Mathematics. — On the fundamental theorem of algebra. (Second communication)¹). By J. G. VAN DER CORPUT.

(Communicated at the meeting of September 21, 1946.)

§ 2. Proof of the transitivity.

Definition. A polynomial is called simple if it is relatively prime to its derivative.

Lemma 1. Every divisor D(X) of a simple polynomial F(X) is simple, for a common divisor of D(X) and $\frac{dD}{dX}$ would occur both in F(X) = D(X) U(X) and in $\frac{dF}{dX} = U \frac{dD}{dX} + D \frac{dU}{dX}$, which is impossible.

Lemma 2. The product of simple relatively prime polynomials is simple.

It is sufficient to give a proof for the product of two polynomials. Be F(X) = U(X) V(X). Then $\frac{dF}{dX} = U \frac{dV}{dX} + V \frac{dU}{dX}$ is relatively prime both to U(X) and V(X), hence also to U(X) V(X).

Lemma 3. Two polynomials, the product of which is simple, are simple and relatively prime.

In fact, a common divisor of U(X) and $\frac{dU}{dX}$ (or V(x)) occurs both in F(X) = U(X) V(X) and in $\frac{dF}{dX} = U\frac{dV}{dX} + V\frac{dU}{dX}$, which is impossible, if U(X) V(X) is simple.

Lemma 4. Any polynomial F(X) may be written in the form $F(X) = F_1^{\alpha_1} \dots F_{\lambda}^{\alpha_{\lambda}},$

where the exponents are positive and $F_1(X), \ldots, F_{\lambda}(X)$ are simple, relatively prime polynomials.

Proof. The theorem is obvious if F(X) is simple. If not, put $F(X) = D(X) F^*(X)$, where D is the greatest common divisor of F(X) and $\frac{dF}{dX}$. In this case the degree of both D(X) and $F^*(X)$ is less than the degree of F(X). We may suppose, that for these polynomials the

¹⁾ Compare Proc. Kon. Ned. Akad. v. Wetensch., Amsterdam, 49, 722-732 (1946) and Indagationes Mathematicae 4 (1946).

proof is already given. Consequently we may write

where the exponents a_i are natural numbers and $F_1 \dots F_{\lambda}$ are simple polynomials. We may even assume that these polynomials are relatively prime. For, if for instance $F_1 = UV$ and $F_2 = UW$, then U, V and W, being divisors of simple polynomials, are simple and

$$F = U^{\alpha_1 + \alpha_2} V^{\alpha_1} W^{\alpha_2} F_3^{\alpha_3} \dots F_{\lambda}^{\alpha_{\lambda}}.$$

Continuing in this manner we finally find F(X) written in a form, similar to (6) with simple, relatively prime polynomials.

Lemma 5. If a polynomial F is written in the form (6), where F_1, \ldots, F_{λ} are simple and relatively prime and the exponents are positive, then $F_1(X) \ldots F_{\lambda}(X)$ is the characteristic divisor of F(X).

Proof. $\frac{dF}{dX}$ is the sum of λ terms; all but one of these terms are divisible by $F_1^{\alpha_1}$. The exceptional term possesses the form $\alpha_1 \frac{F}{F_1}$ and is divisible by $F_1^{\alpha_1-1}$ but has further no common factor with F_1 . The derivative $\frac{dF}{dX}$ is therefore divisible by $F_1^{\alpha_1-1}$ and has no further common factor with F_1 . A similar argument is valid for the polynomials F_2, \ldots, F_{λ} , so that $G = F_1^{\alpha_1-1} \ldots F_{\lambda}^{\alpha_{\lambda}-1}$ is the greatest common divisor of F and $\frac{dF}{dX}$. Hence it follows, that $F_1 \ldots F_{\lambda}$ is equal to the characteristic divisor.

Corollaries. The characteristic divisor $F^*(X)$ of a polynomial F(X) is simple, being the product of simple, relatively prime polynomials.

F(X) divides $(F^*(X))^{\mu}$, where μ denotes the degree of F(X), since each $a_r \leq \mu$.

