Extension of a Theorem of Mason

HAROLD N. SHAPIRO
Courant Institute

AND

GERSON H. SPARER
Pratt Institute

1. Introduction

Letting $N_0(f)$ denote the number of distinct roots of the polynomial $f(x)$, (over the complex numbers), a theorem of Mason (see [1], [2], and [5]) asserts

THEOREM 1.1. *Let $f(x), g(x), h(x)$ be polynomials with complex coefficients such that*

(1.1) f, g, h are not all constants, and

(1.2) f, g, h are relatively prime by pairs, and

(1.3) $f + g = h$.

Then

(1.4) $$\max(\deg f, \deg g, \deg h) \leqq N_0(fgh) - 1 .$$

In Section 2 this is extended in two ways. The condition (1.3) is replaced by

(1.5) $$f_1 + \ldots + f_r = 0 ,$$

for any fixed $r \geqq 3$, and the f_i are polynomials in several variables, i.e., $f_i = f_i(X) = f(x_1, \ldots, x_m)$, with complex coefficients. The definition of N_0 is generalized as follows:

DEFINITION 2.2. Given $f \in \mathbb{C}[x_1, \ldots, x_m], f$ has a factorization

(1.6) $$f(x) = \mathop{\pi}_{j=1}^{s} p_j(X)^{\alpha_j} ,$$

where the p_j are irreducible, distinct, (at most one of them is constant), and the $\alpha_j > 0$ are integers. Then define

(1.7) $$Q(X) = Q[f] = \mathop{\pi}_{j=1}^{s} p_j ,$$

Communications on Pure and Applied Mathematics, Vol. XLVII, 711–718 (1994)
 CCC 0010-3640/94/050711-8

(i.e., $Q[f]$ is the squarefree part of f), and

$$N_0(f) = \text{degree } Q \,. \tag{1.8}$$

In Section 2 we provide the following:

THEOREM 1.2. *For fixed integer* $r \geqq 3$, *let* $f_i \in \mathbb{C}[x_1, \dots, x_m], i = 1, \dots, r$, *such that* (1.5) *holds. Assume also that*

(1.8) the f_i are not all constant, and

(1.9) the f_i are relatively prime by pairs.

Then

$$\max_{i=1,\dots,r} (\deg f_i) \leqq (r-2)\left(N_0\left(\prod_{i=1}^{r} f_i\right) - 1\right). \tag{1.10}$$

(*Note that* (1.8) *implies that at least two of the* f_i *are non-constant.*)

In Section 3 various applications are made to diophantine equations in polynomial unknowns. One such result asserts that for the diophantine equation

$$g_1^{a_1} + g_2^{a_2} + \cdots + g_r^{a_r} = 0 \,, \tag{1.11}$$

$r \geqq 3, a_i$ positive integers, there exists a positive constant $c = c(r)$ such that if the g_i are relatively prime by pairs, and

$$\sigma = a_1^{-1} + \cdots + a_r^{-1} < c \,, \tag{1.12}$$

then all the $g_i, i = 1, \dots, r$, must be constant.

2. Proof of Theorem 1.2

For Δ a differential operator of the form

$$\Delta = (\mu_1 \mu_2 \dots \mu_m)^{-1} \frac{\partial^{\mu_1}}{\partial x_1^{\mu_1}} \cdots \frac{\partial^{\mu_m}}{\partial x_m^{\mu_m}} \,, \tag{2.1}$$

where the $\mu_i \geqq 0$ are integers, we denote the rank of Δ by

$$\rho(\Delta) = \sum_{i=1}^{m} \mu_i \,. \tag{2.2}$$

Given $\Delta_0, \dots, \Delta_s$ such that $\rho(\Delta_i) \leqq i, i = 0, \dots, s$, and polynomials $h_0, \dots, h_s$ in $\mathbb{C}[X]$, a generalized Wronskian has the form

$$W[h_0, \dots, h_s] = \det \left| \Delta_i h_j \right| \,. \tag{2.3}$$

A well-known result (see [3] and [4]) asserts that if the h_i are linearly independent over $\mathbb{C}$, then there exists a generalized Wronskian, of the form (2.3), which does not vanish.

