UNIVERSITÀ DEGLI STUDI DI BERGAMO

International Ph.D. School Human Capital Formation and Labour Relations

XXV[th] Cycle

MATH FOR FREEDOM

AN ORIGINAL PROOF OF THE FUNDAMENTAL THEOREM OF ALGEBRA WITHIN THE AMBIT OF REAL NUMBERS

by

A. J. Cuneo

Mat. 1013952

Supervisor: Prof. Ruggero Ferro

Contents:

"The essence of Mathematics lies in its freedom"
George Cantor

1.0. Freedom in Mathematics and Mathematics for Freedom

George Cantor's quote, ``... the essence of mathematics lies in its freedom'', appears on Ueber unendliche, lineare Punktmannischfaltigkeiten, Mathematische Annalen, volume 21, of 1883. In this paper, Cantor extends the natural numbers to infinite ordinal numbers, with addition and multiplication defined on them. This is the paragraph in which it appears:

'Es ist, wie ich glaube, nicht nöthig in diesen Grundsätzen irgendeine Gefahr für die Wissenschaft zu befürchten, wie dies von Vielen geschieht; einerseits sind die bezeichneten Bedingungen, unter welchen die Freiheit der Zahlenbildung allein geübt werden kann, derartige, dass sie der Willkür einen äussertst geringen Spielraum lassen; dann aber trägt auch jeder mathematische Begriff das nöthige Correctiv in sich selbst einher; ist er unfruchtbar oder unzweckmässig, so zeigt er es sehr bald durch seine Unbrauchbarkeit und er wird alsdann, wegen mangelnden Erfolgs, fallen gelassen. Dagegen scheint mir aber jede überflüssige Einengung des mathematischen Forschungstriebes eine viel grössere Gefahr mit sich zu bringen und eine um so grössere, als dafür aus dem Wesen der Wissenschaft wirklich keinerlei Rechtfertigung gezogen werden kann; denn das Wesen der Mathematik liegt grerade in ihrer Freiheit.'

The translation (by Jörg Peters) in English is: '*These principles, I think, do not represent a danger for science as some have suggested: first, the conditions under which new numbers can be generated leave little space for arbitrariness; second, every mathematical construct comes with a natural corrective; if it is impractical or uninspiring it will be dropped due to lack of impact. Conversely, any unnecessary constraint on the mathematical research impetus carries the far greater risk -- the more so since there is no scientific justification for it; for the essence of mathematics is its lack of constraints.'*

Cantor defends his theory of transfinite numbers from the attacks of many of his contemporaries colleagues based on the principle that mathematical research should keep an open, free of external restrains, attitude. This initial openness allows mathematicians to give some value to every theory as long as free from contradiction. In this sense we can say that there is a lot of freedom in mathematics.

In the same paragraph Cantor reassures his critics saying that '*…every mathematical construct comes with a natural corrective; if it is impractical or uninspiring it will be dropped due to lack of impact.'* This could be read as a restrain to the liberty of mathematics, but it actually is a further aspect of its freedom. The absence of external constraints is not enough for freedom. Think for example: Is there more freedom in a person wandering about randomly with no direction, or within somebody capable of tracking his position and getting to any destination he desires effectively? There are instruments for freedom, and mathematics can be one of them.

If the only limit would be coherence and thus we were to equally accept every non-contradictory theory we would be lost in an ocean of them. We would be left with no criteria to chose among these theories the ones that are more relevant to us, based on their *practicality* and *impact,* to use Cantor's words. So the only limit other than coherence imposed to mathematical theories is that they enhance our freedom. A version of Cantor's quote that takes account of these remarks is: Mathematics is essentially free as long as it is an instrument of freedom.

1.1. A view on Mathematics.

I say that mathematics enhances freedom because it helps us to solve problems[1], and these problems come from the real world. I'm aware that this connection to the real world can at times be hard to spot because of the levels of abstraction. For example sometimes after solving a real world problem we can create a new problem: is the solution given the best possible solution? This new problem is not about real world things, but about solutions of a real world problem. The connection to the real world is still there but at a second level of abstraction.

It can happen that at high levels of abstraction we realize that our solutions also apply to other real situations different from the one we started the process of abstraction with. This happens because the two situations share some common characteristics that only become evident at these high levels of abstraction. Mathematics can give us new insight on reality!

[1] I don't mean that mathematics solves all kind of problems, far from that. In fact I could try to narrow the field with some characterization of the kind of problem that mathematics deals with (a general characteristic for example is that math deals with multiplicity, in a broad sense) but such a characterization is not fundamental to the point I'm trying to make in this work.

Many mathematicians don't immediately share my view. It's usually very hard to keep track of these levels of abstraction and the connections to reality. Even more, having these connections in mind is not always necessary or useful to develop mathematics. Many prefer to think that math objects have no links to real things and can be modified at will, or by chance and still have some meaning and maybe, magically, applications. This view can be sometimes enough to work on math, but it fails to explain why math applications work and why math is developed in certain directions and not others. It also would fail to explain why we call the mathematical results with different names like exercise, theorem, corollary and fundamental theorem. If they were all meaningless, why not call them with the same name?

Here is a fact, we have to live in reality, we have to move in this world, and reality is just too complicated for us to understand it directly. Another way to say that reality is complicated is to say that men have limits in their ability to grasp it directly. If we wouldn't have these limitations we wouldn't need math at all, and we'd just know everything. Math won't certainly help us overcome these limitations, they are part of us (and thus also part of the real world on which math operates) and there is nothing we can do about that. It's just how we are. Anyway mathematics is a mean to help us to do our best within our limitations.

There is an enlightening distinction between the two terms 'complication' and 'complexity', and their role in mathematics, conceived by Ruggero Ferro and communicated at the POM-SIGMAA section of The Joint Mathematics Meeting at New Orleans 2011:

"The world around us with all its most minute aspects shows up in a very complicated manner: a huge amount of disordered, unconnected information, that we are unable to master. The only way out that we have is to neglect the irrelevant pieces of information, and to organize and order those that we consider relevant, since they may still be too many to be managed directly. Thus we overcome complication introducing complexity. I am proposing a sharp distinction between the words complication and complexity, even though they are often viewed as synonymous. Complicate should indicate a situation with a lot, too many, disorganized pieces of information, thus making it impossible to have an idea of what is going on, to understand how things are correlated and could evolve.

On the other hand, complex should indicate a state of affairs difficult to understand and follow, due to a very rich organization of a very large amount of information possibly with several connections of different types and at different levels. In order not to be lost in a complex environment one needs a manner to trace continuously one's position, while in a complicated entourage one is just lost without any point of reference or a direction toward which to move. In *a complicated world the only strategy is random trials and errors, in a complex one the strategy would be quietly track the position and the direction where to go.*

The way from complication to complexity passes through the selection (after trials and errors) of what to consider as relevant for the task at hand, and the organization of the selected items by connections and dependences at different levels and in articulated subordinations. Complexity is not a goal, but something to which to make recourse when simpler pictures, models are not adequate to overcome the complication of the situations at hands."

Thus, I'm proposing a vision were math is a way to solve real problems, problems of problems and so on, introducing ordered complexity to overcome the chaotic complication, coping with our human limitations.

Another important characteristic of mathematics is the treatment that the knowledge acquired by solving problems receives. In order to take account of all the relationships and interconnections among all these pieces of knowledge a lot of the mathematician work is to organize them in structured systems. These systems are meant to facilitate the control of coherence, completeness and precision. The paradigm of this kind of organization is the axiomatic system.

1.2. Math for slaves and math for free people.

But there is still another very important aspect about math problem solving that I have to mention to present a more complete picture. In mathematics we don't want to solve similar problems over and over again. We don't go to higher levels of abstractions to get away from reality, but we do it to grasp the common aspects to many problems and solve all of them once and for all. In order not to redo the solution process each time a similar problem arises, in mathematics is fundamental the creation of automatic procedures of solution. That is to say, algorithms or calculus that work step by step on an artificial representation of the problem and that automatically (without having to think and understand each time the elements of the problem) provides a representation of the solution.

Isn't this a quite reasonable way to proceed? After somebody has spend a lot of effort finding the solution of a problem, instead of using the rest of his time repeating the thought process each time the problem arises he can teach somebody else how to get the solutions, and this second agent doesn't even need to do all the learning process that the first person had to do to solve it. He or it just have to repeat a series of purely mechanical steps to arrive there, without even thinking! In this way the first person can now dedicate his time and his capacity (just increased from his experience) to solve new exciting problems! This makes a lot of sense to me.

These automatic procedures, which can be executed without any knowledge of the concepts involved, are called “calculus”. This is the Latin word for “little stone”, which at first may seem strange, but in fact it makes perfect sense. Latin is the language of the Romans and they used to represent numbers with a system that we still use sometimes to indicate dates in old buildings. But this system of representation of numbers is not as good as our actual system for the creation of automatic algorithms for the basic operations, like sums and products. Have you ever tried to perform a long division using roman numbers? Try and you’ll see what I mean. Despite this, romans still performed these operations, and to do this they used a different system of representation: the abacus. The abacus is a system much more similar to ours and on it you can perform operations following automatic rules. In order to do this, the operator has to move around these little stones (the calculi) placing them in the various sections of a plate.

When calculation involved large numbers (and in a big empire, like the Roman, large numbers were usual) there were a lot of stones to move and the rules had to be followed with a lot of attention to avoid mistakes. As this was quite a boring and tiring job, romans used to train and employ slaves for this work. The slaves didn’t necessarily knew what they were doing, but just mechanically follow the rules to get to the final representation, risking severe punishment in case of error.

Marble funerary relief of young man; Roman, early second century CE. The deceased young man reclines on a couch beneath his dead father's shield-portrait (which may indicate a military career); he holds a scroll in his left hand and a large moneybag in his right. His grieving mother, her palla pulled up over her elegant Flavian hairstyle, is seated on the right, with her arm tenderly resting on her son's shoulder. The slave standing on the left is operating an abacus, which symbolizes the family's success in business. Rome, Palazzo Nuovo (Capitoline Museums). Credits: Barbara McManus, 2003; 2007

This is the origin of the arithmetic calculus. Since then, many other “calculi” came to light: The infinitesimal calculus for derivatives and integrals, the probability calculus, and many more.

Thus, creating calculus algorithms is a very important part of mathematics. For practical purposes is also important to train agents capable of executing calculus effectively. Historically mathematics has developed in both these perspectives. On one side we can find the work of mathematicians that tackle new problem solutions or that organize a set of knowledge in a systematic way. Historically this was the mathematics performed in the academies. On the other hand we can find practical handbooks intended for a specific industry or trade, containing the mathematical automatic procedures necessary for that trade. Historically this was the mathematics performed in the artisan’s workshops. A good example of these two aspects of mathematics can be found in the work of Leonardo Pisano (Fibonacci). Leonardo has the merit of introducing the Hindu-Arabic numeration system in Europe. Everybody knows about his book “Liber Abbaci”, where he explains how and why the new system works, and presents the corresponding algorithms and some examples of problems solved using these instruments. This book is very academic and its main goal is to show a new method for solving the problem of representing numbers and operating with these representations, but at the same time Leonardo understood the need of a more practical approach in order to address non-academic world that needed these instruments. Quoting Keith Devlin in his monthly column:

"…Liber abbaci, completed in 1202, was the first comprehensive description of modern arithmetic (using the Hindu-Arabic numerals, decimal place-value representations of numbers, and the basic arithmetic algorithms we all learn in school) in the western world. But it was not really a textbook. Its mammoth length and the fact that it was written in Latin made it accessible only to scholars.

Smart man that he was, Leonardo recognized that, and wrote a shorter, simpler book, aimed at merchants and businessmen. Unfortunately, though copies of most of Leonardo's works survived (no original Leonardo manuscript did), there are no known copies of his smaller book for merchants. We know he wrote it because he refers to it in Liber abbaci, calling it his liber minoris guise ("book in a smaller manner")…"

Recently a manuscript has been discovered that could be the "liber minoris guise", you can read all about this recent discovery in Keith Devlin's books "The Man of Numbers: Fibonacci's Arithmetic Revolution" and "Leonardo and Steve: The Young Genius Who Beat Apple to Market by 800 Years". The true advantage of the Hindu-Arabic notation is that the calculus algorithms are far more easy to use than the ones performed with the abacus and also more effective. This advantage is somehow represented in Gregor Reisch painting "Arithmetica". I think that the two Leonardo's books (Liber Abbaci and Liber Minoris) illustrate the two perspectives I was just talking about. Firstly the mathematics that solve problems and organize their solutions in a systematic way, creating and justifying algorithms. Secondly, mathematics as training to use these algorithms practically and effectively in everyday life.

Gregor Reisch's Arithmetica, 1503. A symbolic image depicting Boethius and Pythagoras in a mathematical competition. Pythagoras uses an abacus, while Boethius uses numerals from India. Boethius looks very proud, he is ready while the poor Pythagoras still tries to find the solution.

2.0. School Mathematics: the role of calculations.

Given these two perspectives, what kind of mathematics should be taught at school? Is the mathematics performed by academics or the more practical mathematics that was taught at the workshops of the different trades? Is the math that studies how to arrive to the solution of a problem and consequently has to justify its solution methods or is the training on these methods more important? Maybe what we get in school is a little bit of both, but sometimes one of these perspectives can be much more present than the other one.

In particular it seems to me that historically a very important role was always given to the second one. Everyday problems had to be solved and algorithms did a great job, so we needed people using them proficiently. Let me illustrate this point with a personal story.

My grandfather tells me that when he was at school, the teacher used to reward the students that could perform operations correctly faster than their classmates. He grew up in a farm and later he became an “accountant” for the agricultural cooperative. He thinks very highly of his teacher as the abilities that he learned became very important for him. His capacity to perform the operations efficiently helped him move on in this world! At his house I still find the border of the newspaper filled with very long sums of his expenses or retirement payments.

My father had a similar school experience, and when he was in the university studying to be an accountant the ability to perform calculations was fundamental. After graduating from the university he started working in a bank. During his first work years of work the first calculator machines appeared. The efficiency of these machines for performing calculations was far superior of the best-trained employees and allowed them to employ their time in different tasks. My father often says that in order to make progress in his career, he ended up doing something completely different to the preparation he received at school and university, and had to self-teach most of the competences that were actually useful for him.

When I was in school pocket calculators were accessible to anyone, still my teachers preferred to forbid their use because it prevented us from being trained in the use of the algorithms. They didn’t make me compete with my classmates for speed and efficiency in calculations but they still used most of the time in training us in the use of algorithms. I have to confess that I didn’t take much profit of all that training as my grandfather did. I really liked my pocket calculator and I just couldn’t understand why I couldn’t use it and therefore I wasn’t paying much attention. My teachers and parents used to tell me that I was going to make mistakes and be embarrassed when I couldn’t have a calculator at hand and it turned out that they were right. I'm terrible at performing fast calculations efficiently. Anyway, later I became a mathematician. I discovered that there is more to it than being good at performing calculations, and I discovered the whole first perspective about math that I pointed out before. I'm not proud of not being efficient at performing algorithms, in some cases I even forget them altogether, but in my studies I've learned what ideas these algorithms are based on, and from these I can reproduce them when needed in a fairly reasonable time.

As a math teacher I couldn't find an answer of why we insist so much on training for calculation ability. My students claim that nowadays all those things are done with the help of the computer, which is absolutely true. Maybe I should point out to them that the computer may not be available in some situations, but I already know that they won't be satisfied with such an answer.

But what is it that computers do? Quoting again Ruggero Ferro (Representation and Explanation in the Sciences, Colloque de Louvain-la-Neuve, Belgique, 26-28 april 2011) "*Computers are machines that transform their status (which is a manner of representing a situation) according to precise syntactical rules (i.e. rules that consider only the notations and not their meanings) prescribed by a program. There are absolutely no ways for a computer to know what it is doing; it just executes the commands and reaches a final state that could include the lighting of certain pixels on a screen. Humans were able to give a machine a sequence of commands on how to transform its status, and they are able to interpret and read its final status as the representation of the meaning resulting from performing the intended operations on the meanings represented by the initial notations*." As you can see, in many ways, today we can use computers in the same manner as romans utilized their slaves with the abacus, but faster, with greater precision and with no moral issues.

Once, when my grandfather was at school, the motivation to learn calculations was external to mathematics. There was a need of proficient people in the use of the algorithms of mathematics, and that assured a good work prospective. Today computers cause this motivation to vanish. Today it's more reasonable to learn how these algorithms work (and therefore why computers can be trusted) creating a motivation to learn calculations that is internal to mathematics.

Many people think that the fact of having computers shouldn't affect the weight we give to performing calculations "by hand" in school. The reason often given is that performing enough calculations by yourself will help us understand why they are correct and allow us to reproduce them, improve them, etc. In this prospective there is no need to change our practice of training to calculations because this training will also allow us to understand what is beneath these calculations.

But performing calculations is not the best way to understand why the calculations are performed in that way. This is not a good example of "learn by doing". The reason for this resides in a fundamental characteristic of calculations, which is, that they are done to avoid the resource to the concepts they refer to, but these concepts are the reason of why we can rely on these algorithms. Since calculations avoid the use of their related concept, learning the concept from the calculation is not the easiest task and it involves a lot of guessing.

2.1. You don't learn the concept by exercising with the algorithm.

As an example, consider how we learn the first operation that is studied at school, which is addition. You may have already heard criticisms to teaching methods that rely too much on abundant practice of the algorithms, and you probably don't want to hear another one. Most criticisms are based on the poor results obtained by this particular method of teaching, but do they really address the question of why is it that exercising with the algorithm has poor results? If the reasons of the poor results are not specified we could be inclined to think that maybe the approach is correct but it is the implementation that needs to be improved. Many hypotheses (also without fundaments) can be made to explain why the whole approach is wrong, but in order to really answer this question (i.e. make a properly grounded hypothesis) one has to look at the core of the problem and investigate what the concept of algorithms is. I believe that the close and detailed look at the algorithm of addition I'll present, will illustrate many interesting aspects of the concept of algorithm. These characteristics of the algorithms are enough to explain why they cannot be used alone for the teaching of the related concepts.

There is a basic initial distinction that has to be made. Addition and the algorithm of addition are two very different things. The main difference is that the concept of addition is semantic, this means that is meaningful, relates to the solution of an actual problem. In this case the problem is to find the total amount of elements that we have when considering two disjoint sets of known quantity of elements. On the other hand the algorithm of addition is purely syntactical, meaning that relates only to a symbolic representation and to mechanic modifications performed on these symbols.

The first basic strategy for adding that we adopt remains within the semantics: Counting the elements of the union. An improved strategy is to take advantage of what is already known and from the number that represents the quantity of elements of the first set, continue the counting with the elements of the second one.

These strategies are valid for every number, but because of human limitations they are not always practical. In order to count one has to keep in memory every passage to an ulterior element, but our work-memory is quite limited in capacity and many studies have established that it can hold on average from 7 to 10 elements. To keep track of all the passages of counting we can use symbols, some kind of sign or mark corresponding to each element considered. The use of symbols for counting resolves the problem posed by the limit of our memory but introduces another problem that is caused by a different human limitation. Humans can distinguish at a glimpse only small quantities of elements. For example, if somebody shows me a picture, just for a fraction of a second, with two or three elements I have no problem to distinguishing these amounts. However, if the elements are 7 or 8 I'm have difficulties to determine the exact quantity at a glimpse. This effect is known under the name *subitizing*. So even if the symbols we use for counting preserve the information we need, in order to retrieve it, we need to organize these symbols somehow. A commonly used strategy is to form groups of a certain predefined quantity of elements, and also groups of groups and so on, maybe using new symbols to indicate different kinds of groups. Thus we introduce a complex organization to be able to solve the problem within our limitations. Using these instruments, we are not working anymore with the actual quantity that we are unable to master, but we obtain a representation of a quantity that is manageable for us. These organizations are then shared and standardized becoming a numeric system, that is to say a system to represent numbers. There are several substantially different systems to do this, each with their own advantages and disadvantages.

Turning back to the problem of adding numbers, when the quantities involved are large, the counting strategies become too difficult and too long to implement. A solution would be to obtain a representation of the sum from the representations of the numbers that we want to add (So we want to construct a syntactic strategy). In order to do this we need a representation system adequate to this objective and at this particular point the Hindu-Arabic system shows all its power. Using this system we introduce the first serious algorithm that we teach at school: The algorithm of the sum.

Have you ever tried to precisely write all of the instructions you have to follow to get the sum of two generic numbers with the algorithm? You'll notice that even if this algorithm is known for most children around 6 or 7 years, the precise description of every step that has to be followed is quite complex.[2]

As a starting point for the addition algorithm, we need to write the two numbers to be summed in an appropriate notation, this is the modern Hindu-Arabic notation with base 10, but the algorithm is not very different if we use a different base. Both numbers will be represented by a succession of ciphers (0,1,2,3,4,5,6,7,8,9) linked to the number in the usual way. Next step is to write these successions one above the other aligning the corresponding ciphers. Maybe one of the representations is of length *n*, and the other of length *m*. If *n* and *m* are different numbers (lets say $m<n$) then we complete the second representation with $m-n$ zeros to the left. The representation of the result will appear in a third line under them when the process is terminated. The initial setup looks something like the following scheme:

$$\begin{array}{ccccccccccc} & a_n & a_{n-1} & & & \ldots & & & a_2 & a_1 & a_0 \\ & 0 & 0 & \ldots & b_m & & b_{m-1} & \ldots & b_2 & b_1 & b_0 \\ (c_{n+1} & c_n & c_{n-1} & & & \ldots & & & c_2 & c_1 & c_0) \end{array}$$

The idea now is two obtain the final representation by "adding" the ciphers on each column, beginning from the rightmost column and proceeding to the left[3], eventually transporting the exceeding quantities to the next column with rules that will be specified. It seems that somehow we have transformed the problem of adding large numbers into a series of additions of small numbers. But such an interpretation of this algorithm is deceiving because it confuses the numbers with the ciphers representing the number. The algorithm uses the representation of the numbers, therefore it operates on the ciphers and not the numbers themselves, and it can be performed without actually using the concept of addition. The advantage of a purely syntactical algorithm is that it can be programed in a machine that can't understand the "idea" of addition and still obtain the correct result. In order to do this firstly we need to record the results of the sums of the ciphers within a table. This table can then be recorded in the memory of our automatic executor of the algorithm. Such a table would look something like this:

[2] In order to be brief, the following description of the algorithm uses notations that require the reader to have a certain mathematical background, like the concept of natural number. I believe that it's not hard to imagine how this is translated for students without it.

[3] It may be interesting to notice that proceeding from right to left is a reminiscence of Arabic writing.

+	0	1	2	3	4	5	6	7	8	9
0	0	1	2	3	4	5	6	7	8	9
1	1	2	3	4	5	6	7	8	9	0*
2	2	3	4	5	6	7	8	9	0*	1*
3	3	4	5	6	7	8	9	0*	1*	2*
4	4	5	6	7	8	9	0*	1*	2*	3*
5	5	6	7	8	9	0*	1*	2*	3*	4*
6	6	7	8	9	0*	1*	2*	3*	4*	5*
7	7	8	9	0*	1*	2*	3*	4*	5*	6*
8	8	9	0*	1*	2*	3*	4*	5*	6*	7*
9	9	0*	1*	2*	3*	4*	5*	6*	7*	8*

Using this table a cipher is obtained considering the cell that is the intersection of the row and the column corresponding to other two ciphers. The stars on this table are references to specific actions in the instructions of the algorithm. Now we are ready to start writing these instructions. To do this we start considering the first ciphers on the right of each expression.

