CHAPTER 1

Pre-requisites

We will go over the pre-requisites for a second course on geometry in the form of exercises which lead to the important results. If you are willing to assume the results stated then you do not need to do the exercises, however, it *is* worthwhile to go over these at least once in your life!

1. Introduction of co-ordinates

Let us assume that we have the usual axiomatic framework of Euclidean geometry. We will show that the points on a line can be given arithmetic operations and identified with the "usual" decimal numbers. Moreover, we can introduce co-ordinates in space using the Euclidean framework. One important thing to note is that use *only* the parallel postulate—congruence (hence distance and angle) play no role in the introduction of co-ordinates.

We are given some gadget that can draw the line joining two points and the line parallel to a given line through a point outside it. Such a gadget is a *ruler with a roller*. (Alternatively you can use the `xfig` program). In the following constructions (see Figures 1 and 2) we have numbered the lines and points in the sequence in which they are obtained. Assume given a pair of points 0 and 1. We can define addition and multiplication for points a and b on the line l joining 0 and 1 by the constructions given below (the final point in each construction is the sum or product of the the two original points).

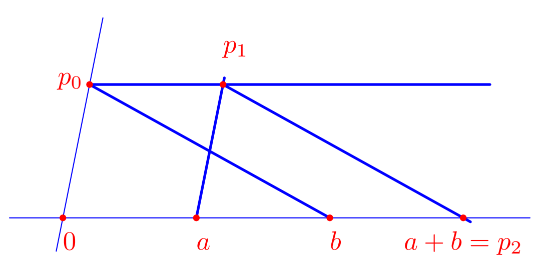

FIGURE 1. Addition

EXERCISE 1. Show that the following (usual) rules of arithmetic hold; in other words the points on a line form a field.

1. The commutative law for addition: $a+b=b+a$.
2. The commutative law for multiplication: $ab=ba$.
3. The associative law for addition: $(a+b)+c=a+(b+c)$.
4. The associative law for multiplication: $a(bc)=(ab)c$.
5. The distributive law: $a(b+c)=ab+ac$.

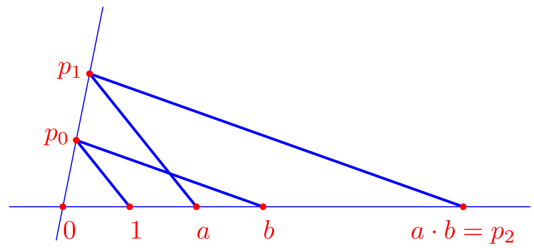

FIGURE 2. Multiplication

6. The identity for addition: $a + 0 = a$.
7. The identity for multiplication: $a \cdot 1 = a$.
8. For any a there is a point $(-a)$ so that $a + (-a) = 0$.
9. For any non-zero a there is a point $(1/a)$ so that $a(1/a) = 1$.
10. If O' and $1'$ are two other points then give a natural correspondence between the points of the line l' joining $0'$ and $1'$ and the line l so that the arithmetic structure is preserved.

In addition, we can use the notion of order on the points of a line to define an order in our arithmetic by saying that a number lies between two other numbers if the corresponding points have the same relation. In particular, we say that $a > 0$ if a is between the points 1 and 0 or if 1 is between a and 0 or if a is 1.

EXERCISE 2. Show in addition that if $a > 0$ and $b > 0$ then $a + b > 0$ and $a \cdot b > 0$.

The following two important axioms are due to Archimedes (but only one carries his name):

AXIOM 1 (Archimedean Property). (Also known as "Big step – Little step") If $x > 0$ (is a Little step) and $y > 0$ (is the Big step) then there is a natural number n (the number of little steps) so that y is less than nx.

The second axiom is perhaps even less "obvious" but is essential.

AXIOM 2 (Least Upper Bound). If A_n is a sequence of points so that for all n, A_{n+1} lies between A_n and D for some fixed point D (i. e. A_n move towards D but do not reach it). Then there is a point B which is the "limit" of A_n. In other words, A_{n+1} is between A_n and B for all n and if C is any other point so that A_{n+1} lies between A_n and C then B lies between A_n and C for all n (see figure 3).

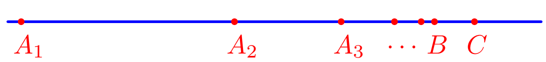

FIGURE 3. The Least Upper Bound

EXERCISE 3. We introduce the decimal representation of a real number as follows.

1. Use the Archimedean Property to show that for any real number x there is an integer n so that $n \leq x < n + 1$. This integer is called the *integer part* $[x]$ of x.

2. Show that the sequence $x_n = [10^n x]/10^n$ is a non-decreasing sequence.
3. Use the Least Upper Bound property to conclude that x_n has a limit y.
4. Using the principle of the excluded middle show that $y = x$.

Finally, we choose four non-coplanar points in space and designate them o, e_1, e_2 and e_3. The point o is called the origin the line through o and e_1 (respectively e_2 or e_3) is called the x-axis (respectively y-axis or z-axis). By drawing lines parallel to the axes we can produce for any point a unique triple of points (x, y, z) one on each axis which uniquely determine the point in space. By the above method we obtain the co-ordinates in decimals as well.

Exercise 4. Show that a line in the plane is the locus of all points with co-ordinates (x, y) such that $ax + by + c = 0$ for some constants a, b and c so that a and b are not both zero. Also show the converse.

2. Conic sections

In the co-ordinate plane one can study more general geometric figures that those desribed by lines. In this section we undertake a rigorous study of conic sections. In particular, we find geometric criteria that distinguish the different conics. We also establish Steiner's construction of conic sections as the locus of intersection of a pair of rotating lines.

The first equation that is more complicated that the equation of a line as given above is one of the form

$$ax^2 + bxy + cy^2 + dx + ey + f = 0$$

where a, b and c are not all zero (in which case the equation would become that of a line). The locus of points (x, y) that satisfy this equation is called a conic or a conic section. By plotting the corresponding curves we find that we have the

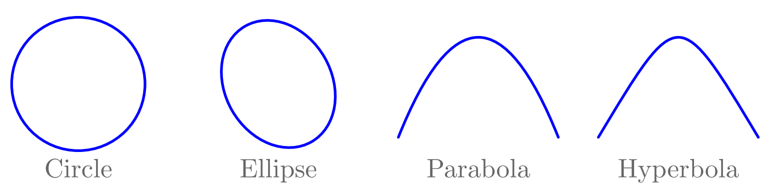

Figure 4. The Smooth Conics

following types of conics:

1. There are no solutions.
2. The solutions all lie on one line.
3. The solutions all lie on a pair of lines.
4. The conic lies within a bounded region of the plane (i. e. the conic is compact). This called an ellipse (of which the circle is a special case).
5. The conic has two parts (i. e. the conic is disconnected).This is called a hyperbola.
6. The conic is connected and not compact. This is called a parabola.

