HISTORICAL ACCOUNT AND ULTRA-SIMPLE PROOFS OF DESCARTES’S RULE OF SIGNS, DE GUA, FOURIER AND BUDAN’S RULES

by Michael Bensimhoun
22th May 2014

ABSTRACT

It may seem a funny notion to write about theorems as old and rehashed as Descartes’s rule of signs, De Gua’s rule or Budan’s. Admittedly, these theorems were proved numerous times over the centuries. However, despite the popularity of these results, it seems that no thorough and up-to-date historical account of their proofs has ever been given, nor has an effort been made to reformulate the oldest demonstrations in modern terms. The motivation of this paper is to put these strongly related theorems back in their historical perspective. More importantly, we suggest a way to understand Descartes’s original statement, which yet remains somewhat of an enigma. We found that this question is related to a certain way of counting the alternations and permanences of signs of the polynomial coefficients, and may have been the convention used by Descartes. Remarkably, this convention not only provides a ultra-simple proof of Descartes’s rule, but it can also be used to simplify the proofs of the titular theorems. Without claiming to be exhaustive, we shall present in this paper an historical account of these theorems and their proofs, and clarify their mutual relation. We will explain how a suitable convention can help understand the original statement of Descartes and greatly simplify its proof, as well as the proofs of the above-mentioned theorems. With the exception of the proof of Budan’s rule, which runs on rudiments of infinitesimal calculus (Taylor’s theorem), the proposed demonstrations are so short and elementary they could be taught at the undergraduate level.

1. Historical Perspective

Of the theorems listed in the title of this paper, the oldest and by far most famous theorem is Descartes’s rule of sign. It was first formulated by Descartes in 1637 in his Geometry ([5]). At the beginning of his exposition, the author gives numerical examples of products of polynomials by \(X - \alpha\). Next, having claimed that a polynomial \(P\) has a root \(\alpha\) if and only if it is divisible by \(X - \alpha\), he added without proof: “As a result, it is possible to know how many true roots\(^{(1)}\) and false

\(^{(1)}\) positive roots
roots\(^{(2)}\) an equation can have\(^{(3)}\). Namely, it can have as many true roots as the signs \(+\) and \(-\)\(^{(4)}\) alternate, & as many false roots as two signs \(+\) or two signs \(-\) follow one another." \(^{(5)}\)

This statement of Descartes’s was attacked by several of his contemporaries, who pointed out that a real polynomial can have fewer positive roots than the number of alternations of signs contained in its coefficients. This counterargument is rather surprising, for it is clear from the words of Descartes that he meant a polynomial has at most as many positive roots as it contains alternations of signs (actually, this is exactly what Descartes objected in his 77th refutation letter, directed against Roberval (\([6]\))). On the other hand, Wallis, another contemporary of Descartes, zealously tried to attribute this rule to his compatriot Harriot, an English geometer postdating Viète and predating Descartes. However, as pointed out by De Gua in his mémoires (\([11]\)), Wallis’s thesis can not be seriously argued (the arguments of De Gua are well documented and extremely convincing).

Undoubtedly, there is a logical flaw in Descartes’s original statement. It seemingly contains two assertions, that should be understood in the following manner:

1. A real polynomial has no more positive roots than alternations of signs between two consecutive coefficients.
2. A real polynomial has no more negative roots than permanences of signs between two consecutive coefficients.

The following question arises: what is the meaning of the terms consecutive coefficients, or in the words of Descartes, “sign that follows one another”. Of course, this is clear if the polynomial is not lacunary, but consider for example the polynomial \(P(X) = +X^2 - 1\), lacunary in \(X^1\). If the term “consecutive” is to be understood in the most obvious manner, then \(P\) has exactly one alternation of signs, and no permanence of signs. Therefore, according to assertion (2) above, \(P\) should have no negative roots. This is obviously false since \(P(X) = (X - 1)(X + 1)\). Could so trivial a counter-example have escaped Descartes? Of course, one can argue that, unlike Fermat, Descartes was not always fastidius regarding the statement of his theorems; he may, after all, have made a mistake. Nevertheless, he was an excellent geometer and algebraist, whose genius profoundly influenced mathematical thought. Could the mathematician whom even Fermat held in great esteem have missed so simple a point? As we shall show below, this may not be the case. We realized that Descartes’s statement is in fact correct if one assigns, in any manner one pleases, signs to the lacunary coefficients. For example, in the aforementioned

\(^{(2)}\) negative roots
\(^{(3)}\) polynomial
\(^{(4)}\) the same as “−”
\(^{(5)}\) free translation
case, we can write symbolically \( P(X) = +X^2 + 0 \cdot X - 1 \); we see that \( P \) has one alternation and one permanence of signs, which corresponds, indeed, to its unique positive and negative root resp. The same would be true, in this case, if we had written \( P(X) = +X^2 - 0 \cdot X - 1 \), but generally, the number of alternations and of permanences of signs is not the same if one assigns signs to the lacunary coefficients in different ways. Thus, in order to find the strictest limits of the number of positive and negative roots, and to apply the rule to its greatest extent, it is natural and judicious to attach signs to the lacunary coefficients in such a way that the number of alternations or of permanences of signs is minimized. This can be performed simply by subscribing to a suitable convention. Furthermore, this has the effect of making the proof of Descartes's rule quite simple, and may reasonably be the basis of the proof Descartes had in mind. This convention will be made explicit in the next section, and a ultra-simple proof will be derived. For the moment, however, we continue the historical account.

As noted above, the meaning of an alternation or a permanence of sign is evident when the polynomial is not lacunary. The situation is more complex when it is. Commonly, lacunary coefficients are ignored during the process of counting the alternations and permanences of signs of \( P \). This is the way Descartes's rule was understood by most of the mathematicians following Descartes, and it is sufficient to assert and prove part (1) of his rule. In fact, we shall see below that this way of counting the alternations of signs coincides with our minimization convention above, as far as alternations of signs are concerned; but the two methods do not coincide when applied to the counting of the permanence of signs. This may explain why part (2) of Descartes's rule was generally ignored by mathematicians, likely regarded as a false statement.

It is common to find versions of Descartes’s rule that agglomerate the genuine rule with other propositions. The first extension, commonly found in old books, is that if all the roots of a polynomial are real, then the number of positive roots is exactly equal to the number of its alternations of signs. Curiously, this proposition is seldom found in modern resources such as Wikipedia, although it could give a useful external criterion to determine the signature of a real quadratic form from its characteristic polynomial, or, in other words, to classify conics. It is difficult to attribute this extension to a specific author, since it was discussed already in Descartes’s time as the way some of his contemporaries understood his rule.

