0, i.e.

$$\nabla u \cdot \nabla v = 0.$$

Proof. The proof is an easy application of the Cauchy-Riemann equations.

$$\nabla u \cdot \nabla v = (u_x, u_y) \cdot (v_x, v_y) = u_x v_x + u_y v_y = v_y v_x - v_x v_y = 0$$

In the last step we used the Cauchy-Riemann equations to substitute v_y for u_x and $-v_x$ for u_y .

The lemma holds whether or not the gradients are 0. To guarantee that the level curves are smooth the next theorem requires that $f'(z) \neq 0$.

Theorem. Let $z = x + iy$ and suppose that $f(z) = u(x, y) + iv(x, y)$ is analytic. If $f'(z) \neq 0$ then the level curve of u through (x, y) is orthogonal to the level curve v through (x, y) .

Proof. The technical requirement that $f'(z) \neq 0$ is needed to be sure that the level curves are smooth. We need smoothness so that it even makes sense to ask if the curves are

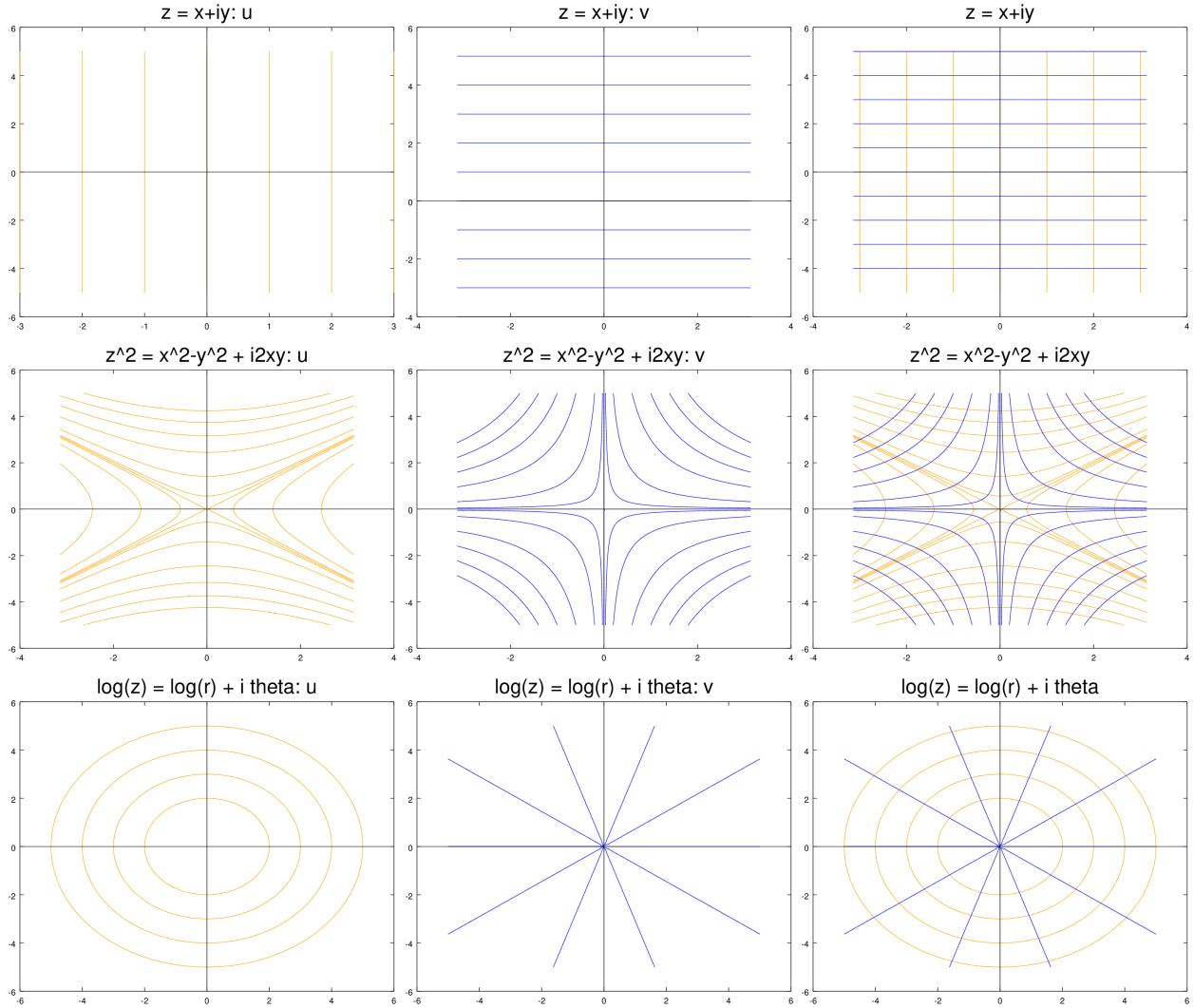
orthogonal. We'll discuss this below. Assuming the curves are smooth the proof of the theorem is trivial: We know from 18.02 that the gradient ∇u is orthogonal to the level curves of u and the same is true for ∇v and the level curves of v . Since, by lemma 5.4, the gradients are orthogonal this implies the curves are orthogonal.

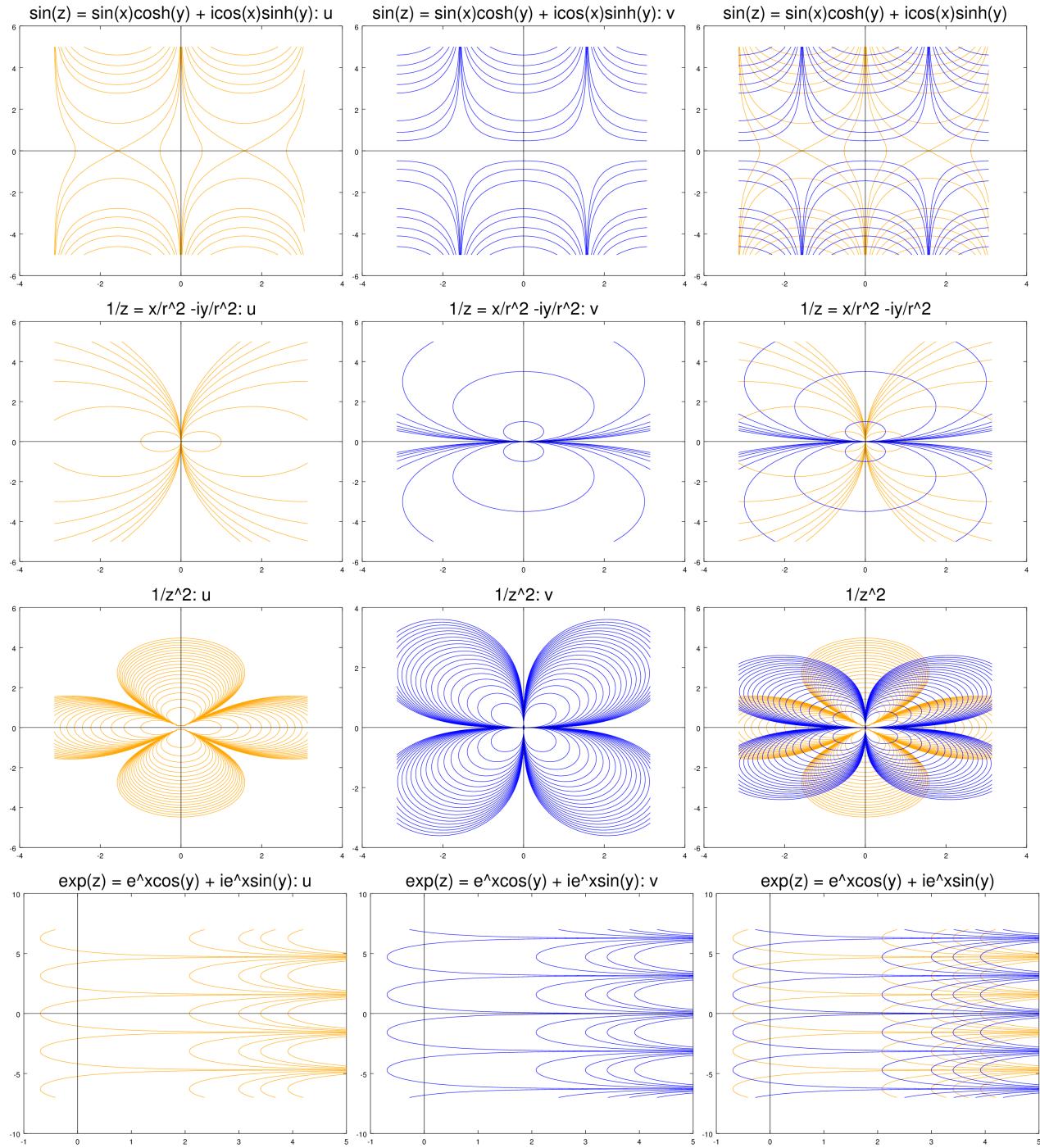
Finally, we show that $f'(z) \neq 0$ means the curves are smooth. First note that

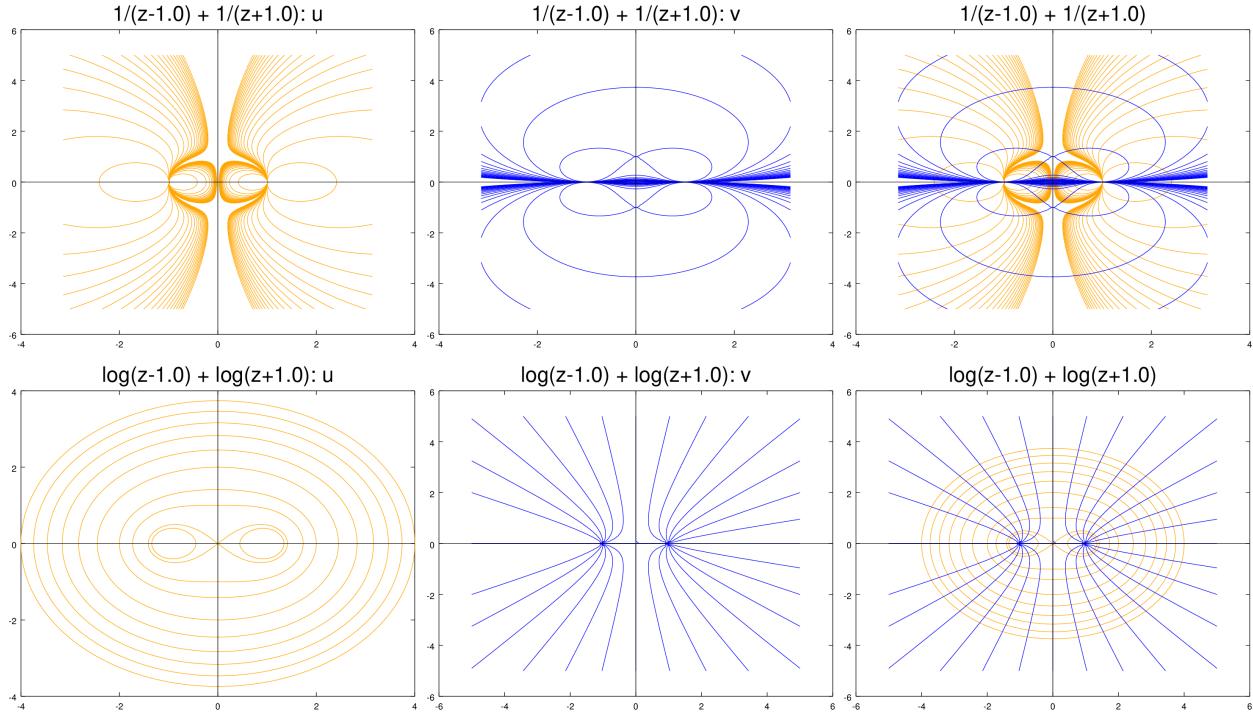
$$f'(z) = u_x(x, y) - iu_y(x, y) = v_y(x, y) + iv_x(x, y).$$

Now, since $f'(z) \neq 0$ we know that $\nabla u = (u_x, u_y) \neq 0$. Likewise, $\nabla v \neq 0$. Thus, the gradients are not zero and the level curves must be smooth.

Example 5.5. The figures below show level curves of u and v for a number of functions. In all cases, the level curves of u are in orange and those of v are in blue. For each case we show the level curves separately and then overlayed on each other.







Example 5.6. Let's work out the gradients in a few simple examples.

(i) Let $f(z) = z^2 = (x^2 - y^2) + i2xy$. So

$$\nabla u = (2x, -2y) \text{ and } \nabla v = (2y, 2x).$$

It's trivial to check that $\nabla u \cdot \nabla v = 0$, so they are orthogonal.

(ii) Let $f(z) = \frac{1}{z} = \frac{x}{r^2} - i\frac{y}{r^2}$. So, it's easy to compute

$$\nabla u = \left(\frac{y^2 - x^2}{r^4}, \frac{-2xy}{r^4} \right) \text{ and } \nabla v = \left(\frac{2xy}{r^4}, \frac{y^2 - x^2}{r^4} \right).$$

Again it's trivial to check that $\nabla u \cdot \nabla v = 0$, so they are orthogonal.

Example 5.7. (Degenerate points: $f'(z) = 0$.) Consider $f(z) = z^2$. From the previous example we have

$$u(x, y) = x^2 - y^2, v(x, y) = 2xy, \nabla u = (2x, -2y), \nabla v = (2y, 2x).$$

At $z = 0$, the gradients are both 0 so the theorem on orthogonality doesn't apply

Let's look at the level curves through the origin. The level curve (really the 'level set') for $u = x^2 - y^2 = 0$ is the pair of lines $y = \pm x$. At the origin this is not a smooth curve.

Look at the figures for z^2 above. It does appear that away from the origin the level curves of u intersect the lines where $v = 0$ at right angles. The same is true for the level curves of v and the lines where $u = 0$. You can see the degeneracy forming at the origin: as the level curves head towards 0 they get pointier and more right angled. So the level curve $u = 0$ is more properly thought of as four right angles. The level curve of $u = 0$, not knowing which leg of $v = 0$ to intersect orthogonally takes the average and comes into the origin at 45° .

5.7 Appendix

5.7.1 Interchanging limits

In our proof that the parts of an analytic function are harmonic based on Cauchy's theorem we needed the following. For $z = x + iy$ write

$$\frac{1}{w - z} = u(x, y, w) + iv(x, y, w)$$

We know that on the circle $C_r : |w - z| = r$, u and v have derivatives of all orders in x and y . (And w , but we only care that it's continuous for w on C_r . Then, if $f(w)$ is continuous on C_r we have

$$\frac{\partial}{\partial x} \int_{C_r} \frac{f(w)}{w - z} dw = \frac{\partial}{\partial x} \int_{C_r} f(w)u(x, y, w) + if(w)v(x, y, w) dw = \int_{C_r} f(w)u_x(x, y, w) + if(w)v_x(x, y, w) dw.$$

That is, we can interchange the order of integration over w and taking partials in x and y .

Proof. To be completed

The basic idea is to see that since $|f(w)|$ is bounded on C_r we have

$$\lim_{x \rightarrow x_0} \int_{C_r} f(w) \left(\frac{u(x, y_0, w) - u(x_0, y_0, w)}{x - x_0} - u_x(x_0, y_0) \right) dw = 0.$$

6 Two dimensional hydrodynamics and complex potentials

6.1 Introduction

Laplace's equation and harmonic functions show up in many physical models. As we have just seen, harmonic functions in two dimensions are closely linked with complex analytic functions. In this section we will exploit this connection to look at two dimensional hydrodynamics, i.e. fluid flow.

Since static electric fields and steady state temperature distributions are also harmonic, the ideas and pictures we use can be repurposed to cover these topics as well.

6.2 Velocity fields

Suppose we have water flowing in a region A of the plane. Then at every point (x, y) in A the water has a velocity. In general, this velocity will change with time. We'll let \mathbf{F} stand for the **velocity vector field** and we can write

$$\mathbf{F}(x, y, t) = (u(x, y, t), v(x, y, t)).$$

The arguments (x, y, t) indicate that the velocity depends on these three variables. In general, we will shorten the name to **velocity field**.

A beautiful and mesmerizing example of a velocity field is at <http://hint.fm/wind/index.html>. This shows the current velocity of the wind at all points in the continental U.S.

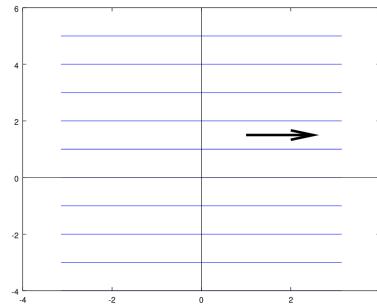
6.3 Stationary flows

If the velocity field is unchanging in time we call the flow a **stationary flow**. In this case, we can drop t as an argument and write:

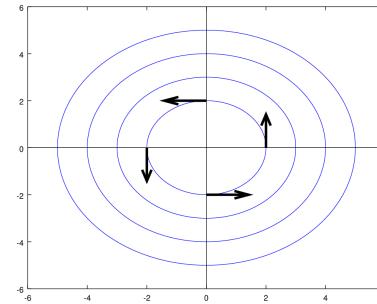
$$\mathbf{F}(x, y) = (u(x, y), v(x, y))$$

Here are a few examples. These pictures show the streamlines from similar figures in topic 5. We've added arrows to indicate the direction of flow.

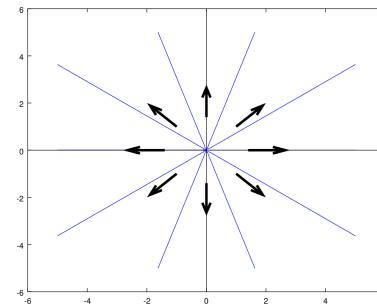
Example 6.1. Uniform flow. $\mathbf{F} = (1, 0)$.



Example 6.2. Eddy (vortex) $\mathbf{F} = (y/r^2, -x/r^2)$



Example 6.3. Source $\mathbf{F} = (x/r^2, y/r^2)$



6.4 Physical assumptions, mathematical consequences

6.4.1 Summary: mathematical consequence of physical assumptions

This is a wordy section, so we'll start by listing the mathematical properties that will follow from our assumptions about the velocity field $\mathbf{F} = u + iv$.

- (A) $\mathbf{F} = \mathbf{F}(x, y)$ is a function of x, y , but not time t .
- (B) $\operatorname{div}\mathbf{F} = 0$ (divergence free).
- (C) $\operatorname{curl}\mathbf{F} = 0$ (curl free).

6.4.2 Physical assumptions

We will make some standard physical assumptions. These don't apply to all flows, but they do apply to a good number of them and they are a good starting point for understanding fluid flow more generally. More important to 18.04, these are the flows that are readily susceptible to complex analysis.

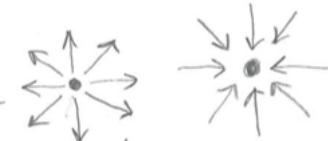
Here are the assumptions about the flow, we'll discuss them further below:

- (A) Stationary.
- (B) Incompressible.
- (C) Irrotational

(B) **Incompressible.** We will assume that the fluid is incompressible. This means that the density of the fluid is constant across the domain. Mathematically this says that the velocity field \mathbf{F} must be **divergence free**, i.e. for $\mathbf{F} = (u, v)$:

$$\operatorname{div}\mathbf{F} \equiv \nabla \cdot \mathbf{F} = u_x + v_y = 0.$$

To understand this, recall that the divergence measures the infinitesimal flux of the field. If the flux is not zero at a point (x_0, y_0) then near that point the field looks like



Left: Divergent field: $\operatorname{div}\mathbf{F} > 0$, right: Convergent field: $\operatorname{div}\mathbf{F} < 0$

If the field is diverging or converging then the density must be changing! That is, the flow is not incompressible.

As a fluid flow the left hand picture represents a source and the right represents a sink. In electrostatics where \mathbf{F} expresses the electric field, the left hand picture is the field of a positive charge density and the right is that of a negative charge density.

If you prefer a non-infinitesimal explanation, we can recall Green's theorem in flux form. It says that for a simple closed curve C and a field $\mathbf{F} = (u, v)$, differentiable on and inside C , the flux of \mathbf{F} through C satisfies

$$\text{Flux of } \mathbf{F} \text{ across } C = \int_C \mathbf{F} \cdot \mathbf{n} ds = \iint_R \operatorname{div}\mathbf{F} dx dy,$$

where R is the region inside C . Now, suppose that $\operatorname{div}\mathbf{F}(x_0, y_0) > 0$, then $\operatorname{div}\mathbf{F}(x, y) > 0$ for all (x, y) close to (x_0, y_0) . So, choose a small curve C around (x_0, y_0) such that $\operatorname{div}\mathbf{F} > 0$ on and inside C . By Green's theorem

$$\text{Flux of } \mathbf{F} \text{ through } C = \iint_R \operatorname{div}\mathbf{F} dx dy > 0.$$

Clearly, if there is a net flux out of the region the density is decreasing and the flow is not incompressible. The same argument would hold if $\operatorname{div}\mathbf{F}(x_0, y_0) < 0$. We conclude that incompressible is equivalent to divergence free.

(C) **Irrational.** We will assume that the fluid is irrotational. This means that there are no infinitesimal vortices in A . Mathematically this says that the velocity field \mathbf{F} must be **curl free**, i.e. for $\mathbf{F} = (u, v)$:

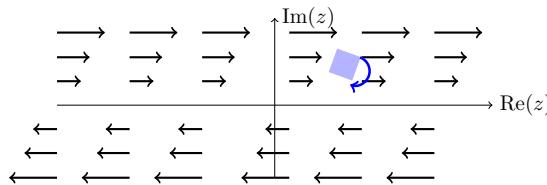
$$\operatorname{curl}\mathbf{F} \equiv \nabla \times \mathbf{F} = v_x - u_y = 0.$$

To understand this, recall that the curl measures the infinitesimal rotation of the field. Physically this means that a small paddle placed in the flow will not spin as it moves with the flow.

6.4.3 Examples

Example 6.4. The eddy is irrotational! The eddy from Example 6.2 is irrotational. The vortex at the origin is not in $A = \mathbf{C} - \{0\}$ and you can easily check that $\operatorname{curl}\mathbf{F} = 0$ everywhere in A . This is not physically impossible: if you placed a small paddle wheel in the flow it would travel around the origin without spinning!

Example 6.5. Shearing flows are rotational. Here's an example of a vector field that has rotation, though not necessarily swirling.



Shearing flow: box turns because current is faster at the top.

The field $\mathbf{F} = (ay, 0)$ is horizontal, but $\operatorname{curl}\mathbf{F} = -a \neq 0$. Because the top moves faster than the bottom it will rotate a square parcel of fluid. The minus sign tells you the parcel will rotate clockwise! This is called a **shearing flow**. The water at one level will be sheared away from the level above it.

6.4.4 Summary

$$\mathbf{F}(x, y) = (u(x, y), v(x, y)).$$

- (A) Stationary: \mathbf{F} depends on x, y , but not t .
- (B) Incompressible: divergence free, $\operatorname{div}\mathbf{F} = u_x + v_y = 0$, i.e. $u_x = -v_y$.
- (C) Irrational: curl free, $\operatorname{curl}\mathbf{F} = v_x - u_y = 0$, i.e. $u_y = v_x$.

For future reference we put these equalities in a numbered equation:

$$u_x = -v_y \text{ and } u_y = v_x \tag{16}$$

These look almost like the Cauchy-Riemann equations (with sign differences)!

6.5 Complex potentials

There are different ways to do this. We'll start by seeing that every complex analytic function leads to an irrotational, incompressible flow. Then we'll go backwards and see that all such flows lead to an analytic function. We will learn to call the analytic function the complex potential of the flow.

Annoyingly, we are going to have to switch notation. Because u and v are already taken by the vector field \mathbf{F} , we will call our complex function $\Phi = \phi + i\psi$.

6.5.1 Analytic functions give us incompressible, irrotational flows

Let $\Phi(z)$ be an analytic function on a region A . For $z = x + iy$ we write

$$\Phi(z) = \phi(x, y) + i\psi(x, y).$$

From this we can define a vector field

$$\mathbf{F} = \nabla\phi = (\phi_x, \phi_y) = (u, v),$$

here we mean that u and v are defined by ϕ_x and ϕ_y .

From our work on analytic and harmonic functions we can make a list of properties of these functions.

1. ϕ and ψ are both harmonic.
2. The level curves of ϕ and ψ are orthogonal.
3. $\Phi' = \phi_x - i\phi_y$.
4. \mathbf{F} is divergence and curl free (proof just below).

Thus the analytic function Φ has given us an incompressible, irrotational vector field \mathbf{F} .

It is standard terminology to call ϕ a **potential function** for the vector field \mathbf{F} . We will also call Φ a **complex potential function** for \mathbf{F} . The function ψ will be called the **stream function** of \mathbf{F} (the name will be explained soon). The function Φ' will be called the **complex velocity**.

Proof. (\mathbf{F} is curl and divergence free.) This is an easy consequence of the definition.

$$\begin{aligned} \text{curl}\mathbf{F} &= v_x - u_y = \phi_{yx} - \phi_{xy} = 0 \\ \text{div}\mathbf{F} &= u_x + v_y = \phi_{xx} + \phi_{yy} = 0 \quad (\text{since } \phi \text{ is harmonic}). \end{aligned}$$

We'll postpone examples until after deriving the complex potential from the flow.

6.5.2 Incompressible, irrotational flows always have a complex potential functions

For technical reasons we need to add the assumption that A is simply connected. This is not usually a problem because we often work locally in a disk around a point (x_0, y_0) .

Theorem. Assume $\mathbf{F} = (u, v)$ is an incompressible, irrotational field on a simply connected region A . Then there is an analytic function Φ which is a complex potential function for \mathbf{F} .

Proof. We have done all the heavy lifting for this in previous topics. The key is to use the property $\Phi' = u - iv$ to guess Φ' . Working carefully we define

$$g(z) = u - iv$$

Step 1. Show that g is analytic. Keeping the signs straight, the Cauchy Riemann equations are

$$u_x = (-v)_y \text{ and } u_y = -(-v)_x = v_x.$$

But, these are exactly the equations in Equation 16. Thus $g(z)$ is analytic.

Step 2. Since A is simply connected, Cauchy's theorem says that $g(z)$ has an antiderivative on A . We call the antiderivative $\Phi(z)$

Step 3. Show that $\Phi(z)$ is a complex potential function for \mathbf{F} . This means we have to show that if we write $\Phi = \phi + i\psi$, then $\mathbf{F} = \nabla\phi$. To do this we just unwind the definitions.

$$\begin{aligned} \Phi' &= \phi_x - i\phi_y && \text{(standard formula for } \Phi') \\ \Phi' &= g = u - iv && \text{(definition of } \Phi \text{ and } g) \end{aligned}$$

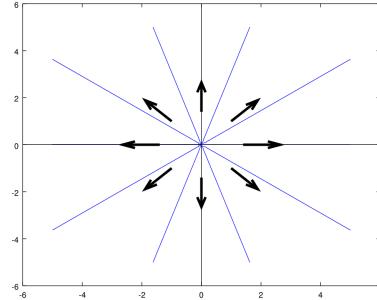
Comparing these equations we get

$$\phi_x = u, \phi_y = v.$$

But this says precisely that $\nabla\phi = \mathbf{F}$. QED

Example 6.6. Source fields. The vector field $\mathbf{F} = a \left(\frac{x}{r^2}, \frac{y}{r^2} \right)$ models a source pushing out water or the 2D electric field of a positive charge at the origin. If you prefer a 3D model, it is the field of an infinite wire along with uniform charge density along the z -axis

Show that \mathbf{F} is curl and divergence free and find its complex potential.



We could compute directly that this is curl and divergence free away from 0. An easier method is to look for a complex potential Φ . If we can find one then this will show \mathbf{F} is curl and divergence free and find ϕ and ψ all at once. If there is no such Φ then we'll know that \mathbf{F} is not both curl and divergence free.

One standard method is to use the formula for Φ' :

$$\Phi' = u - iv = \frac{(x - iy)}{r^2} = \frac{\bar{z}}{(\bar{z} * z)} = \frac{1}{z}.$$

This is analytic and we have $\Phi(z) = \log(z)$.

6.6 Stream functions

In everything we did above poor old ψ just tagged along as the harmonic conjugate of the potential function ϕ . Let's turn our attention to it and see why it's called the stream function.

Theorem. Suppose that $\Phi = \phi + i\psi$ is the complex potential for a velocity field \mathbf{F} . Then the fluid flows along the level curves of ψ , That is, the \mathbf{F} is everywhere tangent to the level curves of ψ . The level curves of ψ are called **streamlines** and ψ is called the stream function.

Proof. Again we have already done most of the heavy lifting to prove this. Since \mathbf{F} is the velocity of the flow at each point, the flow is always tangent to \mathbf{F} . You also need to remember that $\nabla\phi$ is perpendicular to the level curves of ϕ . So we have:

1. The flow is parallel to \mathbf{F} .
2. $\mathbf{F} = \nabla\phi$, so the flow is orthogonal to the level curves of ϕ .
3. Since ϕ and ψ are harmonic conjugates, the level curves of ψ are orthogonal to the level curves of ϕ .

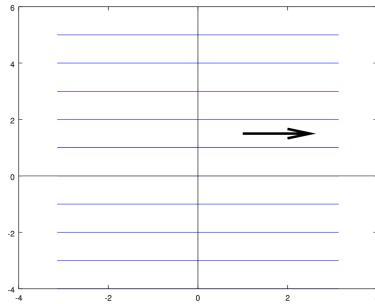
Combining 2 and 3 we see that the flow must be along the level curves of ψ .

6.6.1 Examples

We'll illustrate the streamlines in a series of examples that start by defining the complex potential for a vector field.

Example 6.7. Uniform flow Let $\Phi(z) = z$. Find \mathbf{F} and draw a plot of the streamlines. Indicate the direction of the flow.

answer: $\Phi = x + iy$. So $\phi = x$ and $\mathbf{F} = \nabla\phi = (1, 0)$, which says the flow has uniform velocity and points to the right. We also have $\psi = y$, so the streamlines are the horizontal lines $y = \text{constant}$.



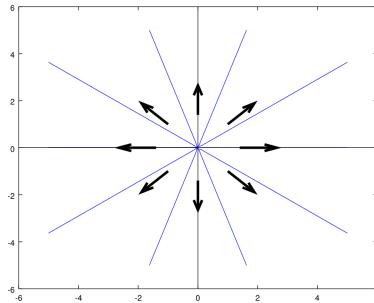
Uniform flow to the right.

Note that another way to see that the flow is to the right is to check the direction that the potential ϕ increases. The topic 5 notes show pictures of this complex potential which show both the streamlines and the equipotential lines.

Example 6.8. Linear source Let $\Phi(z) = \log(z)$. Find \mathbf{F} and draw a plot of the streamlines. Indicate the direction of the flow.

answer: $\Phi = \log(r) + i\theta$. So $\phi = \log(r)$ and $\mathbf{F} = \nabla\phi = (x/r^2, y/r^2)$, which says the flow is radial and decreases in speed as it gets farther from the origin. The field is not defined at

$z = 0$. We also have $\psi = \theta$, so the streamlines are rays from the origin.



Linear source: radial flow from the origin.

6.6.2 Stagnation points

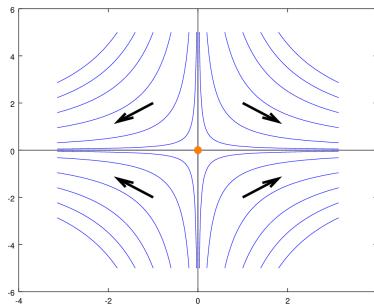
A **stagnation point** is one where the velocity field is 0.

Stagnation points. If Φ is the complex potential for a field \mathbf{F} then the stagnation points $\mathbf{F} = 0$ are exactly the points z where $\Phi'(z) = 0$.

Proof. This is clear since $\mathbf{F} = (\phi_x, \phi_y)$ and $\Phi' = \phi_x - i\phi_y$.

Example 6.9. Stagnation points. Draw the streamlines and identify the stagnation points for the potential $\Phi(z) = z^2$.

answer: (We drew the level curves for this in topic 5.) We have $\Phi = (x^2 - y^2) + i2xy$. So the streamlines are the hyperbolas: $2xy = \text{constant}$. Since $\phi = x^2 - y^2$ increases as $|x|$ increases and decreases as $|y|$ increases, the arrows, which point in the direction of increasing ϕ , are as shown on the figure below.



Stagnation flow: stagnation point at $z = 0$.

The stagnation points are the zeros of $\Phi'(z) = 2z$, i.e. the only stagnation point is at the $z = 0$.

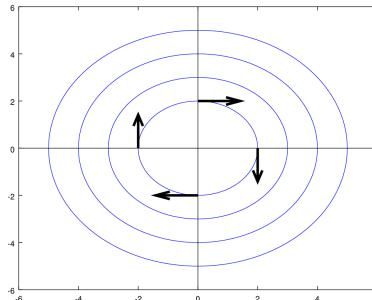
Note. The stagnation points are what we called the critical points of a vector field in 18.03.

6.7 More pretty pictures

Example 6.10. Linear vortex. Analyze the flow with complex potential function $\Phi(z) = i \log(z)$.

answer: Multiplying by i switches the real and imaginary parts of $\log(z)$ (with a sign change). We have $\Phi = -\theta + i \log(r)$. The stream lines are the curves $\log(r) = \text{constant}$, i.e.

circles with center at $z = 0$. The flow is clockwise because the potential $\phi = -\theta$ increases in the clockwise direction.



Linear vortex.

This flow is called a linear vortex. We can find \mathbf{F} using Φ' .

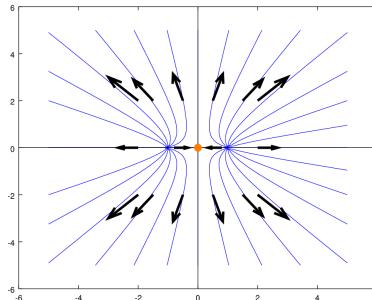
$$\Phi' = \frac{i}{z} = \frac{y}{r^2} + \frac{x}{r^2} = \phi_x - i\phi_y.$$

So $\mathbf{F} = (\phi_x, \phi_y) = (-y/r^2, x/r^2)$ – by now this should be a familiar vector field.

There are no stagnation points, but there is a singularity at the origin.

Example 6.11. Double source. Analyze the flow for with complex potential function $\Phi(z) = \log(z - 1) + \log(z + 1)$.

answer: This is a flow with linear sources at ± 1 . We used Octave to plot the level curves of $\psi = \text{Im}(\Phi)$.



Two sources.

We can analyze this flow further as follows.

1. Near each source the flow looks like a linear source.
2. On the y -axis the flow is along the axis. That is, the y -axis is a streamline. It's worth seeing three different ways of arriving at this conclusion.

Reason 1: By symmetry of vector fields associated with each linear source, the x components cancel and the combined field points along the y -axis.

Reason 2: We can write

$$\Phi(z) = \log(z - 1) + \log(z + 1) = \log((z - 1)(z + 1)) = \log(z^2 - 1).$$

So $\Phi'(z) = 2z/(z^2 - 1)$. On the imaginary axis $\Phi'(iy) = 2iy/(-y^2 - 1)$. Thus, $\mathbf{F} = (0, 2y/(y^2 + 1))$ which is along the axis.

Reason 3: On the imaginary axis $\Phi(iy) = \log(-y^2 - 1)$. Since this has constant imaginary part, the axis is a streamline.

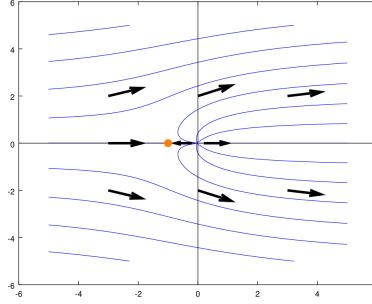
Because of the branch cut for $\log(z)$ we should probably be a little more careful here. First note that the vector field \mathbf{F} comes from $\Phi' = 2z/(z^2 - 1)$, which doesn't have a branch cut. So we shouldn't really have a problem. Now, as z approaches the y -axis from one side or the other, the argument of $\log(z^2 - 1)$ approaches either π or $-\pi$. That is, as such limits, the imaginary part is constant. So the streamline on the y -axis is the limit case of streamlines near the axis.

Since $\Phi'(z) = 0$ when $z = 0$, the origin is a stagnation point. This is where the fields from the two sources exactly cancel each other.

Example 6.12. A source in uniform flow. Consider the flow with complex potential

$$\Phi(z) = z + \frac{Q}{2\pi} \log(z).$$

This is a combination of uniform flow to the right and a source at the origin. The figure below was drawn using Octave. It shows that the flow looks like a source near the origin. Farther away from the origin the flow stops being radial and is pushed to the right by the uniform flow.



A source in uniform flow.

Since the components of Φ' and \mathbf{F} are the same except for signs, we can understand the flow by considering

$$\Phi'(z) = 1 + \frac{Q}{2\pi z}.$$

Near $z = 0$ the singularity of $1/z$ is most important and $\Phi' \approx Q/(2\pi z)$. So, the vector field looks a linear source. Far away from the origin the $1/z$ term is small and $\Phi'(z) \approx 1$, so the field looks like uniform flow.

Setting $\Phi'(z) = 0$ we find one stagnation point $z = -Q/(2\pi)$. It is the point on the x -axis where the flow from the source exactly balances that from the uniform flow. For bigger values of Q the source pushes fluid farther out before being overwhelmed by the uniform flow. That is why Q is called the **source strength**.

7 Taylor and Laurent series

7.1 Introduction

We originally defined an analytic function as one where the derivative, defined as a limit of ratios, existed. We went on to prove Cauchy's theorem and Cauchy's integral formula.

These revealed some deep properties of analytic functions, e.g. the existence of derivatives of all orders.

Our goal in this topic is to express analytic functions as infinite power series. This will lead us to Taylor series. When a complex function has an isolated singularity at a point we will replace Taylor series by Laurent series. Not surprisingly we will derive these series from Cauchy's integral formula.

Although we come to power series representations after exploring other properties of analytic functions, they will be one of our main tools in understanding and computing with analytic functions.

7.2 Geometric series

Having a detailed understanding of geometric series will enable us to use Cauchy's integral formula to understand power series representations of analytic functions. We start with the definition:

Definition. A [finite geometric series](#) has one of the following (all equivalent) forms.

$$\begin{aligned} S_n &= a(1 + r + r^2 + r^3 + \dots + r^n) = a + ar + ar^2 + ar^3 + \dots + ar^n \\ &= \sum_{j=0}^n ar^j = a \sum_{j=0}^n r^j \end{aligned}$$

The number r is called the [ratio of the geometric series](#) because it is the ratio of consecutive terms of the series.

Theorem. The sum of a finite geometric series is given by

$$S_n = a(1 + r + r^2 + r^3 + \dots + r^n) = \frac{a(1 - r^{n+1})}{1 - r}. \quad (17)$$

Proof. This is a standard trick that you've probably seen before.

$$\begin{aligned} S_n &= a + ar + ar^2 + \dots + ar^n \\ rS_n &= \qquad ar + ar^2 + \dots + ar^n + ar^{n+1} \end{aligned}$$

When we subtract these two equations most terms cancel and we get

$$S_n - rS_n = a - ar^{n+1}$$

Some simple algebra now gives us the formula in Equation 17.

Definition. An [infinite geometric series](#) has the same form as the finite geometric series except there is no last term:

$$S = a + ar + ar^2 + \dots = a \sum_{j=0}^{\infty} r^j.$$

Note. We will usually simply say 'geometric series' instead of 'infinite geometric series'.

Theorem. If $|r| < 1$ then the infinite geometric series converges to

$$S = a \sum_{j=0}^{\infty} r^j = \frac{a}{1-r} \quad (18)$$

If $|r| \geq 1$ then the series does not converge.

Proof. This is an easy consequence of the formula for the sum of a finite geometric series. Simply let $n \rightarrow \infty$ in Equation 17.

Note. We have assumed a familiarity with convergence of infinite series. We will go over this in more detail in the appendix to this topic.

7.2.1 Connection to Cauchy's integral formula

Cauchy's integral formula says $f(z) = \frac{1}{2\pi i} \int_C \frac{f(w)}{w-z} dw$. Inside the integral we have the expression $1/(w-z)$ which looks a lot like the sum of a geometric series. We will make frequent use of the following manipulations of this expression.

$$\frac{1}{w-z} = \frac{1}{w} \cdot \frac{1}{1-z/w} = \frac{1}{w} (1 + (z/w) + (z/w)^2 + \dots) \quad (19)$$

The geometric series in this equation has ratio z/w . Therefore, the series converges, i.e. the formula is valid, whenever $|z/w| < 1$, or equivalently when $|z| < |w|$.

$$\frac{1}{w-z} = -\frac{1}{z} \cdot \frac{1}{1-w/z} = -\frac{1}{z} (1 + (w/z) + (w/z)^2 + \dots) \quad (20)$$

The series converges, i.e. the formula is valid, whenever $|w/z| < 1$, or equivalently when $|z| > |w|$.

7.3 Convergence of power series

When we include powers of the variable z in the series we will call it a [power series](#). In this section we'll state the main theorem we need about the convergence of power series. Technical details will be pushed to the appendix for the interested reader.

Theorem 7.1. Consider the power series $f(z) = \sum_{n=0}^{\infty} a_n(z-z_0)^n$. There is a number $R \geq 0$ such that:

1. If $R > 0$ then the series converges absolutely to an analytic function for $|z-z_0| < R$.
2. The series diverges for $|z-z_0| > R$

R is called the [radius of convergence](#). The disk $|z-z_0| < R$ is called the [disk of convergence](#).

3. The derivative is given by term-by-term differentiation

$$f'(z) = \sum_{n=0}^{\infty} n a_n (z-z_0)^{n-1}$$

The series for f' also has radius of convergence R .

4. If γ is a bounded curve inside the disk of convergence then the integral is given by term-by-term integration

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

Notes.

- The theorem doesn't say what happens when $|z - z_0| = R$.
- If $R = \infty$ the function $f(z)$ is entire.
- If $R = 0$ the series only converges at the point $z = z_0$. In this case, the series does not represent an analytic function on any disk around z_0 .
- Often (not always) we can find R using the ratio test.

The proof of this theorem is in the appendix.

7.3.1 Ratio test and root test

Here are two standard tests from calculus on the convergence of infinite series

Ratio test. Consider the series $\sum_0^{\infty} c_n$. If $L = \lim_{n \rightarrow \infty} |c_{n+1}/c_n|$ exists, then:

1. If $L < 1$ then the series converges absolutely.
2. If $L > 1$ then the series diverges.
3. If $L = 1$ then the test gives no information.

Note. In words, L is the limit of the absolute ratios of consecutive terms.

Again the proof will be in the appendix. (It boils down to comparison with a geometric series.)

Example 7.2. Consider the geometric series $1 + z + z^2 + z^3 + \dots$. The limit of the absolute ratios of consecutive terms is

$$L = \lim_{n \rightarrow \infty} \frac{|z^{n+1}|}{|z^n|} = |z|$$

Thus, the ratio test agrees that the geometric series converges when $|z| < 1$. We know this converges to $1/(1-z)$. Note, the disk of convergence ends exactly at the singularity $z = 1$.

Example 7.3. Consider the series $f(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}$. The limit from the ratio test is

$$L = \lim_{n \rightarrow \infty} \frac{|z^{n+1}|/(n+1)!}{|z^n|/n!} = \lim_{n \rightarrow \infty} \frac{|z|}{n+1} = 0.$$

Since $L < 1$ this series converges for every z . Thus, by Theorem 7.1, the radius of convergence for this series is ∞ . That is, $f(z)$ is entire. Of course we know that $f(z) = e^z$.

