ANNALES ACADEMIAE SCIENTIARUM FENNICAE

Series A

I. MATHEMATICA

342

ON RUNGE'S GENERAL RULE OF SIGNS

BY

A. M. OSTROWSKI

HELSINKI 1964
SUOMALAINEN TIEDEAKATEMIA

https://doi.org/10.5186/aasfm.1964.342

Communicated 13 December 1963 by P. J. Myrberg and R. Nevanlinna

KESKUSKIRJAPAINO
HELSINKI 1964

On Runge's General Rule of Signs *

Introduction

1. For a finite sequence of real numbers $u_1, \dots, u_l$ we denote by the symbol $V(u_1, \dots, u_l)$ or $V(u_\lambda)$ the number of the variations of sign in this sequence. In computing this number all vanishing u_λ can be deleted, further the symbol has by definition the values 0 if only one or none of the u_λ is $\neq 0$.

We have then obviously

$$(1) \quad 1 \geqq V(u_1, \dots, u_k, \dots, u_l) - V(u_1, \dots, u_k) - V(u_k, \dots, u_l) \geqq 0 \qquad (1 \leqq k \leqq l).$$

In particular the difference in (1) has the value 0 if $u_k \neq 0$. Further we have, if $u_1 u_l \neq 0$, the relation

$$(2) \qquad V(u_1, \dots, u_l) \equiv \frac{1 - \operatorname{sgn} u_1 u_l}{2} \pmod 2,$$

which is verified immediately.

2. For a fixed real polynomial $f(x)$ of exact degree n we denote by V_a the number of the variations of sign in the complete sequence of the derivatives of f at a,

$$(3) \qquad V_a = V\big(f(a), f'(a), \dots, f^{(n)}(a)\big).$$

Then the famous Theorem of Budan—Fourier consists in the inequality

$$(4) \qquad V_a - V_b \geqq N(a, b) \qquad (a < b)$$

where under $N(a, b)$ is understood the total number of roots of $f(x)$ in the interval $a < x \leqq b$, counted with their multiplicities. As a matter of fact, the Theorem of Budan—Fourier contains a further assertion,

* This investigation was carried out under the contract DA-91-591-EUC-2824 of the US Army with the Mathematical Institute of the University of Basel.

namely, that the difference between both sides of the inequality (4) is *even*, so that we can write

$$V_a - V_b = N(a, b) + 2\,\Delta(a, b), \tag{5}$$

where $\Delta(a, b)$, the *BF-defect* of f, is a non-negative integer. However, this follows easily from (2).

3. Both $V_a - V_b$ and $N(a, b)$ are additive interval functions of the half-open interval $(a, b\rangle$, and the total value of $V_a - V_b$ over the whole real axis is n, while that of $N(a, b)$ is the total number of real roots of $f(x)$. Therefore, $\Delta(a, b)$ is a non-negative interval function whose value over the whole real axis is equal to the number of couples of conjugate non-real roots of $f(x)$.

4. From (5) it follows now that if $f(x)$ has only real roots we have equality in (4). However, the practical usefulness of the formula (5) with this interpretation of $\Delta(a, b)$ goes beyond the above corollary. Indeed, if by some special argument we can localize a set of intervals $J_\varkappa$ over which $\Delta(a, b)$ already assumes its maximal value it follows that for any interval $(a, b\rangle$ without any point in common with the intervals $J_\varkappa$ we have the formula (4) with the equality-sign. A corresponding partial improvement in the formula (4) is possible whenever we succeed to find a set of intervals $J_\varkappa$ over which $\Delta(a, b)$ is positive.

5. The numbers $f^{(\nu)}(a)$ used in (3) are essentially the coefficients of the Taylor development of $f(x)$ at a. This development is a special case of the general Newtonian development of $f(x)$ arising from the general Newton interpolation formula. Consider a sequence of n numbers $x_1, \ldots, x_n$ and put

$$P_\nu(\xi) = (\xi - x_1) \ldots (\xi - x_\nu) \qquad (\nu = 1, 2, \ldots, n), \qquad P_0(\xi) = 1. \tag{6}$$

Then the development in question is

$$f(\xi) = a_0 P_n(\xi) + a_1 P_{n-1}(\xi) + \ldots + a_n P_0(\xi), \tag{7}$$

where the coefficients a_ν are uniquely determined by $f(\xi)$ and the sequence x_ν.

The reader may be reminded here, although we will make no use of this in the following discussion, that the coefficients a_ν are obtained by forming Newton's devided differences both in the case of distinct x_ν and in the case that the values of some of the x_ν coincide.

6. We assume in what follows until the end of this introduction that the numbers x_ν in (7) are all *real* and consider another set of such real numbers $y_1, \dots, y_n$ and the corresponding polynomials

(8) $$Q_\nu(\xi) = (\xi - y_1) \dots (\xi - y_\nu) \quad (\nu = 1, 2, \dots, n), \qquad Q_0(\xi) = 1 .$$

Developping $f(\xi)$ in $Q_\nu(\xi)$ we have

(9) $$f(\xi) = b_0 Q_n(\xi) + b_1 Q_{n-1}(\xi) + \dots + b_n Q_0(\xi) .$$

Then *Runge's General Rule of Signs* is the

Theorem I. *Suppose that we have in (6) and (8)*

(10) $$a = \operatorname*{Max}_\nu x_\nu < b = \operatorname*{Min}_\nu y_\nu ;$$

then it follows for the coefficients a_ν *and* b_ν *in (7) and (9)*[1])

(11) $$V(a_\nu) - V(b_\nu) \geqq N(a, b) .$$

7. The special case of the Budan—Fourier Theorem is obtained from the Theorem I if all x_ν become $= a$ and all y_ν become $= b$.