The second extension consists of the fact that the parities of the number of positive roots and the number of alternation of signs are always equal. This proposition, which several authors have used to prove the rule, is in fact of a very different nature. Indeed, Descartes’s rule generalizes to any function \( f \) for which the \( n \)th derivative is of constant sign, as will be shown below, while the extension above
arises from the fact that irreducible polynomials over \( \mathbb{R} \) are of a degree no higher than 2. We believe that this accidental fact should be deinterlaced from Descartes’s rule, at least regarding its proof. We were unable to determine with certainty which author was the first author to use this extension, but we did find it appears in the 1820 dissertation of Fourier ([8]). More precisely, it was part of Fourier’s statement of the Budan-Fourier theorem in the aforementioned dissertation; since the budan-Fourier theorem is itself a generalization of Descartes’s rule, we shall attribute this extension to Fourier in this paper. Such an attribution should be confirmed, of course, by further research (we were informed that this rule may have been stated by the Chinese mathematician Li Rui in the year 1813).

Several geometers following Descartes tried to prove his theorem. It is not easy to determine who was the first to give a valid proof; in our opinion, the difficulty in this matter is not in vaguely showing that the rule is correct, but in providing a rigorous demonstration. It is also noteworthy that Descartes himself provided a hint on how to prove his rule: Indeed, the words “As a result” just following his remark on the divisibility of a polynomial \( P \) by \( X - \alpha \) (\( \alpha \) root of \( P \)) suggests strongly that he was in possession of an inductive proof, showing that multiplying a polynomial by \( X - \alpha \), with \( \alpha > 0 \), increases the number of its alternations of signs by one unity, while multiplying it by \( X + \alpha \) increases the number of its permanences of signs by one unity.

Roughly speaking, there are two categories of proofs of Descartes’s rule: Proofs belonging to the first category are called algebraic and use neither geometry nor infinitesimal calculus, but only algebra and obvious properties of \( \mathbb{R} \). Proofs belonging to the second category are called analytic, and use geometric curves and extrema to obtain information, or derivatives of the first and higher orders to build inductive arguments. Often, they can be generalized to classes of functions beyond polynomials, which presents an advantage over algebraic proofs.

It is generally acknowledged that the first rigorous demonstration of Descartes’s rule was published by De Gua in 1742 ([10]). In fact, De Gua gave two demonstrations: an algebraic one and an analytic one. His algebraic proof contends that all the roots of the polynomial are real. It is based on a curious lemma: If a polynomial \( P \) contains only real roots, then its coefficients \( a_i \) fulfill \( a_i^2 - a_{i-1}a_{i+1} > 0 \). Following the lines sketched by Descartes, De Gua showed that multiplying a polynomial \( P(X) \) by \( X - \alpha \), where \( \alpha > 0 \), increases the number of its alternations of signs. The second proof of De Gua is analytic, and no assumption is made about the roots of the equation. It is an application of the geometry of curves to algebra, in the spirit of certain recent proofs (see e.g. [18]). But De Gua’s contribution to algebra is not limited to these proofs. In another mémoire ([11]), he gives an extended and precise historical analysis of the theory of equations up to his time, and investigates
means to evaluate the number of complex roots of real polynomials. The so-called *De Gua’s rule* was extracted from this work, a kind of lower-bound on the number of complex roots of real polynomials (see Sec. 2.3 below).

Another algebraic proof of Descartes’s rule was found by Segner in 1756 ([16]). Like the proof of De Gua, it aims to show that multiplying a polynomial $P(X)$ by $X - \alpha$, where $\alpha > 0$, increases the number of its alternations of signs. It was also explained in the Encyclopaedia Britannica in 1824 ([17]). According to this entry, Segner’s proof would be “not only the most simple, but probably the most simple that will ever be invented.” In our opinion, Segner’s proof may be convincing, but it can not be put in rigorous terms without increasing the number of definitions and notations, and is certainly not the simplest possible proof produced since. Moreover, it is not entirely clear how Segner dealt with the signs of the null coefficients in lacunary polynomials.

In the years following De Gua and Segner, several authors presented other demonstrations. For example, a laborious and long analytic proof was given by ÁEpinus in 1758 ([15]), and an algebraic one was given by Lagrange in 1808 ([14]), following the lines of his theory of equations. In spirit, the proof of Lagrange is close to being an analytic proof, showing that analytic arguments can always be turned into algebraic ones, as far as polynomials are concerned; the distinction between algebraic and analytic proofs is more philosophical than mathematical. On the same note, Lagrange mentions a proof of Kæstner, which we were unable to find.

Between the years 1796 and 1822 Descartes’s rule took an unexpected course. Both Budan and Fourier found that it can be generalized in such a way that it provides an upper bound of the number of positive roots of a polynomial, *between* any two given bounds $a$ and $b$. In the literature, this theorem is referred to as Budan’s rule, Fourier’s theorem, or the Budan-Fourier theorem. Actually, Budan formulated his theorem in terms of polynomial coefficients only, while Fourier used derivatives of the first and higher orders. These two formulations are, of course, equivalent, according to Taylor’s Theorem. An ultimate and efficient algorithm, providing the exact number of positive roots between two bounds, was finally found by Sturm. It was strongly inspired by the work of Fourier, as its author acknowledged. Sturm’s algorithm efficiently solves the problem that has preoccupied mathematicians since Descartes: locating the roots of a polynomial, in order to compute them using known approximation algorithms. An excellent account of the origin and influence of Sturm’s algorithm in mathematics, as well as its deep implications as found by Tarski, is given in [12]. Although Sturm’s algorithm completely eclipsed it for centuries, the Budan-Fourier theorem has seen a renewal of interest in the last decade, since it lies at the heart of (through the Vincent algorithm) the most powerful method of polynomial root extraction: the VAS algorithm (2005).
The credit of discovery of Budan’s rule occasioned a dispute of sorts between the partisans of Fourier and those of Budan. Darboux gave an historical account of the work of the two authors in a note ([4]). Since Budan, Fourier and Darboux were all French, the latter can not be suspected of partiality. The salient facts are as follows: During the years 1796-1803, Fourier taught the rule at the école polytechnique, and was most probably in possession of a proof. Nevertheless, he did not publish anything, and several communications to the institut de France were lost. Budan stated his rule in 1806 in a treatise, but was unable to prove it. In 1811, he presented a mémoire on the resolution of equations, including an algebraic proof of the rule, to the Academie des Sciences. This proof was found to be roughly correct by the commissars Lagrange and Legendre. In 1820, Fourier published his paper [8], in which he investigated the theorem and its application to the theory of equations to a greater depth than Budan had. The analytic proof of Fourier turned out to be far superior to that of Budan. In 1822, Budan published a new mémoire on the resolution of equations, in which he included the demonstration of the rule he gave in his first 1811 mémoire ([2]). In the opinion of Darboux, the rule should be, without doubt, attributed to Fourier. The arguments of Darboux may be convincing, but we believe that the only objective criterion to attributing a theorem is that its author published it, or in the very least showed it in a sufficiently wide, public forum. Thus, even if Budan’s rule was discovered and proved by Fourier for the first time, as seems probable, it should nevertheless be attributed to Budan.

