Workshop

ALGORITHMS: AN APPROACH BASED ON HISTORICAL TEXTS IN THE CLASSROOM

Anne Boyé,[a] Martine Bühler[b] & Anne Michel-Pajus[c]

[a]IREM des Pays de Loire, [b]IREM Paris 7 Denis Diderot, [c]APMEP

The new curriculum in French High schools, which is currently being implemented, highlights the importance of algorithms in mathematics, promoting a kind of algorithmic way of thinking. It explicitly requires that elementary but varied works on algorithms be carried out in the classroom.
In this paper, we attempt to show that this work could be carried out on the basis of historical sources, and not only with a computer.

USING HISTORICAL TEXTS: WHY AND HOW?

We have been working on introducing a historical perspective into the mathematics classroom for some thirty years. The main reasons for that are that we think that:

- The history of mathematics allows us to motivate the introduction of a concept and to see the use of it.
- It is an inexhaustible source of problems.
- It shows mathematics in the process of being done, in relation with its time and culture, and not as dogmatic objects.
- It can raise interest in interdisciplinary projects. These enable students to become aware that mathematics contributes to the culture of an age.

Our way of implementing these ideas follows these principles:

- The history of mathematics is integrated into the mathematics curriculum. It is not treated chronologically, but in step with the concepts as they are taught.
- We have students read the original texts because these allow them to become aware of the evolution of the notion of rigour, the multiple attempts which lead to the notations that they use, and the long gestation of concepts. And also because a text written by a mathematician, since it is not written *a priori* for students, demands understanding in depth (in contrast to the automatic reactions when facing ritual and stereotyped exercises).
- The reading of the texts is often supplemented by a set of commentaries presenting the document (context, methods and vocabulary) and exercises restating the problems of the text in modern terms.

A NEW CURRICULUM IN FRANCE

For this workshop, we have chosen the theme "algorithms" because a new curriculum (beginning in 2009) in French High schools highlights the importance of algorithms in mathematics, promoting a kind of algorithmic" way of thinking. It explicitly requires that elementary work on algorithms be carried out in the classroom more or less difficult, in a variety of situations.

> "Algorithmics have a natural place in all fields of mathematics, and the associated problems have connections to other parts of the curriculum (functions, geometry, statistics and probability, logic), as well as to other disciplines or to everyday life."
>
> "The algorithmic process has been, since the beginning of time, an essential part of mathematical activity. In the first years of secondary education, pupils encounter algorithms (Algorithms of Elementary Arithmetic Operations, Euclid's Algorithm, Algorithms in Geometrical Constructions). What is proposed in the curriculum is **formalization** in natural language". (10^{th} grade)

The first use of the word "algorithm" that we know comes from Carmen *de algorismo*, by Alexandre de Villedieu (circa 1220): "This new art is called the algorismus, in which we derive such benefit out of these twice five figures of the Indians: 0 9 8 7 6 5 4 3 2 1." Here "algorismus" refers to the "art" of Algus (or Argus, or Aldus), the latinized name of Al-Kwhārizmī, whose *"The Book of Addition and Subtraction According to the Hindu Calculation"* survived only in its Latin translation. The first words of an untitled manuscript are *Dixit algorizmi* (so said al-Khwārizmī). Beginning in the XIIth century, this positional system of numeration spread into the medieval Europe, in competition with the then-common use of "abacus (counting tables) and tokens." Thus, in the beginning, the word "algorithm" only referred to arithmetical operations. During the course of time, its meaning was extended from routine arithmetic procedures "to mean, in general, the method and notation of all types of calculation. In this sense we say the *algorithm of integral calculus*, the *algorithm of exponential calculus*, the *algorithm of sinus*, etc." as D'Alembert wrote in the article "Algorithme" of the *Encyclopédie* .

Now, here is the definition in our curriculum[1]:

> "An algorithm is defined as an operational method allowing one to solve, with a number of clearly specified steps, all the instances of a given problem. This method can be carried out by a machine or a person".

This definition involves that the number of steps is finite and that the result is the right answer!

[1] 7 th year students majoring in « Informatique et Sciences du Numérique »

AN EXAMPLE OF DIDACTICAL SITUATION: FERMAT ABOUT THE FACTORIZATION OF LARGE NUMBERS.

By 1631, Fermat was councillor at the parliament in Toulouse, and greatly interested in mathematics. He met the mathematician Carcavi, who was also a councilor; it was Carcavi who put Fermat into contact with Mersenne and his group. In the early seventeenth century, there was no scientific journal. Exchanges between scientists were by letters. Mersenne helped to coordinate correspondence between all Europeans scientists; he had nearly 140 correspondents, including astronomers and philosophers as well as mathematicians. Most of Fermat's work in number theory is known by his correspondence with Mersenne.

