

# Counting zeros of generalized polynomials: Descartes' rule of signs and Laguerre's extensions

G.J.O. Jameson (*Math. Gazette* **90**, no. 518 (2006), 223–234)

## 1. Introduction

One of the bedrock theorems of mathematics is the statement that a real polynomial of degree  $n$  has at most  $n$  real zeros. Probably the best-known proof is the algebraic one, by factorization. But there is also a pleasant analytic proof, by deduction from Rolle's theorem.

A slightly different question is how many *positive* zeros a polynomial has. Here the basic result is known as “Descartes' rule of signs”. It says that the number of positive zeros is no more than the number of sign changes in the sequence of coefficients. Descartes included it in his treatise *La Géométrie*, which appeared in 1637. It can be proved by a method based on factorization, but, again, just as easily by deduction from Rolle's theorem.

There are generalizations of Descartes' rule providing more specific information, still for polynomials. The *Budan-Fourier theorem* gives an upper bound for the number of zeros of a polynomial in any interval  $(a, b)$ , and *Sturm's theorem* gives, in principle, a method for determining the exact number of zeros in such an interval. Both theorems are quite laborious to apply in practice. For an account of them, see [Hen, chapter 6] or [Hou, chapter 2].

Here we will be concerned with generalizations of a different sort, to wider classes of functions. First, consider a *generalized polynomial*, that is, a function of the form

$$f(t) = \sum_{j=1}^n a_j t^{p_j}, \quad (1)$$

where the  $p_j$  can be any real numbers (listed in descending order). To ensure that such a function is defined in real numbers, we must restrict  $t$  to the *positive* reals.

The substitution  $t = e^x$  (which, of course, does not alter the number of zeros) transforms  $f(t)$  into the function

$$F(x) = \sum_{j=1}^n a_j e^{p_j x} \quad (x \in \mathbb{R}). \quad (2)$$

Functions of this sort are called (*generalized*) *Dirichlet polynomials*. Note that the further substitution  $b_j = e^{p_j}$  expresses  $F(x)$  in the form  $\sum_{j=1}^n a_j b_j^x$  (now with the  $b_j$  in descending order). The special case  $b_j = 1/j$  gives  $F(x) = \sum_{j=1}^n a_j/j^x$ : such functions are called “ordinary” Dirichlet polynomials (and, if continued to form an infinite series, a *Dirichlet*

series). However, we will drop the word “generalized” when speaking of functions of the form (2).

By adopting a proof based on Rolle’s theorem rather than factorization, Laguerre [La] showed that Descartes’ rule of signs extends to functions of both these types. Furthermore, under certain restrictions, similar statements apply with the sequence  $(a_j)$  replaced by the sequence of its partial sums  $(a_1 + \cdots + a_j)$ . These are powerful variants of Descartes’ rule, because in some cases the new sequence has far fewer sign changes than the original one.

These results of Laguerre were reproduced in the form of exercises, and partly with new methods, in [P-Sz], Part V, chapter 1. However, they have not received much attention in more recent texts: they seem to be in danger of falling into neglect. It is hoped that this article will do something to restore awareness of these interesting results and methods.

I am grateful to Peter Walker for directing me to the relevant literature and to Adam McBride and the *Math. Gazette* referee for helpful comments.

## 2. Rolle’s theorem and zeros

We start by stating Rolle’s theorem. The formal proof is to be found in any text on analysis, but the idea is highly convincing from a picture!

**Rolle’s theorem.** *Suppose that a function  $f$  is differentiable at all points of an interval  $[a, b]$ , and that  $f(a) = f(b)$ . Then there is at least one point  $t_0$  in the open interval  $(a, b)$  such that  $f'(t_0) = 0$ .*

Note that this says that  $t_0$  is *strictly* between  $a$  and  $b$ : this is critical for all the applications that follow. In particular, there is an immediate application to counting zeros of a function. If a (differentiable) function  $f$  has  $n$  distinct zeros on some interval  $I$ , then its derivative  $f'$  must have at least  $n - 1$  zeros there, since it has (at least) one in each of the  $n - 1$  gaps between the zeros of  $f$ .

Here it was implied that the zeros are simply counted as points in the obvious sense. However, without much extra trouble, one can extend it to the case where the zeros are counted with their orders, which is what we will be doing in the theorems to follow. (Readers can ignore this refinement without serious loss; those so inclined should read Theorem 2.2 and go to section 3.)