It is to be observed that Descartes’s rule can be deduced easily from Budan’s, and that it is not only valid for polynomials, but also for functions that own an \( n \)-th derivative of constant sign inside some interval \([a,b]\) (see Thm. 2.4.1 below). In particular, it is valid for polynomial with fractional exponents. A somewhat vague proof of Budan’s rule was given in Dickson ([7]). For the most part, proofs of Budan’s rule fail to be sufficiently rigorous to comply with modern standards. In particular, we believe that an argument equivalent to the least-upper bound property cannot be avoided in general analytic proofs, valid for functions of the above-mentioned type. Nevertheless, in the case of polynomials, algebraic proofs can also be given, in the spirit of Descartes’s strategy above (this is in fact the way Budan proved his theorem). Another modern and simple algebraic treatment, based on Descartes’s rule, can be found in [3].

In 1828, an extremely simple demonstration of Descartes’s rule was found by Gauss ([9]). It is exposed in a somewhat more understandable form in [7]. Despite its simplicity, it does not seem to be well known, even by renowned mathematicians (in [7], Dickson said the proof was communicated to him by Curtiss, and seemed to ignore it is due to Gauss). This may explain why proofs for Descartes’s rule are still being invented. The neatly simple proof of Gauss is algebraic, following, once more,
Descartes’s strategy above, and is based on sign alternation considerations only. In the same paper, Gauss explicitly states the first of the two extensions mentioned above, together with an assertion involving De Gua’s rule (implicitly). Curiously, Abraham Adrian Albert was unsatisfied with the proof of Gauss, and produced a more complicated proof in 1943 ([1]).

A ultra-simple algebraic proof of Descartes’s rule was finally published by Krishnaiah in 1963, based upon Descartes’s strategy ([13]). Since the proofs of Gauss and Krishnaiah, as well as our own, are probably the simplest that will ever be invented, it may be interesting to compare them. Gauss’s proof is non-inductive and may be preferred over the other proofs by persons experienced in algebra. The proof of Krishnaiah can be seen as a formulation of the proof of Gauss in an inductive form; this makes it simpler, though it may be considered inferior because of its inductive nature. Both proofs do not use any special conventions to count the alternations of signs of polynomials: lacunary coefficients are simply ignored. The guiding principle of Krishnaiah’s proof, which is in essence the same as Gauss’s, is the following: assume inductively that multiplying a polynomial with \( r \) alternations of signs \((r > 0)\) increases its number of alternations by 1. Consider a polynomial 
\[ P = a_0 + \cdots + a_n x^n \]
containing \( r + 1 \) alternations of signs, and let \( a_s \) be the last coefficient at which an alternation of sign occurs (say \( a_s a_{s+1} \)). Write the polynomial in the form \( Q + R \), where \( Q \) is the part of \( P \) up to the \( s \)th coefficient, and \( R \) is the remaining part of \( P \). In particular \( Q \) contains \( r \) alternations of signs. Now, 
\[(X - \alpha) P = (X - \alpha) Q + (X - \alpha) R.\]
It is important to note that due to the definition of \( s \), the free and leading coefficients of \( R \) share the same signs. Hence it is not difficult to show that \( P \) must contain at least \( r + 2 \) alternations of signs, achieving the induction. In this proof, it is interesting to see that the induction is performed on the number of alternations of signs, and not on the degree of \( P \).

Not so in our proof: Here, the induction is performed on the degree of \( P \), which is perhaps more natural. As mentioned above, what prevents most of the proofs in the old literature from being rigorous is the lack of a suitable convention of alternations and permanences of signs of the polynomial coefficients. The merit of the proofs of Gauss and Krishnaiah is that they overcome this difficulty without introducing any special convention or external concepts. Nonetheless, once a suitable convention is established, Descartes’s rule follows in a straightforward manner, as does the first extension outlined above; De Gua’s rule can be shown to be an artifact of this convention, while what is referred to below as “Fourier’s rule” appears to be a simple corollary of complex analysis, deinterlaced from Descartes’s and Budan’s rules. Moreover, even the proof of Budan’s rule is simplified by this convention.

In conclusion, the proof of Krishnaiah may be preferred if the goal is to provide a quick proof of Descartes’s rule, free from any other considerations. Our proof of may be preferred over that of Gauss for its rigor and suitability to be taught at a very
elementary level. It may also be preferred over Krishnaiah’s proof if deeper insight of the principles involved is desired, as well as of the extensions of the rule.

2. Descartes’s Presumed Convention and its Applications

This section is devoted to the exposition of the convention mentioned above, and to showing how it simplifies the proofs of Descartes’s rule, De Gua’s rule and Budan’s. From an historical point of view, this may reflect Descartes’s thought more precisely than other methods.

CONVENTION: Consider a polynomial \( P(X) = a_nX^n + a_{n-1}X^{n-1} + \cdots + a_0 \), where some coefficients \( a_i \) may be null. If two contiguous coefficients \( a_{i+1} \) and \( a_i \) are both positive or both negative, then the pair \( (a_{i+1}, a_i) \) contributes 1 permanence and 0 alternation of signs. If one of them is positive and the other is negative, then it contributes 1 alternation and 0 permanence. If now the polynomial has one or more coefficients equal to 0, then in the context of counting the alternations of signs of \( P \), every coefficient \( a_i = 0 \) is considered to be of the same sign as \( a_{i+1} \). Therefore, the sign of a non-zero coefficient limiting a sequence of null coefficients on the left propagates to the entire sequence. For example, if \( P(X) = 3X^4 - X \), we assign the following signs to the lacunary coefficients: \( P(X) = 3X^4 + 0 \cdot X^3 + 0 \cdot X^2 - X + 0 \). In practicality, this is the same as ignoring lacunary coefficients.