We investigate with the students a method of factorization of large numbers developed by Fermat. (Martine Bühler worked about this text with) These students are engaged in a scientific curriculum (17-18 years old). In France, such students have six weekly hours in mathematics, and some of them have two additional weekly hours (called "specialty mathematics"). For this "specialty mathematics", the curriculum has two parts: number theory and matrices.

In 1643, Mersenne challenged Fermat to find "in less than a day" whether the number 100,895,598,169 is prime or not; and, if not, to find the factorization of this number, and to give a general method of factorization. Fermat gave a reply stating "this number is the product of 898,423 by 112,303, which are both prime". He later explained, in another letter, his method of factorization. It's this second letter that we studied in class (in January 2014). The letter is given in Appendix 1. We give below in modern language the method explained in the text.

We want to factorize a non-square odd natural number N. If N is even, it is easy to factorize, and if N is known as a square, N is already factorized. We use ordinary algebraic identities to explain the method.

If we can write N as a difference of two squares, then $N = a^2 - b^2 = (a+b)(a-b)$.Then we have factorized N: $N = p*q$ with $p = a+b$ and $q = a - b$. Conversely, if $N = p*q$, then p and q are odd numbers (because, if p or q is even, the product is even). But N is odd. So, we can write $N = \left(\frac{p+q}{2}\right)^2 - \left(\frac{p-q}{2}\right)^2$: with $p+q$ and $p-q$ even, so that $\frac{p+q}{2}$ and $\frac{p-q}{2}$ are natural numbers. The problem of writing N as a difference of two squares is therefore equivalent to the problem of factorizing N.

We thus have a method for factorizing N. We want now to find natural numbers a and b so that $N = a^2 - b^2$, or to obtain an integer a so that there is a b such as $a^2 - N = b^2$; that is to say, an integer a so that $a^2 - N$ is a square. We must have $a^2 - N \geq 0$, so that $a \geq \sqrt{(N)}$. There are well-known algorithms for the calculation of the square root of an integer N "by hand". We can then calculate the greatest integer less than $\sqrt{(N)}$, namely $\left[\!\left[\sqrt{(N)}\right]\!\right]$. If N is a square, we stop there. If not, we start our searching with

$a = \left[\!\left[\sqrt{(N)}\right]\!\right] + 1$. If $a^2 - N$ is a square, we stop there. If not, we try $a+1$: is $(a+1)^2 - N$ a square? And so on with $a+2$, $a+3$, etc. until we reach a square.

This is clearly an algorithmic method and the algorithm returns a value of a such that $a^2 - N = b^2$, with b = $\sqrt{a^2 - N}$. We can formalize the algorithm as follows:

Input: N, non-square odd integer
Procedure: a is set equal to $\left[\!\left[\sqrt{(N)}\right]\!\right] + 1$
While $\sqrt{a^2 - N}$ is not integer, do
a changes to $a + 1$
WhileEnd
Output: a and $\sqrt{a^2 - N}$.

We now describe the project that was assigned to the students: "***Large numbers factorization***"

Part I: Difference between 2 squares and factorization: the Fermat's method

In 1643, Fermat responded to Mersenne who challenged him – Fermat – to find "in less than a day" a factorization of 100,895,598,169. Fermat found this factorization (898,423 by 112,303), and explained in a later letter a general method for factorizing large numbers. We'll read this letter together.

Let N be an odd non- square natural number.

1) We suppose that $N = a^2 - b^2$ with a and b natural numbers. Determine, depending on a and b, two natural numbers p and q such as $N = pq$.
2) We suppose that $N = pq$ with p and q natural numbers such as $p > q$.
   a) Which is the parity of p et q?
   b) Determine, depending on a and b, two natural numbers p and q such as $N = pq$.
   c) Give all the factorizations of 45 as products of natural numbers, **and** as differences of two squares of natural numbers.
   d) Formulate into one logical equivalence everything that was demonstrated in 1) and 2) b.
3) Reading Fermat's text.

Notice that Fermat uses the following definitions: the *parts* of a number are its *divisors*. And a number is *composed (*product) of its *parts*. For instance, if 45 = 9 * 5, 9 and 5 are the *parts* of 45 and 45 is composed of 9 and 5.

Fragments of a letter by Fermat, <1643>
Translation by Christian Aebi & John Steinig
Every non-square odd number is […] the difference of two squares as many times as it is composed of two numbers, […]
It is quite easy to find the adequate squares, when we are given the number and its parts, and to have its parts when we are given the squares.

Explain the link between these sentences and the questions 1) and 2).

Comments on Student Responses to Part I:
The second question posed difficulties for the students. Finding *a* and *b* in relation to *p* and *q* is difficult: some students did not see that we have a system to solve (with unknown *a* and *b*). Others saw the system, but were not able to solve it.