Runge's method of proof of the Theorem I is of particular interest since it is based on the generalisation of the process of the so-called *synthetic division* to the polynomials given in the form (7) (Theorem III). On the other hand, the detailed working-out of this idea becomes in Runge's presentation very complicated, since Runge tries to prove everything anew inclusive the classical Budan—Fourier case. However, the proof can be considerably simplified if we assume the Budan—Fourier as given. Then Runge's Theorem follows immediately from the following *Fundamental Lemma:*

Theorem II. *Assume that we have in (6)—(9):*

(12) $$y_\mu \geqq x_\nu \qquad (\mu, \nu = 1, 2, \dots, n) ;$$

then it follows

(13) $$V(a_\nu) \geqq V(b_\nu) .$$

[1]) A special case of this Theorem corresponding in a certain sense to the assumption that either all x_ν are $= -\infty$ or all y_ν are $= \infty$ has been published with Runge's permission 1914 by G. Pólya (G. Pólya: Über einige Verallgemeinerungen der Descartesschen Zeichenregel. - Arch. Math. Phys. (3) 23, 1914, pp. 22—32) who gave an elegant direct proof in this case. Following up some indications by G. Pólya in the paper quoted, I unearthed the Theorem I from a manuscript of Runge's course given at the University of Göttingen in Summer 1907 and edited the corresponding parts of this manuscript with the necessary corrections and developments (see Carl Runge: Eine Vorzeichenregel in der Theorie der algebraischen Gleichungen. Aus einem Vorlesungsmanuskript von Carl Runge herausgegeben von Alexander Ostrowski. - Jber. Deutsch. Math. Verein. 66, 1963, pp. 52—66).

8. Beyond that, by means of the Fundamental Lemma we can make use of the advances which have been made in the direction of the Budan—Fourier Theorem in the last 100 years and obtain some corresponding generalisations in the case of the general Newton—Runge set-up. See in particular the Theorems V (sec. 21), VI (sec. 27), VII (sec. 32) and VIII (sec. 41).

Generalized Synthetic Division

9. If the polynomial $f(\xi)$ is given in the form (7) we have, for a parameter $y \neq \xi$ identically

$$\frac{f(\xi)}{\xi-y} = a_0' P_{n-1}(\xi) + a_1' P_{n-2}(\xi) + \ldots + a_{n-1}' P_0(\xi) + \frac{a_n'}{\xi-y} \tag{14}$$

where the a_ν' depend on y.

Multiplying this on both sides by $\xi-y$ we have

$$f(\xi) = \sum_{\nu=0}^{n-1} a_\nu' P_{n-1-\nu}(\xi)(\xi-y) + a_n' . \tag{15}$$

10. Putting in this identity $\xi = y$ we obtain

$$a_n' = f(y) . \tag{16}$$

On the other hand, decomposing the factor $\xi-y$ in the ν-th term of (15) in $(\xi-x_{n-\nu}) - (y-x_{n-\nu})$ we have

$$f(\xi) = \sum_{\nu=0}^{n} a_\nu' P_{n-\nu}(\xi) - \sum_{\nu=0}^{n-1} a_\nu' (y-x_{n-\nu}) P_{n-\nu-1}(\xi) , \tag{17}$$

where the term a_n' has been taken into the first right-hand sum as the term corresponding to $\nu = n$.

The second right-hand sum becomes, if we introduce $\nu+1$ as the new summation variable and denote it again by ν,

$$\sum_{\nu=1}^{n} a_{\nu-1}' (y-x_{n-\nu+1}) P_{n-\nu}(\xi) .$$

Introducing this into (17) and taking the term of the first sum corresponding to $\nu = 0$ separately we obtain

$$f(\xi) = a_0' P_n(\xi) + \sum_{\nu=1}^{n} (a_\nu' - (y-x_{n-\nu+1}) a_{\nu-1}') P_{n-\nu}(\xi) . \tag{18}$$

11. Comparing this development of $f(\xi)$ with (7) we obtain

(19) $$a'_0 = a_0\,, \qquad a'_\nu = a'_{\nu-1}\,(y - x_{n-\nu+1}) + a_\nu \qquad (\nu = 1\,,\,\ldots\,,\,n)\,.$$

We have now immediately the

Theorem III. *In the formula (14) each* a'_ν $(\nu = 0\,,\,\ldots\,,\,n)$ *is a polynomial of exact degree* ν *in* y *and these* a'_ν *can be computed recurrently by (19). In particular, we have* $a'_n = f(y)$.