Instruction 1: c_0 is the cipher that corresponds to a_0 and b_0 using the '+' table.

For the subsequent columns we have two cases, if in the previous column the cipher obtained had no star then there is nothing to transport to the next column and we can proceed like before.

Instruction 2: For $i>0$, if the c_{i-1} had no star then c_i is the cipher that corresponds to a_i and b_i using the '+' table.

If instead the cipher of the previous column had a star, it means that in the next column we have to add 1, so the instruction for this case could be "If the c_{i-1} has a star then c_i is the correspondent cipher to a_i and b_i using the '+' table *plus 1*. But this "plus 1" wouldn't be understood by an automatic operator that doesn't know what addition is. To solve this problem several solutions can be adopted. One could be memorizing a function of the solutions of plus 1, this function would look like this:

x	0	1	2	3	4	5	6	7	8	9	0*	1*	2*	3*	4*	5*	6*	7*	8*
plus1(x)	1	2	3	4	5	6	7	8	9	0*	1*	2*	3*	4*	5*	6*	7*	8*	9*

Another solution is combining this function with the '+' table to get a new table to be used in this case, this table would look like this:

+*	0	1	2	3	4	5	6	7	8	9
0	1	2	3	4	5	6	7	8	9	0*
1	2	3	4	5	6	7	8	9	0*	1*
2	3	4	5	6	7	8	9	0*	1*	2*
3	4	5	6	7	8	9	0*	1*	2*	3*
4	5	6	7	8	9	0*	1*	2*	3*	4*
5	6	7	8	9	0*	1*	2*	3*	4*	5*
6	7	8	9	0*	1*	2*	3*	4*	5*	6*
7	8	9	0*	1*	2*	3*	4*	5*	6*	7*
8	9	0*	1*	2*	3*	4*	5*	6*	7*	8*
9	0*	1*	2*	3*	4*	5*	6*	7*	8*	9*

Therefore, memorizing also this second table, the next instruction would be:

Instruction 3: For $i>0$, if the c_{i-1} has a star then c_i is the cipher that corresponds to a_i and b_i using the '+*' table.

Finally, as the length of the resulting succession of ciphers could be one unit longer than the length of the longer succession that generated it there is a final instruction:

Instruction 4: If the c_n has a star then c_{n+1} is 1 otherwise is 0.

By memorizing the initial scheme and the two tables, and following the four instructions, a completely mechanic operator can get to the final representation even without having any idea of what addition means. Also young students end up using the automatic algorithm, as it is very unlikely that when repeating a large amount of sums they will perform the counting strategy each time they have to add the ciphers. More likely, at some point, they memorize the results of the sum of the ciphers, which is equivalent to memorizing the two tables above.

A student on the other hand is not a mechanical operator (even if it can be forced to behave as one) and therefore he 'does' know the actual concept of addition. The concept of addition had to do with counting, but as you can notice from the previous analysis of the algorithm, at no point of it we have to count. Therefore, once we learn this algorithm it is reasonable to think that we could have the following problem:

For example imagine I add two large numbers, like three hundred and forty five and two hundred and sixty seven, using the algorithm. Subsequently, starting from three hundred and forty five I add one unit two hundred and sixty seven times. How can I be sure that I will get the same result I got using the algorithm? In general the question is:

Using the algorithm we get the representation of a quantity, without using the concept of addition. Would we get the same quantity, directly using the concept of addition (counting the union set)?

The answer is yes, there is a proof that using the proprieties of the Hindu-Arabic notation and the proprieties of addition arrives at the correctness of the algorithm.

But if the accent is put on training to calculus ability, generally the decision is to avoid any hint to this proof and just go forward to exercising. In the presentation of the algorithm I tried to emphasize its mechanical aspects and how purposely it avoids the actual meaning of the operation. This is the reason why exercising alone has a little chance to create a true link between the algorithm and the meaningful operation that can provide justification to our question.

If the correctness of the algorithm is not justified, its acceptance by the students is forced. This decision, and other analogous for many other topics, may be the origin of the generalized poor opinion that students have of this kind of mathematics.

Some students may even decide (and not declare or even be conscious) that the result of the calculation has nothing to do with the "real addition". Somebody not convinced that the algorithm has something to do with addition can still be proficient at its use (even machines can do this), but will be in trouble when facing real problems which solution involves addition. The phenomena of students capable of performing calculations but not capable of solving a problem that is solvable by the operation that corresponds to the same calculations, has been largely observed in our modern schools and I believe that this is the cause of the problem in most cases.

Finally, by training students in the unmotivated acceptance and use of automatic procedures like this one, it's very likely they will feel just like the machines that these algorithms were designed for. To make matters worse, today's easy access to calculation machines makes the insistence on training for calculus ability even more pointless.

2.2. Why the algorithm of addition works

The kind of math taught at school talks about the kind of school we have (and vice versa). The school mathematics intended mainly as training for calculus ability can be read in different ways: In one extreme, as a sign that school is still focused on needs of a recent past and hasn't yet taken account of new realities. On the other extreme a school interested in the formation of obeying citizens capable to do their part in society but in the need of solutions from above for all the rest, rather that creative persons possessing the necessary knowledge to solve most of their problems.

If we accept that the school needs to take account of the new access towards technology that diminishes the importance of calculus training, then consequently we'll want to replace some of the time spent in exercitation with something else. But what is this something else?

Instead of doing calculations to understand the concepts on which they are based, why not using the direct way and teach the concepts and the problems they solve directly? Doing this will probably reduce the time used in calculation training, and therefore the proficiency of students at this task could probably decrease. But this risk can be compensated by the deeper understanding of the problem that allows students to re-build calculations procedures when really needed instead of just remembering them. From the analysis of the example proposed in the last section it follows that an innovative didactic approach to addition should take into account the correctness proof of the algorithm, hopefully proposing activities that will include the main concepts and proprieties involved in it. In this way the students can arrive to a profound understanding of "why" the algorithm works. Let's take a closer look at this.

First, what is the proof of the correctness of the algorithm?[4]

In the algorithm we started by representing the numbers with a succession of ciphers. This means that if a is the number represented by the succession $\boldsymbol{a_n a_{n-1} \dots a_2 a_1 a_0}$ then

$a = a_n \times K^n + a_{n-1} \times K^{n-1} \dots + a_2 \times K^2 + a_1 \times K^1 + a_0 \times K^0$

for some natural number K that is called the base of the number representation system [5] (in the algorithm described before K was 10, but now we can give the general proof for any base). An important constraint in this representation is that all the numbers multiplying the several powers of K must be smaller than K.

---

[4] As in the previous section, for reasons of briefness, I use notations that presuppose some mathematical background.

[5] Previously showing that every number can be represented in this way.

Also if b is the number represented by the succession $b_m b_{m-1} \dots b_2 b_1 b_0$ then

$b = b_m \times K^m + b_{m-1} \times K^{m-1} \dots + b_2 \times K^2 + b_1 \times K^1 + b_0 \times K^0$

but supposing that m≤n we can also write the equivalent expression

$b = b_n \times K^n + b_{n-1} \times K^{n-1} \dots + b_2 \times K^2 + b_1 \times K^1 + b_0 \times K^0$

where all the b_i with i between m and n (n included) are zero.

Now, wanting to consider the addition we can immediately write

$a+b = (a_n \times K^n + a_{n-1} \times K^{n-1} + \dots + a_1 \times K + a_0) + (b_n \times K^n + b_{n-1} \times K^{n-1} + \dots + b_1 \times K + b_0)$,

and thanks to the commutative and associative proprieties of addition and to the distributive propriety of multiplication in addition we can rewrite

$a+b = (a_n + b_n) \times K^n + (a_{n-1} + b_{n-1}) \times K^{n-1} + \dots + (a_1 + b_1) \times K + (a_0 + b_0)$.

If in this last expression all the single factors of the powers of K are all smaller than K, the constraint for our notation is fulfilled, and we get the final representation of the result. But the sum of two numbers smaller than K is smaller than 2K-1 and nothing assures that is smaller than K. To see what happens when one of these factors is greater or equal to K we consider in our current representation of the sum the terms $(a_{i+1} + b_{i+1}) \times K^{i+1} + (a_i + b_i) \times K^i$. If the sum $a_i + b_i$ is greater than K, since it is still smaller that 2K-1, we can express it as K plus a number smaller than K-1, let's call this number c_i. Now our terms are transformed into $(a_{i+1} + b_{i+1}) \times K^{i+1} + (K + c_i) \times K^i$.

Distributing we get $(a_{i+1} + b_{i+1}) \times K^{i+1} + K \times K^i + c_i \times K^i$

which is equal to $(a_{i+1} + b_{i+1}) \times K^{i+1} + K^{i+1} + c_i \times K^i$

and finally factoring together the first two terms $(a_{i+1} + b_{i+1} + 1) \times K^{i+1} + c_i \times K^i$

By working like this, we managed to get a factor c_i of K^i that is smaller than K. While doing this we also affected the factor of K^{i+1} that now is a unit larger. If this factor was equal to $K-1$ now its equal to K and therefore the treatment we give it (which is different for numbers smaller than K) changes depending on what we did with the factor of K^i. Translated in the algorithm, this is the justification of why we proceed from the right column to the left.

The final representation of the sum is $c_{n+1} \times K^{n+1} + c_n \times K^n + c_{n-1} \times K^{n-1} + \dots + c_1 \times K + c_0$. Every c_i in this representation is smaller than K and they are calculated in the following way:

- c_0 is $a_0 + b_0$ if this sum is less than K, while c_0 is $a_0 + b_0 - K$ otherwise;

- for c_i with i between 1 and n we consider several cases

	$a_{i-1}+b_{i-1}<9$	$a_{i-1}+b_{i-1}=9$ and $c_{i-1}=9$	$a_{i-1}+b_{i-1}=9$ and $c_{i-1}=0$	$a_{i-1}+b_{i-1}>9$
$a_i+b_i<9$	$c_i = a_i+b_i$		$c_i = a_i+b_i+1$	
$a_i+b_i=9$	$c_i = a_i+b_i$		$c_i = a_i+b_i+1-10$	
$a_i+b_i>9$	$c_i = a_i+b_i-10$		$c_i = a_i+b_i+1-10$	

- c_{n+1} is equal to 1 if a_n+b_n is greater or equal to K, while c_{n+1} is 0 otherwise.

And this provides justification for the rest of the algorithm.

The proof as presented here is too general, too difficult to be presented to a 7-year old. Anyway it allows me as a teacher to design a successful and meaningful teaching technique. Let me show you now some examples: A fundamental aspect of this demonstration is the role of the positional notation, this knowledge suggests that a good idea is to present a few work-examples were numbers like 2361 are expressed as 2000+300+60+1 revealing the hidden proprieties of the notation. Working the sum of 2361+875 as (2000+300+60+1)+(800+70+5) using the commutative and associative proprieties of addition [6] to get 2000+(300+800)+(60+70)+(1+5) which is 2000+1100+130+6, using the meaning of the notation again 2000+(1000+100)+(100+30)+5, and the associative propriety again to get (2000+1000)+(100+100)+30+5 and finally 3000+200+30+5 to obtain 3235. These kind of examples is already a strong suggestion that the rules of the algorithm are adequate to the meanings and not arbitrary.

The acquisition of a meaningful understanding of the algorithm can be tested, but rather than using more mechanical application of the algorithm teachers can come up with activities that prevent the students to directly apply the rules of the algorithm and forces them to reflect on how it works. For example the student can be presented with worked out examples of additions where some ciphers are missing (like ■9+1◆=98)

Another important observation I can make as a teacher knowing this proof, regards the cases illustrated in the table of c_i. There are four different operations that can be performed between the ciphers of a column to get a cipher of the result: The sum of the ciphers above, just the second cipher of this sum, or one the previous two but adding one. There are twelve different motives to choose among these alternatives. As you can see this can be very

[6] We could go further and also ask why should we accept the commutative and associative proprieties of addition, fundamental to this kind of justification. The arguments for the acceptance of the properties can only be based on experience and on the characteristics of the problematic situations we want to solve.

confusing (and many students become confused) but the knowledge of this complexity allows me to better choose my examples in order to cover all the different alternatives and also to understand with which particular point a student is having trouble. A student's confusion may have many different origins. For example it could be related to a psychological issue, and teachers need to be prepared to deal with these as well. But most times the confusion has to do with mathematical difficulties intrinsic to the concept that is being studied. Mathematical concepts are often very articulated with a complex internal organization. The teacher needs to know and master this articulation and complexity in a very detailed way in order to understand the student's difficulties and help him/her get around them. For example imagine that two students use the algorithm with the numbers 758 and 542 and both make mistakes and obtain respectively 1200 and 1290. Should the teacher consider equally the mistakes of these two students? Stating that the correct answer is 1300 won't probably help much any of them, neither repeating the exercise. If we analyze their solutions at the light of our knowledge of the complexity of the algorithm and its correctness proof, we can say much more:

The first student made a mistake in the third column. In the table of c_i this was the case of the third column and third row, that is to say when in the previous column the sum of the ciphers was nine but the cipher of the result was zero. This means that this third column was affected not only by what happened with the ciphers of the second column but also by what happened in the first column (which caused the cipher of the result in the second column to be zero rather than nine). This student probably has trouble understanding that the result in a column can be affected by other columns other than the one immediately to the right and as a consequence may be confused about how to proceed in this case or also may have created an alternative explanation about what to do in this case. On the other hand the second student made a mistake on the second column, which in the table was the case of the fourth column second row, that is to say when the sum of the cipher of the previous column is greater than nine while on the present column is exactly nine. In this case the result of the first column was ten and probably this student thought that being ten the base of the numeration system, it was after this number that the procedures had to change, and not after nine. Or maybe it's the classical confusion between the relationships 'greater' and 'greater or equal'. The other cipher that is not correct in the work of this second student is a consequence of the mistake on the previous column and in fact shows a correct application of the rules, so the intervention should be only about the previously mentioned problem. The analysis that I've just done was a direct consequence of the knowledge of the correctness proof of the algorithm of the sum and directly suggested very different possible interventions by the teacher in each case. This is why this kind of mathematics is recommended for teachers, not just to widen their mathematical culture, but as a fundamental tool for their daily work.

These kind of activities and teacher interventions provide enough hints to the students so that they can accept for themselves and adopt critically the procedure that will help them solve a great variety of problems. The results achieved can be in great contrast to what we get by training in the use of the algorithm. The student won’t feel that is being treated as a machine since he is given the elements to judge the adequacy of the procedure to the solution of the problem, and also because the focus isn’t put on the proficiency achieved in the use of the algorithm but on the profound understanding of why it works and how, and this task can’t be done automatically but only by an intelligent person thinking with a critical attitude.

3.0. Math for freedom

The example presented was intended to illustrate what is good school mathematics in the case of addition. However, giving a general characterization of what school math should be is tricky, as each particular content has distinctive features and because of that, such characterization could be misleading, so it's only with that in mind that I say the following.
Good school mathematics should:

- Start from real problems, in order to be meaningful;
- have a clear view of its place in culture and its importance;
- identify aspects that seem relevant and discard others (because of human limitations dealing with the complication of reality);
- organize the aspects that were retained relevant, add others that may help[7] and generalize;
- work with precision and arrive to the root of the problems;
- insure correctness by justifying properly;
- encourage a critical attitude when thinking;
- avoid purely mechanical activities intended for machines and favor other activities that need typically human capacities as imagination, creativity and choice;
- be an instrument of freedom by providing good tools for solving problems.

In most cases there is a good indicator for appropriate school mathematics, and that is the question "why?" If the presentation of an argument is deep and precise in answering this question (with the right answer) at every step, it's probably on the good path. Otherwise, if certain passages or procedures are given without any kind of justification (or even forced by false justifications) then they probably need revision.

[7] To manage the complication of reality sometimes we not only consider certain aspects of the problem discarding all others, but also sometimes we add aspects that don't belong to the situation in order to make our representation more manageable. Think about the study of the population dynamics: These problems deal with individuals but the most manageable mathematical approach to it uses continuous exponential functions. Continuity is not a characteristic of the problem that is being studied, as a part of an individual has no meaning, but adding this aspect allows a better manageability of the problem.

3.1. Constructivism in education

Underlying this proposal you can notice the position that concepts are not transmitted by the teacher to a passive student, but constructed actively by the student from his experience and previous knowledge with the guidance of the teacher. This is the foundational argument of the constructivism. I would like to point out that mathematical studies in logic provide a support and justification for this argument. In any kind of learning experience communications is mediating the relationship among students, teachers and the subject. Some results in mathematical logic pose strong limits to the power of communication that we cannot overlook. Here is a non-technical explanation that correctly captures the spirit of these results:

"The language is a communication tool. In a communication there is somebody (say a writer) that speaks asserting something about a world that he knows, and somebody (say a reader) that listens trying to understand the meaning that the stated assertions are intended to convey, by figuring out a situation that respects what is asserted (say the story).

It is easy to accept a plurality of interpretations if not all the details are specified, but just those that are relevant to the story. Thus the reader can let his fantasy free to determine the untold features, and can easily imagine several interpretations. He can also add new names either of individuals or of relations and enjoy choosing their interpretations.

It could be objected that the meanings of the added names are chosen by the reader in such a way as to respect all the statements of the tale, therefore, no matter how the new names are interpreted, the original story remains unchanged. Certainly so, and it should be so, but still different readers can have different interpretations, making the situation envisioned by the writer not univocally determined. Even if the writer was thinking of certain aspects as negligible for the story, a reader may find them relevant. In any case, there is no univocal way to interpreter the description, and the assertions of the story do not force the situation imagined by a reader to coincide with the one envisioned by the writer.

If, instead, every slightest feature is described and the story includes all the sentences of the language that are true in the situation that the writer intended to describe, is it still possible to imagine different interpretations?

If in the story it is stated explicitly that the individuals considered are in a exactly specified finite number, then also the number of all the possible relations among them is well determined by the possible combinations and each one is fully described, according to the present assumption.

Thus, every name has to be interpreted in one of the manners already considered and described, making the reader's interpretation univocally determined.

Whereas, if in the described situation the number of possible individuals, including those that are not explicitly mentioned, may either be any finite number or an infinite number (as it is usually in mathematics), then new names for individuals can be added and each one of them can be interpreted as someone, not yet described, having features that make the interpreted situation definitely different form the one envisioned by the writer. One might object that languages have tools to refer to totalities. True, and this allows for interesting uses of the same totalities, but even thus languages cannot univocally determine all the elements of a totality, leaving room for different interpretations. This statement could appear vague as long as it is not analyzed in details, but, if we do so in due time, we reach the result of logic that no theory with arbitrarily large finite models (or with infinite models) is categorical, i.e. has a unique model up to isomorphism.

Ruggero Ferro (Representation and Explanation in the Sciences, Colloque de Louvain-la-Neuve, Belgique, 26-28 april 2011)

This means that, at least in the case of most mathematical concepts (no matter what language is used and how accurate and extensive the description is) there is no way to know if the idea that the reader develops is the same one that the writer intended to describe.

As a consequence there cannot be a direct transmission of concepts form the teacher to the student.

3.2. Didactics by problems

The math for freedom proposal has also many contact points to what is called 'didactics by problems', I would like to illustrate this with the words of one of the main authors in the field of problem solving: G. Polya: "*Studying the methods of solving problems, we perceive another face of mathematics. Yes, mathematics has two faces; it's the rigorous science of Euclid but its also something else. Mathematics presented in the Euclidean way appears as a systematic, deductive science; but mathematics in the making appears as an experimental, inductive science. Both aspects are as old as the science of mathematics itself. But the second aspect is new in one respect; mathematics in "statu nascendi", in the process of being invented, has never before been presented in quite this manner to the students, or to the teacher himself, or to the general public."* But there is more to be said about the role of problems within didactics. For a good didactics by problems it's not enough to just present to the students problems to solve and focus on the 'heuristics' of problem solving, which are the methods for solving them; but also the focus should be on which problems to present and which knowledge the teachers need in order to assist the student in this task.

Teachers are responsible for presenting centuries of mathematical culture to the students and that won't be achieved by attacking the first problem they find, problems have to be chosen carefully. In order to know which problems to choose to better transmit a whole chunk of culture, the teacher needs to have a deep understanding of the concept as well as an appropriate global view of the subject. Also the knowledge necessary to choose the right problems for each topic is the same that allows the justification of its importance and in this way they are not imposed on the students but properly motivated.

Once the right problems are chosen, students can struggle to solve them, but they cannot be left alone on this task or otherwise it would probably take them forever. Some problems in math took centuries and great ideas to be solved and we cannot expect students to do this on their own in school by just presenting the problem. In his book "How to solve it" G. Polya affirms: *"The student should acquire as much experience of independent work as possible. But if he is left alone with his problem without any help or with insufficient help, he may make no progress at all. If the teacher helps too much nothing is left for the student. The teacher should help, but not too much and not too little, so that the students shall have a reasonable share of the work."* Therefore is up to the teacher to give to the students the right amount of help in the right direction. How do teachers know when to intervene and in which way? This can only be achieved counting on that deep knowledge of the content they have to possess, then they are able to steer the discussion and point out the good ideas that will lead to the solutions.

3.3. Teacher's disciplinary knowledge for teaching mathematics at school

Throughout this entire didactic proposal, there is a big emphasis on the teacher having a deep complete knowledge and understanding of each particular topic. In fact without such an understanding the entire proposal would not make sense. But everybody agrees that to teach mathematics at school it's necessary for the teacher to have a good understanding of the contents involved. Nevertheless, this simple statement's importance is often overlooked, even in teacher's formation. This often happens because people generally consider that a person that has already gone through a certain scholar level has already all the disciplinary knowledge necessary for teaching those contents and that there is not much more to say about them. Therefore the attention is focused only on the general teaching methods or on the learning of more advanced mathematics even if the relationship with school mathematics is not clear.

I argue that focusing on general teaching methods without deepening on the mathematical content to be taught cannot be effective. The teaching strategies won't be incisive if based only on general considerations, overlooking the singularities of the content being taught in a particular situation. If this weren't true all school mathematics could be taught with a single strategy. The didactic of mathematics wouldn't distinguish among the teaching of natural numbers, fractions, real numbers, geometry, calculus, algebra, etc. Instead the intrinsic difficulties of each topic are too relevant to be just ignored. Without the consideration of these, the teaching methods for mathematics become only a manifest of good intentions, with no real incidence on student's meaningful learning.

Another idea about the disciplinary content for school teachers that I've encountered many times is the one which considers that the mathematics a teacher has to know should be a lot more vast than the one that is actually going to be taught to the student. A wide mathematical culture provides to the teacher with a global view of mathematics that is very important. This position is sometimes translated into teacher formation programs that include the study of advanced mathematics that could have few links to the topics actually treated at school. But is there a field of mathematical studies with the characteristic of providing a global view of this science and at the same time remaining among the elementary arguments of mathematics? Maybe the study of such a field could also have the benefit of providing a more detailed knowledge of the contents to be taught.

4.0. The mathematics that come before the axioms.

Studying a particular content in mathematics can be done from several points of view that vary in deepness.