We note that the first three types are distinguished without reference to order among numbers (or separation axioms in geometry) and so make sense over other fields. We will see below how we can distinguish the other conics in a purely algebraic way.

Exercise 5. Find ways of distinguishing the different conics by looking at the equation. (Hint: Examine the discriminant $b^2 - 4ac$).

For any line $ax + by + c = 0$ with a non-zero, we can write the solutions in terms of one parameter as $(-bt - c/a, t)$; similarly when b is not zero. We can also "solve" a conic. Let us suppose that the conic is not of type (1), (2) or (3) above. Fix a point (x_0, y_0) on the conic.

Exercise 6. We will find a parametric solution of a conic. (Hint: Use translation and scaling of co-ordinates to simplify the equations wherever possible).

1. Let $(y - y_0) = t(x - x_0)$ be a line through this point. Show that there is at most one other point of the conic that lines on this line.
2. Find the co-ordinates of this point in terms of the constants a, b, c, d, e, f, x_0, y_0 and the parameter t.
3. Show that this parametric solution is not well defined at two values of t for a hyperbola.
4. Show that this parametric solution misses one point in the case of an ellipse or circle but is well-defined at all values of t.
5. Show that this parametric solution is not well defined for one value of t and misses one point or is well defined and misses no points on a parabola.

This can be carried further through Steiner's construction as follows. Let (x_1, y_1) another point on the conic.

Exercise 7. Show that there are constants A, B, C and D so that for any point (x_2, y_2) of the conic we have

$$s = \frac{At + B}{Ct + D} \text{ when } t = \frac{y_2 - y_0}{x_2 - x_0} \text{ and } s = \frac{y_2 - y_1}{x_2 - x_1}$$

Moreover, these constants are such that if we try to solve for $s = t$, we obtain no solutions when the conic is an ellipse (or circle), one solution for a parabola and two solutions for a hyperbola.

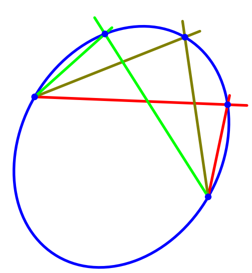

Figure 5. Steiner's Construction

The geometric content of this is the statement that the conic is obtained as the locus of intersection of a pair of rotating lines based at (x_0, y_0) and (x_1, y_1) respectively with respective slopes s and t related by $s(Ct + D) = At + D$.

Exercise 8. Prove the converse that such a locus is always a conic.

3. Polynomials and polynomial functions

We now revise the definition and elementary properties of polynomials and polynomial functions. The fundamental ideas of calculus consist of extending these notions to a larger class of functions.

There are two ways of approaching the concept of function. The ancient way is through formulae, while the modern approach is through the study of functions of sets of points. Most functions that we study arise naturally and can be defined formally (i. e. by formulae or expressions). On the other hand, many of the properties demanded of functions are best defined by thinking of them as set functions. Moreover, most of the formulae have a "life of their own"; the formal expressions have a more general validity than as functions alone. Thus the study of formulae becomes algebra while the study of functions becomes analysis. Calculus is thus seen differently by algebraists and analysts. The fundamental example in both cases is that of a polynomial which we study below.

3.1. Polynomials in one variable. A polynomial P in one variable x is formally defined as a follows

$$P(x) = p_0 + p_1 x + \cdots + p_n x^n$$

where the p_i are constants. If $n = 0$ we identify the polynomial with the constant p_0. If $p_n \neq 0$ then we say the polynomial has degree n. If $p_n = 0$ then we drop the corresponding term unless $n = 0$; the degree of the constant polynomial 0 is considered undefined.

Exercise 9. Define the arithmetic operations on polynomials algorithmically so that polynomial manipulations can be implemented on a computer.

Polynomials can be "evaluated" to give functions; for any choice of a of constant b, we can substitute x by b to obtain the "value" of the polynomial; this gives us the function associated with a polynomial. However, note that when the field of constants is finite (e. g. the field with two elements $\mathbf{F}_2$) a non-constant polynomial might induce a constant function.

Exercise 10. Give an example of a polynomial that is not constant but gives a constant function on the field with three elements.

The points where the function associated with the polynomial vanishes are called solutions or roots of the polynomial.

Exercise 11. Let $P(x)$ be a polynomial of degree n in one variable. The constant b is a root of P (i. e. $P(b) = 0$) if and only if $P(x)$ can be written as a product $(x - b)Q(x)$ where $Q(x)$ has degree $n - 1$. Hence or otherwise show that $P(x)$ has at most n roots.

The polynomial $x^2 + 1$ has no roots over the field of real numbers. The field of complex numbers is collection of numbers of the form $a + b\sqrt{-1}$; where a and b are real numbers.

Exercise 12. Define the arithmetic operations on complex numbers algorithmically in terms of the arithmetic operations on real numbers. Show that any quadratic polynomial $x^2 + ax + b$ is a product $(x - d)(x - c)$ over the field of complex numbers (d and c need not be distinct).

The Fundamental Theorem of Algebra states that *any* polynomial over the field of real numbers is a product of linear and quadratic terms upto a non-zero constant multiple. Thus any polynomial with real coefficients has all its roots over complex numbers.

EXERCISE 13. Assuming the fundamental theorem of algebra show that any polynomial with complex coefficients has all its roots over complex numbers.

In the above discussion, we used the phrase "has all its roots" as a synonym for writing the polynomial as a product of linear terms. Now if some of these factors repeat then we say we have repeated or multiple roots. In particular, we can say that $(x-b)^k$ vanishes k times at b.

EXERCISE 14. Use the Binomial theorem to write a polynomial of P degree n as follows

$$P(x) = \tilde{p}_0 + \tilde{p}_1(x-b) + \cdots + \tilde{p}_n(x-b)^n$$

for some constants $\tilde{p}_i$.

Thus we can say that a polynomial vanishes to order k at b if the terms in the above expression for it satisfy $\tilde{p}_i = 0$ whenever $i < n$.

