132H ESSENTIALS

FRANZ PEDIT, FALL 2018

1. Integral and Anti-derivative

If $f\colon [a,b] \to \mathbb{R}$ is a continuous function, any differentiable function $G\colon [a,b] \to \mathbb{R}$ satisfying

$$G' = f \tag{1}$$

is called an *anti-derivative* of f. The 1st FTC asserts that the "area function"

$$F(x) := \int_a^x f(t)\,dt\,,$$

measuring the area of the graph $y = f(x)$ over the interval $[a,x]$ with $x \in [a,b]$, is an anti-derivative of f, that is, $F' = f$.

We have also seen that any two anti-derivatives of f differ by a constant. Thus, if we have found some anti-derivative G of a function f, we know that

$$G(x) = \int_a^x f(t)\,dt + c$$

for some constant $c \in \mathbb{R}$ (the (in)famous constant of integration). Using this, we obtained the 2nd FTC, which states that

$$\int_a^b f(t)\,dt = G(b) - G(a) \tag{2}$$

for any anti-derivative G of f.

In order to make our notation more concise and intuitive, we have introduced the *indefinite integral* notation

$$\int f(x)\,dx$$

for anti-derivatives of f. Then the relevant relations between differentiation and integration read

$$\frac{d}{dx}\int f(x)\,dx = f(x) \quad \text{and} \quad \int g'(x)\,dx = g(x)\,. \tag{3}$$

Note that the first relation is in essence the FTC (in either version), whereas the second equation just states the obvious, namely that a differentiable function g is an anti-derivative of its derivative g'. This last gives us anti-derivatives of some of the basic functions, by reading the derivative rules backwards.

Using the notation of *differentials*, (3) takes the rather elegant (and possibly more intuitive) form

$$d\int \alpha = \alpha \quad \text{and} \quad \int dg = g \tag{4}$$

for a continuously differentiable function g and an arbitrary continuous differential $\alpha = f(x)dx$. Note that the FTC asserts that every differential $\alpha = f(x)\,dx$ is *exact*, that is, of the form $\alpha = dF$, namely $F = \int \alpha$. This is specific to one variable Calculus and will fail to hold in two and more variables Calculus 233H and 425H.

2. Integration by Parts and Substitution Rule

The Leibniz, or product, rule and the chain rule give us two methods to find anti-derivatives of more complicated functions. The Leibniz rule yields the *integration by parts* formula

$$\int f(x)g'(x)\,dx = -\int f'(x)g(x)\,dx + f(x)g(x)$$

for continuously differentiable functions f and g. Using differentials the formula reads

$$\int f\,dg = -\int df\,g + fg\,. \tag{5}$$

The chain rule provides us with *the substitution rule*

$$\int_{g(a)}^{g(b)} f(y)\,dy = \int_a^b f(g(x))g'(x)\,dx \tag{6}$$

for a continuous function $f\colon I \to \mathbb{R}$ and a continuously differentiable function $g\colon [a,b] \to I$, where $I \subset \mathbb{R}$ is an interval. Using the differentials notation, this formula produces itself via the substitution $y = g(x)$ and thus $dy = dg = g'(x)\,dx$.

3. Length of curves

We discussed in class how the length of a (smooth) curve $\gamma\colon [a,b] \to \mathbb{R}^2$ can be thought of as the limit of finer and finer polygonal approximations of the curve. This process resulted in a Riemann sum for the speed function $||\gamma'||$, and thus it is reasonable to define the *length of a curve* by

$$L(\gamma) = \int_a^b ||\gamma'(t)||\,dt \tag{7}$$

We also introduced notations coming from physical intuitions: the *velocity* (vector) γ' (rate of change of position), the already mentioned *speed*, the length of the velocity $||\gamma'||$, and the *acceleration* (vector) γ'' (rate of change of velocity).

Given a (smooth) function $f\colon I \to \mathbb{R}$, we defined its *graph parametrization* to be the curve

$$\gamma\colon I \to \mathbb{R}^2\,, \quad \gamma(t) := (t, f(t)) \tag{8}$$

which traces out the graph $y = f(x)$ of f over the interval $I \subset \mathbb{R}$. Since $\gamma'(t) = (1, f'(t))$, we obtain from (7) the length of the graph curve

$$L(\gamma) = \int_a^b \sqrt{1 + (f'(t))^2}\,dt$$

where $I = [a,b]$. Disregarding the "rule" that differentials can't be multiplied (in fact, they can....), the above formula for the length of the graph curve can be rewritten as

$$L(\gamma) = \int_a^b \sqrt{dt^2 + df^2}\,.$$

Now, since for the graph parametrization $x = t$ and $y = f(x)$, thus $dt = dx$ and $dy = df = f'(x)dx$, you can rewrite this last as

$$L(\gamma) = \int_a^b \sqrt{dx^2 + dy^2}\,.$$

If you are adventurous, you may want to contemplate the "differential version" for the length formula for an arbitrary curve $\gamma = (\gamma_1, \gamma_2)$, namely

$$L(\gamma) = \int_a^b \sqrt{d\gamma \cdot d\gamma}$$

where we think of $d\gamma = (d\gamma_1, d\gamma_2)$ as an $\mathbb{R}^2$-valued differential—something to contemplate—and the dot indicates the dot-product on $\mathbb{R}^2$. Recalling that $d\gamma_1 = \gamma_1'(t)dt$ and similar for $d\gamma_2$, we obtain

$$\int_a^b \sqrt{d\gamma \cdot d\gamma} = \int_a^b \sqrt{(d\gamma_1)^2 + (d\gamma_2)^2} = \int_a^b \sqrt{(\gamma_1'(t))^2\, dt^2 + (\gamma_2'(t))^2\, dt^2} = \int_a^b \sqrt{(\gamma_1'(t))^2 + (\gamma_2'(t))^2}\, dt$$

and therefore

$$L(\gamma) = \int_a^b \sqrt{(\gamma_1'(t))^2 + (\gamma_2'(t))^2}\, dt$$

which is indeed the formula for the length of the curve γ. Putting $\gamma_1 = x$ and $\gamma_2 = y$, so that $dx = \gamma_1'\, dt$, $dy = \gamma_2'\, dt$, we get the formula

$$L(\gamma) = \int_a^b \sqrt{dx^2 + dy^2}\,.$$

So, no matter what kind of curve we have, this last formula in terms of differentials works, as long as we understand what it all means...

4. Areas and Volumes

A further application of integration is the computation of area and volume for surfaces of revolution. The set up is a s follows: we have a (smooth) function $f\colon [a, b] \to \mathbb{R}$ and rotate its graph around the x-axis, which is the *surface of revolution* with profile curve $y = f(x)$. We want to calculate a formula for the area of this surface (e.g., the area of the piece of paper needed to wrap around an old can to make it a nice looking pencil holder), and the volume of the solid , the bulk inside the surface of revolution (e..g,, how much water does the can hold).