The practical computation of $f(y)$ and of all other coefficients a'_ν can be carried out using Runge's diagram:

(20) $$\begin{array}{cccccccc} a_0 & a_1 & & a_\varkappa & & a_{n-1} & a_n \\ & a'_0\,(y-x_n) & \ldots & a'_{\varkappa-1}\,(y-x_{n-\varkappa+1}) & \ldots & a'_{n-2}\,(y-x_2) & a'_{n-1}\,(y-x_1) \\ \hline a'_0 & a'_1 & \ldots & a'_\varkappa & \ldots & a'_{n-1} & a'_n \end{array}$$

Monotony of the variation number in Synthetic Division

12. We assume now that all x_ν are real and that the parameter y in (14) satisfies the condition

(21) $$y \geqq x_\nu \qquad (\nu = 1\,,\,2\,,\,\ldots\,,\,n)\,.$$

Under this hypothesis we will then compare $V(a'_\nu)$ with $V(a_\nu)$. We use for this purpose the following

Lemma. *If in a sequence* $u_0\,.\,u_1\,,\,\ldots\,,\,u_n$ *we replace for a* $\nu \geqq 1$ *and a* $p \geqq 0$ *the element* u_ν *by*

(22) $$u'_\nu = p\,u_{\nu-1} + u_\nu \qquad (\nu \geqq 1\,,\ \ p \geqq 0)\,,$$

the number of the variations of sign in our sequence does not increase.

13. *Proof.* Indeed, the assertion is evident if $u_\nu = 0$, since then u'_ν is either $= 0$ or has the same sign as $u_{\nu-1}$. Further, if $u_\nu \neq 0$ and u'_ν has the same sign as u_ν the assertion is evident too.

On the other hand, if u_ν is $\neq 0$ and u'_ν is either 0 or has the opposite sign to that of u_ν, then we must have a variation of sign between $u_{\nu-1}$ and u_ν which is lost if we replace u_ν by u'_ν while to the right of u_ν at the most one variation of sign could be won. This proves our Lemma.

14. We consider now a real sequence $c_0\,.\,c_1\,,\,\ldots\,,\,c_n$ undergoing the transformation analogous to (19)

(23) $$c'_0 = c_0\,, \qquad c'_\nu = c'_{\nu-1}\,(y - x_{n-\nu+1}) + c_\nu \qquad (\nu = 1\,,\,\ldots\,,\,n)$$

into the sequence

(24) $$c_0', c_1', \ldots, c_n'.$$

This transformation can be considered as effected by means of n consecutive transformations $T_\varkappa$ $(\varkappa = 1, \ldots, n)$ given by

(25) $$T_\varkappa \quad \left\{ \begin{array}{cccccccc} c_0' & c_1' & \ldots & c_{\varkappa-1}' & c_\varkappa & c_{\varkappa+1} & \ldots & c_n \\ c_0' & c_1' & \ldots & c_{\varkappa-1}' & c_\varkappa' & c_{\varkappa+1} & \ldots & c_n \end{array} \right. .$$

The transformation $T_\varkappa$ consists in replacing the one element $c_\varkappa$ by $c_\varkappa' = c_{\varkappa-1}'(y - x_{n-\varkappa+1}) + c_\varkappa$.

Under the condition (21) the Lemma of the section 12 can be applied then to the transformation $T_\varkappa$ and we see that the number of the variations of sign is not increased by this transformation.

15. But then the same holds for the transformation of all c_ν into the c_ν' and even, more generally, for the transformation of the first row of

(26) $$\left\{ \begin{array}{ccccccc} c_0 & c_1 & \ldots & c_\varkappa & c_{\varkappa+1} & \ldots & c_n \\ c_0' & c_1' & \ldots & c_\varkappa' & c_{\varkappa+1} & \ldots & c_n \end{array} \right. \qquad (\varkappa = 1, \ldots, n)$$

into the second row.

We obtain the

Theorem IV. *If the sequence* $c_0, c_1, \ldots, c_n$ *is transformed by (23) into the sequence* $c_0', c_1', \ldots, c_n'$, *the number of the variations of sign in the sequence* $c_0, c_1, \ldots, c_n$ *is not smaller than the number of the variations of sign in the second row of (26), for* $\varkappa = 1, 2, \ldots, n$.

Proof of the Fundamental Lemma

16. We go from (7) to (9) in n single steps, dividing $f(\xi)$ for $\nu = 0, 1, \ldots, n$ by $Q_\nu(\xi)$:

(27) $$f = A_\nu Q_\nu + B_\nu \qquad (\nu = 0, 1, \ldots, n; \ B_0 \equiv 0),$$

where the degree of $B_\nu(\xi)$ is $\leqq \nu - 1$ and that of $A_\nu(\xi)$ exactly $= n - \nu$.

Develop, putting $n - \nu = m$, $A_\nu(\xi)$ in $P_\varkappa$ and $B_\nu(\xi)$ in $Q_\varkappa$:

$$A_\nu = \alpha_0 P_m + \ldots + \alpha_m P_0, \quad B_\nu = \beta_0 Q_{\nu-1} + \ldots + \beta_{\nu-1} Q_0,$$

and consider the sequence

(28) $$\alpha_0, \ldots, \alpha_{m-1}, \alpha_m, \beta_0, \ldots, \beta_{\nu-1} \qquad (n - \nu = m).$$

For $\nu = 0$, (28) coincides with (a_ν) in (7), and for $\nu = n$ with (b_ν) in (9). It is therefore sufficient to show that the number of variations of sign in (28) does not increase if we go from ν to $\nu+1$.