Root test. Consider the series $\sum_0^{\infty} c_n$. If $L = \lim_{n \rightarrow \infty} |c_n|^{1/n}$ exists, then:

1. If $L < 1$ then the series converges absolutely.
2. If $L > 1$ then the series diverges.
3. If $L = 1$ then the test gives no information .

Note. In words, L is the limit of the n th roots of the (absolute value) of the terms.

The geometric series is so fundamental that we should check the root test on it.

Example 7.4. Consider the geometric series $1 + z + z^2 + z^3 + \dots$. The limit of the n th roots of the terms is

$$L = \lim_{n \rightarrow \infty} |z^n|^{1/n} = \lim |z| = |z|$$

Happily, the root test agrees that the geometric series converges when $|z| < 1$.

7.4 Taylor series

The previous section showed that a power series converges to an analytic function inside its disk of convergence. Taylor's theorem completes the story by giving the converse: around each point of analyticity an analytic function equals a convergent power series.

Theorem 7.5. (Taylor's theorem) Suppose $f(z)$ is an analytic function in a region A . Let $z_0 \in A$. Then,

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

where the series converges on any disk $|z - z_0| < r$ contained in A . Furthermore, we have formulas for the coefficients

$$a_n = \frac{f^{(n)}(z_0)}{n!} \tag{21}$$

$$= \frac{1}{2\pi i} \int_{\gamma} \frac{f(z)}{(z - z_0)^{n+1}} dz \tag{22}$$

We call the series [the power series representing \$f\$ around \$z_0\$](#) .

The proof will be given below. First we look at some consequences of Taylor's theorem.

Corollary. The power series representing an analytic function around a point z_0 is unique. That is, the coefficients are uniquely determined by the function $f(z)$.

Proof. Taylor's theorem gives a formula for the coefficients.

7.4.1 Order of a zero

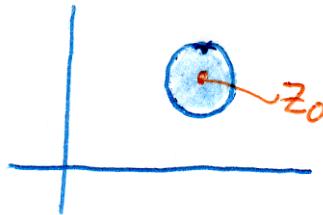
Theorem. Suppose $f(z)$ is analytic on the disk $|z - z_0| < r$ and f is not identically 0. Then there is an integer $k \geq 0$ such that $a_k \neq 0$ and f has Taylor series around z_0 given by

$$f(z) = (z - z_0)^k (a_k + a_{k+1}(z - z_0) + \dots) = (z - z_0)^k \sum_{n=k}^{\infty} a_n (z - z_0)^{n-k}. \tag{23}$$

Proof. Since $f(z)$ is not identically 0, not all the Taylor coefficients are zero. So, we take k to be the index of the first nonzero coefficient.

Theorem 7.6. *Zeros are isolated.* If $f(z)$ is analytic and not identically zero then the zeros of f are isolated.

By isolated we mean that we can draw a small disk around any zeros that doesn't contain any other zeros.



Isolated zero at z_0 : $f(z_0) = 0$, $f(z) \neq 0$ elsewhere in the disk.

Proof. Suppose $f(z_0) = 0$. Write f as in Equation 23. There are two factors: $(z - z_0)^k$ and $g(z) = a_k + a_{k+1}(z - z_0) + \dots$. Clearly $(z - z_0)^k \neq 0$ if $z \neq z_0$. We have $g(z_0) = a_k \neq 0$, so $g(z)$ is not 0 on some small neighborhood of z_0 . We conclude that on this neighborhood the product is only zero when $z = z_0$, i.e. z_0 is an isolated 0.

Definition. The integer k in Theorem 7.6 is called *the order of the zero of f at z_0* .

Note, if $f(z_0) \neq 0$ then z_0 is a zero of order 0.

7.5 Taylor series examples

The uniqueness of Taylor series along with the fact that they converge on any disk around z_0 where the function is analytic allows us to use lots of computational tricks to find the series and be sure that it converges.

Example 7.7. Use the formula for the coefficients in terms of derivatives to give the Taylor series of $f(z) = e^z$ around $z = 0$.

answer: Since $f'(z) = e^z$, we have $f^{(n)}(0) = e^0 = 1$. So,

$$e^z = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \dots = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

Example 7.8. Expand $f(z) = z^8 e^{3z}$ in a Taylor series around $z = 0$.

answer: Let $w = 3z$. So,

$$e^{3z} = e^w = \sum_{n=0}^{\infty} \frac{w^n}{n!} = \sum_{k=0}^{\infty} \frac{3^n}{n!} z^n$$

Thus, $f(z) = \sum_{n=0}^{\infty} \frac{3^n}{n!} z^{n+8}$.

Example 7.9. Find the Taylor series of $\sin(z)$ around $z = 0$ (Sometimes the Taylor series around 0 is called the *Maclaurin series*.)

answer: We give two methods for doing this.

Method 1. $f^{(n)}(0) = \frac{d^n \sin(z)}{dz^n} = \begin{cases} (-1)^m & \text{for } n = 2m + 1 = \text{odd, } m = 0, 1, 2, \dots \\ 0 & \text{for } n \text{ even} \end{cases}$

Method 2. Using $\sin(z) = \frac{e^{iz} - e^{-iz}}{2i}$.

$$\sin(z) = \frac{1}{2i} \left[\sum_{n=0}^{\infty} \frac{(iz)^n}{n!} - \sum_{n=0}^{\infty} \frac{(-iz)^n}{n!} \right] = \frac{1}{2i} \sum_{n=0}^{\infty} [(1 - (-1)^n)] \frac{i^n z^n}{n!}$$

(We need absolute convergence to add series like this.)

Conclusion: $\sin(z) = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!}$, which converges for $|z| < \infty$.

Example 7.10. Expand the rational function $f(z) = \frac{1+2z^2}{z^3+z^5}$ around $z = 0$.

answer: Note that f has a singularity at 0, so we can't expect a convergent Taylor series expansion. We'll aim for the next best thing using the following shortcut.

$$f(z) = \frac{1}{z^3} \frac{2(1+z^2)-1}{1+z^2} = \frac{1}{z^3} \left[2 - \frac{1}{1+z^2} \right].$$

Using the geometric series we have

$$\frac{1}{1+z^2} = \frac{1}{1-(-z^2)} = \sum_{n=0}^{\infty} (-z^2)^n = 1 - z^2 + z^4 - z^6 + \dots$$

Putting it all together

$$f(z) = \frac{1}{z^3} \left(2 - 1 + z^2 - z^4 + \dots \right) = \left(\frac{1}{z^3} + \frac{1}{z} \right) - \sum_{n=0}^{\infty} (-1)^n z^{2n+1}$$

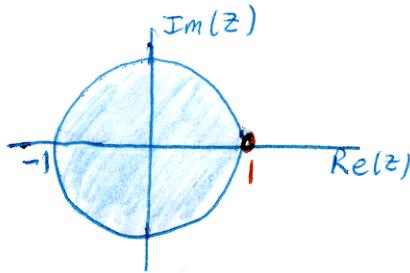
Note. The first terms are called the **singular part**, i.e. those with negative powers of z . The summation is called the regular or analytic part. Since the geometric series for $1/(1+z^2)$ converges for $|z| < 1$, the entire series is valid in $0 < |z| < 1$

Example 7.11. Find the Taylor series for $f(z) = \frac{e^z}{1-z}$ around $z = 0$. Give the radius of convergence.

answer: We start by writing the Taylor series for each of the factors and then multiply them out.

$$\begin{aligned} f(z) &= \left(1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \dots \right) (1 + z + z^2 + z^3 + \dots) \\ &= 1 + (1+1)z + \left(1 + 1 + \frac{1}{2!} \right) z^2 + \left(1 + 1 + \frac{1}{2!} + \frac{1}{3!} \right) z^3 + \dots \end{aligned}$$

The biggest disk around $z = 0$ where f is analytic is $|z| < 1$. Therefore, by Taylor's theorem, the radius of convergence is $R = 1$.



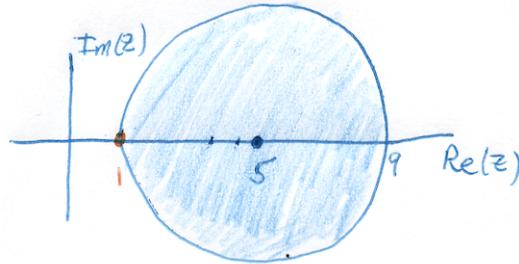
$f(z)$ analytic on $|z| < 1$, has a singularity at $z = 1$.

Example 7.12. Find the Taylor series for $f(z) = \frac{1}{1-z}$ around $z = 5$. Give the radius of convergence.

answer: We have to manipulate this into standard geometric series form.

$$f(z) = \frac{1}{-4(1 + (z-5)/4)} = -\frac{1}{4} \left(1 - \left(\frac{z-5}{4} \right) + \left(\frac{z-5}{4} \right)^2 - \left(\frac{z-5}{4} \right)^3 + \dots \right)$$

Since $f(z)$ has a singularity at $z = 1$ the radius of convergence is $R = 4$. We can also see this by considering the geometric series. The geometric series ratio is $(z-5)/4$. So the series converges when $|z-5|/4 < 1$, i.e. when $|z-5| < 4$, i.e. $R = 4$.



Disk of convergence stops at the singularity at $z = 1$.

Example 7.13. Find the Taylor series for $f(z) = \log(1+z)$ around $z = 0$. Give the radius of convergence.

answer: We know that f is analytic for $|z| < 1$ and not analytic at $z = -1$. So, the radius of convergence is $R = 1$. To find the series representation we take the derivative and use the geometric series.

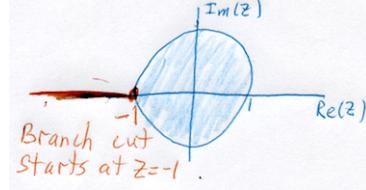
$$f'(z) = \frac{1}{1+z} = 1 - z + z^2 - z^3 + z^4 - \dots$$

Integrating term by term (allowed by Theorem 7.1) we have

$$f(z) = a_0 + z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} + \dots = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{z^n}{n}$$

Here a_0 is the constant of integration. We find it by evaluating at $z = 0$.

$$f(0) = a_0 = \log(1) = 0.$$



Disk of convergence for $\log(1+z)$ around $z = 0$.

Example 7.14. Can the series $\sum a_n(z-2)^n$ converge at $z = 0$ and diverge at $z = 3$.

answer: No! We have $z_0 = 2$. We know the series diverges everywhere outside its radius of convergence. So, if the series converges at $z = 0$, then the radius of convergence is at least 2. Since $|3 - z_0| < 2$ we would also have that $z = 3$ is inside the disk of convergence.

7.5.1 Proof of Taylor's theorem

For convenience we restate Taylor's Theorem 7.5.

Taylor's theorem. (Taylor series) Suppose $f(z)$ is an analytic function in a region A . Let $z_0 \in A$. Then,

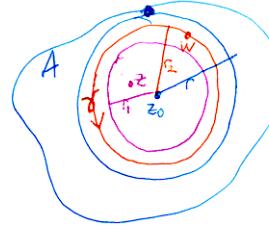
$$f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n,$$

where the series converges on any disk $|z - z_0| < r$ contained in A . Furthermore, we have formulas for the coefficients

$$a_n = \frac{f^{(n)}(z_0)}{n!} \quad (24)$$

$$= \frac{1}{2\pi i} \int_{\gamma} \frac{f(z)}{(z - z_0)^{n+1}} dz \quad (25)$$

Proof. In order to handle convergence issues we fix $0 < r_1 < r_2 < r$. We let γ be the circle $|w - z_0| = r_2$ (traversed counterclockwise).



Disk of convergence extends to the boundary of A
 $r_1 < r_2 < r$, but r_1 and r_2 can be arbitrarily close to r .

Take z inside the disk $|z - z_0| < r_1$. We want to express $f(z)$ as a power series around z_0 . To do this we start with the Cauchy integral formula and then use the geometric series.

As preparation we note that for w on γ and $|z - z_0| < r_1$ we have $|z - z_0| < r_1 < r_2 = |w - z_0|$, so $|z - z_0|/|w - z_0| < 1$. Therefore,

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

Using this and the Cauchy formula gives

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

The last equality follows from Cauchy's formula for derivatives. Taken together the last two equalities give Taylor's formula. QED

7.6 Singularities

Definition. A function $f(z)$ is **singular** at a point z_0 if it is not analytic at z_0

Definition. For a function $f(z)$, the singularity z_0 is an **isolated singularity** if f is analytic on the deleted disk $0 < |z - z_0| < r$ for some $r > 0$.

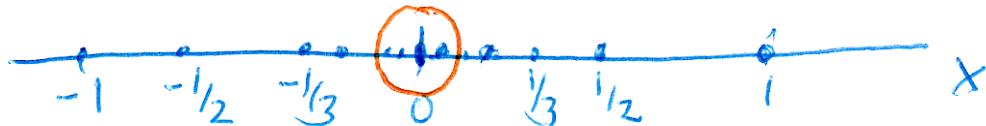
Example 7.15. $f(z) = \frac{z+1}{z^3(z^2+1)}$ has isolated singularities at $z = 0, \pm i$.

Example 7.16. $f(z) = e^{1/z}$ has an isolated singularity at $z = 0$.

Example 7.17. $f(z) = \log(z)$ has a singularity at $z = 0$, but it is not isolated because a branch cut, starting at $z = 0$, is needed to have a region where f is analytic.

Example 7.18. $f(z) = \frac{1}{\sin(\pi/z)}$ has singularities at $z = 0$ and $z = 1/n$ for $n = \pm 1, \pm 2, \dots$

The singularities at $\pm 1/n$ are isolated, but the one at $z = 0$ is not isolated.



Every neighborhood of 0 contains zeros at $1/n$ for large n .

7.7 Laurent series

Theorem 7.19. (Laurent series). Suppose that $f(z)$ is analytic on the annulus $A : r_1 < |z - z_0| < r_2$. Then $f(z)$ can be expressed as a series

$$f(z) = \sum_{n=1}^{\infty} \frac{b_n}{(z-z_0)^n} + \sum_{n=0}^{\infty} a_n (z-z_0)^n.$$

The coefficients have the formulas

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

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

where γ is any circle $|w - z_0| = r$ inside the annulus, i.e. $r_1 < r < r_2$.

Furthermore

- The series $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ converges to an analytic function for $|z - z_0| < r_2$.
- The series $\sum_{n=1}^{\infty} \frac{b_n}{(z - z_0)^n}$ converges to an analytic function for $|z - z_0| > r_1$.
- Together, the series both converge on the annulus A where f is analytic.

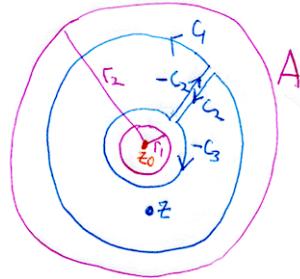
The proof is given below. First we define a few terms.

Definition. The entire series is called the [Laurent series for \$f\$ around \$z_0\$](#) . The series $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ is called the [analytic or regular part of the Laurent series](#). The series $\sum_{n=1}^{\infty} \frac{b_n}{(z - z_0)^n}$ is called the [singular or principal part of the Laurent series](#).

Note. Since $f(z)$ is not defined at z_0 we don't have any formulas for the coefficients using derivatives.

Proof (Laurent series). Choose a point z in A . Now set circles C_1 and C_3 close enough to the boundary that z is inside $C_1 + C_2 - C_3 - C_2$ as shown. Since this curve and its interior are contained in A , Cauchy's integral formula says

$$f(z) = \frac{1}{2\pi i} \int_{C_1 + C_2 - C_3 - C_2} \frac{f(w)}{w - z} dw$$



The contour used for proving the formulas for Laurent series.

The integrals over C_2 cancel, so we have

$$f(z) = \frac{1}{2\pi i} \int_{C_1 - C_3} \frac{f(w)}{w - z} dw.$$

Next, we divide this into two pieces and use our trick of converting to a geometric series. The calculations are just like the proof of Taylor's theorem. On C_1 we have $|z - z_0|/|w - z_0| < 1$, so

$$\begin{aligned} \frac{1}{2\pi i} \int_{C_1} \frac{f(w)}{w - z} dw &= \frac{1}{2\pi i} \int_{C_1} \frac{f(w)}{w - z_0} \cdot \frac{1}{(1 - (z - z_0)/(w - z_0))} dw \\ &= \frac{1}{2\pi i} \int_{C_1} \sum_{n=0}^{\infty} \frac{f(w)}{(w - z_0)^{n+1}} (z - z_0)^n dw \\ &= \sum_{n=0}^{\infty} \left(\frac{1}{2\pi i} \int_{C_1} \frac{f(w)}{(w - z_0)^{n+1}} dw \right) (z - z_0)^n \\ &= \sum_{n=0}^{\infty} a_n (z - z_0)^n. \end{aligned}$$

Here a_n is defined by the integral formula given in the statement of the theorem. Examining the above argument we see that the only requirement on z is that $|z - z_0| < r_2$. So, this series converges for all such z .

Similarly on C_3 we have $|w - z_0|/|z - z_0| < 1$, so

$$\begin{aligned} \frac{1}{2\pi i} \int_{C_3} \frac{f(w)}{w - z} dw &= \frac{1}{2\pi i} \int_{C_3} -\frac{f(w)}{z - z_0} \cdot \frac{1}{(1 - (w - z_0)/(z - z_0))} dw \\ &= -\frac{1}{2\pi i} \int_{C_3} \sum_{n=0}^{\infty} \frac{f(w)}{(z - z_0)^{n+1}} (w - z_0)^n dw \\ &= -\frac{1}{2\pi i} \sum_{n=0}^{\infty} \left(\int_{C_1} \frac{f(w)}{(w - z_0)^n} dw \right) (z - z_0)^{-n-1} \\ &= -\sum_{n=1}^{\infty} \frac{b_n}{(z - z_0)^n}. \end{aligned}$$

In the last equality we changed the indexing to match the indexing in the statement of the theorem. Here b_n is defined by the integral formula given in the statement of the theorem. Examining the above argument we see that the only requirement on z is that $|z - z_0| > r_1$. So, this series converges for all such z .

Combining these two formulas we have

$$f(z) = \frac{1}{2\pi i} \int_{C_1 - C_3} \frac{f(w)}{w - z} dw = \sum_{n=1}^{\infty} \frac{b_n}{(z - z_0)^n} + \sum_{n=0}^{\infty} a_n (z - z_0)^n$$

The last thing to note is that the integrals defining a_n and b_n do not depend on the exact radius of the circle of integration. Any circle inside A will produce the same values. We have proved all the statements in the theorem on Laurent series. QED

7.7.1 Examples of Laurent series

In general, the integral formulas are not a practical way of computing the Laurent coefficients. Instead we use various algebraic tricks. Even better, as we shall see, is the fact

that often we don't really need all the coefficients and we will develop more techniques to compute those that we do need.

Example 7.20. Find the Laurent series for $f(z) = \frac{z+1}{z}$ around $z_0 = 0$. Give the region where it is valid.

answer: $f(z) = 1 + \frac{1}{z}$. This is a Laurent series, valid on the infinite region $0 < |z| < \infty$.

Example 7.21. Find the Laurent series for $f(z) = \frac{z}{z^2 + 1}$ around $z_0 = i$. Give the region where your answer is valid. Identify the singular (principal) part.

answer: Using partial fractions we have $f(z) = \frac{1}{2} \cdot \frac{1}{z-i} + \frac{1}{2} \cdot \frac{1}{z+i}$. Since $\frac{1}{z+i}$ is analytic at $z = i$ it has a Taylor series expansion. We find it using geometric series.

$$\frac{1}{z+i} = \frac{1}{2i} \cdot \frac{1}{1 + (z-i)/(2i)} = \frac{1}{2i} \sum_{n=0}^{\infty} \left(-\frac{z-i}{2i} \right)^n$$

So the Laurent series is

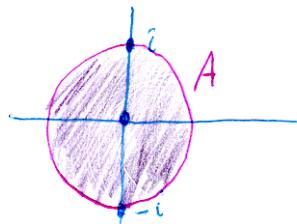
$$f(z) = \frac{1}{2} \cdot \frac{1}{z-i} + \frac{1}{4i} \sum_{n=0}^{\infty} \left(-\frac{z-i}{2i} \right)^n$$

The singular (principal) part is given by the first term. The region of convergence is $0 < |z-i| < 2$.

Note. We could have looked at $f(z)$ on the region $2 < |z-i| < \infty$. This would have produced a different Laurent series. We discuss this further in an upcoming example.

Example 7.22. Compute the Laurent series for $f(z) = \frac{z+1}{z^3(z^2+1)}$ on the region $A : 0 < |z| < 1$ centered at $z = 0$.

answer: This function has isolated singularities at $z = 0, \pm i$. Therefore it is analytic on the region A .



$f(z)$ has singularities at $z = 0, \pm i$.

At $z = 0$ we have $f(z) = \frac{1}{z^3}(1+z)(1-z^2+z^4-z^6+\dots)$. Multiplying this out we get

$$f(z) = \frac{1}{z^3} + \frac{1}{z^2} - \frac{1}{z} - 1 + z + z^2 - z^3 - \dots$$

The following example shows that the Laurent series depends on the region under consideration.

Example 7.23. Find the Laurent series around $z = 0$ for $f(z) = \frac{1}{z(z-1)}$ in each of the following regions:

- (i) the region $A_1 : 0 < |z| < 1$
- (ii) the region $A_2 : 1 < |z| < \infty$.

answer: (i) $f(z) = -\frac{1}{z} \cdot \frac{1}{1-z} = -\frac{1}{z}(1+z+z^2+\dots) = -\frac{1}{z} - 1 - z - z^2 - \dots$

(ii) Since the usual geometric series for $1/(1-z)$ does not converge on A_2 we need a different form.

$$f(z) = \frac{1}{z} \cdot \frac{1}{z(1-1/z)} = \frac{1}{z^2} \left(1 + \frac{1}{z} + \frac{1}{z^2} + \dots \right)$$

Since $|1/z| < 1$ on A_2 our use of the geometric series is justified.

One lesson from this example is that the Laurent series depends on the region as well as the formula for the function.

7.8 Digression to differential equations

Here is a standard use of series for solving differential equations.

Example 7.24. Find a power series solution to the equation $f'(x) = f(x) + 2$, $f(0) = 0$.

answer: We look for a solution of the $f(x) = \sum_{n=0}^{\infty} a_n x^n$. Using the initial condition we find $f(0) = 0 = a_0$. Substituting the series into the differential equation we get

$$f'(x) = a_1 + 2a_2 x + 3a_3 x^3 + \dots = f(x) + 2 = a_0 + 2 + a_1 x + a_2 x^2 + \dots$$

Equating coefficients and using $a_0 = 0$ we have

$$a_1 = a_0 + 2 = 2$$

$$2a_2 = a_1, \text{ so } a_2 = a_1/2 = 1$$

$$3a_3 = a_2, \text{ so } a_3 = 1/3$$

$$4a_4 = a_3, \text{ so } a_4 = 1/(3 \cdot 4)$$

In general

$$(n+1)a_{n+1} = a_n, \text{ so } a_{n+1} = \frac{a_n}{(n+1)} = \frac{1}{(3 \cdot 4 \cdot 5 \cdots (n+1))}.$$

You can check using the ratio test that this function is entire.

7.9 Poles

Poles are about isolated singularities. So, we suppose $f(z)$ is analytic on $0 < |z - z_0| < r$ and has Laurent series

$$f(z) = \sum_{n=1}^{\infty} \frac{b_n}{(z - z_0)^n} + \sum_{n=0}^{\infty} a_n (z - z_0)^n.$$

Definition of poles. If only a finite number of the coefficients b_n are nonzero we say z_0 is a **finite pole** of f . In this case, if $b_k \neq 0$ and $b_n = 0$ for all $n > k$ then we say z_0 is a **pole of order k** .

If z_0 is a pole of order 1 we say it is a **simple pole** of f .

If an infinite number of the b_n are nonzero we say that z_0 is an [essential singularity](#) or a [pole of infinite order](#) of f .

If all the b_k are 0, then z_0 is called a [removable singularity](#). That is, if we define $f(z_0) = a_0$ then f is analytic on the disk $|z - z_0| < r$. The terminology can be a bit confusing. So, imagine that I tell you that f is defined and analytic on the punctured disk $0 < |z - z_0| < r$. Then, a priori we assume f has a singularity at z_0 . But, if after computing the Laurent series we see there is no singular part we can extend the definition of f to the full disk, thereby ‘removing the singularity’.

We can explain the term essential singularity as follows. If $f(z)$ has a pole of order k at z_0 then $(z - z_0)^k f(z)$ is analytic (has a removable singularity) at z_0 . So, $f(z)$ itself is not much harder to work with than an analytic function. On the other hand, if z_0 is an essential singularity then no algebraic trick will change $f(z)$ into an analytic function at z_0 .

7.9.1 Examples of poles

We’ll go back through many of the examples from the previous sections.

Example 7.25. The rational function $f(z) = \frac{1+2z^2}{z^3+z^5}$ expanded to

$$f(z) = \left(\frac{1}{z^3} + \frac{1}{z} \right) - \sum_{n=0}^{\infty} (-1)^n z^{2n+1}.$$

Thus, $z = 0$ is a pole of order 3.

Example 7.26. $f(z) = \frac{z+1}{z} = 1 + \frac{1}{z}$. Thus, $z = 0$ is a pole of order 1, i.e. a simple pole.

Example 7.27. $f(z) = \frac{z}{z^2+1} = \frac{1}{2} \cdot \frac{1}{z-i} + g(z)$, where $g(z)$ is analytic at $z = i$. So, $z = i$ is a simple pole.

Example 7.28. For the function $f(z) = \frac{1}{z(z-1)}$.

(i) On the region $A_1 : 0 < |z| < 1$

$$f(z) = -\frac{1}{z} \cdot \frac{1}{1-z} = -\frac{1}{z}(1+z+z^2+\dots) = -\frac{1}{z} - 1 - z - z^2 - \dots$$

So, for $f(z)$ defined on this region, $z = 0$ is a simple pole

(ii) On the region $A_2 : 1 < |z| < \infty$ the point $z = 0$ is not an isolated singularity, so it makes no sense to talk about it as a pole. In general, when working with poles we will want the function to be analytic on an arbitrarily small curve, around the pole. In this example that would not be possible since the curve must always be outside the circle of radius 1.

Example 7.29. $e^{1/z} = 1 + \frac{1}{z} + \frac{1}{z^2 2!} + \frac{1}{z^3 3!} + \dots$. So, $z = 0$ is an essential singularity.

Example 7.30. $\log(z)$ has a singularity at $z = 0$. Since the singularity is not isolated, it can’t be classified as either a pole or an essential singularity.

7.9.2 Residues

In preparation for discussing the residue theorem in the next topic we give the definition and a few examples here.

Note well, residues have to do with isolated singularities.

Definition 7.31. Consider the function $f(z)$ with an isolated singularity at z_0 , i.e. defined on $0 < |z - z_0| < r$ and with Laurent series

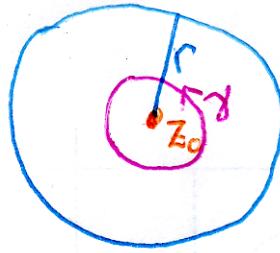
$$f(z) = \sum_{n=1}^{\infty} \frac{b_n}{(z - z_0)^n} + \sum_{n=0}^{\infty} a_n(z - z_0)^n.$$

The residue of f at z_0 is b_1 . This is denoted

$$\text{Res}(f, z_0) \quad \text{or} \quad \underset{z=z_0}{\text{Res}} f = b_1.$$

What is the importance of the residue? If γ is a small, simple closed curve that goes counterclockwise around b_1 then

$$\int_{\gamma} f(z) = 2\pi i b_1.$$



γ is small enough to be inside $|z - z_0| < r$, and surround z_0 .

This is easy to see by integrating the Laurent series term by term. The only nonzero integral comes from the term b_1/z .

Example 7.32. $f(z) = e^{1/2z} = 1 + \frac{1}{2z} + \frac{1}{2(2z)^2} + \dots$ has an isolated singularity at 0. From the Laurent series we see that $\text{Res}(f, 0) = 1/2$.

7.10 Appendix: convergence

This section needs to be completed. It will give some of the careful technical definitions and arguments regarding convergence and manipulation of series. In particular it will define the notion of uniform convergence. The short description is that all of our manipulations of power series are justified on any closed bounded region. Almost, everything we did can be restricted to a closed disk or annulus, and so was valid.

8 Residue Theorem

8.1 Poles and zeros

We remind you of the following terminology.

Suppose $f(z)$ is analytic at z_0 and $f(z) = a_n(z - z_0)^n + a_{n+1}z^{n+1} + \dots$, with $a_n \neq 0$. Then we say f has a **zero of order n at z_0** . If $n = 1$ we say z_0 is a **simple zero**.

Suppose f has an *isolated* singularity at z_0 and Laurent series

$$f(z) = \frac{b_n}{(z - z_0)^n} + \frac{b_{n-1}}{(z - z_0)^{n-1}} + \dots + \frac{b_1}{z} + a_0 + a_1(z - z_0) + \dots$$

which converges on $0 < |z - z_0| < R$ and with $b_n \neq 0$. Then we say f has a **pole of order n at z_0** . If $n = 1$ we say z_0 is a **simple pole**.

There are several examples in the topic 7 notes. Here is one more

Example 8.1. $f(z) = \frac{z+1}{z^3(z^2+1)}$ has isolated singularities at $z = 0, \pm i$ and a zero at $z = -1$. We will show that $z = 0$ is a pole of order 3, $z = \pm i$ are poles of order 1 and $z = -1$ is a zero of order 1. The style of argument is the same in each case.

At $z = 0$: $f(z) = \frac{1}{z^3} \cdot \frac{z+1}{z^2+1}$. Call the second factor $g(z)$. Since $g(z)$ is analytic at $z = 0$ and $g(0) = 1$, it has a Taylor series

$$g(z) = \frac{z+1}{z^2+1} = 1 + a_1z + a_2z^2 + \dots$$

Therefore $f(z) = \frac{1}{z^3} + \frac{a_1}{z^2} + \frac{a_2}{z} + \dots$. This shows $z = 0$ is a pole of order 3.

At $z = i$: $f(z) = \frac{1}{z-i} \cdot \frac{z+1}{z^3(z+i)}$. Call the second factor $g(z)$. Since $g(z)$ is analytic at $z = i$, it has a Taylor series

$$g(z) = \frac{z+1}{z^3(z+i)} = a_0 + a_1(z-i) + a_2(z-i)^2 + \dots$$

where $a_0 = g(i) \neq 0$. Therefore $f(z) = \frac{a_0}{z-i} + a_1 + a_2(z-i) + \dots$. This shows $z = i$ is a pole of order 1.

The arguments for $z = -i$ and $z = -1$ are similar.

8.2 Words: Holomorphic and meromorphic

Definition. A function that is analytic on a region A is called **holomorphic on A** . A function that is analytic on A except for a set of poles of finite order is called **meromorphic on A** .

Example 8.2. Let $f(z) = \frac{z+z^2+z^3}{(z-2)(z-3)(z-4)(z-5)}$. This is meromorphic on \mathbf{C} with (simple) poles at $z = 2, 3, 4, 5$.

8.3 Behavior of functions near zeros and poles

The basic idea is that near a zero of order n , a function behaves like $(z - z_0)^n$ and near a pole of order n , a function behaves like $1/(z - z_0)^n$. The following make this a little more precise.

Behavior near a zero. If f has a zero of order n at z_0 then near z_0 , $f(z) \approx a_n(z - z_0)^n$, for some constant a_n .

Proof. By definition f has a Taylor series around z_0 of the form

$$f(z) = a_n(z - z_0)^n + a_{n+1}(z - z_0)^{n+1} + \dots = a_n(z - z_0)^n \left(1 + \frac{a_{n+1}}{a_n}(z - z_0) + \frac{a_{n+2}}{a_n}(z - z_0)^2 + \dots \right)$$

Since the second factor equals 1 at z_0 , the claim follows.

Behavior near a finite pole. If f has a pole of order n at z_0 then near z_0 , $f(z) \approx \frac{b_n}{(z - z_0)^n}$, for some constant b_n .

Proof. This is nearly identical to the previous argument. By definition f has a Laurent series around z_0 of the form

$$\begin{aligned} f(z) &= \frac{b_n}{(z - z_0)^n} + \frac{b_{n-1}}{(z - z_0)^{n-1}} + \dots + \frac{b_1}{z - z_0} + a_0 + \dots \\ &= \frac{b_n}{(z - z_0)^n} \left(1 + \frac{b_{n-1}}{b_n}(z - z_0) + \frac{b_{n-2}}{b_n}(z - z_0)^2 + \dots \right) \end{aligned}$$

Since the second factor equals 1 at z_0 , the claim follows.

8.3.1 Picard's theorem and essential singularities

Near an essential singularity we have Picard's theorem. We won't prove or make use of this theorem in 18.04. Still, we feel it is pretty enough to warrant showing to you.

Picard's theorem. If $f(z)$ has an essential singularity at z_0 then in every neighborhood of z_0 , $f(z)$ takes on all possible values infinitely many times, with the possible exception of one value.

Example 8.3. It is easy to see that in any neighborhood of $z = 0$ the function $w = e^{1/z}$ takes every value except $w = 0$.

8.3.2 Quotients of functions

We have the following statement about quotients of functions. We could make similar statements if one or both functions has a pole instead of a zero.

Theorem. Suppose f has a zero of order m at z_0 and g has a zero of order n at z_0 . Let $h(z) = f(z)/g(z)$. Then

- If $n > m$ then $h(z)$ has a pole of order $n - m$ at z_0 .
- If $n < m$ then $h(z)$ has a zero of order $m - n$ at z_0 .

- If $n = m$ then $h(z)$ is analytic and nonzero at z_0 .

We can paraphrase this as $h(z)$ has ‘pole’ of order $n - m$ at z_0 . If $n - m$ is negative then the ‘pole’ is actually a zero.

Proof. You should be able to supply the proof. It is nearly identical to the proofs above: express f and g as Taylor series and take the quotient.

Example 8.4. Let $h(z) = \frac{\sin(z)}{z^2}$. We know $\sin(z)$ has a zero of order 1 at $z = 0$ and z^2 has a zero of order 2. So, $h(z)$ has a pole of order 1 at $z = 0$. Of course, we can see this easily using Taylor series

$$h(z) = \frac{1}{z^2} \left(z - \frac{z^3}{3!} + \dots \right)$$

8.4 Residues

In this section we’ll explore calculating residues. We’ve seen enough already to know that this will be useful. We will see that even more clearly when we look at the residue theorem in the next section.

We introduced residues in the previous topic. We repeat the definition here for completeness.

Definition. Consider the function $f(z)$ with an isolated singularity at z_0 , i.e. defined on the region $0 < |z - z_0| < r$ and with Laurent series (on that region)

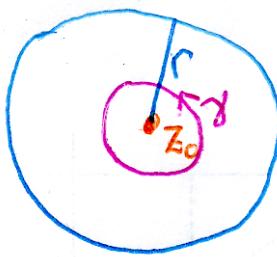
$$f(z) = \sum_{n=1}^{\infty} \frac{b_n}{(z - z_0)^n} + \sum_{n=0}^{\infty} a_n(z - z_0)^n.$$

The residue of f at z_0 is b_1 . This is denoted

$$\text{Res}(f, z_0) = b_1 \quad \text{or} \quad \underset{z=z_0}{\text{Res}} f = b_1.$$

What is the importance of the residue? If γ is a small, simple closed curve that goes counterclockwise around b_1 then

$$\int_{\gamma} f(z) dz = 2\pi i b_1.$$



γ small enough to be inside $|z - z_0| < r$, surround z_0 and contain no other singularity of f .

This is easy to see by integrating the Laurent series term by term. The only nonzero integral comes from the term b_1/z .

Example 8.5. $f(z) = e^{1/2z} = 1 + \frac{1}{2z} + \frac{1}{2(2z)^2} + \dots$ has an isolated singularity at 0. From the Laurent series we see that $\text{Res}(f, 0) = 1/2$.

Example 8.6. 1. Let $f(z) = \frac{1}{z^3} + \frac{2}{z^2} + \frac{4}{z} + 5 + 6z$. f has a pole of order 3 at $z = 0$ and $\text{Res}(f, 0) = 4$.

2. Suppose $f(z) = \frac{2}{z} + g(z)$, where g is analytic at $z = 0$. Then, f has a simple pole at 0 and $\text{Res}(f, 0) = 2$.

3. Let $f(z) = \cos(z) = 1 - z^2/2! + \dots$. Then f is analytic at $z = 0$ and $\text{Res}(f, 0) = 0$.

4. Let

$$f(z) = \frac{\sin(z)}{z} = \frac{1}{z} \left(z - \frac{z^3}{3!} + \dots \right) = 1 - \frac{z^2}{3!} + \dots$$

So, f has a removable singularity at $z = 0$ and $\text{Res}(f, 0) = 0$.

Example 8.7. Using partial fractions. Let $f(z) = \frac{z}{z^2 + 1}$. Find the poles and residues of f .

answer: Using partial fractions we write

$$f(z) = \frac{z}{(z - i)(z + i)} = \frac{1}{2} \cdot \frac{1}{z - i} + \frac{1}{2} \cdot \frac{1}{z + i}.$$

The poles are at $z = \pm i$. We compute the residues at each pole:

At $z = i$: $f(z) = \frac{1}{2} \cdot \frac{1}{z - i} + \text{something analytic at } i$. Therefore the pole is simple and $\text{Res}(f, i) = 1/2$.

At $z = -i$: $f(z) = \frac{1}{2} \cdot \frac{1}{z + i} + \text{something analytic at } -i$. Therefore the pole is simple and $\text{Res}(f, -i) = 1/2$.

Example 8.8. Mild warning! Let $f(z) = -\frac{1}{z(1-z)}$ then we have the following Laurent expansions for f around $z = 0$.

On $0 < |z| < 1$: $f(z) = -\frac{1}{z} \cdot \frac{1}{1-z} = -\frac{1}{z}(1 + z + z^2 + \dots)$. Therefore the pole at $z = 0$ is simple and $\text{Res}(f, 0) = -1$.

On $1 < |z| < \infty$: $f(z) = \frac{1}{z^2} \cdot \frac{1}{1-1/z} = \frac{1}{z^2} \left(1 + \frac{1}{z} + \frac{1}{z^2} + \dots \right)$. Even though this is a valid Laurent expansion you **must not** use it to compute the residue at 0. This is because the definition of residue requires that we use the Laurent series on the region $0 < |z - z_0| < r$.

Example 8.9. Let $f(z) = \log(1+z)$. This has a singularity at $z = -1$, but it is not isolated, so not a pole and therefore there is no residue at $z = -1$.

8.4.1 Residues at simple poles

Simple poles occur frequently enough that we'll study computing their residues in some detail. Here are a number of ways to spot a simple pole and compute its residue. The justification for all of them goes back to Laurent series.

Suppose $f(z)$ has an isolated singularity at $z = z_0$. Then we have the following properties.

Property 1. If the Laurent series for $f(z)$ has the form

$$\frac{b_1}{z - z_0} + a_0 + a_1(z - z_0) + \dots$$

then f has a simple pole at z_0 and $\text{Res}(f, z_0) = b_1$.

Property 2 If $g(z) = (z - z_0)f(z)$ is analytic at z_0 then z_0 is either a simple pole or a removable singularity. In either case $\text{Res}(f, z_0) = g(z_0)$. (In the removable singularity case the residue is 0.)

Proof. Directly from the Laurent series for f around z_0 .

Property 3. If f has a simple pole at z_0 then

$$\lim_{z \rightarrow z_0} (z - z_0)f(z) = \text{Res}(f, z_0)$$

This says that the limit exists and equals the residue. Conversely, if the limit exists then either the pole is simple, or f is analytic at z_0 . In both cases the limit equals the residue.

Proof. Directly from the Laurent series for f around z_0 .

Property 4. If f has a simple pole at z_0 and $g(z)$ is analytic at z_0 then $\text{Res}(fg, z_0) = g(z_0)\text{Res}(f, z_0)$. If $g(z_0) \neq 0$ then $\text{Res}(f/g, z_0) = \frac{1}{g(z_0)}\text{Res}(f, z_0)$.

Proof. Since z_0 is a simple pole,

$$f(z) = \frac{b_1}{z - z_0} + a_0 + a_1(z - z_0)$$

Since g is analytic,

$$g(z) = c_0 + c_1(z - z_0) + \dots, \text{ where } c_0 = g(z_0).$$

Multiplying these series together it is clear that

$$\text{Res}(fg, z_0) = c_0 b_1 = g(z_0) \text{Res}(f, z_0). \quad \text{QED}$$

The statement about quotients f/g follows from the for products because $1/g$ is analytic at z_0 .

Property 5. If $g(z)$ has a simple zero at z_0 then $1/g(z)$ has a simple pole at z_0 and $\text{Res}(1/g, z_0) = 1/g'(z_0)$.

Proof. The algebra for this is similar to what we've done several times above. The Taylor expansion for g is

$$g(z) = a_1(z - z_0) + a_2(z - z_0)^2 + \dots, \text{ where } a_1 = g'(z_0).$$

So

$$\frac{1}{g(z)} = \frac{1}{a_1(z - z_0)} \left(\frac{1}{1 + \frac{a_2}{a_1}(z - z_0) + \dots} \right)$$

The second factor on the right is analytic at z_0 and equals 1 at z_0 . Therefore we know the Laurent expansion of $1/g$ is

$$\frac{1}{g(z)} = \frac{1}{a_1(z - z_0)} (1 + c_1(z - z_0) + \dots)$$

Clearly the residue is $1/a_1 = 1/g'(z_0)$. QED.

Example 8.10. Let $f(z) = \frac{2+z+z^2}{(z-2)(z-3)(z-4)(z-5)}$. Show all the poles are simple and compute their residues.

answer: The poles are at $z = 2, 3, 4, 5$. They are all isolated. We'll look at $z = 2$ the others are similar. Multiplying by $z - 2$ we get

$$g(z) = (z - 2)f(z) = \frac{2+z+z^2}{(z-3)(z-4)(z-5)}.$$

This is analytic at $z = 2$ and $g(z) = \frac{8}{-6} = -\frac{4}{3}$. So the pole is simple and $\text{Res}(f, 2) = -4/3$.

Example 8.11. Let $f(z) = \frac{1}{\sin(z)}$. Find all the poles and their residues.

answer: The poles of $f(z)$ are the zeros of $\sin(z)$, i.e. $n\pi$ for n an integer. Since the derivative $\sin'(n\pi) = \cos(n\pi) \neq 0$, the zeros are simple and by Property 5 above

$$\text{Res}(f, n\pi) = \frac{1}{\cos(n\pi)} = (-1)^n.$$

Example 8.12. Let $f(z) = \frac{1}{z(z^2+1)(z-2)^2}$. Identify all the poles and say which ones are simple.

answer: Clearly the poles are at $z = 0, \pm i, 2$.

At $z = 0$: $g(z) = zf(z)$ is analytic at 0 and $g(0) = 1/4$. So the pole is simple and the residue is $g(0) = 1/4$.

At $z = i$: $g(z) = (z - i)f(z) = \frac{1}{z(z+i)(z-2)^2}$ is analytic at i , the pole is simple and the residue is $g(i)$.

At $z = -i$: This is similar to the case $z = i$. The pole is simple.