EXERCISE 15. If P and Q vanish to order k at b, then so does $R \cdot P + Q$ for any polynomial R.

Note that $\tilde{p}_0$ is the value $P(b)$ of P at b. Moreover, the term $\tilde{p}_1$ depends on P *and* on b that we will now to determine. Suppose that we have

$$\begin{aligned} P(x) &= \tilde{p}_0 + \tilde{p}_1(x-b) + \text{ terms that vanish to order 2 at } b \\ Q(x) &= \tilde{q}_0 + \tilde{q}_1(x-b) + \text{ terms that vanish to order 2 at } b \\ R(x) &= \tilde{r}_0 + \tilde{r}_1(x-b) + \text{ terms that vanish to order 2 at } b \end{aligned}$$

The above algebraic property of vanishing to order n shows us that

$$\begin{aligned} R \cdot P + Q = (\tilde{p}_0\tilde{q}_0 + \tilde{r}_0) + (\tilde{r}_0\tilde{p}_1 + \tilde{r}_1\tilde{p}_0 + \tilde{q}_1)(x-b) \\ + \text{ terms that vanish to order 2 at } b \end{aligned}$$

Thus, if we denote the rule that associates the constant $\tilde{p}_1$ with P as $(d/dx)_{|b}$, then this satisfies

$$(d/dx)_{|b}(R \cdot P + Q) = (d/dx)_{|b}(R)P(b) + R(b)(d/dx)_{|b}(P) + (d/dx)_{|b}(Q)$$

Such a rule (which takes polynomial to constants) is called a (constant) *derivation*. If we formally replace b by the variable x we obtain the requirement for a polynomial derivation (which takes polynomial to polynomial). A polynomial derivation is a rule D which associates to every polynomial P another polynomial $D(P)$ so that

$$D(R \cdot P + Q) = R \cdot D(P) + D(R) \cdot P + D(Q).$$

and $D(c) = 0$ for a constant polynomial c.

EXERCISE 16. For any derivation D and any polynomial P show that $D(P^n) = nP^{n-1}D(P)$. (Hint: Use induction). Hence or otherwise show that a derivation is determined on all polynomials once one knows what it does to the variable x.

In particular, for there is a derivation which takes x to 1. This is called the derivative with respect to x and is denoted by dP/dx.

EXERCISE 17. Show that the value of dP/dx at b is $(d/dx)_{|b}(P)$.

The relation between derivations and the order of vanishing is given by

EXERCISE 18. If P vanishes to order n at b then $D(P)$ vanishes to order $n-1$ (for any derivation D).

One of the aims of calculus is to find a larger class of functions which can be studied in a manner analogous to that given above for polynomials. To do this we need to generalise the notion of "vanishing to order n" and derivations.

A simple way to enlarge the class is to consider "rational functions", constructed from polynomials the same way as fractions are from natural numbers. A rational function is of the form P/Q where P and Q are polynomials with $Q \neq 0$.

EXERCISE 19. Extend the definitions of the arithmetic operations and d/dx to rational functions.

Let P and Q be any polynomials. The usual division algorithm allows us to write an expression $P = RQ + S$, where R and S are polynomials and the degree of S is less than that of Q or S is zero. Using this and the fundamental theorem of algebra it is not difficult to show

EXERCISE 20. Every rational function over real numbers is the sum of terms of the form

$$\frac{a}{(x-b)^n} \text{ and/or } \frac{ax+b}{((x-c)^2+d^2)^n}$$

This the called the partial fraction expansion.

Given a polynomial P consider the problem of trying to find a polynomial Q so that $dQ/dx = P$. This is quite easily solved using the fact (proved above) that $d(x^n)/dx = nx^{n-1}$. When the problem is posed for rational functions it becomes a bit harder.

EXERCISE 21. Show that

$$\begin{aligned} \frac{d}{dx}\frac{1}{(x-b)^n} &= \frac{-n}{(x-b)^{n+1}} \\ \frac{d}{dx}\frac{1}{((x-c)^2+d^2)^n} &= \frac{-2n(x-c)}{((x-c)^2+d^2)^{n+1}} \end{aligned}$$

Hence given any rational function P/Q over reals the only hurdle to solving the problem of finding a function f so that $df/dx = P/Q$ is to solve this when P/Q is either $1/(x-b)$ or $(ax+b)/((x-c)+d^2)$. In the section on integration we will see how these problems can be solved.

3.2. Polynomials in more than one variable. Let us begin by considering polynomials in two variables x and y

$$P(x,y) = p_{0,0} + p_{1,0}x + p_{0,1}y + p_{2,0}x^2 + p_{1,1}xy + p_{0,2}y^2 + \cdots + \\ + \cdots + p_{n,0}x^n + \cdots + p_{0,n}y^n$$

where $p_{k,l}$ are constants. The degree of a term $p_{k,l}x^k y^l$ is defined as $k+l$ if $p_{k,l}$ is non-zero. The degree of a polynomial is the maximum of the degrees of each of its non-zero terms and the degree of the zero polynomial is undefined as before.

Exercise 22. Define the arithmetic operations on polynomials algorithmically so that polynomial manipulations can be implemented on a computer.

Polynomials can be "evaluated" to give functions; for any choice of a pair of constants (a, b), we can substitute x by a and y by b to obtain the "value" of the polynomial. By identifying the pair of constants (a, b) with the corresponding point in the plane, this gives us the function (on the plane) associated with the polynomial. The points where the function associated with the polynomial vanishes are called solutions of the polynomial.

In particular, a polynomial $P(x, y)$ as above has the origin $(0, 0)$ as a solution only if the constant term $p_{0,0}$ is zero. More generally, we say a polynomial vanishes to order m at the origin $(0, 0)$ if all its terms have degree at least m.

The Binomial theorem allows us to extend this notion to points (a, b) other than the origin.

Exercise 23. Use the Binomial theorem to write a polynomial of degree n as a sum of terms of the form

$$\tilde{p}_{k,l}(x - a)^k (y - b)^l$$

where $k + l$ is at most n.

Thus we can say that a polynomial vanishes to order m at (a, b) if the terms in the above expression for it satisfy $k + l \geq m$ whenever the coefficients $\tilde{p}_{k,l}$ are non-zero.

Exercise 24. If P and Q vanish to order m at a point, then so does $R \cdot P + Q$ for any polynomial R.