First level. <u>Technical approach:</u>	The shallowest approach to a mathematical concept is one interested only in how to work with it, in order to use it as a tool for an ulterior objective.
Second level. <u>Intuitive approach</u>:	A second level of deepness is one were the techniques rather than being accepted for motives outside the mathematics, find a justification based on local reasoning, shared or personal experiences and common sense.
Third level. <u>Formal approach</u>:	Even deeper than the previous one is a formal approach, which studies the concept through an organization in an axiomatic system. This organization shows how the several assertions about the concept are interconnected and how they only depend on a short set of assertions called axioms.
Fourth level. <u>Fundamental approach:</u>	There is an even more profound approach that we can call 'fundamental'. At this level we study if the description provided by that axiom system is appropriate to individuate the intended notion and investigate about the completeness and coherence of it. Several other axioms systems may be proposed at this level. An important tool here is mathematical logic that has achieved interesting results (also limitative) to these questions.
Fifth level: <u>Acquisition process approach:</u>	These axiom systems were chosen in order to capture a pre-elaborated concept of the notion, so on an even deeper level, we can study not only if the proposed axioms match the pre-elaborated notion but also how do we come to know this previous notion. Some people call this level meta-mathematic. I don't completely agree with this name as it implies that this kind of studies lie outside of mathematics. I would rather say that they are essential to mathematics. Only at this level we can find clarity and precision about the intuitions, local reasoning, experiences and common sense of level two.

The profoundness of the treatment of a mathematical content at school rarely goes further than level two. However, what level two includes is vague and the teaching at this level may vary greatly in quality. Even in the universities forming research mathematicians most arguments are treated at level three[8].

Instead, the mathematical knowledge necessary to schoolteachers on the elementary topics treated at school should never overlook levels four and five. The reason for this is that they constitute the main (and only) source for the specific teaching methods to a particular topic. These specific teaching methods for a particular topic can then be combined with the suggestions provided by the general teaching methods regarding the whole of mathematics or teaching in school in general which can have their origin from other sources.

I argue that teachers with a good understanding of levels four and five can provide quality teaching of the same contents at level two.

4.1. The two elementary operations have various meanings

The model for the profundity taxonomy proposed is inspired by the evolution of some of the most basic notions in mathematics: The notion of number and the two elementary operations addition and multiplication.

Addition and Multiplication are not only defined among natural numbers but also among integers, rational numbers, real numbers, etc. Nevertheless, in each system the operations don't remain the same. If they would be the same, then they should be, even extensionally, the same functions (i.e. the same set of ordered triplets) but this is impossible because they are total and apply to different sets. Moreover these sets, the number systems, may be introduced as disjoint sets. Indeed integers and positive rational numbers are presented as equivalence classes of pairs of natural numbers and real numbers are presented as equivalence classes of infinite sets of rational numbers. Nonetheless we can define injections of the naturals into the integers, the integers into the rational numbers and the rational into the real numbers. Using these injections we can say that the addition among real numbers is an extension of the

[8] It's quite reasonable that research mathematicians study concepts up to level 3. In order to obtain advanced results in a mathematical field, the questions about the adequacy of the axioms on which that field relies on and the ones about how the notion is acquired are not relevant. These must be taken for granted in order to make further progress. Often this practical attitude can be paired with certain visions of mathematics that ignore the mathematics that come before the axioms. In such visions the notions described by the axioms are not the consequence of a construction process (but instead they are for example 'a priori' or in an ideal world of ideas,) When these visions of mathematics are held in general and not only for pragmatic reasons they result incomplete, because they don't consider a great part of the mathematical work and are consequently inadequate.

addition among rational numbers and this is an extension of the addition among integers, which is itself an extension of the addition among natural numbers.

Not only they can be seen as extensions of one and other, but also these different additions and multiplications share some common meaning: Addition is always some way to "put things together" and multiplication is a "repeated addition" but these characterizations are quite vague and in each context what we interpret as "putting together" and "repeating" changes greatly. Therefore these operations don't even have the same meaning when passing from one system to the next. As an example, remember that the addition among natural numbers meant to proceed, beyond the first element, repeating the passage to the next element as many times as the second number indicates. But what could ever mean to repeat the passage to the next element -5 times or 1/5 times? And what about in the case of rational or real numbers where there is no such thing as a next element? For multiplication the contrasts are even greater. This same question clearly shows that the meaning of the operation is somehow profoundly transformed when we change context. Meanings are different because the operations in each system are meant to deal with substantially different problems. The problems tackled by each system can be summarized as follows:

The natural number system is developed to compare quantities of elements through counting, and therefore manages situations of multiplicity involving discrete quantities of elements.

- The addition of two natural numbers is a way to determine the quantity of elements of a union of two disjoint sets when the quantities of the original sets are known.
- The product in this system is a way to determine the quantity of elements of a set that is the union of two by two disjoint sets, each of them of the same known quantity of elements.

The integer number system is developed to manage situations of multiplicity involving discrete quantities of elements, of two different kinds, one considered positive and one negative. The value of one positive element is opposite to the value of one negative element; hence two of these quantities may have the same value if one can be obtained from the other by cancelling pairs of opposite elements. If all the opposites elements in a certain quantity are cancelled we obtain a quantity featuring either only positive elements, or only negative elements, or empty.

- The addition in this system is a way to determine a quantity equivalent to the one obtained by considering together two of these quantities preserving the distinction between positives and negatives elements.

- The product is a way to determine a quantity equivalent to the one obtained by repeating the addition of the same quantity by itself as many times as the elements of the second quantity, if this second quantity has only positive elements; or the quantity equivalent to the one that becomes empty after the first quantity is repeatedly added to it as many times as the elements of the second quantity, if these were negative.[9]

The rational number system manages situations of multiplicity involving quantities were the elements can be divided into any number of equal parts, and therefore expressed with an ordered couple of numbers, the second one, called denominator, is a natural number indicating in how many parts the unit is divided and the first one, called numerator, is an integer indicating how many of these parts are considered. From one of quantities of this type we obtain an equivalent quantity by subdividing each part into *n* parts and at the same time consider *n* of the new parts for each one of the original parts.

- The addition in this system is a way to determine an equivalent quantity to the one obtained by considering together two of these quantities, taking account that their parts may be of different size. This can be done because is possible to find equivalent expressions of both numbers featuring the same kind of partitions of the unit, that is to say the same denominator.
- The product in this system is a way to determine an equivalent quantity to the one obtained by considering *a fraction of a fraction*. That is to say, by dividing each part of the first quantity into as many parts as indicated by the denominator of the second quantity, and then for each one of the original parts, consider as many of the new parts as indicated by the numerator of the second quantity.

The real number system is constructed to measure continuous magnitudes, and therefore manages situations of multiplicity involving a particular idea of continuity. Each real number is defined by an adequate infinite set of rational approximations.

- The addition in this system is a way of obtaining the measure of a magnitude equivalent to the one obtained by adjoining other two magnitudes of known measure.
- The product in this system is a way of obtaining a measure of a magnitude that has a ratio to a magnitude of given measures that is equal to the ratio between the other magnitudes of given measures and the unit.

[9] Assuming the initial quantities are given in the equivalent form featuring only one kind of elements.

This brief summary of the problems that the operations in each system resolve is enlightening to understand the different meanings that we assign to the expressions "putting together" and "repeated addition" in different contexts.

As these operations extend one another, some of the meaning is lost and some is transferred to the extended operation in the new system. This indicates that the meaning of the operations corresponding to the real number system applies to all the other systems too, so these can be considered as the "essential meanings" of addition and multiplication. Nevertheless, in order to perform the additions or multiplications among real numbers one has to use the approximations, and hence use the operations among rational numbers, also the operations among rational numbers are performed using the operations between integers, and in this manner we realize that the operations in each system rely ultimately on the original operations among natural numbers. Thus, from another point of view, the meanings of the operations among natural numbers, which are also those more manageable for us, are the fundamental ones.

As a consequence all these operations are not only different but also refer to the solution of different problems. Furthermore they have different proprieties, for example: Natural addition and multiplication always produce a number greater or equal than the input numbers (for multiplication the only exception is when one of the factors is zero). This is not true for the operations among integers, rational or real numbers. Also, the operations between two natural, integer and rational numbers are effective in the sense that the output can always be obtained with a procedure with a finite number of steps. This is not true for the operations between two real numbers. On the other hand they have some properties in common, all systems share the associative, and commutative proprieties for addition and multiplication, and the distributive property.

The properties of the operations that are conserved or not in the extensions are determined by the solution of the problems that the extended operations have to reflect. It's usual to present this topic noting that the addition and multiplication in a system are as they are in order to preserve the previously mentioned formal properties, but these properties can't be taken as a starting point without any previous knowledge of the problems that the extended operations have to solve. The problems to be faced come before the operations and their properties. In fact the operations are designed in such a particular way to reflect the specific characteristics of each problem. This is most evident in the multiplication of real numbers, were the problem of the ratio among continues magnitudes was already a Greek one, while the development of the real numbers is much more recent. The efforts on the field of fundaments that regard the number systems allow us to keep track of the connection between these topics in mathematics and real life problems that ultimately justify the choices made in the construction of these concepts.

4.2. Profound school mathematics: An incomplete picture

In the case of the teaching of addition, it was very clear how the deep understanding of all the aspects of the concept was essential to guide effective didactic designs for this argument and also guide teachers in the actual interventions were they can provide help to the students constructing this concept. But where does this understanding come from? Historically most of the results that allowed us such understanding are quite recent.

At profundity level one and two, addition is as old as it gets, but the formalization of addition at level three is much more recent. Peano's axioms for arithmetic are just about 100 years old.

At level four the study of the fundaments of arithmetic begins with Peano's axiomatization and continues through most of the past century with important results as recent as 1970's. The induction principle, In order to be profoundly analyzed, requires second-degree logic that has fully developed only recently.

As for level five, how we acquire the notion of number (and addition) is still a matter of today's discussions.

This is one of the most elemental contents of school mathematics. The situation is similar for the whole area of the number systems mentioned in the previous section. However, if we analyze the profoundness of all the other main topics of school mathematics we'll find that for levels greater than three the knowledge available today is very new, or incomplete, or even inexistent, for the simple reason that the research at fundamental or acquisition levels on these topics has not yet been done.

Even for many elementary topics in mathematics, the development of levels 4 and 5, which are those most useful to didactics (intended as specific teaching strategies for each topic), is still a work in progress, therefore the deep and complete knowledge that teachers need, is not even available for us to acquire!

5.0. The polynomial functions

As an example of an elementary topic of school mathematics in which the available mathematical knowledge about it is not deep enough for the needs of teaching, take the polynomial functions.

Polynomial functions are some of the simplest functions we can build. In fact the study of functions started with polynomials and they became the prototype for the study of every other function. The reason is that polynomials are formed only by additions and multiplications, which are the elementary operations, and thus explore the potentiality of the basic operations to solve problems. Also, by involving only the two fundamental operations polynomials are surely easy to master. In fact, from an analytical point of view the algorithms to calculate limits, derivatives or integrals of polynomials are the simplest of all. Polynomials are not only easy to master, but they also are a good tool to model a very large number of problems. As if this wasn't enough to state the importance of polynomials functions, they also have the characteristic that can be used to approximate many other types of functions.

For all these reasons, polynomial functions are a fundamental topic for school mathematics that intends to be an instrument of freedom. But in order to be so, it's very important that the procedures implemented (for example the basic operations between two polynomials) find a proper justification.

The polynomial functions we use to solve a problem may also be of a high degree, and these have a much higher level of complexity than low-degree polynomials. In fact it has been proved that there cannot be a general solution formula for polynomials of degree five or higher. Thus a basic and very important aspect when handling polynomials is the possibility to express them as a product of simpler polynomials. In the environment of real numbers this is not always possible as we can easily come up with examples of second-degree polynomials that can't be expressed as a product of simpler polynomials (irreducible), take x^2+1 for instance. Nevertheless, all the other real polynomials which degree is greater than 2, can be expressed as a product of real polynomials of the first or second degree. The fact that there are not irreducible polynomials of degree greater than 2 is at the same time comforting and puzzling. Comforting because at least theoretically it will allow us to study complicated high degree polynomials by analyzing their much simpler factors, and puzzling because is difficult to see immediately why this result is true. Given that the degree and coefficients of a polynomial can be chosen at will among infinite possibilities, how come I can't come up with an irreducible polynomial of higher degree?

To give an answer to this question the natural way is to examine the existing mathematical proofs of this fact, and we find out that they all refer to the fundamental theorem of algebra: Every polynomial can be expressed as a product of linear factors. But this result is valid only in a wider field: the complex number system. Real polynomials become now a particular case of complex polynomials and for these also another result holds: If a complex number is a root of a real polynomial, its conjugate is also a root of the same polynomial; and the product of the two linear polynomials corresponding to these conjugate roots, is a real second-degree polynomial.

Notice that to prove a fact about real numbers we had to go to a wider and essentially different field, prove a more general result and then come back to the original environment. By doing this, it's quite difficult to spot in the proof the elements that answer the question of why the result is true and separate them from the ones that have to do only with the more general result. It's reasonable to wonder if the use of complex numbers is really necessary to prove a statement that lies completely within the real numbers, and if whether a proof using only real numbers would have the power to better individuate the fundamental reasons for its validity.

5.1. The Fundamental Theorem of Algebra (Gauss's first proof)

In order to illustrate one proof of the fundamental theorem of Algebra I present a brilliant summary by Harel Cain of Gauss's first proof:

"Gauss started with a real polynomial: $X = x^m + Ax^{m-1} + Bx^{m-2} + \cdots + Lx + M$

treating x as an "unbestimmte Größe", an indeterminate. A real linear factor can be written as $x \pm r$*, with* $r \geq 0$*. An irreducible quadratic factor over the reals can be written as*

$x^2 - 2xr\cos\varphi + r^2$ *again with* $r > 0$*, the roots of which are* $r(\cos\varphi \pm i\sin\varphi)$*.*

By substituting $x = r(\cos\varphi + i\sin\varphi)$ *into the polynomial and separating into real and imaginary parts one can form a pair of expressions*

$U = r^m\cos m\varphi + Ar^{m-1}\cos(m-1)\varphi + \cdots Lr\cos\varphi + M;\ T = r^n\sin m\varphi + Ar^{m-1}\sin(m-1)\varphi + \cdots Lr\sin\varphi$

Gauss now notes that if (r, φ) *satisfy* $T = 0$ *and* $U = 0$ *simultaneously, then the polynomial X would be divisible by* $x \pm r$ *or* $x^2 - 2xr\cos\varphi + r^2$*...*

Gauss regards $T = 0$ *and* $U = 0$ *as algebraic curves of order m given in the polar coordinates r and* φ*, drawn on an orthogonal plane with coordinates* $(r\cos\varphi, r\sin\varphi)$*. To see that the curves are algebraic, one just has to use standard trigonometric formulae to* *express U and T as algebraic expressions in r cos φ and r sinφ. To prove the FTA, Gauss wants to prove that there exists a point of intersection between the two curves T = 0 and U = 0. Using his earlier lemma, such an intersection would be a root of the polynomial X, because if U(r,φ) = Re(X(x)) = 0 and T(r,φ) = Im(X(x)) = 0 then X(x) = 0.*

To look for intersections of the two curves, Gauss now studies their intersections within a circle of radius R, and proves

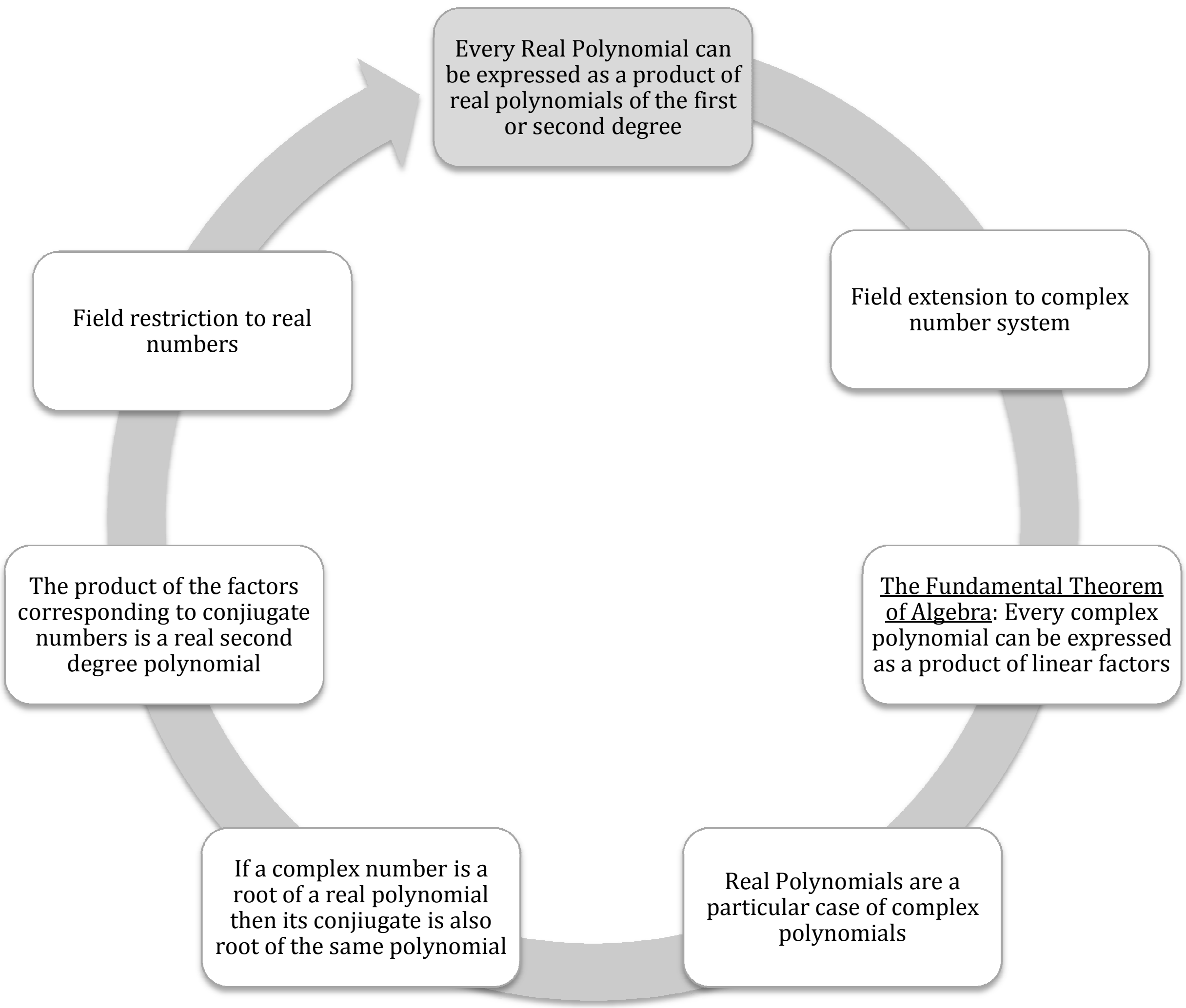
> *For a sufficiently large radius R there are exactly 2m intersections of the circle with T = 0 and 2m intersections with U = 0, and every point of intersection of the second kind lies between two of the first kind.*

This result is proved with complete rigour, but only after the underlying idea is presented intuitively: as R tends to infinity, the curves T = 0 and U = 0 approach the curves Re(x^m) = 0 and Im(x^m) = 0, which are straight lines through the origin. Moreover, the lines where Re(x^m) = 0 alternate with those where Im(x^m) = 0. ... In the next step, Gauss notes that "it can easily be seen" that that the 4m points of intersection of the circle vary very little when R is slightly changed. .... Gauss now reaches the core of his proof: to show that the two curves intersect inside the circle. For this conclusion he gives an intuitive, geometrical proof: he numbers the intersections with the circle sequentially, starting with 0 for the negative x-axis of the plane (which is always part of the solution to T = 0) giving odd numbers to the intersections of U = 0 with the circle, and even numbers to the intersections of T = 0. He now claims that "if a branch of an algebraic curve enters a limited space, it necessarily has to leave it again".

If this remark is to be accepted, then every odd-numbered intersection point on the circle is connected with another odd-numbered point by a branch of the curve U = 0, and similarly for even-numbered points and branches of the curve T = 0. But then, however complicated these connections may be, one can show that a point of intersection exists, in the following way. Suppose no such intersection exists. The point 0 is connected with point 2m through the x-axis, which is always a branch of T = 0. Then 1 cannot be connected with any point on the other side of the axis, so it must be connected with an odd point n, n < 2m. In the same manner, 2 is connected with an even $n' < n$. Note that $n' - 2$ is even. Continuing this way, one ends with an intersection point h connected with h + 2. But then the branch entering the circle at h + 1 must intersect the branch connecting h and h + 2, contrary to our hypothesis.

Thus an intersection point for Re(X) = T = 0 and Im(X) = U = 0 exists, concluding the proof."

As you can see, Gauss starts with the problem of the factorization of real polynomials and uses complex numbers only as a tool to solve this problem. He actually doesn't even need to prove the Fundamental Theorem of Algebra for all complex polynomials but he can limit to complex polynomials with real coefficients. Later, others have shown that the proof can easily be adjusted for all complex polynomials, following the path suggested by this design:



The decision of Gauss doing things on that particular way, reflects that he thought of complex numbers only as a momentary tool to obtain meaningful results among real numbers.

The highlights ideas of the proof are: Dividing the complex polynomial into a real and an imaginary part, which are two surfaces (because they are real polynomials in two variables).

If the zero-level curves for these two surfaces have a common point, the coordinates of this point are the real and imaginary part of a root of the original polynomial. In order to find the intersection of these level curves he considers the terms of highest degree that rule the behavior of the surfaces for big enough values of the variables. Then he considers a big enough circle around the origin and notices that the surfaces outside the circle alternate their zero-level curves. Finally, from this final fact he deducts the existence of the intersection inside the circle.

Analyzing these main ideas of the proof it's difficult to see what is the relationship of the original real polynomial with the polynomials of the real and imaginary part of its correspondent complex polynomial. Moreover, the proof doesn't seem to explain why the zero-level curves of the real part and the imaginary part interweave for large values of the variables. These questions stand in the way of us and our students understanding why the result that every real polynomial can be factored into polynomials of first or second degree holds.

5.2. Complex numbers:

Furthermore, if we were to perform a deep study of the fundamental theorem of algebra in the ambit of real numbers using the complex numbers system then, coherently, we would also need to profoundly study the complex numbers; but the study on these deep levels of this concept is also incomplete.

It's not difficult to formally define complex numbers with its operations. But what about a deeper study of this subject? What results can we find at the level of fundaments? I would say that the most relevant result at this level is the geometric model were a complex number is interpreted as a vector with its origin point at the center of a Cartesian plane. The sum of complex numbers is identified with the usual sum of vectors and the product of two complex numbers can be defined as a vector obtained from other two being its magnitude the product of the magnitudes of the factors and its angle the sum of the angles of the factors. The Cartesian plane model for complex numbers gives us a novel connection between the addition of complex numbers and the already known addition of vectors. On the other hand, the geometric interpretation of the product has nothing to do with the product of vectors. The product of complex numbers, other that being a formal extension of the product of real numbers that conserves the associative, commutative and distributive properties, doesn't seem to have a clear connection to a meaningful problem as the product does in other numeric systems (as discussed in the section 'The two elementary operations have various meanings').

Going even deeper, at the acquisition process level we find ourselves practically with no information: Historically the very idea of complex numbers was retained absurd. There could be no meaning to the solution of the square root of a negative number. Nevertheless it was shown that if we forget for a minute that they are meaningless and we operate with these roots we get to meaningful results. This was the case of the discovery of the first formulas for the third degree equations. Later other formulas were found for this problem that avoided the resource to complex numbers. The fact that using them didn't lead to contradiction but to meaningful results could be read as an indication that there is in fact some meaning to them or it can be read as that they are only useful instruments with no profound meaning.