At $z = 2$: $g(z) = (z - 2)f(z) = \frac{1}{z(z^2+1)(z-2)}$ is not analytic at 2, so the pole is not simple. (It should be obvious that it's a pole of order 2.)

Example 8.13. Let $p(z), q(z)$ be analytic at $z = z_0$. Assume $p(z_0) \neq 0, q(z_0) = 0, q'(z_0) \neq 0$. Find $\text{Res}_{z=z_0} \frac{p(z)}{q(z)}$.

answer: Since $q'(z_0) \neq 0, q$ has a simple zero at z_0 . So $1/q$ has a simple pole at z_0 and

$$\text{Res}(1/q, z_0) = 1/q'(z_0)$$

Since $p(z_0) \neq 0$ we know $\text{Res}(p/q, z_0) = p(z_0) \text{Res}(1/q, z_0) = p(z_0)/q'(z_0)$.

8.4.2 Residues at finite poles

For higher order poles we can make statements similar to those for simple poles, but the formulas and computations are more involved. The general principle is the following

Higher order poles. If $f(z)$ has a pole of order k at z_0 then $g(z) = (z - z_0)^k f(z)$ is analytic at z_0 and if

$$g(z) = a_0 + a_1(z - z_0) + \dots$$

$$\text{then } \text{Res}(f, z_0) = a_{k-1} = \frac{g^{(k-1)}(z_0)}{(k-1)!}.$$

Proof. This is clear using Taylor and Laurent series for g and f .

Example 8.14. Let $f(z) = \frac{\sinh(z)}{z^5}$ find the residue at $z = 0$.

answer: We know the Taylor series for $\sinh(z) = z + z^3/3! + z^5/5! + \dots$ (You can find this using $\sinh(z) = (\mathrm{e}^z - \mathrm{e}^{-z})/2$ and the Taylor series for e^z .) Therefore,

$$f(z) = \frac{1}{z^4} - \frac{1}{3!z^2} + \frac{1}{5!} + \dots$$

We see $\text{Res}(f, 0) = 0$.

Note, we could have seen this by realizing that $f(z)$ is an even function.

Example 8.15. Let $f(z) = \frac{\sinh(z)\mathrm{e}^z}{z^5}$ find the residue at $z = 0$.

answer: It is clear that $\text{Res}(f, 0)$ equals the coefficient of z^4 in the Taylor expansion of $\sinh(z)\mathrm{e}^z$. We compute this directly as

$$\sinh(z)\mathrm{e}^z = \left(z + \frac{z^3}{3!} + \dots\right) \left(1 + z + \frac{z^2}{2} + \frac{z^3}{3!} + \dots\right) = \dots + \left(\frac{1}{4!} + \frac{1}{3!}\right) z^4 + \dots$$

So $\text{Res}(f, 0) = 1/3! + 1/4! = 5/24$.

Example 8.16. Find the residue of $f(z) = \frac{1}{z(z^2+1)(z-2)^2}$ at $z = 2$.

answer: $g(z) = (z-2)^2 f(z) = \frac{1}{z(z^2+1)}$ is analytic at $z = 2$. So, the residue we want is the a_1 term in its Taylor series, i.e. $g'(2)$. This is easy, if dull, to compute

$$\text{Res}(f, 2) = g'(2) = -\frac{13}{100}$$

8.4.3 $\cot(z)$

The function $\cot(z)$ turns out to be very useful in applications. This stems largely from the fact that it has simple poles at all multiples of π and the residue is 1 at each pole. We show that first.

Fact. $f(z) = \cot(z)$ has simple poles at $n\pi$ for n an integer and $\text{Res}(f, n\pi) = 1$.

Proof. $f(z) = \frac{\cos(z)}{\sin(z)}$. This has poles at the zeros of \sin , i.e. at $z = n\pi$. At poles f is of the form p/q where q has a simple zero at z_0 and $p(z_0) \neq 0$. Thus we can use the formula

$$\text{Res}(f, z_0) = \frac{p(z_0)}{q'(z_0)}.$$

In our case, we have $\text{Res}(f, n\pi) = \frac{\cos(n\pi)}{\cos(n\pi)} = 1$ as claimed.

Sometimes we need more terms in the Laurent expansion of $\cot(z)$. There is no known easy formula for the terms, but we can easily compute as many as we need using the following technique.

Example 8.17. Compute the first several terms of the Laurent expansion of $\cot(z)$ around $z = 0$.

answer: Since $\cot(z)$ has a simple pole at 0 we know

$$\cot(z) = \frac{b_1}{z} + a_0 + a_1 z + a_2 z^2 + \dots$$

We also know

$$\cot(z) = \frac{\cos(z)}{\sin(z)} = \frac{1 - z^2/2 + z^4/4! - \dots}{z - z^3/3! + z^5/5! - \dots}$$

Cross multiplying the two expressions we get

$$\left(\frac{b_1}{z} + a_0 + a_1 z + a_2 z^2 + \dots \right) \left(z - \frac{z^3}{3!} + \frac{z^5}{5!} - \dots \right) = 1 - \frac{z^2}{2} + \frac{z^4}{4!} - \dots$$

We can do the multiplication and equate the coefficients of like powers of z .

$$b_1 + a_0 z + \left(-\frac{b_1}{3!} + a_1 \right) z^2 + \left(-\frac{a_0}{3!} + a_2 \right) z^3 + \left(\frac{b_1}{5!} - \frac{a_1}{3!} + a_3 \right) z^4 = 1 - \frac{z^2}{2} + \frac{z^4}{4!}$$

So,

$$b_1 = 1$$

$$a_0 = 0$$

$$-b_1/3! + a_1 = -1/2! \Rightarrow a_1 = -1/3$$

$$-a_0/3! + a_2 = 0 \Rightarrow a_2 = 0$$

$$b_1/5! - a_1/3! + a_3 = 1/4! \Rightarrow a_3 = -1/45.$$

As noted above, all the even terms are 0 as they should be. We have

$$\cot(z) = \frac{1}{z} - \frac{z}{3} - \frac{z^3}{45} + \dots$$

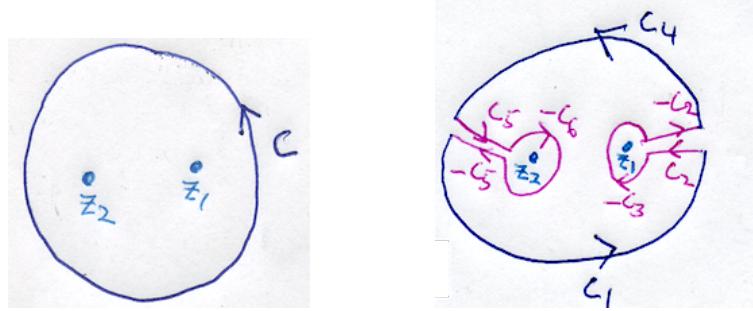
8.5 Cauchy Residue Theorem

This is one of the major theorems in 18.04. It will allow us to make systematic our previous somewhat ad hoc approach to computing integrals on contours that surround singularities.

Cauchy's residue theorem. Suppose $f(z)$ is analytic in the region A except for a set of isolated singularities. Also suppose C is a simple closed curve in A that doesn't go through any of the singularities of f and is oriented counterclockwise. Then

$$\int_C f(z) dz = 2\pi i \sum \text{residues of } f \text{ inside } C$$

Proof. The proof is based on the following figures. They only show a curve with two singularities inside it, but the generalization to any number of singularities is straightforward. In what follows we are going to abuse language and say pole when we mean isolated singularity, i.e. a finite order pole or an essential singularity ('infinite order pole').



The left figure shows the curve C surrounding two poles z_1 and z_2 of f . The right figure shows the same curve with some cuts and small circles added. It is chosen so that there are no poles of f inside it and so that the little circles around each of the poles are so small that there are no other poles inside them. The right hand curve is

$$\tilde{C} = C_1 + C_2 - C_3 - C_4 + C_5 - C_6 - C_5$$

The left hand curve is $C = C_1 + C_4$. Since there are no poles inside \tilde{C} we have, by Cauchy's theorem,

$$\int_{\tilde{C}} f(z) dz = \int_{C_1 + C_2 - C_3 - C_4 + C_5 - C_6 - C_5} f(z) dz = 0$$

Dropping C_2 and C_5 , which are both added and subtracted, this becomes

$$\int_{C_1 + C_4} f(z) dz = \int_{C_3 + C_6} f(z) dz \quad (26)$$

If $f(z) = \dots + \frac{b_2}{(z - z_1)^2} + \frac{b_1}{z - z_1} + a_0 + a_1(z - z_1) + \dots$ is the Laurent expansion of f around z_1 then

$$\int_{C_3} f(z) dz = \int_{C_3} \dots + \frac{b_2}{(z - z_1)^2} + \frac{b_1}{z - z_1} + a_0 + a_1(z - z_1) + \dots dz = 2\pi i b_1 = 2\pi i \operatorname{Res}(f, z_1)$$

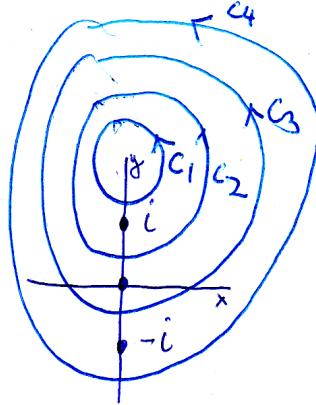
Likewise $\int_{C_6} f(z) dz = 2\pi i \operatorname{Res}(f, z_2)$.

Using these residues and the fact that $C = C_1 + C_4$, Equation 26 becomes

$$\int_C f(z) dz = 2\pi i (\operatorname{Res}(f, z_1) + \operatorname{Res}(f, z_2)).$$

That proves the residue theorem for the case of two poles. As we said, generalizing to any number of poles is straightforward.

Example 8.18. Let $f(z) = \frac{1}{z(z^2 + 1)}$. Compute $\int f(z) dz$ over each of the contours C_1, C_2, C_3, C_4 shown.



answer: The poles of $f(z)$ are at $z = 0, \pm i$. Using the residue theorem we just need to compute the residues of each of these poles.

At $z = 0$: $g(z) = zf(z) = \frac{1}{z^2 + 1}$ is analytic at 0 so the pole is simple and $\text{Res}(f, 0) = g(0) = 1$.

At $z = i$: $g(z) = (z - i)f(z) = \frac{1}{z(z + i)}$ is analytic at i so the pole is simple and $\text{Res}(f, i) = g(i) = -1/2$.

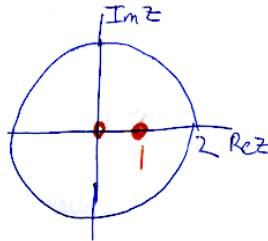
At $z = -i$: $g(z) = (z + i)f(z) = \frac{1}{z(z - i)}$ is analytic at $-i$ so the pole is simple and $\text{Res}(f, -i) = g(-i) = -1/2$.

Using the residue theorem we have

$$\begin{aligned} \int_{C_1} f(z) dz &= 0 \quad (\text{since } f \text{ is analytic inside } C_1) \\ \int_{C_2} f(z) dz &= 2\pi i \text{Res}(f, i) = -\pi i \\ \int_{C_3} f(z) dz &= 2\pi i(\text{Res}(f, i) + \text{Res}(f, 0)) = \pi i \\ \int_{C_4} f(z) dz &= 2\pi i(\text{Res}(f, i) + \text{Res}(f, 0) + \text{Res}(f, -i)) = 0. \end{aligned}$$

Example 8.19. Compute $\int_{|z|=2} \frac{5z - 2}{z(z - 1)} dz$.

answer: Let $f(z) = \frac{5z - 2}{z(z - 1)}$. The poles of f are at $z = 0, 1$ and the contour encloses them both.



At $z = 0$: $g(z) = zf(z) = \frac{5z - 2}{(z - 1)}$ is analytic at 0 so the pole is simple and $\text{Res}(f, 0) = g(0) = 2$.

At $z = 1$: $g(z) = (z - 1)f(z) = \frac{5z - 2}{z}$ is analytic at 1 so the pole is simple and $\text{Res}(f, 1) = g(1) = 3$.

Finally $\int_C \frac{5z - 2}{z(z - 1)} dz = 2\pi i(\text{Res}(f, 0) + \text{Res}(f, 1)) = 10\pi i$.

Example 8.20. Compute $\int_{|z|=1} z^2 \sin(1/z) dz$.

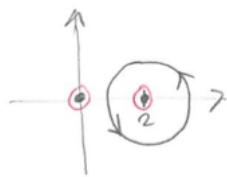
answer: Let $f(z) = z^2 \sin(1/z)$. f has an isolated singularity at $z = 0$. Using the Taylor series for $\sin(w)$ we get

$$z^2 \sin(1/z) = z^2 \left(\frac{1}{z} - \frac{1}{3!z^3} + \frac{1}{5!z^5} - \dots \right) = z - \frac{1/6}{z} + \dots$$

So, $\text{Res}(f, 0) = b_1 = -1/6$. Thus the residue theorem gives

$$\int_{|z|=1} z^2 \sin(1/z) dz = 2\pi i \text{Res}(f, 0) = -\frac{\pi}{3}.$$

Example 8.21. Compute $\int_C \frac{dz}{z(z - 2)^4} dz$, $C : |z - 2| = 1$.



answer: Let $f(z) = \frac{1}{z(z - 2)^4}$. The singularity at $z = 0$ is outside the contour of integration so it doesn't contribute to the integral.

To use the residue theorem we need to find the residue of f at $z = 2$. There are a number of ways to do this. Here's one:

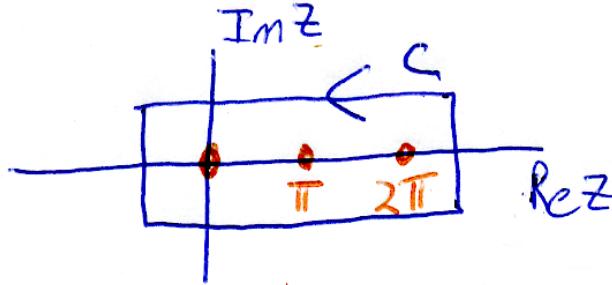
$$\begin{aligned} \frac{1}{z} &= \frac{1}{2 + (z - 2)} \\ &= \frac{1}{2} \cdot \frac{1}{1 + (z - 2)/2} \\ &= \frac{1}{2} \left(1 - \frac{z - 2}{2} + \frac{(z - 2)^2}{4} - \frac{(z - 2)^3}{8} + \dots \right) \end{aligned}$$

This is valid on $0 < |z - 2| < 2$. So,

$$f(z) = \frac{1}{(z-2)^4} \cdot \frac{1}{z} = \frac{1}{2(z-2)^4} - \frac{1}{4(z-2)^3} + \frac{1}{8(z-2)^2} - \frac{1}{16(z-2)} + \dots$$

Thus, $\text{Res}(f, 2) = -1/16$ and $\int_C f(z) dz = 2\pi i \text{Res}(f, 2) = -\frac{\pi i}{8}$.

Example 8.22. Compute $\int_C \frac{1}{\sin(z)} dz$ over the contour C shown.



answer: Let $f(z) = 1/\sin(z)$. There are 3 poles of f inside C at $0, \pi$ and 2π . We can find the residues by taking the limit of $(z - z_0)f(z)$. Each of the limits is computed using L'Hospital's rule. (This is valid, since the rule is just a statement about power series. We could also have used Property 5 from the section on residues of simple poles above.)

At $z = 0$: $\lim_{z \rightarrow 0} \frac{z}{\sin(z)} = \lim_{z \rightarrow 0} \frac{1}{\cos(z)} = 1$. Since the limit exists, $z = 0$ is a simple pole and $\text{Res}(f, 0) = 1$.

At $z = \pi$: $\lim_{z \rightarrow \pi} \frac{z - \pi}{\sin(z)} = \lim_{z \rightarrow \pi} \frac{1}{\cos(z)} = -1$. Since the limit exists, $z = \pi$ is a simple pole and $\text{Res}(f, \pi) = -1$.

At $z = 2\pi$: The same argument shows $\text{Res}(f, 2\pi) = 1$.

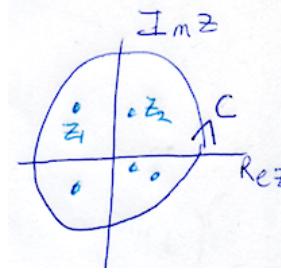
Now, by the residue theorem

$$\int_C f(z) dz = 2\pi i(\text{Res}(f, 0) + \text{Res}(f, \pi) + \text{Res}(f, 2\pi)) = 2\pi i.$$

8.6 Residue at ∞

The residue at ∞ is a clever device that can sometimes allow us to replace the computation of many residues with the computation of a single residue.

Suppose that f is analytic in \mathbf{C} except for a finite number of singularities. Let C be a positively oriented curve that is large enough to contain all the singularities.



All the poles of f are inside C

Definition. We define the residue of f at infinity by

$$\text{Res}(f, \infty) = -\frac{1}{2\pi i} \int_C f(z) dz.$$

We should first explain the idea here. The interior of a simple closed curve is everything to left as you traverse the curve. The curve C is oriented counterclockwise, so its interior contains all the poles of f . The residue theorem says the integral over C is determined by the residues of these poles.

On the other hand, the interior of the curve $-C$ is everything outside of C . There are no poles of f in that region. If we want the residue theorem to hold (which we do – it's that important) then the only option is to have a residue at ∞ and define it as we did.

The definition of the residue at infinity assumes all the poles of f are inside C . Therefore the residue theorem implies

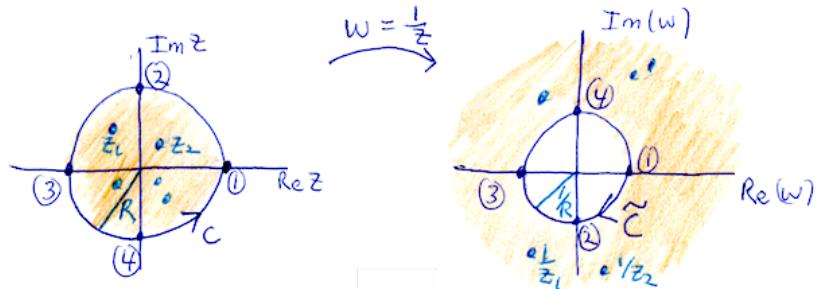
$$\text{Res}(f, \infty) = -\sum \text{the residues of } f.$$

To make this useful we need a way to compute the residue directly. This is given by the following theorem.

Theorem. If f is analytic in \mathbf{C} except for a finite number of singularities then

$$\text{Res}(f, \infty) = -\text{Res}\left(\frac{1}{w^2}f(1/w), 0\right).$$

Proof. The proof is just a change of variables: $w = 1/z$.



Change of variable: $w = 1/z$

First note that $z = 1/w$ and $dz = -(1/w^2)dw$. Next, note that the map $w = 1/z$ carries the positively oriented z -circle of radius R to the negatively oriented w -circle of radius $1/R$. (To see the orientation, follow the circled points 1, 2, 3, 4 on C in the z -plane as they are

mapped to points on \tilde{C} in the w -plane.) Thus,

$$\text{Res}(f, \infty) = -\frac{1}{2\pi i} \int_C f(z) dz = \frac{1}{2\pi i} \int_{\tilde{C}} f(1/w) \frac{1}{w^2} dw$$

Finally, note that $z = 1/w$ maps all the poles inside the circle C to points outside the circle \tilde{C} . So the only possible pole of $(1/w^2)f(1/w)$ that is inside \tilde{C} is at $w = 0$. Now, since \tilde{C} is oriented clockwise, the residue theorem says

$$\frac{1}{2\pi i} \int_{\tilde{C}} f(1/w) \frac{1}{w^2} dw = -\text{Res}\left(\frac{1}{w^2} f(1/w), 0\right)$$

Comparing this with the equation just above finishes the proof.

Example 8.23. Let $f(z) = \frac{5z-2}{z(z-1)}$. Earlier we computed $\int_{|z|=2} f(z) dz = 10\pi i$ by computing residues at $z = 0$ and $z = 1$. Recompute this integral by computing a single residue at infinity.

answer: $\frac{1}{w^2} f(1/w) = \frac{1}{w^2} \frac{5/w-2}{(1/w)(1/w-1)} = \frac{5-2w}{w(w-1)}$. We easily compute that

$$\text{Res}(f, \infty) = -\text{Res}\left(\frac{1}{w^2} f(1/w), 0\right) = -5.$$

Since $|z| = 2$ contains all the singularities of f we have

$$\int_{|z|=2} f(z) dz = -2\pi i \text{Res}(f, \infty) = 10\pi i.$$

This is the same answer we got before!

9 Definite integrals using the residue theorem

9.1 Introduction

In this topic we'll use the residue theorem to compute some real definite integrals.

$$\int_a^b f(x) dx$$

The general approach is always the same

1. Find a complex analytic function $g(z)$ which either equals f on the real axis or which is closely connected to f , e.g. $f(x) = \cos(x)$, $g(z) = e^{iz}$.
2. Pick a closed contour C that includes the part of the real axis in the integral.
3. The contour will be made up of pieces. It should be such that we can compute $\int g(z) dz$ over each of the pieces except the part on the real axis.
4. Use the residue theorem to compute $\int_C g(z) dz$.
5. Combine the previous steps to deduce the value of the integral we want.

9.2 Integrals of functions that decay

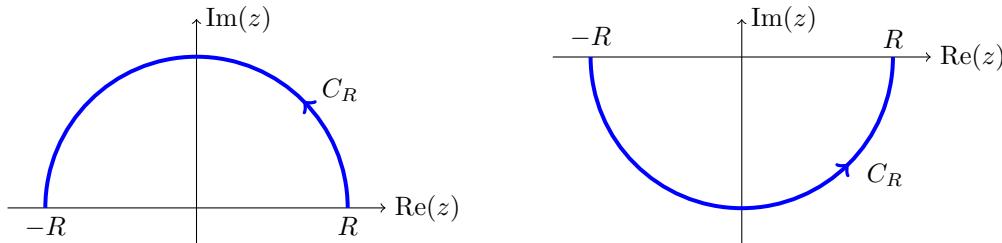
The theorems in this section will guide us in choosing the closed contour C described in the introduction.

The first theorem is for functions that decay faster than $1/z$.

Theorem 9.1. (a) Suppose $f(z)$ is defined in the upper half-plane. If there is an $a > 1$ and $M > 0$ such that $|f(z)| < \frac{M}{|z|^a}$ for $|z|$ large then

$$\lim_{R \rightarrow \infty} \int_{C_R} f(z) dz = 0,$$

where C_R is the semicircle shown below on the left.



Semicircles: left: $Re^{i\theta}, 0 < \theta < \pi$ right: $Re^{i\theta}, \pi < \theta < 2\pi$.

(b) If $f(z)$ is defined in the lower half-plane and $|f(z)| < \frac{M}{|z|^a}$, where $a > 1$ then

$$\lim_{R \rightarrow \infty} \int_{C_R} f(z) dz = 0,$$

where C_R is the semicircle shown above on the right.

Proof. We prove (a), (b) is essentially the same. We use the triangle inequality for integrals and the estimate given in the hypothesis. For R large

$$\left| \int_{C_R} f(z) dz \right| \leq \int_{C_R} |f(z)| |dz| \leq \int_{C_R} \frac{M}{|z|^a} |dz| = \int_0^\pi \frac{M}{R^a} R d\theta = \frac{M\pi}{R^{a-1}}.$$

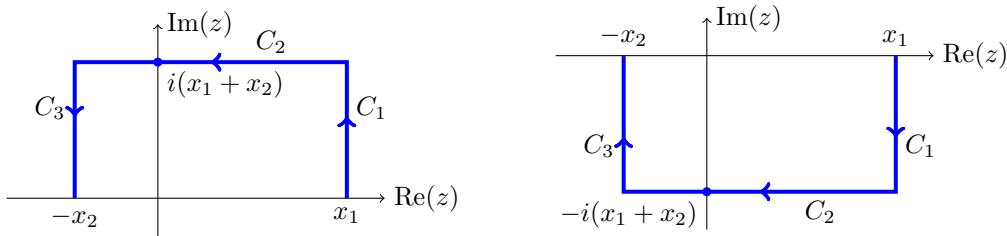
Since $a > 1$ this clearly goes to 0 as $R \rightarrow \infty$. QED

The next theorem is for functions that decay like $1/z$. It requires some more care to prove.

Theorem 9.2. (a) Suppose $f(z)$ is defined in the upper half-plane. If there is an $M > 0$ such that $|f(z)| < \frac{M}{|z|}$ for $|z|$ large then for $a > 0$

$$\lim_{x_1 \rightarrow \infty, x_2 \rightarrow \infty} \int_{C_1 + C_2 + C_3} f(z) e^{iaz} dz = 0,$$

where $C_1 + C_2 + C_3$ is the rectangular path shown below on the left.



Rectangular paths of height and width $x_1 + x_2$.

(b) Similarly, if $a < 0$ then

$$\lim_{x_1 \rightarrow \infty, x_2 \rightarrow \infty} \int_{C_1 + C_2 + C_3} f(z) e^{iaz} dz = 0,$$

where $C_1 + C_2 + C_3$ is the rectangular path shown above on the right.

Proof. (a) We start by parametrizing C_1, C_2, C_3 .

C_1 : $\gamma_1(t) = x_1 + it$, t from 0 to $x_1 + x_2$

C_2 : $\gamma_2(t) = t + i(x_1 + x_2)$, t from x_1 to $-x_2$

C_3 : $\gamma_3(t) = -x_2 + it$, t from $x_1 + x_2$ to 0.

Next we look at each integral in turn. We assume x_1 and x_2 are large enough that $|f(z)| < M/|z|$ on each of the curves C_j .

$$\begin{aligned} \left| \int_{C_1} f(z) e^{iaz} dz \right| &\leq \int_{C_1} |f(z) e^{iaz}| |dz| \leq \int_{C_1} \frac{M}{|z|} |e^{iaz}| |dz| \\ &= \int_0^{x_1+x_2} \frac{M}{\sqrt{x_1^2 + t^2}} |e^{iax_1 - at}| dt \leq \frac{M}{x_1} \int_0^{x_1+x_2} e^{-at} dt \\ &= \frac{M}{x_1} (1 - e^{-a(x_1+x_2)})/a. \end{aligned}$$

Since $a > 0$, it is clear that this last expression goes to 0 as x_1 and x_2 go to ∞ .

$$\begin{aligned} \left| \int_{C_2} f(z) e^{iaz} dz \right| &\leq \int_{C_2} |f(z) e^{iaz}| |dz| \leq \int_{C_2} \frac{M}{|z|} |e^{iaz}| |dz| \\ &= \int_{-x_2}^{x_1} \frac{M}{\sqrt{t^2 + (x_1 + x_2)^2}} |e^{iat - a(x_1 + x_2)}| dt \leq \frac{Me^{-a(x_1+x_2)}}{x_1 + x_2} \int_0^{x_1+x_2} dt \\ &\leq Me^{-a(x_1+x_2)} \end{aligned}$$

Again, clearly this last expression goes to 0 as x_1 and x_2 go to ∞ .

The argument for C_3 is essentially the same as for C_1 , so we leave it to the reader.

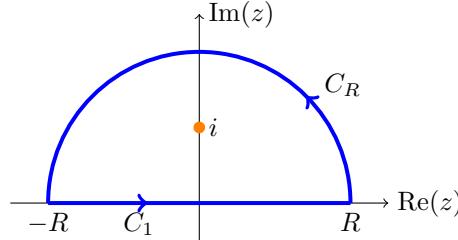
The proof for part (b) is the same. You need to keep track of the sign in the exponentials and make sure it is negative.

9.3 Integrals $\int_{-\infty}^{\infty}$ and \int_0^{∞}

Example 9.3. Compute $I = \int_{-\infty}^{\infty} \frac{1}{(1+x^2)^2}$.

answer: Let $f(z) = 1/(1+z^2)^2$. It is clear that for z large $f(z) \approx 1/z^4$. In particular, the hypothesis of Theorem 9.1 is satisfied. Using the contour shown below we have, by the residue theorem,

$$\int_{C_1+C_R} f(z) dz = 2\pi i \sum \text{residues of } f \text{ inside the contour.} \quad (27)$$



We examine each of the pieces in the above equation.

$\int_{C_R} f(z) dz$: By Theorem 9.1(a), $\int_{C_R} f(z) dz \rightarrow 0$ as $R \rightarrow \infty$.

$\int_{C_1} f(z) du$: Directly, we see that

$$\lim_{R \rightarrow \infty} \int_{C_1} f(z) dz = \lim_{R \rightarrow \infty} \int_{-R}^R f(x) dx = \int_{-\infty}^{\infty} = I.$$

So letting $R \rightarrow \infty$ Equation 27 becomes

$$I = \int_{-\infty}^{\infty} f(x) dx = 2\pi i \sum \text{residues of } f \text{ inside the contour.}$$

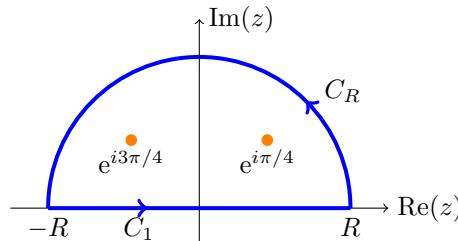
Finally, we compute the needed residues: $f(z)$ has poles of order 2 at $\pm i$. Only $z = i$ is inside the contour, so we compute the residue there. Let $g(z) = (z-i)^2 f(z) = \frac{1}{(z+i)^2}$. Then

$$\text{Res}(f, i) = g'(i) = -\frac{2}{(2i)^3} = \frac{1}{4i}$$

So, $I = 2\pi i \text{Res}(f, i) = \boxed{\frac{\pi}{2}}$.

Example 9.4. Compute $I = \int_{-\infty}^{\infty} \frac{1}{x^4 + 1} dx$.

answer: Let $f(z) = 1/(1+z^4)$. We use the same contour as in the previous example



As in the previous example, $\lim_{R \rightarrow \infty} \int_{C_R} f(z) dz = 0$ and $\lim_{R \rightarrow \infty} \int_{C_1} f(z) dz = \int_{-\infty}^{\infty} f(x) dx = I$. So, by the residue theorem

$$I = \lim_{R \rightarrow \infty} \int_{C_1 + C_R} f(z) dz = 2\pi i \sum \text{ residues of } f \text{ inside the contour.}$$

The poles of f are all simple and at $e^{i\pi/4}, e^{i3\pi/4}, e^{i5\pi/4}, e^{i7\pi/4}$. Only $e^{i\pi/4}$ and $e^{i3\pi/4}$ are inside the contour. We compute their residues as limits using L'Hospital's rule

$$\begin{aligned} z_1 = e^{i\pi/4} : \text{Res}(f, z_1) &= \lim_{z \rightarrow z_1} (z - z_1)f(z) = \lim_{z \rightarrow z_1} \frac{z - z_1}{1 + z^4} = \lim_{z \rightarrow z_1} \frac{1}{4z^3} = \frac{1}{4e^{i3\pi/4}} = \frac{e^{-i3\pi/4}}{4} \\ z_2 = e^{i3\pi/4} : \text{Res}(f, z_2) &= \lim_{z \rightarrow z_2} (z - z_2)f(z) = \lim_{z \rightarrow z_2} \frac{z - z_2}{1 + z^4} = \lim_{z \rightarrow z_2} \frac{1}{4z^3} = \frac{1}{4e^{i9\pi/4}} = \frac{e^{-i\pi/4}}{4} \end{aligned}$$

So,

$$I = 2\pi i(\text{Res}(f, z_1) + \text{Res}(f, z_2)) = 2\pi i \left(\frac{-1 - i}{4\sqrt{2}} + \frac{1 - i}{4\sqrt{2}} \right) = 2\pi i \left(-\frac{2i}{4\sqrt{2}} \right) = \boxed{\pi \frac{\sqrt{2}}{2}}$$

Example 9.5. Suppose $b > 0$. Show $\int_0^\infty \frac{\cos(x)}{x^2 + b^2} dx = \frac{\pi e^{-b}}{2b}$.

answer: The first thing to note is that the integrand is even, so

$$I = \frac{1}{2} \int_{-\infty}^{\infty} \frac{\cos(x)}{x^2 + b^2} dx.$$

Also note that the square in the denominator tells us the integral is absolutely convergent.

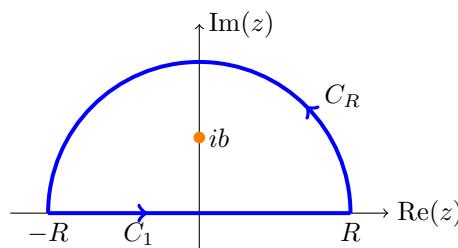
We have to be careful because $\cos(z)$ goes to infinity in either half-plane, so the hypotheses of Theorem 9.1 are not satisfied. The trick is to replace $\cos(x)$ by e^{ix} , so

$$\tilde{I} = \int_{-\infty}^{\infty} \frac{e^{ix}}{x^2 + b^2} dx, \text{ with } I = \frac{1}{2} \operatorname{Re}(\tilde{I}).$$

Now let $f(z) = \frac{e^{iz}}{z^2 + b^2}$. For $z = x + iy$ with $y > 0$ we have

$$|f(z)| = \frac{|e^{i(x+iy)}|}{|z^2 + b^2|} = \frac{e^{-y}}{|z^2 + b^2|}.$$

Since $e^{-y} < 1$, $f(z)$ satisfies the hypotheses of Theorem 9.1 in the upper half-plane. Now we can use the same contour as in the previous examples



We have $\lim_{R \rightarrow \infty} \int_{C_R} f(z) dz = 0$ and $\lim_{R \rightarrow \infty} \int_{C_1} f(z) dz = \int_{-\infty}^{\infty} f(x) dx = \tilde{I}$. So, by the residue theorem

$$\tilde{I} = \lim_{R \rightarrow \infty} \int_{C_1 + C_R} f(z) dz = 2\pi i \sum \text{residues of } f \text{ inside the contour.}$$

The poles of f are at $\pm bi$ and both are simple. Only bi is inside the contour. We compute the residue as a limit using L'Hospital's rule

$$\text{Res}(f, bi) = \lim_{z \rightarrow bi} (z - bi) \frac{e^{iz}}{z^2 + b^2} = \frac{e^{-b}}{2bi}.$$

So, $\tilde{I} = 2\pi i \text{Res}(f, bi) = \frac{\pi e^{-b}}{b}$. Finally, $I = \frac{1}{2} \operatorname{Re}(\tilde{I}) = \frac{\pi e^{-b}}{2b}$, as claimed.

Example. See Example 9.14 below for an example using Theorem 9.2.

9.4 Trigonometric integrals.

The trick here is to put together some elementary properties of $z = e^{i\theta}$ on the unit circle.

$$1. e^{-i\theta} = 1/z.$$

$$2. \cos(\theta) = \frac{e^{i\theta} + e^{-i\theta}}{2} = \frac{z + 1/z}{2}.$$

$$3. \sin(\theta) = \frac{e^{i\theta} - e^{-i\theta}}{2i} = \frac{z - 1/z}{2i}.$$

We start with an example. After that we'll state a more general theorem.

Example 9.6. Compute $\int_0^{2\pi} \frac{d\theta}{1 + a^2 - 2a \cos(\theta)}$. Assume that $|a| \neq 1$.

answer: Notice that $[0, 2\pi]$ is the interval used to parametrize the unit circle as $z = e^{i\theta}$. We need to make two substitutions:

$$\begin{aligned} \cos(\theta) &= \frac{z + 1/z}{2} \\ dz = ie^{i\theta} d\theta &\Leftrightarrow d\theta = \frac{dz}{iz} \end{aligned}$$

Making these substitutions we get

$$I = \int_0^{2\pi} \frac{d\theta}{1 + a^2 - 2a \cos(\theta)} = \int_{|z|=1} \frac{1}{1 + a^2 - 2a(z + 1/z)/2} \cdot \frac{dz}{iz} = \frac{1}{i((1 + a^2)z - a(z^2 + 1))} dz.$$

So, let $f(z) = \frac{1}{i((1 + a^2)z - a(z^2 + 1))}$. The residue theorem implies

$$I = 2\pi i \sum \text{residues of } f \text{ inside the unit circle.}$$

We can factor the denominator: $f(z) = \frac{-1}{ia(z - a)(z - 1/a)}$. The poles are at $a, 1/a$. One

is inside the unit circle and one is outside.

$$\text{If } |a| > 1 \text{ then } 1/a \text{ is inside the unit circle and } \text{Res}(f, 1/a) = \frac{1}{i(a^2 - 1)}$$

$$\text{If } |a| < 1 \text{ then } a \text{ is inside the unit circle and } \text{Res}(f, a) = \frac{1}{i(1 - a^2)}$$

We have

$$I = \begin{cases} \frac{2\pi}{a^2 - 1} & \text{if } |a| > 1 \\ \frac{2\pi}{1 - a^2} & \text{if } |a| < 1 \end{cases}$$

The example illustrates a general technique which we state now.

Theorem 9.7. Suppose $R(x, y)$ is a rational function with no poles on the circle $x^2 + y^2 = 1$ then for $f(z) = R\left(\frac{z+1/z}{2}, \frac{z-1/z}{2i}\right)$ we have

$$\int_0^{2\pi} R(\cos(\theta), \sin(\theta)) d\theta = 2\pi \sum \text{ residues of } f \text{ inside } |z| = 1.$$

Proof. We make the same substitutions as in Example 9.6. So,

$$\int_0^{2\pi} R(\cos(\theta), \sin(\theta)) d\theta = \int_{|z|=1} R\left(\frac{z+1/z}{2}, \frac{z-1/z}{2i}\right) \frac{dz}{iz}$$

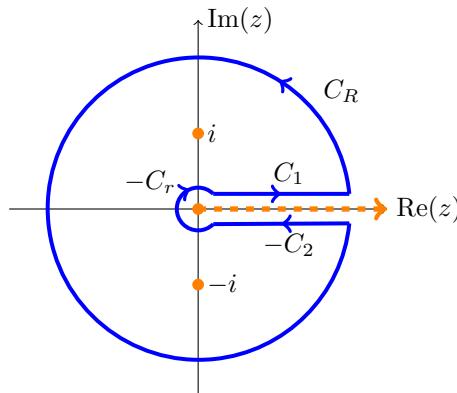
The assumption about poles means that f has no poles on the contour $|z| = 1$. The residue theorem now implies the theorem.

9.5 Integrands with branch cuts

Example 9.8. Compute $I = \int_0^\infty \frac{x^{1/3}}{1+x^2} dx$.

answer: Let $f(z) = z^{1/3}/(1+z^2)$. Since this is asymptotically comparable to $x^{-5/3}$, the integral is absolutely convergent. As a complex function $f(z) = z^{1/3}/(1+z^2)$ needs a branch cut to be analytic (or even continuous), so we will need to take that into account with our choice of contour.

First, we choose the following branch cut: for $z = re^{i\theta}$ take $0 \leq \theta < 2\pi$. Then we use the contour $C_1 + C_R - C_2 - C_r$ shown below.



Contour around branch cut: inner circle of radius r , outer of radius R .

We put convenient signs on the pieces so that the integrals are parametrized in a natural way. You should read this contour as having r so small that C_1 and C_2 are essentially on the x -axis. Note well, that since C_1 and C_2 are on opposite sides of the branch cut, the integral $\int_{C_1-C_2} f(z) dz \neq 0$.

First we analyze the integral over each piece of the curve.

On C_R : Theorem 9.1 says that $\lim_{R \rightarrow \infty} \int_{C_R} f(z) dz = 0$.

On C_r : $|z^{1/3}| = |r^{1/3}e^{i\theta/3}| = r^{1/3}$ and $\frac{1}{|1+z^2|} < \frac{1}{1-r^2}$. Therefore, for small r , $|f(z)| < M$ for some constant M . So,

$$\left| \int_{C_r} f(z) dz \right| \leq \int_0^{2\pi} |f(re^{i\theta})| |ire^{i\theta}| d\theta \leq \int_0^{2\pi} Mr d\theta = 2\pi Mr.$$

Clearly this goes to zero as $r \rightarrow 0$.

On C_1 : $\lim_{r \rightarrow 0, R \rightarrow \infty} \int_{C_1} f(z) dz = \int_0^\infty f(x) dx = I$.

On C_2 : We have (essentially) $\theta = 2\pi$, so $z^{1/3} = e^{i2\pi/3}|z|^{1/3}$. Thus,

$$\lim_{r \rightarrow 0, R \rightarrow \infty} \int_{C_2} f(z) dz = e^{i2\pi/3} \int_0^\infty f(x) dx = e^{i2\pi/3} I.$$

The poles of $f(z)$ are at $\pm i$. Since f is meromorphic inside our contour the residue theorem says

$$\int_{C_1+C_R-C_2-C_r} f(z) dz = 2\pi i (\text{Res}(f, i) + \text{Res}(f, -i)).$$

Letting $r \rightarrow 0$ and $R \rightarrow \infty$ the analysis above shows

$$(1 - e^{i2\pi/3})I = 2\pi i (\text{Res}(f, i) + \text{Res}(f, -i))$$

All that's left is to compute the residues using the chosen branch of $z^{1/3}$

$$\begin{aligned} \text{Res}(f, -i) &= \frac{(-i)^{1/3}}{-2i} = \frac{(e^{i3\pi/2})^{1/3}}{2e^{i3\pi/2}} = \frac{e^{-i\pi}}{2} = -\frac{1}{2} \\ \text{Res}(f, i) &= \frac{i^{1/3}}{2i} = \frac{e^{i\pi/6}}{2e^{i\pi/2}} = \frac{e^{-i\pi/3}}{2} \end{aligned}$$

A little more algebra gives

$$(1 - e^{i2\pi/3})I = 2\pi i \cdot \frac{-1 + e^{-i\pi/3}}{2} = \pi i (-1 + 1/2 - i\sqrt{3}/2) = -\pi i e^{i\pi/3}.$$

Continuing

$$I = \frac{-\pi i e^{i\pi/3}}{1 - e^{i2\pi/3}} = \frac{\pi i}{e^{i\pi/3} - e^{-i\pi/3}} = \frac{\pi/2}{(e^{i\pi/3} - e^{-i\pi/3})/2i} = \frac{\pi/2}{\sin(\pi/3)} = \frac{\pi}{\sqrt{3}}.$$

Whew!

9.6 Cauchy principal value

First an example to motivate defining the principal value of an integral. We'll actually compute the integral in the next section.

Example 9.9. Let $I = \int_0^\infty \frac{\sin(x)}{x} dx$. This integral is not absolutely convergent, but it is conditionally convergent. Formally, of course, we mean $I = \lim_{R \rightarrow \infty} \int_0^R \frac{\sin(x)}{x} dx$.

We can proceed as in Example 9.5. First note that $\sin(x)/x$ is even, so

$$I = \frac{1}{2} \int_{-\infty}^\infty \frac{\sin(x)}{x} dx.$$

Next, to avoid the problem that $\sin(z)$ goes to infinity in both the upper and lower half-planes we replace the integrand by $\frac{e^{ix}}{x}$.

We've changed the problem to computing

$$\tilde{I} = \int_{-\infty}^\infty \frac{e^{ix}}{x} dx.$$

The problems with this integral are caused by the pole at 0. The biggest problem is that the integral doesn't converge! The other problem is that when we try to use our usual strategy of choosing a closed contour problem we can't use one that includes $z = 0$ on the real axis. This is our motivation for defining principal value. We will come back to this example below.