We shall also require the following simple

LEMMA 2.1. *Let* Δ *be of the form* (2.1), *with* $\rho(\Delta) = t$. *Then for* $f(X)$ *not zero we have*

$$f^{-1}\Delta f = \frac{P(X)}{Q[f]^t}, \tag{2.4}$$

where $P = P(X) \in \mathbb{C}[X]$, *and*

$$\deg P \leqq t(\deg Q - 1). \tag{2.5}$$

Proof: If $t = 0, \Delta f = f, P = 1$, and the lemma is trivially true. Also, if $f(X)$ satisfies (1.6), then

$$f^{-1}\frac{\partial f}{\partial x_i} = \sum_{j=1}^{s} \alpha_j \frac{1}{p_j}\frac{\partial p_j}{\partial x_i} = \frac{P_1}{Q}, \tag{2.6}$$

where

$$\deg P_1 \leqq \deg Q - 1. \tag{2.7}$$

Proceeding by induction on $\rho(\Delta)$, so that assuming the lemma for Δ, it suffices to prove it for

$$\tilde{\Delta} = \frac{\partial}{\partial x_i}\Delta. \tag{2.8}$$

From (2.4) we obtain

$$f^{-1}\tilde{\Delta} f = (f^{-1}\Delta f)\,(f_{x_i}/f) + (P_{x_i}Q - tQ_{x_i}P)\,/Q[f]^{t+1}. \tag{2.9}$$

Then from (2.5) and (2.7) the numerator in (2.9) is of degree at most

$$\deg P + \deg Q - 1 \leqq (t+1)(\deg Q - 1),$$

which completes the induction.

The proof of Theorem 1.2 proceeds by induction on r. For $r = 2$ it is vacuously true, since if $f_1 = -f_2, f_1$ and f_2 relatively prime, then they are both constant. Assume the theorem for all cases $r', 2 \leqq r' < r$, and consider that of r polynomials. Note that if two of the f_i are constant, then we may eliminate them if their sum is zero or replace them by their sum when it is not zero. Then the inductive hypothesis could be applied to yield the desired result. Thus we may assume that at most one of the f_i is a constant.

Case I: Rewriting (1.5) as

$$-f_r = f_1 + \cdots + f_{r-1}\,, \tag{2.10}$$

assume that the $f_i, i = 1, \ldots, r-1$, are linearly dependent over $\mathbb{C}$. Note that from the above, at most one of the $f_i, i = 1, \ldots, r-1$, is constant. Let $f_{i_1}, \ldots, f_{i_q}, q < r-1$, be a maximal linearly independent subset of the $f_j, j = 1, \ldots, r-1$. Since $r - 1 \geqq 2$, and the f_j are relatively prime by pairs, it follows that $q \geqq 2$. Then each $f_j, 1 \leqq j \leqq r-1, j$ not one of the i_k, is a linear combination of the f_{i_k}, of the form

$$f_j = \lambda_1 f_{i_1} + \cdots + \lambda_q f_{i_q}\,, \tag{2.11}$$

where the $\lambda_k \in \mathbb{C}$, and at least two of these λ_k are not zero. Using our inductive hypothesis we apply the theorem to (2.11). This yields that if $\lambda_k \neq 0$ then

$$\deg\,(f_{i_k}) \leqq (q-1)\left(N_0\left(f_j \mathop{\pi}_{k=1}^{q} f_{i_k}\right) - 1\right)$$

so that

$$\deg\,(f_{i_k}) \leqq (r-2)\left(N_0\left(\mathop{\pi}_{j=1}^{r} f_j\right) - 1\right). \tag{2.12}$$

From (2.11) the same estimate as in (2.12) follows for $\deg f_j$. Thus the theorem is proved for such f_j and f_{i_k}. Inserting all the relations of the form (2.11) into the right side of (2.10) yields an equation of the form

$$f_r = \kappa_1 f_{i_1} + \cdots + \kappa_q f_{i_q}\,, \tag{2.13}$$

where the $k_j \in \mathbb{C}$. Moreover, if one of these $\kappa_\nu = 0$ then the corresponding f_{i_ν} must have appeared in one of the equations (2.11) with a non-zero λ_ν. Hence, (2.12) has been established for this f_{i_ν}. Finally, for those $\kappa_\nu \neq 0$, we treat (2.13) exactly as we did (2.11), (note that $q + 1 < r$), and obtain the estimate (2.12) for $\deg f_{i_\nu}$, and for $\deg f_r$. This completes the induction in this case.