4.1. **Area.**

In class I have made a conceptual mistake, which is best explained by Picture 1. With the arguments given in class, the area of a portion of the surface where the graph is very steep, will be very badly approximated by a cylinder area. We have to somehow account for the steepness of the graph in our approximation attempts.

A good sanity check is to use the *wrong* formula (based on cylinder approximation) from class and apply it to a cone of base radius R and height h: this cone is obtained by rotating the line $y = \frac{R}{h}x$ around the x-axis, and we obtain for its area

$$A(\text{cone}) = 2\pi \int_0^h \frac{R}{h}\, dx = 2\pi R\,.$$

This is patently not the surface area of a cone, which is known to be $\pi R\sqrt{h^2 + R^2}$ (we only look at the cone mantle, not including the area of the base disk, which would add πR^2). In case you have not done this in high school, the surface area of the cone (mantle) can be calculated without integrals by cutting the cone along a straight line from its base to its vertex, then unrolling it onto

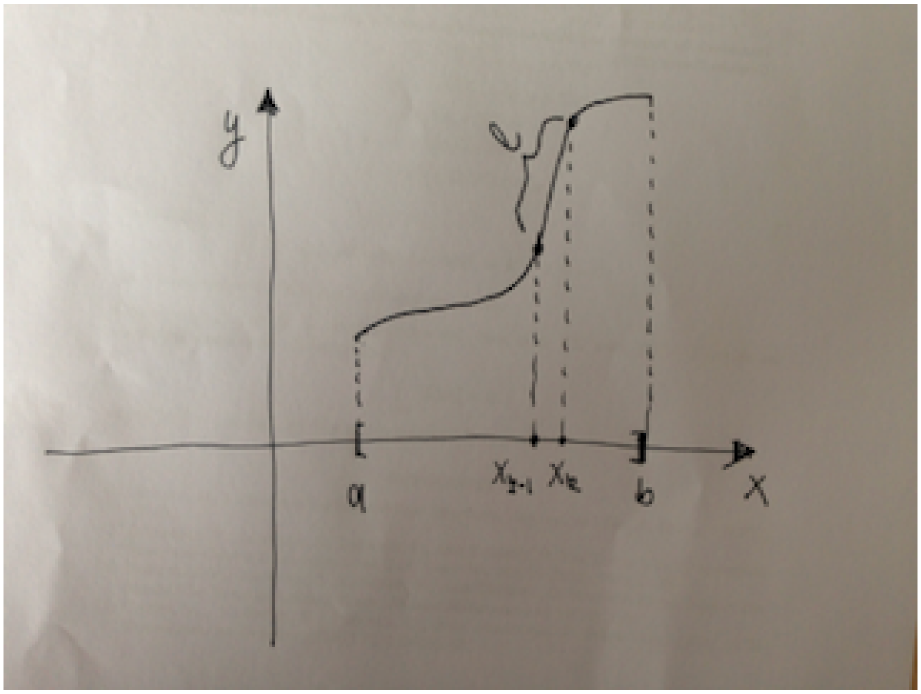

FIGURE 1. The ratio of l to subinterval length $x_k - x_{k-1}$ depends on the steepness of the graph. It converges to $\sqrt{1 + (f'(x_k))^2}$ as $x_{k-1} \to x_k$, and not to 1, as it would be for the cylinder approximation.

the plane, which results in a disk segment (this is how children make cone hats out of paper, and you should do this yourself at least once), see Picture 2. That disk segment (what is the angle of the wedge given R and h?) has the same area as the cone mantle, and is given by $\pi R\sqrt{h^2 + R^2}$ (which can be deduced with or without using integration).

Now that we understand that the steepness of the graph $y = f(x)$ has to be accounted for, we set up a new approximation for the area of the surface obtained by revolving $y = f(x)$ (assuming for simplicity that $f(x) \geq 0$ for all $x \in [a, b]$) around the x-axis: given a partition $\mathcal{P} : (x_0 = a, x_1, \ldots, x_{n-1}, x_n = b)$, we approximate the area over each subinterval $[x_{k-1}, x_k]$ by

$$2\pi f(\eta_k)\sqrt{(x_k - x_{k-1})^2 + (f(x_k) - f(x_{k-1}))^2} = 2\pi f(\eta_k)\sqrt{1 + \left(\frac{f(x_k) - f(x_{k-1})}{x_k - x_{k-1}}\right)^2}\,(x_k - x_{k-1})$$

for some, *yet to be determined*, evaluation point $\eta_k \in [x_{k-1}, x_k]$. Notice that the above expression is nothing but the length of the secant line connecting the points $(x_{k-1}, f(x_{k-1}))$ and $(x_k, f(x_k))$ on the graph $y = f(x)$ times some average radius $f(\eta_k)$ over the subinterval $[x_{k-1}, x_k]$. By the Mean Value Theorem, we can rewrite

$$\frac{f(x_k) - f(x_{k-1}}{x_k - x_{k-1}} = f'(\xi_k)$$

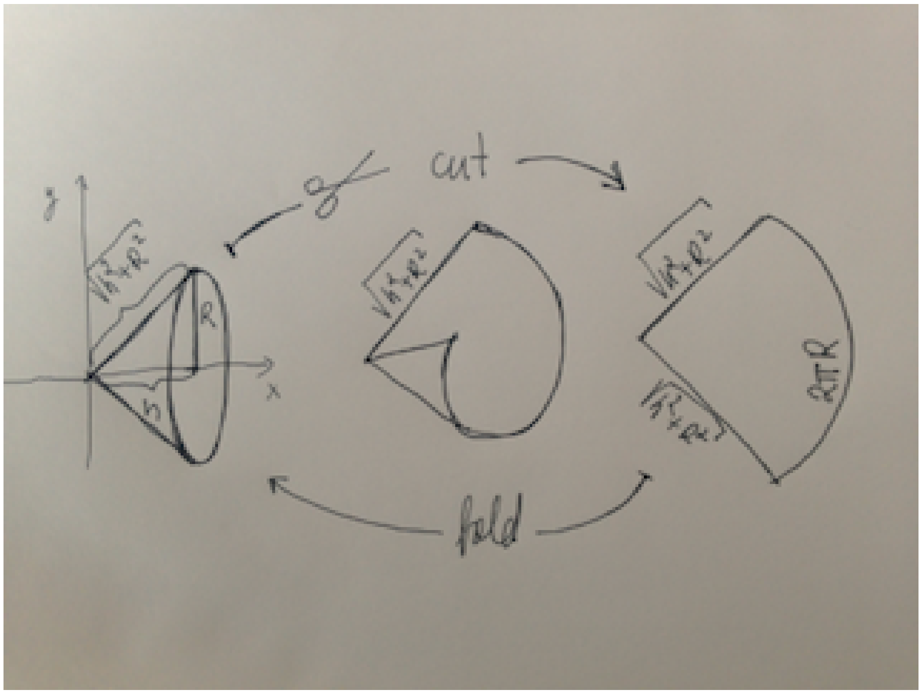

FIGURE 2. The area of a cone.

