

Stellenbosch Camp: Intermediate Algebra

Ralph McDougall

December 2019

“Learn the rules like a pro, so that you can break them like an artist.” - Pablo Picasso

1 Introduction

Questions in Mathematics Olympiads typically focus on four topics; namely: Algebra, Combinatorics, Geometry and Number Theory. As you might have guessed, this document is going to focus on Algebra.

The split in question topics makes it seem that Algebra is only used in a quarter of all of Olympiad Mathematics; however, this cannot be further from the truth. Number Theory questions often use tricks in factorisation to make crucial observations that lead to a full solution. Combinatorics questions can involve subtle algebraic manipulations of complicated sums to greatly simplify questions. All Geometry questions can be solved purely Algebraically ¹.

¹This is left as an exercise in determination and patience to the eager or desperate student.

2 Necessary and Nice-to-Know Tools

This section will cover essentials that must be mastered before any work in Algebra can be performed. This may seem simple, but it cannot be stressed enough how important it is to have a solid foundation before continuing.

2.1 Sets

A set is a mathematical object representing a collection of “things”. Sets can contain numbers, but also cats, dogs, or even other sets. A set is represented by a pair of opening and closing curly brackets ($\{\}$) with the contents of the set separated by commas inside the curly brackets. An example of a set would be $\{2, 3, 5\}$. A further property of sets is that they do not contain duplicates of items. The statement “ $a \in S$ ” is read as “ a is an element of S ” and means that a is in the set S . The statement $A \subseteq B$ is read as “ A is a subset of B ” and means that all elements of A are also elements of B . This terminology is not specific to Algebra, but it is important to know.

You can construct sets of all numbers that follow a certain rule. This is indicated by putting the rule inside of the curly brackets when you define it.

Example 2.1.

$$S = \{x \mid 1 \leq x \leq 1\}$$

Here, S is the set of all numbers from -1 to 1 . You do not need to do this, but it is nice to know.

2.2 Intervals

An interval is like a set, but it contains all of the real numbers between two values. Where sets are indicated with curly brackets, intervals are indicated with round and square brackets. There are 4 different types of intervals.

- $[a, b] = \{x \mid a \leq x \leq b\}$ is called a “closed interval”
- $(a, b] = \{x \mid a < x \leq b\}$ is called an “open-closed interval”
- $[a, b) = \{x \mid a \leq x < b\}$ is called a “closed-open interval”
- $(a, b) = \{x \mid a < x < b\}$ is called an “open interval”

2.3 Types of Numbers

Numbers are likely the first thing that comes into one's mind when one thinks of Mathematics. Everyone is (hopefully) familiar with numbers. A few examples of numbers are 2, 10, 42 and 25. At this point, the reader may think that these are certainly not all of the numbers. There are larger numbers, like 1729 and 2401. There are numbers between numbers, like 1.5, 23.1406926, etc. There are even negative numbers, like -2 . Certainly, from this it is clear that there are many different types of numbers. The different types of numbers are grouped together in sets. A few common sets are given below as well as the notation used to denote those sets.

- $\mathbb{R} = (-\infty, +\infty)$ represents the set of all real numbers. These are all numbers that one typically deals with ².
- $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ represents the set of all integers
- $\mathbb{N} = \{1, 2, 3, 4, 5, \dots\}$ represents the set of all positive integers
- $\mathbb{N}_0 = \{0, 1, 2, 3, 4, \dots\}$ represents the set of all non-negative integers. It is often important to distinguish between this set and the one above.
- $\mathbb{Q} = \{\frac{a}{b} \mid a \in \mathbb{Z}, b \in \mathbb{N}\}$ represents the set of all rational numbers. These are all numbers that can be written as a fraction of two integers.
- \mathbb{Q}' represents the set of all irrationals, that is, numbers that cannot be expressed as a fraction of two integers.

Although this may seem unnecessarily complicated, it is vital that one keeps note of which set one may choose numbers from when looking for solutions to equations or when using functions.