Let  $f$  be a function possessing all derivatives (we shall only be considering functions

of this type). One says that  $f$  has a zero of order (or multiplicity)  $k$  at the point  $t_0$  if

$$f(t_0) = f'(t_0) = \dots = f^{(k-1)}(t_0) = 0 \text{ and } f^{(k)}(t_0) \neq 0.$$

Note that  $f'$  then has a zero of order  $k - 1$  at  $t_0$ ; this even works for  $k = 1$  if a “zero of order 0” means a point that is not a zero! A zero of order 1 is also called a simple zero.

Given a function  $f$  and an interval  $I$ , we shall denote by  $Z(f, I)$  the number of zeros of  $f$  in  $I$ , counted with their orders. In other words, if there are  $n$  zeros, with orders  $k_r$  ( $1 \leq r \leq n$ ), then  $Z(f, I) = k_1 + \dots + k_n$ . We will just write  $Z(f)$  when  $I$  is the whole domain of  $f$  (which will usually be either the whole real line or  $(0, \infty)$ ). The consequence of Rolle’s theorem mentioned above can now be strengthened, as follows.

**Proposition 2.1.** *For any function  $f$  and interval  $I$ , we have  $Z(f', I) \geq Z(f, I) - 1$ .*

*Proof.* Let  $f$  have a zero of order  $k_r$  at  $x_r$  ( $1 \leq r \leq n$ ). Then  $f'$  has a zero of order  $k_r - 1$  at  $x_r$  (with the above comment about order 0): these add up to

$$\sum_{r=1}^n (k_r - 1) = Z(f, I) - n.$$

By Rolle’s theorem,  $f'$  also has at least  $n - 1$  zeros in the gaps between the points  $x_r$ . Together, these two facts give  $Z(f', I) \geq Z(f, I) - 1$ .  $\square$

The basic theorem on zeros of polynomials now follows very easily.

**Theorem 2.2.** *Let  $f$  be a polynomial of degree  $n$ . Then  $Z(f) \leq n$ .*

*Proof.* Induction on  $n$ . A polynomial of degree 0 is a non-zero constant, so has no zeros. In other words, the statement is true when  $n = 0$ . Assume now that the statement is true for a certain value of  $n$ , and let  $f$  be a polynomial of degree  $n + 1$ . Then  $f'$  is a polynomial of degree  $n$ , so, by the induction hypothesis,  $Z(f') \leq n$ . By Proposition 2.1,  $Z(f) \leq n + 1$ . So the statement is true for  $n + 1$ , as required.  $\square$

As a first excursion beyond polynomials, a similar proof (which we leave as an exercise for the reader), gives the following:

*Example 1.* Let  $f$  be a polynomial of degree  $n$ . Then  $e^t - f(t)$  has at most  $n + 1$  zeros, counted with their orders.

For those concerned to count zeros with their orders, we will need (repeatedly) the following lemma about products of functions.

**Lemma 2.3.** *Suppose that  $f$  has a zero of order  $k$  at  $t_0$ . Let  $g$  be another function,*

and let  $h(t) = f(t)g(t)$ . If  $g(t_0) \neq 0$ , then  $h$  has a zero of order  $k$  at  $t_0$ . If  $g$  has a simple zero at  $t_0$ , then  $h$  has a zero of order  $k + 1$  there.

*Proof.* By Leibniz's rule for derivatives of a product,

$$h^{(r)}(t) = \sum_{s=0}^{r-1} \binom{r}{s} f^{(s)}(t)g^{(r-s)}(t) + f^{(r)}(t)g(t).$$

If  $g(t_0) \neq 0$ , it follows that  $h^{(r)}(t_0) = 0$  for  $0 \leq r \leq k - 1$  and  $h^{(k)}(t_0) \neq 0$ . If  $g(t_0) = 0$  and  $g'(t_0) \neq 0$ , then  $h^{(r)}(t_0) = 0$  for  $0 \leq r \leq k$ , while  $h^{(k+1)}(t_0) = (k + 1)f^{(k)}(t_0)g'(t_0) \neq 0$ .  $\square$

We also need the next lemma, showing that the orders of zeros are preserved by the substitution that transforms Dirichlet polynomials into generalized ones.