Contrarily, in the context of counting the permanences of signs of \( P(X) \), we consider the sign of a coefficient \( a_i = 0 \) to be opposite to the sign of \( a_{i+1} \). Thus, the signs of a sequence of null coefficients alternate, starting from the first non-zero coefficient limiting the sequence on the left. For example, using the previous polynomial, the signs are \( P(X) = 3X^4 - 0 \cdot X^3 + 0 \cdot X^2 - X + 0 \). It is important to note that unlike alternations of signs, this convention does not amount to dropping the null coefficients. In fact, the reader should mentally check that the following rule holds: a sequence of coefficients of the form \( a_i, 0, 0, \ldots, 0, a_j \), with \( a_i a_j \neq 0 \), contributes 1 permanence if \( a_i \) and \( a_j \) are of opposite signs and the number of “0” is odd; or, if \( a_i \) and \( a_j \) are of the same signs and the number of “0” is even. Otherwise, it contributes no permanence.

It may seem strange to use two different methods in order to count the number of alternations and of permanences of signs, but the advantages this method offers are twofold: First, with this convention, the number of alternations and sign permanence is always minimized in lacunary polynomial, as can be easily verified. Second, these two methods are dual in the sense that the number of permanences of signs of \( P(X) \) is always equal to the number of alternations of signs of \( P(-X) \). The verification of this simple fact is left to the reader, as well.

In the rest of this section we shall show how this convention can be applied in order to
provide a ultra-simple proof of Descartes’s rule, and to simplify the demonstrations of the other rules cited in the title of this paper.

Notations: Henceforth, we use $z^+(P)$ to denote the number of positive roots of a polynomial $P$, $z^-(P)$ to denote the number of its number of negative roots, and $z^0(P)$ to denote the number of its null roots; roots are always counted with their multiplicities. Of course, the number $z^0(P)$ can be read over from $P$ itself, as it is the largest power of $X$ dividing $P$. We also use $v(P)$ to denote the number of alternations of signs of $P$, and by $c(P)$ the number of its permanences (hence $c(P(X)) = v(P(-X))$). We hold that a coefficient $a_i$ is trailing if $a_i = a_{i-1} = \cdots = a_0 = 0$.

We mark that if the polynomial $P$ is not lacunary, except, perhaps, in its trailing coefficients, then

$$v(P) + c(P) = \deg(P) - z^0(P).$$

Indeed a pair of two contiguous coefficients is either in alternation or in permanence of signs. On the other hand, if $P$ is lacunary, then according to our convention, outlined above, a block of contiguous coefficients of the form $a_i, 0, 0, \ldots, 0, a_j$ (with $a_i a_j \neq 0$) contributes at the most one alternation and one permanence. Since any such block contains at least 3 coefficients (and therefore at least two pairs of contiguous coefficients), the sum of the numbers of alternations and permanences of $P$ is not larger than the overall number of pairs of contiguous (and not trailing) coefficients in $P$. In other words, it is not larger than $\deg(P) - z^0(P)$:

$$v(P) + c(P) \leq \deg(P) - z^0(P).$$

2.1. Descartes’s Rule of Signs

With the notations above, Descartes’s rule and its special case can be formulated as follows:

For any real polynomial $P$,

$$v(P) \geq z^+(P) \quad \text{and} \quad c(P) \geq z^-(P).$$

Furthermore, if all the roots of $P$ are real, then

$$v(P) = z^+(P) \quad \text{and} \quad c(P) = z^-(P).$$

Proof: The first assertion of the theorem implies the second, since if all the roots of $P$ were real and such that $v(P) > z^+(P)$ or $v(P) > z^-(P)$, there would hold

$$\deg(P) - z^0(P) = z^+(P) + z^-(P) < v(P) + c(P),$$

contradicting (2).
To prove that \( v(P) \geq z^+(P) \), note that if \( \alpha \) is a real root of \( P \), then \( P = (X - \alpha)Q(X) \) where \( Q \) is a polynomial of degree smaller than \( P \). Thus, by an evident induction on the degree of \( P \), it suffices to prove that multiplying a polynomial \( Q(X) \) by \( X - \alpha \), with \( \alpha > 0 \), increases the number of its alternations of signs (Descartes’s strategy). We can see that \( c(P) \geq z^-(P) \) follows from \( v(P) \geq z^+(P) \) by changing \( P(X) \) into \( P(-X) \). Let

\[
Q(X) = a_nX^n + \cdots + a_0 \quad \text{and} \quad P(X) = (X - \alpha)Q(X) = b_{n+1}X^{n+1} + \cdots + b_0.
\]

The following relations hold:

\[
b_i = \begin{cases} 
a_{i-1}, & \text{if } i = n + 1, \\
\alpha a_i, & \text{if } n \geq i \geq 1, \\
-\alpha a_i, & \text{if } i = 0.
\end{cases} \quad (1)
\]

Let us denote the sign of \( a_i \) by \( s_i \), and the sign of \( b_i \) by \( s'_i \):

\[
s_i, s'_i \in \{+,-\}.
\]

We input the signs of \( a_i \) and \( b_i \) into a “table of signs” in the following manner:

| \( i \) | \( n + 1 \) | \( n \) | \( n - 1 \) | \( \cdots \) | 0 |
|---|---|---|---|---|---|
| sign of \( a_i \) | \( s_n \) | \( s_{n-1} \) | \( \cdots \) | \( s_0 \) |
| sign of \( b_i \) | \( s'_{n+1} \) | \( s'_n \) | \( s'_{n-1} \) | \( \cdots \) | \( s'_0 \) |

For the sake of simplicity, let us denote such a table by

\[
(s_n, s_{n-1}, \ldots, s_0; s'_{n+1}, s'_n, \ldots, s'_0),
\]

and call the number \( n \) the extent of the table.

Because of relation (1), any such table should satisfy the following three properties:

1. \( s'_{n+1} = s_n \)
2. \( s'_0 \neq s_0 \),
3. if \( s_i \neq s_{i-1} \), then \( s'_i = s_{i-1} \).

It is important to mark that these properties hold even if \( P \) and \( Q \) have lacunary coefficients, due to the convention above. To prove Descartes’s rule, it suffices to prove that the third row in the table contains more alternations of signs than the second row. To this end, we need nothing more than the three aforementioned properties.
If the extent of the table is 1, this is obviously true by the first and second properties. Let us inductively assume this assertion for tables of extent at most \( n - 1 \). Let \( T = (s_0, s_1, \ldots, s_n; s'_0, s'_1, \ldots, s'_n) \) be a table of extent \( n \).

Assume first that no alternation of signs occurs in the sequence \((s_0, \ldots, s_n)\) (that is, it is constant). Then \( s'_0, \ldots, s'_n \) must contain at least one alternation of signs: indeed, by hypothesis, \( s'_n = s_n \) and \( s'_0 \neq s_0 \), hence \( s'_n \neq s'_0 \). Thus, there must exist some \( i \in \{1, \ldots, n+1\} \) such that \( s'_i \neq s'_{i-1} \). This shows that in this case, the induction hypothesis is fulfilled for table \( T \).