Maybe a study of complex number applications could give some insight to these opposite positions. Phillip Spencer from the University of Toronto dedicates part of his time to answer questions on the University web page.

http://www.math.toronto.edu/mathnet/questionCorner/complexinlife.html

A Math teacher submitted the question:

"I've been stumped!

After teaching complex numbers, my students have asked me the obvious question: Where is this math used in real life!

Your assistance would be greatly appreciated."

And this is Dr. Spencer's answer:

"(...)

There are two distinct areas that I would want to address when discussing complex numbers in real life:

Real-life quantities that are naturally described by complex numbers rather than real numbers;

Real-life quantities which, though they're described by real numbers, are nevertheless best understood through the mathematics of complex numbers.

The problem is that most people are looking for examples of the first kind, which are fairly rare, whereas examples of the second kind occur all the time.

Here are some examples of the first kind that spring to mind. In electronics, the state of a circuit element is described by two real numbers (the voltage V across it and the current I flowing through it). A circuit element also may possess a capacitance C and an inductance L that (in simplistic terms) describe its tendency to resist changes in voltage and current respectively.

These are much better described by complex numbers. Rather than the circuit element's state having to be described by two different real numbers V and I, it can be described by a single complex number z = V + i I. Similarly, inductance and capacitance can be thought of as the real and imaginary parts of another single complex number w = C + i L. The laws of electricity can be expressed using complex addition and multiplication.

Another example is electromagnetism. Rather than trying to describe an electromagnetic field by two real quantities (electric field strength and magnetic field strength), it is best described as a single complex number, of which the electric and magnetic components are simply the real and imaginary parts.

What's a little bit lacking in these examples so far is why it is complex numbers (rather than just two-dimensional vectors) that are appropriate; i.e., what physical applications complex multiplication has.[10] *I'm not sure of the best way to do this without getting too far into the physics, but you could talk about a beam of light passing through a medium which both reduces the intensity and shifts the phase, and how that is simply multiplication by a single complex number."*

The lack of a profound explanation of complex multiplication talks about the state of the fundaments research on this field and explains my reluctance to use the complex numbers proof of The Fundamental Theorem of Algebra for a profound explanation of why it's true that real polynomials can be expressed using only linear and quadratic factors.

Maybe a proof of this fact using only real numbers, in order to avoid the unclear connection with the complex numbers could help us with our doubts.

5.3. Mathematical research inspired by didactics

But there is no proof of this fact using only real numbers and real number methods. We could say that the fundaments level of profundity (in my proposed taxonomy) hasn't been developed enough to answer the questions that have been raised. These questions were clearly originated in didactics. It was the need of properly justifying procedures and ideas to students that face the problem of the factorization of polynomials for the first time that generated these interrogatives.

The understanding of the elementary contents of mathematics is incomplete when we look at it at high levels of profundity, with the fundamental approach or even more with the

[10] The underlining is mine

acquisition process approach. As a consequence, in order to study the didactics of mathematics situated in each particular topic, the researcher is driven to develop actual mathematic research, in order to further complete the deepest levels of the topic that constitute the only source for the specific teaching methods.

It's in this spirit that I've developed from scratch a proof that every real polynomial can be expressed as a product of real polynomials of the first or second degree, and in this proof I've avoided any use of the complex number system. This proof is original mathematical work and by itself can be seen merely as a part of a formal system. On the other hand, only the didactic motivations explain why I developed my proof in such a way, acquiring new meaning in an educational framework.

The next chapter will be strictly confined to my proof of the theorem, leaving the interpretation in fundamental and didactic perspectives for the end.

6.0. A really real proof of the Fundamental Theorem of Algebra

A proof using only real numbers and real number methods that any real polynomial can be expressed as a product of real polynomials of degree one or two.

I'll turn my attention only to even degree polynomials, because polynomials of odd degree immediately have at least one root by the Intermediate Value Theorem, therefore a linear factor and an even degree factor. Also, without loosing generality, I can reduce to consider only monic polynomials

I begin with the long division of a generic even degree monic polynomial, by a generic monic quadratic polynomial.

$$(x^n + a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \cdots + a_2x^2 + a_1x^1 + a_0)$$
$$= (x^2 + zx + w) \times \left(\sum_{i=1}^{n-1} k_i x^{n-i-1}\right) + R_1x + R_0$$

The goal is to demonstrate that, for some values of z and w, the quadratic polynomial divides the generic one, and in that case it must be that the remainder $R_1 x + R_0$ is zero, that is to say that R_1 *and* R_0 should be zero simultaneously.

The first step is to calculate the coefficients k_i that appear in the quotient in the previous formula in terms of the coefficients of the generic polynomial. To do this I can begin with a recursive formula extracted from the division algorithm. The "(n)" is to remind that these are being calculated for the case of a polynomial of degree n.

$$k_1^{(n)} = 1, \qquad k_2^{(n)} = a_{n-1} - z \qquad , k_{i+2}^{(n)} = a_{n-(i+1)} - wk_i^{(n)} - zk_{i+1}^{(n)} \text{ for i>0}$$

When I calculate several of these $k_i^{(n)}$, grouping the terms with the same coefficient of the original polynomial, I get:

$$k_1^{(n)} = a_n$$
$$k_2^{(n)} = a_n(-z) + a_{n-1}$$
$$k_3^{(n)} = a_n(z^2 - w) + a_{n-1}(-z) + a_{n-2}$$
$$k_4^{(n)} = a_n(-z^3 + 2zw) + a_{n-1}(z^2 - w) + a_{n-2}(-z) + a_{n-3}$$
$$k_5^{(n)} = a_n(z^4 - 3z^2w + w^2) + a_{n-1}(-z^3 + 2zw) + a_{n-2}(z^2 - w) + a_{n-3}(-z) + a_{n-4}$$
$$k_6^{(n)} = a_n(-z^5 + 4z^3w - 3zw^2) + a_{n-1}(z^4 - 3z^2w + w^2) + a_{n-2}(-z^3 + 2zw) + a_{n-3}(z^2 - w) + a_{n-4}(-z) + a_{n-5}$$
$$k_7^{(n)} = a_n(z^6 - 5z^4w + 6z^2w^2 - w^3) + a_{n-1}(-z^5 + 4z^3w - 3zw^2) + a_{n-2}(z^4 - 3z^2w + w^2) + a_{n-3}(-z^3 + 2zw) + a_{n-4}(z^2 - w) + a_{n-5}(-z) + a_{n-6}$$
$$k_8^{(n)} = a_n(-z^7 + 6z^5w - 10z^3w^2 + 4zw^3) + a_{n-1}(z^6 - 5z^4w + 6z^2w^2 - w^3) + a_{n-2}(-z^5 + 4z^3w - 3zw^2) + a_{n-3}(z^4 - 3z^2w + w^2) + a_{n-4}(-z^3 + 2zw) + a_{n-5}(z^2 - w) + a_{n-6}(-z) + a_{n-7}$$
$$k_9^{(n)} = a_n(z^8 - 7z^6w + 15z^4w^2 - 10z^4w^3 + w^4) + a_{n-1}(-z^7 + 6z^5w - 10z^3w^2 + 4zw^3) + a_{n-2}(z^6 - 5z^4w + 6z^2w^2 - w^3) + a_{n-3}(-z^5 + 4z^3w - 3zw^2) + a_{n-4}(z^4 - 3z^2w + w^2) + a_{n-5}(-z^3 + 2zw) + a_{n-6}(z^2 - w) + a_{n-7}(-z) + a_{n-8}$$

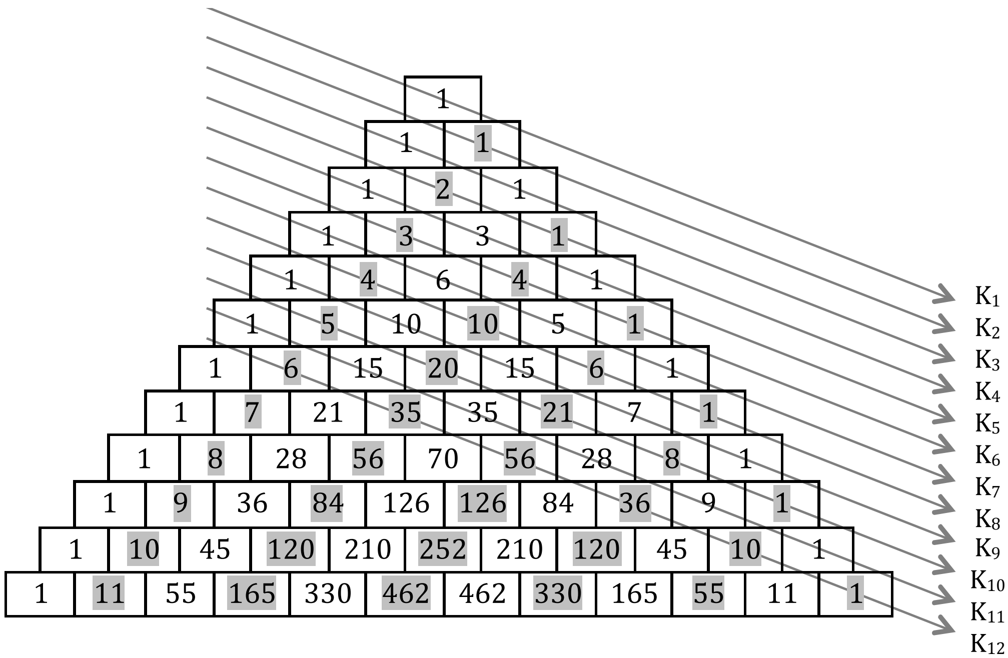
Notice that in these expressions the terms inside the parenthesis repeats, these are:

$$1$$
$$-z$$
$$z^2 - w$$
$$-z^3 + 2zw$$
$$z^4 - 3z^2w + w^2$$
$$-z^5 + 4z^3w - 3zw^2$$
$$z^6 - 5z^4w + 6z^2w^2 - w^3$$
$$-z^7 + 6z^5w - 10z^3w^2 + 4zw^3$$
$$z^8 - 7z^6w + 15z^4w^2 - 10z^2w^3 + w^4$$

There are many interesting properties of these last expressions:

- Each of them corresponds to the coefficient of a_n for a certain $k_i^{(n)}$.
- They alternate in sign: the ones corresponding to an even *i* start with positive sign, and the ones corresponding to an odd *i* start with negative sign.

- Ordering its terms by the power of z, this powers decreases by two from term to term, finding only even powers for even i and odd powers for odd i.
- In this order the power of w increases by one from term to term.
- The number of terms increases every two steps.
- A term with a certain combination of powers of z and w appears in one and only one of these expressions.
- The magnitude of the coefficients of the terms are numbers in offset diagonals of Pascal's triangle, as shown in the next picture:



Each number of Pascal's triangle can be expressed as $\binom{a}{b} = \frac{a!}{(a-b)!b!}$, were the top number is the row, and the second one is the position in the row. A well known fact is that $\binom{a}{b} + \binom{a}{b+1} = \binom{a+1}{b+1}$, that is to say, that a number of the triangle is generated by the sum of the two numbers directly above it.

It seems reasonable to separate the even and odd cases. In the triangle the numbers corresponding to even k_i are shaded, and the ones corresponding to an odd k_i are left unshaded. To obtain a (non)shaded number we consider the three (non)shaded numbers directly above it in the next two rows. Making the subtraction of the two (non)shaded numbers in the upper row, taken from left to right, and adding twice the number in the lower row. With binomial notation it means that:

$$\binom{2p-i}{i-1} + 2\binom{2p-i+1}{i-2} - \binom{2p-i}{i-3} = \binom{2p-i+2}{i-1}.$$

From all these observations follows a non-recursive formula $h_i^{(n)}$:

Definition 1: For *i*>0

$$h_i^{(n)}(z,w) = \left(\sum_{p=1}^{[i+1/2]} a_{n-i+2p-1} \sum_{j=1}^{p} (-1)^{j-1} \binom{2p-j-1}{j-1} z^{2(p-j)} w^{j-1}\right) - \left(\sum_{p=1}^{[i/2]} a_{n-i+2p}\, z \sum_{j=1}^{p} (-1)^{j-1} \binom{2p-j}{j-1} z^{2(p-j)} w^{j-1}\right)$$

I'll prove now, by complete induction, that $h_i^{(n)}$ and $k_i^{(n)}$ are the same.

Theorem 1: $h_i^{(n)} = k_i^{(n)}$ were $h_i^{(n)}(z,w) = \left(\sum_{p=1}^{[i+1/2]} a_{n-i+2p-1} \sum_{j=1}^{p}(-1)^{j-1}\binom{2p-j-1}{j-1} z^{2(p-j)} w^{j-1}\right) - \left(\sum_{p=1}^{[i/2]} a_{n-i+2p}\, z \sum_{j=1}^{p}(-1)^{j-1}\binom{2p-j}{j-1} z^{2(p-j)} w^{j-1}\right)$ and $k_1^{(n)} = 1,\quad k_2^{(n)} = a_{n-1} - z,\ k_{i+2}^{(n)} = a_{n-(i+1)} - wk_i^{(n)} - zk_{i+1}^{(n)}$ for i>0

Proof in appendix

Now from this new k_i formula it's possible to give expressions for R_1 *and* R_0

$$R_1^{(n)}(z,w) = a_1 - wk_{n-2}^{(n)} - zk_{n-1}^{(n)} = k_n^{(n)}(z,w)$$

$$R_1^{(n)}(z,w) = \left(\sum_{p=1}^{n/2} a_{2p-1} \sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i-1}{i-1} z^{2(p-i)} w^{i-1}\right) - \left(\sum_{p=1}^{n/2} a_{2p} z \sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i}{i-1} z^{2(p-i)} w^{i-1}\right)$$

$$R_0^{(n)}(z,w) = a_0 - wk_{n-1}^{(n)}$$

$$R_0^{(n)}(z,w) = a_0 - w\left(\left(\sum_{p=1}^{n/2} a_{2p} \sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i-1}{i-1} z^{2(p-i)} w^{i-1}\right) - \left(\sum_{p=1}^{n/2-1} a_{2p+1}\, z \sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i}{i-1} z^{2(p-i)} w^{i-1}\right)\right)$$

These expressions can be interpreted as surfaces in a Cartesian space of coordinates *z,w,t*; as they are polynomials depending on the two variables *z* and *w*. The intersection of one of these surfaces with a plane is a curve. $R_0^{(n)} = 0$ and $R_1^{(n)} = 0$ are the respective level curves for 0, that is to say that $R_0^{(n)} = 0$ is a curve that is the intersection of the surface $R_0^{(n)}$ with the plane *t*=0 and that $R_1^{(n)} = 0$ is a curve that is the intersection of the surface $R_1^{(n)}$ with the plane *t*=0. If these zero-level curves cross each other, then the remainder vanishes at the intersection, and the coordinates (*z,w*) of the intersection point will be also the coefficients of the second degree polynomial that divides the polynomial I started with, and the answer to the whole theorem.

There are some interesting observations to make about these expressions:

$R_1^{(n)}$ is the algebraic sum of two terms; each of these terms is a nested composition of two sums. Also in $R_0^{(n)}$ we can find two terms that are a nested composition of two sums. Moreover, the inner sums from the terms of $R_1^{(n)}$, and the inner sums of the terms of $R_0^{(n)}$ are the same.

These remainder expressions were made in order to group the terms that have the same coefficient of the original polynomial. Notice that the internal sums that appear in both these expressions don't depend on these coefficients. The result of divisibility that I want to obtain should be valid for every polynomial, and therefore shouldn't depend on its coefficients. This is why we will now concentrate in these inner internal sums, which can be expressed as follows:

Definition 2:

$$S_1^{(p)}(z,w) = \sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i}{i-1} z^{2(p-i)} w^{i-1}$$

$$S_0^{(p)}(z,w) = \sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i-1}{i-1} z^{2(p-i)} w^{i-1}$$

These new expressions of $S_1^{(p)}$ and $S_0^{(p)}$ differ only on their binomial coefficients. Also the variable *z* appears only to even powers. If I substitute Az^2 for *w*, when $z \neq 0$, these expressions become:

$$S_1^{(p)}(z, Az^2) = \sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i}{i-1} z^{2(p-i)} (Az^2)^{i-1} = z^{2p-2} \sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i}{i-1} A^{i-1}$$

$$S_0^{(p)}(z, Az^2) = \sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i-1}{i-1} z^{2(p-i)} (Az^2)^{i-1} = z^{2p-2} \sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i-1}{i-1} A^{i-1}$$

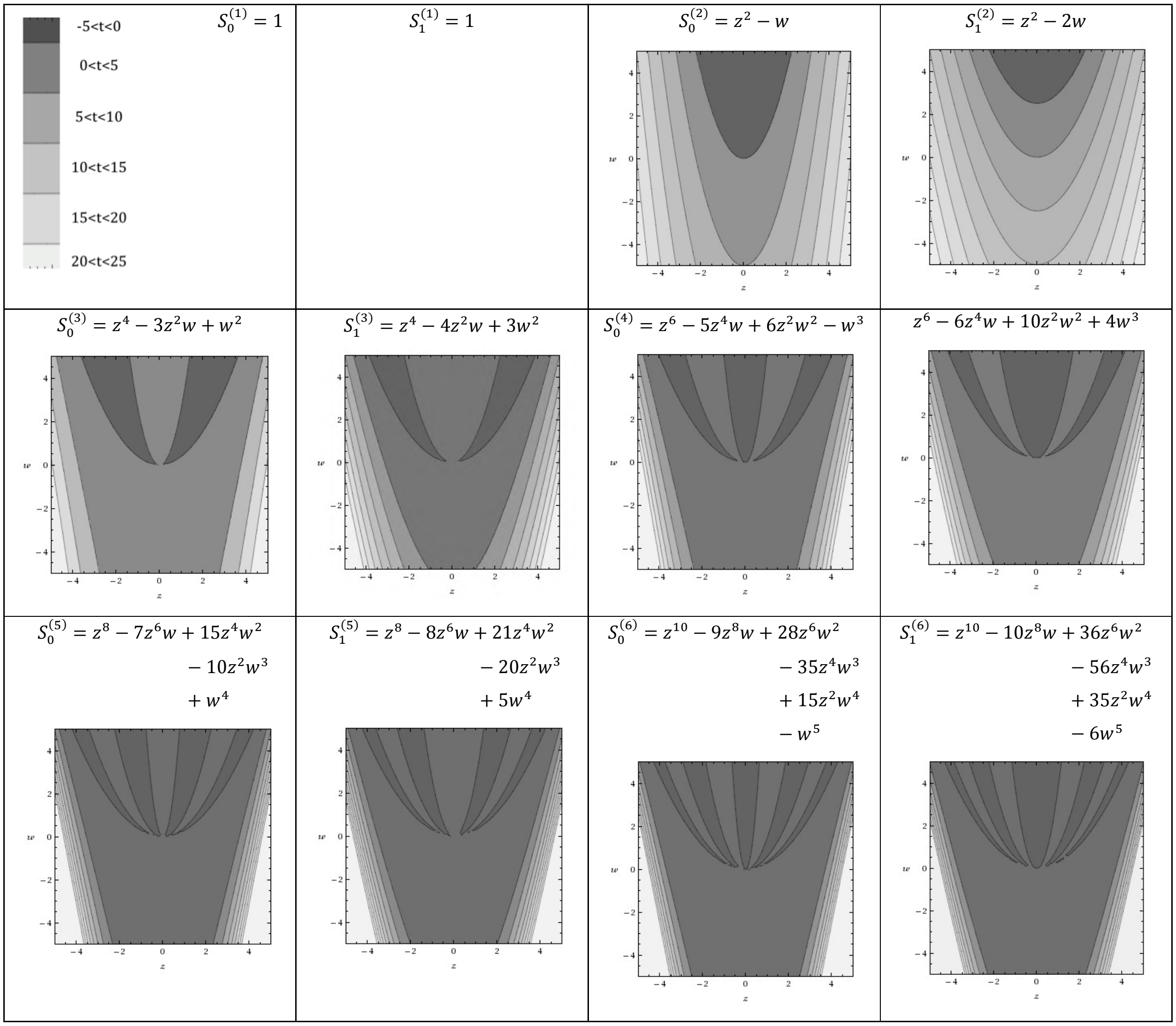
Now I can rewrite:

$$R_1^{(n)}(z, Az^2) = \left(\sum_{p=1}^{n/2} a_{2p-1}\, z^{2p-2} \sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i-1}{i-1} A^{i-1} \right) - \left(\sum_{p=1}^{n/2} a_{2p} z^{2p-1} \sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i}{i-1} A^{i-1} \right)$$

$$R_0^{(n)}(z, w) = a_0 - w \left(\left(\sum_{p=1}^{n/2} a_{2p} z^{2p-2} \sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i-1}{i-1} A^{i-1} \right) - \left(\sum_{p=1}^{n/2-1} a_{2p+1}\, z^{2p-1} \sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i}{i-1} A^{i-1} \right) \right)$$

I'll be interested in studying the behavior of $R_1^{(n)}$ and $R_0^{(n)}$ for large values of the variable *z*, hoping that this study will give enough information to determine the result. This behavior is governed by the terms of the higher degree in each expression. The term of highest degree for $R_1^{(n)}$ in *z* is in its second sum ($S_1^{(p)}$) when *p*=*n*/2. On the other hand, the term of higher degree in *z* of $R_0^{(n)}$ is contained in its first inner sum ($S_0^{(p)}$) when *p*=*n*/2. This is the reason of the sub-indexes 0 and 1 on the previous definition.

But first let's analyze the structure of $S_1^{(p)}$ and $S_0^{(p)}$. Notice that expressed as $S_1^{(p)}(z, Az^2)$ and $S_0^{(p)}(z, Az^2)$ they can be zero only for some values of *A*, which is the curvature of a parabola with vertex at (0,0). Thus there is a set of parabolas with vertex at (0,0) that are zero-level curves of $S_1^{(p)}$ and/or $S_0^{(p)}$.This can be observed in the contour plots of $S_1^{(p)}$ and $S_0^{(p)}$ for *p*=1, *p*=2, *p*=3, and so on:



The observation of these graphs suggests that the level curves are nothing more than this set of nested parabolas with vertex at (0,0). The next theorems will prove this observation.

The substitution $w{=}Az^2$ suggested that the zero-level curves contained parabolas with vertex at (0,0). But in order to prove that there are no other zero-level curves, we have to show that $S_1^{(p)}(z,w)$ can be factored as a product of p-1 polynomials of the type $A^{-1}z^2 - w$. We can accomplish this by writing the long division between $S_1^{(p)}(z,w)$ and $A^{-1}z^2 - w$ with an eventual remainder. Because the variable z appears in $S_1^{(p)}(z,w)$ only to even powers, this remainder doesn't depend on z, and I will show it vanishes for some values of A. The proposed factorization has all the factors linear in w, but the main coefficient of w in the expression that we want to factorize is p, so the final factorization will include a factor p.

Theorem 2-a:

$$S_1^{(p)}(z,w) = \sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i}{i-1} z^{2(p-i)} w^{i-1}$$

$$= (A^{-1}z^2 - w) \times \left(\sum_{q=1}^{p-1} z^{2(p-q)-2} w^{q-1} \sum_{i=1}^{q} (-1)^{i-1} \binom{2p-i}{i-1} A^{q-i+1} \right)$$

$$+ w^{p-1} \sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i}{i-1} A^{p-i}$$

Proof in appendix

The remainder (the last terms) in the right hand side of this equation is 0 when *w*=0 or when

$\sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i}{i-1} A^{p-i} = 0$

This last expression is a shifted Chebyshev polynomial of the second kind. I'll call it $U_p(A)$.