Because of this we can work with arithmetic operations on polynomials "modulo" terms that vanish to order m at a given point (a, b). In particular, any polynomial is like a linear polynomial upto terms that vanish to order two.

$$P(x, y) = \tilde{p}_{0,0} + \tilde{p}_{1,0}(x - a) + \tilde{p}_{0,1}(x - b) + \\ \text{terms that vanish to order 2 at } (a, b))$$

As before $\tilde{p}_{0,0}$ is the value of P at (a, b).

Exercise 25. If we denote the rule $P \mapsto \tilde{p}_{0,0}$ by $(\partial/\partial x)_{|(a,b)}$ then check that this is a constant derivation.

We also have polynomial derivations and as before

Exercise 26. Any polynomial derivation is determined by what it does to the two variables x and y.

In particular we have $(\partial/\partial x)$ which is defined as the derivation that sends x to 1 and y to 0; $\partial/\partial y$ is defined by symmetry.

Exercise 27. Repeat this subsection replacing two variables x and y with n variables $(x_1, \ldots, x_n)$.

4. Sequences

In this section we revise the notion of convergence for real numbers and prove the Bolzano-Weierstrass property.

From section 1 we have a least upper bound (greatest lower bound) for any bounded increasing (respectively decreasing) sequence of real numbers.

EXERCISE 28. Show that any bounded non-decreasing sequence of real numbers has a least upper bound; a bounded non-increasing sequence has a greatest lower bound.

Now if S is any bounded non-empty set of real numbers, let c_1 be an upper bound and $s_1 \in S$. Now we iteratively define a pair of sequences $\{s_n\}$ and $\{c_n\}$ as follows. If $(s_n + c_n)/2$ is an upper bound for S then we define $s_{n+1} = s_n$ and $c_{n+1} = (s_n + c_n)/2$; otherwise let s_{n+1} be any element of S so that $s_{n+1} > (s_n + c_n)/2$ and $c_{n+1} = c_n$.

EXERCISE 29. Show that $\{s_n\}$ is a non-decreasing sequence of elements of S and $\{c_n\}$ is a non-increasing sequence of upper bounds for S so that $(c_{n+1} - s_{n+1}) \leq (s_n - c_n)/2$. Hence show that the greatest lower bound of $\{c_n\}$ is equal to the least upper bound of $\{s_n\}$ and this bound is the least upper bound for S.

In particular, if $\{x_n\}$ is any bounded sequence of real numbers we have a least upper bound and greatest lower bound for this sequence. Let us define

$$\begin{aligned} l_k &= \text{the greatest lower bound of } \{x_n | n \geq k\} \\ u_k &= \text{the least upper bound of } \{x_n | n \geq k\} \\ \liminf\{x_n\} &= \text{the least upper bound of } \{l_k\} \\ \limsup\{x_n\} &= \text{the greatest lower bound of } \{u_k\} \end{aligned}$$

Note that $\{l_k\}$ is a non-decreasing sequence and $\{u_k\}$ is a non-increasing sequence.

EXERCISE 30. Show that $\liminf\{x_n\} \leq \limsup\{x_n\}$.

We say that the sequence $\{x_n\}$ has a limit (is convergent) if these two numbers are equal; this number c is called the limit of this sequence of numbers.

EXERCISE 31. Show that for every positive ϵ there is a index n_0 so that $|x_n - c| < \epsilon$ for all $n > n_0$. Hence, there is an index n_1 so that $|x_n - x_m| < \epsilon$ for all $n > n_1$; this called Cauchy's criterion. Conversely show that any sequence $\{x_n\}$ that satisfies Cauchy's criterion is convergent.

Now for any sequence $\{x_n\}$ we can find subsequences $\{y_k = x_{n_k}\}$ (with $n_1 < n_2 < \cdots$) so that $\{y_k\}$ is convergent (this is called the Bolzano-Weierstrass property).

EXERCISE 32. Show that there are subsequences of $\{x_n\}$ that converge to $\liminf\{x_n\}$ and $\limsup\{x_n\}$.

Finally, let us note some algebraic properties of convergent sequences.

EXERCISE 33. Show that the sum, difference and product of convergent sequences is limit of the sum, difference and product of the terms. If a sequence has a non-zero limit then show that the inverse of the limit is the limit of the inverses of the non-zero terms of the sequence.

5. Functions, continuity and differentiability

In this section we will revise the definition and elementary properties of differentiable functions (n times differentiable functions).

5.1. Definitions. The study of differentiable functions is the study of functions that mimic the behaviour of polynomials "approximately". To begin with we must formally define the notion of approximation.

Exercise 34. For any real number $0 < x < 1$ show that x^n is a decreasing sequence with limit 0.

In particular, we see that a polynomial that vanishes to order $(n+1)$ at 0 satisfies the following condition on functions of one variable.

Definition 1. A function $g(x)$ of one variable is said to be in $o(x^n)$ if for any $\epsilon > 0$ there is a $\delta > 0$ so that

$$|g(x)| < \epsilon |x^n| \text{ for all } x \text{ so that } |x| < \delta.$$

An alternative notion is

Definition 2. A function $g(x)$ one variable is said to be in $O(x^n)$ is there is a $\delta > 0$ and a constant C so that

$$|g(x)| < C|x^n| \text{ for all } x \text{ so that } |x| < \delta$$

Clearly, any polynomial that vanishes to order n is $O(x^n)$. Further, it is clear that an function $g(x)$ that is $O(x^n)$ is $o(x^{n-1})$ and any function that is $o(x^n)$ is $O(x^n)$.

We can extend these notions to many variables as well. A function $g(x_1, \dots, x_n)$ of n variables is said to be in $o(\mathbf{x}^n)$ (respectively $O(\mathbf{x}^n)$) if for all lines $(x_1, \dots, x_n) = (xc_1, \dots, xc_n)$ through the origin the restricted function $f(x) = g(xc_1, \dots, xc_n)$ is in $o(x^n)$ (respectively $O(x^n)$). We can further extend this to define $o((\mathbf{x}-\mathbf{b})^n)$ and $O((\mathbf{x}-\mathbf{b})^n)$ where $\mathbf{b} = (b_1, \dots, b_n)$ is some point, as a way of approximating functions near this point.

We say that g and f agree upto $o((\mathbf{x}-\mathbf{b})^n)$ (or f approximates g upto $o((\mathbf{x}-\mathbf{b})^n)$) if $f-g$ is in $o((\mathbf{x}-\mathbf{b})^n)$. Note in particular, that f and g take the same value at $\mathbf{b}$.