Case II: The $f_1, \ldots, f_{r-1}$ are linearly independent over $\mathbb{C}$. We then have some generalized Wronskian $W[f_1, \ldots, f_{r-1}]$, of the form (2.3), which does not vanish. Applying the operators $\Delta_i, i = 0, \ldots, r-2$, to (2.10) yields

$$-f_r^{-1}\Delta_i f_r = \sum_{j=1}^{r-1} \left(f_j^{-1}\Delta_i f_j\right)\left(f_j/f_r\right)\,, \tag{2.14}$$

for $i = 0, \dots, r-2$. Viewing (2.14) as $r-1$ linear equations in the (f_j/f_r), we solve for these as:

$$\frac{f_j}{f_r} = \frac{\det \left| (\Delta_i f_k)/f_k \text{ for } k \neq j; (-\Delta_i f_r)/f_r \text{ in } j\text{-th col} \right|}{\det \left| (\Delta_i f_k)/f_k \right|}, \tag{2.15}$$

where $j = 1, \dots, r-1, i = 0, \dots, r-2$, and $k = 1, \dots, r-1$. Note that the denominator in (2.15) equals $W[f_1, \dots, f_{r-1}]$ divided by the product of the f_k, and hence is not zero. We let $N = N(X)$ and $D = D(X)$ denote the numerator and denominator, respectively, on the right of (2.15). Thus we have:

$$\frac{f_j}{f_r} = \frac{N}{D}. \tag{2.16}$$

The rational function $D \in \mathbb{C}(X)$ is a determinant which equals a sum of terms of the form

$$\frac{\pm (\Delta_{i_1} f_1)(\Delta_{i_2} f_2) \dots (\Delta_{i_{r-1}} f_{r-1})}{f_1 f_2 \dots f_{r-1}}, \tag{2.17}$$

where $i_1, \dots, i_{r-1}$ is a permutation of $0, \dots, r-2$. From (2.4) of Lemma 2.1, (2.17) is of the form

$$\pm \frac{P_{i_1}}{Q[f_1]^{\rho(\Delta_{i_1})}} \cdots \frac{P_{i_{r-1}}}{Q[f_{r-1}]^{\rho(\Delta_{i_{r-1}})}}, \tag{2.18}$$

where for $k = 1 \dots, r-1, P_{i_k} \in \mathbb{C}[X]$, and

$$\deg P_{i_k} \leqq \rho(\Delta_{i_k})(\deg Q[f_k] - 1). \tag{2.19}$$

Also, since the f_k are relatively prime by pairs, the denominator in (2.18) divides

$$Q[f_1 \dots f_{r-1}]^{r-2}.$$

Thus (2.18) has the form

$$\frac{P_{i_1, \dots, i_{r-1}}}{Q[f_1 \dots f_{r-1}]^{r-2}}, \tag{2.20}$$

where

$$\begin{aligned}
\deg P_{i_1, \dots, i_{r-1}} &\leqq \sum_{k=1}^{r-1} \deg P_{i_k} + \sum_{k=1}^{r-1} (r - 2 - \rho(\Delta_{i_k})) N_0(f_k) \\
&\leqq \sum_{k=1}^{r-1} \rho(\Delta_{i_k})(N_0(f_k) - 1) + (r-2) N_0(f_1 \dots f_{r-1}) \\
&\qquad - \sum_{k=1}^{r-1} \rho(\Delta_{i_k}) N_0(f_k) \\
&\leqq (r-2) N_0(f_1, \dots, f_{r-1}) - \sum_{k=1}^{r-1} \rho(\Delta_{i_k}).
\end{aligned}$$

Since $W[f_1, \dots, f_{r-1}] \neq 0, \rho\,(\Delta_{i_k}) \geqq 1$ for all but one of the k, and this implies

$$\deg P_{i_1,\dots,i_{r-1}} \leqq (r-2)\,(N_0(f_1,\dots,f_{r-1}) - 1)\ . \tag{2.21}$$

Thus the denominator D has the form

$$D = \frac{d(X)}{Q[f_1,\dots,f_r]^{r-2}}\ , \tag{2.22}$$

where $d(X) \in \mathbb{C}\,[X]$, and

$$\deg d(X) \leqq (r-2)\,(N_0(f_1 \dots f_r) - 1)\ . \tag{2.23}$$

By an entirely analogous argument the numerator N can be shown to have the form

$$N = \frac{n(X)}{Q[f_1,\dots,f_r]^{r-2}}\ , \tag{2.24}$$

where $n(X) \in \mathbb{C}\,[X]$, and

$$\deg n(X) \leqq (r-2)\,(N_0\,(f_1,\dots,f_r) - 1)\ . \tag{2.25}$$

From (2.16), (2.22), and (2.24) we have

$$\frac{f_j}{f_r} = \frac{n(X)}{d(X)}\ . \tag{2.26}$$

Finally, since f_j and f_r are relatively prime,

$$\deg f_j \leqq \deg n(X)\ , \qquad \deg f_r \leqq \deg d(X)\ ,$$

so that (2.23) and (2.25) imply (1.10).

3. Applications to Diophantine Equations

We next apply Theorem 1.2 to the diophantine equation (1.11), where the $a_i > 0$ are integers, and $r \geqq 3$.