Definition 3-a:

$$U_p(A) = \sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i}{i-1} A^{p-i}$$

I'll now show that as Chebyshev polynomials U_p can be defined recursively:

Theorem 3:

$U_{p+1}(A) = (A-2)U_p(A) - U_{p-1}(A), \qquad U_1(A) = 1, \qquad U_2(A) = (A-2), \qquad p \geq 2$

Proof:

In this proof we use

$$\binom{2p-i}{i-1} + 2\binom{2p-i+1}{i-2} - \binom{2p-i}{i-3} = \binom{2p-i+2}{i-1}$$

$$U_{p+1}(A) = \sum_{i=1}^{p+1} (-1)^{i-1} \binom{2p-i+2}{i-1} A^{p+1-i}$$

Separating the terms *i*=1, *i*=2 and *i*=*p*+1

$$= (-1)^{1-1} \binom{2p+2-1}{1-1} A^{p+1-1} + (-1)^{2-1} \binom{2p+2-2}{2-1} A^{p+1-2} + \sum_{i=3}^{p} (-1)^{i-1} \binom{2p-i+2}{i-1} A^{p+1-i}$$

$$+ (-1)^{p+1-1} \binom{2p+2-(p+1)}{p+1-1} A^{p+1-p-1}$$

$$= (-1)^0 A^{p+1-1} + (-1)^1 2pA^{p+1-2} + \sum_{i=3}^{p} (-1)^i \binom{2p-i+2}{i-1} A^{p+1-i} + (-1)^p (p+1) A^{p+1-p-1}$$

The second term $(-1)^1 2pA^{p+1-2} = (-1)^1(2p-2)A^{p+1-2} + (-1)^1 2A^{p+1-2}$ and the last term $(-1)^p(p+1)A^{p+1-p-1} = (-1)^p 2pA^{p+1-p-1} - (-1)^p(p-1)A^{p+1-p-1}$ so

$$U_{p+1}(A) = (-1)^0 A^{p+1-1} + (-1)^1(2p-2)A^{p+1-2} + (-1)^1 2A^{p+1-2} + \sum_{i=3}^{p} (-1)^{i-1} \binom{2p-i+2}{i-1} A^{p+1-i} + (-1)^p 2pA^{p+1-p-1}$$
$$-(-1)^p(p-1)A^{p+1-p-1}$$

Using the lemma and then the distributive property,

$$U_{p+1}(A) = (-1)^0 A^{p+1-1} + (-1)^1(2p-2)A^{p+1-2} + (-1)^1 2A^{p+1-2} + \sum_{i=3}^{p} (-1)^{i-1} \left(\binom{2p-i}{i-1} + 2\binom{2p-i+1}{i-2} - \binom{2p-i}{i-3} \right) A^{p+1-i}$$
$$+ (-1)^p 2pA^{p+1-p-1} - (-1)^p(p-1)A^{p+1-p-1}$$
$$= (-1)^0 \binom{2p-1}{1-1} A^{p+1-1} + (-1)^1 \binom{2p-2}{2-1} A^{p+1-2} + (-1)^1 2\binom{2p-2+1}{2-2} A^{p+1-2} + \sum_{i=3}^{p} (-1)^{i-1} \binom{2p-i}{i-1} A^{p+1-i}$$
$$+ \sum_{i=3}^{p} (-1)^{i-1} 2\binom{2p-i+1}{i-2} A^{p+1-i} - \sum_{i=3}^{p} (-1)^{i-1} \binom{2p-i}{i-3} A^{p+1-i} + (-1)^p 2\binom{2p-p-1+1}{p+1-2} A^{p+1-p-1}$$
$$- (-1)^p \binom{2p-p-1}{p+1-3} A^{p+1-p-1}$$

Inserting in the first sum the terms with *i*=1 and *i*=2, in the second sum the terms with *i*=2 and *i* = *p*+1, and in the third sum the term with *i*=*p*+1

$$U_{p+1}(A) = \sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i}{i-1} A^{p+1-i} + \sum_{i=2}^{p+1} (-1)^{i-1} 2\binom{2p-i+1}{i-2} A^{p+1-i} - \sum_{i=3}^{p+1} (-1)^{i-1} \binom{2p-i}{i-3} A^{p+1-i}$$
$$= \sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i}{i-1} A^{p+1-i} - \sum_{i=2}^{p+1} (-1)^{i-2} 2\binom{2p-i+1}{i-1-1} A^{p-i+1} - \sum_{i=3}^{p+1} (-1)^{i-3} \binom{2p-2-i+2}{i-1-2} A^{p-1-i+2}$$

In the second and third sum replacing *i*-1 and *i*-2 respectively by *i*, I get

$$U_{p+1}(A) = \sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i}{i-1} A^{p+1-i} - \sum_{i=1}^{p} (-1)^{i-1} 2\binom{2p-i}{i-1} A^{p-i} - \sum_{i=1}^{p-1} (-1)^{i-1} \binom{2p-2-i}{i-1} A^{p-1-i}$$

$$= (A-2) \sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i}{i-1} A^{p-i} - \sum_{i=1}^{p-1} (-1)^{i-1} \binom{2(p-1)-i}{i-1} A^{p-1-i}$$

$$= (A-2)U_p(A) - U_{p-1}(A)$$

■

The next theorem proves that $U_p(A)$ is symmetrical respect to the line A=2, with odd symmetry for even p and even symmetry for odd p.

Theorem 4:
If p odd then $U_p(2+A) = U_p(2-A)$, if p even then $U_p(2+A) = -U_p(2-A)$

Proof:
U_1 is constantly 1, so $U_1(2+A) = U_1(2-A)$
$U_2(2+A) = (2+A)-2 = A$ and $U_2(2-A) = (2-A)-2 = -A$ so $U_2(2+A) = -U_2(2-A)$
If for every i<p+1 we assume that if i odd then $U_i(2+A) = U_i(2-A)$, and if i even then $i(2+A) = -U_i(2-A)$, then I consider the two cases (1) and (2):
(1) p+1 is even (p odd, p-1 even) then $U_{p+1}(2+A) = (2+A-2)U_p(2+A) - U_{p-1}(2+A)$ and by inductive hypothesis we obtain $(A)U_p(2-A) + U_{p-1}(2-A)$. On the other hand $U_{p+1}(2-A) = (2-A-2)U_p(2-A) - U_{p-1}(2-A)$, which is equivalent to $U_{p+1}(2-A) = -\left((A)U_p(2-A) + U_{p-1}(2-A)\right)$. In this case we obtain that $U_{p+1}(2+A) = -U_{p+1}(2-A)$.
(2) p+1 is odd (p even, p-1 odd) then $U_{p+1}(2+A) = (2+A-2)U_p(2+A) - U_{p-1}(2+A)$ and by inductive hypothesis we obtain $(A)\left(-U_p(2-A)\right) - U_{p-1}(2-A) = -(A)U_p(2-A) - U_{p-1}(2-A)$. On the other hand $U_{p+1}(2-A) = (2-A-2)U_p(2-A) - U_{p-1}(2-A)$, which is equivalent to $U_{p+1}(2-A) = -(A)U_p(2-A) - U_{p-1}(2-A)$; so in this case we obtain that $U_{p+1}(2+A) = U_{p+1}(2-A)$
By complete induction for every p, if p is odd then $U_p(2+A) = U_p(2-A)$, and if p is even then $U_p(2+A) = -U_p(2-A)$.

■

Let's study now $U_p(A) = \sum_{i=1}^{p}(-1)^{i-1}\binom{2p-i}{i-1}A^{p-i}$ when A≤0. In this case, I replace A with –B then B≥0 and I get $U_p(-B) = \sum_{i=1}^{p}(-1)^{i-1+p-i}\binom{2p-i}{i-1}B^{p-i} = (-1)^{p-1}\sum_{i=1}^{p}\binom{2p-i}{i-1}B^{p-i}$ which is always positive when p is odd and always negative when p is even. So I conclude that $U_p(A) \neq 0$ when A≤0, and considering the last theorem also $U_p(A) \neq 0$ when A≥4
As a consequence, all the zeros of $U_p(A)$ lie in the interval 0<A<4. The next theorem is important because it will allows me calculate the zeros in this interval:

Theorem 5:

If $0 < A < 4$ let $A = 2\cos\theta + 2$, $(0 < \theta < \pi)$

Then $U_p(2\cos\theta + 2) = \frac{\sin p\theta}{\sin\theta}$

Proof:

Recall that $U_{p+1}(A) = (A-2)U_p(A) - U_{p-1}(A)$, $U_1(A) = 1$, $U_2(A) = (A-2)$, $p \geq 2$

I'm also going to use the trigonometric angle addition and subtraction formula

$$\sin(\alpha+\beta) = \sin\alpha\cos\beta + \sin\beta\cos\alpha$$

$$\sin(\alpha-\beta) = \sin\alpha\cos\beta - \sin\beta\cos\alpha$$

From these formulas I can derive $\sin(\alpha+\beta) = 2\sin\alpha\cos\beta - \sin(\alpha-\beta)$ (#)

For p=1, $U_1(2\cos\theta+2) = 1$ by definition, in the other hand $\frac{\sin 1\theta}{\sin\theta} = 1$

For p=2, $U_2(2\cos\theta+2) = (A-2)$ by definition, $(A-2) = (2\cos\theta + 2 - 2) = 2\cos\theta$

in the other hand $\frac{\sin 2\theta}{\sin\theta} = \frac{2\cos\theta\sin\theta}{\sin\theta} = 2\cos\theta$

By inductive hypothesis I assume that

$U_p(2\cos\theta+2) = \frac{\sin p\theta}{\sin\theta}$ and also that $U_{p-1}(2\cos\theta+2) = \frac{\sin((p-1)\theta)}{\sin\theta}$

So for p+1

$$U_{p+1}(A) = (A-2)U_p(A) - U_{p-1}(A)$$

$$U_{p+1}(2\cos\theta+2) = (2\cos\theta+2-2)U_p(2\cos\theta+2) - U_{p-1}(2\cos\theta+2)$$

Using the inductive hypothesis

$$U_{p+1}(2\cos\theta+2) = (2\cos\theta)\frac{\sin p\theta}{\sin\theta} - \frac{\sin\big((p-1)\theta\big)}{\sin\theta} = \frac{2\cos\theta\sin p\theta - \sin\big((p-1)\theta\big)}{\sin\theta}$$

In the other side using (#) with $\alpha = p\theta$ and $\beta = \theta$

$$\frac{\sin\big((p+1)\theta\big)}{\sin\theta} = \frac{\sin(p\theta+\theta)}{\sin\theta} = \frac{2\cos\theta\sin p\theta - \sin(p\theta-\theta)}{\sin\theta} = \frac{2\cos\theta\sin p\theta - \sin\big((p-1)\theta\big)}{\sin\theta}$$

Then

$U_{p+1}(2\cos\theta+2) = \frac{\sin\big((p+1)\theta\big)}{\sin\theta}$ and by complete induction I conclude that

$U_p(2\cos\theta+2) = \frac{\sin(p\theta)}{\sin\theta}$ for every natural p.

∎

The zeros of $U_p(A) = \frac{\sin p\theta}{\sin\theta}$, $0 < A < 4$ are $\theta_k = \frac{k\pi}{p}$ $k = 1,2,\dots,p-1$

These are as many as the polynomial degree; so the zeros of $U_p(A)$ are $A_k^{(p)} = 2\cos\frac{k\pi}{p} + 2$

$k = 1,2,\dots,p-1$

For each one of these values $A_i^{(p)}$ I have that $\left({A_i^{(p)}}^{-1}z^2 - w\right)$ is a factor of $S_1^{(p)}(z,w)$. This means that the zero-level curves of the surface $S_1^{(p)}(z,w)$ are nothing more than p-1 parabolas with vertex at (0,0). It also tells me how to calculate these parabolas. In fact I can now rewrite $S_1^{(p)}(z,w) = p\prod_{i=1}^{p-1}\left(\left(A_i^{(p)}\right)^{-1}z^2 - w\right)$.

The same kind of treatment can be done for $S_0^{(p)}(z,w) = \sum_{i=1}^{p}(-1)^{i-1}\binom{2p-i-1}{i-1}z^{2(p-i)}w^{i-1}$

So I write the long division between $S_0^{(p)}(z,w)$ and $A^{-1}z^2 - w$, with an eventual remainder. In this case the expression is monic in w so I don't have to include any coefficient in the final factorization.

Theorem 2-b

$$S_0^{(p)}(z,w) = \sum_{i=1}^{p}(-1)^{i-1}\binom{2p-i-1}{i-1}z^{2(p-i)}w^{i-1}$$

$$= (A^{-1}z^2 - w) \times \left(\sum_{l=1}^{p-1} z^{2(p-l)-2}w^{l-1}\sum_{i=1}^{l}(-1)^{i-1}\binom{2p-i-1}{i-1}A^{l-i+1}\right)$$

$$+ w^{p-1}\sum_{i=1}^{p}(-1)^{i-1}\binom{2p-i-1}{i-1}A^{p-i}$$

Proof analogous to Theorem 2-a

The remainder is 0 when w=0, or when $\sum_{i=1}^{p}(-1)^{i-1}\binom{2p-i-1}{i-1}A^{p-i} = 0$. I'll call it:

Definition 3-b:

$$V_p(A) = \sum_{i=1}^{p}(-1)^{i-1}\binom{2p-i-1}{i-1}A^{p-i}$$

There is a relationship between $V_p(A)$ and $U_p(A)$; it is the following:

Theorem 6:

$$V_p(A) = U_p(A) + U_{p-1}(A)$$

Proof in appendix

From this relationship, using Theorem 3 is easy to see that $V_p(A)$ satisfies the recursive definition

$V_{p+1}(A) = (A-2)V_p(A) - V_{p-1}(A), \qquad V_1(A) = 1, \qquad V_2(A) = (A-1), \qquad p \geq 2$

Theorem 7:

If $0 < A < 4$ let $A = 2\cos\theta + 2, \qquad (0 < \theta < \pi)$

Then $V_p(2\cos\theta + 2) = -\frac{\sin p\theta + \sin(p-1)\theta}{\sin\theta} = -\frac{\sqrt{A}\sin\left(\frac{2p-1}{2}\theta\right)}{\sin\theta}$

Proof:

Because $-V_p(A) = U_p(A) + U_{p-1}(A)$ then

$-V_p(2\cos\theta + 2) = U_p(2\cos\theta + 2) + U_{p-1}(2\cos\theta + 2)$ and since

$U_p(2\cos\theta + 2) = \frac{\sin p\theta}{\sin\theta}$ and $U_{p-1}(2\cos\theta + 2) = \frac{\sin(p-1)\theta}{\sin\theta}$ we get

$-V_p(2\cos\theta + 2) = \frac{\sin p\theta}{\sin\theta} + \frac{\sin(p-1)\theta}{\sin\theta}$, therefore

$$V_p(2\cos\theta + 2) = -\frac{\sin p\theta + \sin(p-1)\theta}{\sin\theta}$$

For the second part I use that

$\sin\alpha + \sin\beta = 2\sin\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2}$, then making $\alpha = p\theta$ and $\beta = (p-1)\theta$

$$\sin p\theta + \sin(p-1)\theta = 2\sin\frac{p\theta + (p-1)\theta}{2}\cos\frac{p\theta - (p-1)\theta}{2}$$

$$\sin p\theta + \sin(p-1)\theta = 2\sin\frac{2p\theta - \theta}{2}\cos\frac{\theta}{2}$$

Now I use that $\cos\frac{\theta}{2} = \sqrt{\frac{\cos\theta + 1}{2}}$ to get

$$\sin p\theta + \sin(p-1)\theta = 2\sin(\frac{(2p-1)}{2}\theta)\sqrt{\frac{\cos\theta + 1}{2}}$$

$$\sin p\theta + \sin(p-1)\theta = 2\sin\frac{(2p-1)}{2}\theta\frac{\sqrt{A}}{2}.$$ So

$$-\frac{\sin p\theta + \sin(p-1)\theta}{\sin\theta} = -\frac{\sqrt{A}\sin\frac{(2p-1)}{2}\theta}{\sin\theta}$$

■

The zeros of $V_p(A) = \frac{\sqrt{A}\sin\left(\frac{2p-1}{2}\theta\right)}{\sin\theta}$, where $0 < A < 4$ are $\theta_k = \frac{2k\pi}{2p-1}$ for $k = 1,2,\dots,p-1$

These are as many as the polynomial degree; so the zeros of $V_p(A)$ are $B_k^{(p)} = 2\cos\frac{2k\pi}{2p-1} + 2$

$k = 1,2,\dots,p-1$

Now I can also rewrite $S_0^{(p)}(z,w) = \prod_{i=1}^{p-1}\left(\left(B_i^{(p)}\right)^{-1} z^2 - w\right)$

By comparing the zeros of $U_p(A)$ and $V_p(A)$ that are also the curvature of the zero-level parabolas of $S_1^{(p)}(z,w) = 0$ and $S_0^{(p)}(z,w) = 0$ respectively, I can see how the parabolas from the graph of $S_1^{(p)}(z,w) = 0$ and $S_0^{(p)}(z,w) = 0$ interweave. Thus I need to prove that

$$0 < A_{p-1}^{(p)} < B_{p-1}^{(p)} < A_{p-2}^{(p)} < B_{p-2}^{(p)} < \cdots < A_2^{(p)} < B_2^{(p)} < A_1^{(p)} < B_1^{(p)}$$

In order to make notation shorter let's call $A_p^{(p)} = 0$

Theorem 8:

$A_{k+1}^{(p)} < B_k^{(p)} < A_k^{(p)}$ with 0<k<p

Proof:

I know that 0<k<p, therefore

$2p - 1 < 2p$, then $(2p-1)k\pi < 2pk\pi$, and $\frac{k\pi}{p} < \frac{2k\pi}{(2p-1)}$. As cos (x) is a decreasing function between 0 and π then: $\cos\frac{2k\pi}{2p-1} < \cos\frac{k\pi}{p}$, therefore $2\cos\frac{2k\pi}{2p-1} + 2 < 2\cos\frac{k\pi}{p} + 2$. That is to say $B_k^{(p)} < A_k^{(p)}$

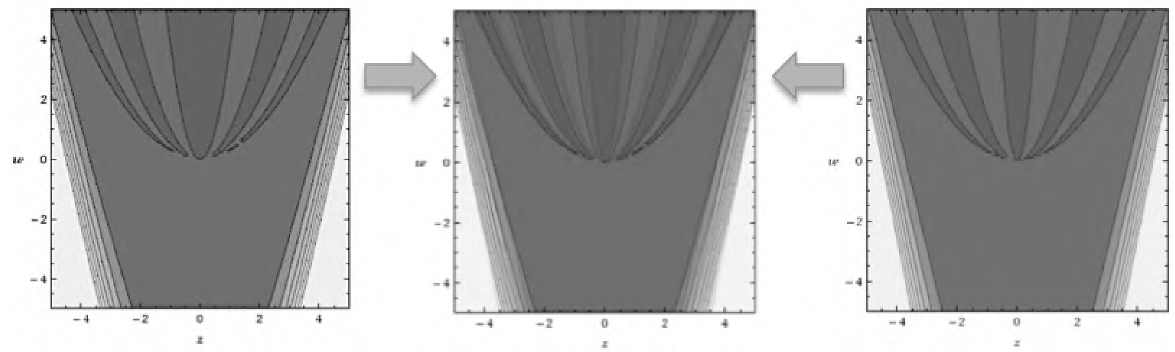
For the first inequality I notice that $k \le p - 1 < 2p - 1$. So $0 < 2p - 1 - k$.

Then $2kp\pi < 2kp\pi + \pi(2p - 1 - k)$. So $2kp\pi < 2kp\pi + 2p\pi - \pi - k\pi$ By factorizing the first two terms and the second two, $2kp\pi < 2p\pi(k+1) - \pi(k+1)$ and factorizing again

$2kp\pi < (k+1)(2p\pi - \pi)$ or $2kp\pi < (k+1)\pi(2p-1)$. Therefore $\frac{2k\pi}{2p-1} < \frac{(k+1)\pi}{p}$. As cos (x) is a decreasing function between 0 and π then $\cos\frac{(k+1)\pi}{p} < \cos\frac{2k\pi}{2p-1}$ and $2\cos\frac{(k+1)\pi}{p} + 2 < 2\cos\frac{2k\pi}{2p-1} + 2$.

That is to say $A_{k+1}^{(p)} < B_k^{(p)}$

■



S_1 — Overlap of S_1 and S_0 showing the structure of interwaved parabolas — S_0

By replacing the new obtained expressions in the remainder I can appreciate its structure, both surfaces are a linear combination of other surfaces that are a set of nested parabolas.

$$R_1^{(n)}(z,w) = \left(\sum_{p=1}^{n/2} a_{2p-1} \prod_{i=1}^{p-1}\left(\left(B_i^{(p)}\right)^{-1} z^2 - w\right)\right) - \left(\sum_{p=1}^{n/2} a_{2p} pz \prod_{i=1}^{p-1}\left(\left(A_i^{(p)}\right)^{-1} z^2 - w\right)\right)$$

$$R_0^{(n)}(z,w) = a_0 + w\left(\left(\sum_{p=1}^{n/2-1} a_{2p+1}\, pz \prod_{i=1}^{p-1}\left(\left(A_i^{(p)}\right)^{-1} z^2 - w\right)\right) - \left(\sum_{p=1}^{n/2} a_{2p} \prod_{i=1}^{p-1}\left(\left(B_i^{(p)}\right)^{-1} z^2 - w\right)\right)\right)$$

The idea for the rest for the proof is that when going far away enough in these surfaces they will behave in a similar way to the set of nested parabolas of the highest degree forming part of each of them.