A function is differentiable n times at the point $\mathbf{c}$ if it is approximated upto $o((\mathbf{x}-\mathbf{b})^n)$ by a polynomial (of degree n). Clearly, a polynomial of any degree is differentiable by the results of the previous section. In the one variable case we write this as follows

$$f(x) = a_0 + a_1 x + \cdots + a_n x^n + o((x-c)^n)$$

Exercise 35. Show that for any two functions f and g in $o(\mathbf{x}^n)$ and a function h which is differentiable n times at the origin, the function $h \cdot f + g$ is in $o(\mathbf{x}^n)$.

Exercise 36. Show that the numbers a_k are uniquely determined by the function f.

Now the number a_1 depends on f and the point c. Now suppose that f is differentiable (1 times) at all points c so that it can be written as above near every point c. Then we can define the derived function f' by letting $f'(c) = a_1$ for each point c; the function f' is also called the derivative of f. Now it clear that if f is the function given by a polynomial P then f' is dP/dx. Thus we also use the notation df/dx for f'. We have the derivation property as well.

EXERCISE 37. If f, g and h are differentiable then so is $hf + g$ and

$$\frac{d}{dx}(hf + g) = \frac{dh}{dx} f + h\frac{df}{dx} + \frac{dg}{dx}$$

However, unlike the condition of vanishing to order n at c, the condition $o((x - c)^n)$ is not very well behaved.

EXERCISE 38. Show that $f(x) = x^2 \sin(1/x)$ is $o(x)$ but the derivative of f' is not $o(x^0)$.

A function $f(\mathbf{x})$ is called continuous at a point $\mathbf{c}$ if $f(\mathbf{x}) - f(\mathbf{c})$ is $o(\mathbf{x} - \mathbf{c})$ (i. e. it is differentiable 0 times!). It is called continuous it it has this property at all points. Thus we would like to study functions f which are differentiable and in addition the derivative f' is continuous. Such functions are provided by the fundamental theorem of calculus.

5.2. Properties. We show the important properties called the intermediate value property and extremal value property of continuous functions. We also deduce the clutch of theorems called mean value theorem, Rolle's theorem and so on for differentiable functions.

The following important property of continuous functions will be used all the time.

EXERCISE 39. Let $f(x)$ be a continuous function and $\{x_n\}$ be a sequence converging to c, then $\{f(x_n\}$ is a sequence converging to $f(c)$. (Hint: Examine the condition for continuity near c).

Let $f(x)$ be continuous for x satisfying $a \leq x \leq b$. Let c be a real number lying between $f(a)$ and $f(b)$ we want to show that c is a value of f; in other words any number intermediate to two values is itself a value.

EXERCISE 40. Let s be the least upper bound of the set

$$\{x | a \leq x \leq b \text{ and } f(x) \leq c\}$$

Show that $f(s) = c$.(Hint: To show that $f(s) \geq c$ take a sequence of points approaching s from above).

Now let C be the least upper bound of the values of $f(x)$, i. e. it is the least upper bound of the set $\{f(x)|a \leq x \leq b\}$. The C is an extremal value for f.

EXERCISE 41. Show that $C = f(x)$ for some x in the range $a \leq x \leq b$. (Hint: We have a sequence $\{x_n\}$ so that $f(x_n)$ converges to C; by the section on sequences this sequence has a convergent subsequence).

The following property of differentiable functions is very important

EXERCISE 42. Let $f(x)$ be differentiable for x satisfying $a \leq x \leq b$. Let s be such that $f(s)$ is an extremal value for f. Then $f'(s) = 0$. (Hint: Examine the condition for differentiability near s).

Now suppose that $f(x)$ is differentiable in the range $a < x < b$ and continuous at the endpoints a and b as well. Suppose that $f(a) = f(b) = 0$. There is a point s where f attains its maximal value; similarly there is a point t where f attains its minimal value. If s is the point a or b then $f(x) \leq 0$ and if t is the point a or b then $f(x) \geq 0$. Thus in case both of these occur then $f(x) = 0$ for all x; then let $c = (a + b)/2$. Otherwise let c be any one of s and t which is not a or b. Thus we have a point where $f'(c) = 0$.

Exercise 43. For a general function $g(x)$ which is differentiable in the range $a < x < b$ and continuous at the endpoints we apply this to the function

$$f(x) = g(x) + \frac{g(b) - g(a)}{b - a} \cdot (x - a)$$

to show that there is a point c where

$$g'(c) = \frac{g(b) - g(a)}{b - a}$$

6. New Functions from old

There are three primary ways in which new functions can be "constructed" from old ones. The first and most familiar one is integration. The second method is the inversion of a monotone function. The third is the implicit definition by an equation in two variables.

Formally, the integral of a function f is a function denoted by $\int f$ which satisfies $d(\int f)/dx = f$; note that such a function is only determined upto a constant since a constant has derivative 0.

Similarly, if $f(x)$ is a function, its formal inverse is a function $g(y)$ so that $g \circ f =$ identity. Finally, if $f(x, y)$ is a function of two variables, we can look for a function $g(x)$ so that $f(x, g(x)) = 0$ identically.

Algebraically this is all we need. We can easily compute the values of derivatives at various orders of the new functions in terms of those of the old functions.

Analytically, we need to show that such functions exists under certain reasonable conditions on the given data. For exmaple, we have already constructed an integral for a polynomial function.

Here is an example of inversion:

Exercise 44. Let $f(x) = x^3$ show that there is a function $g(y)$ so that $g(f(x)) = x$ for all x.

And an example of an implicit function:

Exercise 45. Let $f(x, y) = x^2 + y^2 - 1$ show that there is a function $g(x)$ for $0 < x < 1/2$ so that $f(x, g(x)) = 0$.

In the sections below, we shall construct such solutions in greater generality.

6.1. Inverse functions. We prove the inverse function theorem. A continuously differentiable function which has non-zero derivative at a point has an inverse in a neighbourhood of that point (see Figure 6). A function is called *monotone* if it preserves or reverses order over the domain of its definition. In other words if $(f(x) - f(y))(x - y)$ does not change sign for all x and y in the domain of definition of f. We say f is *strictly* monotone if in addition the above product is zero only when $x = y$. It is clear that a strictly monotone function is one-to-one.