for some $\xi_k \in [x_{k-1}, x_k]$. We now take η_k to be this very ξ_k (whose existence is guaranteed by the MVT), and obtain for our approximate area over the subinterval $[x_{k-1}, x_k]$ the expression

$$2\pi \; f(\xi_k)\sqrt{1 + (f'(\xi_k))^2}\,(x_k - x_{k-1})\,.$$

Now we just need to sum this last expression over all subintervals of the partition, to obtain an approximate area of the surface of revolution

$$A \approx 2\pi \sum_{k=1}^{n} f(\xi_k)\sqrt{1 + (f'(\xi_k))^2}\,(x_k - x_{k-1})\,.$$

But this is just the Riemann sum of the function $2\pi f(x)\sqrt{1 + (f(x)')^2}$, so that by refining the partition and taking the limit as the subinterval length tends to zero, the area of the surface of revolution of the graph $y = f(x)$ becomes

$$A = 2\pi \int_a^b |f(x)|\sqrt{1 + (f'(x))^2}\,dx\,,$$

where we have put the absolute value into our formula to allow the function f to change signs over the interval $[a, b]$.

4.2. **Volume.** The volume formula from our discussion on class remains the same. We approximate, over each subinterval, by the cylinder volume

$$\pi f(\xi_k)^2 (x_k - x_{k-1})$$

of the cylinder of hight $(x_k - x_{k-1})$ and some average radius $f(\xi_k)$ with $\xi_k \in [x_{k-1}, x_k]$. Notice that the steepness of the graph will not really matter here, why? Then, like shown in class, the volume becomes

$$V = \pi \int_a^b f(x)^2 \, dx \, .$$

5. The exponential function—an important example of an infinite series

We discussed in class that the only functions we could in principle compute, when sitting on some island writing symbols in the sand, are polynomials and rational functions. For *transcendental functions* like trig functions we have a geometric definition, but this does not help much with calculating explicit values of those functions. We are slightly better off with the natural logarithm, which could be defined via the integral formula

$$\ln(x) := \int_1^x \frac{dt}{t} = \text{ area under the graph of } y = 1/t \text{ over the interval } [1, x] \, .$$

Using Riemann sums to approximate the integral, or using some more refined numerical integration algorithms, we could in principle calculate values for $\ln(x)$ with reasonable precision. Another well known function is the exponential function $y = e^x$, which somehow appears when looking at powers of a number, i.e. a^n, for integers $n \in \mathbb{Z}$, and generalizing this, in the case $a \geq 0$, to arbitrary real powers a^x for $x \in \mathbb{R}$. This is done by a limiting process, where one approximates a real number x by a sequence of rational numbers $r_n = p_n/q_n$, where $p_n, q_n \in \mathbb{Z}$, which limit to x, that is, $\lim_{n\to\infty} r_n = x$. Then one sets

$$a^x = \lim_{n\to\infty} a^{r_n}$$

where we interpret $a^{p/q}$ as the q^{th}-root of $a^p = a \cdots a$, multiplied p-many times. Perhaps less obvious in this approach is why the number $e = 2.71828182846\ldots$, called *Euler's number*, plays a central role. This is why a more fundamental approach to the exponential function is needed.

It is interesting that a process called *beta decay*, which is responsible for radioactivity, started more or less immediately after the Big Bang. It is governed by a fundamental law, namely that the average rate of change of the material (which is subjected to beta decay) over some time interval is proportional to how much material is present on average over that time interval. Expressed in a formula this becomes

$$\frac{Q(t+\Delta t) - Q(t)}{\Delta t} = k \frac{Q(t+\Delta t) + Q(t)}{2}$$

where $k \in \mathbb{R}$ is the proportionality factor. In particular, for a decay process, that is, the quantity $Q(t)$ decreases over time, $k < 0$ and for a growth process, when $Q(t)$ increases over time, $k > 0$. Taking the limit as $\Delta t \to 0$, we arrive at

$$Q'(t) = kQ(t) \tag{9}$$

which says that the instantaneous rate of change of the function $Q(t)$ at time t is proportional to its value $Q(t)$ at that time. Now beta decay has been happening over the last 10 Billion or so years, most likely long before there were sentient beings able to conceive the mathematical frame work to express this process in a formula.

The question for us of course is, what function(s) satisfy the relation (9), which is called a *first order differential equation (relation)* for the unknown function $Q(t)$? In order to not get bogged down by unnecessary constants etc., we choose $k = 1$ and look at the equation $Q'(t) = Q(t)$. Since the only functions we really understand are polynomials, we tried the Ansatz (German for "educated guess") that $Q(t)$ is a polynomial, but this failed. But we learned that if we don't limit the degree of the polynomial, we would succeed, namely the function

$$Q(t) = \sum_{k=0}^{\infty} \frac{t^k}{k!}$$

is a solution of $Q'(t) = Q(t)$ for the *initial condition* $Q(0) = 1$, as can be checked by term by term differentiation. The issue here is, that we need to add terms infinitely often, and we don't really know how to make sense of that. Postponing that for a moment, we make

Definition 1. *The exponential function is defined by the infinite series*

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

In general, "infinite degree polynomials", such as $\exp(x)$, are called *power series*:

Definition 2. *A power series centered at $x_0 \in \mathbb{R}$ is an expression of the form*

$$\sum_{k=0}^{\infty} a_k (x - x_0)^k$$

where $a_k \in \mathbb{R}$ are the coefficients of the power series.

The exponential function thus is a power series centered at $x_0 = 0$ with coefficients $a_k = \frac{1}{k!}$. What we hope/wish for at this stage is the following:

(i) Under reasonable and checkable circumstances, a power series defines a function

$$f(x) := \sum_{k=0}^{\infty} a_k (x - x_0)^k$$

on some interval around x_0. Moreover, this function is infinitely often differentiable and integrable and one computes derivatives and integrals term by term. Since each term is, up to the constant a_k, a monomial function $(x - x_0)^k$, we can explicitly carry this out.

(ii) Any reasonable function $g\colon I \to \mathbb{R}$ defined on some interval $I \subset \mathbb{R}$ should have a description in terms of a power series, that is, choosing some $x_0 \in I$ then $g(x) = \sum_{k=0}^{\infty} a_k (x - x_0)^k$. If this is at all true, then, by taking repeated derivatives of this equation and evaluating at $x = x_0$, one immediately checks $a_k = \frac{g^{(k)}(x_0)}{k!}$. So this premeditates the Tayler series of a function....

Returning to the exponential function $\exp(x)$ and assuming (i) for now, we could easily show the following important properties of the exponential function:

(i) $\exp(x)$ is defined for all $x \in \mathbb{R}$, in particular, $e := \exp(1) \in \mathbb{R}$ is a well-defined number, which we call *Euler's number*. We used the infinite series to calculate $e = 2.71828182846\ldots$ up to as many decimals as we wanted.

(ii) $\exp(x + y) = \exp(x)\exp(y)$, the *functional equation*.