Another difficulty: we are looking for two adequate integers (i.e. integers *p* and *q* which are solutions for the problem) but not **all** adequate integers. Thus, it is sufficient to take $a = \frac{p+q}{2}$ and $b = \frac{p-q}{2}$, but it is not necessary. This provides an opportunity to work on logic, which is also part of the curriculum.

None of the students (even the best) managed to write the logical equivalence in 2d. We discussed this in class and, finally, all the students actually saw that this is what Fermat asserts in his letter.

After correction of Part I, we worked on the questions in Part II. We answered question 1 together without any difficulty, and students then worked in groups for the other questions.

Part II. The factorization algorithm:
In this Part II, *N* is a odd non- square natural number.

1) Explain why determining a factorization of *N* reduces to determining an integer *a* such that N-a^2 is the square of a nonzero natural number *b*.
2) Write an algorithm that determines an integer *a* such that *N*-a^2 is the square of an integer *b*, with *N* being an input from the user and *a* and *b* given as outputs.
3) Does the algorithm end for every odd non-square natural number *N* ?
4) What is the output of the algorithm if *N* is a prime number?
5) In his letter Fermat uses his method in order to factorize *N* = 2,027,651,281.
    a) Do the same by running the algorithm "by hand". You can use a table or your calculator if you wish.
    b) How many steps are required?

c) The algorithm requires the computation of a square at each step of the conditional loop. By using the expansion of $(a + 1)^2$, modify the algorithm in order to that the test requires only a first grade computation.

6) When running the algorithm, you have to test whether some numbers are the squares of primes. Fermat asserts in his letter: "the *remainder is 49619, which is not a square, because no square ends with 19*"

a) Is it possible that a square ends with 7? Justify your answer.

b) How would you justify the assertion of Fermat?

Comments on Student Responses to Part II:

1) Writing the algorithm:

One group immediately wrote an appropriate algorithm, initializing the variable *a* appropriately. Three groups started pretty quickly, after the teacher (Martine Bühler) told them to think how they would deal with 45 "by hand". Two groups found it difficult to start.

Of the latter five groups: one group tried a conditional testing through an IF/THEN/ELSE structure to be sure that a^2-N is greater than 0, but the group did not succeed in this way. It is easier to initialize *a* correctly. One group initialized *a* at 1, and did not see any problem. One group initialized *a* at 1, but became aware of the problem when they tried question 5a and then made the correction. One group had great difficulties to write the WHILE loop. One group used another variable *b*, equal to a^2-N, initialized *b*, but did not change the value of *b* in the WHILE loop.

2) Other observations

- Question 3: some groups immediately saw that the algorithm always stops, as $N = 1 \times N$, hence *N* is always the difference of two squares with $a=\frac{N+1}{2}$ and $b=\frac{N-1}{2}$. So, even if *N* is a prime number, the algorithm stops and then the output is $a=\frac{N+1}{2}$and $b=\frac{N-1}{2}$. The converse was studied at the following class session, but I gave the result. As we ran out of time, I asked the students to answer questions 5a and 5b at home and to go directly to questions 5c and 6.
- Question 5c: We can optimize the initial algorithm; to do so, we must calculate a square at each step. But, having calculated a^2-N, it is easier to calculate $(a+1)^2-N = a^2-N + 2a + 1$.Thus, it is sufficient to add $2a + 1$ to the previous number and it is not necessary to calculate $(a+1)^2$. We need an additional variable, but the calculation is simpler. Some groups had no problem in revising the algorithm after I told them that they need another variable, but others could not see where to put the loop and we went over this together.

- Question 6: Students had no difficulty with this. It's a traditional question using congruences and "disjunctive cases". Students are used to make congruences tables. To justify Fermat's assertion, you may work modulo 100.

Optimized algorithm
Input: *N* odd integer, not a square
Procedure: *a* changes to $\left[\left[\sqrt{N}\right]\right]+1$
c changes to a^2- N
While $\sqrt{c}$ is not an integer, do:
c changes to $c + 2a + 1$
a changes to $a + 1$
WhileEnd
Return: *a* and $\sqrt{c}$

We remarked that, at each step, we add a term of an arithmetic sequence with 2 as common difference.

Thus, with this problem of factorization, we can deal with the problem of time complexity. The efficiency of an algorithm can be measured by time complexity and by space complexity. Run time analyses is a classification that estimates the increase in running time of an algorithm as its input increases. This is a topic of great interest in computer theory.

At the end of the session, I gave students Fermat's text (Appendix 1: we read it at the following session in class) and the second part of the problem (see below), as homework. The aim of the problem is the factorization of 250,507, using a sieving method, and the study of Carissan's device.

Part III. Factorization of large numbers and Carissan's device.