Now assume that an alternation of signs occurs at some rank \( i > 0 \) inside the sequence \( s_0, \ldots, s_n \):

\[ s_i \neq s_{i-1}. \]

By the third property above, \( s'_i = s_{i-1} \), hence \( s'_i \neq s_i \). We can extract from \( T \) the following table, of smaller extent:

\[ T' = (s_0, \ldots, s_i; s'_{n+1}, \ldots, s'_n). \]

This table obviously satisfies the three properties above. Therefore, using \( k_1 \) to denote the number of alternations of \( s_0, \ldots, s_i \) and \( k'_1 \) to denote that of \( s'_{n+1}, \ldots, s'_n \), the induction hypothesis implies \( k'_1 > k_1 \). On the other hand, we can extract another table of smaller extent from \( T \), namely

\[ T'' = (s_{i-1}, \ldots, s_0; s'_1, \ldots, s'_0). \]

It satisfies the properties above as well, hence the sequence \( s'_1, \ldots, s'_0 \) must also contain more alternations than the sequence \( s_{i-1}, \ldots, s_0 \), say \( k'_2 > k_2 \). But the number of alternations of \( s_{n+1}, \ldots, s_0 \) is at most equal to \( k_1 + k_2 + 1 \), because this sequence, being the concatenation of \( s_0, \ldots, s_i \) and \( s_{i-1}, \ldots, s_0 \), contains only the alternations of these sequences, and possibly an additional alternation \( s_is_{i-1} \) (if \( s_i \neq s_{i-1} \)). Since \( k'_1 + k'_2 \geq k_1 + 1 + k_2 + 1 > k_1 + k_2 + 1 \), it follows that the sequence \( s'_{n+1}, \ldots, s'_0 \) contains more alternations than the sequence \( s_0, \ldots, s_n \), in accordance to the induction hypothesis. \( \blacksquare \)

2.2. Fourier’s Rule

We will now show that Fourier’s rule is essentially independent of Descartes’s rule, since each can be proved algebraically without the help of the other. The rule is:

For any real polynomial \( P \), \( v(P) \) and \( z^+(P) \) are of the same parity, as are \( c(P) \) and \( z^-(P) \).
Proof: The second assertion follows immediately from the first by changing $P(X)$ into $P(-X)$.

We can assume without loss of generality that $P(0) \neq 0$, because if $P(0) = 0$, then the polynomial $Q$ obtained by dividing $P$ by a suitable power $X^k$ fulfills $Q(0) \neq 0$ and contains the same number of alternations of signs and the same number of positive roots as $P$.

Let us write $P = AP_1P_2$, where $P_1$ and $P_2$ are monic, $A \in \mathbb{R}$, $P_1$ has no real positive roots, and all the roots of $P_2$ are real positive.

Let us put $\deg(P_2) = n$. Since the coefficients of $P_1$ are real, its complex roots come in conjugate pairs $\gamma, \bar{\gamma}$. Therefore the polynomial $P_1$ can be written

$$P_1(X) = (X - \gamma_1)(X - \bar{\gamma}_1)(X - \gamma_2)(X - \bar{\gamma}_2) \cdots (X + \alpha_1)(X + \alpha_2) \cdots$$

where the $\gamma_i, \bar{\gamma}_i$ are the complex roots of $P_1$, and $-\alpha_i$ are its negative roots ($\alpha_i > 0$).

The free coefficient of $P_1$ is clearly equal to

$$P_1(0) = (-\gamma_1)(-\bar{\gamma}_1)(-\gamma_2)(-\bar{\gamma}_2) \cdots \alpha_1\alpha_2 \cdots = |\gamma_1|^2|\gamma_2|^2 \cdots \alpha_1\alpha_2 \cdots > 0.$$

Thus, the free coefficient of $P_1$ is strictly positive: $P_1(0) > 0$. Put

$$P_2 = (X - \beta_1)(X - \beta_2) \cdots,$$

where $\beta_1, \beta_2, \cdots > 0$. The free coefficient of $P$ is $P(0) = AP_1(0)P_2(0)$, hence is of the same sign as $AP_2(0) = A(-\beta_1)(-\beta_2) \cdots$. Consequently, the signs of $A$ and $P(0)$ are equal if and only if $P$ contains an even number of positive roots.

Thus, we see that the theorem is equivalent to the fact that a polynomial contains an even number of alternations of signs if and only if its leading and free coefficients have like signs. To this end, we need only show that a sequence of signs contains an even number of alternations if and only if its extremities are equal. This is obviously true if the length of the sequence is 2. Assume inductively this assertion is true for sequences of lengths $n - 1$, and consider a sequence of signs $S = (s_1, s_2, \ldots, s_n)$ of length $n$ ($s_i \in \{+,-\}$). Let $S' = (s_1, s_2, \ldots, s_{n-1})$. If $s_{n-1} = s_n$, then the number of alternations of $S$ and $S'$ are equal, and the extremities of $S$ and $S'$ are pairwise equal; since $S'$ fulfills the induction hypothesis, it is clear that $S$ fulfills it as well. If $s_{n-1} \neq s_n$, then the number of alternations of $S'$ is larger than the number of alternations of $S$ by 1; hence the parities of these numbers are opposite. But the extremities of $S$ are either opposite or equal, depending on whether the extremities of $S'$ are equal or opposite, resp. This implies that the induction hypothesis holds for $S$ in this case, as well.  

2.3. De Gua’s Rule

De Gua’s rule is:
If, in a polynomial \( P \), a group of \( r \) consecutive terms is missing, then \( P \) has at least
\( r \) imaginary roots if \( r \) is even, or it has at least \( r + 1 \) or \( r - 1 \) imaginary roots if \( r \) is odd, depending on whether the terms immediately preceding and following the group have like or unlike signs resp.

It turns out that De Gua’s rule is weaker than the following rule, that is an immediate corollary of Descartes’s rule of signs:

2.3.1. Corollary: The number of imaginary roots of a polynomial \( P \) is at least equal to
\( \text{deg}(P) - z^0(P) - v(P) - c(P) \).

Therefore, all we must prove is that this corollary implies De Gua’s rule.

Proof: In the process of counting the number of alternations and permanences of signs of \( P \), only pairs of contiguous non-zero coefficients \( (a_i, a_{i-1}) \) are involved, or blocks of contiguous coefficients of the form \( (a_i, 0, \ldots, 0, a_j) \) with \( a_ia_j \neq 0 \), containing one or more “0”, say \( r \) times “0”. Now, a pair \( (a_i, a_{i-1}) \) contributes either one alternation, or one permanence of signs in the overall sum \( v(P) + c(P) \).