If $z \neq 0$, by changing to parabolic coordinates using $w = Az^2$

$S_1^{(p)}(z, \mathrm{A}z^2) = pz^{2p-2} \prod_{i=1}^{p-1}\left(\left(A_i^{(p)}\right)^{-1} - A\right)$ and $S_0^{(p)}(z, \mathrm{A}z^2) = z^{2p-2} \prod_{i=1}^{p-1}\left(\left(B_i^{(p)}\right)^{-1} - A\right)$

Therefore if $z \neq 0$ then I can rewrite (remembering that $a_n = 1$)

$$R_1^{(n)}(z, \mathrm{A}z^2) = \left(\sum_{p=1}^{n/2} a_{2p-1}\, z^{2p-2} \prod_{i=1}^{p-1}\left(\left(B_i^{(p)}\right)^{-1} - A\right)\right) - \left(\sum_{p=1}^{n/2} a_{2p} p z^{2p-1} \prod_{i=1}^{p-1}\left(\left(A_i^{(p)}\right)^{-1} - A\right)\right)$$

$$R_0^{(n)}(z, \mathrm{A}z^2) = a_0 + A\left(\left(\sum_{p=1}^{n/2-1} a_{2p+1}\, p z^{2p+1} \prod_{i=1}^{p-1}\left(\left(A_i^{(p)}\right)^{-1} - A\right)\right) - \left(\sum_{p=1}^{n/2} a_{2p}\, z^{2p} \prod_{i=1}^{p-1}\left(\left(B_i^{(p)}\right)^{-1} - A\right)\right)\right)$$

When $|z|$ is sufficiently large I can study these surfaces by considering only the terms of the highest degree in z.

$$R_1^{(n)}(z,\mathrm{A}z^2) \cong -\tfrac{n}{2}z^{n-1}\prod_{i=1}^{\frac{n}{2}-1}\left(\left(A_i^{\left(\frac{n}{2}\right)}\right)^{-1}-A\right) = -zS_1^{\left(\frac{n}{2}\right)}(z,\mathrm{A}z^2).$$

$$R_0^{(n)}(z,\mathrm{A}z^2) \cong -Az^n\prod_{i=1}^{\frac{n}{2}-1}\left(\left(B_i^{\left(\frac{n}{2}\right)}\right)^{-1}-A\right) = -Az^2S_0^{\left(\frac{n}{2}\right)}(z,\mathrm{A}z^2).$$

More precisely this means that $\lim_{z\to\infty}\frac{R_1^{(n)}(z,\mathrm{A}z^2)}{-zS_1^{\left(\frac{n}{2}\right)}(z,\mathrm{A}z^2)} = 1$ and $\lim_{z\to\infty}\frac{R_0^{(n)}(z,\mathrm{A}z^2)}{-\mathrm{A}z^2S_0^{\left(\frac{n}{2}\right)}(z,\mathrm{A}z^2)} = 1$

I'll show now that $R_1^{(n)}$ and $R_0^{(n)}$ zero-level curves interweave in the same way that $S_1^{\left(\frac{n}{2}\right)}$ and $S_0^{\left(\frac{n}{2}\right)}$ for large values of z. To do this let's now study the sign of $-zS_1^{\left(\frac{n}{2}\right)}(z,\mathrm{A}z^2)$ and $-\mathrm{A}z^2S_0^{\left(\frac{n}{2}\right)}(z,\mathrm{A}z^2)$

To simplify the notation and reduce the number of cases, I'll adopt

$$\left(A_0^{\left(\frac{n}{2}\right)}\right)^{-1} = -\infty\,,\left(A_{\frac{n}{2}}^{\left(\frac{n}{2}\right)}\right)^{-1} = \infty,\ \left(B_0^{\left(\frac{n}{2}\right)}\right)^{-1} = 0\,,\left(B_{\frac{n}{2}}^{\left(\frac{n}{2}\right)}\right)^{-1} = \infty\,, 0 \le i \le \tfrac{n}{2}$$

For $-zS_1^{\left(\frac{n}{2}\right)}(z,\mathrm{A}z^2) = -\frac{n}{2}z^{n-1}\prod_{i=1}^{\frac{n}{2}-1}\left(\left(A_i^{\left(\frac{n}{2}\right)}\right)^{-1}-A\right)$. There are four cases:

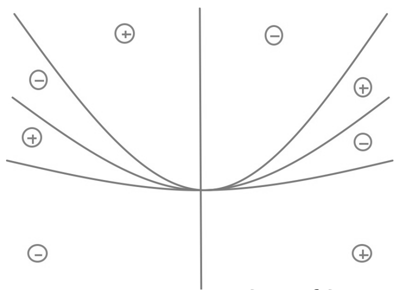
1) $\left(A_i^{\left(\frac{n}{2}\right)}\right)^{-1} < A < \left(A_{i+1}^{\left(\frac{n}{2}\right)}\right)^{-1}$, i is odd and z>0, then $-zS_1^{\left(\frac{n}{2}\right)}(z,\mathrm{A}z^2) > 0$
2) $\left(A_i^{\left(\frac{n}{2}\right)}\right)^{-1} < A < \left(A_{i+1}^{\left(\frac{n}{2}\right)}\right)^{-1}$, i is odd and z<0, then $-zS_1^{\left(\frac{n}{2}\right)}(z,\mathrm{A}z^2) < 0$
3) $\left(A_i^{\left(\frac{n}{2}\right)}\right)^{-1} < A < \left(A_{i+1}^{\left(\frac{n}{2}\right)}\right)^{-1}$, i is even, and z>0, then $-zS_1^{\left(\frac{n}{2}\right)}(z,\mathrm{A}z^2) < 0$
4) $\left(A_i^{\left(\frac{n}{2}\right)}\right)^{-1} < A < \left(A_{i+1}^{\left(\frac{n}{2}\right)}\right)^{-1}$, i is even, and z<0, then $-zS_1^{\left(\frac{n}{2}\right)}(z,\mathrm{A}z^2) > 0$

For $-Az^2S_0^{\left(\frac{n}{2}\right)}(z,\mathrm{A}z^2)$=$-Az^n\prod_{i=1}^{\frac{n}{2}-1}\left(\left(B_i^{\left(\frac{n}{2}\right)}\right)^{-1}-A\right)$ There are three cases:

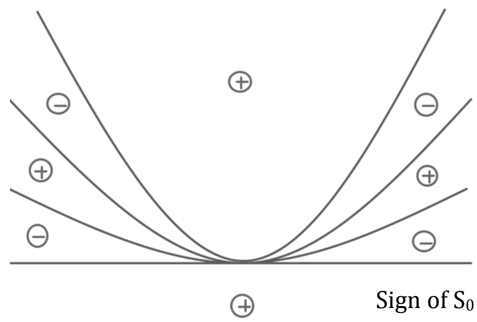
1) $-\infty < A < 0$, then $-Az^2 S_0^{\left(\frac{n}{2}\right)}(z, Az^2) > 0$

2) $\left(B_i^{\left(\frac{n}{2}\right)}\right)^{-1} < A < \left(B_{i+1}^{\left(\frac{n}{2}\right)}\right)^{-1}$, i is even, then $-Az^2 S_0^{\left(\frac{n}{2}\right)}(z, Az^2) < 0$

3) $\left(B_i^{\left(\frac{n}{2}\right)}\right)^{-1} < A < \left(B_{i+1}^{\left(\frac{n}{2}\right)}\right)^{-1}$, i is odd, then $-Az^2 S_0^{\left(\frac{n}{2}\right)}(z, Az^2) > 0$



Sign of S_1



Sign of S_0

From Theorem 8 I know that $A_{i+1}^{(p)} < B_i^{(p)} < A_i^{(p)}$ with $0 < i < p$ and $A_p^{(p)} = 0$

so $A_i^{\left(\frac{n}{2}\right)^{-1}} < B_i^{\left(\frac{n}{2}\right)^{-1}} < A_{i+1}^{\left(\frac{n}{2}\right)^{-1}}$ with $0 < i < \frac{n}{2}$ and $\left(A_{\frac{n}{2}}^{\left(\frac{n}{2}\right)}\right)^{-1} = \infty$. I will now choose points

$C_0, C_1, \ldots, C_{n-1}$ between them as follows: $C_0 < 0 < C_1 < A_1^{\left(\frac{n}{2}\right)^{-1}} < C_2 < B_1^{\left(\frac{n}{2}\right)^{-1}} < C_3 <$
$A_2^{\left(\frac{n}{2}\right)^{-1}} < \cdots < A_i^{\left(\frac{n}{2}\right)^{-1}} < C_{2i} < B_i^{\left(\frac{n}{2}\right)^{-1}} < C_{2i+1} < A_{i+1}^{\left(\frac{n}{2}\right)^{-1}} < \cdots < B_{\frac{n}{2}-1}^{\left(\frac{n}{2}\right)^{-1}} < C_{n-1}$

Now I can calculate the limits for these particular values of A

$\lim_{z\to\infty} R_1^{(n)}(z, C_i z^2) = -\infty$ if $i = 4j$ or $i = 4j+1$, and $\lim_{z\to\infty} R_1^{(n)}(z, C_i z^2) = \infty$ if $i = 4j+2$ or $i = 4j+3$

$\lim_{z\to\infty} R_0^{(n)}(z, C_i z^2) = -\infty$ if $i = 4j+1$ or $i = 4j+2$, and $\lim_{z\to\infty} R_0^{(n)}(z, C_i z^2) = \infty$ if $i = 4j$ or $i = 4j+3$

This means that there is an N_{Ci} such that for every z greater than N_{Ci}, $R_1^{(n)}(z, C_i z^2) < -1$ for $i = 4j$ or $i = 4j+1$ and $R_1^{(n)}(z, C_i z^2) > 1$ for $i = 4j+2$ or $i = 4j+3$

There is an M_{Ci} such that for every z greater than M_{Ci}, $R_0^{(n)}(z, C_i z^2) < -1$ for $i = 4j+1$ or $i = 4j+2$ and $R_0^{(n)}(z, C_i z^2) > 1$ for $i = 4j$ or $i = 4j+3$

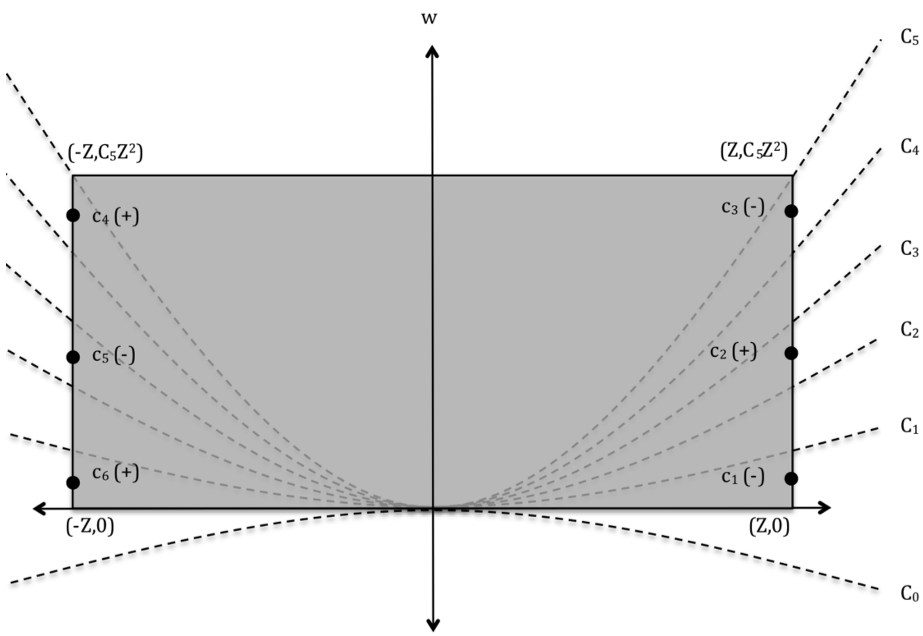
Let now Z be such that $Z > \max(N_{C0}, N_{C1}, \dots, N_{Cn-1}, M_{C0}, M_{C1}, \dots, M_{Cn-1})$. For $0 \le i \le n-1$
If $i = 4j$ then $R_1^{(n)}(Z, C_iZ^2) < 0$ and $R_0^{(n)}(Z, C_iZ^2) > 0$
If $i = 4j + 1$ then $R_1^{(n)}(Z, C_iZ^2) < 0$ and $R_0^{(n)}(Z, C_iZ^2) < 0$
If $i = 4j + 2$ then $R_1^{(n)}(Z, C_iZ^2) > 0$ and $R_0^{(n)}(Z, C_iZ^2) < 0$
If $i = 4j + 3$ then $R_1^{(n)}(Z, C_iZ^2) > 0$ and $R_0^{(n)}(Z, C_iZ^2) > 0$

The vertical segment $\overline{(Z, C_0Z^2)(Z, C_{n-1}Z^2)}$ is divided by the points (Z, C_iZ^2) into n-1 parts. In each of these parts either $R_0^{(n)}(Z, \mathrm{A}Z^2)$ changes sign between the extremes of the segment while $R_1^{(n)}(Z, \mathrm{A}Z^2)$ doesn't change sign between the extremes, or $R_1^{(n)}(Z, \mathrm{A}Z^2)$ changes sign between the extremes of the segment while $R_0^{(n)}(Z, \mathrm{A}Z^2)$ doesn't change sign between the extremes. As n-1 is odd, there are $\frac{n}{2}$ of these segments where $R_0^{(n)}(Z, \mathrm{A}Z^2)$ changes sign and $\frac{n}{2} - 1$ where $R_1^{(n)}(Z, \mathrm{A}Z^2)$ changes sign.
Because of the Intermediate Value Theorem there is at least one value of A for each segment for which $R_0^{(n)}(Z, \mathrm{A}Z^2) = 0$ or $R_1^{(n)}(Z, \mathrm{A}Z^2) = 0$ alternatively. But the degree in A of $R_0^{(n)}(Z, \mathrm{A}Z^2) = 0$ is $\frac{n}{2}$, and the degree in A of $R_1^{(n)}(Z, \mathrm{A}Z^2) = 0$ is $\frac{n}{2} - 1$. So there is exactly one value of A in each segment for which $R_0^{(n)}(Z, \mathrm{A}Z^2) = 0$ or $R_1^{(n)}(Z, \mathrm{A}Z^2) = 0$ alternatively.
By an analogous reasoning we have the same situation for $R_0^{(n)}(-Z, \mathrm{A}Z^2) = 0$ and $R_1^{(n)}(-Z, \mathrm{A}Z^2) = 0$.
$R_0^{(n)}(z, \mathrm{A}z^2) = 0$ and $R_1^{(n)}(z, \mathrm{A}z^2) = 0$, are the 0-level curves of $R_0^{(n)}$and $R_1^{(n)}$, and since they are continuous functions, for every value of z greater than Z or less than –Z its graphs are between the parabolas C_iZ^2 and therefore interweaved as I wanted to show.

Let's now consider the line passing through the points $(Z, C_{n-1}Z^2)$ and $(-Z, C_{n-1}Z^2)$.
$R_0^{(n)}(z, \mathrm{A}z^2) = 0$ intersects the line n times, $\frac{n}{2}$ to the right of $(Z, C_{n-1}Z^2)$ and $\frac{n}{2}$ to the left of $(-Z, C_{n-1}Z^2)$, these are as many as its degree in z, so there are no further intersections between the points $(Z, C_{n-1}Z^2)$ and $(-Z, C_{n-1}Z^2)$.
Notice also that $R_0^{(n)}(z, 0) = a_0$ so the graph of $R_0^{(n)}(z, \mathrm{w}) = 0$ doesn't cross the horizontal axis.
Finally, let's consider now the rectangle with vertex at $(Z, C_{n-1}Z^2)$, $(-Z, C_{n-1}Z^2)$, $(Z, 0)$ and $(-Z, 0)$.

$R_0^{(n)}(z, \mathrm{A}z^2) = 0$ intersects the rectangle in *n* points if $a_0 > 0$ or in *n*-2 points if $a_0 < 0$. Half of them are on the right vertical side and half of them in the left one. Also in these points the sign of $R_1^{(n)}(z, \mathrm{A}z^2)$ alternates from one vertical point to the next one and also from a point in one line to the corresponding point in the other. Let's call these points $c_1, c_2, \dots, c_n$ ordered by the angle they for with the horizontal axis. $R_1^{(n)}(z, \mathrm{A}z^2)$ is positive in c_i if *i* is even and negative if *i* is odd.

The perimeter of the rectangle is divided into *n* parts by the points $c_1, c_2, \dots, c_n$, between which the sign of $R_0^{(n)}$ alternates. As we expand the rectangle, the segments $\overline{c_1c_2}$, $\overline{c_2c_3}$, ... etc. sweep out *n* regions extending to infinity within which $R_0^{(n)}$ has alternately positive and negative values and these regions are separated by curves on which $R_0^{(n)} = 0$.



Borrowing the terminology from Gauss treatment in a similar situation arisen in his proof of the fundamental theorem of algebra, call the regions outside of the rectangle where $R_0^{(n)} > 0$ **lands,** and the regions where $R_0^{(n)} < 0$ **seas**. The curves where $R_0^{(n)}$=0 are then **seashores**. The *n* seas and *n* lands extend themselves inside the rectangle across the *n* parts of its perimeter. If a I enter the rectangle in c_1 moving always along a seashore I will eventually exit the rectangle through another point c_i. If the shoreline passing through c_2 intersects the previous one, then there is also a shore line inside the rectangle connecting c_1 and c_2 (starting from c_1, moving along the first line to the intersection and the along the second one to c_2) On the other hand if it doesn't intersect it, then it must exit the rectangle in one of the points before c_i. Either way, repeating the process we arrive to a shoreline connecting one of the points with the next one.

So let´s assume that y is a continuous path inside the rectangle joining c_i and c_{i+1} such that $R_0^{(n)} = 0$ along y (a seashore). $R_1^{(n)}(c_i)$ and $R_1^{(n)}(c_{i+1})$ have different sign. But $R_1^{(n)}$ is a continuous function along the path y, so by the Intermediate Value Theorem at some point on y it takes the value 0. At this point $R_1^{(n)}$ and $R_0^{(n)}$ are both zero and is the point I was searching for.

6.1. Corollary. $\prod_{i=1}^{p-1} A_i^{(p)} = p$

Proof:

Instead of writing the long division to find a factored expression of S_1 I could have just made the following transformations to the earlier expression

$$S_1^{(p)}(z, Az^2) = \sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i}{i-1} z^{2(p-i)} (Az^2)^{i-1} = z^{2p-2} \sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i}{i-1} A^{i-1}$$

Dividing and multiplying by A^{p-1}.

$$S_1^{(p)}(z, Az^2) = A^{p-1} z^{2p-2} \sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i}{i-1} \left(\frac{1}{A}\right)^{p-i}$$

And this is

$$S_1^{(p)}(z, Az^2) = A^{p-1} z^{2p-2} U_p\left(\frac{1}{A}\right)$$

All the zeros of $U_p(A)$ are $A_k^{(p)} = 2\cos\frac{k\pi}{p} + 2$ with $k = 1, 2, \dots, p-1$, so factoring U_p.

$$S_1^{(p)}(z, Az^2) = A^{p-1} z^{2p-2} \prod_{i=1}^{p-1} \left(\left(\frac{1}{A}\right) - A_i^{(p)}\right)$$

Returning to w with $A = w/z^2$.

$$S_1^{(p)}(z, w) = w^{p-1} \prod_{i=1}^{p-1} \left(\left(\frac{z^2}{w}\right) - A_i^{(p)}\right) = \prod_{i=1}^{p-1} \left(z^2 - wA_i^{(p)}\right)$$

Dividing and multiplying by each A_i

$$S_1^{(p)}(z, w) = \left(\prod_{i=1}^{p-1} A_i^{(p)}\right) \left(\prod_{i=1}^{p-1} \left(\left(A_i^{(p)}\right)^{-1} z^2 - w\right)\right)$$

But I had already obtained by the long division method that

$$S_1^{(p)}(z, w) = p \prod_{i=1}^{p-1} \left(\left(A_i^{(p)}\right)^{-1} z^2 - w\right)$$

So it must be that

$$\prod_{i=1}^{p-1} A_i^{(p)} = p$$

As I wanted to show.

Remark:

$\prod_{i=1}^{p-1} A_i^{(p)} = p$ is equivalent to $2^{p-1}\prod_{i=1}^{p-1} sin\frac{i\pi}{p} = p$ (were *i* is the index of the product operator and not the imaginary unit) but the proof of this fact uses the complex numbers, even if it is a statement completely among real numbers. This is the reason for the approach I choose in the main proof to find a factored expression of S_1. Also the previous proof is a completely real proof of this trigonometric fact.

6.2. Appendix 1: Proofs of theorems 1, 2 and 6.

Proof of Theorem 1:

This is the proof that $k_i^{(n)}(z,w) = h_i^{(n)}(z,w)$ for every natural number i

For $i = 1$ and $i = 2$ the result is achieved by simple substitution, but in the first case I have to consider that the sum $\sum_{p=1}^{0} \dots = 0$

Assuming $k_i^{(n)} = h_i^{(n)}$ and $k_{i+1}^{(n)} = h_{i+1}^{(n)}$

$$k_{i+2}^{(n)} = a_{n-(i+1)} - w\left(\sum_{p=1}^{\left[\frac{i+1}{2}\right]} a_{n-i+2p-1} \sum_{j=1}^{p} (-1)^{j-1} \binom{2p-j-1}{j-1} z^{2(p-j)} w^{j-1}\right)$$
$$+ \left(\sum_{p=1}^{\left[\frac{i}{2}\right]} a_{n-i+2p}\, z \sum_{j=1}^{p} (-1)^{j-1} \binom{2p-j}{j-1} z^{2(p-j)} w^{j-1}\right)$$
$$- z\left(\sum_{p=1}^{\left[\frac{i+2}{2}\right]} a_{n-i+2p-2} \sum_{j=1}^{p} (-1)^{j-1} \binom{2p-j-1}{j-1} z^{2(p-j)} w^{j-1}\right)$$
$$+ \left(\sum_{p=1}^{\left[\frac{i+1}{2}\right]} a_{n-i+2p-1}\, z \sum_{j=1}^{p} (-1)^{j-1} \binom{2p-j}{j-1} z^{2(p-j)} w^{j-1}\right)$$

Distributing w and z.

$$k_{i+2}^{(n)} = a_{n-(i+1)} - \left(\sum_{p=1}^{\left[\frac{i+1}{2}\right]} a_{n-i+2p-1} \sum_{j=1}^{p} (-1)^{j-1} \binom{2p-j-1}{j-1} z^{2(p-j)} w^{j}\right)$$
$$+ \left(\sum_{p=1}^{\left[\frac{i}{2}\right]} a_{n-i+2p}\, z \sum_{j=1}^{p} (-1)^{j-1} \binom{2p-j}{j-1} z^{2(p-j)} w^{j}\right)$$
$$- \left(\sum_{p=1}^{\left[\frac{i+2}{2}\right]} a_{n-i+2p-2} \sum_{j=1}^{p} (-1)^{j-1} \binom{2p-j-1}{j-1} z^{2(p-j)+1} w^{j-1}\right)$$
$$+ \left(\sum_{p=1}^{\left[\frac{i+1}{2}\right]} a_{n-i+2p-1}\, z \sum_{j=1}^{p} (-1)^{j-1} \binom{2p-j}{j-1} z^{2(p-j)+1} w^{j-1}\right)$$

Separating the term with p=1 from the first sum, changing j for j-1 in the first sum and factoring -1 from the first sum. Also changing p for p-1 and j for j-1 in the second sum and factoring -1 from the second sum.

$$k_{i+2}^{(n)} = a_{n-(i+1)} - a_{n-i+1}w + \left(\sum_{p=2}^{[^{i+1}/_2]} a_{n-i+2p-1} \sum_{j=2}^{p+1} (-1)^{j-1} \binom{2p-j}{j-2} z^{2(p-j)+2} w^{j-1} \right)$$
$$- \left(\sum_{p=2}^{[^{i}/_2]+1} a_{n-i+2p-2}\, z \sum_{j=2}^{p} (-1)^{j-1} \binom{2p-j-1}{j-2} z^{2(p-j)} w^{j-1} \right)$$
$$- \left(\sum_{p=1}^{[^{i+2}/_2]} a_{n-i+2p-2} \sum_{j=1}^{p} (-1)^{j-1} \binom{2p-j-1}{j-1} z^{2(p-j)+1} w^{j-1} \right)$$
$$+ \left(\sum_{p=1}^{[^{i+1}/_2]} a_{n-i+2p-1}\, z \sum_{j=1}^{p} (-1)^{j-1} \binom{2p-j}{j-1} z^{2(p-j)+1} w^{j-1} \right)$$

Bringing the fourth sum to the second place. Then separating the terms with j=p+1 from the first sum, separating the term with p=1 from the second sum and also separating the terms with j=1 from the third and fourth sum.