This was verified by regarding y as fixed, and differentiating the functions $\exp(x + y)$ and $\exp(x)\exp(y)$ w.r.t. x. This showed that both of these functions satisfy the relation $Q' = Q$ with $Q(0) = \exp(y)$. We then convinced ourselves that all solutions of $Q' = Q$ are of the form $c\exp(x)$ for some constant $c \in \mathbb{R}$. Therefore, $\exp(x + y) = c_1 \exp(x)$ and $\exp(x)\exp(y) = c_2 \exp(x)$. Putting $x = 0$, we see $c_1 = c_2 = \exp(y)$ and we verified the functional equation.

(iii) Since $1 = \exp(0) = \exp(x + (-x)) = \exp(x)\exp(-x)$, we also deduced
$$\frac{1}{\exp(x)} = \exp(-x)$$
.

(iv) From the above, we deduce $\exp(x) > 0$ and $\exp\colon \mathbb{R} \to (0, \infty)$ is strictly monotone increasing and concave up with $\lim_{x\to-\infty} \exp(x) = 0$ and $\lim_{x\to\infty} = \infty$. It is thus bijective and its inverse function $\exp^{-1}(x) = \ln(x)$ is the natural logarithm. To check this last, we needed to recall the formula for the derivative of the inverse function, which gives $[\exp^{-1}(x)]' = 1/x$. Thus, from the definition of $\ln(x) = \int_1^x \frac{dt}{t}$, both $\exp^{-1}(x)$ and $\ln(x)$ are antiderivatives of the function $y = 1/x$. So they differ by a constant, but $\ln(1) = 0$ and also $\exp^{-1}(1) = 0$ (since $\exp(0) = 1$), so that constant has to be zero. Being the inverse function of the exponential function, $\ln\colon (0, \infty) \to \mathbb{R}$ is bijective, strictly monoton increasing, and concave down, with $\lim_{x\to 0} \ln(x) = -\infty$ and $\lim_{x\to\infty} \ln(x) = \infty$.

With this at hand and the defining formulas for the inverse function
$$\exp(\ln(x)) = x \quad \text{and} \quad \ln(\exp(x)) = x$$
we deduce the functional relations for the natural logarithm:

(i) $\ln(1) = 0$
(ii) $\ln(x + y) = \ln(x) + \ln(y)$
(iii) $\ln(1/x) = -\ln(x)$
(iv) $\ln(x^\alpha) = \alpha \ln(x)$ for any $\alpha \in \mathbb{R}$.

The properties of exp and ln make it possible to define what it means to take an arbitrary power $x \in \mathbb{R}$ of a non-negative real number $a \geq 0$:
$$a^x := \exp(x \ln(a))$$
In light of the properties listed, this last definition makes a lot of sense. We checked that for $x = n \in \mathbb{Z}$ this reproduces the n-fold product of a with itself. And for rational $x = p/q$ it gives the usual expression of the q-th root of a^p, just as we expected. But notice, we derived everything from the exponential function, and Euler's number appears naturally. With this last definition, we now can write
$$e^x = \exp(x)$$
and everything you ever learned about e^x is contained in the above couple of paragraphs, fairly rigorously proven, from a first principle as old as the universe....The caveat is that we are dealing with infinite series and we have, so far, not really defined what it means to sum up infinitely many numbers. We worked in good faith that things pan out as we expect them to. The next section fills in the necessary theory of infinite series.

6. Complex numbers

When we discussed partial fraction decomposition, we needed to find the zeros of polynomials. Unfortunately, not all polynomials have real zeros as $P(x) = x^2 + 1$ shows. This problem has a long history and eventually got resolved by the introduction of complex numbers $\mathbb{C}$. A complex number is given by
$$z = x + iy \quad \text{where} \quad x, y \in \mathbb{R} \quad \text{and} \quad i^2 = -1$$

The "imaginary unit" $i = \sqrt{-1}$ is formally introduced to build a new number system which extends the real numbers $\mathbb{R} \subset \mathbb{C}$. We call

$$\mathrm{Re}(z) := x \quad \text{the real part and} \quad \mathrm{Im}(z) := y \quad \text{the imaginary part}$$

of the complex number $z = x + iy$. The pictorial representation of $\mathbb{C}$, given by Gauss, is the plane $\mathbb{R}^2$, where the x-axis presents the real numbers and the y-axis presents all the real multiples of i, the imaginary numbers. A complex number z then is just a point in the plane, whose x-coordinate is its real part and whose y-coordinate is its imaginary part. The usual operations of numbers hold for complex numbers (commutativity, associativity, distributivity) and the only new feature is to replace $i \cdot i = i^2 = -1$ in all calculations.

We list some new notation needed when calculating with complex numbers $z = x + iy$:

(i) $|z| := \sqrt{x^2 + y^2}$ is called the *absolute value* of $z \in \mathbb{C}$, which is just the Euclidean length of z viewed as a vector in the plane. Notice that $|z|$ becomes the absolute value in $\mathbb{R}$ when $z \in \mathbb{R}$.

(ii) $\bar{z} := x - iy$ is called the *conjugate complex number*, which is reflection of z in the real axis, the x-axis. Since

$$z\bar{z} = (x + iy)(x - iy) = x^2 + y^2 = |z|^2 \quad \text{we have} \quad |z| = \sqrt{z\bar{z}}\,.$$

Furthermore,

$$z \in \mathbb{R} \quad \text{is real, if and only if} \quad z = \bar{z}$$

and

$$z \in i\mathbb{R} \quad \text{is imaginary, if and only if} \quad z = -\bar{z}$$

Addition/subtraction and scaling by $c \in \mathbb{R}$ of complex numbers $z = x + iy$, $w = u + iv$ is just vector addition/subtraction and scaling:

$$z \pm w = (x \pm u) + i(y \pm v) \qquad cz = cx + icy$$

Multiplication of complex numbers

$$zw = (xu - yv) + i(xv + yu)$$

can be interpreted as first counter clockwise rotating w by the angle z makes with the x-axis and scaling the result by $|z|$. In particular, iz is the vector perpendicular to z of same length in counter clockwise direction.

The reciprocal of a non-zero complex number $z \neq 0$ gets calculated by

$$\frac{1}{z} = \frac{\bar{z}}{z\bar{z}} = \frac{\bar{z}}{|z|^2} = \frac{1}{|z|^2}(x - iy)$$

which shows that the reciprocal $\frac{1}{z}$ of a non-zero complex number is the reflection of z in the x-axis scaled by $\frac{1}{|z|^2}$. These operations are compatible with taking complex conjugation and absolute value in the following way:

$$\overline{z \pm w} = \bar{z} \pm \bar{w} \qquad \overline{zw} = \bar{z}\bar{w} \qquad \overline{\left(\frac{z}{w}\right)} = \frac{\bar{z}}{\bar{w}}$$

$$|z \pm w| \leq |z| + |w| \qquad |zw| = |z||w| \qquad \left|\frac{z}{w}\right| = \frac{|z|}{|w|}$$

Finally we note that the dot product between vectors z, w in the plane is expressed using complex numbers by

$$z \bullet w = \mathrm{Re}(\bar{z}w) = \mathrm{Re}(z\bar{w})$$

We have extended our number universe to complex numbers by taking all combinations of real numbers and imaginary numbers. One of the wonderful results of this expanded view of numbers is that polynomials now have as many zeros as their degree (which is false over real numbers, as the example $P(x) = x^2 + 1$ shows, which is a polynomial of degree 2 without any real zeros, but with the two complex zeros $\pm i$).