2.4 Functions

If you don't care too much about formality, you can see a function as a magical black box that takes stuff in and gives other stuff out. A function is just a rule. More specifically, it takes a value from one set and corresponds it with a unique value in another set. These sets don't necessarily have to be the same, so a distinction must be made. The set that the function takes

²Chances are, if you can give a number that is not a real number, then you already know what is meant by a real number.

a value from is called the “domain” and the set that it corresponds a value to is called the “codomain”. In the following case:

$$f : A \rightarrow B$$

means that the function f takes values from A and corresponds them with values in B . In other words, f has the domain A and codomain B . The statement is read as “ f is a function from A to B ”.

When using functions, the notation that is used is $f(a) = b$ where $a \in A$ and $b \in B$. This means that f matches the value a with the value b . a can only map to one value in the codomain, but multiple values in the domain can map to the same value in the codomain. (Consider the example $f(x) = x^2$ and notice that $f(1) = f(-1)$).

2.5 Expressions

Whereas numbers represent fixed values, expressions represent a more abstract relationship between some unknown quantities. In the past, Mathematicians would write expressions out in words. This led to rather clumsy notation. Our modern notation was developed over several centuries. Consider the following statement: “Take a value and multiply it by itself and then take that result and multiply it by 3. Use that answer and add to it the value of the original value multiplied by 7. Finally consider the value that is 2 less than what has been calculated.”³ This would be written much more succinctly as “ $3x^2 + 7x - 2$ ” today.

The example brings up an important concept, that of a variable. A variable really just represents “some value”. A variable is conventionally represented by a symbol from the Latin alphabet(since that is what we use most often), however, it is sometimes nice to use symbols from the Greek alphabet(particularly common in Geometry)⁴. A variable can generally take any value from the set that it is defined in. That value is consistent through the expression though. For example, if x represent 18 in one position in the expression, it represents 18 in all positions in the expression.

Using variables, all expressions can be expressed as a string of variables and operations such as addition, subtraction, multiplication, division, exponentiation, etc. Expressions are not limited to a single variable. Oftentimes it

³Just an example, not historically accurate, nor significant.

⁴There is nothing stopping you from using symbols from other alphabets like Cyrillic or Hebrew, but it is best to be considerate to the people reading your solutions.

is necessary to consider expressions in multiple variables such as “ $20x + 19y$ ”.

In general, performing any operation on some expressions results in another expressions. That is, you can add or multiply any two expressions and always get an expression as an answer. Addition of expressions follows the idea of “adding like-terms”. That is, when you add two expressions, you add the coefficients of terms that match.

Example 2.2. *Expression addition:*

$$\begin{aligned} & (4x + 3y - 8xy - 1x^2y) + (1x - 1y + 6xy + 2xy^2) \\ &= (4 + 1)x + (3 + (-1))y + ((-8) + 6)xy + ((-1) + 0)x^2y + (0 + 2)xy^2 \\ &= 5x + 2y - 2xy - 1x^2y + 2xy^2 \end{aligned}$$

Multiplication of expressions is slightly more cumbersome, but should become second nature. If one multiplies two expressions, A and B , then every term in A must be multiplied with every term in B .

Example 2.3. *Expression multiplication:*

$$\begin{aligned} & (x + 2y) \times (xy + y^2 + 3) \\ &= x \times xy + x \times y^2 + x \times 3 + 2y \times xy + 2y \times y^2 + 2y \times 3 \\ &= x^2y + xy^2 + 3x + 2xy^2 + 2y^3 + 6y \\ &= x^2y + 3xy^2 + 3x + 6y + 2y^3 \end{aligned}$$

If you are ever unsure of whether you did the addition or multiplication correctly, you can substitute in values for x and y and verify that the answers match.

The reverse process of multiplication of expressions is factorisation. That is, where multiplying out takes several expressions and gives one expression, factorisation takes one expression and shows it as a product of multiple expressions. This is an essential skill that must be practised if one is unsure.