**Lemma 2.4.** *Suppose that  $F$  is defined on  $\mathbb{R}$  and  $f(t) = F(\log t)$  for  $t > 0$ . If  $F$  has a zero of order  $k$  at  $x_0$ , then  $f$  has a zero of order  $k$  at  $e^{x_0}$ . Hence  $Z(f) = Z(F)$ .*

*Proof.* It is easily checked by induction that for each  $r \geq 1$ ,

$$f^{(r)}(t) = \frac{1}{t^r} \sum_{s=1}^r \alpha_{s,r} F^{(s)}(\log t)$$

for certain coefficients  $\alpha_{s,r}$ , with  $\alpha_{r,r} = 1$ . The first statement follows. So corresponding zeros have the same order, and hence  $Z(f) = Z(F)$ .  $\square$

A further point about the order of zeros is helpful. If  $f$  has a simple zero at  $t_0$ , so that  $f'(t_0) \neq 0$ , then  $f(t)$  has opposite signs on either side of  $t_0$ , since its graph crosses the axis with non-zero gradient. More generally, we have:

**Lemma 2.5.** *Let  $f$  have a zero of order  $k$  at  $t_0$ . If  $k$  is odd, then  $f(t)$  changes sign at  $t_0$ . If  $k$  is even, it does not change sign there.*

*Proof.* The first non-zero term in the Taylor series is  $(1/k!)(t - t_0)^k f^{(k)}(t_0)$ . More exactly, Lagrange's form of Taylor's theorem says that for  $t \neq t_0$ ,

$$f(t) = \frac{(t - t_0)^k}{k!} f^{(k)}(u)$$

for some  $u$  between  $t$  and  $t_0$ . By continuity,  $f^{(k)}(u)$  has the same sign as  $f^{(k)}(t_0)$  for all  $u$  close enough to  $t_0$ . Meanwhile,  $(t - t_0)^k$  changes sign at  $t_0$  only if  $k$  is odd.  $\square$

Once we know all the zeros of  $f$ , with their orders, Lemma 2.5 enables us to determine its sign at all points. In particular, if all the zeros of  $f$  are simple, then the sign of  $f$  alternates on the intervals between them.

### 3. Descartes' rule of signs

Let  $(a_j) = (a_1, a_2, \dots, a_n)$  be a finite sequence of real numbers. We denote by  $S[(a_j)]$  the *number of sign changes* in the sequence, in other words, the number of terms that have the opposite sign to the previous term (leaving out any zero terms). For example,  $(2, -1, 1, -2, -2, 1, -1, 2)$  has six sign changes. Three obvious facts:

- (i)  $S(a_1, \dots, a_n) \leq n - 1$ ;
  - (ii)  $S(a_n, \dots, a_1) = S(a_1, \dots, a_n)$ ;
  - (iii)  $S[(a_j)]$  is *even* if  $a_n$  has the same sign as  $a_1$ , *odd* if it has the opposite sign.
- (This is like crossing a river an even or odd number of times!)

We now prove Descartes' rule of signs for Dirichlet and generalized polynomials.

**Theorem 3.1.** *Let  $F$  be defined by (2), and  $f$  by (1), with  $p_1 > \dots > p_n$ . Then  $Z(F)$  and  $Z(f)$  are not greater than  $S[(a_j)]$ .*

*Proof.* We prove the statement for  $F$  first. The proof is by induction on the number of sign changes. If this number is zero, then all the  $a_j$  have the same sign (say  $a_j > 0$ ): then, clearly,  $F(x) > 0$  for all  $x$ , so  $Z(F) = 0$ .

So assume that the statement is true when there are  $m$  sign changes, and suppose that  $S[(a_j)] = m + 1$ . Let one of the sign changes (say the last one) occur at the term  $a_k$ , so that  $a_k$  has the opposite sign to  $a_{k-1}$ . Choose  $p$  such that  $p_k < p < p_{k-1}$ , and let

$$F_0(x) = e^{-px} F(x) = \sum_{j=1}^n a_j e^{(p_j - p)x}.$$

By Lemma 2.3,  $F_0$  has the same zeros (with the same orders) as  $F$ . Now

$$F_0'(x) = \sum_{j=1}^n (p_j - p) a_j e^{(p_j - p)x}.$$

Since  $p_{k-1} - p > 0$  and  $p_k - p < 0$ , it is clear that  $(p_j - p)a_j$  does *not* change sign at  $j = k$ ; otherwise, it has the same sign changes as  $(a_j)$ . So  $[(p_j - p)a_j]$  has  $m$  sign changes, and, by the induction hypothesis,  $Z(F_0') \leq m$ . By Proposition 2.1, it follows that  $Z(F) = Z(F_0) \leq m + 1$ , as required.

The statement for  $f$  follows, by Lemma 2.4. Alternatively, one could give a direct proof on the above lines, with  $t^{-p}f(t)$  replacing  $F_0(x)$ .  $\square$

The theorem has numerous consequences and applications. Firstly, it applies to ordinary polynomials, thereby giving Descartes' original result. The algebraic proof based on

factorization ([Hen], section 6.2) is no shorter. For ordinary polynomials, one can apply the rule again to  $f(-t)$  to obtain a bound for the number of *negative* zeros.