Definition. Suppose we have a function $f(x)$ that is continuous on the real line except at the point x_1 , then we define the [Cauchy principal value](#) as

$$\text{p.v. } \int_{-\infty}^\infty f(x) dx = \lim_{R \rightarrow \infty, r_1 \rightarrow 0} \int_{-R}^{x_1 - r_1} f(x) dx + \int_{x_1 + r_1}^R f(x) dx.$$

Provided the limit converges. You should notice that the intervals around x_1 and around ∞ are symmetric. Of course, if the integral $\int_{-\infty}^\infty f(x) dx$ converges, then so does the principal value and they give the same value.

We can make the definition more flexible by including the following cases.

1. If $f(x)$ is continuous on the entire real line then we define the principal value as

$$\text{p.v. } \int_{-\infty}^\infty f(x) dx = \lim_{R \rightarrow \infty} \int_{-R}^R f(x) dx$$

2. If we have multiple points of discontinuity, $x_1 < x_2 < x_3 < \dots < x_n$, then

$$\text{p.v. } \int_{-\infty}^\infty f(x) dx = \lim \int_{-R}^{x_1 - r_1} f(x) dx + \int_{x_1 + r_1}^{x_2 - r_2} + \int_{x_2 + r_2}^{x_3 - r_3} + \dots \int_{x_n + r_n}^R f(x) dx.$$

Here the limit is taken as $R \rightarrow \infty$ and each of the $r_k \rightarrow 0$.



Intervals of integration for principal value are symmetric around x_k and ∞

The next example shows that sometimes the principal value converges when the integral itself does not. The opposite is never true. That is, we have the following theorem.

Theorem 9.10. If $f(x)$ has discontinuities at $x_1 < x_2 < \dots < x_n$ and $\int_{-\infty}^{\infty} f(x) dx$ converges then so does p.v. $\int_{-\infty}^{\infty} f(x) dx$.

Proof. The proof amounts to understanding the definition of convergence of integrals as limits. The integral converges means that each of the limits

$$\lim_{R_1 \rightarrow \infty, a_1 \rightarrow 0} \int_{-R_1}^{x_1 - a_1} f(x) dx, \quad \lim_{b_1 \rightarrow 0, a_2 \rightarrow 0} \int_{x_1 + b_1}^{x_2 - a_2} f(x) dx, \quad \dots \lim_{R_2 \rightarrow \infty, b_n \rightarrow 0} \int_{x_n + b_n}^{R_2} f(x) dx. \quad (28)$$

converges. There is no symmetry requirement, i.e. R_1 and R_2 are completely independent, as are a_1 and b_1 etc.

The principal value converges means

$$\lim \int_{-R}^{x_1 - r_1} + \int_{x_1 + r_1}^{x_2 - r_2} + \int_{x_2 + r_2}^{x_3 - r_3} + \dots \int_{x_n + r_n}^R f(x) dx \quad (29)$$

converges. Here the limit is taken over all the parameter $R \rightarrow \infty$, $r_k \rightarrow 0$. This limit has symmetry, e.g. we replaced both a_1 and b_1 in Equation 28 by r_1 etc. Certainly if the limits in Equation 28 converge then so do the limits in Equation 29. QED

Example 9.11. Consider both $\int_{-\infty}^{\infty} \frac{1}{x} dx$ and p.v. $\int_{-\infty}^{\infty} \frac{1}{x} dx$. The first integral diverges since

$$\int_{-R_1}^{-r_1} \frac{1}{x} dx + \int_{r_2}^{R_2} \frac{1}{x} dx = \ln(r_1) - \ln(R_1) + \ln(R_2) - \ln(r_2).$$

This clearly diverges as $R_1, R_2 \rightarrow \infty$ and $r_1, r_2 \rightarrow 0$.

On the other hand the symmetric integral

$$\int_{-R}^{-r} \frac{1}{x} dx + \int_r^R \frac{1}{x} dx = \ln(r) - \ln(R) + \ln(R) - \ln(r) = 0.$$

This clearly converges to 0.

We will see that the principal value occurs naturally when we integrate on semicircles around points. We prepare for this in the next section.

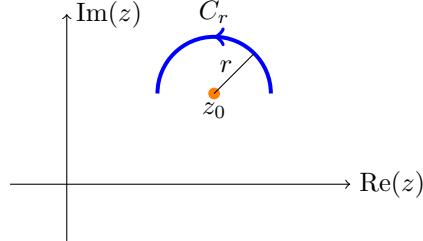
9.7 Integrals over portions of circles

We will need the following theorem in order to combine principal value and the residue theorem.

Theorem 9.12. Suppose $f(z)$ has a simple pole at z_0 . Let C_r be the semicircle $\gamma(\theta) =$

$z_0 + re^{i\theta}$, with $0 \leq \theta \leq \pi$. Then

$$\lim_{r \rightarrow 0} \int_{C_r} f(z) dz = \pi i \operatorname{Res}(f, z_0) \quad (30)$$



Small semicircle of radius r around z_0

Proof. Since we take the limit as r goes to 0, we can assume r is small enough that $f(z)$ has a Laurent expansion of the punctured disk of radius r centered at z_0 . That is, since the pole is simple,

$$f(z) = \frac{b_1}{z - z_0} + a_0 + a_1(z - z_0) + \dots \quad \text{for } 0 < |z - z_0| \leq r.$$

Thus,

$$\int_{C_r} f(z) dz = \int_0^\pi f(z_0 + re^{i\theta}) rie^{i\theta} d\theta = \int_0^\pi b_1 i + a_0 ire^{i\theta} + a_1 ir^2 e^{i2\theta} + \dots d\theta$$

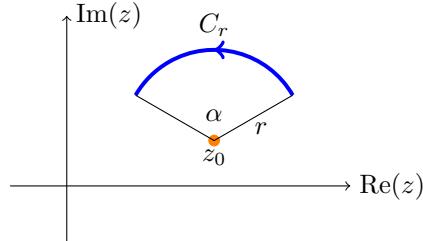
The b_1 term gives $\pi i b_1$. Clearly all the other terms go to 0 as $r \rightarrow 0$. QED.

If the pole is not simple the theorem doesn't hold and, in fact, the limit does not exist.

The same proof gives a slightly more general theorem.

Theorem. Suppose $f(z)$ has a **simple** pole at z_0 . Let C_r be the circular arc $\gamma(\theta) = z_0 + re^{i\theta}$, with $\theta_0 \leq \theta \leq \theta_0 + \alpha$. Then

$$\lim_{r \rightarrow 0} \int_{C_r} f(z) dz = \alpha i \operatorname{Res}(f, z_0)$$

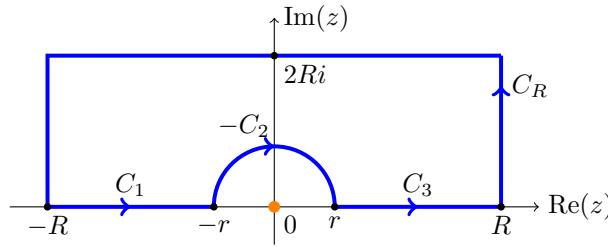


Small circular arc of radius r around z_0

Example 9.13. (Return to Example 9.9.) A long time ago we left off Example 9.9 to define principal value. Let's now use the principal value to compute $\tilde{I} = \text{p.v.} \int_{-\infty}^{\infty} \frac{e^{ix}}{x} dx$.

answer: We use the **indented contour** shown below. The indentation is the little semicircle that goes around $z = 0$. There are no poles inside the contour so the residue theorem implies

$$\int_{C_1 - C_2 + C_3 + C_R} \frac{e^{iz}}{z} dz = 0.$$



Next we break the contour into pieces.

$$\lim_{R \rightarrow \infty, r \rightarrow 0} \int_{C_1 + C_3} \frac{e^{iz}}{z} dz = \tilde{I}.$$

Theorem 9.2(a) implies $\lim_{R \rightarrow \infty} \int_{C_R} \frac{e^{iz}}{z} dz = 0$.

Equation 30 in Theorem 9.12 tells us that

$$\lim_{r \rightarrow 0} \int_{C_2} \frac{e^{iz}}{z} dz = \pi i \operatorname{Res}\left(\frac{e^{iz}}{z}, 0\right) = \pi i$$

Combining all this together we have

$$\lim \int_{C_1 - C_2 + C_3 + C_R} \frac{e^{iz}}{z} dz = \tilde{I} - \pi i = 0,$$

so $\tilde{I} = \pi i$. Thus, looking back at Example 30, where $I = \int_0^\infty \frac{\sin(x)}{x} dx$, we have

$$I = \frac{1}{2} \operatorname{Im}(\tilde{I}) = \frac{\pi}{2}.$$

There is a subtlety about convergence we alluded to above. That is, I is a genuine (conditionally) convergent integral, but \tilde{I} only exists as a principal value. However since I is a convergent integral we know that computing the principle value as we just did is sufficient to give the value of the convergent integral.

9.8 Fourier transform

Definition. The **Fourier transform** of a function $f(x)$ is defined by

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(x) e^{-ix\omega} dx$$

This is often read as ‘ f -hat’.

Theorem. (**Fourier inversion formula.**) We can recover the original function $f(x)$ with the Fourier inversion formula

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega) e^{ix\omega} d\omega.$$

So, the Fourier transform converts a function of x to a function of ω and the Fourier inversion converts it back. Of course, everything above is dependent on the convergence of the various integrals.

Proof. We will not give the proof here. We may get to it later in the course.

Example 9.14. Let $f(t) = \begin{cases} e^{-at} & \text{for } t > 0 \\ 0 & \text{for } t < 0 \end{cases}$, where $a > 0$. Compute $\hat{f}(\omega)$ and verify the Fourier inversion formula in this case.

answer: Computing \hat{f} is easy:

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt = \int_0^{\infty} e^{-at}e^{-i\omega t} dt = \frac{1}{a+i\omega} \quad (\text{recall } a > 0).$$

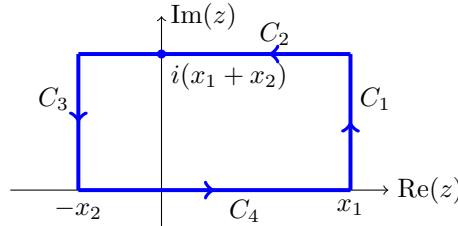
We should first note that the inversion integral converges. To avoid distraction we show this at the end of this example.

Now, let

$$g(z) = \frac{1}{a+iz}$$

Note that $\hat{f}(\omega) = g(\omega)$ and $|g(z)| < \frac{M}{|z|}$ for large $|z|$.

To verify the inversion formula we consider the cases $t > 0$ and $t < 0$ separately. For $t > 0$ we use the standard contour.



Theorem 9.2(a) implies that

$$\lim_{x_1 \rightarrow \infty, x_2 \rightarrow \infty} \frac{1}{2\pi} \int_{C_1 + C_2 + C_3} g(z)e^{izt} dz = 0$$

Clearly

$$\lim_{x_1 \rightarrow \infty, x_2 \rightarrow \infty} \frac{1}{2\pi} \int_{C_4} g(z)e^{izt} dz = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega) d\omega$$

The only pole of $g(z)e^{izt}$ is at $z = ia$, which is in the upper half-plane. So, applying the residue theorem to the entire closed contour, we get for large x_1, x_2 :

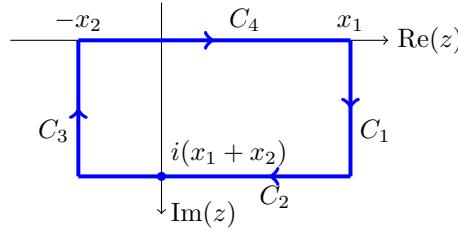
$$\frac{1}{2\pi} \int_{C_1 + C_2 + C_3 + C_4} g(z)e^{izt} dz = \text{Res} \left(\frac{e^{izt}}{a+iz}, ia \right) = e^{-at}.$$

Combining the above three equations we have

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega) d\omega = e^{-at} \quad \text{for } t > 0$$

This shows the inversion formula holds for $t > 0$.

For $t < 0$ we use the contour



Theorem 9.2(b) implies that

$$\lim_{x_1 \rightarrow \infty, x_2 \rightarrow \infty} \frac{1}{2\pi} \int_{C_1 + C_2 + C_3} g(z)e^{izt} dz = 0$$

Clearly

$$\lim_{x_1 \rightarrow \infty, x_2 \rightarrow \infty} \frac{1}{2\pi} \int_{C_4} g(z)e^{izt} dz = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega) d\omega$$

There are no poles of $g(z)e^{izt}$ in the lower half-plane. So, applying the residue theorem to the entire closed contour, we get for large x_1, x_2 :

$$\frac{1}{2\pi} \int_{C_1 + C_2 + C_3 + C_4} g(z)e^{izt} dz = \text{Res}\left(\frac{e^{izt}}{a + iz}, ia\right) = 0.$$

Combining the above three equations we have

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\omega) d\omega = 0 \quad \text{for } t < 0$$

This shows the inversion formula holds for $t < 0$.

Finally, we give the promised argument that the inversion integral converges. By definition

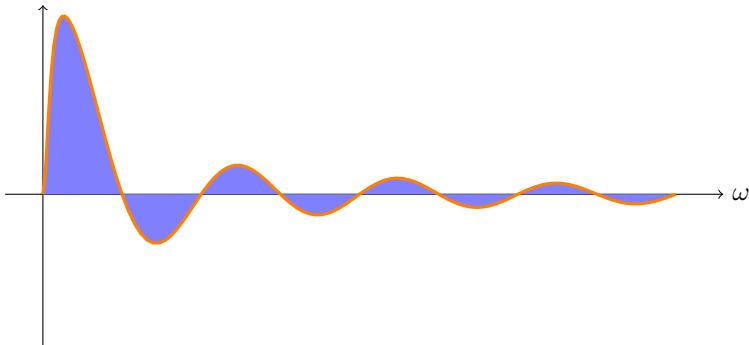
$$\begin{aligned} \int_{-\infty}^{\infty} \hat{f}(\omega)e^{i\omega t} d\omega &= \int_{-\infty}^{\infty} \frac{e^{i\omega t}}{a + i\omega} d\omega \\ &= \int_{-\infty}^{\infty} \frac{a \cos(\omega t) + \omega \sin(\omega t) - i\omega \cos(\omega t) + ia \sin(\omega t)}{a^2 + \omega^2} d\omega \end{aligned}$$

The terms without a factor of ω in the numerator converge absolutely because of the ω^2 in the denominator. The terms with a factor of ω in the numerator do not converge absolutely.

For example, since $\frac{\omega \sin(\omega t)}{a^2 + \omega^2}$ decays like $1/\omega$, its integral is not absolutely convergent. However, we claim that the integral does converge conditionally. That is, both limits

$$\lim_{R_2 \rightarrow \infty} \int_0^{R_2} \frac{\omega \sin(\omega t)}{a^2 + \omega^2} d\omega \text{ and } \lim_{R_1 \rightarrow \infty} \int_{-R_1}^0 \frac{\omega \sin(\omega t)}{a^2 + \omega^2} d\omega$$

exist and are finite. The key is that, as $\sin(\omega t)$ alternates between positive and negative arches, the function $\frac{\omega}{a^2 + \omega^2}$ is decaying monotonically. So, in the integral, the area under each arch adds or subtracts less than the arch before. This means that as R_1 (or R_2) grows the total area under oscillates with a decaying amplitude around some limiting value.



Total area oscillates with a decaying amplitude.

10 Conformal transformations

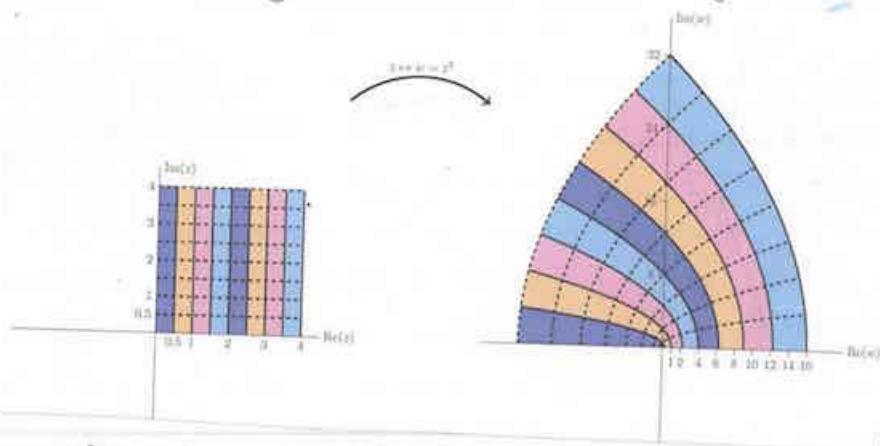
10.1 Introduction

In this topic we will look at the notion of conformal maps. This is a geometric notion and it will turn out that analytic functions are automatically conformal. Once we have the general notion, we will look at a specific family of conformal maps: linear fractional transformations. We will look at their geometric properties. As an application we will use them to solve the Dirichlet problem for harmonic functions on the unit disk with specified values on the unit circle. At the end we will return to some questions of fluid flow.

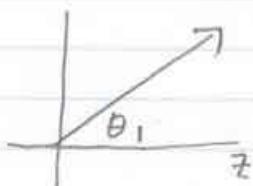
Topic 10 Conformal Maps

Definition (Informal) Conformal maps are functions that preserve the angles between curves

Example If f is analytic we have seen that f maps grid lines to orthogonal curves



Example. Consider multiplication by $c = a e^{i\phi}$



Ray at angle θ_1

$$w = cz$$



Ray at angle $\phi + \theta_1$

Multiplication by c rotates curves through an angle ϕ .

2

Topic 10

Definition (Precise) A function f defined

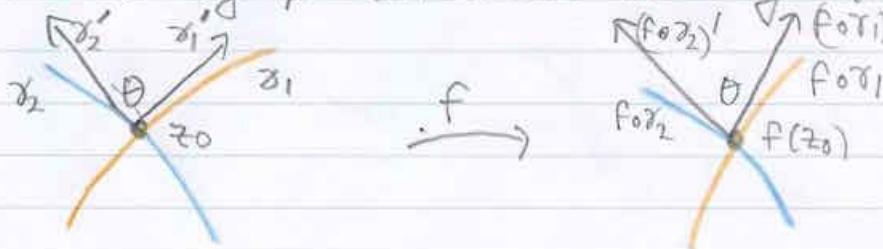
on a region A is conformal if at every

z_0 in A there is an $a > 0$ and an angle ϕ

(depending on z_0) such that for any curve

γ through z_0 , f rotates the tangent vector

at z_0 by ϕ and scales it by a ,



Same angle between tangent vectors
(Also they are both scaled by the same amount)

- Infinitesimally f looks like multiplication by $c = ae^{i\phi}$.

Theorem (Operational definition of conformal)

If f is analytic on A and $f'(z_0) \neq 0$ then

f is conformal and at z_0 we have $c = ae^{i\phi} = f'(z_0)$

3

Topic 10

PROOF Suppose we have the curve γ

and $\gamma(t_0) = z_0$. Then by the chain rule

$$(f \circ \gamma)'(t_0) = f'(\gamma(t_0)) \gamma'(t_0) = f'(z_0) \cdot \gamma'(t_0)$$

So f rotates $\gamma'(t_0)$ by $\text{Arg}(f'(z_0))$ and scales it by $|f'(z_0)|$.

- Linear approximation: $f(z) \approx f(z_0) + f'(z_0)z$

so $f(\gamma(t)) \approx \boxed{f(z_0)} + \boxed{f'(z_0)} \gamma(t)$

Shift Scale and rotate.

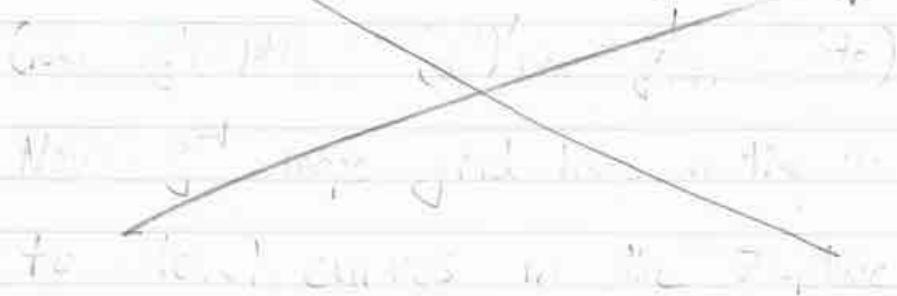
For us Conformal \iff Analytic $f'(z) \neq 0$.

Theorem If u, v are harmonic conjugates

and $g = u + iv$ has $g(z_0) \neq 0$ then the level

curves of u and v through z_0 are orthogonal

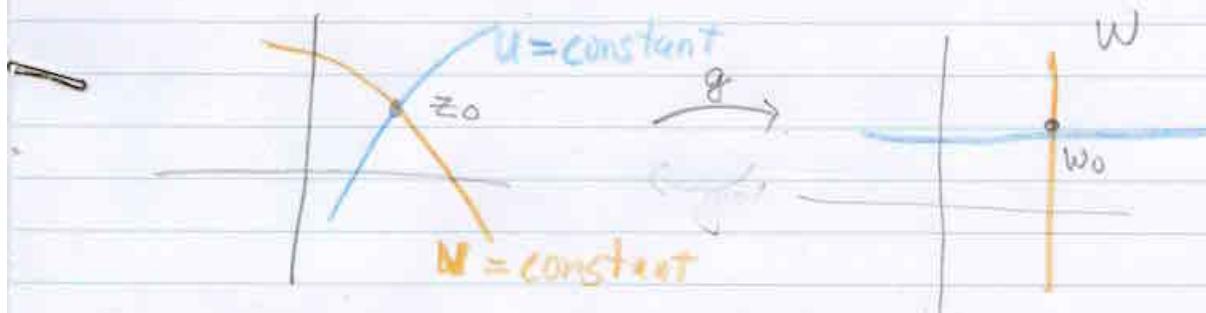
(We knew this already using Cauchy-Riemann.)



4.

Since g maps level curves to grid lines

$$(U(x,y)=c \rightarrow \operatorname{Re}(g(x+iy))=c)$$



(g takes level curves to grid lines)

and grid lines are orthogonal. The conformality
of g implies the level curves are orthogonal

Riemann Mapping Theorem

If a region A is simply connected and not
the whole plane then there is a bijective
conformal map from A to the unit disk.

Corollary Any two such regions are conformally
equivalent.

Proof Will skip. In practice we will write down
explicit conformal maps between regions.

10.2 Fractional linear transformations

Definition. A [fractional linear transformation](#) is a function of the form

$$T(z) = \frac{az + b}{cz + d}, \text{ where } a, b, c, d \text{ are constants and } ad - bc \neq 0$$

These are also called [Möbius transforms](#) or [bilinear transforms](#). We will abbreviate fractional linear transformation as [FLT](#).

Note. If $ad - bc = 0$ then $T(z)$ is a constant function.

Proof. The full proof requires that we deal with all the cases where some of the coefficients are 0. We'll give the proof assuming $c \neq 0$ and leave the case $c = 0$ to you. Assuming $c \neq 0$, the condition $ad - bc = 0$ implies $\frac{d}{c}(c, d) = (a, b)$. So,

$$T(z) = \frac{(a/c)(cz + d)}{cz + d} = \frac{a}{c}.$$

That is, $T(z)$ is constant.

Note 2. It will be convenient to consider linear transformations to be defined on the [extended complex plane](#) $\mathbf{C} \cup \{\infty\}$. We define

$$T(\infty) = \begin{cases} a/c & \text{if } c \neq 0 \\ \infty & \text{if } c = 0. \end{cases} \quad \text{and} \quad T(-d/c) = \infty \text{ if } c \neq 0.$$

10.2.1 Examples

Example 10.1. Scale and rotate. Let $T(z) = az$. If $a = r$ is real this scales the plane. If $a = e^{i\theta}$ it rotates the plane. If $a = re^{i\theta}$ it does both at once.

You supply a figure showing this rotation.

Example 10.2. Scale and rotate and translate. Let $T(z) = az + b$. Adding the b term introduces a translation to the previous example.

You supply a figure showing this LFT.

Example 10.3. Inversion. Let $T(z) = 1/z$. This is called an inversion. It turns the unit circle inside out. Note that $T(0) = \infty$ and $T(\infty) = 0$

You supply a figure showing inversion.

Example 10.4. Let $T(z) = \frac{z-i}{z+i}$. We claim that this maps the x -axis to the unit circle and the upper half-plane to the unit disk.

You supply a figure showing $T(z)$.

Proof. First take x real, then $|T(x)| = \frac{|x-i|}{|x+i|} = \frac{\sqrt{x^2+1}}{\sqrt{x^2+1}} = 1$. So, T maps the x -axis to the unit circle.

Next take $z = x + iy$ with $y > 0$, i.e. z in the upper half-plane. Clearly $|y+1| > |y-1|$, so $|z+i| = |x+i(y+1)| > |x+i(y-1)| = |z-i|$, that is $|T(z)| = \frac{|z-i|}{|z+i|} < 1$. So, T maps the upper half-plane to the unit disk.

We will use this map frequently, so for the record we note that

$$T(i) = 0, \quad T(\infty) = 1, \quad T(-1) = i, \quad T(0) = -1, \quad T(1) = -i.$$

These computations show that the real axis is mapped counterclockwise around the unit circle starting at 1 and coming back to 1.

10.2.2 Lines and circles

Theorem. A linear fractional transformation maps lines and circles to lines and circles.

Before proving this, note that it does not say lines are mapped to lines and circles to circles. For example, in Example 10.4 the real axis is mapped the unit circle. You can also check that inversion $w = 1/z$ maps the line $z = 1 + iy$ to the circle $|z - 1/2| = 1/2$.

Proof. We start by showing that inversion maps lines and circles to lines and circles. Given z and $w = 1/z$ we define x, y, u and v by

$$z = x + iy \quad \text{and} \quad w = \frac{1}{z} = \frac{x - iy}{x^2 + y^2} = u + iv$$

So, $u = \frac{x}{x^2 + y^2}$ and $v = -\frac{y}{x^2 + y^2}$. Now, every circle or line can be described by the equation

$$Ax + By + C(x^2 + y^2) = D$$

(If $C = 0$ it describes a line, otherwise a circle.) We convert this to an equation in u, v as follows.

$$Ax + By + C(x^2 + y^2) = D \Leftrightarrow \frac{Ax}{x^2 + y^2} + \frac{By}{x^2 + y^2} + C = \frac{D}{x^2 + y^2} \Leftrightarrow Au - Bv + C = D(u^2 + v^2).$$

In the last step we used the fact the $u^2 + v^2 = |w|^2 = 1/|z|^2 = 1/(x^2 + y^2)$. We have shown that a line or circle in x, y is transformed to a line or circle in u, v . This shows that inversion maps lines and circles to lines and circles.

We note that for the inversion $w = 1/z$.

1. Any line not through the origin is mapped to a circle through the origin.
2. Any line through the origin is mapped to a line through the origin.
3. Any circle not through the origin is mapped to a circle not through the origin.
4. Any circle through the origin is mapped to a line not through the origin.

Now, to prove that an arbitrary linear fractional transformation maps lines and circles to lines and circles, we factor it into a sequence of simpler transformations.

First suppose that $c = 0$. So, $T(z) = (az + b)/d$. Since this is just translation, scaling and rotating, it is clear it maps circles to circles and lines to lines.

Now suppose that $c \neq 0$. Then,

$$T(z) = \frac{az + b}{cz + d} = \frac{\frac{a}{c}(cz + d) + b - \frac{ad}{c}}{cz + d} = \frac{a}{c} + \frac{b - ad/c}{cz + d}$$

So, $w = T(z)$ can be computed as a composition of transforms

$$z \rightarrow w_1 = cz + d \rightarrow w_2 = 1/w_1 \rightarrow w = \frac{a}{c} + (b - ad/c)w_2$$

We know that each of the transforms in this sequence maps lines and circles to lines and circles. Therefore the entire sequence does also. QED

10.2.3 Mapping z_j to w_j

It turns out that for two sets of three points z_1, z_2, z_3 and w_1, w_2, w_3 there is a linear fractional transformation that takes z_j to w_j . We can construct this map as follows.

Let $T_1(z) = \frac{(z - z_1)(z_2 - z_3)}{(z - z_3)(z_2 - z_1)}$. Notice that $T_1(z_1) = 0$, $T_1(z_2) = 1$, $T_1(z_3) = \infty$.

Likewise let $T_2(w) = \frac{(w - w_1)(w_2 - w_3)}{(w - w_3)(w_2 - w_1)}$. Notice that $T_2(w_1) = 0$, $T_2(w_2) = 1$, $T_2(w_3) = \infty$.

Now $T(z) = T_2^{-1} \circ T_1(z)$ is the required map.

10.2.4 Correspondence with matrices

We can identify the transformation $T(z) = \frac{az + b}{cz + d}$ with the matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$. This identification is useful because of the following algebraic facts.

1. If $r \neq 0$ then $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $r \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ correspond to the same FLT.

Proof. This follows from the obvious equality

$$\frac{az + b}{cz + d} = \frac{raz + rb}{rcz + rd}.$$

2. If $T(z)$ corresponds to $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $S(z)$ corresponds to $B = \begin{bmatrix} e & f \\ g & h \end{bmatrix}$ then composition $T \circ S(z)$ corresponds to matrix multiplication AB .

Proof. The proof is just a bit of algebra.

$$\begin{aligned} T \circ S(z) &= T\left(\frac{ez + f}{gz + h}\right) = \frac{a((ez + f)/(gz + h)) + b}{c((ez + f)/(gz + h)) + d} = \frac{(ae + bg)z + af + bh}{(ce + dg)z + cf + dh} \\ AB &= \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} = \begin{bmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{bmatrix} \end{aligned}$$

The claimed correspondence is clear from the last entries in the two lines above.

3. If $T(z)$ corresponds to $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ then T has an inverse and $T^{-1}(w)$ corresponds to A^{-1} and also to $\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$, i.e. to A^{-1} without the factor of $1/\det(A)$.

Proof. Since $AA^{-1} = I$ it is clear from the previous fact that T^{-1} corresponds to A^{-1} . Since

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Fact 1 implies A^{-1} and $\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$ both correspond to the same FLT, i.e. to T^{-1} .

Example 10.5.

1. The matrix $\begin{bmatrix} a & b \\ 0 & 1 \end{bmatrix}$ corresponds to $T(z) = az + b$.
2. The matrix $\begin{bmatrix} e^{i\alpha} & 0 \\ 0 & e^{i\alpha} \end{bmatrix}$ corresponds to rotation by 2α .
3. The matrix $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ corresponds to the inversion $w = 1/z$.

10.3 Reflection and symmetry

10.3.1 Reflection and symmetry in a line

Example 10.6. Suppose we have a line S and a point z_1 not on S . The reflection of z_1 in S is the point z_2 so that S is the perpendicular bisector to the line segment $\overline{z_1 z_2}$. Since there is exactly one such point z_2 , the reflection of a point in a line is unique.

You supply the figure.

Definition. We also say that z_1 and z_2 are [symmetric with respect to the line \$S\$](#) .

In order to define the reflection of a point in a circle we need to work a little harder. Looking back at the previous example we can show the following.

Fact. Any circle through z_1 and z_2 intersects S orthogonally.

Proof. Call the circle C . Since S is the perpendicular bisector of a chord of C , the center of C lies on S . Therefore S is a radial line, i.e. it intersects C orthogonally.

You supply the figure.

10.3.2 Reflection and symmetry in a circle

We will adapt this for our definition of reflection in a circle. So that the logic flows correctly we need to start with the definition of symmetric pairs of points.

Definition. Suppose S is a line or circle. A pair of points z_1, z_2 is called [symmetric with respect to \$S\$](#) if every line or circle through the two points intersects S orthogonally.

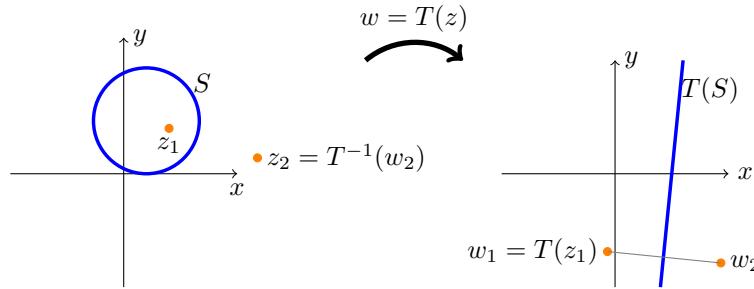
First we state an almost trivial fact.

Fact. Linear fractional transformations preserve symmetry. That is, if z_1 and z_2 are symmetric in a line or circle S , then, for an LFT T , $T(z_1)$ and $T(z_2)$ are symmetric in $T(S)$.

Proof. The definition of symmetry is in terms of lines and circles, and angles. Linear fractional transformations map lines and circles to lines and circles and, being conformal, preserve angles. QED

Theorem. Suppose S is a line or circle and z_1 a point not on S . There is a unique point z_2 such the pair z_1, z_2 is symmetric in S .

Proof. Let T be a linear fractional transformation that maps S to a line. We know that $w_1 = T(z_1)$ has a unique reflection w_2 in this line. Since T^{-1} preserves symmetry, z_1 and $z_2 = T^{-1}(w_2)$ are symmetric in S . Since w_2 is the unique point symmetric to w_1 the same is true for z_2 vis-a-vis z_1 . This is all shown in the figure below.



We can now define reflection in a circle.

Definition. The point z_2 in the theorem is called [the reflection of \$z_1\$ in \$S\$](#) .

10.3.3 Reflection in the unit circle

Using the symmetry preserving feature of linear fractional transformations, we start with a line and transform to the circle. Let R be the real axis and C the unit circle. We know the LFT

$$T(z) = \frac{z - i}{z + i}$$

maps R to C . We also know that the points z and \bar{z} are symmetric in R . Therefore

$$w_1 = T(z) = \frac{w - i}{w + i} \quad \text{and} \quad w_2 = T(\bar{z}) = \frac{\bar{z} - i}{\bar{z} + i}$$

are symmetric in D . Looking at the formulas, it is clear that $w_2 = 1/\bar{w}_1$. This is important enough that we highlight it as a theorem.

Theorem. Reflection in the unit circle. The reflection of z in the unit circle is $1/\bar{z}$. That is, the reflection of $r e^{i\theta}$ is $\frac{e^{i\theta}}{r}$.

Note. It is possible, but more tedious and less insightful, to arrive at this theorem by direct calculation.

You supply a figure displaying all this.

Example 10.7. Reflection in the circle of radius R . Suppose S is the circle $|z| = R$ and z_1 is a point not on S . Find the reflection of z_1 in S .

answer: Our strategy is to map S to the unit circle, find the reflection and then map the unit circle back to S .

You supply a flow diagram illustrating this.

Start with the map $T(z) = w = z/R$. Clearly T maps S to the unit circle and $w_1 = T(z_1) = z_1/R$. The reflection of w_1 is $w_2 = 1/\bar{w}_1 = R/\bar{z}_1$. Mapping back to the unit circle by T^{-1} we have $z_2 = T^{-1}(w_2) = R w_2 = R^2/\bar{z}_1$. Therefore the reflection of z_1 is R^2/\bar{z}_1 .

Example 10.8. Find the reflection of z_1 in the circle of radius R centered at c .

answer: Let $T(z) = z - c$. T maps the circle centered at c to one centered at 0. So, the reflection of z_1 is $z_2 = \frac{1}{\bar{z}_1 - c} + c$.

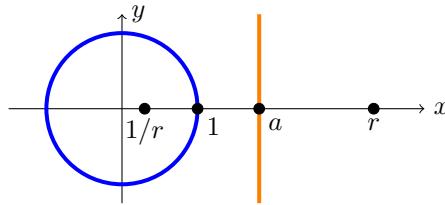
We can now record the following important fact.

Fact. Reflection of the center. For a circle S with center c , the pair c, ∞ is symmetric with respect to the circle.

Proof. This is an immediate consequence of the formula for the reflection of a point in a circle. For example, the reflection of z in the unit circle is $1/\bar{z}$. So, the reflection of 0 is infinity.

Example 10.9. Show that if a circle and a line don't intersect then there is a pair of points z_1, z_2 that is symmetric with respect to both the line and circle.

answer: By shifting, scaling and rotating we can find a fractional linear transformation T that maps the circle and line to the following configuration: The circle is mapped to the unit circle and the line to the vertical line $x = a > 1$.



For any real r , $w_1 = r$ and $w_2 = 1/r$ are symmetric in the unit circle. We can choose a specific r so that r and $1/r$ are equidistant from a , i.e. also symmetric in the line $x = a$. It is clear geometrically that this can be done. Algebraically we solve the equation

$$\frac{r + 1/r}{2} = a \Rightarrow r^2 - 2ar + 1 = 0 \Rightarrow r = a + \sqrt{a^2 - 1}, \frac{1}{r} = a - \sqrt{a^2 - 1}.$$

Thus $z_1 = T^{-1}(a + \sqrt{a^2 - 1})$ and $z_2 = T^{-1}(a - \sqrt{a^2 - 1})$ are the required points.

Example 10.10. Show that if two circles don't intersect then there is a pair of points z_1, z_2 that is symmetric with respect to both circles.

answer: Using a linear fractional transformation that maps one of the circles to a line (and the other to a circle) we can reduce the problem to that in the previous example.

Example 10.11. Show that any two circles that don't intersect can be mapped conformally to concentric circles.

answer: Call the circles S_1 and S_2 . Using the previous example start with a pair of points z_1, z_2 which are symmetric in both circles. Next, pick a linear fractional transformation T that maps z_1 to 0 and z_2 to infinity. For example, $T(z) = \frac{z - z_1}{z - z_2}$. Since T preserves symmetry 0 and ∞ are symmetric in the circle $T(S_1)$. This implies that 0 is the center of $T(S_1)$. Likewise 0 is the center of $T(S_2)$. Thus, $T(S_1)$ and $T(S_2)$ are concentric.

10.4 Solving the Dirichlet problem for harmonic functions

In general, a Dirichlet problem in a region A asks you to solve a partial differential equation in A where the values of the solution on the boundary of A are specified.

Example 10.12. Find a function u harmonic on the unit disk such that

$$u(e^{i\theta}) = \begin{cases} 1 & \text{for } 0 < \theta < \pi \\ 0 & \text{for } -\pi < \theta < 0 \end{cases}$$

This is a Dirichlet problem because the values of u on the boundary are specified. The partial differential equation is implied by requiring that u be harmonic, i.e. we require $\nabla^2 u = 0$. We will solve this problem in due course.

10.4.1 Harmonic functions on the upper half-plane

Our strategy will be to solve the Dirichlet problem for harmonic functions on the upper half-plane and then transfer these solutions to other domains.

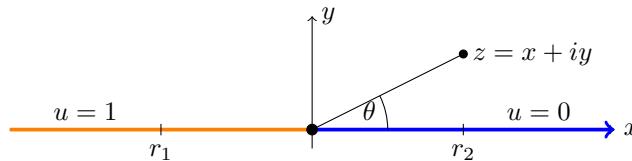
Example 10.13. Find a harmonic function $u(x, y)$ on the upper half-plane that satisfies the boundary condition

$$u(x, 0) = \begin{cases} 1 & \text{for } x < 0 \\ 0 & \text{for } x > 0 \end{cases}$$

answer: We can write down a solution explicitly as

$$u(x, y) = \frac{1}{\pi} \theta,$$

where θ is the argument of $z = x + iy$. Since we are only working on the upper half-plane we can take any convenient branch with branch cut in the lower half-plane, say $-\pi/2 < \theta < 3\pi/2$.



To show u is truly a solution, we have to verify two things:

1. u satisfies the boundary conditions
2. u is harmonic.

Both of these are straightforward. First, look at the point r_2 on the positive x -axis. This has argument $\theta = 0$, so $u(r_2, 0) = 0$. Likewise $\arg(r_1) = \pi$, so $u(r_1, 0) = 1$. Thus, we have shown point (1).

To see point (2) remember that $\log(z) = \log(r) + i\theta$. So, $u = \operatorname{Re}\left(\frac{1}{\pi i} \log(z)\right)$. Since it is the real part of an analytic function, u is harmonic.

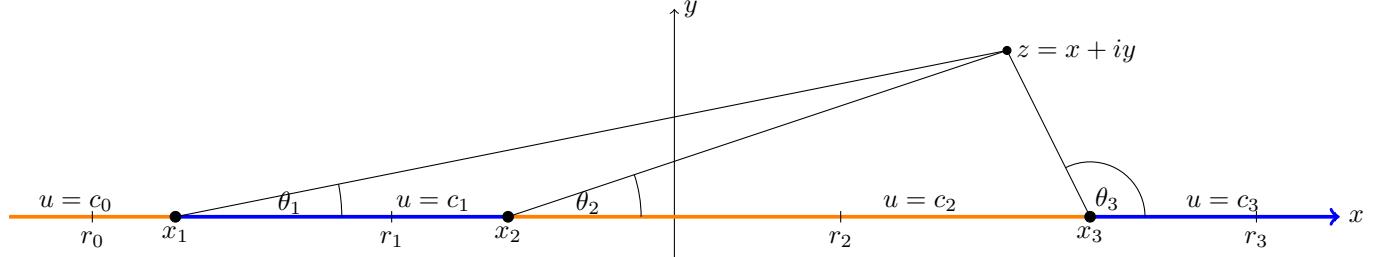
Example 10.14. Suppose $x_1 < x_2 < x_3$. Find a harmonic function u on the upper half-plane that satisfies the boundary condition

$$u(x, 0) = \begin{cases} c_0 & \text{for } x < x_1 \\ c_1 & \text{for } x_1 < x < x_2 \\ c_2 & \text{for } x_2 < x < x_3 \\ c_3 & \text{for } x_3 < x \end{cases}$$

answer: We mimic the previous example and write down the solution

$$u(x, y) = c_3 + (c_2 - c_3) \frac{\theta_3}{\pi} + (c_1 - c_2) \frac{\theta_2}{\pi} + (c_1 - c_0) \frac{\theta_1}{\pi}.$$

Here, the θ_j are the angles shown in the figure. We again choose the branch $-\pi/2 < \theta_j < 3\pi/2$.



To convince yourself that u satisfies the boundary condition test a few points:

At r_3 : all the $\theta_j = 0$. So, $u(r_3, 0) = c_3$ as required.

At r_2 : $\theta_1 = \theta_2 = 0$, $\theta_3 = \pi$. So, $u(r_2, 0) = c_3 + c_2 - c_3 = c_2$ as required.

Likewise, at r_1 and r_0 , u have the correct values.

As before, u is harmonic because it is the real part of the analytic function

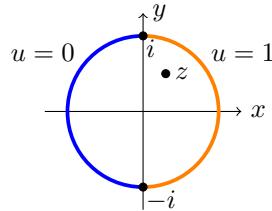
$$\Phi(z) = c_3 + \frac{(c_2 - c_3)}{\pi i} \log(z - x_3) + \frac{(c_1 - c_2)}{\pi i} \log(z - x_2) + \frac{(c_1 - c_0)}{\pi i} \log(z - x_1).$$

10.4.2 Harmonic functions on the unit disk

Let's try to solve a problem similar to the one in Example 10.12.