17. Indeed, dividing $A_\nu(\xi)$ by $\xi - y_{\nu+1}$ we can write

$$A_\nu = (\xi - y_{\nu+1}) A'_{\nu+1} + \alpha'_m , \tag{29}$$

and, putting this into (27),

$$f = A'_{\nu+1} Q_{\nu+1} + \alpha'_m Q_\nu + B_\nu ,$$

so that we have $B_{\nu+1} = \alpha'_m Q_\nu + B_\nu$, $A'_{\nu+1} = A_{\nu+1}$.

If we put then

$$A'_{\nu+1} = \alpha'_0 P_{m-1} + \ldots + \alpha'_{m-1} ,$$

the sequence (28) corresponding to $\nu+1$ becomes

$$\alpha'_0 , \ldots , \alpha'_{m-1} , \alpha'_m , \beta_0 , \beta_1 , \ldots , \beta_{\nu-1} . \tag{30}$$

And here the α'_μ are obtained from the α_μ by (29), that is by Synthetic Division.

If we identify the sequence (28) with the sequence $c_0, c_1, \ldots, c_n$ in (23) and apply to the c the transformation (23) with $y = \xi$ and $x_{n-\nu+1} = y_{\nu-1}$, we obtain a sequence $c'_0, c'_1, \ldots, c'_n$ and the sequence (30) is identical with

$$c'_0 , \ldots , c'_m , c_{m+1} , \ldots , c_n .$$

From the Theorem IV, applied for $\varkappa = m$, it follows then that the number of the variations of sign is not increased, going from (28) to (30), and the Theorem II is proved.

Proof of Theorem I

18. We have, under the conditions of the Theorem I, if the developments (7) and (9) are compared with the Taylor developments at a and b, by the Theorem II:

$$V(a_\nu) \geqq V_a \geqq V_b \geqq V(b_\nu) ,$$

where V_a, V_b are defined by (3).

We have therefore

$$V(a_\nu) - V(b_\nu) \geqq V_a - V_b ,$$

and (11) follows immediately from (4). The Theorem I is proved.

The Parity Discussion

19. The inequality (11) contains though (4) but, of course, not the relation (5). It is therefore of interest to discuss under what conditions both sides of the inequality (11) have the same parity:

$$V(a_\nu) - V(b_\nu) \equiv N(a\,,b) \pmod{2}\,. \tag{31}$$

We assume in this discussion that $a_n\, b_n \neq 0$ that is

$$f(x_1) \neq 0\,, \qquad f(y_1) \neq 0\,. \tag{32}$$

Further it can be assumed without loss of generality that $a_0 = b_0 \neq 0$.

20. Then we have from (2)

$$V(a_\nu) - V(b_\nu) \equiv \frac{\operatorname{sgn} a_n - \operatorname{sgn} b_n}{2} \pmod{2}\,.$$

On the other hand, obviously

$$\frac{\operatorname{sgn} a_n - \operatorname{sgn} b_n}{2} = \frac{\operatorname{sgn} f(x_1) - \operatorname{sgn} f(y_1)}{2} \equiv N(x_1\,,y_1) \pmod{2}\,.$$

and therefore

$$V(a_\nu) - V(b_\nu) \equiv N(x_1\,,y_1) \pmod{2}\,. \tag{33}$$

We see that necessary and sufficient for (31) is the congruence

$$N(x_1\,,y_1) \equiv N(a\,,b) \pmod{2}\,,$$

that is

$$N(x_1\,,a) \equiv N(b\,,y_1) \pmod{2}\,. \tag{34}$$

21. (34) is certainly satisfied if we have in particular

$$N(x_1\,,a) = N(b\,,y_1) = 0\,. \tag{35}$$

A more symmetric sufficient condition is obtained if we introduce the smallest closed interval $\langle x_1\,,\ldots,x_n\rangle$ containing all $x_1\,,\ldots,x_n$ and the correspondingly defined interval $\langle y_1\,,\ldots,y_n\rangle$; then we have the

Theorem V. *Under the conditions of Theorem I and in the hypothesis (32), the condition (34) is necessary and sufficient for the relation (31). In particular, the relation (31) certainly holds, if the intervals* $\langle x_1\,,\ldots,x_n\rangle$ *and* $\langle y_1\,,\ldots,y_n\rangle$ *are free of the roots of* $f(x)$.

22. As the relation (5) holds also if $f(a)$ or $f(b)$ vanishes one could expect that the condition (32) is not essential in the above discussion. However, the following counter example shows that the analogy can break down if (32) is not satisfied.

Consider the polynomial

$$f(\xi) = \xi (\xi-1)(\xi-3)$$

and take

$$x_1 = x_2 = x_3 = -1; \quad y_1 = 0, \quad y_2 = y_3 = 2.$$

Then we have on the one side $N(-1, 0) = 1$. On the other side, we have from the Taylor development at -1: $V(a_\nu) = 3$, while the development corresponding to (9) is

$$f = Q_3 - Q_1 = \xi (\xi-2)^2 - \xi$$

and therefore $V(b_\nu) = 1$. We see that in this case both sides of (31) have not the same parity.