Theorem 1 (Fundamental Theorem of Algebra)**.** *Let $P(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0$ be a polynomial with complex coefficients $a_k \in \mathbb{C}$ with $a_n \neq 0$, so $\deg P = n$. Then*

$$P(z) = a_n (z - z_1)^{n_1} \dots (z - z_k)^{n_k}$$

with $z_1, \dots, z_k$ the mutually distinct zeros of $P(z)$ with multiplicities (how often the zero repeats) $n_1, \dots, n_k$ satisfying $n = n_1 + \dots + n_k$. In other words, every polynomial decomposes into linear factors over $\mathbb{C}$ (as compared to the more tedious decomposition over $\mathbb{R}$).

In our simple example $z^2 + 1 = (z + i)(z - i)$, with two distinct zeros each of multiplicity 1. Another example if $P(z) = z^2 + 4z + 13$ whose zeros are $z_1 = -2 + 3i$ and $z_2 = -2 - 3i$, so that $P(z) = (z - (-2 + 3i))(z - (-2 - 3i))$. We notice here that the two zeros $z_2 = \bar{z_1}$ are complex conjugate to each other. This is a general fact for polynomials with *real coefficients*:

Lemma 1. *Let $P(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0$ with $a_k \in \mathbb{R}$ real, then the zeros of $P(z)$ are either real or, if non-real complex, come in complex conjugate pairs. Therefore,*

$$P(z) = a_n (z - x_1)^{n_1} \dots (z - x_k)^{n_k} (z - z_1)^{m_1} (z - \bar{z_1})^{m_1} \dots (z - z_l)^{m_l} (z - \bar{z_l})^{m_l}$$

where $x_1, \dots, x_k \in \mathbb{R}$ are the distinct real zeros, $z_1, \bar{z_1}, \dots, z_k, \bar{z_k} \in \mathbb{C}$ are the distinct complex conjugate pairs of zeros, and $n = n_1 + \dots + n_k + 2m_1 + \dots + 2m_l$.

Proof. Let $z_0 \in \mathbb{C}$ be a zero of $P(z)$ which is *not* real. Then $P(z_0) = 0$ implies also that $\overline{P(z_0)} = 0$, which, since all the coefficient $a_k \in \mathbb{R}$ of $P(z)$ are real, that is, $\bar{a_k} = a_k$, implies $P(\bar{z_0}) = 0$. But this says that both, z_0 and $\bar{z_0}$ are zeros of $P(z)$. □

Notice that for a non-real zero z_0 of a polynomial with real coefficients the term

$$(z - z_0)(z - \bar{z_0}) = z^2 - 2\,\mathrm{Re}(z_0) z + |z_0|^2$$

is a polynomial of degree 2 with real coefficients, but without any real zeros (compare this to the example $P(z) = z^2 + 4z + 13$ above). So the decomposition in the lemma above of a real coefficient polynomial $P(z)$ is exactly the one on the HW sheet about partial fraction decompositions

$$P(z) = a_n (z - x_1)^{n_1} \dots (z - x_k)^{n_k} (z^2 + a_1 z + b_1)^{n_1} \dots (z^2 + a_l z + b_l)^{n_l}$$

where $a_j = -2\,\mathrm{Re}(z_j)$ and $b_j = |z_j|^2$ for $j = 1, \dots, m$.

7. Infinte series

Definition 3. *An infinite series of numbers (which can be real or complex at this stage, we have expanded our world view lately) is an expression of the form*

$$\sum_{k=1}^{\infty} z_k = z_1 + z_2 + z_3 + \dots$$

where the $z_k \in \mathbb{C}$ are complex numbers (note, this includes also real numbers $\mathbb{R} \subset \mathbb{C}$), the summands *or* terms *in the sum. Some series will start at $k = 1$, others at $k = 0$, or sometimes even at $k = N$, for some suitable integer number N.*

Since we can sum up finitely many terms, we do this first, and define the *n-th partial sum* of the infinite series $\sum_{k=1}^{\infty} z_k$ to be

$$s_n := \sum_{k=1}^{n} z_k = z_1 + \cdots + z_n$$

So we can calculate (at least in principle) every partial sum, no matter how big n is. An important example is the n-th partial sum

$$\sum_{k=0}^{n} q^k = \frac{1 - q^{n+1}}{1 - q}$$

of the geometric series $\sum_{k=0}^{\infty} q^k$, which we derived in class. Since this derivation never used the fact that q was a real number (but only the usual arithmetic rules), this formula also holds generally for $q \in \mathbb{C}$.

The crucial notion is that of *convergency* of an infinite series: I think it is intuitively clear that the only reasonable notion for convergency of an infinite series is that we continue summing more an more terms and hope that this process tends to some number, that is, the *sequence* of partial sums $\{s_n\}$ converges to some number $s \in \mathbb{C}$.

Definition 4. *The infinite series* $\sum_{k=1}^{\infty} z_k$ converges *(a better term I think would be "is summable") to some number* $s \in \mathbb{C}$*, if and only if the sequence of partial sums* $s_n \to s$ *as* $n \to \infty$*, that is*

$$\lim_{n\to\infty} \sum_{k=1}^{n} z_k = s$$

To test whether $s_n \to s$*, one reduces this to the test whether* $(s_n - s) \to 0$*, which is equivalent to the distances* $|s_n - s| \to 0$ *as* $n \to \infty$*.*

If a series does not converge, we say it diverges*.*

To diverge does not necessarily mean that the result of summing all terms becomes infinite (which is mostly the case though), but it could be that summing more and more terms will generate some "jumping all over the place" of the partial sums, like in the series $1-1+1-1+1\ldots$, where $s_{2n} = 0$ but $s_{2n+1} = 1$, so that the partial sum sequence is $\{s_n\} = 1, -1, 1, -1, \ldots$, which certainly does not converge to anything, it jumps between 1 and -1.

Make sure you clearly understand the difference between "a sequence converges" and "an infinite series converges", the latter meaning that summing up more and more terms in the series (that is, calculating partials sums s_n for larger and larger n) leads to a finite number s—as I said, summable would be a better terminology for "convergency of a series".

Another important issue to understand is, that when checking the convergency/divergency of an infinite series it does not matter to totally disregard *finitely many terms* of the series. The sum of those finitely many terms will always be a finite number, so only the infinitely many other terms matter. But once you have decided that a series does converge and you want to calculate its value, all terms have to be summed up to get the correct value of the sum.