Example 10.15. Find a function u harmonic on the unit disk such that

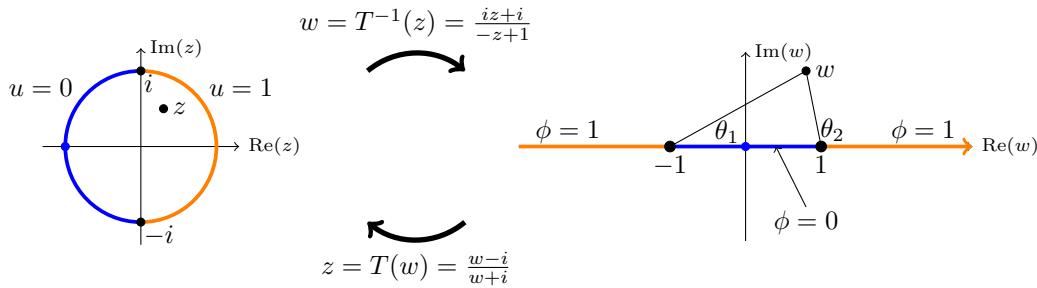
$$u(e^{i\theta}) = \begin{cases} 1 & \text{for } -\pi/2 < \theta < \pi/2 \\ 0 & \text{for } \pi/2 < \theta < 3\pi/2 \end{cases}$$



answer: Our strategy is to start with a conformal map T from the upper half-plane to the unit disk. We can use this map to pull the problem back to the upper half-plane. We solve it there and then push the solution back to the disk.

Let's call the disk D , the upper half-plane H . Let z be the variable on D and w the variable on H . Back in Example 10.4 we found a map from H to D . The map and its inverse are

$$z = T(w) = \frac{w - i}{w + i}, \quad w = T^{-1}(z) = \frac{iz + i}{-z + 1}.$$



The function u on D is transformed by T to a function ϕ on H . The relationships are

$$u(z) = \phi \circ T^{-1}(z) \quad \text{or} \quad \phi(w) = u \circ T(w)$$

These relationships determine the boundary values of ϕ from those we were given for u . We compute:

$$T^{-1}(i) = -1, \quad T^{-1}(-i) = 1, \quad T^{-1}(1) = \infty, \quad T^{-1}(-1) = 0.$$

This shows the left hand semicircle bounding D is mapped to the segment $[-1, 1]$ on the real axis. Likewise, the right hand semicircle maps to the two half-lines shown. (Literally, to the ‘segment’ 1 to ∞ to -1 .)

We know how to solve the problem for a harmonic ϕ on H :

$$\phi(w) = 1 - \frac{1}{\pi} \theta_2 + \frac{1}{\pi} \theta_1 = \operatorname{Re} \left(1 - \frac{1}{\pi i} \log(w-1) + \frac{1}{\pi i} \log(w+1) \right).$$

Transforming this back to the disk we have

$$u(z) = \phi \circ T^{-1}(z) = \operatorname{Re} \left(1 - \frac{1}{\pi i} \log(T^{-1}(z)-1) + \frac{1}{\pi i} \log(T^{-1}(z)+1) \right).$$

If we wanted to, we could simplify this somewhat using the formula for T^{-1} .

10.5 Flows around cylinders

10.5.1 Milne-Thomson circle theorem

The Milne-Thomson theorem allows us to insert a circle into a two-dimensional flow and see how the flow adjusts. First we’ll state and prove the theorem.

Theorem. Milne-Thomson circle theorem. If $f(z)$ is a complex potential with all its singularities outside $|z| = R$ then

$$\Phi(z) = f(z) + \overline{f\left(\frac{R^2}{\bar{z}}\right)}$$

is a complex potential with streamline on $|z| = R$ and the same singularities as f in the region $|z| > R$.

Proof. First note that R^2/\bar{z} is the reflection of z in the circle $|z| = R$.

Next we need to see that $\overline{f(R^2/\bar{z})}$ is analytic for $|z| > R$. By assumption $f(z)$ is analytic for $|z| \leq R$, so it can be expressed as a Taylor series

$$f(z) = a_0 + a_1 z + a_2 z^2 + \dots \quad (31)$$

Therefore,

$$\overline{f\left(\frac{R^2}{\bar{z}}\right)} = a_0 + a_1 \frac{R^2}{z} + a_2 \left(\frac{R^2}{z}\right)^2 + \dots \quad (32)$$

All the singularities of f are outside $|z| = R$, so the Taylor series in Equation 31 converges for $|z| \leq R$. This means the Laurent series in Equation 32 converges for $|z| \geq R$. That is, $\overline{f(R^2/\bar{z})}$ is analytic for $|z| \geq R$, i.e. it introduces no singularities to $\Phi(z)$ outside $|z| = R$.

The last thing to show is that $|z| = R$ is a streamline for $\Phi(z)$. This follows because for $z = Re^{i\theta}$.

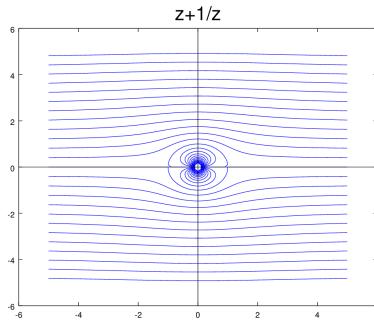
$$\Phi(Re^{i\theta}) = f(Re^{i\theta}) + \overline{f(Re^{i\theta})}$$

is real. Therefore $\psi(Re^{i\theta}) = \text{Im}(\Phi(Re^{i\theta})) = 0$.

10.5.2 Examples

Think of $f(z)$ as representing flow, possibly with sources or vortices outside $|z| = R$. Then $\Phi(z)$ represents the new flow when a circular obstacle is placed in the flow. Here are a few examples.

Example 10.16. Uniform flow around a circle. We know $f(z) = z$ is the complex potential for uniform flow to the right. So, $\Phi(z) = z + R^2/z$ is the potential for uniform flow around a circle of radius R centered at the origin.



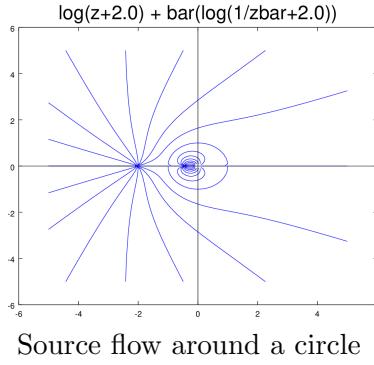
Uniform flow around a circle

Just because they look nice, the figure includes streamlines inside the circle. These don't interact with the flow outside the circle.

Note, that as z gets large flow looks uniform. We can see this analytically because $\Phi'(z) = 1 - R^2/z^2$ goes to 0 as z gets large. (Recall that the velocity field is $(\phi_x, -\phi_y)$, where $\Phi = \phi + i\psi \dots$)

Example 10.17. Source flow around a circle. Here the source is at $z = -2$ (outside the unit circle) with complex potential $f(z) = \log(z + 2)$. With the appropriate branch cut the singularities of f are also outside $|z| = 1$. So, we can apply Milne-Thomson and

$$\Phi(z) = \log(z + 2) + \overline{\log\left(\frac{1}{\bar{z}} + 2\right)}$$



We know that far from the origin the flow should look the same as a flow with just a source at $z = -2$. Let's see this analytically. First we state a useful fact:

Useful fact. If $g(z)$ is analytic then so is $h(z) = \overline{g(\bar{z})}$ and $h'(z) = \overline{g'(\bar{z})}$.

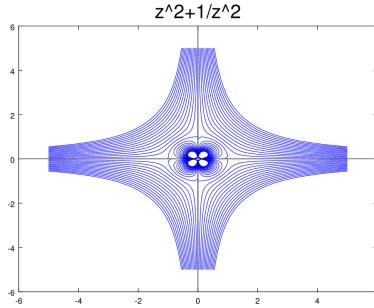
Proof. Use the Taylor series for g to get the Taylor series for h and then compare $h'(z)$ and $\overline{g'(\bar{z})}$.

Using this we have

$$\Phi'(z) = \frac{1}{z+2} - \frac{1}{z(1+2z)}$$

For large z the second term decays much faster than the first, so $\Phi'(z) \approx 1/(z+2)$. That is, the velocity field looks just like the velocity field for $f(z)$, i.e. the velocity field of a source at $z = -2$.

Example 10.18. Transforming flows. If we use $g(z) = z^2$ we can transform a flow from the upper half-plane to the first quadrant



Source flow around a quarter circular corner

11 Argument Principle

11.1 Introduction

The argument principle (or principle of the argument) is a consequence of the residue theorem. It connects the winding number of a curve with the number of zeros and poles inside the curve. This is useful for applications (mathematical and otherwise) where we want to know the location of zeros and poles.

11.2 Principle of the argument

Setup.

γ a simple closed curve

$f(z)$ analytic on and inside γ , except for (possibly) some finite poles inside (not on) γ and some zeros inside (not on) γ .

Let p_1, \dots, p_m be the poles of f inside γ .

Let z_1, \dots, z_m be the zeros of f inside γ .

Write $\text{mult}(z_k) =$ the multiplicity of the zero at z_k . Likewise write $\text{mult}(p_k) =$ the order of the pole at p_k .

Theorem 11.1. $\int_{\gamma} \frac{f'(z)}{f(z)} dz = 2\pi \left(\sum \text{mult}(z_k) - \sum \text{mult}(p_k) \right).$

Proof. Simple algebra with finite order poles and zeros. (This was given in class and will be added here at some point.)

Definition. We write $Z_{f,\gamma}$ for the sum of multiplicities of the zeros of f inside γ . Likewise for $P_{f,\gamma}$.

So the theorem says, $\int_{\gamma} \frac{f'}{f} dz = 2\pi i (Z_{f,\gamma} - P_{f,\gamma}).$

Definition. **Winding number.** We have an intuition of what this means. We define it formally via Cauchy's formula. If γ is a closed curve the its winding number (or **index**) about z_0 is defined as

$$\text{Ind}(\gamma, z_0) = \frac{1}{2\pi i} \int_{\gamma} \frac{1}{z - z_0} dz.$$

In class I drew some pictures. You should draw a few now.

11.2.1 Mapping curves: $f \circ \gamma$

One of the key notions in this topic is mapping one curve to another. That is, if $z = \gamma(t)$ is a curve and $w = f(z)$ is a function, then $w = f \circ \gamma(t) = f(\gamma(t))$ is another curve. We say f maps γ to $f \circ \gamma$. We have done this frequently in the past, but it is important enough to us now, so that we will stop here and give a few examples. This is a key concept in the argument principle and you should make sure you are very comfortable with it.

Example 11.2. Let $\gamma(t) = e^{it}$ with $0 \leq t \leq 2\pi$ (the unit circle). Let $f(z) = z^2$. Describe the curve $f \circ \gamma$.

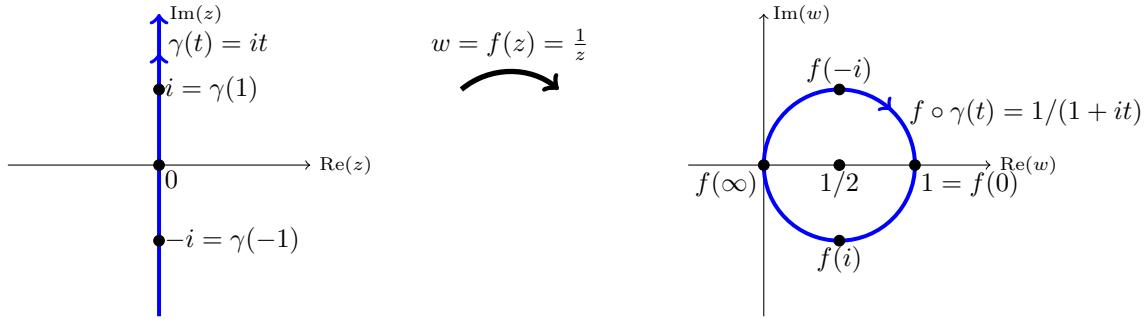
answer: Clearly $f \circ \gamma(t) = e^{2it}$ traverses the unit circle twice as t goes from 0 to 2π .

Example 11.3. Let $\gamma(t) = it$ with $-\infty < t < \infty$ (the y -axis). Let $f(z) = 1/(z + 1)$. Describe the curve $f \circ \gamma(t)$.

answer: $f(z)$ is a linear fractional transformation and maps the line given by γ to the circle through the origin centered at $1/2$. By checking at a few points:

$$f(-i) = \frac{1}{-i+1} = \frac{1+i}{2}, \quad f(0) = 1, \quad f(i) = \frac{1}{i+1} = \frac{1-i}{2}, \quad f(\infty) = 0.$$

We see that the circle is traversed in a clockwise manner as t goes from $-\infty$ to ∞ .



The curve $z = \gamma(t) = it$ is mapped to $w = f \circ \gamma(t) = 1/(it + 1)$.

11.2.2 Argument principle

You will also see this called [the principle of the argument](#).

Theorem 11.4. [Argument principle](#). For f and γ with the same setup as above

$$\int_{\gamma} \frac{f'(z)}{f(z)} dz = 2\pi i \text{Ind}(f \circ \gamma, 0) = 2\pi i (Z_{f,\gamma} - P_{f,\gamma}) \quad (33)$$

Proof. The previous theorem showed that

$$\int_{\gamma} \frac{f'(z)}{f(z)} dz = 2\pi i (Z_{f,\gamma} - P_{f,\gamma})$$

So all we have to show is that the integral also equals the winding number given. This is simply the change of variables $w = f(z)$. With this change of variables the contour $z = \gamma(t)$ becomes $w = f \circ \gamma(t)$ and $dw = f'(z) dz$ so

$$\int_{\gamma} \frac{f'(z)}{f(z)} dz = \int_{f \circ \gamma} \frac{dw}{w} = 2\pi i \text{Ind}(f \circ \gamma, 0)$$

The last equality in the above equation comes from the definition of winding number.

Note that by assumption γ does not go through any zeros of f , so $w = f(\gamma(t))$ is never zero and $1/w$ in the integral is not a problem.

Here is an easy corollary to the argument principle that will be useful to us later.

Corollary. Assuming that $f \circ \gamma$ does not go through -1 , i.e. there are no zeros of $1 + f(z)$ on γ then

$$\int_{\gamma} \frac{f'}{f+1} = 2\pi i \text{Ind}(f \circ \gamma, -1) = Z_{1+f,\gamma} - P_{1+f,\gamma} \quad (34)$$

Proof. Applying the argument principle in Equation 33 to the function $1 + f(z)$, we get

$$\int_{\gamma} \frac{(1+f)'(z)}{1+f(z)} dz = 2\pi i \text{Ind}(1 + f \circ \gamma, 0) = 2\pi i (Z_{1+f,\gamma} - P_{1+f,\gamma})$$

Now, we can compare each of the terms in this equation to those in Equation 34:

$$\begin{aligned} \int_{\gamma} \frac{(1+f)'(z)}{1+f(z)} dz &= \int_{\gamma} \frac{f'(z)}{1+f(z)} dz && (\text{because } (1+f)' = f') \\ \text{Ind}(1+f \circ \gamma, 0) &= \text{Ind}(f \circ \gamma, -1) && (1+f \text{ winds around } 0 \Leftrightarrow f \text{ winds around } -1) \\ Z_{1+f,\gamma} &= Z_{f,\gamma} && (\text{same in both equations}) \\ P_{1+f,\gamma} &= P_{f,\gamma} && (\text{poles of } f = \text{poles of } 1+f) \end{aligned}$$

Example 11.5. Let $f(z) = z^2 + z$. Find the winding number of $f \circ \gamma$ around 0 for each of the following curves.

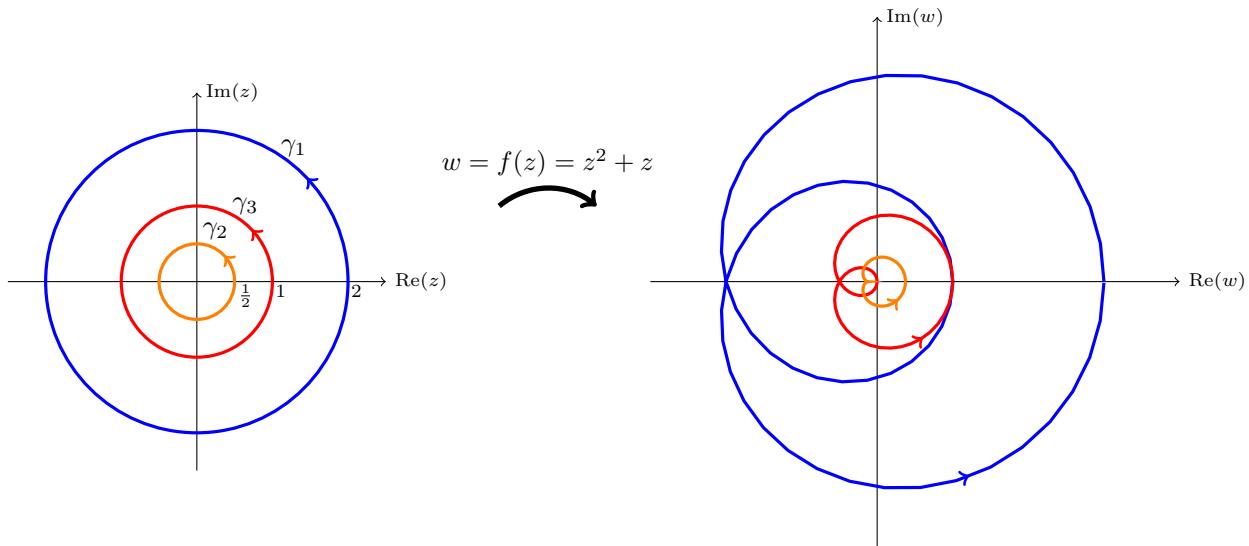
1. γ_1 = circle of radius 2.
2. γ_2 = circle of radius $1/2$.
3. γ_3 = circle of radius 1.

answers. $f(z)$ has zeros at 0, 1. It has no poles.

So, f has no poles and two zeros inside γ_1 . The argument principle says $\text{Ind}(f \circ \gamma_1, 0) = Z_{f,\gamma_1} - P_{f,\gamma_1} = 2$

Likewise f has no poles and one zero inside γ_2 , so $\text{Ind}(f \circ \gamma_2, 0) = 1 - 0 = 1$

For γ_3 a zero of f is on the curve, i.e. $f(-1) = 0$, so the argument principle doesn't apply. The image of γ_3 is shown in the figure below – it goes through 0.



The image of 3 different circles under $f(z) = z^2 + z$.

11.2.3 Rouché's theorem.

Theorem 11.6. Rouché's theorem. Make the following assumptions:

γ is simple closed curve

f, h are analytic functions on and inside γ , except for some finite poles.

There are no poles of f and h on γ .

$|h| < |f|$ everywhere on γ .

Then $\text{Ind}(f \circ \gamma, 0) = \text{Ind}((f + h) \circ \gamma, 0)$. That is,

$$Z_{f,\gamma} - P_{f,\gamma} = Z_{f+h,\gamma} - P_{f+h,\gamma} \quad (35)$$

Proof. In class we gave a heuristic proof involving a person walking a dog around $f \circ \gamma$ on a leash of length $h \circ \gamma$. Here is the analytic proof.

The argument principle requires the function to have no zeros or poles on γ . So we first show that this is true of f , $f + h$, $(f + h)/f$. The argument is quick:

Zeros: The fact that $0 \leq |h| < |f|$ on γ implies f has no zeros on γ . It also implies $f + h$ has no zeros on γ , since the value of h is never big enough to cancel that of f . Since f and $f + h$ have no zeros, neither does $(f + h)/f$.

Poles: By assumption f and h have no poles on γ , so $f + h$ has no poles there. Since f has no zeros on γ , $(f + h)/f$ has no poles there.

Now we can apply the argument principle to f and $f + h$

$$\frac{1}{2\pi i} \int_{\gamma} \frac{f'}{f} dz = \text{Ind}(f \circ \gamma, 0) = Z_{f,\gamma} - P_{f,\gamma}. \quad (36)$$

$$\frac{1}{2\pi i} \int_{\gamma} \frac{(f + h)'}{f + h} dz = \text{Ind}((f + h) \circ \gamma, 0) = Z_{f+h,\gamma} - P_{f+h,\gamma}. \quad (37)$$

Next, by assumption $\left| \frac{h}{f} \right| < 1$, so $\left(\frac{h}{f} \right) \circ \gamma$ is inside the unit circle. This means that $1 + \frac{h}{f} = \frac{f + h}{f}$ maps γ to the inside of the unit disk centered at 1. (You should draw a figure for this.) This implies that

$$\text{Ind} \left(\left(\frac{f + h}{f} \right) \circ \gamma, 0 \right) = 0.$$

Let $g = \frac{f + h}{f}$. The above says $\text{Ind}(g \circ \gamma, 0) = 0$. So, $\int_{\gamma} \frac{g'}{g} dz = 0$. (We showed above that g has no zeros or poles on γ .)

Now, it's easy to compute that $\frac{g'}{g} = \frac{(f + h)'}{f + h} - \frac{f'}{f}$. So, using

$$\text{Ind}(g \circ \gamma, 0) = \int_{\gamma} \frac{g'}{g} dz = \int_{\gamma} \frac{(f + h)'}{f + h} dz - \int_{\gamma} \frac{f'}{f} dz = 0 \Rightarrow \text{Ind}((f + h) \circ \gamma, 0) = \text{Ind}(f \circ \gamma, 0).$$

Now equations 36 and 37 tell us $Z_{f,\gamma} - P_{f,\gamma} = Z_{f+h,\gamma} - P_{f+h,\gamma}$, i.e. we have proved Rouché's theorem.

Corollary. Under the same hypotheses, If h and f are analytic (no poles) then $Z_{f,\gamma} = Z_{f+h,\gamma}$.

Proof. Since the functions are analytic $P_{f,\gamma}$ and $P_{f+h,\gamma}$ are both 0. So Equation 35 shows $Z_f = Z_{f+h}$. QED.

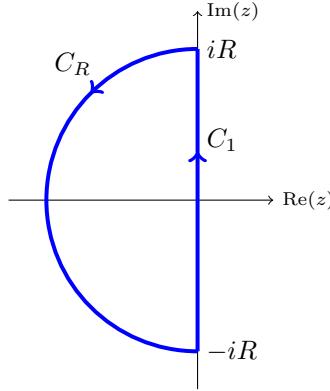
We think of h as a small perturbation of f .

Example 11.7. Show all 5 zeros of $z^5 + 3z + 1$ are inside the curve $C_2 : |z| = 2$.

answer: Let $f(z) = z^5$ and $h(z) = 3z + 1$. Clearly all 5 roots of f (really one root with multiplicity 5) are inside C_2 . Also clearly, $|h| < 7 < 32 = |f|$ on C_2 . The corollary to Rouch  s theorem says all 5 roots of $f + h = z^5 + 3z + 1$ must also be inside the curve.

Example 11.8. Show $z + 3 + 2e^z$ has one root in the left half-plane.

answer: Let $f(z) = z + 3$, $h(z) = 2e^z$. Consider the contour from $-iR$ to iR along the y -axis and then the left semicircle of radius R back to $-iR$. That is, the contour $C_1 + C_R$ shown below.



To apply the corollary to Rouch  s theorem we need to check that (for R large) $|h| < |f|$ on $C_1 + C_R$. On C_1 , $z = iy$, so

$$|f(z)| = |3 + iy| \geq 3, \quad |h(z)| = 2|e^{iy}| = 2.$$

So $|h| < |f|$ on C_1 .

On C_R , $z = x + iy$ with $x < 0$ and $|z| = R$. So,

$$|f(z)| > R - 3 \text{ for } R \text{ large,} \quad |h(z)| = 2|e^{x+iy}| = 2e^x < 2 \text{ (since } x < 0\text{).}$$

So $|h| < |f|$ on C_R .

The only zero of f is at $z = -3$, which lies inside the contour.

Therefore, by the Corollary to Rouch  s theorem, $f + h$ has the same number of roots as f inside the contour, that is 1. Now let R go to infinity and we see that $f + h$ has only one root in the entire half-plane.

Theorem. Fundamental theorem of algebra.

Rouch  s theorem can be used to prove the fundamental theorem of algebra as follows.

Proof. Let

$$P(z) = z^n + a_{n-1}z^{n-1} + \dots + a_0$$

be an n th order polynomial. Let $f(z) = z^n$ and $h = P - f$. Choose an R such that $R > \max(1, n|a_{n-1}|, \dots, n|a_0|)$. Then on $|z| = R$ we have

$$|h| \leq |a_{n-1}|R^{n-1} + |a_{n-2}|R^{n-2} + \dots + |a_0| \leq \frac{R}{n}R^{n-1} + \frac{R}{n}R^{n-2} + \dots + \frac{R}{n} < R^n.$$

On $|z| = R$ we have $|f(z)| = R^n$, so we have shown $|h| < |f|$ on the curve. Thus, the corollary to Rouch  s theorem says $f + h$ and f have the same number of zeros inside

$|z| = R$. Since we know f has exactly n zeros inside the curve the same is true for the polynomial $f + h$. Now let R go to infinity, we've shown that $f + h$ has exactly n zeros in the entire plane.

11.3 Nyquist criterion for stability

The Nyquist criterion is a graphical technique for telling whether an unstable linear time invariant system can be stabilized using a negative feedback loop. We will look a little more closely at such systems when we study the Laplace transform in the next topic. For this topic we will content ourselves with a statement of the problem with only the tiniest bit of physical context.

Note. You have already encountered linear time invariant systems in 18.03 (or its equivalent) when you solved constant coefficient linear differential equations.

11.3.1 System functions

A linear time invariant systems has a **system function** which is a function of a complex variable. Typically, the complex variable is denoted by s and a capital letter is used for the system function.

Let $G(s)$ be such a system function. We will make a standard assumption that $G(s)$ is meromorphic with a finite number of (finite) poles. This assumption holds in many interesting cases. For example, quite often $G(s)$ is a rational function $Q(s)/P(s)$ (Q and P are polynomials).

We will be concerned with the stability of the system.

Definition. The system with system function $G(s)$ is called **stable** if all the poles of G are in the left half-plane. That is, if all the poles of G have negative real part.

The system is called **unstable** if any poles are in the right half-plane, i.e. have positive real part.

For the edge case where no poles have positive real part, but some are pure imaginary we will call the system **marginally stable**. This case can be analyzed using our techniques. For our purposes it would require and an indented contour along the imaginary axis. If we have time we will do the analysis.

Example 11.9. Is the system $G(s) = \frac{s}{(s+2)(s^2+4s+5)}$ stable?

answer: The poles are $-2, -2 \pm i$. Since they are all in the left half-plane, the system is stable.

Example 11.10. Is the system $G(s) = \frac{s}{(s^2-4)(s^2+4s+5)}$ stable?

answer: The poles are $\pm 2, -2 \pm i$. Since one pole is in the right half-plane, the system is unstable.

Example 11.11. Is the system $G(s) = \frac{s}{(s+2)(s^2+4)}$ stable?

answer: The poles are $-2, \pm 2i$. There are no poles in the right half-plane. Since there are poles on the imaginary axis, the system is marginally stable.

Terminology. So far, we have been careful to say ‘the system with system function $G(s)$ ’. From now on we will allow ourselves to be a little more casual and say ‘the system $G(s)$ ’. It is perfectly clear and rolls off the tongue a little easier!

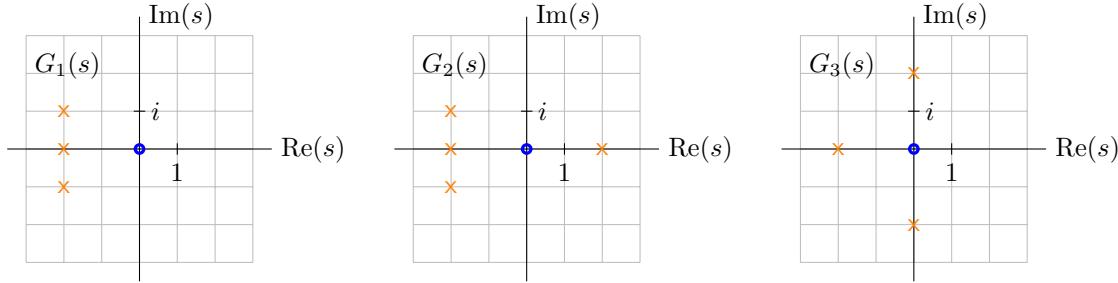
11.3.2 Pole-zero diagrams

We can visualize $G(s)$ using a [pole-zero diagram](#). This is a diagram in the s -plane where we put a small cross at each pole and a small circle at each zero.

Example 11.12. Give zero-pole diagrams for each of the systems

$$G_1(s) = \frac{s}{(s+2)(s^2+4s+5)}, \quad G_2(s) = \frac{s}{(s^2-4)(s^2+4s+5)}, \quad G_3(s) = \frac{s}{(s+2)(s^2+4)}$$

answer: These are the same systems as in the examples just above. We first note that they all have a single zero at the origin. So we put a circle at the origin and a cross at each pole.



Pole-zero diagrams for the three systems.

11.3.3 A bit about stability

This is just to give you a little physical orientation. Given our definition of stability above, we could, in principle, discuss stability without the slightest idea what it means for physical systems.

The poles of $G(s)$ correspond to what are called modes of the system. A simple pole at s_1 corresponds to a mode $y_1(t) = e^{s_1 t}$. The system is stable if the modes all decay to 0, i.e. if the poles are all in the left half-plane.

Physically the modes tell us the behavior of the system when the input signal is 0, but there are initial conditions. A pole with positive real part would correspond to a mode that goes to infinity as t grows. It is certainly reasonable to call a system that does this in response to a zero signal (often called ‘no input’) unstable.

To connect this to 18.03: if the system is modeled by a differential equation, the modes correspond to the homogeneous solutions $y(t) = e^{st}$, where s is a root of the characteristic equation. In 18.03 we called the system stable if every homogeneous solution decayed to 0. That is, if the unforced system always settled down to equilibrium.

11.3.4 Closed loop systems

If the system with system function $G(s)$ is unstable it can sometimes be stabilized by what is called a [negative feedback loop](#). The new system is called a [closed loop system](#). Its system function is given by Black's formula

$$G_{CL}(s) = \frac{G(s)}{1 + kG(s)}, \quad (38)$$

where k is called the [feedback factor](#). We will just accept this formula. Any class or book on control theory will derive it for you.

In this context $G(s)$ is called the [open loop system function](#).

Since G_{CL} is a system function, we can ask if the system is stable.

Theorem. The poles of the closed loop system function $G_{CL}(s)$ given in Equation 38 are the zeros of $1 + kG(s)$.

Proof. Looking at Equation 38, there are two possible sources of poles for G_{CL} .

1. The zeros of the denominator $1 + kG$. The theorem recognizes these.
2. The poles of G . Since G is in both the numerator and denominator of G_{CL} it should be clear that the poles cancel. We can show this formally using Laurent series. If G has a pole of order n at s_0 then

$$G(s) = \frac{1}{(s - s_0)^n} (b_n + b_{n-1}(s - s_0) + \dots + a_0 + \dots),$$

where $b_n \neq 0$. So,

$$G_{CL}(s) = \frac{\frac{1}{(s-s_0)^n} (b_n + b_{n-1}(s - s_0) + \dots + a_0 + \dots)}{1 + \frac{k}{(s-s_0)^n} (b_n + b_{n-1}(s - s_0) + \dots + a_0 + \dots)} = \frac{(b_n + b_{n-1}(s - s_0) + \dots + a_0 + \dots)}{(s - s_0)^n + k(b_n + b_{n-1}(s - s_0) + \dots + a_0 + \dots)},$$

which is clearly analytic at s_0 . (At s_0 it equals $b_n/(kb_n) = 1/k$.)

Example 11.13. Set the feedback factor $k = 1$. For what values of a is the open loop system $G(s) = \frac{1}{s+a}$ stable? For what values of a is the corresponding closed loop system $G_{CL}(s)$ stable?

answer: $G(s)$ has one pole at $s = -a$. Thus, it is stable when the pole is in the left half-plane, i.e. for $a > 0$.

The closed loop system function is

$$G_{CL}(s) = \frac{1/(s+a)}{1 + 1/(s+a)} = \frac{1}{s+a+1}.$$

This has a pole at $s = -a - 1$, so it's stable if $a > -1$. The feedback loop has stabilized the unstable open loop systems with $-1 < a \leq 0$. (Actually, for $a = 0$ the open loop is marginally stable, but it is fully stabilized by the closed loop.)

Note. The algebra involved in canceling the $s + a$ term in the denominators is exactly the cancellation that makes the poles of G removable singularities in G_{CL} .

Example 11.14. Suppose $G(s) = \frac{s+1}{s-1}$. Is the open loop system stable? Is the closed loop system stable when $k = 2$.

answer: $G(s)$ has a pole in the right half-plane, so the open loop system is not stable. The closed loop system function is

$$G_{CL}(s) = \frac{G}{1+kG} = \frac{(s+1)/(s-1)}{1+2(s+1)/(s-1)} = \frac{s+1}{3s+1}.$$

The only pole is at $s = -1/3$, so the closed loop system is stable. This is a case where feedback stabilized an unstable system.

Example 11.15. $G(s) = \frac{s-1}{s+1}$. Is the open loop system stable? Is the closed loop system stable when $k = 2$.

answer: The only pole of $G(s)$ is in the left half-plane, so the open loop system is stable. The closed loop system function is

$$G_{CL}(s) = \frac{G}{1+kG} = \frac{(s-1)/(s+1)}{1+2(s-1)/(s+1)} = \frac{s-1}{3s-1}.$$

This has one pole at $s = 1/3$, so the closed loop system is unstable. This is a case where feedback destabilized a stable system. It can happen!

11.3.5 Nyquist plot

For the Nyquist plot and criterion the curve γ will always be the imaginary s -axis. That is

$$s = \gamma(\omega) = i\omega, \text{ where } -\infty < \omega < \infty.$$

For a system $G(s)$ and a feedback factor k , the **Nyquist plot** is the plot of the curve

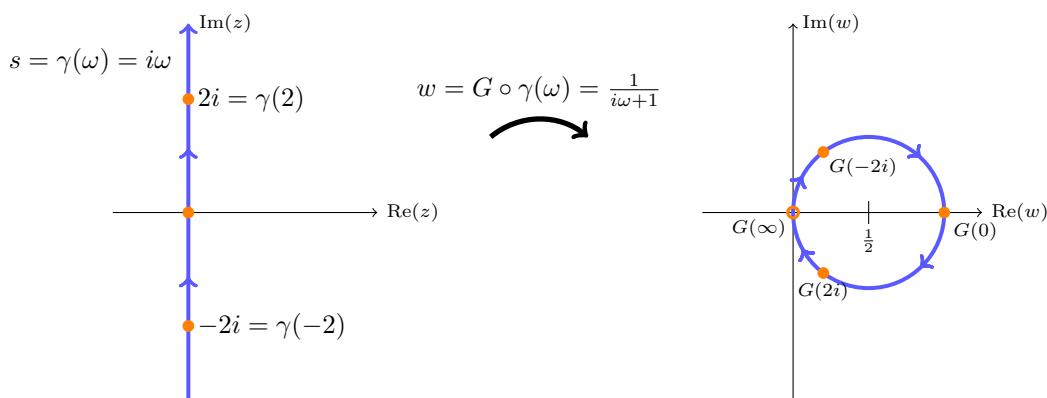
$$w = kG \circ \gamma(\omega) = kG(i\omega).$$

That is, the Nyquist plot is the image of the imaginary axis under the map $w = kG(s)$.

Note. In $\gamma(i\omega)$ the variable is a greek omega and in $w = G \circ \gamma$ we have a double-u.

Example 11.16. Let $G(s) = \frac{1}{s+1}$. Draw the Nyquist plot with $k = 1$.

answer: In this case $G(s)$ is a linear fractional transformation, so we know it maps the imaginary axis to a circle. It is easy to check it is the circle through the origin with center $w = 1/2$. You can also check that it is traversed clockwise.



Nyquist plot of $G(s) = 1/(s + 1)$.

Example 11.17. Take $G(s)$ from the previous example. Describe the Nyquist plot with gain factor $k = 2$.

answer: The Nyquist plot is the graph of $kG(i\omega)$. The factor $k = 2$ will scale the circle in the previous example by 2. That is, the Nyquist plot is the circle through the origin with center $w = 1$.

In general, the feedback factor will just scale the Nyquist plot.

11.3.6 Nyquist criterion

The Nyquist criterion gives a graphical method for checking the stability of the closed loop system.

Theorem 11.18. Nyquist criterion. Suppose that $G(s)$ has a finite number of zeros and poles in the right half-plane. Also suppose that $G(s)$ decays to 0 as s goes to infinity. Then the closed loop system with feedback factor k is stable if and only if the winding number of the Nyquist plot around $w = -1$ equals the number of poles of $G(s)$ in the right half-plane.

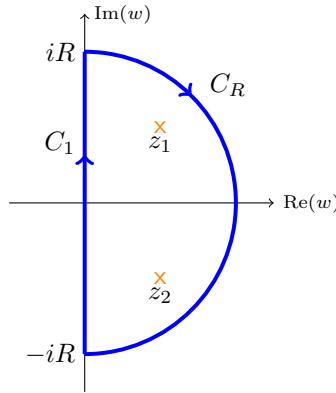
More briefly,

$$G_{CL}(s) \text{ is stable} \Leftrightarrow \text{Ind}(G \circ \gamma, -1) = P_{G,\text{RHP}}$$

Here, γ is the imaginary s -axis and $P_{G,\text{RHP}}$ is the number of poles of **the original open loop system function $G(s)$** in the right half-plane.

Proof. G_{CL} is stable exactly when all its poles are in the left half-plane. Now, recall that the poles of G_{CL} are exactly the zeros of $1 + kG$. So, stability of G_{CL} is exactly the condition that the number of zeros of $1 + kG$ in the right half-plane is 0.

Let's work with a familiar contour.



Let $\gamma_R = C_1 + C_R$. Note that γ_R is traversed in the *clockwise* direction. Choose R large enough that the (finite number) of poles and zeros of G in the right half-plane are all inside γ_R . Now we can apply Equation 34 in the corollary to the argument principle to $kG(s)$ and γ to get

$$-\text{Ind}(kG \circ \gamma_R, -1) = Z_{1+kG, \gamma_R} - P_{G, \gamma_R}$$

(The minus sign is because of the clockwise direction of the curve.) Thus, for all large R

$$\text{the system is stable} \Leftrightarrow Z_{1+kG, \gamma_R} = 0 \Leftrightarrow \text{Ind}(kG \circ \gamma_R, -1) = P_{G, \gamma_R}.$$

Finally, we can let R go to infinity. The assumption that $G(s)$ decays to as s goes to ∞ implies that in the limit, the entire curve $G \circ C_R$ becomes a single point at the origin. So in the limit $G \circ \gamma_R$ becomes $G \circ \gamma$. QED

11.3.7 Examples using the Nyquist Plot mathlet

The Nyquist criterion is a visual method which requires some way of producing the Nyquist plot. For this we will use one of the MIT Mathlets (slightly modified for our purposes).

Open the Nyquist Plot applet at

<http://web.mit.edu/jorloff/www/jmoapplets/nyquist/nyquistPlot.html>

Play with the applet, read the help.

Now refresh the browser to restore the applet to its original state. Check the *Formula* box. The formula is an easy way to read off the values of the poles and zeros of $G(s)$. The applet should have a zero at $s = 1$ and poles at $s = 0.33 \pm 1.75i$.

The left hand graph is the pole-zero diagram. The right hand graph is the Nyquist plot.

Example 11.19. To get a feel for the Nyquist plot. Look at the pole diagram and use the mouse to drag the yellow point up and down the imaginary axis. Its image under $kG(s)$ will trace out the Nyquist plot.

Notice that when the yellow dot is at either end of the axis its image on the Nyquist plot is close to 0.

Example 11.20. Refresh the page, to put the zero and poles back to their original state. There are two poles in the right half-plane, so the open loop system $G(s)$ is unstable. With $k = 1$, what is the winding number of the Nyquist plot around -1? Is the closed loop system stable?

answer: The curve winds twice around -1 in the counterclockwise direction, so the winding number $\text{Ind}(kG \circ \gamma, -1) = 2$. Since the number of poles of G in the right half-plane is the same as this winding number, the closed loop system is stable.

Example 11.21. With the same poles and zeros, move the k slider and determine what range of k makes the closed loop system stable.

answer: When k is small the Nyquist plot has winding number 0 around -1. For these values of k , G_{CL} is unstable. As k increases, somewhere between $k = 0.65$ and $k = 0.7$ the winding number jumps from 0 to 2 and the closed loop system becomes stable. This continues until k is between 3.10 and 3.20, at which point the winding number becomes 1 and G_{CL} becomes unstable.

Answer: The closed loop system is stable for k (roughly) between 0.7 and 3.10.

Example 11.22. In the previous problem could you determine analytically the range of k where $G_{CL}(s)$ is stable?

answer: Yes! This is possible for small systems. It is more challenging for higher order systems, but there are methods that don't require computing the poles.

In this case, we have

$$G_{CL}(s) = \frac{G(s)}{1 + kG(s)} = \frac{\frac{s-1}{(s-0.33)^2+1.75^2}}{1 + \frac{k(s-1)}{(s-0.33)^2+1.75^2}} = \frac{s-1}{(s-0.33)^2 + 1.75^2 + k(s-1)}$$

So the poles are the roots of

$$(s - 0.33)^2 + 1.75^2 + k(s - 1) = s^2 + (k - 0.66)s + 0.33^2 + 1.75^2 - k$$

For a quadratic with positive coefficients the roots both have negative real part. This happens when

$$0.66 < k < 0.33^2 + 1.75^2 \approx 3.17.$$

Example 11.23. What happens when k goes to 0.

answer: As k goes to 0, the Nyquist plot shrinks to a single point at the origin. In this case the winding number around -1 is 0 and the Nyquist criterion says the closed loop system is stable if and only if the open loop system is stable.

This should make sense, since with $k = 0$,

$$G_{CL} = \frac{G}{1 + kG} = G.$$

Example 11.24. Make a system with the following zeros and poles:

A pair of zeros at $0.6 \pm 0.75i$

A pair of poles at $-0.5 \pm 2.5i$.

A single pole at 0.25.

Is the corresponding closed loop system stable when $k = 6$.

answer: The answer is no, G_{CL} is not stable. G has one pole in the right half plane. The mathlet shows the Nyquist plot winds once around $w = -1$ in the *clockwise* direction. So the winding number is -1, which does not equal the number of poles of G in the left half-plane.

If we set $k = 3$, the closed loop system is stable.

11.4 A bit on negative feedback

Given Equation 38, in 18.04 we can ask if there are any poles in the right half-plane without needing any underlying physical model. Still, it's nice to have some sense of where this fits into science and engineering.

In a negative feedback loop the output of the system is looped back and subtracted from the input.

Example 11.25. The heating system in my house is an example of a system stabilized by feedback. The thermostat is the feedback mechanism. When the temperature outside (input signal) goes down the heat turns on. Without the thermostat it would stay on and

overheat my house. The thermostat turns the heat up or down depending on whether the inside temperature (the output signal) is too low or too high (negative feedback).

Example 11.26. Walking or balancing on one foot are examples negative feedback systems. If you feel yourself falling you compensate by shifting your weight or tensing your muscles to counteract the unwanted acceleration.

12 Laplace transform

12.1 Introduction

The Laplace transform takes a function of time and transforms it to a function of a complex variable s . Because the transform is invertible, no information is lost and it is reasonable to think of a function $f(t)$ and its Laplace transform $F(s)$ as two views of the same phenomenon. Each view has its uses and some features of the phenomenon are easier to understand in one view or the other.

We can use the Laplace transform to transform a linear time invariant system from the time domain to the s -domain. This leads to the system function $G(s)$ for the system –this is the same system function used in the Nyquist criterion for stability.

One important feature of the Laplace transform is that it can transform analytic problems to algebraic problems. We will see examples of this for differential equations.