Specialisations of Theorem I

23. Take in the formula (11) all b_ν equal, $b_1 = b_2 = \ldots = b_n = b$ with a b greater than all roots of $f(\xi), f'(\xi), \ldots, f^{(n-1)}(\xi)$. Then we have

$$V(b_\nu) = V_b = 0, \qquad N(a, b) = N(a, \infty)$$

and (11) becomes

$$N(a, \infty) \leqq V(a_\nu), \tag{36}$$

the special case of Runge's Rule of Signs published and proved directly in G. Pólya's above mentioned article. From our discussion in the sections 19—21 it follows further that if $f(x_1) \neq 0$, then the necessary and sufficient conditions for the both sides of (36) to have the same parity is

$$N(x_1, a) \equiv 0 \pmod 2. \tag{37}$$

This is in any case satisfied if we have

$$x_1 \geqq x_\nu \quad (\nu = 2, \ldots, n), \qquad f(x_1) \neq 0. \tag{38}$$

24. In the completely symmetric manner we obtain another specialisation of Runge's Rule (cf. Pólya's above mentioned article)

$$N(-\infty, b) \leqq V(b_\nu), \tag{39}$$

where the b_ν are given by (9) and b by (10). If $f(y_1) \neq 0$, the condition

$$N(b, y_1) \equiv 0 \pmod 2 \tag{40}$$

is necessary and sufficient in order that both sides of (39) have the same parity. And this is certainly the case if

$$y_1 \leqq y_\nu \quad (\nu = 2, \ldots, n), \qquad f(y_1) \neq 0 .$$

25. Consider now the case where for a fixed $h > 0$

$$x_1 = p, \qquad x_2 = p + h, \quad \ldots, \quad x_n = p + (n-1)\,h = a,$$
$$y_1 = q = b, \quad y_2 = q + h, \quad \ldots, \quad y_n = q + (n-1)\,h .$$

Since we have the general Newton—Gregory formula

$$f(\xi) = \sum_{\nu=0}^{n} \frac{\Delta^\nu f(\xi)}{\nu!\, h^\nu} (\xi-\alpha)(\xi-\alpha-h)\ldots(\xi-\alpha-(\nu-1)\,h),$$

where the differences $\Delta^\nu f(\alpha)$ correspond to the step h, it follows for the coefficients in the developments (6) and (8):

$$a_{n-\nu} = \frac{\Delta^\nu f(p)}{\nu!\, h^\nu}, \qquad b_{n-\nu} = \frac{\Delta^\nu f(q)}{\nu!\, h^\nu}$$

We obtain now from the Theorem I the

Corollary. *If* $q - p > (n-1)\,h$ *for a positive* h *then we have*

$$N(p + (n-1)\,h, q) \leqq V(\Delta^\nu f(p)) - V(\Delta^\nu f(q)) \qquad (q-p > (n-1)\,h). \tag{41}$$

In the special case that the difference $q-p$ is a multiple of h this result has been stated by N. Obreschkoff (N. Obreschkoff: Sur les racines des équations algébriques. - Annuaire Univ. Sofia Fac. Sci. Phys. Math. Livre 1 Math. 23, 1927, pp. 177—200; cf. also Obreschkoff's book: Verteilung und Berechnung der Nullstellen reeller Polynome. - Hochschulbücher für Mathematik 55, Deutscher Verlag der Wissenschaften, Berlin, 1963, pp. 116—118).

Reality Discussion

26. It has been already mentioned in the section 4 that we have equality in the Budan—Fourier formula (4) if $f(\xi)$ has only real roots. Since the inequality (11) of Runge's Rule has been obtained from (4) it can be expected that for polynomials with only real roots the formula (11) could be proved *with the equality-sign* in some important cases.

As a matter of fact, more general results can be derived if we use the formula (5) and restrict ourselves to the subintervals of the real axis on which Δ, the BF-defect of f, vanishes.

Put under the hypothesis of Theorem I

$$A = \operatorname*{Min}_{\nu} x_\nu , \qquad B = \operatorname*{Max}_{\nu} y_\nu . \tag{42}$$

Then we have, considering the Taylor developments of $f(\xi)$ at A, a, b, B, and using the Theorem II

$$V_A \geqq V(a_\nu) \geqq V_a \geqq V_b \geqq V(b_\nu) \geqq V_B$$

and therefore

$$V_A - V_B \geqq V(a_\nu) - V(b_\nu) \geqq V_a - V_b . \tag{43}$$

From (43) it follows, subtracting $N(a, b)$ from the 2-nd and the 3-rd term of the inequality and using (5),

$$V(a_\nu) - V(b_\nu) - N(a, b) \geqq 2\Delta(a, b) , \tag{44}$$

and in the same way, subtracting $N(A, B)$ from the first and the 2-nd term of the inequality (43),

$$2\Delta(A, B) \geqq V(a_\nu) - V(b_\nu) - N(A, B) . \tag{45}$$

27. We have now the following

Theorem VI. *Assume that under the hypothesis of Theorem II all* x_ν *and* y_ν *are contained in an interval* $(c, C\rangle$, *open from the left and closed from the right, for which the BF-defect vanishes,* $\Delta(c, C) = 0$.

Then we have

$$N(A, B) \geqq V(a_\nu) - V(b_\nu) \geqq N(a, b) , \tag{46}$$

and if

$$N(A, a) = N(b, B) = 0 \tag{47}$$

holds, even

$$V(a_\nu) - V(b_\nu) = N(a, b) . \tag{48}$$

28. Indeed, from our assumption about $\Delta(c, C)$ follows that $\Delta(A, B) = \Delta(a, b) = 0$ and therefore (46) is an immediate consequence of (44) and (45).