The aim of this part is the factorization of N =250,507. *This is done by determining an integer x such that* $x^2 - N$ *is the square of an integer.*

1°) Working modulo 7.

a) Complete the following table with the remainder of X^2 modulo 7 depending on the values taken by X modulo 7.

X mod7	0	1	2	3	4	5	6
X^2 mod7							

Is it possible for the number 7 x 113 + 3 to be a square? Why? (It is imperative to use the previous table, but **not** the calculator; the numbers 0, 1, 2, 4 are called quadratic residues modulo 7).

b) Determine the remainder of the Euclidean division of 250,507 by 7.

c) We are looking for an integer x such as $x^2 - N$ is the square of an integer. Thus, if

the integer x is suitable, then the number $x^2 - 250{,}507$ must be a square. Using the precedent table, determine the possible values of $x^2 - 250{,}507$ modulo 7. Deduce the possible values for x^2 modulo 7.

d) But x^2 must be a square; thus the same table allows us to restrict again the possible values of x^2 modulo 7. Do this, then give the possible values for x modulo 7. Could the number 778 be a solution of the given problem?

2°) Do the same work modulo 9.

3°) Do the same work modulo 15.

4°) Determination of an integer x such that $x^2 - 250{,}507$ is a square.

a) Justify the assertion: if x is a solution of the problem, then, $x^2 \geq 250{,}507$. What is the smallest possible value for x ?

b) Let $x_0 = 501$. Calculate the remainders of x_0 modulo 7, modulo 9 and modulo 15. Is the number x_0 a solution of the problem?

c) Complete the following table until you find a value of x which fits the conditions found in the questions 1°), 2°), 3°).

x	501	502	503	504	...
Mod 7					
Mod 9					
Mod 15					

Can we be sure that the value found with this method is a solution of the given problem?

Verify that this value is actually a solution and deduce a factorization of 250,507.

Comments on Student Responses to Part III: We started with the end of Fermat's text.

$N = 2{,}027{,}651{,}281.$

$\left[\left[\sqrt{N}\right]\right] = A = 45029$

$N = A^2 + R$ with $R = 40{,}440$.

We have to calculate $(A + 1)^2 - N$ to see whether it is a square or not.

$(A + 1)^2 - N = 2A + 1 - R$.

That is: We subtract the remainder R = 40,440 from the double plus 1 of the square root of N (the translation in the appendix 1 is not quite correct there).

$(A + 1)^2 - N = 49{,}619$ (this number is not a square).

We add 90,061= 90,059 + 2. Students understood this because they knew that, at each step, we add a term of an arithmetical sequence of common difference 2 (as had been seen in the part II of the exercise).

Another question that naturally arises about the algorithm concerns the number of steps (namely the question of time complexity).

Set $N = pq = a^2 - b^2$, with $a = \frac{p+q}{2}$ and $b = \frac{p-q}{2}$. One starts with a = $\left[\left[\sqrt{N}\right]\right]+1$

The number of steps is

$$\frac{p+q}{2} - \left(\left[\left[\sqrt{N}\right]\right]+1\right) \approx \frac{p+q-2\sqrt{pq}}{2} \approx \frac{\left(\sqrt{p}-\sqrt{q}\right)^2}{2}$$

This number is even smaller than the difference of the divisors *p* and *q*. So the method gives the divisors which are the closest to the square root of *N*. We can thus deduce that, if we obtain, $a = \frac{N+1}{2}$(that is, $a = \frac{p+q}{2}$ with $p = N$ and $q = 1$), then we are sure that *N* is a prime number, because, if not, we would have obtained a divisor *p* of *N* closer to the square root of *N*.

This method can be used as a primality test, but it is a bad primality test. If *N* is prime, the number of steps is about *N*/2, far more than with the elementary method of trying of odd integers between 3 and *N*.

But, in the case studied by Fermat, the example is well selected, because we need only 11 steps. It is easy to create such examples, by choosing two prime numbers close to one another and calculating their product.

Principle of Carissan's Device
The method is a sieving method.

X mod7	0	1	2	3	4	5	6
X^2 mod7	0	1	4	2	2	4	1

$N \equiv 5(mod7)$
If $x^2 - N$ is a square, then $x^2 - N$ is equivalent to 0 or 1 or 2 or 4 modulo 7. namely, $x^2 \equiv 5$ or 6 or 0 or 2 mod 7
These are the possible values modulo 7. We can do the same work modulo 9 and 15, and we will obtain the possible values mod 9 and mod 15.
It remains to complete the last table, and to stop when the **three** value are possible values for *x*.
This is useful because we can then mechanize the algorithm.
The choice of the moduli is a pedagogical choice. I wanted three moduli, because it is necessary to understand the method to have at least 3 moduli (2 are not enough). I also

did not want the preparatory work to be too long. I thus choose numbers that are not too large.