If D(X) is a divisor of F(X), then $D^*(X)$ is a divisor of $F^*(X)$. In fact, we may write $D(X) = F_1^{\beta_1} \dots F_{\lambda}^{\beta_{\lambda}}$ where $0 \leq \beta_r \leq a_r$, so that the characteristic divisor $\prod_{\beta_{\varrho} > 0} F_{\varrho}$ of D(X) is a divisor of the characteristic divisor $\prod_{\beta_{\varrho} > 0} F_{\varrho}$ of D(X) is a divisor of the characteristic divisor $\prod_{\varrho=1}^{\lambda} F_{\varrho}$ of F(X).

Lemma 6. If U is the greatest common divisor of F and G, then U^* is the greatest common divisor of F^* and G^* .

Proof. U is a divisor of both F and G. By corollary of lemma 5 the polynomial U^* is a divisor of both F^* and G^* , hence also of their greatest common divisor.

The greatest common divisor D of F^* and G^* is a divisor of U. Hence the characteristic divisor D^* of D is a divisor of U^* . But D is a divisor of the simple polynomial $F^*(X)$, so that it is also simple, consequently identical with its own characteristic divisor. Hence D is a divisor of U^* . Consequently U^* and D are identical.

Lemma 7. If q denotes a positive element of Ω , then every interval can be divided in a finite number of subintervals, each with length $\leq q$. In fact, Ω being Archimedeanly ordered, a natural number ϱ exists, such that $\varrho > \frac{b-a}{q}$, where a and b denote the endpoints of the interval. The points $a + \frac{\sigma(b-a)}{q}$ where σ runs through $0, 1, \dots \varrho$, divide the given interval in subintervals, each with length $\leq q$.

Lemma 8. To any two relatively prime polynomials F(X) and G(X) corresponds a positive element $p = p(F, G) \leq e$ of Ω , satisfying the inequality

$$|F(X)|+|G(X)|>p$$

for all elements X of Ω .

Proof. We can restrict ourselves to an interval Γ outside which |F(X)| > e. Since F(X) and G(X) are relatively prime, the unit element of Ω can be written in the form

$$e = U(X) F(X) + V(X) G(X),$$

where U(X) and V(X) denote appropriate polynomials. The absolute value of U(X) and V(X) in Γ is less than $\frac{e}{p}$, where p is a suitable positive element of Ω . Then we have for all X of Ω in Γ

$$e < \frac{e}{p} \left(|F(X)| + |G(X)| \right),$$

which implies the required result.

Lemma 9. To any two relatively prime polynomials F(X) and G(X) corresponds a positive element q = q(F, G) of Ω , such that G(X) is definite in every interval with length $\cong q$, where F(X) changes sign, and F(X) is definite in every interval with length $\cong q$, where G(X) changes sign ("definite" means: always positive or always negative).

Proof. We may restrict ourselves to a bounded interval \triangle . Consider the elements $p(F^*, G)$ and $p(F, G^*)$ of lemma 8. Be p_1 the smaller of) these two elements. Ω contains a positive element $m \ge e$, such that each derivative of $F^*(X) = H(X)$, of F(X), of $G^*(X)$ and of G(X)has in \triangle an absolute value $\le m$. To prove that the element $q = \frac{p_1}{4m}$ possesses the required property we consider a subinterval \triangle_1 of \triangle with length $\le q$, where F(X) changes sign. For any two elements x and x+h of \triangle_1 we have $|h| \equiv q \equiv e$, hence, if μ denotes the degree of F,

$$|H(x+h)-H(x)| = |h| \left| \frac{H'(x)}{1!} + \ldots + \frac{h^{\mu-1}H^{(\mu)}(x)}{\mu!} \right|$$

$$\equiv |h| m \left(\frac{1}{1!} + \frac{1}{2!} + \ldots + \frac{1}{\mu!} \right) < 2|h| m \equiv 2q m = \frac{p_1}{2} \right\}.$$
(7)

Similarly

$$|G(x+h)-G(x)| < \frac{p_1}{2}$$
. (8)

Since F(X) changes sign in \triangle_1 , this interval contains two elements u and v with $F^*(v) \leq 0 \leq F^*(u)$.