Further, (48) follows from (46) if $N(A, B) = N(a, b)$ and this is certainly true under the hypothesis (47).

In particular, (46) holds if $f(\xi)$ has only real roots.

29. Consider for instance the case where

$$x_1 = a-n+1\,, \quad x_2 = a-n+2\,, \quad \ldots, \quad x_{n-1} = a-1\,, \quad x_n = a\,;$$
$$y_1 = b\,, \quad y_2 = b+1\,, \quad \ldots, \quad y_{n-1} = b+n-2\,, \quad y_n = b+n-1\,.$$

Then the developments (7) and (9) are the corresponding Gregory – Newton developments with the step $h = 1$

$$f(\xi) = \sum_{\nu=0}^{n} \Delta^{\nu} f(a-n+1) \binom{\xi-a+n-1}{\nu}, \qquad f(\xi) = \sum_{\nu=0}^{n} \Delta^{\nu} f(b) \binom{\xi-b}{\nu},$$

and therefore

$$(n-\nu)!\, a_{\nu} = \Delta^{n-\nu} f(a-n+1)\,, \qquad (n-\nu)!\, b_{\nu} = \Delta^{n-\nu} f(b)\,.$$

Since $A = a-n+1$, $B = b+n-1$ we have

$$V(\Delta^{\nu} f(a-n+1)) - V(\Delta^{\nu} f(b)) = N(a\,, b)\,,$$

if $f(\xi)$ has only real roots, if $b-a$ is greater than $n-1$ and if there are no roots of $f(\xi)$ in the intervals $(a-n+1\,, a\rangle$, $(b\,, b+n-1\rangle$.

Generalisation of Obreschkoff's Theorems

30. For a fixed real polynomial $f(\xi)$ of exact degree n, a real a and a non-negative m_a, we denote by $M(a\,, m_a)$ the number of roots w of f satisfying the condition:

$$(49) \qquad |\arg\,(w-a)| < \frac{\pi}{m_a}$$

where $w = a$ is *not* counted even if a is a root, and by $\bar{M}(b\,, m_b)$ the number of roots w, satisfying the condition:

$$(50) \qquad |\arg\,(b-w)| \leq \frac{\pi}{m_b}\,,$$

where $b = w$ is counted with its multiplicity, if b is a root.

Further, the number of roots w of $f(\xi)$ satisfying *both conditions* (49) and (50) will be denoted by $N(a\,, b\,, m_a\,, m_b)$.

If m_a in (49) or m_b in (50) becomes $= 0$, the corresponding conditions become meaningless and have to be disregarded in the definition of the above symbols.

31. Using these notations and those of sections 2 and 6, Obreschkoff's Theorems in question are given by:

A. *We have for real* $a < b$

(51) $$N(a\,,b\,,n - V_a\,,V_b) \leqq V_a - V_b \qquad (a < b)\,;$$

B. *We have for any real* $a\,,b$

(52) $$M(a\,,n - V_a) \leqq V_a\,, \qquad \bar{M}(b\,,V_b) \leqq n - V_b\,;$$

C. *If for a real* a *we have*

(53) $$p = M(a\,,n+2-p)\,, \qquad n-p = \bar{M}(a\,,p+2)\,,$$

then it follows $V_a = p$ [2]).

Obreschkoff's relation (51) generalizes the Budan—Fourier inequality (4) very considerably; while $N(a\,,b)$ in (4) is the number of real roots in the interval $(a\,,b\rangle$, the left side expression in (51) is equal to $N(a\,,b)$ plus the number of the non-real roots *inside* of the quadrangle of the Fig. 1, symmetric to the real axis.

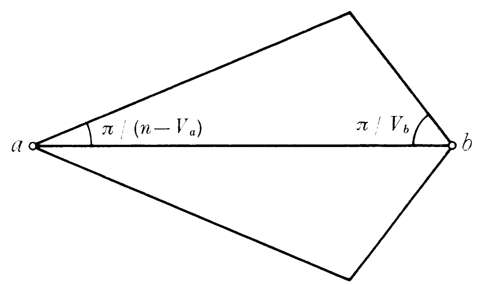

Fig. 1

32. From the Theorems **A**, **B**, **C** we will derive, applying the Fundamental Lemma:

Theorem VII. *We have in the notations of the sections 2, 6 and 30:*

(54) $$N\big(a\,,b\,,n - V(b_\nu)\,,V(a_\nu)\big) \leqq V(a_\nu) - V(b_\nu) \qquad (a < b)\,,$$

and

(55) $$M(a\,,n) \leqq V(a_\nu)\,, \qquad \bar{M}(A\,,n) \leqq n - V(a_\nu)\,;$$

[2]) Cf. the exposition in Obreschkoff's book quoted in sec. 25, which contains also a complete bibliography. As a matter of fact, in Obreschkoff's formulation l. c. p. 83, b, if it is a root, is still not counted in the left side expression in (51). However, the inequality including b is obtained immediately from that given by Obreschkoff by a continuity argument.

further, if we have

(56) $$M(a\,,2+n-p) = p\,, \qquad \bar{M}(A\,,2+p) = n-p\,,$$

and no roots of f *satisfy the condition* $|\arg(A-w)| = \pi\,/\,(2+p)$ *or lie in* A, *then* $a_0\,a_n \neq 0$, $V(a_\nu) = p$, *and for any* $a_\varkappa = 0$ *we have* $a_{\varkappa-1}\,a_{\varkappa+1} < 0$.