2.6 Factorisation

When it comes to expressions in 2 variables (x and y in this case), the following list represents the most important factorisations in Olympiads.

- $x^2 + 2xy + y^2 = (x + y)^2$
- $x^2 - y^2 = (x - y)(x + y)$

- $x^3 + y^3 = (x + y)(x^2 - xy + y^2)$
- $x^3 - y^3 = (x - y)(x^2 + xy + y^2)$
- $x^4 + y^4 = (x^2 - \sqrt{2}xy + y^2)(x^2 + \sqrt{2}xy + y^2)$
- $x^4 - y^4 = (x - y)(x + y)(x^2 + y^2)$
- $x^n + y^n = (x + y)(x^{n-1} - x^{n-2}y + x^{n-3}y^2 - \dots - xy^{n-2} + y^{n-1})$ when n is an odd positive integer
- $x^n - y^n = (x - y)(x^{n-1} + x^{n-2}y + x^{n-3}y^2 + \dots + xy^{n-2} + y^{n-1})$ where n is any positive integer

Although not as frequently used, the following ones are good to memorise if you can.

- $a^2 + b^2 + c^2 + 2ab + 2bc + 2ca = (a + b + c)^2$
- $a^3 + b^3 + c^3 - 3abc = (a + b + c)(a^2 + b^2 + c^2 - ab - bc - ca)$
- $x^5 + x^4 + 1 = (x^3 - x + 1)(x^2 + x + 1)$
- $2x^2y^2 + 2y^2z^2 + 2z^2x^2 - x^4 - y^4 - z^4 = (x + y + z)(x + y - z)(x - y + z)(-x + y + z)$

If this document teaches you anything, it should be that $(x+y)^2 \neq x^2 + y^2$ unless x or y is 0.

Probably the most commonly factorised expressions, certainly in South African schools, is the quadratic. This is any expression of the form $ax^2 + bx + c$, where a , b and c are real numbers and $a \neq 0$. These can be factorised mechanically by “completing the square”. The idea behind completing the square is that you take the expression that is given, and make it look like a square since those are relatively easy to factorise. The procedure is demonstrated with an example:

Example 2.4. *Completing the square:*

$$\begin{aligned}
 & x^2 + 6x + 8 \\
 &= x^2 + 6x + 8 + 9 - 9 \\
 &= x^2 + 6x + 9 + (8 - 9) \\
 &= (x + 3)^2 - 1 \\
 &= (x + 3)^2 - 1^2
 \end{aligned}$$

$$\begin{aligned}
&= (x + 3 - 1)(x + 3 + 1) \\
&= (x + 2)(x + 4)
\end{aligned}$$

Exercise 2.1. Find all integers a and b satisfying: $ab + 5a - 3b = 19$

Exercise 2.2. Find the value of $x^2 + \frac{1}{x^2}$ if $x + \frac{1}{x} = 17$. What about $x^3 + \frac{1}{x^3}$?

Exercise 2.3. Show that 9 is a factor of $2^{2019} + 1$ and that 7 is a factor of $2^{2019} - 1$.

3 Polynomials

3.1 Useful facts

A polynomial is any expression of the form $a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x^1 + a_0 x^0$ where $a_n \neq 0$ and a_0, a_1, \dots, a_n are all constant real numbers. This is usually written as:

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$$

The name P is normally used to describe some arbitrary polynomial, but you can call the polynomial whatever you want.

Example 3.1. *Some polynomials:*

$$A(x) = x$$

$$B(x) = x^2 - 2x + 1$$

$$C(x) = \pi x^3 - 17.4x^2 + 1$$

Polynomials are very common in Maths Olympiads. Whether the question is directly relating to some polynomial or a polynomial is used in solving the question, it is important to be able to manipulate them appropriately.