12.2 A brief introduction to linear time invariant systems

Let's start by defining our terms.

Signal. A **signal** is any function of time.

System. A **system** is some machine or procedure that takes one signal as **input** does something with it and produces another signal as **output**.

Linear system. A **linear system** is one that acts linearly on inputs. That is, $f_1(t)$ and $f_2(t)$ are inputs to the system with outputs $y_1(t)$ and $y_2(t)$ respectively, then the input $f_1 + f_2$ produces the output $y_1 + y_2$ and, for any constant c , the input cf_1 produces output cy_1 .

This is often phrased in one sentence as input $c_1f_1 + c_2f_2$ produces output $c_1y_1 + c_2y_2$, i.e. linear combinations of inputs produces a linear combination of the corresponding outputs.

Time invariance. Suppose a system takes input signal $f(t)$ and produces output signal $y(t)$. The system is called **time invariant** if the input signal $g(t) = f(t - a)$ produces output signal $y(t - a)$.

LTI. We will call a linear time invariant system an **LTI system**.

Example 12.1. Consider the constant coefficient differential equation

$$3y'' + 8y' + 7y = f(t)$$

This equation models a damped harmonic oscillator, say a mass on a spring with a damper, where $f(t)$ is the force on the mass and $y(t)$ is its displacement from equilibrium. If we

consider f to be the input and y the output, then this is a linear time invariant (LTI) system.

Example 12.2. There are many variations on this theme. For example, we might have the LTI system

$$3y'' + 8y' + 7y = f'(t),$$

where we call $f(t)$ the input signal and $y(t)$ the output signal.

12.3 Laplace transform

Definition. The [Laplace transform](#) of a function $f(t)$ is defined by the integral

$$\mathcal{L}(f; s) = \int_0^\infty e^{-st} f(t) dt,$$

for those s where the integral converges. Here s is allowed to take complex values.

Important note. The Laplace transform is only concerned with $f(t)$ for $t \geq 0$. Generally, speaking we can require $f(t) = 0$ for $t < 0$.

Standard notation. Where the notation is clear, we will use an upper case letter to indicate the Laplace transform, e.g., $\mathcal{L}(f; s) = F(s)$.

The Laplace transform we defined is sometimes called the one-sided Laplace transform. There is a two-sided version where the integral goes from $-\infty$ to ∞ .

12.3.1 First examples

Let's compute a few examples. We will also put these results in the Laplace transform table at the end of these notes.

Example 12.3. Let $f(t) = e^{at}$. Compute $F(s) = \mathcal{L}(f; s)$ directly. Give the region in the complex s -plane where the integral converges.

$$\begin{aligned} \mathcal{L}(e^{at}; s) &= \int_0^\infty e^{at} e^{-st} dt = \int_0^\infty e^{(a-s)t} dt = \frac{e^{(a-s)t}}{a-s} \Big|_0^\infty \\ &= \begin{cases} \frac{1}{s-a} & \text{if } \operatorname{Re}(s) > \operatorname{Re}(a) \\ \text{divergent} & \text{otherwise} \end{cases} \end{aligned}$$

The last formula comes from plugging ∞ into the exponential. This is 0 if $\operatorname{Re}(a - s) < 0$ and undefined otherwise.

Example 12.4. Let $f(t) = b$. Compute $F(s) = \mathcal{L}(f; s)$ directly. Give the region in the complex s -plane where the integral converges.

$$\begin{aligned} \mathcal{L}(b; s) &= \int_0^\infty b e^{-st} dt = \frac{b e^{-st}}{-s} \Big|_0^\infty \\ &= \begin{cases} \frac{b}{s} & \text{if } \operatorname{Re}(s) > 0 \\ \text{divergent} & \text{otherwise} \end{cases} \end{aligned}$$

The last formula comes from plugging ∞ into the exponential. This is 0 if $\operatorname{Re}(-s) < 0$ and undefined otherwise.

Example 12.5. Let $f(t) = t$. Compute $F(s) = \mathcal{L}(f; s)$ directly. Give the region in the complex s -plane where the integral converges.

$$\begin{aligned}\mathcal{L}(t; s) &= \int_0^\infty te^{-st} dt = \frac{te^{-st}}{-s} - \frac{e^{-st}}{s^2} \Big|_0^\infty \\ &= \begin{cases} \frac{1}{s^2} & \text{if } \operatorname{Re}(s) > 0 \\ \text{divergent} & \text{otherwise} \end{cases}\end{aligned}$$

Example 12.6. Compute $\mathcal{L}(\cos(\omega t))$

answer: We use the formula $\cos(\omega t) = \frac{e^{i\omega t} + e^{-i\omega t}}{2}$. So,

$$\mathcal{L}(\cos(\omega t); s) = \frac{1/(s - i\omega) + 1/(s + i\omega)}{2} = \frac{s}{s^2 + \omega^2}$$

12.3.2 Connection to Fourier transform

The Laplace and Fourier transforms are intimately connected. In fact, the Laplace transform is often called the Fourier-Laplace transform. To see the connection we'll start with the Fourier transform of a function $f(t)$. $\hat{f}(\omega) = \int_{-\infty}^\infty f(t)e^{-i\omega t} dt$. If we assume $f(t) = 0$ for $t < 0$, this becomes

$$\hat{f}(\omega) = \int_0^\infty f(t)e^{-i\omega t} dt. \quad (39)$$

Now if $s = i\omega$ then the Laplace transform is

$$\mathcal{L}(f; s) = \mathcal{L}(f; i\omega) = \int_0^\infty f(t)e^{-i\omega t} dt \quad (40)$$

Comparing these two equations we see that $\hat{f}(\omega) = \mathcal{L}(f; i\omega)$. We see the transforms are basically the same things using different notation –at least for functions that are 0 for $t < 0$.

12.4 Exponential type

The Laplace transform is defined when the integral for it converges. Functions of exponential type are a class of functions for which the integral converges for all s with $\operatorname{Re}(s)$ large enough.

Definition. We say that $f(t)$ has **exponential type a** if there exists an M such that $|f(t)| < M e^{at}$ for all $t \geq 0$.

Note. As we've defined it, the exponential type of a function is not unique. For example, a function of exponential type 2 is clearly also of exponential type 3. It's nice, but not always necessary, to find the smallest exponential type for a function.

Theorem. If f has exponential type a then $\mathcal{L}(f)$ converges absolutely for $\operatorname{Re}(s) > a$.

Proof. We prove absolute convergence by bounding $|f(t)e^{-st}|$. The key here is that $\operatorname{Re}(s) > a$ implies $\operatorname{Re}(a - s) < 0$. So, we can write

$$\int_0^\infty |f(t)e^{-st}| dt \leq \int_0^\infty |M e^{(a-s)t}| dt = \int_0^\infty M e^{\operatorname{Re}(a-s)t} dt$$

The last integral clearly converges when $\operatorname{Re}(a - s) < 0$. QED

Example 12.7. Here is a list of some functions of exponential type.

$$\begin{aligned} f(t) = e^{at} : \quad &|f(t)| < 2e^{\operatorname{Re}(a)t} \quad (\text{exponential type } \operatorname{Re}(a)) \\ f(t) = 1 : \quad &|f(t)| < 2 = 2e^{0 \cdot t} \quad (\text{exponential type 0}) \\ f(t) = \cos(\omega t) : \quad &|f(t)| \leq 1 \quad (\text{exponential type 0}) \end{aligned}$$

In the above, all of the inequalities are for $t \geq 0$.

For $f(t) = t$, it is clear that for any $a > 0$ there is an M depending on a such that $|f(t)| \leq M e^{at}$ for $t \geq 0$. In fact, it is a simple calculus exercise to show $M = 1/(ae)$ works. So, $f(t) = t$ has exponential type a for any $a > 0$.

The same is true of t^n . It's worth pointing out that this follows because, if f has exponential type a and g has exponential type b then fg has exponential type $a + b$. So, if t has exponential type a then t^n has exponential type na .

12.5 Properties of Laplace transform

We have already used the linearity of Laplace transform when we computed $\mathcal{L}(\cos(\omega t))$. Let's officially record it as a property.

Property 1. The Laplace transform is linear. That is, if a and b are constants and f and g are functions then

$$\mathcal{L}(af + bg) = a\mathcal{L}(f) + b\mathcal{L}(g). \quad (41)$$

(The proof is trivial – integration is linear.)

Property 2. A key property of the Laplace transform is that, with some technical details,

Laplace transform transforms derivatives in t to multiplication by s (plus some details).

This is proved in the following theorem.

Theorem. If $f(t)$ has exponential type a and Laplace transform $F(s)$ then

$$\mathcal{L}(f'(t); s) = sF(s) - f(0), \text{ valid for } \operatorname{Re}(s) > a. \quad (42)$$

Proof. We prove this using integration by parts.

$$\mathcal{L}(f'; s) = \int_0^\infty f'(t)e^{-st} dt = f(t)e^{-st}\Big|_0^\infty + \int_0^\infty sf(t)e^{-st} dt = -f(0) + sF(s).$$

In the last step we used that at $t = \infty$, $f(t)e^{-st} = 0$, which follows from the assumption about exponential type.

Equation 42 gives us formulas for all derivatives of f .

$$\mathcal{L}(f''; s) = s^2 F(s) - sf(0) - f'(0) \quad (43)$$

$$\mathcal{L}(f'''; s) = s^3 F(s) - s^2 f(0) - sf'(0) - f''(0) \quad (44)$$

Proof. For Equation 43:

$$\mathcal{L}(f''; s) = \mathcal{L}((f')'; s) = s\mathcal{L}(f'; s) - f'(0) = s(sF(s) - f(0)) - f'(0) = s^2 F(s) - sf(0) - f'(0). \text{ QED}$$

The proof Equation 44 is similar. Also, similar statements hold for higher order derivatives.

Note. There is a further complication if we want to consider functions that are discontinuous at the origin or if we want to allow $f(t)$ to be a generalized function like $\delta(t)$. In these cases $f(0)$ is not defined, so our formulas are undefined. The technical fix is to replace 0 by 0^- in the definition and all of the formulas for Laplace transform. You can learn more about this by taking 18.031.

Property 3. Theorem. If $f(t)$ has exponential type a , then $F(s)$ is an analytic function for $\operatorname{Re}(s) > a$ and

$$F'(s) = -\mathcal{L}(tf(t); s). \quad (45)$$

Proof. We take the derivative of $F(s)$. The absolute convergence for $\operatorname{Re}(s)$ guarantees that we can interchange the order of integration and taking the derivative.

$$F'(s) = \frac{d}{ds} \int_0^\infty f(t)e^{-st} dt = \int_0^\infty -tf(t)e^{-st} dt = \mathcal{L}(-tf(t); s).$$

This proves Equation 45.

Equation 45 is called the **s-derivative rule**. We can extend it to more derivatives in s : Suppose $\mathcal{L}(f; s) = F(s)$. Then,

$$\mathcal{L}(tf(t); s) = -F'(s) \quad (46)$$

$$\mathcal{L}(t^n f(t); s) = (-1)^n F^{(n)}(s) \quad (47)$$

Equation 46 is the same as Equation 45 above. Equation 47 follows from this.

Example 12.8. Use the s-shift rule and the formula $\mathcal{L}(1; s) = 1/s$ to compute the Laplace transform of t^n for n a positive integer.

answer: Let $f(t) = 1$ and $F(s) = \mathcal{L}(f; s)$. Using the s-shift rule we get

$$\mathcal{L}(t; s) = \mathcal{L}(tf; s) = -F'(s) = \frac{1}{s^2}$$

$$\mathcal{L}(t^2; s) = \mathcal{L}(t^2 f; s) = (-1)^2 F''(s) = \frac{2}{s^3}$$

$$\mathcal{L}(t^n; s) = \mathcal{L}(t^n f; s) = (-1)^n F^{(n)}(s) = \frac{n!}{s^{n+1}}$$

Property 4. t-shift rule. As usual, assume $f(t) = 0$ for $t < 0$. Suppose $a > 0$. Then,

$$\mathcal{L}(f(t-a); s) = e^{-as} F(s) \quad (48)$$

Proof. We go back to the definition of the Laplace transform and make the change of variables $\tau = t - a$.

$$\begin{aligned}\mathcal{L}(f(t-a); s) &= \int_0^\infty f(t-a)e^{-st} dt = \int_a^\infty f(t-a)e^{-st} dt \\ &= \int_0^\infty f(\tau)e^{-s(\tau+a)} d\tau = e^{-sa} \int_0^\infty f(\tau)e^{-s\tau} d\tau = e^{-sa} F(s).\end{aligned}$$

The properties in Equations 41-48 will be used in examples below. They are also in the table at the end of these notes.

12.6 Differential equations

Coverup method. We are going to use partial fractions and the coverup method. We will assume you have seen partial fractions. If you don't remember them well or have never seen the coverup method, you should read the note *Partial fractions and the coverup method* posted with the class notes.

Example 12.9. Solve $y'' - y = e^{2t}$, $y(0) = 1$, $y'(0) = 1$ using Laplace transform.

answer: Call $\mathcal{L}(y) = Y$. Apply the Laplace transform to the equation:

$$(s^2Y - sy(0) - y'(0)) - Y = \frac{1}{s-2}$$

Algebra: $(s^2 - 1)Y = \frac{1}{s-2} + s + 1$, so

$$Y = \frac{1}{(s-2)(s^2-1)} + \frac{s+1}{s^2-1} = \frac{1}{(s-2)(s^2-1)} + \frac{1}{s-1}$$

Use partial fractions to write

$$Y = \frac{A}{s-2} + \frac{B}{s-1} + \frac{C}{s+1} + \frac{1}{s-1}.$$

The coverup method gives $A = 1/3$, $B = -1/2$, $C = 1/6$.

We recognize $\frac{1}{s-a}$ as the Laplace transform of e^{at} , so

$$y(t) = Ae^{2t} + Be^t + Ce^{-t} + e^t = \frac{1}{3}e^{2t} - \frac{1}{2}e^t + \frac{1}{6}e^{-t} + e^t.$$

Example 12.10. Solve $y'' - y = 1$, $y(0) = 0$, $y'(0) = 0$.

answer: The rest (zero) initial conditions are nice because they won't add any terms to the algebra. As in the previous example we apply the Laplace transform to the entire equation.

$$s^2Y - Y = \frac{1}{s}, \text{ so } Y = \frac{1}{s(s^2-1)} = \frac{1}{s(s-1)(s+1)} = \frac{A}{s} + \frac{B}{s-1} + \frac{C}{s+1}$$

The coverup method gives $A = -1$, $B = 1/2$, $C = 1/2$. So,

$$y = A + Be^t + Ce^{-t} = -1 + \frac{1}{2}e^t + \frac{1}{2}e^{-t}.$$

12.7 System functions and the Laplace transform

When we introduced the Nyquist criterion for stability we stated without any justification that the system was stable if all the poles of the system function $G(s)$ were in the left half-plane. We also asserted that the poles corresponded to exponential modes of the system. In this section we'll use the Laplace transform to more fully develop these ideas for differential equations.

12.7.1 Lightning review of 18.03

Definitions.

1. $D = \frac{d}{dt}$ is called a **differential operator**. Applied to a function $f(t)$ we have $Df = \frac{df}{dt}$.

We read Df as ‘ D applied to f ’.

Example 12.11. If $f(t) = t^3 + 2$ then $Df = 3t^2$, $D^2f = 6t$.

2. If $P(s)$ is a polynomial then $P(D)$ is called a **polynomial differential operator**.

Example 12.12. Suppose $P(s) = s^2 + 8s + 7$. What is $P(D)$? Compute $P(D)$ applied to $f(t) = t^3 + 2t + 5$. Compute $P(D)$ applied to $g(t) = e^{2t}$.

answer: $P(D) = D^2 + 8D + 7I$. (The I in $7I$ is the identity operator.) To compute $P(D)f$ we do the following.

$$\begin{aligned} f(t) &= t^3 + 2t + 5 \\ Df(t) &= 3t^2 + 2 \\ D^2f(t) &= 6t \end{aligned}$$

Therefore,

$$(D^2 + 8D + 7I)f = 6t + 8(3t^2 + 2) + 7(t^3 + 2t + 5) = 7t^3 + 24t^2 + 20t + 51.$$

$$\begin{aligned} g(t) &= e^{2t} \\ Dg(t) &= 2e^{2t} \\ D^2g(t) &= 4e^{2t} \end{aligned}$$

Therefore,

$$(D^2 + 8D + 7I)g = 4e^{2t} + 8(2)e^{2t} + 7e^{2t} = (4 + 16 + 7)e^{2t} = P(2)e^{2t}.$$

The substitution rule is a straightforward statement about the derivatives of exponentials.

Substitution rule:

$$P(D)e^{st} = P(s)e^{st}. \quad (49)$$

Proof. This is obvious. We ‘prove it’ by example. Let $P(D) = D^2 + 8D + 7I$. Then

$$P(D)e^{at} = a^2e^{at} + 8ae^{at} + 7e^{at} = (a^2 + 8a + 7)e^{at} = P(a)e^{at}.$$

Let's continue to work from this specific example. From it we'll be able to remind you of the general approach to solving constant coefficient differential equations.

Example 12.13. Suppose $P(s) = s^2 + 8s + 7$. Find the exponential modes of the equation $P(D)y = 0$.

answer: The exponential modes are solutions of the form $y(t) = e^{s_0 t}$. Using the substitution rule

$$P(D)e^{s_0 t} = 0 \Leftrightarrow P(s_0) = 0.$$

That is, $y(t) = e^{s_0 t}$ is a mode exactly when s_0 is a root of $P(s)$. The roots of $P(s)$ are $-1, -7$. So the modal solutions are $y_1(t) = e^{-t}$ and $y_2(t) = e^{-7t}$.

Example 12.14. Redo the previous example using the Laplace transform.

answer: For this we solve the differential equation with arbitrary initial conditions:

$$P(D)y = y'' + 8y' + 7y = 0; \quad y(0) = c_1, \quad y'(0) = c_2.$$

Let $Y(s) = \mathcal{L}(y; s)$. Applying the Laplace transform to the equation we get

$$(s^2 Y(s) - sy(0) - y'(0)) + 8(sY(s) - y(0)) + 7Y(s) = 0$$

Algebra:

$$(s^2 + 8s + 7)Y(s) - sc_1 - c_2 - 8c_1 = 0 \Leftrightarrow Y = \frac{sc_1 + 8c_1 + c_2}{s^2 + 8s + 7}$$

Factoring the denominator and using partial fractions, we get

$$Y(s) = \frac{sc_1 + 8c_1 + c_2}{s^2 + 8s + 7} = \frac{sc_1 + 8c_1 + c_2}{(s+1)(s+7)} = \frac{A}{s+1} + \frac{B}{s+7}.$$

We are unconcerned with the exact values of A and B . Taking the Laplace inverse we get

$$y(t) = Ae^{-t} + Be^{-7t}.$$

That is, $y(t)$ is a linear combination of the exponential modes.

You should notice that the denominator in the expression for $Y(s)$ is none other than the characteristic polynomial $P(s)$.

12.7.2 The system function

Example 12.15. With the same $P(s)$ as in Example 12.12 solve the inhomogeneous DE with rest initial conditions: $P(D)y = f(t)$, $y(0) = 0$, $y'(0) = 0$.

answer: Taking the Laplace transform of the equation we get

$$P(s)Y(s) = F(s).$$

Therefore

$$Y(s) = \frac{1}{P(s)} F(s)$$

We can't find $y(t)$ explicitly because $f(t)$ isn't specified.

But, we can make the following definitions and observations. Let $G(s) = 1/P(s)$. If we declare f to be the input and y the output of this linear time invariant system, then $G(s)$ is called the [system function](#). So, we have

$$Y(s) = G(s) \cdot F(s) \quad (50)$$

the formula $Y = G \cdot F$ can be phrased as

$$\text{output} = \text{system function} \times \text{input}.$$

Note well, the roots of $P(s)$ correspond to the exponential modes of the system, i.e. the poles of $G(s)$ correspond to the exponential modes.

The system is called [stable](#) if the modes all decay to 0 as t goes to infinity. That is, if all the poles have negative real part.

Example 12.16. This example is to emphasize that not all system functions are of the form $1/P(s)$. Consider the system modeled by the differential equation

$$P(D)x = Q(D)f,$$

where P and Q are polynomials. Suppose we consider f to be the input and x to be the output. Find the system function.

answer: If we start with rest initial conditions for x and f then the Laplace transform gives $P(s)X(s) = Q(s)F(s)$ or

$$X(s) = \frac{Q(s)}{P(s)} \cdot F(s)$$

Using the formulation

$$\text{output} = \text{system function} \times \text{input},$$

we see that the system function is $G(s) = \frac{Q(s)}{P(s)}$.

Note that when $f(t) = 0$ the differential equation becomes $P(D)x = 0$. So the modes of the system still correspond to the roots of $P(s)$, i.e. the poles of the system function.

Comments. All LTI systems have system functions. They are not all of the form $Q(s)/P(s)$. But, in the s -domain, the output is always the system function times the input. If the system function is meromorphic then the system is stable if all the poles are in the left half-plane.

The system function is also called the [transfer function](#). You can think of it as describing how the system transfers the input to the output.

12.8 Laplace inverse

Up to now we have computed the inverse Laplace transform by table lookup. For example, $\mathcal{L}^{-1}(1/(s - a)) = e^{at}$. To do this properly we should first check that the Laplace transform has an inverse.

We start with the bad news: Unfortunately this is not strictly true. There are many functions with the same Laplace transform. We list some of the ways this can happen.

1. If $f(t) = g(t)$ for $t \geq 0$, then clearly $F(s) = G(s)$. Since the Laplace transform only concerns $t \geq 0$, the functions can differ completely for $t < 0$.

2. Suppose $f(t) = e^{at}$ and

$$g(t) = \begin{cases} f(t) & \text{for } t \neq 1 \\ 0 & \text{for } t = 1. \end{cases}.$$

That is, f and g are the same except we arbitrarily assigned them different values at $t = 1$. Then, since the integrals won't notice the difference at one point, $F(s) = G(s) = 1/(s - a)$. In this sense it is impossible to define $\mathcal{L}^{-1}(F)$ uniquely.

The good news is that the inverse exists as long as we consider two functions that only differ on a negligible set of points the same. In particular, we can make the following claim.

Theorem. Suppose f and g are continuous and $F(s) = G(s)$ for all s with $\operatorname{Re}(s) > a$ for some a . Then $f(t) = g(t)$ for $t \geq 0$.

This theorem can be stated in a way that includes piecewise continuous functions. Such a statement takes more care, which would obscure the basic point that the Laplace transform has a unique inverse up to some, for us, trivial differences.

We start with a few examples that we can compute directly.

Example 12.17. Let $f(t) = e^{at}$. So, $F(s) = \frac{1}{s - a}$. Show

$$f(t) = \sum \operatorname{Res}(F(s)e^{st}) \quad (51)$$

$$f(t) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} F(s)e^{st} ds \quad (52)$$

The sum is over all poles of $e^{st}/(s - a)$. As usual, we only consider $t > 0$.

Here, $c > \operatorname{Re}(a)$ and the integral means the path integral along the vertical line $x = c$.

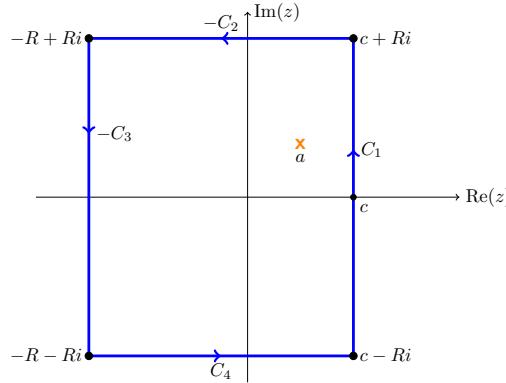
answer: Proving Equation 51 is straightforward: It is clear that $\frac{e^{st}}{s - a}$ has only one pole which is at $s = a$. Since, $\sum \operatorname{Res}\left(\frac{e^{st}}{s - a}, a\right) = e^{at}$ we have proved Equation 51.

Proving Equation 52 is more involved. We should first check the convergence of the integral. In this case, the integral is

$$\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} F(s)e^{st} ds = \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{e^{(c+iy)t}}{c+iy-a} dy = \frac{e^{ct}}{2\pi i} \int_{-\infty}^{\infty} \frac{e^{iyt}}{c+iy-a} dy.$$

The (conditional) convergence of this integral follows using exactly the same argument as in the example near the end of Topic 9 on the Fourier inversion formula for $f(t) = e^{at}$. That is, the integrand is a decaying oscillation, around 0, so its integral is also a decaying oscillation around some limiting value.

Now we use the contour shown below.



We will let R go to infinity and use the following steps to prove Equation 52.

1. The residue theorem guarantees that if the curve is large enough to contain a then

$$\frac{1}{2\pi i} \int_{C_1-C_2-C_3+C_4} \frac{e^{st}}{s-a} ds = \sum \text{Res} \left(\frac{e^{st}}{s-a}, a \right) = e^{at}.$$

2. In a moment we will show that the integrals over C_2 , C_3 , C_4 all go to 0 as $R \rightarrow \infty$.

3. Clearly as R goes to infinity, the integral over C_1 goes to the integral in Equation 52

Putting these steps together we have

$$e^{at} = \lim_{R \rightarrow \infty} \int_{C_1-C_2-C_3+C_4} \frac{e^{st}}{s-a} ds = \int_{c-i\infty}^{c+i\infty} \frac{e^{st}}{s-a} ds.$$

Except for proving the claims in step 2, this proves Equation 52.

To verify step 2 we look at one side at a time.

C_2 : C_2 is parametrized by $s = \gamma(u) = u + iR$, with $-R \leq u \leq c$. So,

$$\left| \int_{C_2} \frac{e^{st}}{s-a} ds \right| = \int_{-R}^c \left| \frac{e^{(u+iR)t}}{u+iR-a} \right| \leq \int_{-R}^c \frac{e^{ut}}{R} du = \frac{e^{ct} - e^{-Rt}}{tR}.$$

Since c and t are fixed, it's clear this goes to 0 as R goes to infinity.

The bottom C_4 is handled in exactly the same manner as the top C_2 .

C_3 : C_3 is parametrized by $s = \gamma(u) = -R + iu$, with $-R \leq u \leq R$. So,

$$\left| \int_{C_3} \frac{e^{st}}{s-a} ds \right| = \int_{-R}^R \left| \frac{e^{(-R+iu)t}}{-R+iu-a} \right| \leq \int_{-R}^R \frac{e^{-Rt}}{R+a} du = \frac{e^{-Rt}}{R+a} \int_{-R}^R du = \frac{2Re^{-Rt}}{R+a}.$$

Since a and $t > 0$ are fixed, it's clear this goes to 0 as R goes to infinity.

Example 12.18. Repeat the previous example with $f(t) = t$ for $t > 0$, $F(s) = 1/s^2$.

We did this in class. It is similar to the previous example. Since F decays like $1/s^2$ we can actually allow $t \geq 0$

Theorem 12.19. Laplace inversion 1. Assume f is continuous and of exponential type a . Then for $c > a$ we have

$$f(t) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} F(s)e^{st} ds. \quad (53)$$

As usual, this formula holds for $t > 0$.

Proof. The proof uses the Fourier inversion formula. We will just accept this theorem for now. Example 12.17 above illustrates the theorem.

Theorem 12.20. **Laplace inversion 2.** Suppose $F(s)$ has a finite number of poles and decays like $1/s$ (or faster). Define

$$f(t) = \sum \text{Res}(F(s)e^{st}, p_k), \text{ where the sum is over all the poles } p_k. \quad (54)$$

Then $\mathcal{L}(f; s) = F(s)$

Proof. **Proof to be added.** The basic ideas are present in the examples above, though it requires a fairly clever choice of contours.

The integral inversion formula in Equation 53 can be viewed as writing $f(t)$ as a ‘sum’ of exponentials. This is extremely useful. For example, for a linear system if we know how the system responds to input $f(t) = e^{at}$, then we know how it responds to any input by writing it as a ‘sum’ of exponentials.

12.9 Delay and feedback.

Let $f(t) = 0$ for $t < 0$. Fix $a > 0$ and let $h(t) = f(t - a)$. So, $h(t)$ is a delayed version of the signal $f(t)$. The Laplace property Equation 48 says $H(s) = e^{-as}F(s)$, where H and F are the Laplace transforms of h and f respectively.

Now, suppose we have a system with system function $G(s)$. (Again, called the open loop system.) As before, can feed the output back to through the system. But, instead of just multiplying the output by a scalar we can delay it also. This is captured by the feedback factor ke^{-as} .

The system function for the closed loop system is

$$G_{CL}(s) = \frac{G}{1 + ke^{-as}G}$$

Note even if you start with a rational function the system function of the closed loop with delay is not rational. Usually it has an infinite number of poles.

Example 12.21. Suppose $G(s) = 1$, $a = 1$ and $k = 1$ find the poles of $G_{CL}(s)$.

answer: $G_{CL}(s) = \frac{1}{1 + e^{-s}}$. So the poles occur where $e^{-s} = -1$, i.e. at $in\pi$, where n is an odd integer. There are an infinite number of poles on the imaginary axis.

Example 12.22. Suppose $G(s) = 1$, $a = 1$ and $k = 1/22$ find the poles of $G_{CL}(s)$. Is the closed loop system stable?

answer: $G_{CL}(s) = \frac{1}{1 + e^{-s}/2}$. So the poles occur where $e^{-s} = -2$, i.e. at $-\log(2) + in\pi$, where n is an odd integer. Since $-\log(2) < 0$, there are an infinite number of poles in the left half-plane. With all poles in the left half-plane, the system is stable.

Example 12.23. Suppose $G(s) = 1$, $a = 1$ and $k = 2$ find the poles of $G_{CL}(s)$. Is the closed loop system stable?

answer: $G_{CL}(s) = \frac{1}{1+2e^{-s}}$. So the poles occur where $e^{-s} = -1/2$, i.e. at $\log(2) + in\pi$, where n is an odd integer. Since $\log(2) > 0$, there are an infinite number of poles in the right half-plane. With poles in the right half-plane, the system is not stable.

Remark. If $\operatorname{Re}(s)$ is large enough we can express the system function $G(s) = \frac{1}{1+ke^{-as}}$ as a geometric series

$$\frac{1}{1+ke^{-as}} = 1 - ke^{-as} + k^2 e^{-2as} - k^3 e^{-3as} + \dots$$

So, for input $F(s)$, we have output

$$X(s) = G(s)F(s) = F(s) - ke^{-as}F(s) + k^2 e^{-2as}F(s) - k^3 e^{-3as}F(s) + \dots$$

Using the shift formula Equation 48, we have

$$x(t) = f(t) - kf(t-a) + k^2 f(t-2a) - k^3 f(t-3a) + \dots$$

(This is not really an infinite series because $f(t) = 0$ for $t < 0$.) If the input is bounded and $k < 1$ then even for large t the series is bounded. So bounded input produces bounded output –this is also what is meant by stability. On the other hand if $k > 1$, then bounded output can lead to unbounded output –this is instability.

12.10 Table of Laplace transforms

Properties and Rules

We assume that $f(t) = 0$ for $t < 0$.

<u>Function</u>	<u>Transform</u>	
$f(t)$	$F(s) = \int_0^\infty f(t)e^{-st} dt$	(Definition)
$a f(t) + b g(t)$	$a F(s) + b G(s)$	(Linearity)
$e^{at} f(t)$	$F(s-a)$	(s -shift)
$f'(t)$	$sF(s) - f(0)$	
$f''(t)$	$s^2 F(s) - sf(0) - f'(0)$	
$f^{(n)}(t)$	$s^n F(s) - s^{n-1} f(0) - \dots - f^{(n-1)}(0)$	
$tf(t)$	$-F'(s)$	
$t^n f(t)$	$(-1)^n F^{(n)}(s)$	
$f(t-a)$	$e^{-as} F(s)$	(t -translation or t -shift)
$\int_0^t f(\tau) d\tau$	$\frac{F(s)}{s}$	(integration rule)
$\frac{f(t)}{t}$	$\int_s^\infty F(\sigma) d\sigma$	

Function Table

<u>Function</u>	<u>Transform</u>	<u>Region of convergence</u>
1	$1/s$	$\operatorname{Re}(s) > 0$
e^{at}	$1/(s - a)$	$\operatorname{Re}(s) > \operatorname{Re}(a)$
t	$1/s^2$	$\operatorname{Re}(s) > 0$
t^n	$n!/(s^{n+1})$	$\operatorname{Re}(s) > 0$
$\cos(\omega t)$	$s/(s^2 + \omega^2)$	$\operatorname{Re}(s) > 0$
$\sin(\omega t)$	$\omega/(s^2 + \omega^2)$	$\operatorname{Re}(s) > 0$
$e^{at} \cos(\omega t)$	$(s - a)/((s - a)^2 + \omega^2)$	$\operatorname{Re}(s) > \operatorname{Re}(a)$
$e^{at} \sin(\omega t)$	$\omega/((s - a)^2 + \omega^2)$	$\operatorname{Re}(s) > \operatorname{Re}(a)$
$\delta(t)$	1	all s
$\delta(t - a)$	e^{-as}	all s
$\cosh(kt) = \frac{e^{kt} + e^{-kt}}{2}$	$s/(s^2 - k^2)$	$\operatorname{Re}(s) > k$
$\sinh(kt) = \frac{e^{kt} - e^{-kt}}{2}$	$k/(s^2 - k^2)$	$\operatorname{Re}(s) > k$
$\frac{1}{2\omega^3}(\sin(\omega t) - \omega t \cos(\omega t))$	$\frac{1}{(s^2 + \omega^2)^2}$	$\operatorname{Re}(s) > 0$
$\frac{t}{2\omega} \sin(\omega t)$	$\frac{s}{(s^2 + \omega^2)^2}$	$\operatorname{Re}(s) > 0$
$\frac{1}{2\omega}(\sin(\omega t) + \omega t \cos(\omega t))$	$\frac{s^2}{(s^2 + \omega^2)^2}$	$\operatorname{Re}(s) > 0$
$t^n e^{at}$	$n!/(s - a)^{n+1}$	$\operatorname{Re}(s) > \operatorname{Re}(a)$
$\frac{1}{\sqrt{\pi t}}$	$\frac{1}{\sqrt{s}}$	$\operatorname{Re}(s) > 0$
t^a	$\frac{\Gamma(a+1)}{s^{a+1}}$	$\operatorname{Re}(s) > 0$

13 Analytic continuation and the Gamma function

13.1 Introduction

In this topic we will look at the Gamma function. This is an important and fascinating function that generalizes factorials from integers to all complex numbers. We look at a few of its many interesting properties. In particular, we will look at its connection to the Laplace transform.

We will start by discussing the notion of analytic continuation. We will see that we have, in

fact, been using this already without any comment. This was a little sloppy mathematically speaking and we will make it more precise here.

13.2 Analytic continuation

Example 13.1. Consider the function

$$F(z) = \int_0^\infty e^{3t} e^{-zt}. \quad (55)$$

We recognize this as the Laplace transform of $f(t) = e^{3t}$ (though we switched the variable from s to z). The integral converges absolutely and F is analytic in the region $A = \{\operatorname{Re}(z) > 3\}$.

Can we extend $F(z)$ to be analytic on a bigger region B ? That is we want a function $\tilde{F}(z)$ that is

1. analytic on B
2. agrees with F on A , i.e. $\tilde{F}(z) = F(z)$ for $z \in A$.

answer: Yes. We know that $F(z) = \frac{1}{z-3}$ –valid for any z in A . So we can define $\tilde{F}(z) = \frac{1}{z-3}$ for any z in $B = \mathbf{C} - \{3\}$.

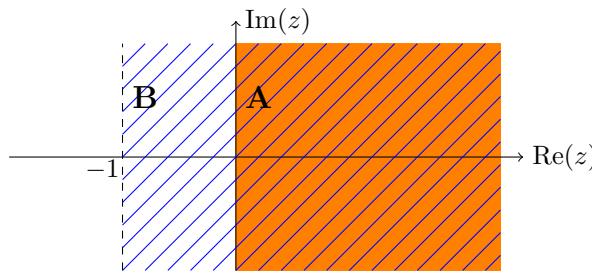
We say that we have analytically continued F on A to \tilde{F} on B .

Note. Usually we don't rename the function. We would just say $F(z)$ defined by Equation 55 can be continued to $F(z) = \frac{1}{z-3}$ on B .

Definition. Suppose $f(z)$ is analytic on a region A . Suppose also that A is contained in a region B . We say that f can be **analytically continued** from A to B if there is a function $\tilde{f}(z)$ such that

1. $\tilde{f}(z)$ is analytic on B .
2. $\tilde{f}(z) = f(z)$ for all z in A .

As noted above, we usually just use the same symbol f for the function on A and its continuation to B .



The region $A = \operatorname{Re}(z) > 0$ is contained in $B = \operatorname{Re}(z) > -1$.

Note. We used analytic continuation implicitly in, for example, the Laplace inversion formula involving residues of $F(s) = \mathcal{L}(f; s)$. Recall that we wrote that for $f(t) = e^{3t}$, $F(s) = \frac{1}{s-3}$ and

$$f(t) = \sum \text{residues of } F.$$

As an integral, $F(s)$ was defined for $\operatorname{Re}(s) > 3$, but the residue formula relies on its analytic continuation to $\mathbf{C} - \{3\}$.

13.2.1 Analytic continuation is unique

Theorem 13.2. Suppose f, g are analytic on a connected region A . If $f = g$ on an open subset of A then $f = g$ on all of A .

Proof. Let $h = f - g$. By hypothesis $h(z) = 0$ on an open set in A . Clearly this means that the zeros of h are not isolated. Back in topic 7 we showed that for analytic h on a connected region A either the zeros are isolated or else h is identically zero on A . Thus, h is identically 0, which implies $f = g$ on A .

Corollary. There is at most one way to analytically continue a function from a region A to a connected region B .

Proof. Two analytic continuations would agree on A and therefore must be the same.

Extension. Since the proof of the theorem uses the fact that zeros are isolated, we actually have the stronger statement: if f and g agree on a nondiscrete subset of A then they are equal. In particular, if f and g are two analytic functions on A and they agree on a line or ray in A then they are equal.

Example 13.3. A little care must be taken. Suppose we define f and g as follows

$$\begin{aligned} f(z) &= \log(z) \text{ with } 0 < \theta < 2\pi \\ g(z) &= \log(z) \text{ with } -\pi < \theta < \pi \end{aligned}$$

Clearly f and g agree on the first quadrant. But we can't use the theorem to conclude that $f = g$ everywhere. The problem is that the regions where they are defined are different. f is defined on \mathbf{C} minus the positive real axis, and g is defined on \mathbf{C} minus the negative real axis. The region where they are both defined is \mathbf{C} minus the real axis, which is not connected.

Because they are both defined on the upper half-plane, we can conclude that they are the same there. (It's easy to see this is true.) But (in this case) being equal in the first quadrant doesn't imply they are the same in the lower half-plane.

13.3 Definition and properties of the Gamma function

Definition. The [Gamma function](#) is defined by the integral formula

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} dt \tag{56}$$

The integral converges absolutely for $\operatorname{Re}(z) > 0$.

Properties

1. $\Gamma(z)$ is defined and analytic in the region $\operatorname{Re}(z) > 0$.
2. $\Gamma(n+1) = n!$, for integers $n \geq 0$.

3. $\Gamma(z+1) = z\Gamma(z)$ ([functional equation](#))

This property and Property 2 characterize the factorial function. Thus, $\Gamma(z)$ generalizes $n!$ to complex numbers z . Some authors will write $\Gamma(z+1) = z!$.

4. $\Gamma(z)$ can be analytically continued to be meromorphic on the entire plane with simple poles at $0, -1, -2 \dots$. The residues are

$$\text{Res}(\Gamma, -m) = \frac{(-1)^m}{m!}$$

5. $\Gamma(z) = \left[z e^{\gamma z} \prod_1^\infty \left(1 + \frac{z}{n}\right) e^{-z/n} \right]^{-1}$, where γ is Euler's constant

$$\gamma = \lim_{n \rightarrow \infty} 1 + \frac{1}{2} + \frac{1}{3} + \dots \frac{1}{n} - \log(n) \approx 0.577$$

This property uses an infinite product. Unfortunately we won't have time, but infinite topics represent an entire topic on their own. Note that the infinite product makes the positions of the poles of Γ clear.

6. $\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin(\pi z)}$

With Property 5 this gives a product formula for $\sin(\pi z)$.

7. $\Gamma(z+1) \approx \sqrt{2\pi} z^{z+1/2} e^{-z}$ for $|z|$ large, $\text{Re}(z) > 0$.

In particular, $n! \approx \sqrt{2\pi} n^{n+1/2} e^{-n}$. ([Stirling's formula](#))

8. $2^{2z-1}\Gamma(z)\Gamma(z+1/2) = \sqrt{\pi}\Gamma(2z)$ ([Legendre duplication formula](#))

Notes. These are just some of the many properties of $\Gamma(z)$. As is often the case, we could have chosen to define $\Gamma(z)$ in terms of some of its properties and derived Equation 56 as a theorem.

We will prove (some of) these properties below.

Example 13.4. Use the properties of Γ to show that $\Gamma(1/2) = \sqrt{\pi}$ and $\Gamma(3/2) = \sqrt{\pi}/2$.

answer: From Property 2 we have $\Gamma(1) = 0! = 1$. The Legendre duplication formula with $z = 1/2$ then shows

$$2^0\Gamma\left(\frac{1}{2}\right)\Gamma(1) = \sqrt{\pi}\Gamma(1) \Rightarrow \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

Now, using the functional equation Property 3 we get

$$\Gamma\left(\frac{3}{2}\right) = \Gamma\left(\frac{1}{2} + 1\right) = \frac{1}{2}\Gamma\left(\frac{1}{2}\right) = \frac{\sqrt{\pi}}{2}.$$

13.4 Connection to Laplace

Claim. For $\text{Re}(z) > 1$ and $\text{Re}(s) > 0$, $\mathcal{L}(t^{z-1}; s) = \frac{\Gamma(z)}{s^z}$.

Proof. By definition $\mathcal{L}(t^{z-1}; s) = \int_0^\infty t^{z-1} e^{-st} dt$. It is clear that if $\operatorname{Re}(z) > 1$, then the integral converges absolutely for $\operatorname{Re}(s) > 0$.

Let's start by assuming that $s > 0$ is real. Use the change of variable $\tau = st$. The Laplace integral becomes

$$\int_0^\infty t^{z-1} e^{-st} dt = \int_0^\infty \left(\frac{\tau}{s}\right)^{z-1} e^{-\tau} \frac{d\tau}{s} = \frac{1}{s^z} \int_0^\infty \tau^{z-1} e^{-\tau} = \frac{\Gamma(z)}{s^z}.$$

This shows that $\mathcal{L}(t^{z-1}; s) = \frac{\Gamma(z)}{s^z}$ for s real and positive. Since both sides of this equation are analytic on $\operatorname{Re}(s) > 0$, the extension to Theorem 13.2 guarantees they are the same.

Corollary. $\Gamma(z) = \mathcal{L}(t^{z-1}; 1)$.

13.5 Proofs of (some) properties of Γ

Property 1. This is clear since the integral converges absolutely for $\operatorname{Re}(z) > 0$.