Hence from (7) we infer

$$0 \leq F^{*}(u) \leq F^{*}(u) - F^{*}(v) < \frac{p_{1}}{2}.$$

From the preceeding lemma it follows then $|G(u)| > \frac{p_1}{2}$. From (8) we find for each element x of Δ_1

$$G(x) > G(u) - \frac{p_1}{2} > 0$$
, if $G(u) > \frac{p_1}{2}$,

and

$$G(x) < G(u) + \frac{p_1}{2} < 0$$
, if $G(u) < -\frac{p_1}{2}$.

Hence G(X) is definite in \triangle_1 .

The second part of the theorem follows by exchanging F and G.

Definition. Two polynomials F(X) and G(X) are said to be equivalent in an interval Φ , if two polynomials P and Q, both definite in Φ , exist, such that $F^*P = G^*Q$.

Lemma 10. If a polynomial F(X) changes sign in an interval Φ , then every multiple G(X) of F(X) changes also sign in Φ . If further is given, that G(X) changes sign at most once in Φ , then F(X) and G(X) are equivalent in Φ (and therefore both change sign exactly once in Φ).

Proof. G^* is a multiple of F^* . Be $G^*(X) = F^*(X) U(X)$. Since $F^*(X)$ changes sign in Φ , by lemma 9 and 3 we can divide Φ in subintervals, each with length $\cong q$, where q denotes the element $q(F^*, U)$ of lemma 9. In at least one of these subintervals, say Φ_1 , the polynomial $F^*(X)$ changes sign. In virtue of the choice of q the polynomial U(X) is definite in Φ_1 , hence $G^*(X) = F^*(X) U(X)$ changes sign in Φ_1 .

Suppose that G(X) changes sign only once in Φ . If U(X) were not definite in Φ , in one of the constructed subintervals, say Φ_2 , the polynomial U(X) would change sign and there $F^*(X)$ would be definite. Hence $G^*(X) = F^*(X) U(X)$ would change sign in Φ_2 . Since U(X) is definite in Φ_1 , but not in Φ_2 , the intervals Φ_1 and Φ_2 would be different and $G^*(X)$ would change sign in Φ more than once. This not being the case, we find U(X) to be definite in Φ and therefore F(X) and G(X) to be equivalent in Φ .

Remark. From $(\Gamma, C) = (\Delta, D)$ it follows, that (C, D) changes sign at most once in (Γ, Δ) , for (C, D) changes sign there and has a multiple D(X), which changes sign at most once in (Γ, Δ) .

Lemma 11. If F(X) and G(X) are equivalent in Φ , they are also equivalent to their greatest common divisor in Φ .

Proof. Put $F^* = U^*P$ and $G^* = U^*Q$, where U^* is the greatest common divisor of F^* and G^* , so that P and Q are relatively prime. By lemma 6 the polynomial U^* is the characteristic divisor of the greatest common divisor U of F and G. If P(X) is not definite in Φ , then a subinterval of Φ exists, in which P(X) changes sign and which has a length $< q(U^*, P)$ and < q(P, Q), where q denotes the element of Ω , introduced in lemma 9. By this lemma U^* and Q are definite in this subinterval; hence also G^* , while F^* changes sign in that subinterval. This is impossible, since F(X) and G(X) are equivalent in Φ .

Proof of the transitivity. From $(\Gamma, C) = (\Delta, D)$ and $(\Delta, D) = (\Lambda, L)$ we must deduce $(\Gamma, C) = (\Lambda, L)$. The greatest common divisor (C, D) of C and D changes sign in (Γ, Δ) and possesses a multiple D(X), which changes sign at most once in that interval. From lemma 10 these polynomials are equivalent in (Γ, Δ) and both change sign there. Similarly (D, L) and D are equivalent in (Δ, Λ) and both change sign there. By lemma 11 the polynomial D(X) is in $(\Gamma, \Delta, \Lambda)$ equivalent to the greatest common divisor (C, D, L) of (C, D) and (D, L). In $(\Gamma, \Delta, \Lambda)$ the polynomial D(X) changes sign, for otherwise D(X) would change sign both in (Γ, Δ) and (Δ, Λ) outside $(\Gamma, \Delta, \Lambda)$, consequently more than once in Δ , which is impossible. Hence the polynomial (C, D, L) which is equivalent with D(X), changes sign in $(\Gamma, \Delta, \Lambda)$, and therefore also its multiple (C, L) by lemma 10. Then (C, L) changes sign also in (Γ, Λ) .