Some important terminology regarding $P(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$:

- The degree of P is the highest power in the expression. This is denoted by $\deg(P)$. In this case, $\deg(P) = n$.
- The values a_0, a_1, \dots, a_n are referred to as the coefficients.
- Any value r satisfying $P(r) = 0$ is referred to as a root of P .
- P is called a monic polynomial iff $a_n = 1$.

In general, if P and Q are polynomials, then so are $P(x) + Q(x)$, $P(x) - Q(x)$, $P(x)Q(x)$ and $P(Q(x))$. $\frac{P(x)}{Q(x)}$ is not always a polynomial; it is called a rational function (similar to a rational number).

One more theorem is necessary before you have all of the knowledge you need to use polynomials to start solving problems.

Theorem 3.1 (The Fundamental Theorem of Algebra). *Every non-constant polynomial with real coefficients can be expressed as a product of polynomials with real coefficients with degree at most two⁵.*

This theorem is significant because no matter how complicated a polynomial may look, it is always possible to express it as a product of relatively simple polynomials. In fact, this allows us to write a polynomial in another form:

$$P(x) = a_n R_1(x) R_2(x) R_3(x) \dots R_k(x)$$

where all R_i are either of the form $R_i(x) = x + p_i$ or $R_i(x) = x^2 + p_i x + q_i$ where p_i and q_i are real numbers.

Example 3.2.

$$x^3 - 2x - 4 = (x - 2)(x^2 + 2x + 2)$$

Example 3.3.

$$x^3 - 2x^2 - 5x + 6 = (x - 1)(x + 2)(x - 3)$$

Example 3.4.

$$x^4 - 1 = (x^2 + 1)(x - 1)(x + 1)$$

Example 3.5.

$$x^5 - 1 = (x - 1)\left(x^2 + \frac{1 - \sqrt{5}}{2}x + 1\right)\left(x^2 + \frac{1 + \sqrt{5}}{2}x + 1\right)$$

This is incredibly useful to know when one is trying to find the roots of a polynomial. To see this, consider the situation $A \times B = 0$ where A and B are some values. It is fairly easy to see that one of A and B must be 0. This idea can be extended to polynomials. Suppose that r is a root of P ; that is, $P(r) = 0$. From this we get that $a_n R_1(r) R_2(r) \dots R_k(r) = 0$. Since $a_n \neq 0$ in the definition of a polynomial, we must have that one of $R_1(r)$, $R_2(r)$, ... $R_k(r)$ must be 0. This is significant since in theory, if we could find the roots of the monic linear and quadratic factors, we would find all roots of the polynomial.

⁵This definition is sufficient for Intermediate-level Algebra. The full definition is left as an exercise to the eager student.

3.2 Roots of linear and quadratic polynomials

Let $P(x) = ax + b$. Let r be a root of P .

$$\iff P(r) = 0$$

$$\iff ar + b = 0$$

$$\iff ar = -b$$

$$\iff r = \frac{-b}{a}$$

Let $P(x) = ax^2 + bx + c$ and r be a root of P .

$$\iff P(r) = 0$$

$$\iff ar^2 + br + c = 0$$

$$\iff r^2 + \frac{b}{a}r + \frac{c}{a} = 0$$

$$\iff r^2 + \frac{b}{a}r = \frac{-c}{a}$$

$$\iff r^2 + \frac{b}{a}r + \frac{b^2}{4a^2} = \frac{-c}{a} + \frac{b^2}{4a^2}$$

$$\iff \left(r + \frac{b}{2a}\right)^2 = \frac{-c}{a} + \frac{b^2}{4a^2}$$

$$\iff \left(r + \frac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a^2}$$

$$\iff r + \frac{b}{2a} = \pm \sqrt{\frac{b^2 - 4ac}{4a^2}}$$

$$\iff r + \frac{b}{2a} = \frac{\pm \sqrt{b^2 - 4ac}}{2a}$$

$$\iff r = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

It is worth memorising this equation as soon as possible. There are also formulas for the solutions to general cubic and quartic equations, but they are not worth memorising. Furthermore, it has been proven that no formula exists for the solutions to general equation with a degree higher than 4. However, that does not mean that it is impossible to solve *every* equation that has a high degree.