Property 2. We know (see the Laplace table) $\mathcal{L}(t^n; s) = \frac{n!}{s^{n+1}}$. Setting $s = 1$ and using the corollary to the claim above we get

$$\Gamma(n+1) = \mathcal{L}(t^n; 1) = n!.$$

Property 3. We could do this relatively easily using integration by parts, but let's continue using the Laplace transform. Let $f(t) = t^z$. We know

$$\mathcal{L}(f, s) = \frac{\Gamma(z+1)}{s^{z+1}}$$

Now assume $\operatorname{Re}(z) > 0$, so $f(0) = 0$. Then $f' = zt^{z-1}$ and we can compute $\mathcal{L}(f'; s)$ two ways.

$$\begin{aligned} \mathcal{L}(f'; s) &= \mathcal{L}(zt^{z-1}; s) = \frac{z\Gamma(z)}{s^z} \\ \mathcal{L}(f'; s) &= s\mathcal{L}(t^z; s) = \frac{\Gamma(z+1)}{s^z} \end{aligned}$$

Comparing these two equations we get property 3 for $\operatorname{Re}(z) > 0$.

Property 4. We'll need the following notation for regions in the plane.

$$\begin{aligned} B_0 &= \{\operatorname{Re}(z) > 0\} \\ B_1 &= \{\operatorname{Re}(z) > -1\} - \{0\} \\ B_2 &= \{\operatorname{Re}(z) > -2\} - \{0, -1\} \\ B_n &= \{\operatorname{Re}(z) > -n\} - \{0, -1, \dots, -n+1\} \end{aligned}$$

So far we have $\Gamma(z)$ defined and analytic on B_0 . Our strategy is to use Property 3 to analytically continue Γ from B_0 to B_n . Along the way we will compute the residues at 0 and the negative integers.

Rewrite Property 3 as

$$\Gamma(z) = \frac{\Gamma(z+1)}{z} \quad (57)$$

The right side of this equation is analytic on B_1 . Since it agrees with $\Gamma(z)$ on B_0 it represents an analytic continuation from B_0 to B_1 . We easily compute

$$\text{Res}(\Gamma, 0) = \lim_{z \rightarrow 0} z\Gamma(z) = \Gamma(1) = 1.$$

Similarly, Equation 57 can be expressed as $\Gamma(z+1) = \frac{\Gamma(z+2)}{z+1}$. So,

$$\Gamma(z) = \frac{\Gamma(z+1)}{z} = \frac{\Gamma(z+2)}{(z+1)z} \quad (58)$$

The right side of this equation is analytic on B_2 . Since it agrees with Γ on B_0 it is an analytic continuation to B_2 . The residue at -1 is

$$\text{Res}(\Gamma, -1) = \lim_{z \rightarrow -1} (z+1)\Gamma(z) = \frac{\Gamma(1)}{-1} = -1.$$

We can iterate this procedure as far as we want

$$\Gamma(z) = \frac{\Gamma(z+m+1)}{(z+m)(z+m-1) + \dots + (z+1)z} \quad (59)$$

The right side of this equation is analytic on B_{m+1} . Since it agrees with Γ on B_0 it is an analytic continuation to B_{m+1} . The residue at $-m$ is

$$\text{Res}(\Gamma, -m) = \lim_{z \rightarrow -m} (z+m)\Gamma(z) = \frac{\Gamma(1)}{(-1)(-2)\dots(-m)} = \frac{(-1)^m}{m!}.$$

We'll leave the proofs of Properties 5-8 to another class!

14 Review of multivariable calculus (18.02) constructs

14.1 Introduction

These notes are a terse summary of what we'll need from multivariable calculus. If, after reading these, some parts are still unclear, you should consult your notes or book from your multivariable calculus or ask about it at office hours. We've also posted a more detailed review of line integrals and Green's theorem. You should consult that if needed.

We've seen that complex exponentials make trigonometric functions easier to work with and give insight into many of the properties of trig functions. Similarly, we'll eventually reformulate some material from 18.02 in complex form. We'll see that it's easier to present and the main properties are more transparent in complex form.

14.2 Terminology and notation

Vectors. We'll denote vectors in the plane by (x, y)

Note. In physics and in 18.02 we usually write vectors in the plane as $x\mathbf{i} + y\mathbf{j}$. This use of \mathbf{i} and \mathbf{j} would be confusing in 18.04, so we will write this vector as (x, y) .

In 18.02 you might have used angled brackets $\langle x, y \rangle$ for vectors and round brackets (x, y) for points. In 18.04 we will adopt the more standard mathematical convention and use round brackets for both vectors and points. It shouldn't lead to any confusion.

Orthogonal. Orthogonal is a synonym for perpendicular. Two vectors are orthogonal if their dot product is zero, i.e. $\mathbf{v} = (v_1, v_2)$ and $\mathbf{w} = (w_1, w_2)$ are orthogonal if

$$\mathbf{v} \cdot \mathbf{w} = (v_1, v_2) \cdot (w_1, w_2) = v_1 w_1 + v_2 w_2 = 0.$$

Composition. Composition of functions will be denoted $f(g(z))$ or $f \circ g(z)$, which is read as 'f composed with g'

14.3 Parametrized curves

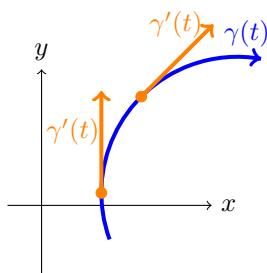
We often use the greek letter gamma for a parametrized curve, i.e.

$$\gamma(t) = (x(t), y(t)).$$

We think of this as a moving point tracing out a curve in the plane. The tangent vector

$$\gamma'(t) = (x'(t), y'(t))$$

is tangent to the curve at the point $(x(t), y(t))$. Its length $|\gamma'(t)|$ is the instantaneous speed of the moving point.



Parametrized curve $\gamma(t)$ with some tangent vectors $\gamma'(t)$.

Example Rev.1. Parametrize the straight line from the point (x_0, y_0) to (x_1, y_1) .

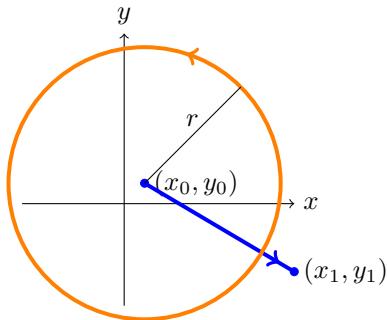
answer: There are always many parametrizations of a given curve. A standard one for straight lines is

$$\gamma(t) = (x, y) = (x_0, y_0) + t(x_1 - x_0, y_1 - y_0), \text{ with } 0 \leq t \leq 1.$$

Example Rev.2. Parametrize the circle of radius r around the point (x_0, y_0) .

answer: Again there are many parametrizations. Here is the standard one with the circle traversed in the counterclockwise direction:

$$\gamma(t) = (x, y) = (x_0, y_0) + r(\cos(t), \sin(t)), \text{ with } 0 \leq t \leq 2\pi.$$



Line from (x_0, y_0) to (x_1, y_1) and circle around (x_0, y_0) .

14.4 Chain rule

For a function $f(x, y)$ and a curve $\gamma(t) = (x(t), y(t))$ the chain rule gives

$$\frac{df(\gamma(t))}{dt} = \frac{\partial f}{\partial x}\Big|_{\gamma(t)} x'(t) + \frac{\partial f}{\partial y}\Big|_{\gamma(t)} y'(t) = \nabla f(\gamma(t)) \cdot \gamma'(t) \text{ dot product of vectors.}$$

Here ∇f is the gradient of f defined in the next section.

14.5 Grad, curl and div

Gradient. For a function $f(x, y)$, the gradient is defined as $\text{grad } f = \nabla f = (f_x, f_y)$. A vector field \mathbf{F} which is the gradient of some function is called a **gradient vector field**.

Curl. For a vector in the plane $\mathbf{F}(x, y) = (M(x, y), N(x, y))$ we define

$$\text{curl } \mathbf{F} = N_x - M_y.$$

Note. **The curl is a scalar.** In 18.02 and in general, the curl of a vector field is another vector field. However, for vectors fields in the plane the curl is always in the \hat{k} direction, so we have simply dropped the \hat{k} and made curl a scalar.

Divergence. The divergence of the vector field $\mathbf{F} = (M, N)$ is

$$\text{div } \mathbf{F} = M_x + N_y.$$

14.6 Level curves

Recall that the level curves of a function $f(x, y)$ are the curves given by $f(x, y) = \text{constant}$.

Recall also that the gradient ∇f is orthogonal to the level curves of f

14.7 Line integrals

The ingredients for line (also called path or contour) integrals are the following:

- A vector field $\mathbf{F} = (M, N)$

- A curve $\gamma(t) = (x(t), y(t))$ defined for $a \leq t \leq b$

Then the line integral of \mathbf{F} along γ is defined by

$$\int_{\gamma} \mathbf{F} \cdot d\mathbf{r} = \int_a^b \mathbf{F}(\gamma(t)) \cdot \gamma'(t) dt = \int_{\gamma} M dx + N dy.$$

Example Rev.3. Let $\mathbf{F} = (-y/r^2, x/r^2)$ and let γ be the unit circle. Compute line integral of \mathbf{F} along γ .

answer: You should be able to supply the answer to this example

14.7.1 Properties of line integrals

1. Independent of parametrization.
2. Reverse direction on curve \Rightarrow change sign. That is,

$$\int_{-C} \mathbf{F} \cdot d\mathbf{r} = - \int_C \mathbf{F} \cdot d\mathbf{r}.$$

(Here, $-C$ means the same curve traversed in the opposite direction.)

3. If C is closed then we sometimes indicate this with the notation $\oint_C \mathbf{F} \cdot d\mathbf{r} = \oint_C M dx + N dy$.

14.7.2 Fundamental theorem for gradient fields

Theorem Rev.4. (Fundamental theorem for gradient fields)

If $\mathbf{F} = \nabla f$ then $\int_{\gamma} \mathbf{F} \cdot d\mathbf{r} = f(P) - f(Q)$, where Q, P are the beginning and endpoints respectively of γ .

Proof. By the chain rule we have

$$\frac{df(\gamma(t))}{dt} = \nabla f(\gamma(t)) \cdot \gamma'(t) = \mathbf{F}(\gamma(t)) \cdot \gamma'(t).$$

The last equality follows from our assumption that $\mathbf{F} = \nabla f$. Now we can this when we compute the line integral:

$$\begin{aligned} \int_{\gamma} \mathbf{F} \cdot d\mathbf{r} &= \int_a^b \mathbf{F}(\gamma(t)) \cdot \gamma'(t) dt \\ &= \int_a^b \frac{df(\gamma(t))}{dt} dt \\ &= f(\gamma(b)) - f(\gamma(a)) \\ &= f(P) - f(Q) \end{aligned}$$

Notice that the third equality follows from the fundamental theorem of calculus.

Definition. If a vector field \mathbf{F} is a gradient field, with $\mathbf{F} = \nabla f$, then we call f a **a potential function** for \mathbf{F} .

Note: the usual physics terminology would be to call $-f$ the potential function for \mathbf{F} .

14.7.3 Path independence and conservative functions

Definition. For a vector field \mathbf{F} , the line integral $\int \mathbf{F} \cdot d\mathbf{r}$ is called **path independent** if, for any two points P and Q , the line integral has the same value for *every* path between P and Q .

Theorem. $\int_C \mathbf{F} \cdot d\mathbf{r}$ is path independent is equivalent to $\oint_C \mathbf{F} \cdot d\mathbf{r} = 0$ for any closed path.

Sketch of proof. Draw two paths from Q to P . Following one from Q to P and the reverse of the other back to P is a closed path. The equivalence follows easily. We refer you to the more detailed review of line integrals and Green's theorem for more details.

Definition. A vector field with path independent line integrals, equivalently a field whose line integrals around any closed loop is 0 is called a **conservative vector field**.

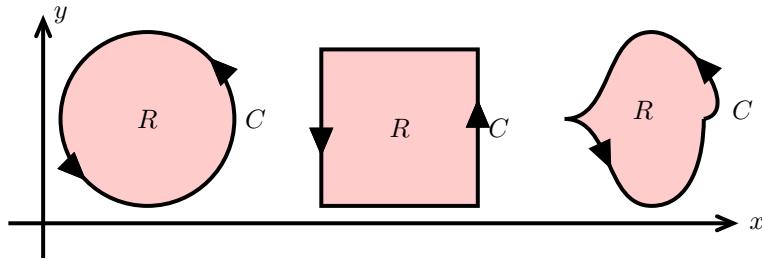
Theorem Rev.5. We have the following equivalence: On a connected region, a gradient field is conservative and a conservative field is a gradient field.

Proof. Again we refer you to the more detailed review for details. Essentially, if \mathbf{F} is conservative then we can define a potential function $f(x, y)$ as the line integral of \mathbf{F} from some base point to (x, y) .

14.8 Green's Theorem

Ingredients: C a *simple* closed curve (i.e. no self-intersection), and R the interior of C .

C must be *positively oriented* (traversed so interior region R is on the left) and *piecewise smooth* (a few corners are okay).



Theorem Rev.6. Green's Theorem: If the vector field $\mathbf{F} = (M, N)$ is defined and differentiable on R then

$$\oint_C M dx + N dy = \iint_R N_x - M_y dA.$$

In vector form this is written

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \iint_R \operatorname{curl} \mathbf{F} dA.$$

where the curl is defined as $\operatorname{curl} \mathbf{F} = (N_x - M_y)$.

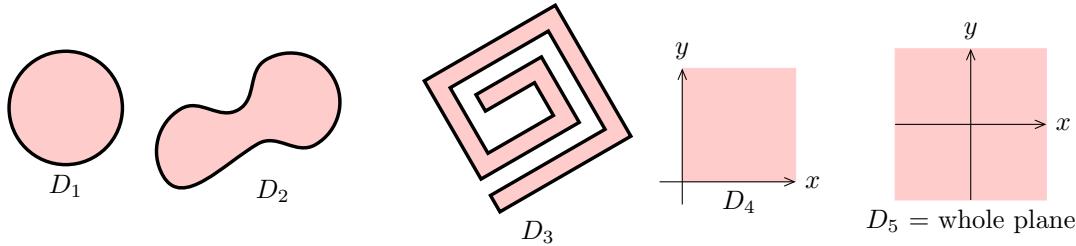
Proof of Green's Theorem. See the more detailed notes on Green's theorem and line integrals for the proof.

14.9 Extensions and applications of Green's theorem

14.9.1 Simply connected regions

Definition: A region D in the plane is **simply connected** if it has “no holes”. Said differently, it is simply connected for every simple closed curve C in D , the interior of C is fully contained in D .

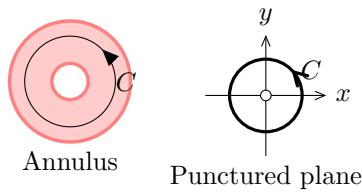
Examples:



D₁-D₅ are simply connected. For any simple closed curve C inside any of these regions the interior of C is entirely inside the region.

Note. Sometimes we say any curve can be shrunk to a point without leaving the region.

The regions below are not simply connected. For each, the interior of the curve C is not entirely in the region.



14.9.2 Potential Theorem

Here is an application of Green's theorem which tells us how to spot a conservative field on a simply connected region. The theorem does not have a standard name, so we choose to call it the Potential Theorem.

Theorem Rev.7. (Potential Theorem) Take $\mathbf{F} = (M, N)$ defined and differentiable on a region D .

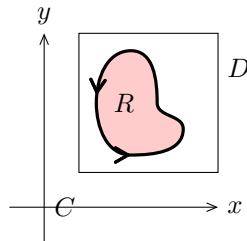
- (a) If $\mathbf{F} = \nabla f$ then $\text{curl}\mathbf{F} = N_x - M_y = 0$.
- (b) If D is simply connected and $\text{curl}\mathbf{F} = 0$ on D , then $\mathbf{F} = \nabla f$ for some f .

We know that on a connected region, being a gradient field is equivalent to being conservative. So we can restate the Potential Theorem as: on a simply connected region, \mathbf{F} is conservative is equivalent to $\text{curl}\mathbf{F} = 0$.

Proof of (a): $\mathbf{F} = (f_x, f_y)$, so $\text{curl}\mathbf{F} = f_{yx} - f_{xy} = 0$.

Proof of (b): Suppose C is a simple closed curve in D . Since D is simply connected the interior of C is also in D . Therefore, using Green's theorem we have,

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \iint_R \operatorname{curl} \mathbf{F} \, dA = 0.$$



This shows that \mathbf{F} is conservative in D . Therefore, by Theorem Rev.5 \mathbf{F} is a gradient field.

Summary: Suppose the vector field $\mathbf{F} = (M, N)$ is defined on a simply connected region D . Then, the following statements are equivalent.

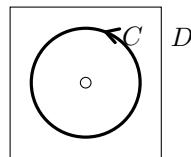
- (1) $\int_P^Q \mathbf{F} \cdot d\mathbf{r}$ is path independent.
- (2) $\oint_C \mathbf{F} \cdot d\mathbf{r} = 0$ for any closed path C .
- (3) $\mathbf{F} = \nabla f$ for some f in D
- (4) \mathbf{F} is conservative in D .

If \mathbf{F} is continuously differentiable then 1,2,3,4 all imply 5:

- (5) $\operatorname{curl} \mathbf{F} = N_x - M_y = 0$ in D

14.9.3 Why we need simply connected in the Potential Theorem

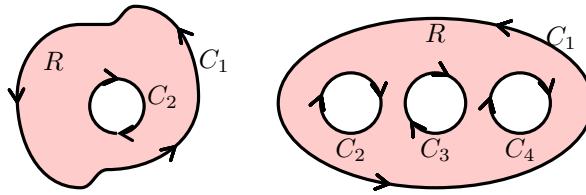
If there is a hole then \mathbf{F} might not be defined on the interior of C . (See the example on the tangential field below.)



14.9.4 Extended Green's Theorem

We can extend Green's theorem to a region R which has multiple boundary curves.

Suppose R is the region between the two simple closed curves C_1 and C_2 .



(Note R is always to the left as you traverse either curve in the direction indicated.)

Then we can extend Green's theorem to this setting by

$$\oint_{C_1} \mathbf{F} \cdot d\mathbf{r} + \oint_{C_2} \mathbf{F} \cdot d\mathbf{r} = \iint_R \operatorname{curl} \mathbf{F} dA.$$

Likewise for more than two curves:

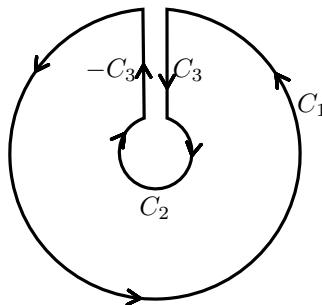
$$\oint_{C_1} \mathbf{F} \cdot d\mathbf{r} + \oint_{C_2} \mathbf{F} \cdot d\mathbf{r} + \oint_{C_3} \mathbf{F} \cdot d\mathbf{r} + \oint_{C_4} \mathbf{F} \cdot d\mathbf{r} = \iint_R \operatorname{curl} \mathbf{F} dA.$$

Proof. The proof is based on the following figure. We ‘cut’ both C_1 and C_2 and connect them by two copies of C_3 , one in each direction. (In the figure we have drawn the two copies of C_3 as separate curves, in reality they are the same curve traversed in opposite directions.)

Now the curve $C = C_1 + C_3 + C_2 - C_3$ is a simple closed curve and Green's theorem holds on it. But the region inside C is exactly R and the contributions of the two copies of C_3 cancel. That is, we have shown that

$$\iint_R \operatorname{curl} \mathbf{F} dA = \int_{C_1+C_3+C_2-C_3} \mathbf{F} \cdot d\mathbf{r} = \int_{C_1+C_2} \mathbf{F} \cdot d\mathbf{r}.$$

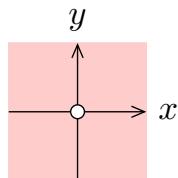
This is exactly Green's theorem, which we wanted to prove.



The punctured plane.

Example Rev.8. Let $\mathbf{F} = \frac{(-y, x)}{r^2}$ (“tangential field”)

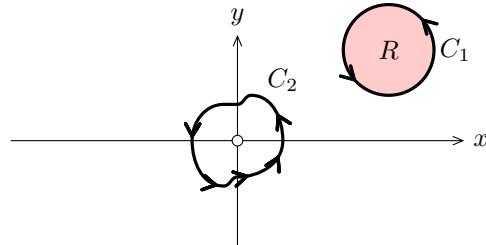
\mathbf{F} is defined on $D = \text{plane} - (0,0) = \text{the punctured plane}$. (Shown below.)



It's easy to compute (we've done it before) that $\operatorname{curl}\mathbf{F} = 0$ in D .

Question: For the tangential field \mathbf{F} what values can $\oint_C \mathbf{F} \cdot d\mathbf{r}$ take for C a simple closed curve (positively oriented)?

answer: We have two cases (i) C_1 not around 0 (ii) C_2 around 0



In case (i) Green's theorem applies because the interior does not contain the problem point at the origin. Thus,

$$\oint_{C_1} \mathbf{F} \cdot d\mathbf{r} = \iint_R \operatorname{curl}\mathbf{F} dA = 0.$$

For case (ii) we will show that $\oint_{C_2} \mathbf{F} \cdot d\mathbf{r} = 2\pi$.

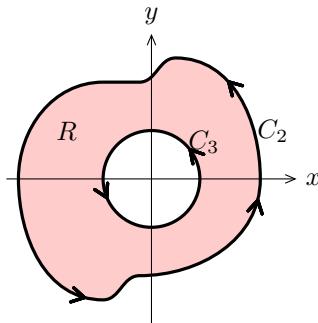
Let C_3 be a small circle of radius a , entirely inside C_2 . By the extended Green's theorem we have

$$\oint_{C_2} \mathbf{F} \cdot d\mathbf{r} - \oint_{C_3} \mathbf{F} \cdot d\mathbf{r} = \iint_R \operatorname{curl}\mathbf{F} dA = 0.$$

Thus, $\oint_{C_2} \mathbf{F} \cdot d\mathbf{r} = \oint_{C_3} \mathbf{F} \cdot d\mathbf{r}$.

Using the usual parametrization of a circle we can easily compute that the line integral is

$$\int_{C_3} \mathbf{F} \cdot d\mathbf{r} = \int_0^{2\pi} 1 dt = 2\pi. \quad QED.$$

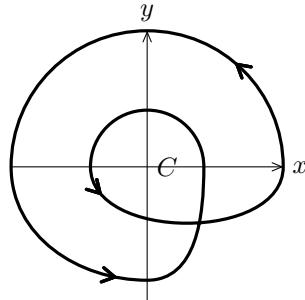


Answer to the question: The only possible values are 0 and 2π .

We can extend this answer in the following way:

If C is not simple, then the possible values of $\oint_C \mathbf{F} \cdot d\mathbf{r}$ are $2\pi n$, where n is the number of times C goes (counterclockwise) around $(0,0)$.

Not for class: n is called the *winding number* of C around 0. n also equals the number of times C crosses the positive x -axis, counting +1 from below and -1 from above.



15 Line integrals and Green's theorem

15.1 Vector Fields

Vector notation. In 18.04 we will mostly use the notation $(v) = (a, b)$ for vectors. The other common notation $(v) = ai + bj$ runs the risk of \mathbf{i} being confused with $i = \sqrt{-1}$ –especially if I forget to make \mathbf{i} boldfaced.

Definition. A **vector field** (also called a **vector-valued function**) is a function $\mathbf{F}(x, y)$ from \mathbf{R}^2 to \mathbf{R}^2 . That is,

$$\mathbf{F}(x, y) = (M(x, y), N(x, y)),$$

where M and N are regular functions on the plane. In standard physics notation

$$\mathbf{F}(x, y) = M(x, y)\mathbf{i} + N(x, y)\mathbf{j} = (M, N).$$

Examples: (a) Force, constant gravitational field $\mathbf{F}(x, y) = (0, -g)$.

(b) Velocity

$$\mathbf{V}(x, y) = \left(\frac{x}{x^2 + y^2}, \frac{y}{x^2 + y^2} \right) = \left(\frac{x}{r^2}, \frac{y}{r^2} \right).$$

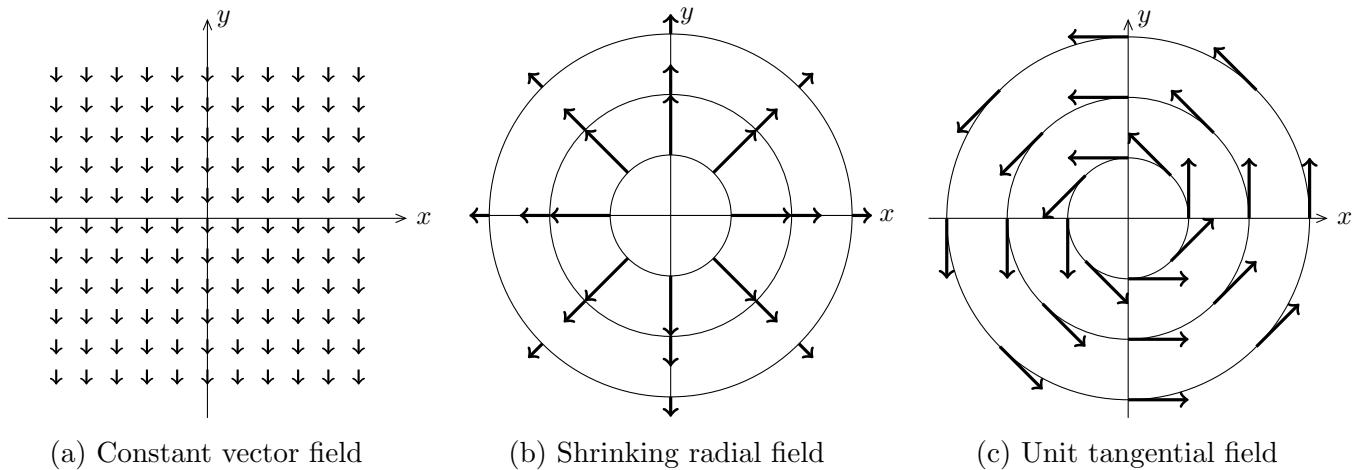
(This is a shrinking radial field –like water pouring from a source at (0,0).)

(c) Unit tangential field $\mathbf{F} = (-y, x)/r$.

(d) Gradient field $\mathbf{F} = \nabla f$, e.g., $f(x, y) = xy^2 \Rightarrow \nabla f = (y^2, 2xy)$.

15.1.1 Visualization of vector fields

Draw little arrows in the plane.

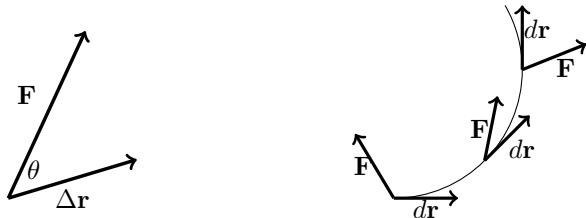


15.1.2 Work done by a force over a curve

Say \mathbf{F} is a constant force, applied over a displacement $\overrightarrow{\Delta r}$

Work = $|\mathbf{F}| \cos \theta |\Delta r| = \mathbf{F} \cdot \overrightarrow{\Delta r}$. (Note $|\mathbf{F}| \cos \theta$ = force in direction of motion = only part of force that does work.)

For a variable force over a curve the total work is found by 'summing' over pieces, i.e., integrating: $\int_C \mathbf{F} \cdot d\mathbf{r}$.



The force vector is at an angle to the curve. This angle varies over the curve.

15.1.3 Grad, curl and div

Gradient. For a function $f(x, y)$: $\text{grad } f = \nabla f = (f_x, f_y)$.

Curl. For a vector in the plane $\mathbf{F}(x, y) = (M(x, y), N(x, y))$ we define

$$\text{curl } \mathbf{F} = N_x - M_y.$$

NOTE. This is a scalar. In general, the curl of a vector field is another vector field. For vectors fields in the plane the curl is always in the \hat{k} direction, so we simply drop the \hat{k} and make curl a scalar.

Div. The divergence of the vector field $\mathbf{F} = (M, N)$ is

$$\text{div } \mathbf{F} = M_x + N_y.$$

This is just for completeness. We probably won't see this in 18.04.

15.2 Line integrals

(Also called [path or contour integrals](#))

Ingredients

Vector field $\mathbf{F}(x, y) = (M, N)$

Curve C : $\mathbf{r}(t) = (x(t), y(t))$

Definition. The [Line integral](#) of \mathbf{F} over C is computed by

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_C (M, N) \cdot (dx, dy) = \int_C M dx + N dy.$$

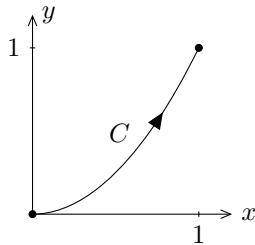
($\mathbf{F} \cdot d\mathbf{r}$ –used in physics, $M dx + N dy$ –used in thermodynamics)

15.2.1 Computing line integrals

Example GT.1. Evaluate $I = \int_C x^2 y dx + (x - 2y) dy$ over the part of the parabola $y = x^2$ from $(0, 0)$ to $(1, 1)$.

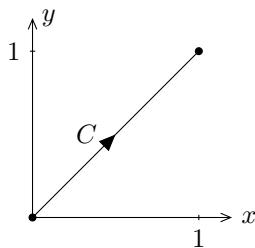
answer: Parametrize curve: $x = t$, $y = t^2$, $0 \leq t \leq 1$. $\Rightarrow dx = dt$, $dy = 2t dt$
 $\Rightarrow I = \int_0^1 t^2(t^2) dt + (t - 2t^2)2t dt = \int_0^1 t^4 + 2t^2 - 4t^3 dt = -\frac{2}{15}$.

Reason for substitution: $\int M dx = \int M \frac{dx}{dt} dt$ etc.



Example GT.2. Same integral as previous example except C is the straight line from $(0, 0)$ to $(1, 1)$.

answer: Parametrize curve: $x = t$, $y = t$, $0 \leq t \leq 1$. $\Rightarrow dx = dt$, $dy = dt$
 $\Rightarrow I = \int_0^1 t^2 \cdot t dt + (t - 2t) dt = \int_0^1 t^3 - t dt = -\frac{1}{4}$.
(Different value from Example GT.1.)



Example GT.3. Same as Example GT.1 except we use a different parametrization of C .

$$x = \sin t, \quad y = \sin^2 t, \quad 0 \leq t \leq \pi/2 \Rightarrow dx = \cos t dt, dy = 2 \sin t \cos t dt.$$

$$\begin{aligned} \Rightarrow I &= \int_0^{\pi/2} \sin^4 t \cos t dt + (\sin t - 2 \sin^2 t) 2 \sin t \cos t dt \\ &= \int_0^{\pi/2} (\sin^4 t + 2 \sin^2 t - 4 \sin^3 t) \cos t dt = \frac{\sin^5 t}{5} + \frac{2}{3} \sin^3 t - \sin^4 t \Big|_0^{\pi/2} = -\frac{2}{15}. \end{aligned}$$

This is the same value as in Example GT.1. That is, **the parametrization doesn't affect the value of the line integral.**

15.2.2 Properties of line integrals

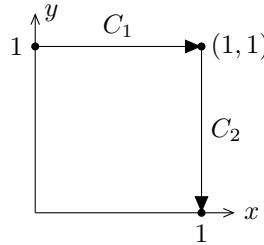
1. Independent of parametrization.
2. Reverse direction on curve \Rightarrow change sign. That is,

$$\int_{-C} \mathbf{F} \cdot d\mathbf{r} = - \int_C \mathbf{F} \cdot d\mathbf{r}.$$

(Here, $-C$ means the same curve traversed in the opposite direction.)

3. If C is closed we use the notation $\oint_C \mathbf{F} \cdot d\mathbf{r} = \oint_C M dx + N dy$.

Example GT.4. Evaluate $I = \int_C y dx + (x + 2y) dy$ where $C = C_1 + C_2$ is the curve shown.



answer: Have 2 pieces so $I = \int_{C_1} y dx + (x + 2y) dy + \int_{C_2} y dx + (x + 2y) dy$.

Note, we don't always need to introduce the variable t . In this case we can easily use x or y as parameters.

$C_1 : y = 1$, use x as parameter. $0 \leq x \leq 1 \Rightarrow dx = dx, dy = 0$. So,

$$\int_{C_1} y dx + (2 + 2y) dy = \int_0^1 dx = 1.$$

$C_2 : x = 1$, use y as parameter. y goes from 1 to 0. So,

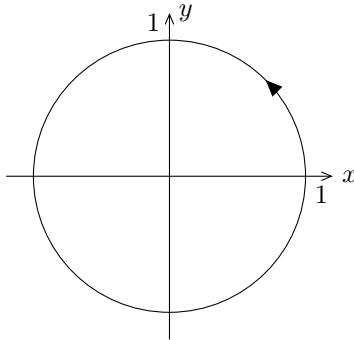
$$\int_{C_2} y dx + (1 + 2y) dy = \int_1^0 (1 + 2y) dy = - \int_0^1 1 + 2y dy = -2.$$

answer: $I = 1 - 2 = -1$.

Example GT.5. Evaluate $I = \oint_C -y \, dx + x \, dy$ where C is the unit circle traversed in a counterclockwise (CCW) direction.

answer: Parametrization: $x = \cos t$, $y = \sin t$, $0 \leq t \leq 2\pi$. So, $dx = \cos t \, dt$, $dy = -\sin t \, dt$. We get

$$I = \int_0^{2\pi} -\sin t(-\sin t) \, dt + \cos t(\cos t) \, dt = \int_0^{2\pi} dt = 2\pi.$$



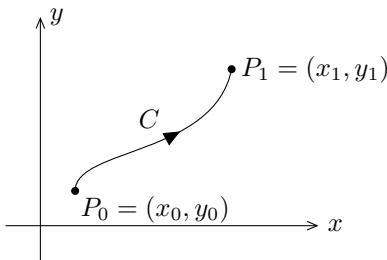
15.3 Gradient and conservative fields

15.3.1 The fundamental theorem for gradient fields

Theorem GT.6. (Fundamental theorem for gradient fields) Suppose that $\mathbf{F} = \nabla f$ is a gradient field. Then,

$$\boxed{\int_C \mathbf{F} \cdot d\mathbf{r} = f(x, y)|_{P_0}^{P_1} = f(P_1) - f(P_0) = f(x_1, y_1) - f(x_0, y_0)}$$

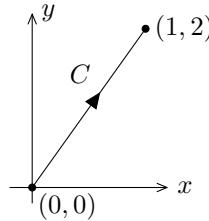
where $P_0 = (x_0, y_0)$ and $P_1 = (x_1, y_1)$ are the endpoints of the curve C . That is, for *gradient fields* the line integral is independent of the path taken.



The proof of the fundamental theorem is given after the next example

Example GT.7. Let $f(x, y) = xy^3 + x^2$. So, $\mathbf{F} = \nabla f = (y^3 + 2x, 3xy^2)$.

Let C be the curve shown and compute $I = \int_C \mathbf{F} \cdot d\mathbf{r}$. Do this both directly and using the fundamental theorem.



answer: Method 1: parametrize C : $x = x$, $y = 2x$, $0 \leq x \leq 1$. Then,

$$I = \int_C (y^3 + 2x) dx + 3xy^2 dy = \int_0^1 (8x^3 + 2x) dx + 12x^3 \cdot 2 dx = \int_0^1 32x^3 + 2x dx = 9.$$

Method 2: $\int_C \nabla f \cdot d\mathbf{r} = f(1, 2) - f(0, 0) = 9$.

15.3.2 Proof of the fundamental theorem

This just requires some algebra and the chain rule:

$$\begin{aligned} \int_C \nabla f \cdot d\mathbf{r} &= \int_C f_x dx + f_y dy = \int_{t_0}^{t_1} \left[f_x(x(t), y(t)) \frac{dx}{dt} + f_y(x(t), y(t)) \frac{dy}{dt} \right] dt \\ &= \int_{t_0}^{t_1} \frac{d}{dt} f(x(t), y(t)) dt = f(x(t), y(t))|_{t_0}^{t_1} = f(P_1) - f(P_0) \quad \blacksquare \end{aligned}$$

15.3.3 Path independence

Definition. For a vector field \mathbf{F} , the line integral $\int \mathbf{F} \cdot d\mathbf{r}$ is called **path independent** if, for any two points P and Q , the line integral has the same value for *every* path between P and Q .

Theorem GT.8. $\int_C \mathbf{F} \cdot d\mathbf{r}$ is path independent is equivalent to $\oint_C \mathbf{F} \cdot d\mathbf{r} = 0$ for any closed path.

Note. These statements are technically a little sloppy. They should all include the qualifying phrase ‘in a region A ’ for \mathbf{F} , the points and the paths.

Proof. We have to show two things:

- (i) Path independence implies the line integral around any closed path is 0.
- (ii) If the line integral around all closed paths is 0 then we have path independence.

To see (i), assume path independence and consider the closed path C shown in figure (i) below. Since the starting point P_0 is the same as the endpoint P_1 the line integral $\oint_C \mathbf{F} \cdot d\mathbf{r}$ must have the same value as the line integral over the curve consisting of the single point P_0 . Since that is clearly 0 we must have the integral over C is 0.

To see (ii), assume $\oint_C \mathbf{F} \cdot d\mathbf{r} = 0$ for any closed curve. Consider the two curves C_1 and C_2 shown in figure (ii). Both start at P_0 and end at P_1 . By the assumption that integrals

around closed paths are 0, we have $\oint_{C_1-C_2} \mathbf{F} \cdot d\mathbf{r} = 0$. So,

$$\int_{C_1} \mathbf{F} \cdot d\mathbf{r} = \int_{C_2} \mathbf{F} \cdot d\mathbf{r}.$$

That is, any two paths from P_0 to P_1 have the same line integral. We have shown that the line integrals are path independent.

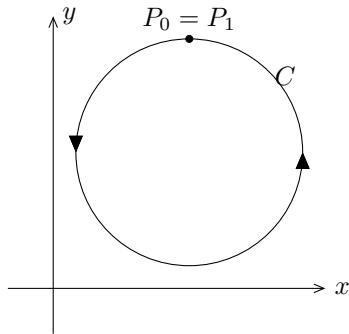


Figure (i)

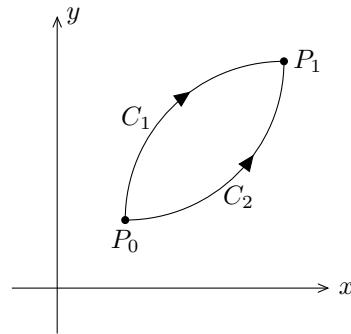


Figure (ii)

15.3.4 Conservative vector fields

Given a vector field \mathbf{F} , Theorem GT.8 in the previous section said that the line integrals of \mathbf{F} were path independent if and only if the line integral of \mathbf{F} around any closed path is 0. We call such a vector field a **conservative vector field**.

The fundamental theorem says that if \mathbf{F} is a gradient field: $\mathbf{F} = \nabla f$, then the line integral $\int_C \mathbf{F} \cdot d\mathbf{r}$ is path independent. That is,

a gradient field is conservative.

The following theorem says that the converse is also true on connected regions. By a **connected region** we mean that any two points in the region can be connected by a continuous path that lies entirely in the region.

Theorem GT.9. A conservative field on a connected region is a gradient field.

Proof. Call the region D . We have to show that if we have a conservative field $\mathbf{F} = (M, N)$ on D there is a potential function f with $\mathbf{F} = \nabla f$.

The easy part will be defining f . The trickier part will be showing that its gradient is \mathbf{F} . So, first we define f : Fix a point (x_0, y_0) in D . Then for any point (x, y) in D we define

$$f(x, y) = \int_{\gamma} \mathbf{F} \cdot d\mathbf{r}, \text{ where } \gamma \text{ is any path from } (x_0, y_0) \text{ to } (x, y).$$

Path independence guarantees that $f(x, y)$ is well defined, i.e. it doesn't depend on the choice of path.

Now we need to show that $\mathbf{F} = \nabla f$, i.e. if $\mathbf{F} = (M, N)$, then we need to show that $f_x = M$ and $f_y = N$. We'll show the first case, the case $f_y = N$ is essentially the same. First, note

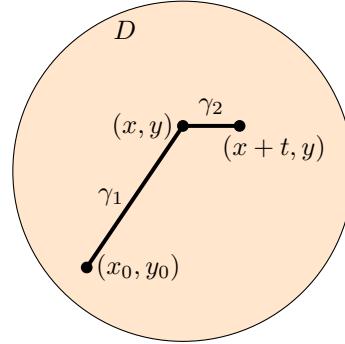
that by definition

$$f_x(x, y) = \frac{df(x + t, y)}{dt} \Big|_{t=0}.$$

This means we need to write down the line integral for $f(x + t, y)$.

In the figure below, $f(x, y)$ is the integral along the path γ_1 from (x_0, y_0) to (x, y) and $f(x + t, y)$ is the integral along γ_1 followed by the horizontal line γ_2 to $(x + t, y)$. This means that

$$f(x + t, y) = \int_{\gamma_1} \mathbf{F} \cdot d\mathbf{r} + \int_{\gamma_2} \mathbf{F} \cdot d\mathbf{r} = f(x, y) + \int_{\gamma_2} \mathbf{F} \cdot d\mathbf{r}.$$



We need to parameterize the horizontal segment γ_2 . Notationwise, the letters x , y and t are already taken, so we let $\gamma_2(s) = (u(s), v(s))$, where

$$u(s) = x + s, \quad v(s) = y; \quad \text{with } 0 \leq s \leq t$$

On this segment $du = ds$ and $dy = 0$. So,

$$f(x + t, y) = f(x, y) + \int_0^t M(x + s, y) ds$$

The piece $f(x, y)$ is constant as t varies, so the fundamental theorem of calculus says that

$$\frac{df(x + t, y)}{dt} = M(x + t, y).$$

Setting $t = 0$ we have

$$f_x(x, y) = \frac{df(x + t, y)}{dt} \Big|_{t=0} = M(x, y).$$

This is exactly what we needed to show! We summarize in a box:

On a connected region a conservative field is a gradient field.

Example GT.10. If \mathbf{F} is the electric field of an electric charge it is conservative.

Example GT.11. A gravitational field of a mass is conservative.

15.4 Potential functions

Definition. If \mathbf{F} is a gradient field with $\mathbf{F} = \nabla f$, then we call f a **potential function for \mathbf{F}** .

Note. The usual physics terminology would be to call $-f$ the potential function for \mathbf{F} .

We have just seen that if $\mathbf{F} = \nabla f$ (a gradient field) then \mathbf{F} is conservative. That is,

$\int_{P_0}^{P_1} \mathbf{F} \cdot d\mathbf{r}$, is independent of the path from P_0 to P_1 and, $\oint_C \mathbf{F} \cdot d\mathbf{r} = 0$, where C is any closed path.

Questions: 1. How can we tell if a vector field \mathbf{F} is a gradient field?

2. If \mathbf{F} is a gradient field, how do we find a potential function f .

We start to answer these questions in the next theorem.

Theorem GT.12. Take $\mathbf{F} = (M, N)$

(a) If $\mathbf{F} = \nabla f$ then $\text{curl}\mathbf{F} = N_x - M_y = 0$.

(b) If \mathbf{F} is differentiable and $\text{curl}\mathbf{F} = 0$ in the *whole* plane then $\mathbf{F} = \nabla f$ for some f , i.e. \mathbf{F} is conservative.

Notes. The restriction that \mathbf{F} is defined on the whole plane is too stringent for our needs. Below, we will give what we call the Potential Theorem, which only requires that \mathbf{F} be defined and differentiable on what's called a simply connected region.