In the last classroom session, I showed the beginning of a film explaining the historical context of the Carissan's device. The film is an amateur's one, lasting about 15 minutes. You can see the film on the IREM site:
http://www.irem.univ-paris-diderot.fr/videos/la_machine_des_freres_carissan/
We quickly corrected the problem, rather successfully completed by the students. I showed them a Carissan's device with 3 disks using an overhead projector[2]. We then watched the last 10 minutes of the film, showing the Carissan's device at work.

SOME OTHER EXAMPLES OF HISTORICAL ALGORITHMS WHICH CAN BE CARRIED OUT IN THE CLASSROOMS

Algorithms for geometrical constructions based on Euclid Elements

You will find some very basic geometrical constructions, in Euclid *Elements* which are clearly algorithms. Each of them is a step by step procedure, with a set of rules. The problem is enounced in natural language, then the data are named, and the procedure you apply. You have only to be sure that you know (or the machine knows) the elementary tools. If not, you have to insert a subprogramme. We have chosen some propositions from Euclid Elements as illustrations.

Proposition 1, Book 1

To construct an equilateral triangle on a given finite straight line.

Let *AB* be the given finite straight line. It is required to construct an equilateral triangle on the straight line *AB.*

Construction:

Describe the circle *BCD* with centre *A* and radius *AB.* (I post 3)

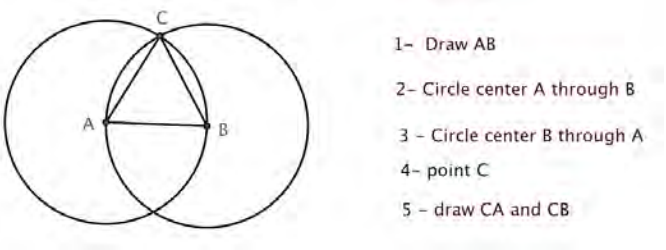
Again describe the circle *ACE* with centre *B* and radius *BA (*I post 3*).*

Join the straight lines *CA* and *CB* from the point *C* at which the circles cut one another to the points *A* and *B. (*I post 1)

Conclusion: Therefore the triangle *ABC* (I, Def 20) is equilateral, and it has been constructed on the given finite straight line *AB.* Q.E.F.

Then, you will find, of course a demonstration. But this is apart from the algorithmic construction. The tools you need are pointed as, for instance (I post 3) or (I, Def 20), that means: Book 1 postulate 3 or definition 20.

[2] See the website of the irem Paris (groupe MATH).



Proposition 11, Book 1

To draw a straight line at right angles to a given straight line from a given point on it.

Let *AB* be the given straight line, and *C* the given point on it.

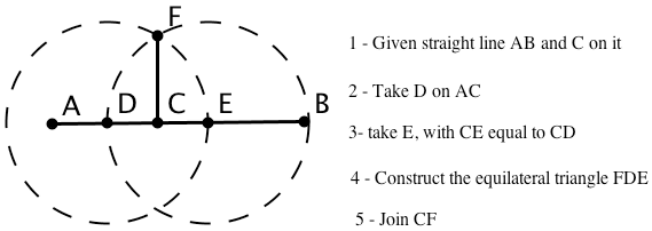
It is required to draw a straight line at right angles to the straight line *AB* from the point *C*.

Construction:

Take an arbitrary point *D* on *AC*. Make *CE* equal to *CD*. (I 3). Construct the equilateral triangle *FDE* on *DE*, and join *CF* (I 1). (I post 1)

I say that the straight line *CF* has been drawn at right angles to the given straight line *AB* from *C* the given point on it.

Conclusion: Therefore the straight line *CF* has been drawn at right angles to the given straight line *AB* from the given point *C* on it. Q. E. F.



Proposition 9, Book 1

To bisect a given rectilinear angle.

Let the angle *BAC* be the given rectilinear angle.

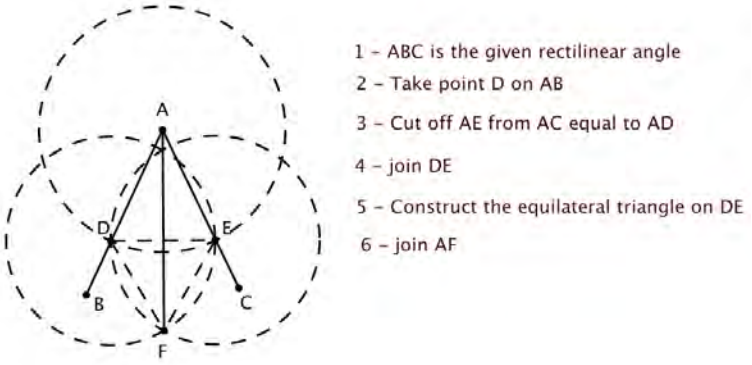
It is required to bisect it.