Using these two results, it is possible to find the roots of any polynomial with real coefficients if the factorisation of it is provided. However, finding the factorisation is the hard part of finding the roots of a polynomial. This is where your factorisation skills become useful.

There is an elephant in the room in this section. The term $\sqrt{b^2 - 4ac}$ should have raised a few red flags. If $b^2 - 4ac < 0$, then $\sqrt{b^2 - 4ac}$ doesn't make sense. Indeed, in this case, the polynomial has no real-valued roots. However, if one considers solutions in the complex numbers, then it has two, distinct roots. It is not necessary to know this at an Intermediate level, but it can only help you to know about them now.

3.3 Roots of polynomials from specific sets

At the end of the previous subsection, it was mentioned that a quadratic polynomial does not always have real roots. If one is trying to find properties of real solutions of a polynomial, then it is often a waste of time to find the complex solutions. Similarly, if one knows that the size of some set is a root of some polynomial, then one does not care about roots that are not non-negative integers. As a result, there are a number of theorems and properties relating to roots from specific sets.

The number of complex roots of P is equal to $\deg(P)$. The eager reader is urged to research Viète's Formula if they are interested in learning more properties of all of the roots of a polynomial.

The numerical values of the real roots of a polynomial can always be approximated, but finding their exact values is extremely difficult. The number of real roots of P is less than or equal to $\deg(P)$. However, if $\deg(P)$ is odd, then P has at least one real root. A further upper-bound on the number of real roots of P can be found with the next theorem.

Theorem 3.2 (Descartes' Rule of Signs). *The total number of real roots is at most equal to the number of times that the signs of the coefficients of $P(x)$ and $P(-x)$ change.*

Example 3.6.

$$x^3 - 3x - 2 = 0$$

The coefficients are 1, -3 and -2 . The sign changes once between 1 and -3 , so there is at most one positive root. Replacing x with $-x$ gives the

expression:

$$-x^3 + 3x - 2$$

which has coefficients -1 , 3 and -2 which changes signs twice. Thus, there are at most two negative roots. Factorising the original expressions gives:

$$x^3 - 3x - 2 = (x + 1)^2(x - 2)$$

from where we can see that the polynomial does indeed have a single positive root and a single unique negative root which is less than the two which was predicted.

Theorem 3.3 (Rational Root Theorem). *If P is a polynomial with integer coefficients and $P(\frac{s}{t}) = 0$, then $s \mid a_0$ and $t \mid a_n$.*

Using Theorem 3.3, one can create all possible values that can be roots and manually determine which ones are actually roots.

Integer solutions can be found by the same technique by forcing $t = 1$. If this does not work, one can always try using a Number Theoretical approach such as looking at the remainder of the polynomial after division by some integer k . This text will not explore this idea further however.

Example 3.7 (Find all rational solutions to $2x^3 + 3x^2 - 8x + 3 = 0$). *The Rational Root Theorem tells us that any rational root can be found by looking at factors of the leading coefficient as well as the constant coefficient. The factors of 2 are -2 , -1 , 1 and 2 (note the inclusion of negative factors). The factors of 3 are -3 , -1 , 1 and 3 . From here we evaluate the value of $P(x) = 2x^3 + 3x^2 - 8x + 3$ at each of the possible fractions:*

- $P(\frac{-3}{1}) = 0$
- $P(\frac{-3}{2}) = \text{something}$
- $P(\frac{-1}{1}) = 12$
- $P(\frac{-1}{2}) = \text{something}$
- $P(\frac{1}{1}) = 0$
- $P(\frac{1}{2}) = 0$
- $P(\frac{3}{1}) = 60$
- $P(\frac{3}{2}) = \text{something}$

From here we see that the rational roots can only be $\frac{1}{2}$, -3 and -1 . In fact, these are all of the solutions since it can be checked that $2x^3 + 3x^2 - 8x + 3 = (2x - 1)(x + 3)(x - 1)$.