Exercise 46. Let f be monotone and *continuous* in for x in the interval $a < x < b$. Show that there is a function g on the range of f so that $g(f(x)) = x$ for all x in the interval $a < x < b$. (Hint: Use the intermediate value theorem and extremal value theorem).

One way to ensure that a function is monotone is to use Rolle's theorem in reverse to prove:

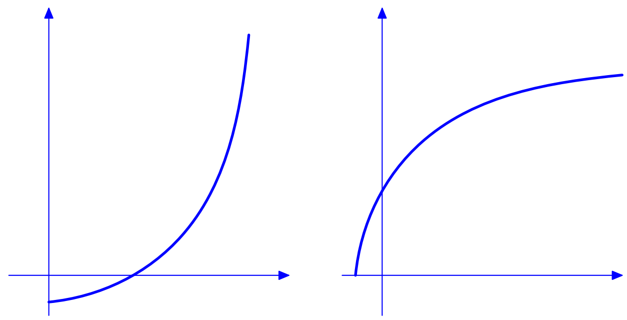

FIGURE 6. The Inverse Function

EXERCISE 47. If f is continuously differentiable and $f'(x_0) \neq 0$ then show that f is monotonic in some interval around x_0. Hence show that f has a inverse g (as in the exercise above) in some small enough interval around $f(x_0)$.

We can also compute the formal inverse

EXERCISE 48. If f can be expressed as

$$f(x) = f(x_0) + f_1(x - x_0) + \cdots + f_n(x - x_0)^n + o((x - x_0)^n)$$

with $f_1 \neq 0$, then show that the inverse function $g(y)$ has the following form where $y_0 = f(x_0)$.

$$g(y) = x_0 + \frac{1}{f_1}(y - y_0) - \frac{f_2}{f_1^3}(y - y_0)^2 + \cdots + g_n(y - y_0)^n + o((y - y_0)^n)$$

where g_n is of the form $P_n(f_1, \ldots, f_n)/f_1^{n+1}$, where P_n is a polynomial function.

6.2. Implicit functions. We will prove the implicit function theorem.

Let $f(x, y)$ be a continuous function of two variables so that it is continuous, and differentiable with respect to y when x is kept fixed; in particular, we have expression

$$f(x, y) = f_0(x, y_0) + f_1(x, y_0)(y - y_0) + o_x(y - y_0)$$

where the subscript in the o_x denotes the dependence of the condition on x. The function $f_1(x, y)$ is denoted by $\partial f/\partial y$ and is called the partial derivative of f with respect to y (we have already seen this for the case of polynomials). We further assume that f_1 is continuous.

Now suppose that $f(a, b) = 0$ and $c = (\partial f/\partial y)(a, b) \neq 0$. We want to find the implicit function defined by $f = 0$. We do this by shoing that for each x near a there is a unique y so that $f(x, y) = 0$. Equivalently, we need to show that the function $g(x, y) = y - f(x, y)/c$ has a unique fixed point for any chosen x near a.

Contractions give rise to functions with a unique fixed point.

EXERCISE 49. Let $0 < c < 1$ be a constant. Let $g(x)$ be a function so that g send the interval $[a, b]$ to itself and $|g(x) - g(y)| < c|x - y|$ for all x and y in $[a, b]$. Show that $g(x_0) = x_0$ for exactly one point x_0 in $[a, b]$.

Consider the function $g(x, y) = y - f(x, y)/c$ as above; we have $g(a, b) = b$ and $(\partial g/\partial y) = 0$. For each fixed x we would like $g(x, y)$ to be a contraction on some interval around b. As a first step:

EXERCISE 50. Using the continuity of f_1 show that there is an interval $[a-r, a+r]$ around a and an interval $[b-s, b+s]$ around b so that $|(\partial g/\partial y)(x, y)| \leq 1/2$ for x and y in these respective intervals. (Hint: Write $\partial g/\partial y$ in terms of f_1 to show that it is continuous). In particular, by the mean value theorem show that $|g(x, y) - g(x, y')| \leq 1/2(y - y')$ on these intervals.

Now by the continuity of f (and thus of g) we can choose a smaller r so that $|g(x, b) - g(a, b)| \leq s/2$ for x in the interval $[a-r, a+r]$. Since $g(a, b) = b$, it follows that $g(x, y) = b + (g(x, y) - g(x, b)) + (g(x, b) - g(a, b))$ lies in the interval $[b-s, b+s]$. Applying the above exercise it follows that for every x there is a unique point y so that $g(x, y) = y$ or equivalently $f(x, y) = 0$. We denote this point y as $h(x)$. This function h is the required implicit function.

We have the identity,

$$\begin{aligned} h(x) - h(x') = g(x, h(x)) - g(x', h(x')) = \\ (g(x, h(x)) - g(x', h(x))) + (g(x', h(x)) - g(x', h(x'))) \end{aligned}$$

Applying the mean value theorem to $g(x', y)$ we obtain

$$|h(x) - h(x')| \leq |g(x, h(x)) - g(x', h(x)| + \frac{1}{2}|h(x) - h(x')|$$

EXERCISE 51. Use the above inequality and the continuity of $g(x, y_0)$ for every fixed y_0 to conclude the $h(x)$ is continuous.

Now if we assume in addition that f is differentiable to order k then it follows that h is also differentiable to order k by an entirely similar reasoning to the one in the above exercise.

EXERCISE 52. If $f(x, y)$ has the form (near the point (a, b))

$$\begin{aligned} f(x, y) = f_{1,0}(x-a) + f_{0,1}(y-b) + \\ f_{2,0}(x-a)^2 + 2f_{1,1}(x-a)(y-b) + f_{0,2}(y-b)^2 + o((x-a, y-b)^2) \end{aligned}$$

with $f_{0,1} \neq 0$, then show that the implicit function $g(x)$ has the form (near $x = a$),

$$g(x) = b - \frac{f_{1,0}}{f_{0,1}}(x-a) + \frac{f_{0,2}f_{1,0}^2 + 2f_{1,1}f_{1,0}f_{0,1} - f_{2,0}f_{0,1}^2}{f_{0,1}^3}(x-a)^2 + o((x-a)^2)$$

6.3. Integration. We develop the theory of integration of continuous functions and prove the fundamental theorem of calculus.