Finally, the usual rules for adding/subtracting and scaling *convergent* series $\sum_{k=1}^{\infty} z_k$, $\sum_{k=1}^{\infty} w_k$ hold:

$$\sum_{k=1}^{\infty}(z_k \pm w_k) = \sum_{k=1}^{\infty} z_k \pm \sum_{k=1}^{\infty} z_k \quad \text{and} \quad \sum_{k=1}^{\infty} c\, z_k = c \sum_{k=1}^{\infty} z_k$$

for any $c \in \mathbb{C}$. It is important to keep in mind, that these "obvious formulas" only hold for *convergent* series, and generally give non-sensical results for divergent series, as the example $z_k = 1$ and $w_k = -1$ shows.

Before looking at examples, we strengthen the notion of convergency somewhat to the notion of *absolute convergency*, which will play a decisive role for power series:

Definition 5. *An infinite series $\sum_{k=1}^{\infty} z_k$* converges absolutely*, if the infinite series $\sum_{k=1}^{\infty} |z_k|$ of absolute values of each of the terms in the original series converges. Note the latter is a series of* non-negative real numbers*, even when $z_k \in \mathbb{C}$.*

Here a few important

Remarks.

- It can be checked that if a series converges absolutely, then it converges. The converse is false, as the alternating harmonic series shows: $\sum_{k=1}^{\infty}(-1)^{k+1}\frac{1}{k}$ converges by Leibniz Test, but the sum of the absolute value terms $|(-1)^{k+1}\frac{1}{k}| = \frac{1}{k}$ is the harmonic series $\sum_{k=1}^{\infty}\frac{1}{k}$, which diverges (to verify these statements see the various tests below).
- Absolute convergence and convergence are the same for series of non-negative real numbers.
- Most tests mentioned below will test for absolute convergence, except the Leibniz and integral tests, were assumptions on the signs of the series elements are necessary.
- Also note, that complex numbers are not ordered like real numbers, so it makes no sense to compare two complex numbers, other than by their length.

An important example of a convergent series is the geometric series:

Example 1. The geometric series $\sum_{k=0}^{\infty} q^k$ converges absolutely for $|q| < 1$. We have for the partial sums

$$\lim_{n\to\infty}\sum_{k=0}^{n} q^k = \lim_{n\to\infty}\frac{1-q^{n+1}}{1-q} = \frac{1-\lim_{n\to\infty} q^{n+1}}{1-q} = \frac{1}{1-q}$$

provided that $|q| < 1$, since then the repeated powers q^n as $n \to \infty$ get smaller and smaller and converge to zero. Therefore

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1-q} \quad \text{if} \quad |q| < 1$$

Notice, that this says that the geometric series converges for any complex number q inside the disk of unit radius around zero (but not on its boundary) in the complex plane. In fact, this is the largest domain on which the series converges.

8. Convergency tests of infinite series

Theorem 2 (Divergency Test)**.** *If the series $\sum_{k=1}^{\infty} z_k$ converges, then $\lim_{k\to\infty} z_k = 0$. This is often used in the equivalent negation: If $\lim_{k\to\infty} z_k \neq 0$ then the series $\sum_{k=1}^{\infty} z_k$ diverges. Note that $z_k \in \mathbb{C}$ can be complex numbers.*

Important is to understand that $\lim_{k\to\infty} z_k = 0$ *does not* imply that the series $\sum_{k=1}^{\infty} z_k$ converges, as shown by the harmonic series. The Divergency Test is the first test to apply. It either gives you divergency (and thus nothing more needs to be done), or it is inconclusive, at which point you look for another test.

Theorem 3 (Integral Test)**.** *Let $f\colon [1,\infty] \to [0,\infty)$ be a non-negative real valued function $f(x) \geq 0$ which can be integrated over every finite subinterval $[1,b] \in [1,\infty)$ (e.g., any continues function $f(x) \geq 0$ satisfies this assumption). Then the infinite series $\sum_{k=1}^{\infty} f(k)$ converges if and only if the improper integral $\int_1^{\infty} f(x)dx < \infty$, that is, the area under the graph $y = f(x)$ is finite over the whole interval $[1,\infty)$.*

The standard example where this test can be applied to is the series $\sum_{k=1}^{\infty} \frac{1}{n^\alpha}$: it converges for $\alpha > 1$ and diverges for $\alpha \leq 1$.

Theorem 4 (Leibnitz Test)**.** *Consider the alternating series* $\sum_{k=1}^{\infty}(-1)^k a_k$ *where*

(i) $a_k \geq 0$ *and*
(ii) $\lim_{k\to\infty} a_k = 0$.

Then the series $\sum_{k=1}^{\infty}(-1)^k a_k$ *converges.*

The standard example is the alternating harmonic series $\sum_{k=1}^{\infty}(-1)^k \frac{1}{k}$, which sums to ln(2) as we will soon prove.

The last two tests only apply to series of real numbers. The next two criterions, which test for absolute convergency, can be applied to series of complex numbers.

Theorem 5 (Ratio Test)**.** *Let* $\sum_{k=1}^{\infty} z_k$ *be a series of complex numbers. Then the following holds:*

(i) *If* $\lim_{k\to\infty} \frac{|z_{k+1}|}{|z_k|} < 1$, *the series is absolutely convergent.*
(ii) *If* $\lim_{k\to\infty} \frac{|z_{k+1}|}{|z_k|} > 1$, *the series is divergent.*
(iii) *If* $\lim_{k\to\infty} \frac{|z_{k+1}|}{|z_k|} = 1$, *nothing can be concluded.*

Sometimes the successive ratios of absolute values $\frac{|z_{k+1}|}{|z_k|}$ *of the terms of the original series do not form a convergent sequence. In this case a modified, but slightly more tedious, criterion can be formulated:*

(i) *If there exists a number* $0 \leq M < 1$ *so that* $\frac{|z_{k+1}|}{|z_k|} \leq M$ *for all but finitely many* k, *then the series is absolutely convergent. Important here is that* $M < 1$ *is strictly less than* 1.
(ii) *If there exists a number* $M > 1$ *so that* $\frac{|z_{k+1}|}{|z_k|} \geq M$ *for infinitely many* k, *then the series is divergent. Again, it is important that* $M > 1$ *is strictly larger than* 1.

The standard example where this test can be applied is the exponential series $\sum_{k=0}^{\infty} \frac{z^k}{k!}$, which we now also have for complex numbers $z \in \mathbb{C}$: in this case $z_k = \frac{z^k}{k!}$ and thus $\frac{|z_{k+1}|}{|z_k|} = \frac{|z|}{k+1}$. Since $\frac{1}{k+1} \to 0$ as $k \to \infty$, we have

$$\lim_{k\to\infty} \frac{|z_{k+1}|}{|z_k|} = \lim_{k\to\infty} \frac{|z|}{k+1} = |z| \lim_{k\to\infty} \frac{1}{k+1} = 0$$

Since $0 < 1$, we conclude from the Ratio Test that the exponential series converges absolutely for all $z \in \mathbb{C}$. Note that we now can raise any real non-negative number $a \geq 0$ to complex powers via

$$a^z := e^{z \ln a}$$

So we came a long way from $a^n = a \cdots a$ to a^z.