Exercise 3.1. Let P be a polynomial of degree 5 with a leading coefficient of 2, and let:

$$P(1) = 1, P(2) = 2, P(3) = 3, P(4) = 4, P(5) = 5$$

What is the value of $P(6)$?

4 Inequalities

4.1 The only inequality you ever need to know

Almost all inequalities are derived from a single, powerful inequality:

Theorem 4.1 (The only inequality you ever need to know).

$$x^2 \geq 0 \quad \forall x \in \mathbb{R}$$

with equality only when $x = 0$.

Using this inequality, you can prove most Olympiad inequalities if you are resourceful enough. Doing this is known as “proving by first-principles”. However, this does require very strong Algebraic skill, which can only be built through practice.

An important property to know about inequalities is that if $a \geq b$ and $c \geq d$, then $a + c \geq b + d$. Using this, we can combine $x^2 \geq 0$ and $y^2 \geq 0$ to get $x^2 + y^2 \geq 0$. One can be more subtle as well.

Example 4.1. *Consider $a^2 + b^2 \geq 2ab$. This may appear daunting at first if you have not seen inequalities before, but notice the following:*

$$a^2 + b^2 \geq 2ab$$

$$\iff a^2 - 2ab + b^2 \geq 0$$

$$\iff (a - b)^2 \geq 0$$

and this is of the same form as $x^2 \geq 0$, so it is true for all $a, b \in \mathbb{R}$.

Although proving by first-principles is extremely powerful, it can be quite cumbersome. Some inequalities (that can be proved by first principles) appear so often that they are worth remembering.

4.2 AM-GM

Theorem 4.2 (Arithmetic Mean-Geometric Mean Inequality).

$$\frac{a_1 + a_2 + a_3 + \dots + a_n}{n} \geq \sqrt[n]{a_1 a_2 a_3 \dots a_n}, \forall n \in \mathbb{N}, a_i \in \mathbb{R}_{>0}$$

with equality occurring if and only if $a_1 = a_2 = a_3 = \dots = a_n$.

The left part is known as the Arithmetic Mean, another name for the average. The right part is known as the Geometric Mean, which is less frequently used. Proving AM-GM uses a combination of First Principles and Induction.

Example 4.2. Consider Example 4.1:

$$\begin{aligned} a^2 + b^2 &\geq 2ab \\ \iff \frac{a^2 + b^2}{2} &\geq ab \\ \iff \frac{a^2 + b^2}{2} &\geq \sqrt[2]{a^2 b^2} \end{aligned}$$

which is AM-GM with $a_1 = a^2$ and $a_2 = b^2$.

4.3 Power-Mean inequality

The Power-Mean inequality is a generalisation of the AM-GM inequality.

Theorem 4.3 (Power-Mean Inequality). *Define:*

$$P_r = \left(\frac{a_1^r + a_2^r + a_3^r + \dots + a_n^r}{n} \right)^{\frac{1}{r}} \quad \forall r \in \mathbb{R}$$

$P_i \leq P_j \iff i \leq j$, with equality if and only if $a_1 = a_2 = a_3 = \dots = a_n$.

Table 1: A few important power-mean values

r	P_r
$-\infty$	Minimum
-1	Harmonic mean
0	Geometric mean
1	Arithmetic mean
2	Quadratic mean
∞	Maximum

Notice that P_0 is not really properly defined since there is division by 0. However, as r becomes closer and closer to 0, P_r becomes the Geometric Mean. For this reason, P_0 is defined to be the Geometric Mean. Thus, the AM-GM inequality is directly equivalent to $P_1 \geq P_0$, which is true by the Power-Mean inequality.