Construction

Take an arbitrary point *D* on *AB*. Cut off *AE* from *AC* equal to *AD*,(I 3) and join *DE*.(I post 1*)* Construct the equilateral triangle *DEF* on *DE* (I 1*)*, and join *AF*.

I say that the angle *BAC* is bisected by the straight line *AF*.

Conclusion: Therefore the given rectilinear angle *BAC* is bisected by the straight line *AF*.



You can find here the useful definitions and postulates for these three algorithmic constructions:

> I def 20: Of trilateral figures, an *equilateral triangle* is that which has its three sides equal, an *isosceles triangle* that which has two of its sides alone equal, and a *scalene triangle* that which has its three sides unequal.
> I post 1: To draw a straight line from any point to any point.
> I post 3: To describe a circle with any centre and radius.
> I prop 3: To cut off from the greater of two given unequal straight lines a straight line equal to the less.
> I prop 8: If two triangles have the two sides equal to two sides respectively, and also have the base equal to the base, then they also have the angles equal which are contained by the equal straight lines.

HERON OF ALEXANDRIA: FORMULA VERSUS ALGORITHM

Heron of Alexandria was an important geometer and worker in mechanics. A major difficulty regarding Heron was to establish the date at which he lived. From Heron's writings it is reasonable to deduce that he taught at the Museum in Alexandria. His works look like lecture notes from courses he must have given there on mathematics, physics, pneumatics, and mechanics. Some are clearly textbooks while others are perhaps drafts of lecture notes not yet worked into final form for a student textbook.

We propose here some excerpts from *Metrica* (ca 50 A. D.)

About area of triangles:

Book I of his treatise *Metrica* deals with areas of triangles, quadrilaterals, regular polygons of between 3 and 12 sides, surfaces of cones, cylinders, prisms, pyramids, spheres etc. Usually you will find Heron's formula about the area A of a triangle whose sides of length are a, b, c, and the half-perimeter $p=\frac{(a+b+c)}{2}$ is given as follows:

$$A = \sqrt{p(p-a)(p-b)(p-c)}$$

Actually in Heron's *Metrica* you find the result formulated as an algorithm on a generic example:

Heron's *Metrica*	Modern algebra
Let the sides of the triangle be of 7, 8, 9 units.	$a = 7, b = 8, c = 9$
Compose [add] the 7 and the 8 and the 9: the result is 24;	$a+b+c = 24$
from this take the half: the result is 12;	$(a+b+c)/2 = 12 \quad p = 12$
subtract the 7: 5 remaining.	$(a+b+c)/2 - a = 5 \quad p - a = 5$
Again from the 12, substract the 8: 4 remaining;	$(a+b+c)/2 - b = 4 \quad p - b = 4$
and again the 9: 3 remaining.	$(a+b+c)/2 - c = 3 \quad p - c = 3$
Make the 12 by the 5: the result is 60;	$p(p-a) = 60$
these by the 4: the result is 240;	$p(p-a)(p-b) = 240$
these by the 3: the result is 720;	$p(p-a)(p-b)(p-c) = 720$
from these take a side and it will be the area of the triangle.	$\sqrt{p(p-a)(p-b)(p-c)} = A$

Then, Heron gives a method in natural language, on a generic example, to approximate the square root

Heron gives this in the following form:

> Since 720 has not its side rational, we can obtain its side within a very small difference as follows. Since the next succeeding square number is 729, which has 27 for its side, divide 720 by 27. This gives $26\ ^{2}/_{3}$. Add 27 to this, making $53\ ^{2}/_{3}$, and take half this or $26\ ^{5}/_{6}$. The side of 720 will therefore be very nearly $26\ ^{5}/_{6}$. In fact, if we multiply $26\ ^{5}/_{6}$ by itself, the product is $720\ ^{1}/_{36}$, so the difference in the square is $^{1}/_{36}$. If we desire to make the difference smaller still than $^{1}/_{36}$, we shall take $720\ ^{1}/_{36}$ instead of 729 (or rather we should take $26\ ^{5}/_{6}$ instead of 27), and by proceeding in the same way we shall find the resulting difference much less than $^{1}/_{36}$.

So you can notice:

The explanation on a generic example
The algorithm is iterative
The explanation is given in natural language
The quantities are expressed with fractions (epistemic context)
The basic idea for the process is the notion of arithmetical mean
The algorithm gives the same results as Newton's method[3]

[3] See an example on the website of the APMEP (French Association of Teachers of Mathematics), by Martine Buhler

ABOUT HORNER'S ALGORITHM

In mathematics, the algorithm known as Horner's method is described in many textbooks. Horner's method is an economical way of evaluating a polynomial for a given value of the argument. It is efficient, too, for polynomial division, polynomial root finding, and very fast for derivatives evaluation.