However, according to what has been explained at the beginning of Sec. 2, a block such as the one above contributes one alternation if \( a_ia_j < 0 \), 0 alternation if \( a_i a_j > 0 \), one permanence if \( a_ia_j > 0 \) and \( r \) is even or if \( a_ia_j < 0 \) and \( r \) is odd, and 0 permanence otherwise. Let \( q \) be the number of consecutive pairs in the block: \( q = r + 1 \). If \( r \) is even, we have just seen that the block contributes either one alternation, or one permanence. Therefore, the difference between the number of consecutive pairs and the sum of the numbers of alternations and permanences occurring in this block is \( q - 1 = r \), a positive loss. Similarly, if \( r \) is odd and \( a_ia_j < 0 \), then the block contributes one alternation and one permanence, hence the loss is \( q - 2 = r - 1 \). Finally, if \( r \) is odd and \( a_ia_j > 0 \), then it contributes no alternation and no permanence, hence the loss is \( q = r + 1 \). In any case, the loss is \( \geq 0 \) (in the case of a normal contiguous pair \( (a_i, a_{i-1}) \) it is null). Furthermore, it is clear that the process of computing the loss associated with a block coincides exactly with De Gua’s method. Since \( \text{deg}(P) - z^0(P) \) is equal to the number of contiguous pairs of non-trailing coefficients, it follows that \( \text{deg}(P) - z^0(P) - v(P) - c(P) \) is the sum of the losses of each block as above. This particularly implies De Gua’s rule, and is even more powerful, as one can apply this rule to each sequence of contiguous zero (non-trailing) coefficients, summing the losses over the sequences.

2.4. Budan’s Rule

As stated earlier in the paper, Budan’s rule can be given a simple algebraic proof based on Descartes’s rule. Even so, it cannot be extended to more general classes
of functions than polynomials. Admittedly, the proof we present is not *ultra-simple*, but it has the merit of being valid for every function $f$ whose $n$-th differential does not cancel and is of constant sign inside an interval (see Thm. 2.4.1 below). We did not find this extension explicitly in the work of Fourier, but in regard to his proof, it is likely that he had an extension of this type in mind. Once more, the minimization convention allows the proof to be simplified.

Let $P$ be a polynomial of degree $n$. We continue using our previous minimization convention, and use $v(P, t)$ to denote the number of variations of signs of the polynomial

$$Q(X) = P(X + t) \quad (t \in \mathbb{R}).$$

According to Taylor’s theorem, this is also the number of alternations of signs in the sequence

$$P^{(n)}(t), P^{(n-1)}(t), \ldots, P'(t), P(t).$$

This suggests using $v(f, t, n)$ to denote, for every function $f$ differentiable up to the order $n$, the number of alternations of signs of the sequence

$$f^{(n)}(t), f^{(n-1)}(t), \ldots, f'(t), f(t),$$

where we adopt the same minimization convention as we did above for sequences of polynomial coefficients. We also say that a zero $\alpha$ of $f$ is of multiplicity $\mu$ if $f^{(i)}(\alpha) = 0$ for all $0 \leq i \leq \mu$, and use $z(f, u, v)$ to denote the number of zeros of $f$ inside $[u, v]$, counted with their multiplicities. Finally, we assert that a function is differentiable $n$ times in $[a, b]$ if it is differentiable $n$ times in $]a, b[$, and is differentiable $n$ times to the right of $a$ and to the left of $b$. Since the genuine Taylor’s Theorem is based only on L’Hospital’s rule, it is clear that right and left versions of Taylor’s theorem hold in this case at $a$ and $b$. Budan’s rule is:

*Let $P(X)$ be a polynomial of degree $n$, and $a, b \in \mathbb{R}$ with $a \leq b$. If $a$ is not a root of $P$, then the number of real roots of $P$ inside the interval $[a, b]$ is not larger than $v(P, a) - v(P, b)$, and is always of the same parity.*

**Remarks:** Because of their polynomial nature, the functions $P^{(i)}(x) \to \pm \infty$ as $x \to +\infty$, for all $i \in \{0, \ldots, n\}$. From this, elementary considerations show that whenever $x \to +\infty$, either $P^{(i)}(x) \to +\infty$ for all $i$, or $P^{(i)}(x) \to -\infty$ for all $i$. In any case, $v(P, x) \to 0$ as $x \to +\infty$, or, in other words, $v(P, x) = 0$ for all $x$ sufficiently large. Thus the above theorem implies Descartes’s rule of signs, and provides an analytic proof.

Notice also that the final assertion of the above rule is simply the extension of Fourier’s rule, and cannot be extended to other classes of functions, essentially because of the algebraic structure of polynomials.
Proof: Put \( x = x' + t \); then \( x > t \) if and only if \( x' > 0 \). Hence the roots of \( P \) larger than \( t \) correspond univocally to the positive roots of \( Q(X) = P(X + t) \).

Consequently, Fourier's rule (Sec. 2.2) implies that the number of real roots of \( P \) larger than \( a \) is equal to \( v(P, a) + 2k_1 \), and the number of real roots of \( P \) larger than \( b \) is equal to \( v(P, b) + 2k_2 \) \( (k_1, k_2 \in \mathbb{Z}) \). Since \( a \) is not a root of \( P \), \( v(P, a) + 2k_1 \) is also the number of roots of \( P \) larger than, or equal to \( a \), hence

\[
z(P, a, b) = v(P, a) - v(P, b) + 2(k_1 - k_2).
\]

Consequently, \( z(P, a, b) \) is of the same parity as \( v(P, a) - v(P, b) \), which proves the final assertion of the theorem. We now show that \( z(P, a, b) \leq v(P, a) - v(P, b) \).

In order to allow the proof a broader generality, let \( f \) be a function \( [a, b] \rightarrow \mathbb{R} \), differentiable \( n \) times, and assume that the \( n \)th derivative \( f^{(n)}(x) \) does not cancel and is of constant sign inside \( [a, b] \). It can be shown that this implies \( f^{(i)}(x) \) has at most a finite number of zeros in \( [a, b] \), for all \( i < n \). It can be assumed here since this last property obviously holds if \( f \) is a polynomial. Notice that polynomials are a particular case of this setting, since the \( n \)th derivative of a (non zero) polynomial of degree \( n \) is constant and not null. For such a function \( f \), we prove that the number of zeros of \( f \) inside \( [a, b] \) is at most equal to \( v(f, a, n) - v(f, b, n) \), provided \( a \) is not a zero of \( f \).