4.4 Cauchy-Schwartz

Theorem 4.4 (Cauchy-Schwartz Inequality). *For any real numbers a_1, a_2, \dots, a_n and b_1, b_2, \dots, b_n*

$$\left(\sum_{i=1}^n a_i b_i\right)^2 \leq \left(\sum_{i=1}^n a_i^2\right) \left(\sum_{i=1}^n b_i^2\right)$$

with equality when there exists some nonzero constant k such that $ka_i = b_i$ for all i .

Theorem 4.5 (Cauchy-Schwartz Inequality in Engel Form / Titu's Lemma). *For positive reals a_i and b_i*

$$\frac{a_1^2}{b_1} + \frac{a_2^2}{b_2} + \dots + \frac{a_n^2}{b_n} \geq \frac{a_1^2 + a_2^2 + a_3^2 + \dots + a_n^2}{b_1 + b_2 + b_3 + \dots + b_n}$$

Engel form is derived by substituting $a_i = \frac{x_i}{\sqrt{y_i}}$ and $b_i = \sqrt{y_i}$.

Exercise 4.1 (PAMO 2015 Q1). *Show that for all x where the expression is well-defined:*

$$\sqrt{x-1} + \sqrt{2x+9} + \sqrt{19-3x} < 9$$

Exercise 4.2 (PAMO 2016 Q4). *Let x, y, z be positive real numbers such that $xyz = 1$. Prove that*

$$\frac{1}{(x+1)^2 + y^2 + 1} + \frac{1}{(y+1)^2 + z^2 + 1} + \frac{1}{(z+1)^2 + x^2 + 1} \leq \frac{1}{2}$$

5 Functional Equations

5.1 The basic process

Example 5.1. Find all functions $f : \mathbb{R} \rightarrow \mathbb{R}$ such that

$$f(x - y) = f(x) + f(y) - 2xy \quad \forall x, y \in \mathbb{R}$$

Functional Equations may seem highly intimidating at first, but the ideas behind the solutions are generally always the same. Functional Equation questions always start by giving the name, the domain and the codomain of the function. The remainder of the question is the condition that needs to be met. In the Example 5.1, there is a function, f , that takes values from the real numbers as inputs and gives real numbers as outputs. The condition that must be satisfied is $f(x - y) = f(x) + f(y) - 2xy$, no matter which values of x and y are chosen from the domain.

The first step is to try and guess solutions. This may seem counter-intuitive since we're trying to rigourously find all possible functions satisfying the condition, but guessing solutions puts a post in the ground. It is normally possible to spot a "simple" solution fairly quickly and with a bit more thought, a "weird" solution sometimes pops out too. Don't stress too much about guessing *all* of the solutions, but the more you find, the better. These solutions help you to plan your approach to solving the function. For example, if you know that $f(x) = 1 \quad \forall x$ is a solution, then there is no point trying to find a k such that $f(k) = 0$. You can try to spot properties of the solutions that you have guessed and from there prove that all other solutions must have those properties too, but more on that later.

Some useful functions to try as solutions are:

- $f(x) = 0$
- $f(x) = c$
- $f(x) = x$
- $f(x) = cx$
- $f(x) = x + c$
- $f(x) = x^2$

Sometimes you can spot a solution by looking at the form of the constraint condition, but that is up to your Algebraic intuition.

In the Example 5.1, the only “simple” function that works is $f(x) = x^2$. Note that it is not enough to just guess one solution, one must prove rigourously that no other “weird” function exists. Some simple-looking functional equations have what are known as “pathological solutions” that are impossible to write down concisely.

The next step in solving a functional equation is to poke and prod it to see what it does. This is done by substituting in values for the inputs. Always be sure to write down clearly what substitution you have made. If the marker is unsure of what you have done, they will likely assume that you did something wrong or skipped a step and then start taking marks off. In the example, look at what happens when $x = 0$ and $y = 0$. The condition becomes:

$$\begin{aligned} f(0 - 0) &= f(0) + f(0) - 2 \times 0 \times 0 \\ \Rightarrow f(0) &= 2f(0) - 0 \\ \Rightarrow f(0) &= 0 \end{aligned}$$

This is very useful to know since now, any function f that is a solution to the equation *must* satisfy $f(0) = 0$.