*Example 2.* Let  $f(t) = t^{12} - 6t^5 + 4$ . Then  $f$  has at most two positive zeros. In fact, the intermediate value theorem shows that it has exactly two, since  $f(0) = 4$ ,  $f(1) = -1$  and  $f(2) > 0$ . Also,  $f(-t)$  has entirely positive coefficients, so  $f$  has no negative zeros.

The *length* of a generalized or Dirichlet polynomial is the number of non-zero terms in its defining expression (1) or (2). Note that an ordinary polynomial of degree  $n$  may have length  $n + 1$  (it may also be much less, as in Example 2). By remark (i) above, we have:

**Corollary 3.2.** *A generalized polynomial, or a Dirichlet polynomial, of length  $n$ , has at most  $n - 1$  zeros (counted with their orders).  $\square$*

To prove this statement directly, without bothering with sign changes, one would modify the proof of Theorem 3.1 by taking  $F_0(x) = e^{-p_k x} F(x)$ , so that one term becomes constant and  $F'_0$  has length  $n - 1$ .

**Corollary 3.3.** *Let  $F$  be defined by (2), with  $S[(a_j)] = m$ . Then  $F$  assumes any particular value at most  $r$  times, where  $r = \min(m + 2, n)$ . Similarly for the function  $f$  defined by (1). If  $p_1 \leq 0$  or  $p_n \geq 0$ , the same holds with  $r = m + 1$ .*

*Proof.* Choose  $c$ , and let  $G(x) = F(x) - c = F(x) - ce^0$ . Then  $G$  is a Dirichlet polynomial of length  $n + 1$  (or still  $n$  if one of the  $p_j$  is 0). Also, when  $-c$  is inserted into the sequence  $(a_j)$  in the appropriate position, it introduces at most two new sign changes, or at most one if  $p_1 \leq 0$  or  $p_n \geq 0$ , since it is then in the first or last position (then recall that  $m + 1 \leq n$ ).  $\square$

In the usual way, we can reformulate Corollary 3.2 in terms of linear algebra:

**Theorem 3.4.** *Given distinct real numbers  $p_i$  ( $1 \leq i \leq n$ ) and distinct real numbers  $x_j$  ( $1 \leq j \leq n$ ), the matrix with entries  $e^{p_i x_j}$  is non-singular.*

*Proof.* Let  $u_i$  be row  $i$  of the matrix. Suppose that  $\sum_{i=1}^n a_i u_i = 0$ , with some  $a_i \neq 0$ . Let  $F(x) = \sum_{i=1}^n a_i e^{p_i x}$ . Then the length of  $F$  is  $n$  (or less, if some of the  $a_i$  are zero), and  $F(x_j) = 0$  for each  $j$ . By Corollary 3.2, this is not possible.  $\square$

**Corollary 3.5.** *Let distinct real numbers  $p_i$  ( $1 \leq i \leq n$ ) be given. Then there is a unique Dirichlet polynomial  $F(x) = \sum_{i=1}^n a_i e^{p_i x}$  taking specified values at  $n$  distinct points.*

*Proof.* This also follows from linear independence of the columns.  $\square$

The following further property is often included in the statement of Descartes' rule.

**Proposition 3.6.** *Let  $F$  be defined by (2). Then  $S[(a_j)] - Z(F)$  is an even integer.*

*Proof.* Suppose first that  $a_1$  and  $a_n$  have the same sign. Then  $S[(a_j)]$  is even. For sufficiently large  $x$ ,  $F(x)$  is dominated by the term  $a_1 e^{p_1 x}$ , so has the same sign as  $a_1$ . For sufficiently large  $-x$ , it has the same sign as  $a_n$ . So the total number of sign changes of  $F(x)$  is even. By Lemma 2.5, this means that there are an even number of zeros of odd order, so that  $Z(F)$  is even. If, instead,  $a_1$  and  $a_n$  have opposite signs, then both quantities are odd. In either case, the difference is even.  $\square$

*Two further remarks on Theorem 3.1.* (1) If the formula defining  $F(x)$  or  $f(t)$  is replaced by an infinite series of the same type, then the result still holds (with the same proof) for  $x$  or  $t$  within the interval of convergence, assuming that the sequence  $(a_j)$  has only finitely many sign changes.

(2) As a function of a *complex* variable, a Dirichlet polynomial can have infinitely many zeros: for example,  $e^x - 1 = 0$  when  $x = 2n\pi i$  for any integer  $n$ .