Theorem 6 (Root Test)**.** *Let* $\sum_{k=1}^{\infty} z_k$ *be a series of complex numbers. Then the following holds:*

(i) *If* $\lim_{k\to\infty} \sqrt[k]{|z_k|} < 1$, *the series is absolutely convergent.*
(ii) *If* $\lim_{k\to\infty} \sqrt[k]{|z_k|} > 1$, *the series is divergent.*
(iii) *If* $\lim_{k\to\infty} \sqrt[k]{|z_k|} = 1$, *nothing can be concluded.*

Sometimes the k-th roots of the absolute values $\sqrt[k]{|z_k|}$ *of the terms of the original series do not form a convergent sequence. In this case a modified, but slightly more tedious, criterion can be formulated:*

(i) *If there exists a number* $0 \leq M < 1$ *so that* $\sqrt[k]{|z_k|} \leq M$ *for all but finitely many* k*, then the series is absolutely convergent. Important here is that* $M < 1$ *is strictly less than* 1.
(ii) *If there exists a number* $M > 1$ *so that* $\sqrt[k]{|z_k|} \geq M$ *for infinitely many* k*, then the series is divergent. Again, it is important that* $M > 1$ *is strictly larger than* 1.

The Ratio and Root Tests may give different results when applied to the same series, so if one of the tests is inconclusive, try the other.

9. Power series

We are now applying all the things we know about series of numbers to power series, and we will do it immediately over complex numbers since there is not much of a difference. Suggestive notation (not always religiously adhered to and motivated by $z = x + iy$) is to use z for complex numbers and x for real numbers.

Recall that a power series with (complex) coefficients $a_k \in \mathbb{C}$ centered at $z_0 \in \mathbb{C}$ is just an infinite degree polynomial

$$\sum_{k=0}^{\infty} a_k (z - z_0)^k$$

The first question arising is for which values of z does this series converge? For the two examples we have discussed the geometric series $\sum_{k=0}^{\infty} z^k$ converges only in the interior of the unit disk $|z| < 1$ and the exponential series $\sum_{k=0}^{\infty} \frac{z^k}{k!}$ converges for al $z \in \mathbb{C}$.

In general there are three possibilities when discussing the convergency behavior of a power series $\sum_{k=0}^{\infty} a_k (z - z_0)^k$:

(i) The power series only converges for $z = z_0$, in which case its value is a_0.
(ii) There is some number $R > 0$ so that the power series absolutely converges for $|z - z_0| < R$, that is, inside the disk of radius R centered at z_0 in the complex plane, and diverges for $|z - z_0| > R$, that is, outside this disk.
(iii) The power series converges for all $z \in \mathbb{C}$.

We call the number R the *convergency radius* of the power series. We set $R = 0$ in case (i) and $R = \infty$ in case (iii). Notice that nothing is said what happens along the boundary circle $|z - z_0| = R$. There the power series may converge of diverge.

The convergency radius of a power series $\sum_{k=0}^{\infty} a_k (z - z_0)^k$ is calculated using the Ratio (or Root) Tests as follows: since $z_k = a_k (z - z_0)^k$ we have $\frac{z_{k+1}}{z_k} = \frac{a_{k+1}}{a_k}(z - z_0)$, and therefore the power series converges absolutely provided that

$$\lim_{k \to \infty} |z - z_0| \frac{|a_{k+1}|}{|a_k|} < 1 \quad \text{or equivalently} \quad |z - z_0| < \frac{1}{\lim_{k \to \infty} \frac{|a_{k+1}|}{|a_k|}}$$

and diverges if $|z - z_0| > \frac{1}{\lim_{k \to \infty} \frac{|a_{k+1}|}{|a_k|}}$. Thus, we obtain the following formulas for the convergency radius, depending whether it is computed by the Ratio or Root Test:

Theorem 7 (Convergency radius)**.** *The convergency radius of the power series* $\sum_{k=0}^{\infty} a_k (z - z_0)^k$ *is given by*

$$R = \frac{1}{\lim_{k \to \infty} \frac{|a_{k+1}|}{|a_k|}} \text{ (from Ratio Test)} \quad \textit{and} \quad R = \frac{1}{\lim_{k \to \infty} \sqrt[k]{|a_k|}} \text{ (from Root Test)}$$

One of the reasons power series are important is that every "reasonable" function can be expressed/represented as a power series (even though we have encountered a function in HW 11 which cannot be expressed by a power series). A second reason is that power series can be differentiated and integrated term by term.

Theorem 8. *Let $\sum_{k=0}^{\infty} a_k(z-z_0)^k$ be a power series with convergency radius $R>0$ and define the function*

$$f(z) := \sum_{k=0}^{\infty} a_k(z-z_0)^k$$

that is, the function values $f(z)$ are given by summing up the power series with input z. Then,

(i) $f'(z) = \sum_{k=1}^{\infty} k\, a_k(z-z_0)^{k-1}$ *with same convergency radius $R>0$.*
(ii) $\int f(z)dz = \sum_{k=0}^{\infty} \frac{a_k}{k+1}(z-z_0)^{k+1}$ *with same convergency radius $R>0$.*

In particular, when restricted to the real axis (and assuming the coefficients a_k are real numbers), $f\colon (x_0-R, x_0+R) \to \mathbb{R}$ is an infinitely often differentiable and integrable function whose derivatives and integrals are computed term by term.

This leaves us with the question which functions have power series representations. From (i) in the previous theorem we note that $f'(z_0) = a_1$ and by definition of $f(z)$ we have $f(z_0) = a_0$. Calculating $f''(z)$ term by term, we see that $f''(z_0) = 2a_2$, and generally $f^{(k)}(z_0) = k!a_k$. Therefore, if some arbitrary function $g(z)$ can be expressed as a power series centered at some point z_0 at all, then this series must have the form

$$g(z) = \sum_{k=0}^{\infty} \frac{g^{(k)}(z_0)}{k!}(z-z_0)^k$$

Definition 6. *Given an infinitely often differentiable function $f\colon I \to \mathbb{R}$ on an interval $I \subset \mathbb{R}$ and a point $z_0 \in I$, its* Taylor series *$T(f)(z)$ is the power series*

$$T(f)(z) = \sum_{k=0}^{\infty} \frac{f^{(k)}(z_0)}{k!}(z-z_0)^k\,.$$

How can we decide whether the Taylor series on its interval of convergency is equal to the function, that is, $T(f)(z) = f(z)$ for $|z-z_0| < R$? This is answered by Taylor's Theorem, which expands any infinitely often differentiable function in its Taylor polynomial and a remainder term:

Theorem 9 (Taylor polynomial with remainder term)**.** *Let $f\colon I \to \mathbb{R}$ be an infinitely often differentiable function and let $z_0 \in I$. Then*

$$f(z) = \sum_{k=0}^{n} \frac{f^{(k)}(z_0)}{k!}(z-z_0)^k + \begin{cases} \int_{z_0}^{z} \frac{(z-t)^k}{k!} f^{(k+1)}(t)\,dt \\ \frac{(z-z_0)^{k+1}}{(k+1)!} f^{(k+1)}(\xi) \end{cases}$$

where $\xi \in [z_0, z]$ is between z_0 and z. We call the first expression on the right hand side the Taylor polynomial *of degree n, and the second term the n-th remainder term $R_n(z)$, for which we have two versions. In this theorem everything is on the real number line, even though we notate by z etc.*

From this presentation of a function $f(z)$ we deduce that the remainder term has to tend to zero as $n \to \infty$ in order for the Taylor series to equal the function.