This obviously holds true if \( n = 0 \), since the sequence \( f^{(n)}(t), \ldots, f(t) \) reduces to \( f(t) \) for all \( t \in [a, b] \), and hence has no alternations in this interval. Let us assume inductively that this assertion is true for any interval \( [u, v] \), and any function \( f \) differentiable \( n - 1 \) times in \( [u, v] \), with \( f(u) \neq 0 \), such that \( f^{(n-1)} \) does not cancel and is of constant sign inside \( [u, v] \).

Let \( f \) be a function differentiable \( n \) times, whose \( n \)-th differential satisfies the same properties, and such that \( f(a) \neq 0 \). As usual, we assert that a property holds for all \( t \) sufficiently close to \( a \) if there exists \( \eta > 0 \) of such that it holds for all \( t \) with \( |t - a| < \eta \).

**Claim 1:** If \( \alpha \in [a, b] \) is not a zero of \( f \), then for every \( u, v \in [a, b] \) sufficiently close to \( \alpha \), with \( u \leq \alpha \leq v \),

\[
z(f, u, v) = 0 \quad \text{and} \quad v(f, u, n) - v(f, v, n) \geq 0.
\]

**Proof:** Because of the continuity of \( f \), for all \( u, v \in [a, b] \) with \( u \leq \alpha \) and \( v \geq \alpha \) sufficiently close to \( \alpha \), \( f \) does not cancel inside \([u, v]\); it is therefore of constant sign in this interval. Assume that \( f'(\alpha) \neq 0 \); then, with the same reasoning, \( f' \) does not cancel inside \([u, v]\), for all \( u, v \in [u, v] \) with \( u \leq \alpha \) and \( v \geq \alpha \) sufficiently close to \( \alpha \).

So, \( f \) and \( f' \) do not cancel and are of constant sign inside such intervals \([u, v]\). (2)
By the induction hypothesis, \( v(f', u, n - 1) - v(f', v, n - 1) \geq 0 \). In other words, the number of alternations of signs in the sequence

\[ f^{(n)}(u), \ldots, f'(u) \]

is not smaller than the number of alternations of signs in the sequence

\[ f^{(n)}(v), \ldots, f'(v). \]

But according to (2), the pair of signs determined by \((f'(u), f(u))\) and \((f'(v), f(v))\) should be equal, and hence the number of alternations of signs in

\[ f^{(n)}(u), \ldots, f'(u), f(u) \]

remains not smaller than the number of alternation of signs in

\[ f^{(n)}(v), \ldots, f'(v), f(v). \]

This shows that \( v(f, u, n) \geq v(f, v, n) \), and the claim is established in this case.

If, in contrast, \( f'(\alpha) = 0 \), then for all \( u, v \in [a, b] \) with \( u \leq \alpha \) and \( v \geq \alpha \) sufficiently close to \( \alpha \), \( f' \) does not cancel inside \([u, v]\), except, of course, at \( \alpha \): this is a derivative of the fact that the zeros of \( f' \) are supposed to be in finite number, and so are isolated. The induction hypothesis now implies that \( v(f', u, n - 1) \geq v(f', v, n - 1) + 1 \), and since the pair of signs determined by \((f'(v), f(v))\) can add only one more alternation to the sequence

\[ f^{(n)}(v), \ldots, f'(v), f(v) \]

(recall our minimization convention), it follows that \( v(f, u, n) \geq v(f, v, n) \). Hence, the claim is established in this case, as well.

**Claim 2:** If \( \alpha \in [a, b] \) is a zero of multiplicity \( \mu \) of \( f \), then for all \( u, v \in [a, b] \) sufficiently close to \( \alpha \), with \( u < \alpha \leq v \),

\[ z(f, u, v) = \mu \quad \text{and} \quad v(f, u, n) - v(f, v, n) \geq \mu. \]

**Proof:** Notice first that necessarily, \( \mu < n \). Since the zeros of \( f \) are finite in number, they are isolated and \( z(f, u, v) = \mu \) for all \( u, v \in [a, b] \) sufficiently close to \( \alpha \), with \( u < \alpha \leq b \).

If, for some \( i \) with \( \mu < i < n \), \( f^{(i)}(t) \) does not cancel inside any interval \([u, v]\) with \( u, v \) sufficiently close to \( \alpha \) and \( u < \alpha \leq v \), then it is of constant sign inside \([u, v]\), since it is continuous. According to the induction hypothesis,

\[ v(f, u, i) \geq v(f, v, i) + \mu \quad \text{and} \quad v(f^{(i)}, u, n - i) \geq v(f^{(i)}, v, n - i). \]
However, the sequence \( f^{(n)}, \ldots, f', f \) is the product of combining the sequences, \( f^{(n)}, \ldots, f^{(i)} \) and \( f^{(i)}, \ldots, f \). Therefore
\[
v(f, u, n) \geq v(f, v, n) + \mu.
\]
The claim, therefore, holds true in this case.

Now, if for all \( i \) with \( \mu < i < n \) and all \( u, v \in [a, b] \) sufficiently close to \( \alpha \), with \( u < \alpha \leq v \), \( f^{(i)} \) cancels at some point inside \([u, v]\), then due to the continuity of \( f^{(i)} \), \( f^{(i)}(\alpha) = 0 \) for all \( \mu < i < n \). So, \( f^{(i)}(\alpha) = 0 \) for all \( 0 \leq i < n \). Since the zeros of \( f, f', \ldots, f^{(n-1)} \) are isolated, we can even suppose that the functions \( f, f' \ldots f^{(n-1)} \) cancel only at \( \alpha \) inside \([u, v]\). Taylor’s theorem implies that
\[
\begin{align*}
f(u) &= \frac{1}{n!} f^{(n)}(\alpha)(-1)^n(\alpha - u)^n + o((\alpha - u)^n), \\
f'(u) &= \frac{1}{(n-1)!} f^{(n)}(\alpha)(-1)^{\mu}(\alpha - u)^{n-1} + o((\alpha - u)^{n-1}), \\
&\vdots \\
f^{(n)}(u) &= \frac{1}{(n-\mu)!} f^{(n)}(\alpha)(-1)^{n-\mu}(\alpha - u)^{n-\mu} + o((\alpha - u)^{n-\mu}).
\end{align*}
\]
This shows that the signs of \( f^{(i)}(u) \) and \( f^{(i+1)}(u) \) are opposite for every \( u < \alpha \) sufficiently close to \( \alpha \), and \( 0 \leq i < \mu \). Hence the sequence \( f^{(n)}(u), \ldots, f(u) \) has at least \( \mu \) alternations of signs. On the other hand, Taylor’s theorem again leads to
\[
f^{(i)}(v) = \frac{1}{(n-i)!} f^{(n)}(\alpha)(v-\alpha)^{n-i} + o((v-\alpha)^{n-i}),
\]
for all \( v \geq \alpha \) sufficiently close to \( \alpha \) and \( 0 \leq i < n \). This shows that the sequence \( f^{(n)}(v), \ldots, f(v) \) has no alternations of signs. In conclusion, \( v(f, u, n) \geq v(f, v, n) + \mu \) in this case, too.