William George Horner (1786?-1837) was a school master who ran his own school at Bath, from 1809 until his death . In 1819, he published a paper on the numerical solution of equations: *A new method of solving numerical equations of all orders, by continuous approximation*.

In this paper (philosophical transactions, 1819), he gave a tabular scheme for computing a real root of a polynomial equation, but it was substantially different from the basic algorithm so called now Horner's method. The paper was also written in a very obscure style.

Anyway, Julian Coolidge (1949) wrote:

> A great mathematician he certainly was not. He offers a fine example of what an amateur can accomplish by dogged industry, and his method is surely the best we have for solving numerical equations.

In fact, the basic algorithm conventionally called Horner's method was first given in England by Theophilus Holdred (a londonian clock maker) in 1820, and was not described by Horner until 1830 (published in 1845, post mortem). The method had been anticipated by Paolo Ruffini(1765-1822) in Italy(1804), and François-Désiré Budan de Boislaurent (1761-1840) in France (1807). And long before, related techniques were known to chinese and arabic mathematicians.

Asking why this method is primarily known as Horner's method, we note that the popularization of Horner's process was due to Augustus de Morgan (1806-1871), a prominent figure in 19th century English mathematics.

Description of the Algorithm

The so called basic method, of this algorithm, as explained by Thomas Stephen Davies, in *The mathematician*, 1845 is**:**

Given the polynomial:

$$A_n x^n + A_{n-1} x^{n-1} + A_{n-2} x^{n-2} + etc. + A_2 x^2 + A_1 x + A_0$$

We wish to evaluate it at a specific numeral value of x, say x_0.

To accomplish this, we define a new sequence of constants as follows:

$$B_n = A_n$$

$$B_{n-1} = A_{n-1} + B_n x_0$$

...

$$B_0 = A_0 + B_1 x_0$$

And B_0 is the value searched.

To see why this works, note that the polynomial can be written as follows:
$A_0 + x(A_1 + x(A_2 + etc. + x(A_{n-1} + A_n x)))$
Thus, by iteratively substituting, you obtain:
$A_0 + x_0(A_1 + x_0(A_2 + etc. + x_0(A_{n-1} + A_n x_0)))$
$= A_0 + x_0(A_1 + x_0(A_2 + etc. + x_0 B_{n-1}))$
$etc.$
$= A_0 + x_0 B_1$
$= B_0$

Examples

First: give the value of $x^4 + 2x^3 - 22x^2 + 7x + 42$, for x = 3.

We use the so called ***synthetic division***, as follows

	1	**2**	**- 22**	**7**	**42**
3		**3**	**15**	**- 21**	**- 42**
	1	**5**	**- 7**	**- 14**	**0**
	B_4	B_3	B_2	B_1	B_0

The entries in the third row are the sum of those in the first two. Each entry in the second row is the product of the *x*-value (3 in this example) with the third-row entry immediately to the left. The entries in the first row are the coefficients of the polynomial to be evaluated.
The value for 3 is 0.
And the remainder in the division by (x – 3) is 0.
3 is a root.
And you can write: $x^4 + 2x^3 - 22x^2 + 7x + 42 = (x-3)(x^3 + 5x^2 - 7x - 14)$
Second example:
Divide $x^3 - 6x^2 + 11x - 7$ by $(x - 2)$

	1	**- 6**	**11**	**- 7**
2		**2**	**-8**	**6**
	1	**- 4**	**3**	**- 1**
	B_3	B_2	B_1	B_0

The quotient is $x^2 - 4x + 3$, and the remainder is -1, so that:
$x^3 - 6x^2 + 11x - 7 = (x^2 - 4x + 3)(x - 2) - 1$
Horner's method is optimal, in the sense that any algorithm to evaluate an arbitrary polynomial must use at least as many operations. Alexander Ostrowski proved in 1954 that the number of additions required is minimal. Victor Pan proved in 1966 that the number of multiplications is minimal. In fact this method is very efficient, so that it is used for computers and calculators.

Use for Derivatives

In their first publications, Ruffini and Horner used the differential calculus in expounding their methods. Later, both authors gave simplified explanations, using ordinary algebra, as it is shown above.
But it can be used to evaluate derivatives.
Given a polynomial $P(x)$, and a real number x_0.
Using the preceding method, you write $P(x)$ as: $P(x) = (x\text{-}x_0)Q(x) + P(x_0)$**,** where $Q(x)$ is a new polynomial.
Now, you can repeat with $Q(x)$, and obtain: $Q(x) = (x\text{-}x_0)Q_1(x) + Q(x_0)$**.**
That gives: $P(x) = P(x_0) + (x\text{-}x_0)Q(x_0) + (x\text{-}x_0)^2 Q_1(x)$**.**
Then you'll obtain:
$P(x) = P(x) = P(x_0) + (x\text{-}x_0)Q(x_0) + (x\text{-}x_0)^2 Q_1(x_0) + (x - x_0)^3 Q_2(x)$.
And so on.
Actually the Taylor's theorem, gives
$P'(x_0) = Q(x_0)$; $P''(x_0) = 2!\, Q_1(x_0)$; … ; $P^{(k)} = k!\, Q_{k-1}(x_0)$

Example:
Given: $P(x) = x^5 - 6x^4 + 8x^3 + 8x^2 + 4x - 40$
We wish to evaluate $P(3)$, then $P'(3)$, $P''(3)$.