Among Dirichlet polynomials, a special case of particular interest is when each  $a_j$  is either 1 or  $-1$ , with equally many of each occurring. We will call this type *bipartite*. Note that this implies  $F(0) = 0$ . One such example, for those who know about such things, is given by the periodic blocks of a real Dirichlet  $L$ -series. Another (with a change of notation) is as follows. Given a vector  $b = (b_1, \dots, b_n)$  (with  $b_j \geq 0$  for all  $j$ ) and  $p \geq 1$ , its  $\ell_p$ -norm is  $\|b\|_p = (\sum_{j=1}^n b_j^p)^{1/p}$ . (This is a scale of alternative measurements of "length":  $p = 2$  gives the usual "Euclidean" length). Given now two vectors  $b$  and  $c$ , which has larger  $\ell_p$ -norm? The answer may be different for different  $p$ . The comparison is represented by the bipartite Dirichlet polynomial  $F(p) = \sum_{j=1}^n (b_j^p - c_j^p)$ .

*Example 3.* Let  $b = (5, 2, 2)$  and  $c = (4, 4, 1)$ . Then (with notation as above),

$$F(p) = 5^p - 2 \cdot 4^p + 2 \cdot 2^p - 1^p,$$

with three sign changes (to show it explicitly as bipartite, we would write  $4^p + 4^p$  instead of  $2 \cdot 4^p$ ). Simple calculation shows that  $\sum_{j=1}^3 b_j = \sum_{j=1}^3 c_j$  and  $\sum_{j=1}^3 b_j^2 = \sum_{j=1}^3 c_j^2$ , so that  $F(0) = F(1) = F(2) = 0$ . By Theorem 3.1, these are the only zeros of  $F$ , and they are all simple zeros. Since  $F(3) = 12 > 0$ , we can deduce that  $F(p) > 0$  on  $(2, \infty)$  and  $(0, 1)$ , and  $F(p) < 0$  on  $(-\infty, 0)$  and  $(1, 2)$ . In terms of  $\ell_p$ -norms,  $\|b\|_p < \|c\|_p$  for  $1 < p < 2$  and  $\|b\|_p > \|c\|_p$  for  $p > 2$ .

#### 4. Bounds in terms of the sequence of partial sums of $(a_j)$

Write  $A_j = a_1 + a_2 + \cdots + a_j$  (and similarly for another sequence  $(b_j)$ ). In this section, we present two powerful variations of Descartes' rule, relating the number of zeros to  $S[(A_j)]$  instead of  $S[(a_j)]$ . As the following example shows,  $S[(A_j)]$  can be much smaller.

*Example 4.* Consider our previous example  $(a_j) = (2, -1, 1, -2, -2, 1, -1, 2)$ , which has six sign changes. Then  $(A_j)$  is  $(2, 1, 2, 0, -2, -1, -2, 0)$ , with one sign change.

The first theorem requires the extra condition  $A_n = 0$ . This is clearly satisfied in the bipartite case. With  $f$  and  $F$  defined by (1) and (2), we have  $A_n = f(1) = F(0)$ , so the condition  $A_n = 0$  means that  $f(1) = 0$  or  $F(0) = 0$ . If the condition is not satisfied, but  $f(t_0)$  or  $F(x_0)$  is zero at another point  $t_0$  or  $x_0$ , then the substitution  $f_1(t) = f(tt_0)$  or  $F_1(x) = F(x + x_0)$  gives another function of the same sort satisfying the condition.

We follow the elegant method outlined in [P-Sz] (Part V, exercises 80, 83). First, some elementary facts about  $S[(A_j)]$  in general.

**Lemma 4.1.** *In all cases,  $S[(A_j)] \leq S[(a_j)]$ . If  $A_n = 0$ , then  $S[(a_j)] - S[(A_j)]$  is an odd integer, so  $S[(A_j)] \leq S[(a_j)] - 1$ .*

*Proof.* Suppose that  $(A_j)$  has  $k$  sign changes, at  $r_1, r_2, \dots, r_k$ , and that  $A_1 > 0$ . Then  $A_{r_1} < 0$ ,  $A_{r_2} > 0$ , and so on. Also,  $A_{r_1-1} \geq 0$ , so  $a_{r_1} < 0$ . Similarly,  $a_{r_2} > 0$ , and so on. So  $(a_j)$  has at least one sign change between 1 and  $r_1$  (possibly more, and not necessarily at  $r_1$  itself!). Again, it has at least one sign change between  $r_1$  and  $r_2$ , and so on. Hence  $S[(a_j)] \geq k$ .