§ 3. Definition of sum and product.

Consider two polynomials

$$F(X) = f_0 + f_1 X + \ldots + f_{\mu} X^{\mu}$$
 and $G(X) = g_0 + g_1 X + \ldots + g_{\nu} X^{\nu}$,

where $f_{\mu} = g_r = e$. The products $\Pi(X - Y_e - Z_r)$ and $\Pi(X - Y_e Z_r)$, where ϱ runs through $1, \ldots, \mu$ and where σ runs through $1, \ldots, \nu$, and where X, Y_{ϱ} and Z_{τ} denote indeterminates, may be written as integral rational functions of X, of the elementary symmetric functions of Y_1, \ldots, Y_{μ} and of the elementary symmetric functions of Z_1, \ldots, Z_{ν} . If we replace the elementary symmetric functions ΣY_1 , $\Sigma Y_1 Y_2, \ldots, Y_1 Y_2, \ldots, Y_1 Y_2, \ldots, \Sigma Z_1$, $\Sigma Z_1 Z_2, \ldots, Z_1 Z_2 \ldots Z_r$ respectively by $-f_{\mu-1}, f_{\mu-2}, \ldots, (-1)^{\mu} f_0, -g_{\nu-1}, g_{\nu-2}, \ldots, (-1)^{\nu} g_0$, these products become polynomials in X, which I denote by F(X) + G(X) and $F(X) \times G(X)$.

If we put $G(X) = G_1(X) G_2(X)$, the left side of the identity

$$\prod_{\varrho=1}^{\mu} G(X - Y_{\varrho}) = \prod_{\varrho=1}^{\mu} G_{1}(X - Y_{\varrho}) \cdot \prod_{\varrho=1}^{\mu} G_{2}(X - Y_{\varrho})$$

becomes F(X) + G(X), hence

$$F(X) + G(X) = (F(X) + G_1(X)) (F(X) + G_2(X)).$$

Herefrom it follows: If $G_1(X)$ is a divisor of G(X), then $F(X) + G_1(X)$ is a divisor of F(X) + G(X). In the same way we get: If also $F_1(X)$ is a divisor of F(X), then $F_1(X) + G_1(X)$ is a divisor of $F(X) + G_1(X)$, hence of F(X) + G(X). In particular:

Lemma 12. If u is a root of F(X) and v is a root of G(X), then (X-u) + (X-v) = X-u-v is a divisor of F(X) + G(X), hence u+v is a root of F(X) + G(X).

Lemma 13. A polynomial F(X), the derivative of which is always ≥ 0 in an interval Φ , satisfies the inequality $F(u) \leq F(v)$ for all elements u and v of Φ with $u \leq v$.

Proof. The second and higher derivatives of F(X) in Φ are absolutely less than a suitably chosen element m of Ω . Divide the interval with endpoints u and v into σ equal parts, each of length $l = \frac{v-u}{\sigma} \leq e$. For the endpoints a and b (a < b) of such a part we have

$$F(b) - F(a) = \frac{b-a}{1!} F'(a) + \frac{(b-a)^2}{2!} F''(a) + \ldots + \frac{(b-a)^{\mu} F^{(\mu)}(a)}{\mu!}$$

where μ denotes the degree of F. From $F'(a) \ge 0$ it follows

$$F(b)-F(a) \cong -m l^2\left(\frac{1}{2!}+\ldots+\frac{1}{\mu!}\right) \cong -m l^2,$$

and adding we obtain

$$F(v)-F(u) \cong -m l^2 \sigma = -\frac{m (v-u)^2}{\sigma}$$

Since the number σ may be taken arbitrary large, we find $F(v) - F(u) \ge 0$, for otherwise the number σ could be taken as large as to contradict the inequality.