Proof of (a): If $\mathbf{F} = (M, N) = \nabla f$ then we have $M = f_x$ and $N = f_y$. Thus,

$$M_y = f_{xy}, \quad \text{and} \quad N_x = f_{yx}.$$

This proves (a) (provided f has continuous second partials).

The proof of (b) will be postponed until after we have proved Green's theorem and we can state the Potential Theorem. The examples below will show how to find f

Example GT.13. For which values of the constants a and b will $\mathbf{F} = (axy, x^2 + by)$ be a gradient field?

answer: $M_y = ax$, $N_x = 2x \Rightarrow a = 2$ and b is arbitrary.

Example GT.14. Is $\mathbf{F} = ((3x^2 + y), e^x)$ conservative?

answer: $M_y = 1$, $N_x = e^x$. Since $M_y \neq N_x$, \mathbf{F} is not conservative.

Example GT.15. Is $\frac{(-y, x)}{x^2 + y^2}$ conservative?

answer: NO! The reasoning is a little trickier than in the previous example. First,

$$M_y = \frac{y^2 - x^2}{(x^2 + y^2)^2} = N_x.$$

BUT, since the field is not defined for all (x, y) Theorem GT.12b does not apply.

The answer turns out to be no, \mathbf{F} is not conservative. We can see that as follows. Consider C = unit circle parametrized by $x = \cos t$, $y = \sin t$. Computing the line integral directly we have

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \oint_C -\frac{y}{x^2 + y^2} dx + \frac{x}{x^2 + y^2} dy = \int_0^{2\pi} dt = 2\pi.$$

Since this is *not* 0 the field is not conservative.

Example GT.16. Is $\mathbf{F} = \frac{(x, y)}{x^2 + y^2}$ conservative?

answer: Again we find $N_x = M_y$, BUT since \mathbf{F} is not defined at $(0, 0)$ Theorem GT.12b does not apply. HOWEVER, it turns out that

$$\mathbf{F} = \nabla \ln(\sqrt{x^2 + y^2}) = \nabla \ln r$$

Since we found a potential function we have shown directly that \mathbf{F} is conservative on the region consisting of the plane minus the origin.

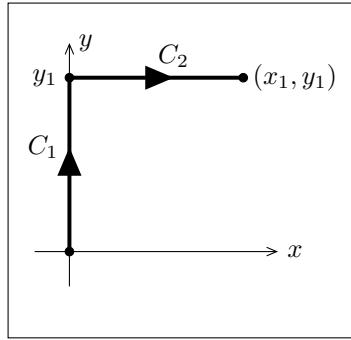
15.4.1 Finding the potential function

Example GT.17. Show $\mathbf{F} = (3x^2 + 6xy, 3x^2 + 6y)$ is conservative and find the potential function f such that $\mathbf{F} = \nabla f$.

answer: First, $M_y = 6x = N_x$. Since \mathbf{F} is defined for all (x, y) Theorem GT.12 implies \mathbf{F} is conservative.

Method 1 (for finding f): Since $\int_C \mathbf{F} \cdot d\mathbf{r} = f(P_1) - f(P_0)$ we have $f(P_1) = f(P_0) + \int_C \mathbf{F} \cdot d\mathbf{r}$ for any path from P_0 to P_1 .

We let $P_1 = (x_1, y_1)$ be an arbitrary point. We can fix P_0 and C any way we want. For this problem take $P_0 = (0, 0)$ and $C = C_1 + C_2$ as the path shown.



$C_1 : x = 0, y = y$, so $dx = 0, dy = dy$,

$C_2 : x = x, y = y_1$, so $dx = dx, dy = 0$

$$\begin{aligned} f(x_1, y_1) - f(0, 0) &= \int_C \mathbf{F} \cdot d\mathbf{r} = \int_C M dx + N dy \\ &= \int_0^{x_1} M(x, 0) dx + \int_0^{y_1} N(x_1, y) dy \\ &= \int_0^{y_1} 6y dy + \int_0^{x_1} 3x^2 + 6y_1 dx \\ &= 3y_1^2 + x_1^3 + 3x_1^2 y_1 \end{aligned}$$

Thus,

$$f(x_1, y_1) - f(0, 0) = 3y_1^2 + x_1^3 + 3x_1^2 y_1 = 3y_1^2 + x_1^3 + 3x_1^2 y_1.$$

Switching from (x_1, y_1) to (x, y) and letting $f(0, 0) = C$, we have

$$f(x, y) = 3y^2 + x^3 + 3x^2 y + C.$$

Method 2: We start with $f_x = 3x^2 + 6xy$.

Integrating with respect to x gives

$$f(x, y) = x^3 + 3x^2y + g(y).$$

We need to add $g(y)$ as the ‘constant’ of integration since y is a constant when differentiating with respect to x .

Now integrating $f_y = 3x^2 + 6y$ with respect to y gives

$$f(x, y) = 3x^2y + 3y^2 + h(x)$$

Comparing the two expressions for f we see $g(y) = 3y^2 + C$ and $h(x) = x^3 + C$. So,

$$f(x, y) = x^3 + 3x^2y + 3y^2 + C.$$

(The same as method 1.)

In general I prefer method 1.

Example GT.18. Let $\mathbf{F} = ((x + y^2), (2xy + 3y^2))$. Show that \mathbf{F} is a gradient field and find the potential function using both methods.

answer: Testing the partials we have: $M_y = 2y = N_x$, \mathbf{F} defined on all (x, y) . Thus, by Theorem GT.12, \mathbf{F} is conservative.

Method 1: Use the path shown.

$$f(P_1) - f(0, 0) = \int_C \mathbf{F} \cdot d\mathbf{r} = \int_C M dx + N dy.$$

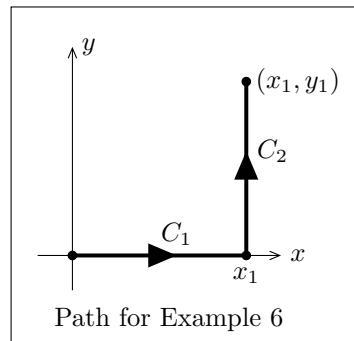
$$C_1 : x = x, y = 0, \Rightarrow dx = dx, dy = 0 \Rightarrow M(x, 0) = x.$$

$$C_2 : x = x_1, y = y, \Rightarrow dx = 0, dy = dy \Rightarrow N(x_1, y) = 2x_1y + 3y^2.$$

$$\int_C \mathbf{F} \cdot d\mathbf{r} = \int_0^{x_1} x dx + \int_0^{y_1} 2x_1y + 3y^2 dy = x_1^2/2 + x_1y_1^2 + y_1^3.$$

Thus, $f(x_1, y_1) - f(0, 0) = x_1^2/2 + x_1y_1^2 + y_1^3$. Equivalently

$$f(x, y) = x^2/2 + xy^2 + y^3 + C.$$



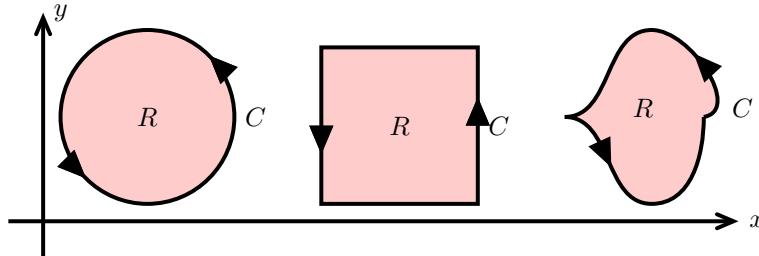
Method 2: $f_x = x + y^2$, so $f = x^2/2 + xy^2 + g(y)$. $f_y = 2xy + 3y^2$, so $f = xy^2 + y^3 + h(x)$. Comparing the two expressions for f we get

$$f(x, y) = x^2/2 + xy^2 + y^3 + C.$$

15.5 Green's Theorem

Ingredients: C a *simple* closed curve (i.e. no self-intersection), and R the interior of C .

C must be *positively oriented* (traversed so interior region R is on the left) and *piecewise smooth* (a few corners are okay).



Theorem GT.19. **Green's Theorem:** If the vector field $\mathbf{F} = (M, N)$ is defined and differentiable on R then

$$\oint_C M \, dx + N \, dy = \iint_R N_x - M_y \, dA. \quad (60)$$

In vector form this is written

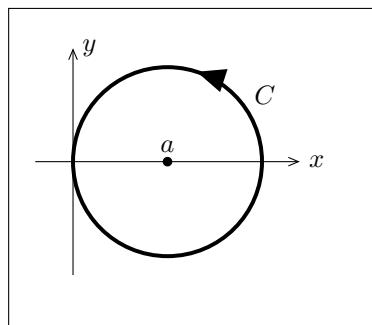
$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \iint_R \operatorname{curl} \mathbf{F} \, dA.$$

where the curl is defined as $\operatorname{curl} \mathbf{F} = (N_x - M_y)$.

Example GT.20. Use the right hand side of Equation 60 to find the left hand side. Use Green's Theorem to compute

$$I = \oint_C 3x^2y^2 \, dx + 2x^2(1+xy) \, dy$$

where C is the circle shown.



answer: By Green's Theorem

$$I = \iint_R 6x^2y + 4x - 6x^2y \, dA = 4 \iint_R x \, dA.$$

We could compute this directly, but here's a trick. We know the x center of mass is

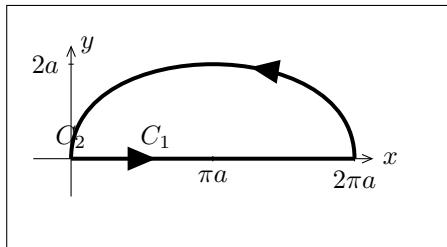
$$x_{cm} = \frac{1}{A} \iint_R x \, dA = a.$$

Thus $\iint_R x \, dA = \pi a^3$, so $I = 4\pi a^3$.

Example GT.21. (Use the left hand side of Equation 60 to find the right hand side.)

Use Green's Theorem to find the area under one arch of the cycloid.

answer: The cycloid is $x = a(\theta - \sin(\theta))$, $y = a(1 - \cos(\theta))$. The picture shows the curve $C = C_1 + C_2$ surrounding the area under one arch.



By Green's Theorem, $\oint_C -y \, dx = \iint_R dA = \text{area}$. So,

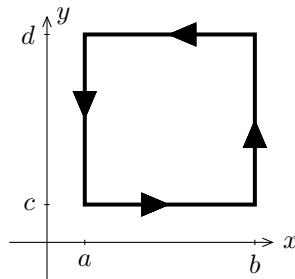
$$\begin{aligned} \text{area} &= \oint_{C_1+C_2} -y \, dx = \int_{C_1} 0 \cdot dx + \int_{C_2} -y \, dx \\ \text{area} &= \int_0^{2\pi} a^2(1 - \cos(\theta))^2 \, d\theta = 3\pi a^2. \end{aligned}$$

Other ways to compute area:

$$\iint_R dA = \oint_C -y \, dx = \oint_C x \, dy = \frac{1}{2} \oint_C -y \, dx + x \, dy.$$

15.5.1 'Proof' of Green's Theorem

(i) First we'll work on the rectangle shown. Later we'll use a lot of rectangles to approximate an arbitrary region.



(ii) We'll simplify a bit and assume $N = 0$. The proof when $N \neq 0$ is essentially the same with a bit more writing.

By direct calculation the right hand side of Green's Theorem (Equation 60) is

$$\iint_R -\frac{\partial M}{\partial y} \, dA = \int_a^b \int_c^d -\frac{\partial M}{\partial y} \, dy \, dx.$$

Inner integral:

$$-M(x, y)|_c^d = -M(x, d) + M(x, c)$$

Putting this into the outer integral we have shown that

$$\iint_R -\frac{\partial M}{\partial y} dA = \int_a^b M(x, c) - M(x, d) dx. \quad (61)$$

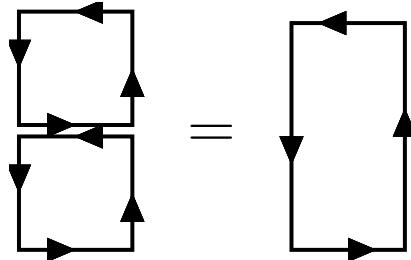
For the left hand side of Equation 60 we have

$$\begin{aligned} \oint_C M dx &= \int_{bottom} M dx + \int_{top} M dx \quad (\text{since } dx = 0 \text{ along the sides}) \\ &= \int_a^b M(x, c) dx + \int_b^a M(x, d) dx = \int_a^b M(x, c) - M(x, d) dx. \end{aligned} \quad (62)$$

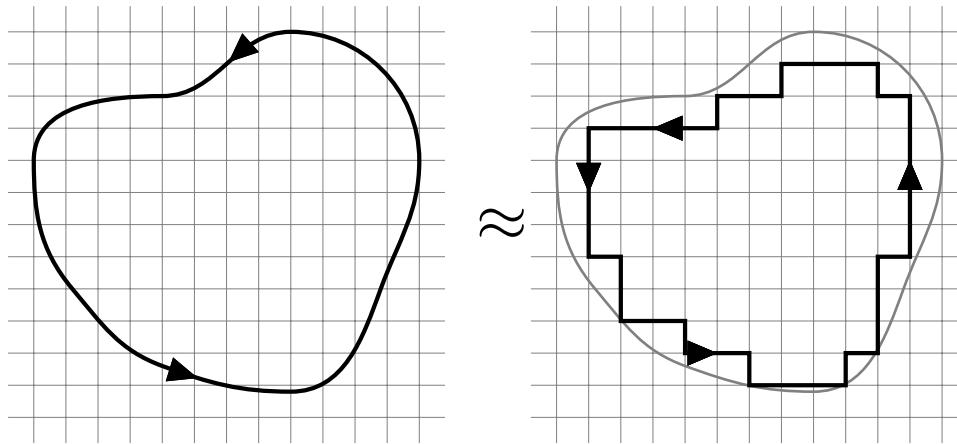
Comparing Equations 61 and 62 we find that we have proved Green's Theorem.

Next we'll use rectangles to build up an arbitrary region. We start by stacking two rectangles on top of each other.

For line integrals when adding two rectangles with a common edge the common edges are traversed in opposite directions so the sum of the line integrals of the two rectangles equals the line integral over the outside boundary.



Similarly when adding a lot of rectangles: everything cancels except the outside boundary. This extends Green's Theorem on a rectangle to Green's Theorem on a sum of rectangles. Since any region can be approximated as close as we want by a sum of rectangles, Green's Theorem must hold on arbitrary regions.

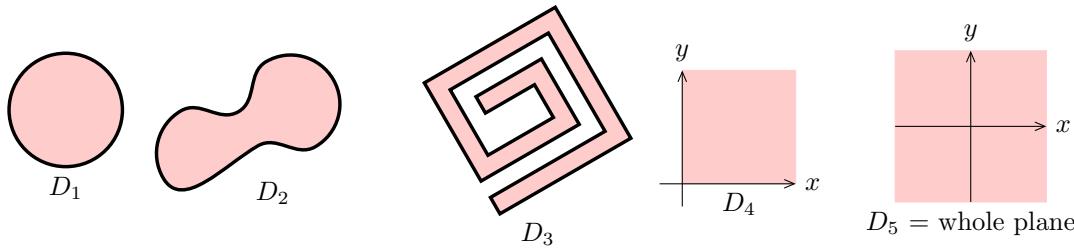


15.6 Extensions and applications of Green's theorem

15.6.1 Simply connected regions

Definition. A region D in the plane is [simply connected](#) if it has “no holes”. Said differently, it is simply connected for every simple closed curve C in D , the interior of C is fully contained in D .

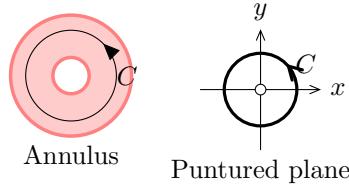
Examples:



D₁-D₅ are simply connected. For any simple closed curve C inside any of these regions the interior of C is entirely inside the region.

Note. Sometimes we say the region is simply connected if any curve can be shrunk to a point without leaving the region.

The regions below are not simply connected. For each, the interior of the curve C is not entirely in the region.



15.6.2 Potential Theorem

Here is an application of Green's theorem which tells us how to spot a conservative field on a simply connected region. The theorem does not have a standard name, so we choose to call it the Potential Theorem. You should check that it is largely a restatement of Theorem GT.12 above.

Theorem GT.22. (Potential Theorem) Take $\mathbf{F} = (M, N)$ defined and differentiable on a region D .

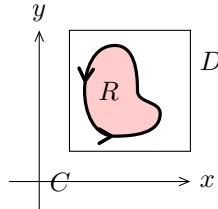
- (a) If $\mathbf{F} = \nabla f$ then $\text{curl}\mathbf{F} = N_x - M_y = 0$.
- (b) If D is simply connected and $\text{curl}\mathbf{F} = 0$ on D , then $\mathbf{F} = \nabla f$ for some f .

We know that on a connected region, being a gradient field is equivalent to being conservative. So we can restate the Potential Theorem as: on a simply connected region, \mathbf{F} is conservative if and only if $\text{curl}\mathbf{F} = 0$.

Proof of (a): This was proved in Theorem GT.12.

Proof of (b): Suppose C is a simple closed curve in D . Since D is simply connected the interior of C is also in D . Therefore, using Green's theorem we have,

$$\oint_C \mathbf{F} \cdot d\mathbf{r} = \iint_D \operatorname{curl} \mathbf{F} \, dA = 0.$$



This shows that \mathbf{F} is conservative in D . Therefore by Theorem GT.9 \mathbf{F} is a gradient field.

Summary: Suppose the vector field $\mathbf{F} = (M, N)$ is defined on a simply connected region D . Then, the following statements are equivalent.

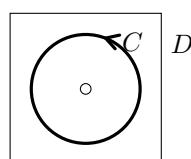
- (1) $\int_P^Q \mathbf{F} \cdot d\mathbf{r}$ is path independent.
- (2) $\oint_C \mathbf{F} \cdot d\mathbf{r} = 0$ for any closed path C .
- (3) $\mathbf{F} = \nabla f$ for some f in D
- (4) \mathbf{F} is conservative in D .

If \mathbf{F} is continuously differentiable then 1,2,3,4 all imply 5:

- (5) $\operatorname{curl} \mathbf{F} = N_x - M_y = 0$ in D

15.6.3 Why we need simply connected in the Potential Theorem

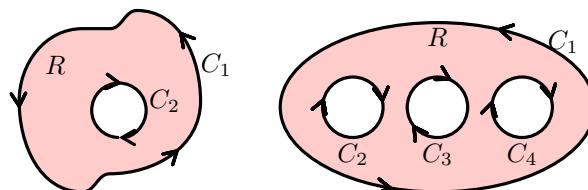
If there is a hole then \mathbf{F} might not be defined on the interior of C . (See the example on the tangential field below.)



15.6.4 Extended Green's Theorem

We can extend Green's theorem to a region R which has multiple boundary curves.

Suppose R is the region between the two simple closed curves C_1 and C_2 .



(Note R is always to the left as you traverse either curve in the direction indicated.)

Then we can extend Green's theorem to this setting by

$$\oint_{C_1} \mathbf{F} \cdot d\mathbf{r} + \oint_{C_2} \mathbf{F} \cdot d\mathbf{r} = \iint_R \operatorname{curl} \mathbf{F} dA.$$

Likewise for more than two curves:

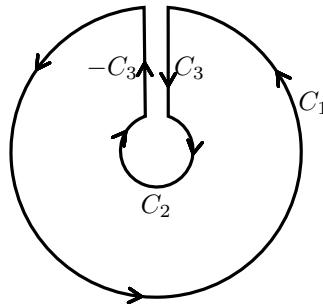
$$\oint_{C_1} \mathbf{F} \cdot d\mathbf{r} + \oint_{C_2} \mathbf{F} \cdot d\mathbf{r} + \oint_{C_3} \mathbf{F} \cdot d\mathbf{r} + \oint_{C_4} \mathbf{F} \cdot d\mathbf{r} = \iint_R \operatorname{curl} \mathbf{F} dA.$$

Proof. The proof is based on the following figure. We 'cut' both C_1 and C_2 and connect them by two copies of C_3 , one in each direction. (In the figure we have drawn the two copies of C_3 as separate curves, in reality they are the same curve traversed in opposite directions.)

Now the curve $C = C_1 + C_3 + C_2 - C_3$ is a simple closed curve and Green's theorem holds on it. But the region inside C is exactly R and the contributions of the two copies of C_3 cancel. That is, we have shown that

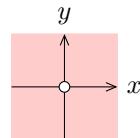
$$\iint_R \operatorname{curl} \mathbf{F} dA = \int_{C_1+C_3+C_2-C_3} \mathbf{F} \cdot d\mathbf{r} = \int_{C_1+C_2} \mathbf{F} \cdot d\mathbf{r}.$$

This is exactly Green's theorem, which we wanted to prove.



Example GT.23. Let $\mathbf{F} = \frac{(-y, x)}{r^2}$ ("tangential field")

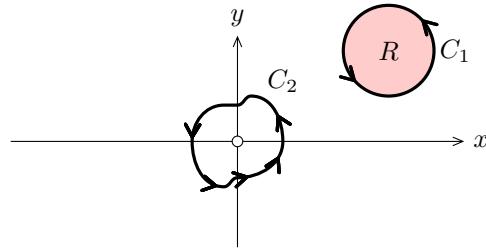
\mathbf{F} is defined on $D = \text{plane} - (0,0) =$ the punctured plane. (Shown below.)



It's easy to compute (we've done it before) that $\operatorname{curl} \mathbf{F} = 0$ in D .

Question: For the tangential field \mathbf{F} what values can $\oint_C \mathbf{F} \cdot d\mathbf{r}$ take for C a simple closed curve (positively oriented)?

answer: We have two cases (i) C_1 not around 0 (ii) C_2 around 0



In case (i) Green's theorem applies because the interior does not contain the problem point at the origin. Thus,

$$\oint_{C_1} \mathbf{F} \cdot d\mathbf{r} = \iint_R \operatorname{curl} \mathbf{F} \, dA = 0.$$

For case (ii) we will show that $\oint_{C_2} \mathbf{F} \cdot d\mathbf{r} = 2\pi$.

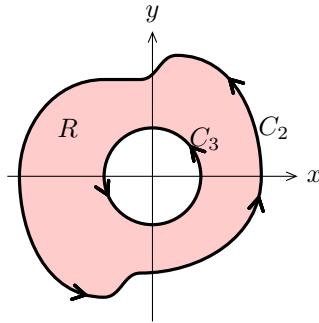
Let C_3 be a small circle of radius a , entirely inside C_2 . By the extended Green's theorem we have

$$\oint_{C_2} \mathbf{F} \cdot d\mathbf{r} - \oint_{C_3} \mathbf{F} \cdot d\mathbf{r} = \iint_R \operatorname{curl} \mathbf{F} \, dA = 0.$$

Thus, $\oint_{C_2} \mathbf{F} \cdot d\mathbf{r} = \oint_{C_3} \mathbf{F} \cdot d\mathbf{r}$.

Using the usual parametrization of a circle we can easily compute that the line integral is

$$\oint_{C_3} \mathbf{F} \cdot d\mathbf{r} = \int_0^{2\pi} 1 \, dt = 2\pi. \quad QED.$$

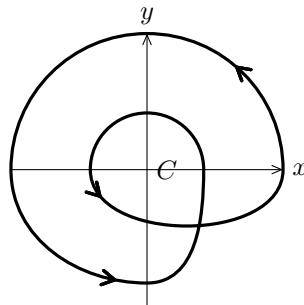


Answer to the question: The only possible values are 0 and 2π .

We can extend this answer in the following way:

If C is not simple, then the possible values of $\oint_C \mathbf{F} \cdot d\mathbf{r}$ are $2\pi n$, where n is the number of times C goes (counterclockwise) around $(0,0)$.

Not for class: n is called the *winding number* of C around 0 . n also equals the number of times C crosses the positive x -axis, counting $+1$ from below and -1 from above.



Example GT.24. Let $\mathbf{F} = r^n(x, y)$. Use extended Green's theorem to show that \mathbf{F} is conservative for all integers n . Find a potential function.

answer: First note, $M = r^n x$, $N = r^n y$. Thus, $M_y = nr^{n-2}xy = N_x$. Said differently: $\text{curl}\mathbf{F} = 0$.

We can't yet claim that \mathbf{F} is conservative because the origin is a possible problem point for \mathbf{F} . So, we'll show \mathbf{F} is conservative by showing $\int_C \mathbf{F} \cdot d\mathbf{r} = 0$ for all simple closed curves C .

Suppose C_1 is a simple closed curve not around 0 then Green's theorem applies and we have

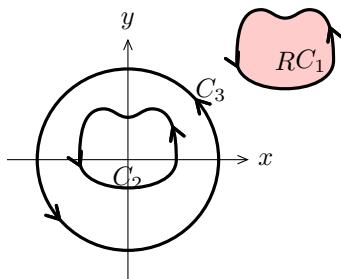
$$\int_{C_1} \mathbf{F} \cdot d\mathbf{r} = 0.$$

Next, suppose C_3 is a circle centered on $(0,0)$ then, since \mathbf{F} is radial and $d\mathbf{r}$ is tangential to the curve we know $\mathbf{F} \cdot d\mathbf{r} = 0$. So, $\oint_{C_3} \mathbf{F} \cdot d\mathbf{r} = 0$. (If you don't like this argument, parametrize the circle in the usual way and show that the resulting integrand is 0.)

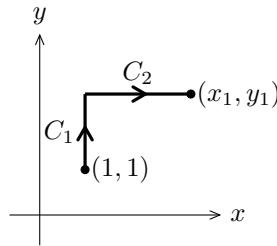
Now take the circle C_3 so large that it completely surrounds C_2 . In this case, extended Green's theorem implies (since $\text{curl}\mathbf{F} = 0$ between the curves) that

$$\oint_{C_2} \mathbf{F} \cdot d\mathbf{r} = \oint_{C_3} \mathbf{F} \cdot d\mathbf{r} = 0.$$

Thus $\oint_C \mathbf{F} \cdot d\mathbf{r} = 0$ for all closed loops. As we know, this implies that \mathbf{F} is conservative.



To find the potential function we use method 1 over the curve C shown. The calculation works for $n \neq -2$. For $n = -2$ everything is the same except we get natural logs instead of powers. (We also ignore the fact that if (x_1, y_1) is on the negative x -axis we should use a different path that doesn't go through the origin. This isn't really an issue since we already know a potential function exists, so continuity would handle these points without using an integral.)



Curve for computing the potential function.

Here's the calculation of the potential:

$$\begin{aligned}
 f(x_1, y_1) &= \int_C r^n x \, dx + r^n y \, dy \\
 &= \int_1^{y_1} (1 + y^2)^{n/2} y \, dy + \int_1^{x_1} (x^2 + y_1^2)^{n/2} x \, dx \\
 &= \frac{(1 + y^2)^{(n+2)/2}}{n+2} \Big|_1^{y_1} + \frac{(x^2 + y_1^2)^{(n+2)/2}}{n+2} \Big|_1^{x_1} \\
 &= \frac{(1 + y_1^2)^{(n+2)/2} - 2^{(n+2)/2}}{n+2} + \frac{(x_1^2 + y_1^2)^{(n+2)/2} - (1 + y_1^2)}{(n+2)/2} \\
 &= \frac{(x_1^2 + y_1^2)^{(n+2)/2} - 2^{(n+1)/2}}{n+2}
 \end{aligned}$$

Thus, our potential function is

$$f(x, y) = \frac{r^{n+2}}{n+2} + C.$$

In the special case of $n = -2$ we get $f(x, y) = \ln r + C$.

*Much of this note is freely borrowed from an MIT 18.01 note written by Arthur Mattuck.

16 Partial fractions and the coverup method

16.1 Heaviside Cover-up Method

16.1.1 Introduction

The cover-up method was introduced by Oliver Heaviside as a fast way to do a decomposition into partial fractions. This is an essential step in using the Laplace transform to solve differential equations, and this was more or less Heaviside's original motivation.

The cover-up method can be used to make a partial fractions decomposition of a proper rational function $\frac{P(s)}{Q(s)}$ whenever the denominator can be factored into distinct linear factors.

We will see that we can extend the coverup method to also deal with distinct quadratic factors.

Note. We put this section first, because the coverup method is so useful and many people have not seen it. Some of the later examples rely on the full algebraic method of undetermined coefficients presented in the next section. If you have never seen partial fractions you should read that section first.

16.1.2 Linear Factors

We first show how the method works on a simple example, and then show why it works.

Example PF.1. Decompose $\frac{s-7}{(s-1)(s+2)}$ into partial fractions.

answer: We know the answer will have the form

$$\frac{s-7}{(s-1)(s+2)} = \frac{A}{s-1} + \frac{B}{s+2}. \quad (63)$$

To determine A by the cover-up method, on the left-hand side we mentally remove (or cover up with a finger) the factor $s - 1$ associated with A , and substitute $s = 1$ into what's left; this gives A :

$$\left. \frac{s-7}{(s+2)} \right|_{s=1} = \frac{1-7}{1+2} = -2 = A. \quad (64)$$

Similarly, B is found by covering up the factor $s + 2$ on the left, and substituting $s = -2$ into what's left. This gives

$$\left. \frac{s-7}{(s-1)} \right|_{s=-2} = \frac{-2-7}{-2-1} = 3 = B.$$

Thus, our answer is

$$\frac{s-7}{(s-1)(s+2)} = \frac{-2}{s-1} + \frac{3}{s+2}. \quad (65)$$

16.1.3 Why does the method work?

The reason is simple. The “right” way to determine A from Equation 63 would be to multiply both sides by $(s - 1)$; this would give

$$\frac{s-7}{(s+2)} = A + \frac{B}{s+2}(s-1). \quad (66)$$

Now if we substitute $s = 1$, what we get is exactly Equation 64, since the term on the right with B disappears. The cover-up method therefore is just an easy and efficient way of doing the calculations.

In general, if the denominator of the proper rational function factors into the product of distinct linear factors:

$$\frac{P(s)}{(s-a_1)(s-a_2)\cdots(s-a_r)} = \frac{A_1}{s-a_1} + \cdots + \frac{A_r}{s-a_r}, \quad a_i \neq a_j,$$

then A_i is found by covering up the factor $s - a_i$ on the left, and setting $s = a_i$ in the rest of the expression.

Example PF.2. Decompose $\frac{1}{s^3 - s}$ into partial fractions.

answer: Factoring, $s^3 - s = s(s^2 - 1) = s(s - 1)(s + 1)$. By the cover-up method,

$$\frac{1}{s(s-1)(s+1)} = \frac{-1}{s} + \frac{1/2}{s-1} + \frac{1/2}{s+1}.$$

To be honest, the real difficulty in all of the partial fractions methods (the cover-up method being no exception) is in factoring the denominator.

16.1.4 Quadratic Factors

There are multiple ways to deal with a quadratic factors that don't have real roots.

1. Factor the quadratic factor into complex linear factors.
2. Use 'complex coverup'
3. Use algebraic techniques with just real numbers

We will show all three methods using the same example

Example PF.3. Decompose $G(s) = \frac{s-1}{(s+1)(s^2+4)}$ by factoring the denominator into complex linear terms

$$\text{answer: } \frac{s-1}{(s+1)(s^2+4)} = \frac{s-1}{(s+1)(s-2i)(s+2i)} = \frac{A}{s+1} + \frac{B}{s-2i} + \frac{C}{s+2i}.$$

Using coverup we see that $A = -2/5$, $B = \frac{2i-1}{(2i+1)(4i)} = \frac{4-3i}{20}$, $C = \frac{4+3i}{20}$. So,

$$G(s) = \frac{-2/5}{s+1} + \frac{(4-3i)/20}{s-2i} + \frac{(4+3i)/20}{s+2i} = \frac{-2/5}{s+1} + \frac{(8s+12)/20}{s^2+4} = \frac{-2/5}{s+1} + \frac{(2s/5+3/5)}{s^2+4}.$$

Example PF.4. Decompose $G(s) = \frac{s-1}{(s+1)(s^2+4)}$ without using complex techniques.

answer: Notice that in the previous example in the last expression for $G(s)$ the numerator of the s^2+4 term in the partial fraction decomposition is a linear term instead of a constant. This is the general rule for quadratic terms.

$$\frac{s-1}{(s+1)(s^2+4)} = \frac{A}{s+1} + \frac{Bs+C}{s^2+4}. \quad (67)$$

Notice the quadratic factor gets a linear term in the numerator. Notice also that the number of unknown coefficients is the same as the degree of the denominator in the original fraction.

Using coverup we find $A = -2/5$. Now we can do some algebra to find B and C . Multiplying Equation 67 through by the denominator gives

$$s-1 = A(s^2+4) + (Bs+C)(s+1) = (A+B)s^2 + (B+C)s + (4A+C).$$

Equate the coefficients on both sides:

$$\begin{aligned} s^2 : \quad 0 &= A + B \\ s : \quad 1 &= B + C \\ s^2 : \quad -1 &= 4A + C \end{aligned}$$

Since we already know A , the first equation gives $B = 2/5$, and then the second gives $C = 3/5$. Comparing this with the previous example we see we've found the same answer.

Example PF.5. Decompose $\frac{s-1}{(s+1)(s^2+4)}$ using ‘complex coverup’.

answer: This starts the same as the previous example.

$$\frac{s-1}{(s+1)(s^2+4)} = \frac{A}{s+1} + \frac{Bs+C}{s^2+4}. \quad (68)$$

Coverup gives us $A = -2/5$. To do complex coverup, mentally multiply the above equation by $s^2 + 4$ and then substitute $s = 2i$. We get

$$\frac{2i-1}{2i+1} = B(2i) + C \Leftrightarrow \frac{3+4i}{5} = C + 2Bi.$$

So, $C = 3/5$ and $B = 2/5$. (Same as before.)

Example PF.6. Don't be fooled by quadratic terms that factor into linear ones.

$$\frac{1}{(s+1)(s^2-4)} = \frac{1}{(s+1)(s+2)(s-2)} = \frac{A}{s+1} + \frac{B}{s+2} + \frac{C}{s-2}.$$

Example PF.7. Don't forget that the rational function must be proper. For example, decompose $\frac{s^3+2s+1}{s^2+s-2}$ using partial fractions.

answer: First, we must use long-division to make this proper.

$$\frac{s^3+2s+1}{s^2+s-2} = s-1 + \frac{5s-1}{s^2+s-2} = s-1 + \frac{5s-1}{(s+2)(s-1)} = s-1 + \frac{A}{s+2} + \frac{B}{s-1}.$$

Using coverup we get $A = 11/3$, $B = 4/3$.

16.1.5 Repeated Linear Factors

The cover-up method can also be used if a linear factor is repeated, but it gives just partial results. It applies only to the *highest power of the linear factor*.

Example PF.8. Decompose $\frac{1}{(s-1)^2(s+2)}$.

answer: We write

$$\frac{1}{(s-1)^2(s+2)} = \frac{A}{(s-1)^2} + \frac{B}{s-1} + \frac{C}{s+2}. \quad (69)$$

To find A cover up $(s-1)^2$ and set $s = 1$; you get $A = 1/3$. To find C , cover up $s+2$, and set $s = -2$; you get $C = 1/9$.

This leaves B which cannot be found by the cover-up method. But since A and C are already known in Equation 69, B can be found by substituting any numerical value (other

than 1 or -2) for s in (69). For instance, if we put $s = 0$ and remember that $A = 1/3$ and $C = 1/9$, we get

$$\frac{1}{2} = \frac{1/3}{1} + \frac{B}{-1} + \frac{1/9}{2},$$

giving $B = -1/9$.

B could also be found by applying the method of undetermined coefficients to the Equation 69; note that since A and C are known, it is enough to get a single linear equation in order to determine B — simultaneous equations are no longer needed.

The fact that the cover-up method works for just the *highest power* of the repeated linear factor can be seen just as before. In the above example for instance, the cover-up method for finding A is just a short way of multiplying Equation 71 through by $(s - 1)^2$ and then substituting $s = 1$ into the resulting equation.

16.2 Partial Fractions: Undetermined Coefficients

16.2.1 Introduction

Logically this section should precede the previous one on coverup since it explains what we are doing with partial fractions and shows an algebraic method that never fails. However, since most students in this course will have seen partial fractions before it seemed reasonable to start with the coverup method.

16.2.2 Rational Functions

A **rational function** is one that is the ratio of two polynomials. For example

$$\frac{s+1}{s^2+7s+9} \quad \text{and} \quad \frac{s^2+7s+9}{s+1}$$

are both rational functions.

A rational function is called **proper** if the degree of the numerator is strictly smaller than the degree of the denominator; in the examples above, the first is proper while the second is not.

Long-division: Using long-division we can always write an improper rational function as a polynomial plus a proper rational function. The partial fraction decomposition only applies to proper functions.

Example PF.9. Use long-division to write $\frac{s^3+2s+1}{s^2+s-2}$ as the sum of a polynomial and a proper rational function.

answer:

$$\begin{array}{r|rrr} & s-1 & & \\ s^2+s-2 & \overline{s^3 + 2s + 1} & +1 & \\ & s^3 & +s^2 & -2s \\ \hline & -s^2 & +4s & +1 \\ & -s^2 & -s & +2 \\ \hline & 5s & & -1 \end{array}$$

Therefore,

$$\frac{s^3 + 2s + 1}{s^2 + s - 2} = s - 1 + \frac{5s - 1}{s^2 + s - 2}.$$

16.2.3 Linear Factors

Here we assume the denominator factors into distinct linear factors. We start with a simple example. We will explain the general principle immediately afterwards.

Example PF.10. Decompose $R(s) = \frac{s-3}{(s-2)(s-1)}$ using partial fractions.

answer:

$$\frac{s-3}{(s-2)(s-1)} = \frac{A}{s-2} + \frac{B}{s-1}.$$

Multiplying both sides by the denominator on the left gives

$$s-3 = A(s-1) + B(s-2) \quad (70)$$

The sure algebraic way is to expand out the right hand side and equate the coefficients with those of the polynomial on the left.

$$s-3 = (A+B)s + (-A-2B) \Rightarrow \begin{cases} \text{coeff. of } s: & 1 = A+B \\ \text{coeff. of } 1: & -3 = -A-2B \end{cases}$$

We solve this system of equations to find the undetermined coefficients A and B : $A = -1$, $B = 2$.

Answer: $R(s) = -1/(s-2) + 2/(s-1)$.

Note. Not surprisingly, using the coverup method by plugging the roots of each factor into Equation 70 would be easier because when you do this every term except one becomes 0.

$$\begin{array}{ll} \text{Plug in } s=1 & \Rightarrow -2 = B(-1) \Rightarrow B=2 \\ \text{Plug in } s=2 & \Rightarrow -1 = A(1) \Rightarrow A=-1. \end{array}$$

In general, if $P(s)/Q(s)$ is a proper rational function and $Q(s)$ factors into distinct linear factors $Q(s) = (s-a_1)(s-a_2)\cdots(s-a_n)$ then

$$\frac{P(s)}{Q(s)} = \frac{A_1}{s-a_1} + \frac{A_2}{s-a_2} + \cdots + \frac{A_n}{s-a_n}.$$

The proof of this is not hard, but we will not give it. Remember you must have a *proper* rational function and each of the factors must be distinct. Repeated factors are discussed below.

Example PF.11. Decompose $\frac{3}{s^3 - 3s^2 - s + 3}$.

answer: The hardest part of this problem is to factor the denominator. For higher order polynomials this might be impossible. In this case you can check

$$s^3 - 3s^2 - s + 3 = (s-1)(s+1)(s-3).$$

The partial fractions decomposition is

$$\frac{3}{(s-1)(s+1)(s-3)} = \frac{A}{s-1} + \frac{B}{s+1} + \frac{C}{s-3}.$$

Multiplying through by the denominator gives

$$3 = A(s+1)(s-3) + B(s-1)(s-3) + C(s-1)(s+1).$$

Plugging in $s = 1$ gives $A = -3/4$, likewise $s = -1$ gives $B = 3/2$ and $s = 3$ gives $C = -3/4$.

16.2.4 Quadratic Factors

Example PF.12. Decompose $\frac{5s+6}{(s^2+4)(s-2)}$.

answer: We write

$$\frac{5s+6}{(s^2+4)(s-2)} = \frac{As+B}{s^2+4} + \frac{C}{s-2}. \quad (71)$$

We first determine C by the cover-up method, getting $C = 2$. Then A and B can be found by the method of undetermined coefficients; the work is greatly reduced since we need to solve only two simultaneous equations to find A and B , not three.

Following this plan, using $C = 2$, we combine terms on the right of (71) so that both sides have the same denominator. The numerators must then also be equal, which gives us

$$5s+6 = (As+B)(s-2) + 2(s^2+4). \quad (72)$$

Comparing the coefficients of s^2 and of the constant terms on both sides of (72) gives the two equations

$$0 = A + 2 \quad \text{and} \quad 6 = -2B + 8,$$

from which $A = -2$ and $B = 1$.

In using Equation 72, one could have instead compared the coefficients of s , getting $5 = -2A + B$, leading to the same result, but providing a valuable check on the correctness of the computed values for A and B .

In Example PF.12, an alternative to undetermined coefficients would be to substitute two numerical values for s into the original Equation 71, say $s = 0$ and $s = 1$ (any values other than $s = 2$ are usable). Again one gets two simultaneous equations for A and B . This method requires addition of fractions, and is usually better when only one coefficient remains to be determined (as in the example just below).

16.2.5 Repeated Linear Factors

For repeated linear factors we need one partial fraction term for each power of the factor as illustrated by the following example.

Example PF.13. Decompose $G(s) = \frac{2s}{s^3(s+1)^2(s+2)}$ using partial fractions.

answer:

$$G(s) = \frac{2s}{s^3(s+1)^2(s+2)} = \frac{A}{s} + \frac{B}{s^2} + \frac{C}{s^3} + \frac{D}{s+1} + \frac{E}{(s+1)^2} + \frac{F}{s+2}$$

Here the denominator has a linear factor s repeated three times (term s^3), and a linear factor $(s+1)$ repeated twice (term $(s+1)^2$); hence three partial fractions are associated with the first, while two are associated with the latter. The term $(s+2)$ which is not repeated leads to one partial fraction as previously seen. You can check that the coefficients are

$$A = -5/2, B = 1, C = 0, D = 2, E = 2, F = 1/2.$$

16.2.6 Repeated Quadratic Factors

Just like repeated linear factors, quadratic factors have one term for each power of the factor as illustrated in the following example.

Example PF.14. Find $G(s) = \frac{2s}{s(s^2+1)^2(s^2+4s+6)}$ using partial fractions.

answer: The partial fractions decomposition is

$$G(s) = \frac{2s}{s(s^2+1)^2(s^2+4s+6)} = \frac{A}{s} + \frac{Bs+C}{s^2+1} + \frac{Ds+E}{(s^2+1)^2} + \frac{Fs+G}{s^2+4s+6}$$

Note the repeated factor $(s^2+1)^2$ lead to two partial fraction terms.

We won't compute the coefficients – you can do this by going through the algebra.

16.2.7 Complex Factors

As in Example PF.3 we can allow complex roots. In this case all quadratic terms factor into linear terms.

Example PF.15. Decompose $s/(s^2+\omega^2)$ using complex partial fractions.

answer:

$$\frac{s}{s^2+\omega^2} = \frac{s}{(s-i\omega)(s+i\omega)} = \frac{A}{s-i\omega} + \frac{B}{s+i\omega}.$$

Multiplying through by the denominator gives $s = A(s+i\omega) + B(s-i\omega)$.

Plug in $s = i\omega \Rightarrow A = 1/2$.

Plug in $s = -i\omega \Rightarrow B = 1/2$.