If  $A_n = 0$ , then  $A_{n-1} = -a_n$ , while  $A_1 = a_1$ , so  $S[(A_j)]$  must differ from  $S[(a_j)]$  by an odd integer.  $\square$

**Lemma 4.2.** *Suppose that  $A_n = 0$  and let  $(b_j)$  be  $(a_j)$  in reverse order, so that  $(b_1, b_2, \dots, b_n) = (a_n, a_{n-1}, \dots, a_1)$ . Then  $S[(B_j)] = S[(A_j)]$ .*

*Proof.* We have  $B_j + A_{n-j} = a_1 + a_2 + \cdots + a_n = 0$ , hence

$$(B_1, B_2, \dots, B_n) = (-A_{n-1}, -A_{n-2}, \dots, -A_1, 0),$$

which clearly has the same number of sign changes as  $(A_j)$ .  $\square$

Nothing of the sort is true when  $A_n \neq 0$ , as we shall see in examples below.

**Lemma 4.3.** *If  $(a_j)$  is bipartite of length  $2n$ , then  $S[(A_j)] \leq n - 1$ .*



*Proof.* Suppose that  $(A_j)$  has a sign change at  $j = k$ . Since each  $a_j$  is 1 or  $-1$ , this means that  $A_{k-1} = 0$ , and the next sign change cannot occur before  $j = k + 2$ . The first sign change cannot occur until  $j = 3$ , so the total number is at most  $n - 1$ . This number occurs when  $(a_j)$  consists of pairs  $(1, -1)$  alternating with  $(-1, 1)$ .  $\square$

The starting point for the stated result is the following useful identity, known as *Abel summation*. Note that  $a_1 = A_1$  and  $a_j = A_j - A_{j-1}$  for  $j \geq 2$ . Substituting this, we obtain

$$\begin{aligned} a_1b_1 + a_2b_2 + \cdots + a_nb_n &= A_1b_1 + (A_2 - A_1)b_2 + \cdots + (A_n - A_{n-1})b_n \\ &= \sum_{j=1}^{n-1} A_j(b_j - b_{j+1}) + A_nb_n. \end{aligned} \quad (3)$$

When  $A_n = 0$ , the final term in (3) disappears.

This time there is a much easier proof for ordinary polynomials, so we present it first. For this purpose, relying on Lemma 4.2, we list the powers  $t^j$  in the usual *increasing* order. Also, the first term is  $a_0$ , so  $A_j$  means  $a_0 + a_1 + \cdots + a_j$ .

**Proposition 4.4.** *Let  $f(t) = \sum_{j=0}^n a_j t^j$ , with  $A_n = 0$ . Then  $Z[f, (0, \infty)] \leq S[(A_j)] + 1$ .*

*Proof.* By (3), we have

$$f(t) = \sum_{j=0}^{n-1} A_j(t^j - t^{j+1}) = (1 - t)g(t),$$

where  $g(t) = \sum_{j=0}^{n-1} A_j t^j$ . By Descartes' rule,  $Z(g) \leq S[(A_j)]$ . Now  $1 - t$  has a simple zero at 1, and no other zeros, so Lemma 2.3 gives  $Z(f) = Z(g) + 1$ .  $\square$

*Example 5.* Let

$$f(t) = 2t^{12} - t^{10} + t^9 - 2t^8 - 2t^5 + t^3 - t + 2.$$

As we saw in Example 4,  $(A_j)$  has one sign change. Hence  $Z[f, (0, \infty)] \leq 2$ . In fact, it equals 2, since  $f(0) = 2$ ,  $f(1) = 0$  and  $f'(1) > 0$ , so that  $f(t) < 0$  for  $t$  just less than 1.

We now turn to the general case. Let  $F(x)$  be defined by (2), with  $p_1 > \cdots > p_n$ . By (3), together with the condition  $A_n = 0$ , we have

$$F(x) = \sum_{j=1}^{n-1} A_j(e^{p_j x} - e^{p_{j+1} x}). \quad (4)$$

If  $(A_j)$  has no sign changes, so that (say)  $A_j \geq 0$  for each  $j$ , this shows that  $F(x) > 0$  for  $x > 0$  and  $F(x) < 0$  for  $x < 0$ , so that  $F$  has no zeros except the one at 0.

For fixed  $x \neq 0$ , we can rewrite (4) as an integral with respect to another variable  $t$ , using the fact that  $\frac{d}{dt}e^{tx} = xe^{tx}$ :

$$F(x) = x \sum_{j=1}^{n-1} A_j \int_{p_{j+1}}^{p_j} e^{tx} dt. \quad (5)$$

Equation (5) also holds when  $x = 0$ , because  $F(0) = A_n = 0$ . We can rewrite the identity again as follows:  $F(x) = xG(x)$ , where

$$G(x) = \int_{p_n}^{p_1} \phi(t)e^{tx} dt, \quad (6)$$

in which

$$\phi(t) = A_j \quad \text{for } p_{j+1} < t < p_j \quad (1 \leq j \leq n-1). \quad (7)$$

It doesn't matter what values are assigned to  $\phi(t)$  at the points  $p_j$ , because the value at one point does not affect the integral. In terms of the values assumed on successive intervals  $(p_{j+1}, p_j)$ , the function  $\phi$  has sign changes exactly corresponding to those of  $(A_j)$ . Of course,  $\phi$  is discontinuous at the points  $p_j$ , but continuous (in fact, constant) on the open intervals between them.