In order to prove the theorem, first note that the induction hypothesis holds for \( f \) inside \([a, v]\), for every \( v \geq 0 \) sufficiently close to \( u \) (Claim 1 with \( u = 0 \)). Let \( s \) be the supremum of the numbers \( t \leq b \) such that the theorem holds for every interval \([a, v]\) with \( v \leq t \) (notice that \( s > a \)). Assume, in order to obtain contradiction, that \( s < b \), or that \( s = b \) but the theorem does not hold in \([a, b]\). Then Claim 1 and Claim 2 imply that for all \( u, v \in [a, b] \) with \( u < s \) and \( v \geq s \) sufficiently close to \( s \),
\[
v(f, a, n) - v(f, v, n) = v(f, a, n) - v(f, u, n) + v(f, u, n) - v(f, v, n) \\
\geq z(f, a, u) + z(f, u, v) = z(f, a, v),
\]
where the final equality results from the fact that \( u \) is never a zero of \( f \) in Claim 1 and Claim 2. Thus, the induction hypothesis holds for \( f \) inside \([a, v]\), for every \( v \in [a, b] \) sufficiently close to \( v \), with \( v \geq s \), a contradiction.
The proof above contains the promised generalization of Budan’s theorem, which does not seem to be well known:

2.4.1. Theorem (generalized Budan’s rule): Let \( f : [a, b] \to \mathbb{R} \) be a function differentiable \( n \) times inside \([a, b]\). Assume that the \( n \)-th derivative \( f^{(n)} \) does not cancel and is of constant sign inside \([a, b]\), and that \( f(a) \neq 0 \). Then the number of zeros of \( f \) inside \([a, b]\), including multiplicities, is not larger than

\[
v(f, a, n) - v(f, b, n).
\]

Proof: The proof is contained within the previous proof, excepting the fact that the derivatives \( f^{(i)} \) have a finite number of zeros inside \([a, b]\), for every \( 0 \leq i \leq n \). Consider the property \( \mathcal{P}(n) \), depending on \( n \), defined by:

For every function \( f \) \( n \) times differentiable, if the \( n \)-th derivative \( f^{(n)} \) does not cancel and is of constant sign in \([a, b]\), then the number of zeros of \( f \) is at most finite.

Clearly, \( \mathcal{P}(0) \) holds. Assume inductively that \( \mathcal{P}(n - 1) \) holds for all \( n \geq 1 \), and let \( f \) be a function \( n \) times differentiable, such that \( f^{(n)} \) does not cancel and is of constant sign in \([a, b]\). If the number of zeros of \( f \) were infinite inside \([a, b]\), then By Rolle’s theorem, between two zeros of \( f \), would lie at least one zero of \( f' \), hence the number of zeros of \( f' \) would be infinite in \([a, b]\) as well. This would contradict \( \mathcal{P}(n - 1) \). Therefore, \( \mathcal{P}(n) \) holds and the theorem is proved.

Bibliography

[1] A. A. Albert. An inductive proof of Descartes’s rule of sign. The American Mathematical Monthly, 50(3):178–180, Mar. 1943.

[2] F. D. Budan. Nouvelle méthode pour la résolution des equations d’un degré quelconque. Bondé-Dupré, Bachelier, Paris, 1822.

[3] N. B. Conkwright. An elementary proof of the Budan-Fourier theorem. The American Mathematical Monthly, 50(10):603–605, Dec. 1943.

[4] G. Darboux. Œuvres de Fourier, volume 2, page 310. Gauthier-Villars & fils, Paris, 1890.

[5] R. Descartes. La géométrie (Discours de la méthode, third part), page 373. Ed. of Leyde, 1637.

[6] R. Descartes. Lettres de Mr Descartes, volume 3. Charles Angot, Paris, 1667.

[7] L. E. Dickson. First Course in the Theory of Equations, pages 71–74 and 83–85. John Willey & sons, New York, 1922.

[8] J. Fourier. Sur l’usage du théorème de Descartes dans la recherches des limites des racines. Bulletin des Sciences par la Société philomatique de Paris, pages 156–165 and 181–187, Oct. and Dec. 1820.

[9] C. F. Gauss. Beweis eines algebraischen lehrsatzes. Crelle’s Journal fur die reine und angewandte Mathematik, 3, Jan. 1828.
[10] J. P. De Gua. Sur le nombre des racines réelles ou imaginaires, réelles positives ou réelles négatives, qui se trouvent dans les équations de tous les degrés. In *Histoire de l’Académie royale des sciences (sec. mémoires)*, pages 72–95. Imprimerie Royale, Paris, 1741.

[11] J. P. De Gua. Recherches du nombre des racines réelles ou imaginaires, réelles positives ou réelles négatives, qui peuvent se trouver dans les équations de tous les degrés. In *Histoire de l’Académie royale des sciences (sec. mémoires)*, pages 435–494. Imprimerie Royale, Paris, 1741.

[12] B. S. Hourya. Deux moments dans l’histoire du théorème d’algèbre de Ch. F. sturm. *Revue d’histoire des sciences*, 41(2):99–132, Mar. 1988.

[13] P. V. Krishnaiah. A simple proof of Descartes’s rule of signs. *Mathematics Magazine*, 36(3):135, May–Jun. 1963.

[14] J. L. Lagrange. *Traité de la résolution des équations numériques de tous les degrés*, pages 150–166. Bachelier, Paris, second edition, 1826.

[15] F. U. TÆpinus. Démonstration du théorème de Harriot avec une méthode de chercher si une équation algébrique a toutes les racines possibles ou non. In *Histoire de l’Académie Royale des Sciences et des Belles-Lettres de Berlin*, pages 354–366. Haude & Spener, Berlin, 1758.

[16] J. A. Segner. Démonstration de la règle de Descartes, pour connoitre le nombre des racines affirmatives et négatives qui peuvent se trouver dans les équations. In *Histoire de l’Académie Royale des Sciences et des Belles-Lettres de Berlin*, pages 292–299. Haude & Spener, Berlin, 1756.

[17] J. A. Segner. Equations. In *Supplement to the Fourth, Fifth, and Sixth editions of the Encyclopædia Britannica*, volume 4, pages 674–675. Edinburgh, 1824.

[18] X. Wang. A simple proof of Descartes’s rule of signs. *The American Mathematical Monthly*, 111(6):525–526, Jun.–Jul. 2004.