33. *Proof.* Defining a and b by (10) we have (51). The right hand expression of (51) is, in virtue of (43), majorized by $V(a_\nu) - V(b_\nu)$. On the other hand, by the Theorem II, $V_b \leqq V(a_\nu)$, $n - V_a \leqq n - V(b_\nu)$, so that the quadrangle corresponding to the left side expression in (54) is contained in that corresponding to (51). This proves (54).

To derive (55) from (52), take in (52) $b = A$ and observe that we have

$$n - V_a \leqq n\,, \qquad V_A \leqq n$$

and that, by the Theorem II,

$$V_a \leqq V(a_\nu) \leqq V_A\,.$$

34. To prove the last assertion of the Theorem VII observe that under the hypothesis (56) all roots of f lie inside of the two angular domains of Fig. 2, symmetric to the real axis.

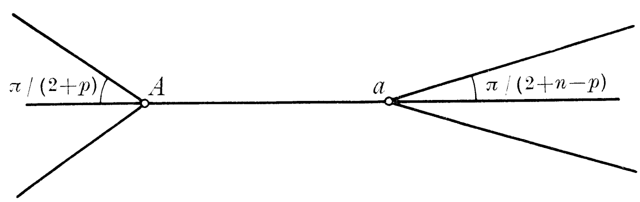

Fig. 2

But then we have from (56)

$$\bar{M}(a\,,2+p) = n-p\,,$$

and **C**, applied to a, gives $V_a = p$.

On the other hand it follows from (56)

$$N(A\,,2+n-p) = p\,,$$

and **C**, applied to A, gives $V_A = p$. Since, by the Theorem II, $V(a_\nu)$ is contained between V_a and V_A, we see that $V(a_\nu) = p$.

35. Introduce now the polynomial

$$\bar{f}(\xi) \equiv (-1)^n f(-\xi)\,,$$

for which we have obviously from (56):

$$\bar{M}(-a\,,\,2+n-p) \;=\; p\,, \qquad M(-A\,,\,2+p) \;=\; n-p$$

so that, for $\bar{f}$, we have to replace p by $n-p$.

On the other hand, replacing ξ by $-\xi$, in (7), and multiplying on both sides by $(-1)^n$ we obtain:

$$\bar{f}(\xi) \;=\; \sum_{\nu=0}^{n} (-1)^\nu a_\nu \bar{P}_{n-\nu}(\xi)\,, \qquad \bar{P}_\varkappa(\xi) \;=\; (-1)^\varkappa P_\varkappa(-\xi) \;=\; \prod_{\nu=1}^{\varkappa} (\xi+x_\nu)\,,$$

so that here the x_ν are to be replaced by the $-x_\nu$ and the a_ν by $(-1)^\nu a_\nu$. Applying then to $\bar{f}(\xi)$ the part of the Theorem VII which is already proved, we obtain $V((-1)^\nu a_\nu) \;=\; n-p$, and therefore

$$V(a_\nu) + V((-1)^\nu a_\nu) \;=\; n\,. \tag{57}$$

The last assertion of the Theorem VII will therefore follow immediately from the corollary of sec. 40 from Gauss' Lemma which we now proceed to formulate and to prove.

Symmetric Variation Numbers

36. For a sequence

$$a_0\,,\; a_1\,,\; \ldots\,,\; a_n \tag{58}$$

we put

$$W(a_\nu) \;=\; V(a_\nu) + V((-1)^\nu a_\nu) \tag{59}$$

and call $W(a_\nu)$ the *Symmetric Variation Number* of the sequence a_ν. Here, the exponent of $(-1)^\nu$ at a_ν is the order number of a_ν in the sequence, diminished by 1.

If none of the a_ν vanishes, $V((-1)^\nu a_\nu)$ is the number of permanences in the sequence a_ν, so that then $W(a_\nu) = n$.

If we replace then some of the a_ν by zeros the variation numbers cannot increase, so that we have:

$$W(a_\nu) \leqq n\,, \qquad W(a_\nu) \;=\; n \quad (a_\nu \neq 0)\,. \tag{60}$$

Further, if an $a_k \neq 0$ we have obviously

$$W(a_\nu) \;=\; W(a_0\,,\ldots,a_k) + W(a_k\,,\ldots,a_n) \qquad (a_k \neq 0)\,. \tag{61}$$

37. If we have in the sequence of the a_ν

$$a_k \neq 0\,, \quad a_{k+1} = \ldots = a_{k+p} = 0\,, \quad a_{k+p+1} \neq 0\,. \tag{62}$$

we have here a *gap* of the *length* p, *bridged over* by a variation if $a_k a_{k+p+1} < 0$, or by a *permanence* if $a_k a_{k+p+1} > 0$; this gap is called *odd* or *even* according its length p is odd or even.