What we now need is a general result analogous to Theorem 3.1 for functions defined by an integral in this way. We change the notation slightly for this purpose. In fact, suppose that  $\phi$  is a function on  $(a, b)$  such that

- (A) there exist points  $a = t_0 < t_1 < \dots < t_n = b$  such that  $\phi$  is bounded, continuous and non-zero on each open interval  $(t_{j-1}, t_j)$ ,

We count the point  $t_j$  as a sign change of  $\phi$  if it has opposite signs on  $(t_{j-1}, t_j)$  and  $(t_j, t_{j+1})$ .

**Lemma 4.5.** *Suppose that  $\phi$  satisfies (A) and has  $m$  sign changes in  $(a, b)$ . Let*

$$G(x) = \int_a^b \phi(t)e^{tx} dt \quad (x \in \mathbb{R}).$$

*Then  $Z(G) \leq m$ .*

*Proof.* Induction on  $m$ , copying the proof of Theorem 3.1. If  $m = 0$ , then either  $\phi(t) > 0$  or  $\phi(t) < 0$  on each interval  $(t_{j-1}, t_j)$ : assume the first. Condition (A) now ensures that  $G(x) > 0$  for all  $x$ , so  $Z(G) = 0$ .

Now assume that the statement is correct for a certain value  $m$  and that  $\phi$  has  $m + 1$  sign changes. Let one of them be at  $t_k$ , and let

$$G_0(x) = e^{-t_k x} G(x) = \int_a^b \phi(t)e^{(t-t_k)x} dt.$$

By differentiation under the integral sign (but see the note below!),

$$G'_0(x) = \int_a^b (t - t_k)\phi(t)e^{(t-t_k)x} dt.$$

The function  $(t - t_k)\phi(t)$  still satisfies condition (A), and does not have a sign change at  $t_k$  (because  $t - t_k$  changes sign there), so it has  $m$  sign changes. By the induction hypothesis,  $Z(G'_0) \leq m$ . Rolle's theorem gives the required statement  $Z(G) = Z(G_0) \leq m + 1$ .  $\square$

*Note.* The general theorem on differentiation under the integral sign is distinctly more advanced than everything else we have used. What is really required here is only the following special case: if  $G(x) = \int_0^a t^n e^{tx} dt$ , then  $G'(x) = \int_0^a t^{n+1} e^{tx} dt$ . One can verify this directly, as follows (we leave the details to the sufficiently determined reader). Substitute  $tx = u$  to get  $G(x) = x^{-n-1} \int_0^{ax} u^n e^u du$ . Differentiate by the product rule, using the fundamental theorem of calculus in the form  $\frac{d}{dx} \int_0^x f(u) du = f(x)$ . Use integration by parts to show that the supposed expression for  $G'(x)$  (after substituting  $tx = u$  again) agrees. There is no need for an explicit evaluation of the integral defining  $G(x)$ !

Putting all this together, we have our promised theorem:

**Theorem 4.6.** *Let  $F$  be defined by (2), and  $f$  by (1), with  $p_1 > \dots > p_n$  and  $A_n = 0$ . Then  $Z(F)$  and  $Z(f)$  are not greater than  $S[(A_j)] + 1$ .*

*Proof.* Recall that  $F(x) = xG(x)$ , where  $G(x)$  is defined by (6) and  $\phi(t)$  by (7). By Lemma 4.5,  $Z(G) \leq S[(A_j)]$ . The factor  $x$  is zero at 0, so, by Lemma 2.3, has the effect that  $Z(F) = Z(G) + 1$ .  $\square$

*Example 6.* Let  $G(p) = \sum_{j=1}^4 (b_j^p - c_j^p)$ , where  $b = (9, 6, 5, 2)$  and  $c = (8, 8, 3, 3)$ . Then  $G$  is bipartite of length 8, so our theorems show that  $Z(G) \leq 4$ . It is easy to check the remarkable fact that  $\sum_{j=1}^4 b_j^r = \sum_{j=1}^4 c_j^r$  for  $r = 0, 1, 2, 3$ , so that  $G(0) = G(1) = G(2) = G(3) = 0$ . So, by Theorem 4.6, these are the only zeros of  $G$ , and they are simple, so the sign of  $G(p)$  alternates on the intervals between them.