	1	- 6	8	8	4	- 40	
3		3	- 9	- 3	15	57	
	1	- 3	- 1	5	19	17	P(3) = 17
3		3	0	- 3	6		
	1	0	- 1	2	25	← Q(3)	P'(3) = 25
3		3	9	24			
	1	3	8	26	← $Q_1(3)$	P"(3) = 2! 26	P"(3) = 52
3		3	18	78			
	1	6	26	← $Q_2(3)$	$P^{(3)}(3) = 3!\, 26$	$P^{(3)}(3) = 156$	

Appendix 1

Translation of *Fragment d'une lettre de Fermat*, *Œuvres*, éd. Tannery et Henry, tome II, 1894, pp.256-258. There are a few annotations between square brackets to help the reader. Translation by Christian Aebi and John Steinig.

Site: http://gradelle.educanet2.ch/christian.aebi/.ws_gen/9/Fermat_english.pdf

Fragments of a letter by Fermat <1643>

Every non-square odd number is […] the difference of two squares as many times as it is composed [*written as a product*] of two numbers, and if the squares are prime to one another then the same may be said of the two composition numbers [*factors*]. But if the squares have a common divisor, the number in question will also be divisible by the same common divisor, and the composition numbers will be divisible by the side [*square root*] of the common divisor.

For example: 45 is composed of 5 and of 9, of 3 and 15, of 1 and 45. So, it will be thrice the difference of two squares: according of 4 and 49, who are prime to one another, as are the corresponding composers 5 and 9; plus of 36 and 81, which have 9 as common divisor, and the corresponding composers, 3 and 15, have the side of 9, meaning 3, as common divisor; finally 45 is the difference of 484 and of 529, which have 1 and 45 as corresponding composers.

It is quite easy to find the adequate squares, when we are given the number and its parts[*divisors*], and to have its parts when we are given the squares.

[…]

That settled, let a number be given to me, for example 2 027 651 281; we are asked if it is prime or composed, and in that case, of which composers.

I extract the square root to find the smallest of the preceding numbers, and find 45 029 with 40 440 as remainder, from which I substract the double plus 1 from the preceding root, meaning 90 059: the remainder is 49 619, which is not a square, because no square ends with 19, from there I add 90 061 to it, meaning 2 more than 90 059 which is the double plus 1 of the root 45 029. And since the sum 139 680 is still not a square, as can be seen by the ending [*final digits*], I add to it once again the same number increased by 2, meaning 90 063, and I continue to add in that manner until the sum is a square, as can be seen here. This happens only at 1 040 400, which is the square of 1 020, and thus the given number is composed; because it is easy, by the examination of the preceding sums, to see that there is no other that is square except the last, because squares cannot bear the ending they have, except for 499 944 which nevertheless is not a square.

Now, to find out the numbers that compose 2 027 651 281, I remove the number that I had first added, meaning 90 061, from the last added 90 081. There remains 20, from which half plus 2, meaning from 12, I add the root previously found, 45 029. The sum is 45 041, to which number by adding and removing 1020(the root of the last sum 1 040 400), we have 46 061 and 44 021, which are the two nearest [*side-by-side*] numbers that compose 2 027 651 281. These are also the only ones [*factors*], for they are, one as well the other, prime.

REFERENCES

Aebi, C & Steining J. From elementary algebraic identities to Fermat's factorization
http://gradelle.educanet2.ch/christian.aebi/.ws_gen/9/Fermat_english.pdf

Budan de Boislaurent, F.D. (1807). *Nouvelle méthode de resolution des équations numériques d'un degré quelconque d'après laquelle tout le calcul exigé pour cette resolution se réduit à l'emploi des deux premières règles de l'arithmétique*, Courcier, Paris.

Coolidge, J.L. (1949). *The mathematics of great Amateurs*, Oxford Science Publications, 186-194.

Horner, W.G. (1819). A new method of solving numerical equations of all orders, by continuous approximation, *The philosophical transactions of the Royal Society of London*, july 1819, part II, 308-335.

Joyce, D.E. (1996-1997). *Euclid's Elements*.
http://aleph0.clarku.edu/~djoyce/java/elements/elements.html

Ruffini, P. (1804). *Sopra la determinazione delle radici nelle equazioni numeriche di qualunque grado*. Societa typografica, Modena.