The values that you substitute do not need to be explicit numbers. Consider letting $x = k$ and $y = 0$, where k is some fixed real number.

$$\begin{aligned} \Rightarrow f(k - 0) &= f(k) + f(0) - 2 \times k \times 0 \\ \Rightarrow f(k) &= f(k) + f(0) = f(k) + 0 \end{aligned}$$

This result is not helpful. You will encounter this very often when you practice functional equations. This tells us that the substitution we made did not provide any new information. Let us consider instead: $x = k, y = k$.

$$\begin{aligned} \Rightarrow f(k - k) &= f(k) + f(k) - 2 \times k \times k \\ \Rightarrow f(0) &= 2f(k) - 2k^2 \\ \Rightarrow f(k) &= k^2 \end{aligned}$$

This is incredibly useful. Since k was not given any explicit value, we must have that no matter which k you choose, $f(k) = k^2$. This is exactly the solution we guessed, and since we followed a progression of true statements,

this must be the *only* solution that can possibly work.

We are not done with the question, however. There is a subtlety in the process we have followed. We have said that *if* a solution exists, it must be $f(x) = x^2$. However, this does not mean that that is actually a solution. The only way to check if something is a solution is to check that it satisfies the condition. This is done by considering the left-hand-side (LHS) and right-hand-side (RHS) of the constraint and checking that they are the same.

If $f(x) = x^2 \forall x \in \mathbb{R}$:

$$LHS = f(x - y) = (x - y)^2 = x^2 - 2xy + y^2$$

$$RHS = f(x) + f(y) - 2xy = x^2 + y^2 - 2xy$$

$$\Rightarrow LHS = RHS$$

Therefore, $f(x) = x^2$ checks successfully, so it is the only solution.

If you do not include this check at the end, you will lose at least 1 point out of 7. People have missed out on IMO medals because they didn't check the function at the end.

5.2 Cauchy equation

Although every functional equation question is different, occasionally, they can be reduced to a few well-known functions. These don't appear very often, but it is worth knowing about them.

$f(x + y) = f(x) + f(y)$ is called the Cauchy Equation. If the domain is \mathbb{Q} , then $f(x) = cx$ is the only solution. Over the reals, it is not so easy to solve. However, if one of the following conditions holds, then $f(x) = cx$ is the only solution.

- f is monotonic on some interval
- f is continuous
- f is bounded on some interval
- $f(x) > 0 \forall x \geq 0$

Don't worry too much about what "continuous" and "bounded" mean. They will not appear in any Maths Olympiad questions at school level.

Exercise 5.1. Find all continuous functions $f : \mathbb{R} \rightarrow (0, +\infty)$ satisfying

$$f(x+y) = f(x)f(y)$$

Exercise 5.2. Find all continuous functions $f : (0, +\infty) \rightarrow \mathbb{R}$ satisfying

$$f(xy) = f(x) + f(y)$$

Exercise 5.3. Find all continuous functions $f : (0, +\infty) \rightarrow (0, +\infty)$ satisfying

$$f(xy) = f(x)f(y)$$

Exercise 5.4 (IMO 2019 Q1). Let \mathbb{Z} be the set of integers. Determine all functions $f : \mathbb{Z} \rightarrow \mathbb{Z}$ such that, for all integers a and b ,

$$f(2a) + 2f(b) = f(f(a+b))$$

Exercise 5.5 (PAMO 2013 Q2). Find all functions $f : \mathbb{R} \rightarrow \mathbb{R}$ such that

$$f(x)f(y) + f(x+y) = xy$$

6 Conclusion

These are all of the basic tools required to get started with Algebra in Olympiads. This is by no means the extent of the Algebra required, but it does cover most of the essentials. From here, the best way to improve is to practice as much as you can.

Good luck!