$$k_{i+2}^{(n)} = a_{n-(i+1)} - a_{n-i+1}w + \sum_{p=2}^{[^{i+1}/_2]} a_{n-i+2p-1} (-1)^p w^p$$
$$+ \left(\sum_{p=2}^{[^{i+1}/_2]} a_{n-i+2p-1} \sum_{j=2}^{p} (-1)^{j-1} \binom{2p-j}{j-2} z^{2(p-j)+2} w^{j-1} \right)$$
$$+ \left(\sum_{p=2}^{[^{i+1}/_2]} a_{n-i+2p-1} \sum_{j=2}^{p} (-1)^{j-1} \binom{2p-j}{j-1} z^{2(p-j)+2} w^{j-1} \right) + a_{n-i+1} z^2 + \sum_{p=2}^{[^{i+1}/_2]} a_{n-i+2p-1}\, z^{2p}$$
$$- \left(\sum_{p=2}^{[^{i}/_2]+1} a_{n-i+2p-2} \sum_{j=2}^{p} (-1)^{j-1} \binom{2p-j-1}{j-2} z^{2(p-j)+1} w^{j-1} \right)$$
$$- \left(\sum_{p=2}^{[^{i+2}/_2]} a_{n-i+2p-2} \sum_{j=2}^{p} (-1)^{j-1} \binom{2p-j-1}{j-1} z^{2(p-j)+1} w^{j-1} \right) - \sum_{p=1}^{[^{i+2}/_2]} a_{n-i+2p-2}\, z^{2p-1}$$

Reordering terms and merging together the first and second sum and the third and forth sum by taking common factors.

$$k_{i+2}^{(n)} = a_{n-(i+1)} + a_{n-i+1}z^2 - a_{n-i+1}w + \sum_{p=2}^{[i+1/2]} a_{n-i+2p-1}\,(-1)^p w^p$$
$$+\left(\sum_{p=2}^{[i+1/2]} a_{n-i+2p-1} \sum_{j=2}^{p}(-1)^{j-1} z^{2(p-j)+2} w^{j-1}\left(\binom{2p-j}{j-2}+\binom{2p-j}{j-1}\right)\right) + \sum_{p=2}^{[i+1/2]} a_{n-i+2p-1}\,z^{2p}$$
$$-\left(\sum_{p=2}^{[i/2]+1} a_{n-i+2p-2} \sum_{j=2}^{p}(-1)^{j-1} z^{2(p-j)+1} w^{j-1}\left(\binom{2p-j-1}{j-2}+\binom{2p-j-1}{j-1}\right)\right)$$
$$-\sum_{p=1}^{[i+2/2]} a_{n-i+2p-2}\,z^{2p-1}$$

Using $\binom{a}{b}+\binom{a}{b+1}=\binom{a+1}{b+1}$

$$k_{i+2}^{(n)} = a_{n-(i+1)} + a_{n-i+1}z^2 - a_{n-i+1}w + \sum_{p=2}^{[i+1/2]} a_{n-i+2p-1}\,(-1)^p w^p$$
$$+\left(\sum_{p=2}^{[i+1/2]} a_{n-i+2p-1} \sum_{j=2}^{p}(-1)^{j-1}\binom{2p-j+1}{j-1} z^{2(p-j)+2} w^{j-1}\right) + \sum_{p=2}^{[i+1/2]} a_{n-i+2p-1}\,z^{2p}$$
$$-\left(\sum_{p=2}^{[i/2]+1} a_{n-i+2p-2} \sum_{j=2}^{p}(-1)^{j-1}\binom{2p-j}{j-1} z^{2(p-j)+1} w^{j-1}\right) - \sum_{p=1}^{[i+2/2]} a_{n-i+2p-2}\,z^{2p-1}$$

Merging together the last two terms, changing *p* for *p*-1 in the first, second and third sums and writing the first and third sum in the format of the second one (*j*=*p* and *j*=1)

$$k_{i+2}^{(n)} = a_{n-(i+1)} + a_{n-i+1}z^2 - a_{n-i+1}w + \sum_{p=3}^{\left[{}^{i+1}/_{2}\right]+1} a_{n-i+2p-3}\,(-1)^{p-1}\binom{2p-p-1}{p-1} z^{2(p-p)} w^{p-1}$$

$$+\left(\sum_{p=3}^{\left[{}^{i+1}/_{2}\right]+1} a_{n-i+2p-3}\sum_{j=2}^{p-1}(-1)^{j-1}\binom{2p-j-1}{j-1} z^{2(p-j)} w^{j-1}\right)$$

$$+\sum_{p=3}^{\left[{}^{i+1}/_{2}\right]+1} a_{n-i+2p-3}\,(-1)^{1-1}\binom{2p-1-1}{1-1} z^{2(p-1)}\, w^{1-1}$$

$$-\left(\sum_{p=1}^{\left[{}^{i}/_{2}\right]+1} a_{n-i+2p-2}\sum_{j=1}^{p}(-1)^{j-1}\binom{2p-j}{j-1} z^{2(p-j)+1} w^{j-1}\right)$$

$$k_{i+2}^{(n)} = a_{n-(i+1)} + a_{n-i+1}z^2 - a_{n-i+1}w + \left(\sum_{p=3}^{\left[{}^{i+1}/_{2}\right]+1} a_{n-i+2p-3}\sum_{j=1}^{p}(-1)^{j-1}\binom{2p-j-1}{j-1} z^{2(p-j)} w^{j-1}\right)$$

$$+-\left(\sum_{p=1}^{\left[{}^{i}/_{2}\right]+1} a_{n-i+2p-2}\sum_{j=1}^{p}(-1)^{j-1}\binom{2p-j}{j-1} z^{2(p-j)+1} w^{j-1}\right)$$

The first three terms can be inserted into the first sum (p=1 j=1; p=2 j=1 and p=2 j=2)

$$k_{i+2}^{(n)} = \left(\sum_{p=1}^{\left[{}^{i+1}/_{2}\right]+1} a_{n-i+2p-3}\sum_{j=1}^{p}(-1)^{j-1}\binom{2p-j-1}{j-1} z^{2(p-j)} w^{j-1}\right)$$

$$+-\left(\sum_{p=1}^{\left[{}^{i}/_{2}\right]+1} a_{n-i+2p-2}\sum_{j=1}^{p}(-1)^{j-1}\binom{2p-j}{j-1} z^{2(p-j)+1} w^{j-1}\right)$$

Therefore $k_{i+2}^{(n)} = h_{i+2}^{(n)}$ and using complete induction I conclude that $k_i^{(n)}(z,w) = h_i^{(n)}(z,w)$ for every natural number *i*.

Proof of Theorem 2:

This is the proof that

$$S_1^{(p)}(z,w) = \sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i}{i-1} z^{2(p-i)} w^{i-1}$$
$$= (A^{-1}z^2 - w) \times \left(\sum_{q=1}^{p-1} z^{2(p-q)-2} w^{q-1} \sum_{i=1}^{q} (-1)^{i-1} \binom{2p-i}{i-1} A^{q-i+1} \right)$$
$$+ w^{p-1} \sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i}{i-1} A^{p-i}$$

In the right hand side of the equation I'll distribute, considering that $2(p-q)-2 = 2(p-q-1)$

$$\left(\sum_{q=1}^{p-1} z^{2(p-q)} w^{q-1} \sum_{i=1}^{q} (-1)^{i-1} \binom{2p-i}{i-1} A^{q-i} \right) - \left(\sum_{q=1}^{p-1} z^{2(p-q-1)} w^{q} \sum_{i=1}^{q} (-1)^{i-1} \binom{2p-i}{i-1} A^{q-i+1} \right)$$
$$+ \left(w^{p-1} \sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i}{i-1} A^{p-i} \right)$$

Changing in the second sum in *q*, *q'*=*q*+1

$$\left(\sum_{q=1}^{p-1} z^{2(p-q)} w^{q-1} \sum_{i=1}^{q} (-1)^{i-1} \binom{2p-i}{i-1} A^{q-i} \right) - \left(\sum_{q=2}^{p} z^{2(p-q)} w^{q-1} \sum_{i=1}^{q-1} (-1)^{i-1} \binom{2p-i}{i-1} A^{q-i} \right)$$
$$+ \left(w^{p-1} \sum_{i=1}^{p} (-1)^{i-1} \binom{2p-i}{i-1} A^{p-i} \right)$$

Separating in the first sum the term with *q*=1 and the term with *i*=*q*; and also separating in the second sum the term with *q*=*p* and in the third sum the term with *i*=*p*

$$\left(z^{2(p-1)}(-1)^0\binom{2p-1}{1-1}\right)$$
$$+\left(\sum_{q=2}^{p-1} z^{2(p-q)}w^{q-1}\left(\sum_{i=1}^{q-1}(-1)^{i-1}\binom{2p-i}{i-1}A^{q-i}+(-1)^{q-1}\binom{2p-q}{q-1}\right)\right)$$
$$-\left(\sum_{q=2}^{p-1} z^{2(p-q)}w^{q-1}\sum_{i=1}^{q-1}(-1)^{i-1}\binom{2p-i}{i-1}A^{q-i}\right)$$
$$-\left(z^{2(p-p)}w^{p-1}\sum_{i=1}^{p-1}(-1)^{i-1}\binom{2p-i}{i-1}A^{p-i}\right)+\left(w^{p-1}\sum_{i=1}^{p-1}(-1)^{i-1}\binom{2p-i}{i-1}A^{p-i}\right)$$
$$+\left(w^{p-1}(-1)^{p-1}\binom{2p-p}{p-1}\right)=$$

Distributing on the second sum

$$\left((-1)^0\binom{2p-1}{1-1}z^{2(p-1)}\right)+\left(\sum_{q=2}^{p-1} z^{2(p-q)}w^{q-1}\sum_{i=1}^{q-1}(-1)^{i-1}\binom{2p-i}{i-1}A^{q-i}\right)$$
$$+\left(\sum_{q=2}^{p-1}(-1)^{q-1}\binom{2p-q}{q-1}z^{2(p-q)}w^{q-1}\right)$$
$$-\left(\sum_{q=2}^{p-1} z^{2(p-q)}w^{q-1}\sum_{i=1}^{q-1}(-1)^{i-1}\binom{2p-i}{i-1}A^{q-i}\right)$$
$$-\left(w^{p-1}\sum_{i=1}^{p-1}(-1)^{i-1}\binom{2p-i}{i-1}A^{p-i}\right)+\left(w^{p-1}\sum_{i=1}^{p-1}(-1)^{i-1}\binom{2p-i}{i-1}A^{p-i}\right)$$
$$+\left((-1)^{p-1}\binom{2p-p}{p-1}z^{2(p-p)}w^{p-1}\right)=$$

Now, cancelling the opposite terms.

$$(-1)^{1-1}\binom{2p-1}{1-1}z^{2(p-1)}w^{1-1}+\sum_{q=2}^{p-1}(-1)^{q-1}\binom{2p-q}{q-1}z^{2(p-q)}w^{q-1}$$
$$+(-1)^{p-1}\binom{2p-p}{p-1}z^{2(p-p)}w^{p-1}=$$

And finally inserting the first and last terms into the middle sum.

$$\sum_{q=1}^{p}(-1)^{q-1}\binom{2p-q}{q-1}z^{2(p-q)}w^{q-1}$$

I get the left hand-side of the equation.

■

Proof of theorem 6:

I want to prove that

$$\sum_{i=1}^{p}(-1)^{i-1}\binom{2p-i-1}{i-1}A^{p-i}=\sum_{i=1}^{p}(-1)^{i-1}\binom{2p-i}{i-1}A^{p-i}+\sum_{i=1}^{p-1}(-1)^{i-1}\binom{2p-2-i}{i-1}A^{p-i}$$

In the right hand part I separate the term i=1 from the first sum

$$(-1)^{1-1}\binom{2p-1}{1-1}A^{p-1}+\sum_{i=2}^{p}(-1)^{i-1}\binom{2p-i}{i-1}A^{p-i}+\sum_{i=1}^{p-1}(-1)^{i-1}\binom{2p-2-i}{i-1}A^{p-1-i}$$

Then change in the second sum i=i+1

$$(-1)^{1-1}\binom{2p-1}{1-1}A^{p-1}+\sum_{i=2}^{p}(-1)^{i}\binom{2p-i}{i-1}A^{p-i}+\sum_{i=2}^{p}(-1)^{i-2}\binom{2p-2-i+1}{i-1-1}A^{p-1-i+1}=$$

$$(-1)^{1-1}A^{p-1}+\sum_{i=2}^{p}(-1)^{i-1}\binom{2p-i}{i-1}A^{p-i}-\sum_{i=2}^{p}(-1)^{i-1}\binom{2p-1-i}{i-2}A^{p-i}=$$

$$(-1)^{1-1}A^{p-1}+\sum_{i=2}^{p}(-1)^{i-1}\left(\binom{2p-i}{i-1}-\binom{2p-1-i}{i-2}\right)A^{p-i}=$$

Because $\binom{a}{b}+\binom{a}{b+1}=\binom{a+1}{b+1}$:

$$(-1)^{1-1}A^{p-1} + \sum_{i=2}^{p}(-1)^{i-1}\binom{2p-1-i}{i-1}A^{p-i} =$$

Re-inserting the first term in the sum

$$(-1)^{1-1}\binom{2p-1-1}{1-1}A^{p-1} + \sum_{i=2}^{p}(-1)^{i-1}\binom{2p-1-i}{i-1}A^{p-i} =$$

$$\sum_{i=1}^{p}(-1)^{i-1}\binom{2p-1-i}{i-1}A^{p-i}$$

■

6.3 Appendix 2: An alternative to the use of trigonometric functions

Given that:

- $U_{p+1}(A) = (A-2)U_p(A) - U_{p-1}(A),\ \ p > 2 \qquad U_1(A) = 1,\ \ U_2(A) = (A-2),$
- If p odd then $U_p(2+A) = U_p(2-A)$, if p even then $U_p(2+A) = -U_p(2-A)$
- All the zeros of $U_p(A)$ lie in the interval *0<A<4*
- $V_p(A) = U_p(A) + U_{p-1}(A)$; $V_1(A) = 1$
- $U_p(0) = (-1)^{p-1}p$ and $V_p(0) = (-1)^{p-1}$
- The degree of $V_p(A)$ and $U_p(A)$ is *p-1*

Call A_i^p the zeros of U_p with $0 < i < p$ and $A_i^p < A_{i+1}^p$

Call B_i^p the zeros of V_p with $0 < i < p$ and $B_i^p < B_{i+1}^p$

I want to prove that:

I. U_p has exactly p-1 zeros, V_p has exactly p-1 zeros

II. If p is even then $A_{\frac{p}{2}}^p = 2$ and $0 < A_1^p < A_1^{p-1} < A_2^p < \cdots < A_{\frac{p}{2}-1}^p < A_{\frac{p}{2}-1}^{p-1} < 2$; if p is odd then $A_{\frac{p-1}{2}}^{p-1} = 2$ and $0 < A_1^p < A_1^{p-1} < A_2^p < \cdots < A_{\frac{p-1}{2}-1}^{p-1} < A_{\frac{p-1}{2}}^p < 2$

III. And finally that $B_1^p < A_1^p < B_2^p < \cdots < A_{p-2}^p < B_{p-1}^p < A_{p-1}^p$

For *p=1*

U_1 and V_1 have no zeros

For *p=2*

U_2 has one zero $A_1^2 = 2$ and V_2 has one zero $B_1^2 = 1$

if we suppose that *I, II* and *III* are true for p and *p-1* then

if p is even

$U_p(0) = -p$ and $U_{p-1}(0) = p-1$, therefore $U_{p+1}(0) = (0-2)U_p(0) - U_{p-1}(0) = 2p - p + 1 = p + 1$

Also $V_{p+1}(0) = (0-1)U_p(0) - U_{p-1}(0) = p - p + 1 = 1$

$U_p(A_1^p) = 0$ and $U_{p-1}(A_1^p) > 0$, therefore $U_{p+1}(A_1^p) = (A_1^p - 2)U_p(A_1^p) - U_{p-1}(A_1^p) = -U_{p-1}(A_1^p) < 0$

Also $V_{p+1}(A_1^p) = U_{p+1}(A_1^p) + U_p(A_1^p) < 0$

Thus, by the medium value theorem, in the interval $(0, A_1^p)$ exists the point A_1^{p+1} such that $U_{p+1}(A_1^{p+1}) = 0$ and the point B_1^{p+1} such that $V_{p+1}(B_1^{p+1}) = 0$. Also in this interval $U_p(A) < 0$ and being that $V_{p+1}(A) = U_{p+1}(A) + U_p(A)$ it means that throughout the hole interval $V_{p+1}(A) < U_{p+1}(A)$,

Now consider the points A_i^p, $A_i^{p-1}$and A_{i+1}^p for odd (even) i. $0 < i < \frac{p}{2}$

$U_p(A_i^p) = 0$ and with $U_{p-1}(A_i^p)$ positive (negative) then $U_{p+1}(A_i^p) = -U_{p-1}(A_i^p)$ is negative (positive).

Also $V_{p+1}(A_i^p) = U_{p+1}(A_i^p) + U_p(A_i^p)$ is negative (positive)

$U_{p-1}(A_i^{p-1}) = 0$ and $U_p(A_i^{p-1})$ is positive (negative) then $U_{p+1}(A_i^{p-1}) = (A_i^{p-1} - 2)U_p(A_i^{p-1})$ is negative (positive).

$U_p(A_{i+1}^p) = 0$ and $U_{p-1}(A_{i+1}^p)$ negative (positive) then $U_{p+1}(A_{i+1}^p) = -U_{p-1}(A_{i+1}^p)$ is positive (negative).

Also $V_{p+1}(A_{i+1}^p) = U_{p+1}(A_{i+1}^p) + U_p(A_{i+1}^p)$ is positive (negative)

Thus, by the medium value theorem, in the interval (A_i^{p-1}, A_{i+1}^p) exists the point A_{i+1}^{p+1} such that $U_{p+1}(A_{i+1}^{p+1}) = 0$ and in the interval (A_i^p, A_{i+1}^p) the point B_{i+1}^{p+1} such that $V_{p+1}(B_{i+1}^{p+1}) = 0$.

Also in this interval $U_p(A)$ is positive (negative) and being that $V_{p+1}(A) = U_{p+1}(A) + U_p(A)$ it means that throughout the hole interval $V_{p+1}(A) > U_{p+1}(A)$ ($V_{p+1}(A) < U_{p+1}(A)$)

Thus there is a zero A_{i+1}^{p+1} of U_{p+1} corresponding to each interval (A_i^{p-1}, A_{i+1}^p) with $0 < i < \frac{p}{2}$, plus the zero A_1^{p+1} of U_{p+1} in the interval $(0, A_1^p)$. These are *p/2* zeros of U_{p+1} between *0* and *2*. By symmetry there are p zeros of U_{p+1} between *0* and *4.* The degree of U_{p+1} is p, therefore there is one and only one of these zeros in each interval.

Also there is a zero B_{i+1}^{p+1} of V_{p+1} corresponding to each interval (A_i^p, A_{i+1}^p) with $0 < i < \frac{p}{2}$, plus the zero B_1^{p+1} of V_{p+1} in the interval $(0, A_1^p)$. These are *p/2* zeros of V_{p+1} between *0* and *2*. By symmetry there are p zeros of V_{p+1} between *0* and *4.* The degree of V_{p+1} is p, therefore there is one and only one of these zeros in each interval.

Thus, for even values of p, U_{p+1} and V_{p+1} have exactly p zeros.

In the interval $(0, A_1^p)$, $U_{p+1}(0)$ and $V_{p+1}(0)$ are positive while $U_{p+1}(A_1^p)$ and $V_{p+1}(A_1^p)$ are negative, also in this interval $V_{p+1}(A) < U_{p+1}(A)$ thus $B_1^{p+1} < A_1^{p+1}$. For odd (even) i. $0 < i < \frac{p}{2}$, $U_{p+1}(A_i^p)$ and $V_{p+1}(A_i^p)$ are both negative (positive) and $U_{p+1}(A_{i+1}^p)$ and

$V_{p+1}(A_{i+1}^p)$ are both positive (negative). Also in these intervals, $V_{p+1}(A) > U_{p+1}(A)$ ($V_{p+1}(A) < U_{p+1}(A)$), therefore $B_{i+1}^{p+1} < A_{i+1}^{p+1}$

As p is even $U_p(2-A) = -U_p(2+A)$ while $U_{p-1}(2-A) = U_{p-1}(2+A)$. The interval $(0, A_1^p)$ corresponds in both symmetries to the interval $(A_{p-1}^p, 4)$ in this interval $U_{p+1}(A_{p-1}^p)$ and $V_{p+1}(0)$ are positive while $U_{p+1}(4)$ and $V_{p+1}(4)$ are negative, also in this interval $V_{p+1}(A) < U_{p+1}(A)$ thus $B_p^{p+1} < A_p^{p+1}$. Each interval (A_i^p, A_{i+1}^p) with $0 < i < \frac{p}{2}$ corresponds in both symmetries to another interval (A_{p-i-1}^p, A_{p-i}^p) where $U_{p-1}(A)$ has the same sign that in the corresponding interval while the sign of $U_p(A)$ changes. Consequently, in the intervals corresponding to odd (even) i. $0 < i < \frac{p}{2}$, $U_{p+1}(A_{p-i-1}^p)$ and $V_{p+1}(A_{p-i-1}^p)$ are both negative (positive) and $U_{p+1}(A_{p-i}^p)$ and $V_{p+1}(A_{p-i}^p)$ are both positive (negative). Also in these intervals, $V_{p+1}(A) > U_{p+1}(A)$ ($V_{p+1}(A) < U_{p+1}(A)$), therefore $B_{p-i-1}^{p+1} < A_{p-i-1}^{p+1}$.

If p is odd

$U_p(0) = p$ and $U_{p-1}(0) = -p+1$, therefore $U_{p+1}(0) = (0-2)U_p(0) - U_{p-1}(0) = -2p + p - 1 = -p - 1$

Also $V_{p+1}(0) = (0-1)U_p(0) - U_{p-1}(0) = -p + p - 1 = -1$

$U_p(A_1^p) = 0$ and $U_{p-1}(A_1^p) < 0$, therefore $U_{p+1}(A_1^p) = (A_1^p - 2)U_p(A_1^p) - U_{p-1}(A_1^p) = -U_{p-1}(A_1^p) > 0$

Also $V_{p+1}(A_1^p) = U_{p+1}(A_1^p) + U_p(A_1^p) > 0$

Thus, by the medium value theorem, in the interval $(0, A_1^p)$ exists the point A_1^{p+1} such that $U_{p+1}(A_1^{p+1}) = 0$ and the point B_1^{p+1} such that $V_{p+1}(B_1^{p+1}) = 0$. Also in this interval $U_p(A) > 0$ and being that $V_{p+1}(A) = U_{p+1}(A) + U_p(A)$ it means that throughout the hole interval $V_{p+1}(A) > U_{p+1}(A)$,

Now consider the points A_i^p, $A_i^{p-1}$and A_{i+1}^p for odd (even) i. $0 < i < \frac{p-1}{2} + 1$

$U_p(A_i^p) = 0$ and with $U_{p-1}(A_i^p)$ negative (positive) then $U_{p+1}(A_i^p) = -U_{p-1}(A_i^p)$ is positive (negative).

Also $V_{p+1}(A_i^p) = U_{p+1}(A_i^p) + U_p(A_i^p)$ is positive (negative)

$U_{p-1}(A_i^{p-1}) = 0$ and $U_p(A_i^{p-1})$ is negative (positive) then $U_{p+1}(A_i^{p-1}) = (A_i^{p-1} - 2)U_p(A_i^{p-1})$ is positive (negative).