Not all is lost when  $A_n \neq 0$ . Retracing our steps, we obtain the following variant of the theorem, counting *positive* zeros for  $F$  and zeros in  $(1, \infty)$  for  $f$ .

**Theorem 4.7.** *Let  $F$  be defined by (2), and  $f$  by (1), with  $p_1 > \dots > p_n$ . Then (without the condition  $A_n = 0$ ),  $Z[F, (0, \infty)]$  and  $Z[f, (1, \infty)]$  are not greater than  $S[(A_j)]$ .*

*Proof.* In (4), we must add the term  $A_n e^{p_n x}$ . For  $x > 0$ , this can be written as  $x \int_{-\infty}^{p_n} e^{tx} dt$ , which must be added to the expression in (5). Then continue as before, but the factor  $x$  no longer adds 1, since 0 is not in  $(0, \infty)$ .  $\square$

By applying this theorem to  $F(-x)$  or  $f(1/t)$ , we obtain at once the following corollary, in which the order of the  $a_j$  is reversed (which amounts to considering the  $p_j$  in *increasing* order):

**Corollary 4.8.** *Let  $F$  and  $f$  be as in Theorem 4.7. Let  $(b_1, b_2, \dots, b_n) = (a_n, a_{n-1}, \dots, a_1)$ . Then  $Z[F, (-\infty, 0)]$  and  $Z[f, (0, 1)]$  are not greater than  $S[(B_j)]$ .  $\square$*

By substituting  $F_1(x) = F(x+r)$  and  $f_1(t) = f(rt)$ , we can apply Theorem 4.7 to give bounds for  $Z[F, (r, \infty)]$  and  $Z[f, (r, \infty)]$ : the coefficients  $a_j$  are replaced by  $a_j e^{rp_j}$  and  $a_j r^{p_j}$  in the two cases.

Theorem 4.7 and Corollary 4.8 can be applied to  $F(x) - c$  to give bounds for the number of times a non-zero value is assumed (but Theorem 4.6 can't, because the condition  $A_n = 0$  does not survive the insertion of an extra term). The next example illustrates these results, using an ordinary polynomial.

*Example 7.* Let

$$f(t) = (t - \frac{1}{10})(t - \frac{2}{10})(t - \frac{3}{10})(t - \frac{4}{10}) = t^4 - t^3 + \frac{7}{20}t^2 - \frac{1}{20}t + \frac{24}{10^4},$$

with four zeros in  $(0, 1)$  and none in  $(1, \infty)$ . When the terms are taken in descending order,  $A_j \geq 0$  for all  $j$ , but when they are taken in ascending order, the partial sums  $(B_j)$  have four sign changes. Further, let  $g(t) = t^4 - t^3 + \frac{7}{20}t^2 - \frac{1}{20}t - \frac{3}{10}$ . For  $g$ , we have  $A_j \geq 0$  for all  $j$  and  $A_5 = 0$ , so  $g$  has only one positive zero (at 1). However,  $g(t)$  assumes the value  $-\frac{3}{10} - \frac{24}{10^4}$  when  $f(t)$  is zero, hence four times.

*Example 8.* Let  $f(t) = t^6 - 5t^5 + 2t^3 - 2t^2 + 8$ . By 4.7 and 4.8,  $f$  has at most two zeros in  $(1, \infty)$  (in fact, there are two), and none in  $(0, 1)$ .

*Historical note.* Theorem 4.7 is stated explicitly in both [La] and [P-Sz], but Theorem 4.6 isn't. Laguerre's original proof of Theorem 4.7, for polynomials, can be seen in [Hou, Theorem 2.4.5]. For the extension to other functions, his reasoning seems to be incomplete, because it depends on an unexplained (and, to this author, unconvincing) limiting process [La, p. 9].

## REFERENCES

[Hen] P. Henrici, *Applied and Computational Complex Analysis*, vol. 1, Wiley, New York (1974).

[Hou] A.S. Householder, *The Numerical Treatment of a Single Nonlinear Equation*, McGraw Hill, New York (1970).

[La] E.N. Laguerre, Sur la théorie des équations numériques, *J. Math. Pures et Appl.* **9** (1883); (in *Oeuvres de Laguerre*, vol. 1, Paris (1898), 3–47).

[P-Sz] G. Pólya and G. Szegő, *Aufgaben und Lehrsätze aus der Analysis, Band II*, Springer-Verlag, Berlin (1925, 1964).

email: [g.jameson@lancaster.ac.uk](mailto:g.jameson@lancaster.ac.uk)