Then the gist of an argument due to Gauss and published in 1828 (C. F. Gauss: Beweis eines algebraischen Lehrsatzes. - J. Reine Angew. Math. 3, 1828, pp. 1—4, reprinted in Carl Friedrich Gauss Werke III, Königliche Gesellschaft der Wissenschaften zu Göttingen, 1876, pp. 67—70) can be formulated as the

Gauss' Lemma. *Assume that we have in (58)* $a_0 a_n \neq 0$ *and denote by* α *the total number of odd gaps bridged over by variations, by* β *the total number of odd gaps bridged over by permanences and by* λ *the sum of lengths of all gaps in (58). Then we have*

$$n - W(a_\nu) = \lambda - \alpha + \beta . \tag{63}$$

Proof of Gauss' Lemma

38. If the gap (62) is an even one we leave out all vanishing elements $a_{k+1}, \ldots, a_{k+p}$ and reduce thereby n by p. On the other hand, as the parity of the indices is not changed, the value of $W(a_\nu)$ is not changed either, while λ is diminished by p, so that on both sides of (63) p is substracted. We can therefore assume, proving (63), that there are *no even gaps*.

If the gap (62) is odd and its length p is > 1, then we leave from (58) $p-1$ first elements of this gap, $a_{k+1}, \ldots, a_{k+p-1}$. Again, n and λ are both diminished by $p-1$, while, as $p-1$ is even, $W(a_\nu)$ is not changed. As we can proceed in this way with all odd gaps, we can assume now that there exist only odd gaps and that each of them has the length 1. But then we have obviously $\lambda = \alpha + \beta$ so that the relation to prove becomes

$$n - W(a_\nu) = 2\beta \tag{64}$$

39. If we have now $a_k = 0$, consider the decomposition

$$W(a_\nu) = W(a_0, \ldots, a_{k-1}) + W(a_{k-1}, a_k, a_{k+1}) + W(a_{k+1}, \ldots, a_n) . \tag{65}$$

If this gap is bridged over by a variation, so that $a_{k-1} a_{k+1} < 0$ we have obviously

$$W(a_{k-1}, a_k, a_{k+1}) = 2 ,$$

and this remains true, if we replace a_k by any number $\neq 0$, say by 1. We see that all odd gaps bridged over by a variation can be filled up without changing $W(a_\nu)$. We can therefore assume, in proving (64) that all gaps $a_k = 0$ which remain are bridged over by a permanence, $a_{k-1}\, a_{k+1} > 0$. But then, if we replace a_k by 1 we have

$$W(a_{k-1}, 0, a_{k+1}) = 0, \qquad W(a_{k-1}, 1, a_{k+1}) = 2$$

so that $W(a_\nu)$ is increased by 2, while β is diminished by 1 and both sides of (64) are diminished by 2. In this way we can reduce the proof of (64) to the case that there are no gaps at all and then (64) follows immediately from (60). Gauss' Lemma is proved.

40. **Corollary.** *Necessary and sufficient in order that we have* $W(a_\nu) = n$ *is that* $a_0\, a_n \neq 0$ *and that the only gaps in (58) are gaps of the length* 1 *bridged over by a variation.*

Proof. If $a_0 = 0$ or $a_n = 0$, n could be replaced in (60) by $n-1$, so that obviously $a_0\, a_n \neq 0$ is indeed necessary. But then it follows from (63) that we must have

$$\lambda - \varkappa + \beta = 0$$

while the total length λ of all gaps is certainly $\geqq \varkappa + \beta$. Therefore, we must have in any case $\beta = 0$ and $\lambda = \varkappa$, and this is already also sufficient. But $\lambda = \varkappa$ signifies that the only gaps which exist are those of the length 1 counted in $\varkappa$, which is the assertion of our corollary.

41. In Gauss' paper referred to in sec. 37 the expression $\lambda - \varkappa + \beta$ in (63) was derived as a lower limit for the number of the non-real roots of the equation $a_0 x^n + a_1 x^{n-1} + \ldots + a_n = 0$. Correspondingly a similar result holds for the equation $f(\xi) = 0$, if $f(\xi)$ is given by (7):

Theorem VIII. *If* $f(\xi)$ *is given by (7) with* $a_0\, a_n \neq 0$ *and according to (10) and (42)*

$$a = \operatorname*{Max}_\nu x_\nu, \qquad A = \operatorname*{Min}_\nu x_\nu$$

then the number of non-real roots of $f(\xi)$ *plus the number of real roots from the interval* $\langle A, a\rangle$ *is not less than* $\lambda - \varkappa + \beta$, *where* λ *is the total length of all gaps in (58),* $\varkappa$ *the number of odd gaps bridged over by a variation and* β *the number of odd gaps bridged over by a permanence.*

42. *Proof.* Indeed, in virtue of (36), the number of real roots from the interval (a, ∞) is $\leqq V(a_\nu)$.

Replacing ξ in (7) by $-\xi$ and proceeding as in the section 35 we obtain for the number of real roots of $f(\xi)$ in the interval $(-\infty, A)$ the upper limit $V((-1)^\nu a_\nu)$. Therefore a lower limit for the number of all other roots is

$$n - V(a_\nu) - V((-1)^\nu a_\nu) = n - W(a_\nu),$$

and this is by Gauss' Lemma equal to $\lambda - \alpha + \beta$.

University of Basel
Switzerland

Printed August 1964