Remark. From this lemma it follows immediately: If a polynomial has a definite derivative in an interval, the polynomial changes sign there at most once.

Lemma 14. If the polynomial C(X) changes sign in Γ and the polynomial D(X) in Δ , then the polynomial C(X) + D(X) changes sign in $\Gamma + \Delta$.

Proof. Suppose first that C and D are simple. Put C(X) + D(X) = F(X). If a and b denote the endpoints of $\Gamma + \Delta$, we may assume without loss of generality $F^*(a) \neq 0$ and $F^*(b) \neq 0$, for otherwise the lemma is evident. Ω contains a positive element m, such that the second and higher derivatives of $F^*(X)$ are all absolutely $\cong m$ in $\Gamma + \Delta$. Choose in Ω a positive element l, satisfying the inequalities

$$l \leq e; \ l \leq \frac{p}{4m}; \ l \leq \frac{2}{p} F^*(a); \ l \leq \frac{2}{p} F^*(b), \ . \ . \ . \ (9)$$

(where p denotes the element $p\left(F^*, \frac{dF^*}{dX}\right)$, introduced in lemma 8),

such that in the interval with endpoints a and a+l the inequality $|F^*(x)| > \frac{1}{2} |F^*(a)|$ is valid and in the interval with endpoints b-l and b similarly $|F^*(x)| > \frac{1}{2} |F^*(b)|$. From lemma 5, corollary we know $(F^*)^{\mu} = FG$, where μ denotes the degree of F and G is a suitable polynomial. Ω contains a positive element g, such that in $\Gamma + \Delta$ the polynomial G possesses an absolute value $\equiv g$.

If s and t are arbitrary elements of Ω , then

$$H(X, s, t) = \{(C(X) - s) + (D(X) - t)\} - \{C(X) + D(X)\}$$

is a polynomial in X, s and t. In each term of H either a factor s or a factor t occurs, for H(X, 0, 0) is identically equal to 0. Hence a positive element k of Ω exists, such that from |s| < k and |t| < k it follows

$$|H(w, s, t)| < \frac{1}{g} \left(\frac{p l}{4}\right)^{\mu}, \ldots \ldots \ldots \ldots \ldots$$
(10)

for all elements w of $\Gamma + \Delta$. Finally we choose the positive element h of Ω such that in every subinterval of Γ with length $\cong h$, the oscillation of C(X) is less than k, and that also in each subinterval of Δ with length $\cong h$, the oscillation of D(X) is less than k. Divide Γ and Δ into subintervals, each of length $\cong h$. In at least one of these subintervals of Γ , say Γ_1 , the polynomial C(X) changes sign and in at least one of the subintervals of Δ , say Δ_1 , the polynomial D(X) changes sign.

Then the interval Γ_1 contains two elements u and u_1 with $C(u_1) \leq 0 \leq C(u)$. Since the oscillation of C(X) in Γ_1 is less than k, it follows

$$0 \leq C(u) \leq C(u) - C(u_1) < k.$$

Similarly \triangle_1 contains a point v with $0 \leq D(v) < k$. Hence inequality (10) is valid for w = u + v, s = C(u) and t = D(v). Then u is a root of C(X) - s and v is a root of D(X) - t, hence w = u + v is a root of (C(X) - s) + (D(X) - t) by lemma 12. Therefore

$$|F(w)| = |C(w) + D(w)| = |H(w, s, t)| < \frac{1}{g} \left(\frac{p l}{4}\right)^{\mu}.$$

Hence

$$|F^*(w)|^{\mu} = |F(w) G(w)| < \left(\frac{pl}{4}\right)^{\mu}.$$

Consequently

$$|F^*(w)| < \frac{pl}{4} < \frac{1}{2}p.$$
 (11)

From the definition of $p = p\left(F^*, \frac{dF^*}{dX}\right)$ it follows

From (11) and (9) we infer

$$|F^*(w)| < \frac{1}{2} |F^*(a)|$$
 and $|F^*(w)| < \frac{1}{2} |F^*(b)|$.