Let R be any (bounded) region in the plane which we want to measure the area of. We can tile the plane with squares of unit length and count the number of such squares that are contained in the region to obtain an approximation to the area from below. On the other hand we can count the number of squares that meet to region to obtain an approximation to the area from above. We can repeat this with squares of smaller size and appropriately scale the count it seems clear that the approximant from below will increase and the approximant from above will decrease. The least upper bound of the former is called the inner measure and the greatest lower bound of the latter the outer measure. To obtain an area for the region we must show that these two numbers are the same; moreover, we would like these numbers to be independent of the placement of the grid as well as rotation and/or shearing of the grid.

EXERCISE 53. Show that the measure of any rectangle is the product of the two sides. More generally, show that the area of a parallelogram is the product of the height and the base.

Let $f(x)$ be a (non-negative) continuous function for $a \leq x \leq b$. Let R be the region bounded by the lines $x = a$ and $x = b$ on the left and right, by the x-axis below and the curve $y = f(x)$ at the top. By the previous exercise we can calculate the measure by using rectangles instead of squares. We do this in the following exercise.

EXERCISE 54. By a partition P of the interval $[a, b]$ we mean a (finite) collection of points $a = t_0 < t_1 < \cdots < t_n = b$. For any such partition we define m_i (respectively M_i) to be the minimum (respectively maximum) value of $f(x)$ for $t_{i-1} \leq x \leq t_i$.

1. Show that the sum $L(P, f)$ (respectively $U(P, f)$) approximate the area of the region from below (respectively above), i. e. are sums of areas of rectangles enclosed by (respectively enclosing) the region R.

$$\begin{aligned} L(P, f) &= \sum_{i=1}^{n} m_i(t_i - t_{i-1}) \\ U(P, f) &= \sum_{i=1}^{n} M_i(t_i - t_{i-1}) \end{aligned}$$

2. If P' is a finer partition than P (i. e. each point of P is also a point of P') then show that

$$L(P, f) \leq L(P', f) \leq U(P', f) \leq U(P, f)$$

3. Let P_n denote the partition of $[a, b]$ into n equal parts. Show that

$$\sup\{L(P, f) | P \text{ a partition }\} \geq \sup\{L(P_n, f)\}$$

   Similarly for the infimum of the $U(P, f)$,

$$\inf\{U(P, f) | P \text{ a partition }\} \leq \inf\{U(P_n, f)\}$$

4. Let $i = i(P, f)$ be such that the difference $M_i - m_i$ is maximum. Then show that

$$U(P, f) - L(P, f) \leq (M_i - m_i)(b - a)$$

   Let $x(P, f)$ denote the mid point of the interval $[i_{i-1}, t_i]$ for this i.
5. Let c be any point of the interval $[a, b]$. For any positive ϵ, show that there is a $\delta > 0$ so that the difference between the maximum and minimum values of $f(x)$ on the interval $[c - \delta, c + \delta]$ is less than $\epsilon/(b - a)$. (Hint: use continuity of f at c).
6. The sequence $\{x_n = x(P_n, f)\}$ has a convergent subsequence $\{y_k = x_{n_k}\}$, with limit point c. Show that there is a k_0 so that if $k \geq k_0$ and $i = i(P_{n_k}, f)$ then the entire sub-interval $[t_{i-1}, t_i]$ of the partition P_{n_k} is contained in $[c - \delta, c + \delta]$.
7. Deduce that $\sup\{L(P_{n_k}, f)\} = \inf\{U(P_{n_k}, f)\}$.
8. Conclude that the inner and outer measure of the region R coincide.

This show that the area of the region R is well-defined. It is denoted by $\int_a^b f$ to denote its depends on a, b and f (we will justify the use of the $\int$ symbol below. First of all note that

EXERCISE 55. Let $a < c < b$ and $f(x)$ and $g(x)$ be (positive) functions that are continuous in the interval $[a, b]$; let $d > 0$ be any positive constant. Then we have

$$\int_a^c f + \int_c^b f = \int_a^b f$$

and

$$\int_a^b (d \cdot f + g) = d \cdot \int_a^b f + \int_a^b g$$

Due to this we are justified in extending the definition as follows. If $f(x)$ is continuous in the interval $[a, b]$ and $c > d$ are points in this interval we define $\int_c^d f := -\int_d^c f$. Moreover, if f is not everywhere positive we define

$$f_+(x) = \max\{f(x,), 0\} = (f(x) + |f(x)|)/2$$

and

$$f_-(x) = \max\{-f(x,), 0\} = (-f(x) + |f(x)|)/2$$

Then clearly $f_\pm$ are positive functions and $f = f_+ - f_-$. We then define the integral of f by the formula $\int_a^b f := \int_a^b f_+ - \int_a^b f_-$. This definition too is justified by the additive property given above.

Now if f is continuous on an interval $[a, b]$ where its minimum value is m and its maximum value is M it is clear that

$$m(d - c) \leq \int_c^d f \leq M(d - c)$$

EXERCISE 56. Let $g(x) = \int_a^x f$, then g is a differentiable function for all points x so that $a < x < b$ and its derivative is f. (Hint: Apply the addition rule and use continuity of f at x)

This justifies the use of the $\int$ symbol. We have shown how to analytically compute the function which we formally defined above. In addition we have a way of constructing a function which is differentiable and its derivative is continuous—a problem that was raised in the section on continuous and differentiable functions.

7. Curves

Now that we have a basic understanding of differentiability of functions we can begin the study of plane curves which can be defined parametrically or as the locus of vanishing of a function of two variables. At the very least we need the function to be (piecewise) differentiable with "good" first order properties. We shall see later that for any interesting (and æsthetic!) study we will need second derivatives as well.

A plane curve is defined (locally) as the locus of points where a "good" function f of two variables vanishes. In particular, let $p = (a, b)$ be a point where f vanishes, then we assume that $f(x, y) = f_{1,0}(x - a) + f_{0,1}(y - b) + o(\mathbf{x})$ is *continuously differentiable* at this point. The curve is said to be *singular* at p if both the above coefficients are 0; otherwise we call the curve *non-singular* or *smooth*. For a smooth curve through p the line $f_{1,0}(x-a)+f_{0,1}(y-b)=0$ is called the tangent line. It is the "best" linear approximation to the curve in an obvious way:

EXERCISE 57. Consider the natural parametrisation of the line $\alpha(x-a)+\beta(y-b)=0$. The restriction of the function f to this line can then be thought of as a function of one variable. Show that this function vanishes to order 2 if and only if the line is the one above *or* the curve $f=0$ is singular at the point $p=(a,b)$.

In what follows we restrict our attention to smooth curves. Singular curves are very interesting and are studied extensively in algebraic geometry.