Theorem 10 (Taylor series)**.** *Let $f\colon I \to \mathbb{R}$ be an infinitely often differentiable function and let $z_0 \in I$. Assume that the Taylor series $T(f)(z)$ has convergency radius $R > 0$. Then*

$$f(z) = T(f)(z) = \sum_{k=0}^{\infty} \frac{f^{(k)}(z_0)}{k!}(z-z_0)^k$$

on the interval $|z - z_0| < R$ if and only if $\lim_{n\to\infty} R_n(z) = 0$ for all $|z - z_0| < R$.

Moreover, if the Taylor series converges at any of the boundary points $y = z_0 \pm R$ of the interval $(z_0 - R, z_0 + R)$ then its value at y equals to the function value $f(y)$, that is,

$$f(y) = \sum_{k=0}^{\infty} \frac{f^{(k)}(z_0)}{k!}(y-z_0)^k\,.$$

One example, where this theorem applies, is the Taylor series of the natural logarithm

$$\ln(1+z) = \sum_{k=1}^{\infty}(-1)^{k+1}\frac{z^k}{k}$$

which has convergency radius $R = 1$, and thus absolutely converges for $|z| < 1$. But by the Leibniz Test the series also converges at $z = 1$, so that we obtain

$$\ln(2) = \sum_{k=1}^{\infty}(-1)^{k+1}\frac{1}{k}$$

as the value of the alternating harmonic series.

10. Euler's Formula

As a first application of complex numbers, we investigate the exponential series

$$e^z = \sum_{k=0}^{\infty}\frac{z^k}{k!} \tag{10}$$

Since $|z^k| = |z|^k$ and the series of non-negative real numbers $\sum_{k=0}^{\infty}\frac{|z|^k}{k!}$ converges for any value of $|z|$, we conclude that (10) converges absolutely for all $z \in \mathbb{C}$. Moreover, the same arguments as for the real exponential series show the functional equation

$$e^{z+w} = e^z e^w$$

Next note that

$$\overline{e^z} = \overline{\sum_{k=0}^{\infty}\frac{z^k}{k!}} = \sum_{k=0}^{\infty}(\overline{\frac{z^k}{k!}}) = \sum_{k=0}^{\infty}\frac{\overline{z^k}}{k!} = \sum_{k=0}^{\infty}\frac{\bar{z}^k}{k!} = e^{\bar{z}}$$

and hence, for real $\alpha \in \mathbb{R}$,

$$\overline{e^{i\alpha}} = e^{\overline{i\alpha}} = e^{-i\alpha}\,.$$

Therefore

$$|e^{i\alpha}| = \sqrt{e^{i\alpha}\overline{e^{i\alpha}}} = \sqrt{e^{i\alpha}e^{-i\alpha}} = \sqrt{e^{i\alpha-i\alpha}} = \sqrt{e^0} = 1$$

so that $e^{i\alpha} \in S^1$ is a point on the the circle S^1 of radius 1 centered at the origin. The small miracle happening here is that $e^{i\alpha}$ makes the angle (in radiants) $\alpha \in [0, 2\pi)$ with the x-axis in counter clockwise direction. This means, using the Greek's geometric definition of the trig functions, that we obtain *Euler's Formula*

$$e^{i\alpha} = \cos\alpha + i\sin\alpha \tag{11}$$

which relates the trig functions and the exponential function.

To see why $e^{i\alpha} \in S^1$ has angle α with the x-axis, we view $\gamma(t) = e^{it}$ as a parameterized curve starting at $\gamma(0) = 1$ tracing out the unit circle. Since

$$|\gamma'(t) = |ie^{it}| = 1$$

this curve has unit speed, and therefore its length between 1 and $e^{i\alpha}$ is given by

$$L(\gamma; [0, \alpha]) = \int_0^\alpha |\gamma'(t)|\, dt = \int_0^\alpha dt = \alpha$$

which shows that $e^{i\alpha} \in S^1$ makes angle α in radiants (which is just arclength on the unit circle) with the x-axis.

From our geometric interpretation of $e^{i\alpha}$ and the periodicity of the trig functions, we read off the periodicity of the exponential function on the imaginary line:

$$e^0 = 1 \quad e^{i\frac{\pi}{2}} = i \quad e^{i\pi} = -1 \quad e^{i\frac{3\pi}{2}} = -i \quad e^{i2\pi} = 1$$

and therefore

$$e^{i2\pi n} = 1$$

for all integers $n \in \mathbb{Z}$.

Euler's formula (11) can be used, for example, to derive the power series expansions of the trig functions and the angle sum formulas for the trig functions from the functional equation of the exponential function.

A further application is the polar decomposition of a complex number

$$z = |z|e^{i\alpha}$$

into its length and angle $\alpha \in [0, 2\pi)$ in radiants measured counter clockwise from the x-axis. The polar presentation is particularly useful when multiplying complex numbers: let $z = |z|e^{i\alpha}$ and $w = |w|e^{i\beta}$, then

$$zw = |z||w|e^{i(\alpha+\beta)}\,.$$

This verifies our interpretation of multiplication of complex numbers: multiplication by z rotates w by the angle α which z makes with the x-axis and scales w by the length $!z!$ of z.

Besides their algebraic significance—polynomials have as many zeros as their degree—complex numbers present planar Euclidean geometry: the absolute value $|z|$ is the Euclidean length, addition of complex numbers is vector addition, and multiplication by a complex number is a stretch rotation.

There are two more number systems, in each of which one has to give up familiar properties, which describe Euclidean geometry of Euclidean 3- and 4-space, the *Hamilton numbers* or *Quaternions* $\mathbb{H}$, and the *Cayley numbers* or *Octonions* $\mathbb{O}$, which present Euclidean geometry in 7- and 8-dimensional space. In the case of Quaternions multiplication presents rotation is 3- and 4-space, which we know is non-commutative (e.g., the Rubik's cube). For the Cayley numbers multiplication is non-commutative but also non-associative. One can show that those are the only number systems which extend the real numbers, so we have a tower of number systems

$$\mathbb{R} \subset \mathbb{C} \subset \mathbb{H} \subset \mathbb{O}\,.$$