Therefore it is impossible that w lies either in the interval with endpoints a and a+l or in the interval with endpoints b-l and b. As w lies in the interval with endpoints a and b, it lies in the interval with endpoints a+l and b-l. The interval $\Gamma + \Delta$ contains consequently the elements w-l and w+l.

Since the second and higher derivatives of F^* are absolutely $\leq m$ in $\Gamma + \triangle$ and since $l \leq e$, the Taylor development gives

$$\left|F^{*}(w \mp l) - F^{*}(w) \pm l \frac{dF^{*}(w)}{dw}\right| \leq m l^{2} \left(\frac{1}{2l} + \frac{1}{3l} + \dots + \frac{1}{\mu l}\right) < m l^{2} \leq \frac{1}{4} p l_{0}$$

hence by (11) we get

$$\left|F^*(w\pm l)\mp l\frac{dF^*(w)}{dw}\right| < \frac{1}{2}pl.$$

From (12) we infer that $F^*(w-l)$ and $F^*(w+l)$ have different sign; consequently F(X) changes sign in $\Gamma + \Delta$.

Suppose now that C and D are not both simple. Then C^* and D^* , and therefore $C^* + D^*$ change sign resp. in Γ , Δ and $\Gamma + \Delta$; consequently the multiple C + D of $C^* + D^*$ changes sign also in $\Gamma + \Delta$.

Now we pass to the definition of the sum of $\gamma = (\Gamma, C)$ and $\delta = (\Delta, D)$. Put F = C + D. A subinterval Γ' of Γ and a subinterval Δ' of Δ can be found, both with length $\leq \frac{1}{2}q$, where q denotes the element $q\left(F^*, \frac{dF^*}{dX}\right)$, introduced in lemma 9, such that C changes sign in Γ' and D in Δ' . In the interval $\Gamma' + \Delta'$ with length $\leq q$ the polynomial F changes sign by lemma 14. By lemma 9 the derivative $\frac{dF^*}{dX}$ is definite throughout that interval; therefore F^* changes sign there at most once. So we have proved the existence of subintervals Γ' of Γ and Δ' of Δ ,

$$\gamma + \delta = (\Gamma' + \Delta', C + D),$$

if we show that the couple $(\Gamma' + \Delta', C + D)$ is uniquely determined. Suppose

$$(\Gamma, C) = (\Gamma_1, C_1)$$
 and $(\Delta, D) = (\Delta_1, D_1)$.

We have to prove

$$(\Gamma' + \Delta', C + D) = (\Gamma_1' + \Delta_1', C_1 + D_1),$$

where $\Gamma', \Delta', \Gamma_1'$ and Δ_1' are subintervals respectively of Γ, Δ, Γ_1 and Δ_1 , such that C changes sign only once in Γ', D in Δ', C_1 in Γ_1', D_1 in Δ_1' , F = C + D in $\Phi = \Gamma' + \Delta'$ and $F_1 = C_1 + D_1$ in $\Phi_1 = \Gamma_1' + \Delta_1'$. We must prove that the greatest common divisor L of F = C + D and $F_1 = C_1 + D_1$ changes sign in the common part Λ of Φ and Φ_1 . Since $S = (C, C_1) + (D, D_1)$ is a divisor both of F = C + D and $F_1 = C_1 + D_1$, the polynomial S is also a divisor of their greatest common divisor L. By assumption the couples $(\Gamma, C) = (\Gamma' C)$ and $(\Gamma_1, C_1) = (\Gamma_1', C_1)$ are equal, so that (C, C_1) changes sign in (Γ', Γ_1') and similarly (D, D_1) in (Δ', Δ_1') . By lemma 14 S changes sign in Σ , hence certainly in Λ , which contains Σ ; in fact each point w of Σ may be written in the form u + v, where u lies both in Γ' and $\Gamma_1' + \Delta_1'$, consequently also in their common part Λ . This establishes the proof.

In a similar way we define the product of two couples.