A different way of representing curves is to think of a curve as a "moving point". A curve can be given in parametric form by writing a pair of functions $(x(t), y(t))$ so that as t varies we will trace out a curve. As before we will insist on the two functions being continuously differentiable. We say that our curve is non-singular at "time" $t=t_0$ we need at least of the pair $(x'(t_0), y'(t_0))$ to be non-zero; otherwise we call the curve singular.

EXERCISE 58. Consider the function of t given by substituting the above pair of functions in the linear form $l(x,y)=\alpha x+\beta y+\gamma$. Show that this function vanishes to order 2 at $t=t_0$ if and only if the curve is singular *or* the form $l(x,y)$ is the *tangent form* $y'(t_0)(x-x(t_0))-x'(t_0)(y-y(t_0))$.

We need to have some way of going from the parametric form of a curve to the equation and vice verse. For the first problem, let us assume (without loss of generality) that $x'(t_0)\neq 0$. Then, by the inverse function theorem, we have $g(x)$ so that $g(x(t))=t$, so that we can re-parametrise the curve to get $(x, y(g(x))$. The curve is the (locally) given by the equation $y-h(x)=0$ where $h(x)=y(g(x)$.

To go from the equation to the parametric form we need to show that lines parallel to a line which is *not* tangent to the curve will meet the curve in exactly one point near the given point. This done through the implicit function theorem. (Note to author: Exercises to be added here).

One of the advantanges of working with "orders" of vanishing is that make these theorems "explicit" if we only need our equations to be satisfied upto terms of some order. For example, we say that $(x(t), y(t))$ is a parametrisation at $t=t_0$ upto order r of the curve $f(x(t), y(t))=o((t-t_0)^r)$. Similarly, two curves, f and g are said to *osculate* upto order r if $f-g=o(\mathbf{x}^r)$. In particular, any curve osculates upto order 2 with a conic; thus it is possible to write a parametrisation upto order two quite explicitly.

Finally, there is one distinguished parametrisatisation. Let $(x(t), y(t))$ be a curve. Thinking of this as a moving point we have not only a tangent line but a tangent vector (called the velocity vector) $(x'(t), y'(t))$. It is thus natural to define the *speed* of the curve as length of this vector. We can ask for a constant speed (or more accurately constant energy) parametrisation. In other words, can we find $t=u(s)$ so that

$$(x'(t)^2+y'(t)^2)_{|t=u(s)}u'(s)^2 = \text{constant}$$

EXERCISE 59. Use the inverse function theorem to show that the function $s(t)=\sqrt{x'(t)^2+y'(t)^2}$ has an inverse. Show that this inverse function satisfies the above equation.

We will see that such a parametrisation called parametrisation by arc-length plays an important role in geometry. Meanwhile,

EXERCISE 60. Show that to obtain such a parametrisation for the circle, we need to solve the equation

$$u'(t)^2 \frac{1}{(1+t^2)^2} = 1$$

(Hint: Use the following parametrisation.)

$$(x(t), y(t)) = \left(\frac{2t}{1+t^2}, \frac{1-t^2}{1+t^2} \right)$$

We will study the solution of this and related equations in the next section.

8. Elementary functions

We are now in a position to introduce many interesting functions. In particular, we define and obtain properties of the exponential and trigonometric functions.

Warning As a consequence of integration theory it is clear that there are *more* functions that can be integrated than those that can be differentiated. This runs contrary to the feeling one gets in high school calculus due to the requirement there of expressing all integrals as a "formula" in terms of elementary functions.

Recall, that the rational functions that could not be integrated in terms of rational functions were of the form $1/(x-b)$ or $(ax+b)/((x-c)^+d^2)$. The former is a continuous function when $x \neq b$ and the latter is continuous for all x. Thus we study the two functions

$$A(x) = \int_1^x \frac{dt}{t} \text{ and } B(x) = \int_0^x \frac{dt}{1+t^2}$$

The function $A(x)$ is called the logarithm denoted by $\log(x)$ and the function $B(x)$ is the inverse of the tangent function denoted by $\tan^{-1}(x)$ or $\arctan(x)$.

EXERCISE 61. Show that the function $A(x)$ is monotonic and goes to infinity as x goes to infinity.

In particular, it follows that A has an inverse. This function is called the exponential function and denoted by $\exp(x)$. We can define the number e as $\exp(1)$; it is called Napier's natural base for the logarithm.

EXERCISE 62. Show that e is between 2 and 4.

Now the fundamental property of the logarithm is

EXERCISE 63. Show that $A(x+y) = A(x) + A(y)$. Consequently we obtain $\exp(xy) = \exp(x)\exp(y)$.

This property and the monotonicity of exp characterise it.

EXERCISE 64. Let $f(x)$ be a monotonic function so that $f(xy) = f(x)f(y)$ and $f(x) = a$. Show that $f(x) = \exp(x\log(a))$.

Finally, we can define the hyperbolic functions as usual, $\cosh(x) = 1/2(\exp(x) + \exp(-x))$ and so on.

We noe study the functions related with the function B.

EXERCISE 65. Show that the function $B(x)$ is monotonic and remains bounded as x goes to infinity.

The least upper bound of the numbers $2B(x)$ is denoted by π. We can thus define the function $\tan(x)$ on the range $(-\pi/2, \pi/2)$ as the inverse of B.

EXERCISE 66. Show that an arclength parametrisation of the unit circle $x^2 + y^2 = 1$ is given by $t \mapsto (\sin(t), \cos(t))$ where we define these functions by the formulae

$$\begin{aligned}\sin(x) &= \frac{2\tan(x/2)}{1+\tan(x/2)^2}\\ \cos(x) &= \frac{1-\tan(x/2)^2}{1+\tan(x/2)^2}\end{aligned}$$

Note that it follows that the perimeter of the circle $x^2 + y^2 = 1$ is 2π.

It is thus natural to extend these functions *periodically* for all t. We can *define* a multiplication on the unit circle by

$$(x_1, y_1) \cdot (x_2, y_2) = (x_1x_2 - y_1y_2, x_1y_2 + x_2y_1)$$

EXERCISE 67. Show that the arclength parametrisation gives a group homomorphism from the additive group of real numbers to the circle.

This exercise provides the link between the trigonometric and exponential function which is expressed by the formula

$$\exp(\sqrt{-1}x) = \cos(x) + \sqrt{-1}\sin(x)$$

which the reader can try to prove!