THEOREM 3.1. *If the σ defined in* (1.12) *satisfies*

$$\sigma \leqq (r-2)^{-1}\ , \tag{3.1}$$

then the only solutions to (1.11) *for $g_i \in \mathbb{C}\,[X]$, which are relatively prime by pairs, are those for which the g_i are all constant.*

Proof: Assuming that the g_i are not all constant, Theorem 1.2 applies and (1.10) asserts that for $j = 1, \ldots, r$,

$$a_j \deg g_j \leqq (r-2)\left(N_0\left(g_1, \ldots, g_r\right) - 1\right) . \tag{3.2}$$

Dividing by a_j and summing over j yields

$$(1 - \sigma(r-2)) N_0\left(g_1, \ldots, g_r\right) \leqq -\sigma(r-2) < 0 ,$$

which contradicts (3.1).

COROLLARY . *In the case where all the* $a_i = n$, *i.e.,* (1.11) *becomes*

$$g_1^n + \ldots + g_r^n = 0 . \tag{3.3}$$

Then if $r \geqq 3$ *and* $n \geqq r(r-2)$, *the only solutions to* (3.3), *for which the* g_i *are relatively prime by pairs, are those for which the* g_i *are all constant.*

Remark. For $r = 3$, Theorem 3.1, yields that for $\sigma = a_1^{-1} + a_2^{-1} + a_3^{-1} \leqq 1$, the equation

$$g_1^{a_1} + g_2^{a_2} + g_3^{a_3} = 0 \tag{3.4}$$

has no non-constant solutions for which the g_i are relatively prime by pairs. The question arises as to whether or not if $\sigma > 1$, (3.4) has infinitely many "independent" non-constant solutions of this type. This can be shown to be so *except* for one isolated case:

$$g_1^2 + g_2^3 + g_3^5 = 0 . \tag{3.5}$$

In this case the question remains unresolved.

The diophantine equation (1.11) can be further generalized in that the powers which appear can be replaced by "almost powers."

DEFINITION 3.1. For $\eta > 0$ a polynomial $f \in \mathbb{C}[X]$ is called an η-almost-sth-power if

$$f = u^s h , \tag{3.6}$$

where

$$\deg h \leqq \eta \deg u . \tag{3.7}$$

(Note that a 0-almost-sth-power is an sth power.)

THEOREM 3.2. *If for $i = 1, \ldots, r, g_i \in \mathbb{C}[X]$ is an η-almost-a_ith-power, and*

$$\sigma = \sum_{i=1}^{r} a_i^{-1} \leqq (r-2)^{-1}(1+\eta)^{-1} , \tag{3.8}$$

then the only such sets of g_i which are relatively prime by pairs and satisfy (1.11) *are those for which the g_i are all constant.*

Proof: Since $g_i = u_i^{a_i} h_i$, where $\deg h_i \leqq \eta \deg u_i$, we have

$$\begin{aligned} a_j \deg u_j \leqq \deg g_j &\leqq (r-2)\left(N_0\left(\mathop{\pi}_{i=1}^{r} g_i\right) - 1\right) \\ &\leqq (r-2)\left(\sum_{i=1}^{r}(N_0(u_i) + N_0(h_i)) - 1\right) \\ &\leqq (r-2)\left(\sum_{i=1}^{r}(\deg u_i + \deg h_i) - 1\right) \\ &< (r-2)\left(\sum_{i=1}^{r}\deg u_i\right)(1+\eta) . \end{aligned}$$

Dividing by a_j, and summing over $j = 1, \ldots, r$, yields

$$\left(\sum_{i=1}^{r}\deg u_i\right)(1 - \sigma(r-2)(1+\eta)) < 0 ,$$

which contradicts (3.8).

Bibliography

[1] Lang, S., *Old and new conjectured diophantine equations,* Bull. Amer. Math. Soc. (n.s.), 23, 1990, pp. 37–75.

[2] Mason, R. C., *Equations over function fields,* pp. 149–157 in: *Number Theory, Noordwijerhout 1983,* Proceedings of the Journées Arithmétiques, H. Jager, ed., Lecture Notes in Mathematics No. 1068, Springer-Verlag, Berlin–New York, 1984.

[3] Roth, K. F., *Rational approximation to algebraic numbers,* Mathematika 2, 1955, pp. 1–20.

[4] Schneider, T., *Einführung in die Tranzsendenten Zahlen,* Springer-Verlag, Berlin, 1957, pp. 15–16.

[5] Vojta, P., *Diophantine Approximation and Value Distribution Theory,* Lecture Notes in Mathematics No. 1239, Springer-Verlag, Berlin–New York, 1987.

Received June 1992.