$U_p(A_{i+1}^p) = 0$ and $U_{p-1}(A_{i+1}^p)$ positive (negative) then $U_{p+1}(A_{i+1}^p) = -U_{p-1}(A_{i+1}^p)$ is negative (positive).

Also $V_{p+1}(A_{i+1}^{p}) = U_{p+1}(A_{i+1}^{p}) + U_p(A_{i+1}^{p})$ is negative (positive)

Thus, by the medium value theorem, in the interval (A_i^{p-1}, A_{i+1}^{p}) exists the point A_{i+1}^{p+1} such that $U_{p+1}(A_{i+1}^{p+1}) = 0$ and in the interval (A_i^p, A_{i+1}^p) the point B_{i+1}^{p+1} such that $V_{p+1}(B_{i+1}^{p+1}) = 0$. Also in this interval $U_p(A)$ is negative (positive) and being that $V_{p+1}(A) = U_{p+1}(A) + U_p(A)$ it means that throughout the hole interval $V_{p+1}(A) < U_{p+1}(A)$ ($V_{p+1}(A) > U_{p+1}(A)$)

Thus there is a zero A_{i+1}^{p+1} of U_{p+1} corresponding to each interval (A_i^{p-1}, A_{i+1}^{p}) with $0 < i < \frac{p-1}{2}$, plus the zero A_1^{p+1} of U_{p+1} in the interval $(0, A_1^p)$. These are $\frac{p-1}{2}$ zeros of U_{p+1} between *0* and $A_{\frac{p-1}{2}}^{p}$ wich is the greatest zero of U_p smaller than *2*. By symmetry there are *p-1* zeros of U_{p+1} plus another zero in the interval $\left(A_{\frac{p-1}{2}}^{p}, A_{\frac{p-1}{2}+1}^{p}\right)$. The degree of U_{p+1} is p, therefore there is one and only one of these zeros in each interval.

Also there is a zero B_{i+1}^{p+1} of V_{p+1} corresponding to each interval (A_i^p, A_{i+1}^p) with $0 < i < \frac{p-1}{2}$, plus the zero B_1^{p+1} of V_{p+1} in the interval $(0, A_1^p)$. These are $\frac{p-1}{2}$ zeros of V_{p+1} between *0* and $A_{\frac{p-1}{2}}^{p}$ wich is the greatest zero of U_p smaller than *2*. By symmetry there are p zeros of V_{p+1} plus another zero in the interval $\left(A_{\frac{p-1}{2}}^{p}, A_{\frac{p-1}{2}+1}^{p}\right)$. The degree of V_{p+1} is p, therefore there is one and only one of these zeros in each interval.

Thus, for odd values of p, U_{p+1} and V_{p+1} have exactly p zeros.

In the interval $(0, A_1^p)$, $U_{p+1}(0)$ and $V_{p+1}(0)$ are negative while $U_{p+1}(A_1^p)$ and $V_{p+1}(A_1^p)$ are positive, also in this interval $V_{p+1}(A) > U_{p+1}(A)$ thus $B_1^{p+1} < A_1^{p+1}$. For odd (even) i. $0 < i < \frac{p-1}{2} + 1$, $U_{p+1}(A_i^p)$ and $V_{p+1}(A_i^p)$ are both positive (negative) and $U_{p+1}(A_{i+1}^p)$ and $V_{p+1}(A_{i+1}^p)$ are both negative (positive). Also in these intervals, $V_{p+1}(A) < U_{p+1}(A)$ ($V_{p+1}(A) > U_{p+1}(A)$), therefore $B_{i+1}^{p+1} < A_{i+1}^{p+1}$.

As p is odd $U_p(2 - A) = U_p(2 + A)$ while $U_{p-1}(2 - A) = -U_{p-1}(2 + A)$. The interval $(0, A_1^p)$ corresponds in both symmetries to the interval $(A_{p-1}^p, 4)$ in this interval $U_{p+1}(A_{p-1}^p)$ and $V_{p+1}(0)$ are negative while $U_{p+1}(4)$ and $V_{p+1}(4)$ are positive, also in this interval $V_{p+1}(A) > U_{p+1}(A)$ thus $B_p^{p+1} < A_p^{p+1}$. Each interval (A_i^p, A_{i+1}^p) with $0 < i < \frac{p-1}{2}$ corresponds in both symmetries to another interval (A_{p-i-1}^p, A_{p-i}^p) where $U_{p-1}(A)$ has the opposite sign that in the corresponding interval while the sign of $U_p(A)$ remains the same. Consequently, in the intervals corresponding to odd (even) i. $0 < i < \frac{p-1}{2}$, $U_{p+1}(A_{p-i-1}^p)$ and $V_{p+1}(A_{p-i-1}^p)$ are both positive (negative) and $U_{p+1}(A_{p-i}^{p})$ and $V_{p+1}(A_{p-i}^{p})$ are both negative (positive). Also in these intervals, $V_{p+1}(A) < U_{p+1}(A)$ ($V_{p+1}(A) > U_{p+1}(A)$), therefore $B_{p-i-1}^{p+1} < A_{p-i-1}^{p+1}$

Summarizing I've showed that:

I, II and *III* are valid in the case *p=1* and *p=2*.

Assuming *I, II* and *III* are valid for *p* and *p-1* then

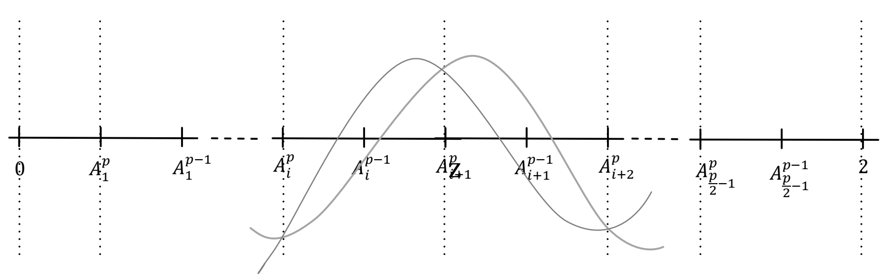
If *p* is even dividing the interval (0,2) with the points A_i^p, were $0 < i < \frac{p}{2}$

-In each one of these $p/2$ intervals there is one and only one point A_i^{p+1} that is a zero of $U_{p+1}(A)$, which proofs statement *II* for *p+1* in the case of even *p*.
-In each one of these $p/2$ intervals there is one and only one point of B_i^{p+1} that is a zero of $V_{p+1}(A)$
-By symmetry we get *p* intervals preserving the previous properties and because they are as much as the degree of the polynomials, the zeros of $U_{p+1}(A)$ and $V_{p+1}(A)$ are exactly *p*, proving *I* for *p+1* in the case of even *p*.
-On each interval between *0* and *4* $B_i^{p+1} < A_i^{p+1}$, proving statement *III* for *p+1* in the case of even p.

If *p* is odd consider the intervals (A_i^p, A_{i+1}^p) were $0 < i < \frac{p-1}{2}$ as $A_{\frac{p-1}{2}}^p$ is the greatest zero of U_p smaller than *2*.

-In each one of these $\frac{p-1}{2}$ intervals there is one and only one point A_i^{p+1} that is a zero of $U_{p+1}(A)$, which proofs statement *II* for *p+1* in the case of odd *p*.
-In each one of these $\frac{p-1}{2}$ intervals there is one and only one point of B_i^{p+1} that is a zero of $V_{p+1}(A)$
-By symmetry we get *p-1* intervals preserving the previous properties. Also in the interval $(A_{\frac{p-1}{2}}^p, A_{\frac{p-1}{2}+1}^p)$ there are zeros of $U_{p+1}(A)$ and $V_{p+1}(A)$. As the number of intervals considered are as much as the degree of the polynomials, the zeros of $U_{p+1}(A)$ and $V_{p+1}(A)$ are exactly *p*, proving *I* for *p+1* in the case of odd *p*.
-On each interval between *0* and *4* $B_i^{p+1} < A_i^{p+1}$, proving statement *III* for *p+1* in the case of odd p.

Thus *I, II* and *III* are valid for *p+1* so by complete induction the assertions *I, II and III* are valid for all natural *p* completing the proof.



7.0. Summary of the motivations for this proof

There are many visions of mathematics but some of them are more adequate for teaching than others. A particularly important view for teaching considers that Mathematics is originated in reality, which ultimately justifies its applicability and importance, and therefore his role as a school subject of studies.

The mathematical process from reality to the solution of problems has to take into account the human limitations to understand and master the complication of reality. The strategies to do this include several choices that will show their adequacy only at the end of the process, as the adequacy of these choices is tied to whether the solutions to the original problems are achieved. Some of these choices are: to consider only the aspects of reality that seem relevant to the problem and discard all others, to organize these aspects in a certain manner, to add other aspects that doesn't belong to the original situation in order to make it more manageable, to generalize, etc.

On a second stage, the knowledge obtained by these processes is then organized into axiomatic systems. Also the solutions strategies for a certain type of problem are mechanized into algorithms. The processes on the first stage are subjective or inter-subjective while the processes on the second stage are objective. Other visions of mathematics tend to consider only the processes on the second stage, but on these visions the link to reality that is fundamental for teaching, is lost.

The mathematical concepts can be approached with several levels of profoundness, technical, intuitive, formal, fundamental or focused on the acquisition process. The study of the profound levels (fundaments and acquisition process) which relate to the intuitive part of the mathematical concept is essential to design effective teaching methods (at intuition level) for these very concepts at school.

Many examples can be found to illustrate how the knowledge at these profound levels can effectively orient didactical practices. Nevertheless for many other concepts we find that a deep study of it is unavailable or incomplete.

In the case of polynomial functions, which is an elementary and very important argument in school mathematics, most justifications of the procedures that involve polynomials can be found making reference to the corresponding properties of addition and multiplication, operations on which polynomials are founded. Hence, when it comes to the factorization of polynomials in the environment of real numbers (which is the natural ambit to study them at school) there is one result that can be a cause of perplexity.

Every real polynomial can be expressed as the product of real polynomials of the first or second degree. So far, every single proof of this fact makes explicit or implicit use of the complex number system, and this raises at least two important questions:

1) Are the complex numbers essential to prove that any real polynomial can be expressed by linear and quadratic factors?

2) If they are not, which are the essential ideas on which such a fact rests?

7.1. The essential ideas of the proof

My proof provides a negative answer to question 1) because without any loss of rigor it successfully avoids any reference to the complex numbers. Moreover this proof is a font of valuable information to answer question 2). These are the essential ideas of this proof:

- Odd degree polynomials surely have one root as a consequence of the intermediate value theorem, therefore they can be expressed as the product of a linear factor and an even degree polynomial. Thus is enough to see that every even degree polynomial has a quadratic factor.
- A polynomial has a quadratic factor if and only if the division between the polynomial and this factor has remainder zero.
- In general the remainder of the division between a polynomial of even degree and a quadratic polynomial is itself a polynomial with degree at most 1 and its coefficients can be interpreted as two surfaces. The two independent variables of these surfaces are the coefficients of the quadratic polynomial.
- The goal then is to see if these surfaces zero-level curves intersect each other, as the intersection point represents the coefficients of a quadratic polynomial for which the division has remainder zero.
- The algebraic formula for these surfaces is a two variable polynomial. The terms of these polynomials can be arranged grouping together the terms that multiply the same coefficient of the original polynomial. Expressed in this particular way, they become a linear combination of other surfaces. Let’s call these new surfaces the composing surfaces
- The remainder has many recursive proprieties that are inherited from the division that generates it when performed at stages. These recursive proprieties are the ones that

allowed me to prove that the previously mentioned composing surfaces have zero-level curves in the form of nested parabolas with the vertex at the origin of coordinates.

- Using parabolic coordinates I showed that far away enough from the origin of coordinates, the behavior of the surfaces of the remainder are controlled only by one of their composing surfaces that contain the term of higher degree.
- Each of these two composing surfaces (one for each surface of the remainder) have as zero-level curves a set of nested parabolas. The parabolas corresponding to one of these surfaces interweave with the parabolas corresponding to the other surface.
- As a consequence, the zero-level curves of the remainder surfaces present an interwoven behavior far away from the origin and this is enough to deduce the existence of an intersection of these curves of zeros (essentially using continuity).

7.2. Ideas that were already present in Gauss's proof

Some of the proving techniques I used are based on underlying ideas that are very similar to the ones that I pointed out when analyzing Gauss's proof:

- Gauss also used two surfaces, but instead of being the remainder of the division they were the real and imaginary part of the polynomial considered among complex numbers.
- He also considered the terms of higher degree to study the behavior of the surface in a region far away enough from the origin of coordinates.
- The interwoven pattern is also present between these surfaces, and from this point on, the proofs are almost identical.

These coincidences support the opinion that all these ideas were somehow essential, not only for the fundamental theorem of algebra for complex polynomials but also for the fact that every real polynomial can be factored using first or second degree polynomials. This observation was difficult to do by just analyzing Gauss proof because in that proof it could be that these elements were only important for complex polynomials.

Real polynomials and complex polynomials are two very different things. A real polynomial involves two real variables (one independent and one dependent) and its graphic is a curve in a Cartesian plane. The respective complex polynomial involves four real variables (two independent and two dependent) and its geometric interpretation is a particular transformation of the whole plane, this transformation includes the particular product between complex numbers for which a profound explanation has yet to be found. The real and imaginary parts of this complex polynomial involve three real variables (two independent and one dependent) and these are the surfaces used in Gauss's proof.

Given one real polynomial, there is no difficulty on calculating the real and imaginary polynomials of the corresponding complex polynomial. Nevertheless, this gives no hint to answer the question: What do these polynomials, that we obtain from the original real polynomial, have to do with its factorization? This question could have a merely technical answer: Using these tools we get to a fine proof of the theorem, so at the end everything works despite these open questions. Such an answer is satisfactory for someone controlling the correctness of the proof, but if we want to really understand how the proof works this cannot be the only answer.

On the other hand in my proof the surfaces I use are obtained by calculating the remainder of the division between a real polynomial and a quadratic polynomial. The remainder of a division has a strong meaningful link with the factorization of the polynomial that is being divided: if the remainder is zero, the divisor is a factor. By doing this I not only get a rigorous proof that works, but also a proof that can meaningfully justify every single step.

7.3. New ideas present in this proof

Even if there were some similarities with the complex proof, in this new proof there are new essential aspects that came to light. The most important of them is the recursive structure of the expressions involved. On many steps of this proof I use induction relying on these recursive structures. A central point of the new proof is the part where I show how some substructures can be expressed trigonometrically and subsequently I used the information provided by these trigonometric expressions. However, the information provided by these trigonometric expressions is overabundant and the theorem depends only on the fact that the curves of zeros are interweaved, and the knowledge of the exact position of these zeros isn't fundamental. This is why in the second annex I show how this fundamental property of being interweaved can be obtained directly from the recursive properties without using the trigonometric functions. This means that trigonometric functions are not essential for this theorem.

If we take a closer look at Gauss's proof we notice that in order to separate the real and imaginary polynomials he has to use the DeMoivre's theorem. In DeMoivre's theorem similar trigonometric functions are used, therefore similar recursive structures. Thus recursive structures were also present in Gauss's proof but they don't seem to have such a central role as in my proof.

At this point, after this last observation, it seems every important idea I needed to prove my theorem was also necessary to prove the FTA among complex numbers, and no other fundamental idea is needed, other than the very basic definition of a complex number. One could have hoped to see the opposite situation, were some ideas were the same and thus fundamental to the real numbers ambit, and some were exclusive to the complex numbers ambit. In such a case these ideas could help to better understand the fundaments of the complex numbers. But this is not the case, so the suggestion is that as far as the fundamental theorem of algebra goes, complex numbers remain only a useful, manageable and powerful tool to describe real number situations.

8.0. Ideas for further research

Fundaments of mathematics:

If I'm correct about the fundamental ideas of the theorems being almost the same, could it be possible to use my theorem to prove the fundamental theorem of Algebra for complex numbers?

How come the polynomials of the real part and imaginary part corresponding to a real polynomial have so many properties in common with the polynomials obtained as the remainder of the division between the same polynomial and a quadratic polynomial? Could it be found a clear relation between the two approaches? Maybe the real and imaginary part can be expressed as the remainder of some division?

Could a profound study of DeMoivre's theorem reveal that the recursive properties are also fundamental to the complex number system?

In the corollary I achieved a real proof of another fact that is expressed completely inside the field of the real numbers but which proof was always given using complex numbers. Are there many other examples like this one? Could the creation of real proofs to these problems illuminate on the true nature of the complex numbers?

Didactics of mathematics:

With the desire to work thoroughly, I tried to be as detailed as possible in my proof, sometimes using several pages to go through non-fundamental aspects with the danger to cause the reader to get lost. I believe much shorter versions of the proof can evince the main ideas and avoid the peripheral aspects of it. I wonder if a didactic transposition is possible for students in the last years of school. I realize that the prerequisites are quite heavy (induction, continuity) and they need to be learned with a certain level of profundity. Anyhow the purely didactic aspects for the concept of polynomial factorization are still an open matter of study and I hope my work can be a valuable resource for it.

What is the actual status of the research at the fundaments level and acquisition process level of my taxonomy for the profundity of the contents, for the elementary topics of mathematics? Could this study lead to more interesting open questions as the one I have investigated? For example I'm thinking about the vector product. The original meaning of product is a repeated application of addition. However, the sense in which we interpret the word "repeated" changes greatly from field to field. In which sense the product of vectors is a repetition of the addition of vectors?

Instead, when the profound studies of a concept are available, a didactics research project could work on specifying the suggestions these studies imply for the teaching methods of that concept at school.

BIBLIOGRAPHY:

- **Ahlfors L. V.**
  - Complex Analysis, McGraw Hill 1966.
- **Baratella S. –Ferro R.**
  - Non Standard Regular Finite Set Theory, Math. Log. Quart., 41 (1995), pp. 161-172.
- **Bell J.T. –Machover M.**:
  - A first course in mathematical Logic, North Holland
- **Bertagna G.**
  - Cultura e pedagogia per la scuola di tutti, La Scuola, 1992
  - Dall'Educazione alla Pedagogia, Editrice La Scuola, 2010
  - Da «uomini» a «persone». Itinerari di epistemologia pedagogica, La Scuola - 2010
- **Cantor G.**
  - Ueber unendliche, lineare Punktmannischfaltigkeiten, Mathematische Annalen, volume 21, of 1883.
- **Cassari E.**
  - Questioni di Filosofia della Matematica, Feltrinelli 1964.
- **Celluci C.**
  - La Filosofia della Matematica, Laterza 1967.
- **DeMoivre A., Jones G.**
  - De Reductione Radicalium ad Sempliciores Terminus, Seu de extrahenda radice quacunque data ex Binomio $a + \sqrt[2]{b}$, vel $a + \sqrt[2]{-b}$. Philosphical Transactions of The Royal Society 1737-1738 40, 463-478
- **Devlin K.**
  - Dove va la matematica, Bollati Boringhieri, Torino, 1999.
  - The Man of Numbers: Fibonacci's Arithmetic Revolution, Walker Books (2011).
  - Leonardo and Steve: the Young Genius Who Beat Apple to Market by 800 Years, Ted Weinstein (2011), e-book.
  - The Math Gene: How Mathematical Thinking Evolved and Why Numbers Are Like Gossip, Basic Books (2000).
  - Mathematics Education for a New Era: Video Games as a Medium for Learning, AK Peters/CRC Press (2011).
- **Eisermann M.**
  - The Fundamental Theorem of Algebra Made Effective: An Elementary Real-Algebric Proof via Sturm Chains. The American Mathematical Monthly November 2012.
- **Enriques F.**

- Le matematiche nella storia e nella cultura, Zanichelli, Bologna 1938 (Ristampa anastatica 1971).
- Insegnamento dinamico, Periodico di matematica, 1921, pp.6-16.

- **Ferro R.**
  - Conoscere umanamente la conoscenza umana in H. Puyau e al. eds.: "Epistemologia de las ciencias: la vida humana, su especificidad", CIAFIC Ediciones, Buenos Aires, 2003, pp. 203-258.
  - Realtà, conoscenza della realtà e rappresentazione della conoscenza della realtà, negli Atti del Simposio de Epistemología de las Ciencias Naturales su "La filosofía actual y su relación con el estado contemporáneo de la Ciencia" Buenos Aires 2007, CIAFIC Ediciones, Buenos Aires, in stampa.
  - Abstraction and objectivity in Mathematics. MAA Session on Philosophy of Mathematics in Teaching and Learning. JMM New Orleans 2011.
  - Representation and Explanation in the Sciences, Colloque de Louvain-la-Neuve, Belgique, 26-28 april 2011.
- **Fine Benjamin, Rosenberg Gerhard;**
  - The Fundamental Theorem of Algebra, Springer
- **Fox L. and Parker I.B**
  - Chebyshev Polynomials in numerical analysis. Oxford University Press.
- **Gauss C.F**
  - First Proof, doctoral thesis Helmstadt- (in-werke,3,1-30), 1799
  - Second Proof, Comm. Soc. Goett. 3, 1814/15, 107-142
  - Third Proof, Comm. Soc. Goett. 3 (in-werke, 3) 1816, 59-64
  - Fourth Proof, Abhand. Der Ges. Der Wiss. Zu Goett. 4. 1848/50, 3-34
- **Gentile G.**
  - Sommario di pedagogia come scienza filosofica, Sansoni, vol. I-II, Firenze 1959.
- **Giusti E.**
  - Ipotesi sulla natura degli oggetti matematici, Bollati Boringhieri, Torino, 1999.
- **Hardy G.H.**
  - Apologia di un matematico, Garzanti Elefanti, Bergamo 2008.
- **Krantz S. G.**
  - Complex Analysis the Geometric View Point, The Carus Mathematical Monographs MAA Number 23, 1990.
- **Lakoff G. - Nunez R.**
  - Where mathematics comes from. How the embodied mind brings mathematics into being, Basic Books, 2000.

- **Lolli G.**
  - Filosofia della matematica, Il Mulino, Bologna, 2002.
  - QED - Fenomenologia della dimostrazione, Bollati Boringhieri, Torino, 2005
  - "La questione dei fondamenti tra matematica e filosofia", in S. Albeverio e F. Minazzi (a cura di), Matematica e filosofia, nn. 14-15 di Note di Matematica, Storia, Cultura (Pristem/Storia, Università Bocconi, Milano), 2006, pp. 17-35.
- **Margiotta U.**
  - Professione Docente, Formazione e Insegnamento SSIS Veneto, IV-1/2 2006.
- **Mazur B.**
  - Imagining Numbers (particularly the square root of minus 15)
- **Nahin P.l**
  - An Imaginary Tale, the story of $\sqrt[2]{-1}$, Princeton University Press.
- **Peano G.**
  - Opere Scelte, a cura dell' UMI, Edizioni Cremonese, Roma 1957.
- **Poincarè H.**
  - L‟invenzione matematica‟ In Opere epistemologiche, vol.II., Piovan Editore, Abano Terme 1989.
- **Polya G.**
  - How to Solve It. Princeton University Press, 1973.
- **Rio Branco de Oliveira O.**
  - The Fundamental Theorem of Algebra: From the four basic operations. The American Mathematical Monthly November 2012.
- **Rivlin T. J.**
  - The Chebyshev Polynomials, Jhon Wiley & sons.
- **Sfard A.**
  - Psicologia del Pensiero Matematico, Erikson 2008.
- **Spencer P.**
  - Question Corner, University of Toronto web page.