ACTA ARITHMETICA
136.4 (2009)

Generalized ABC theorems for non-Archimedean entire functions of several variables in arbitrary characteristic

by

William Cherry (Denton, TX) and
Cristina Toropu (Albuquerque, NM)

1. Introduction. The well-known ABC theorem for polynomials, first proved by Stothers [31] and often called "Mason's theorem" [22], says that if $a+b=c$, where a and b are relatively prime univariate polynomials with at least one of the derivatives a' or b' not identically zero, then

$$\max\{\deg a, \deg b, \deg c\} \leq \deg R(abc) - 1.$$

Here if f is a polynomial, we use $R(f)$ to denote $f/\gcd(f, f')$. In characteristic zero, $\deg R(f)$ is simply the degree of the square free part of f, also called the *degree* of the radical of f, and is the number of distinct zeros of f in an algebraically closed field containing the coefficients of f. We note that the above theorem immediately extends to polynomials of several variables by replacing the ordinary derivative with a partial derivative.

The existence of an appropriately analogous inequality for $c = a + b$ with a, b, and c relatively prime integers is the famous ABC conjecture of Masser and Oesterlé (see [25]), which states that for each $\varepsilon > 0$, there exists a constant $C(\varepsilon)$ such that

$$\max\{|a|, |b|, |c|\} \leq C(\varepsilon) S(abc)^{1+\varepsilon},$$

where we use $S(abc)$ to denote the square free part of abc. The ABC conjecture for integers has spectacular consequences in number theory—see e.g. [16] and [21]. To date, in the case of integers, the best proven upper bounds

2000 *Mathematics Subject Classification*: 11D04, 11D88, 32P05.
Key words and phrases: ABC theorems, non-Archimedean entire functions, several variables, Hasse derivatives, vanishing sums, Mason's theorem, generalized Wronskians.

Partial financial support for this research was provided by a Faculty Research Grant from the University of North Texas and by the United States National Security Agency under Grant Number H98230-07-1-0037. The United States Government is authorized to reproduce and distribute preprints notwithstanding any copyright notation herein.

DOI: 10.4064/aa136-4-4

on $\max\{|a|,|b|,|c|\}$ in terms of $S(abc)$ are super-polynomial in $S(abc)$—see e.g. [30].

The ABC theorem for polynomials has been generalized in a variety of directions, including: to sums in one-dimensional function fields by Mason [23], by Voloch [34] and by Brownawell and Masser [10], to sums of pairwise relatively prime polynomials of several variables by Shapiro and Sparer [29], to sums in higher-dimensional function fields by Hsia and Wang [18], and to quantum deformations of polynomials by Vaserstein [32]. Motivated by the analogy between Diophantine approximation and Nevanlinna theory [33], the ABC theorem has also been proven for complex entire functions by Van Frankenhuysen [13], [14] and for p-adic entire functions by Hu and Yang [19].

In a recent article, An and Manh [3] gave an ABC type theorem for p-adic entire functions of several variables $c = a + b$, but under some rather restrictive hypotheses, including the assumption that a, b, and c have no common zeros, which in several variables is a much stronger assumption than simply supposing that a, b, and c are pairwise relatively prime in the ring of entire functions. The purpose of this article is to prove general ABC theorems for sums

$$f_n = f_0 + \cdots + f_{n-1}$$

of non-Archimedean entire functions on affine m-space $\mathbb{A}^m$ in arbitrary characteristic analogous to the existing theorems for several-variable polynomials. We do this for two reasons. First, we illustrate that if the existing polynomial proofs for ABC theorems are correctly interpreted, then they provide immediate proofs for the analogous statements for non-Archimedean entire functions without the need for any fundamentally new ideas or for additional technical assumptions as in [3], with the exception of one slight subtlety discussed just before Corollary 6.5 below. Second, Cherry and Ye [12] developed a several-variable non-Archimedean Nevanlinna theory in a form that was intended to be easy to use. However, in applications, it is often convenient to work with, or at least to think in terms of, truncated counting functions, which were not discussed in [12]. Thus, our second purpose is to illustrate how to define and work with truncated counting functions in several variables and in positive characteristic; this is well-known to the experts, but we thought it helpful to illustrate here for the novice's benefit.

The plan of this paper is as follows. In Section 2 we set up some notation and recall some basic facts we will need. We also prove the most basic form of the ABC theorem in Section 2, so readers only interested in the basic idea do not need to read past Section 2. In Section 3, we recall the notion of Hasse derivative and generalized Wronskians necessary to work in positive characteristic and in several variables. In Section 4, we discuss how to define truncated counting functions in positive characteristic. We recall some linear algebra from Brownawell and Masser [10] in Section 5. Finally, in Section 6, we derive our general ABC theorems and indicate how various ABC theorems in the literature can be derived as corollaries. Our method of proof is essentially that of Hu and Yang [20], who adapted the argument of Brownawell and Masser [10] to the context of non-Archimedean entire functions of one variable. Our presentation in Section 6 was also influenced by the recent work of De Bondt [5], who formulated generalized ABC theorems for complex polynomials in several variables from which the other various versions in the literature can be derived.

2. Preliminaries, notation and warm-up. We will find it convenient to use some Nevanlinna notation. We will use [12] as our basic reference and mostly follow the notation there, although here we will not assume characteristic zero, as was done in [12].

Throughout, $\mathbb{F}$ will denote an algebraically closed field complete with respect to a non-Archimedean absolute value $|\ |$. We make no assumption about the characteristic of $\mathbb{F}$. Let $\mathbb{F}^\times$ denote $\mathbb{F}\setminus\{0\}$, and let $|\mathbb{F}^\times|$ be the subset of the positive real numbers defined by

$$|\mathbb{F}^\times| = \{|a| : a \in \mathbb{F}^\times\}.$$

Let $\mathbb{F}^m$ denote the mth Cartesian product of $\mathbb{F}$, which is the set of $\mathbb{F}$-points of affine m-space $\mathbb{A}^m$. By an *entire function* on $\mathbb{A}^m$ or $\mathbb{F}^m$, we mean a formal power series in m variables with coefficients in $\mathbb{F}$ and with infinite radius of convergence. We will use $\mathcal{E}_m$ to denote the ring of entire functions on $\mathbb{A}^m$.

If $z_1, \ldots, z_m$ are $\mathbb{F}$-valued variables, we use z to refer collectively to the m-tuple $(z_1, \ldots, z_m)$. When convenient, we will use multi-index notation. If $\gamma = (\gamma_1, \ldots, \gamma_m)$ is a multi-index, i.e., an m-tuple of non-negative integers, then by definition

$$z^\gamma = z_1^{\gamma_1} \cdots z_m^{\gamma_m}, \qquad |\gamma| = \gamma_1 + \cdots + \gamma_m, \qquad \partial^\gamma f = \frac{\partial^{|\gamma|} f}{\partial z^\gamma}.$$

Similarly, if $\mathbf{r} = (r_1, \ldots, r_m)$ is an m-tuple of non-negative real numbers, we define

$$\mathbf{r}^\gamma = r_1^{\gamma_1} \cdots r_m^{\gamma_m}.$$

We can therefore write an entire function f in $\mathcal{E}_m$ as

$$f(z) = \sum_\gamma a_\gamma z^\gamma$$

where a_γ are in $\mathbb{F}$, and for all m-tuples of non-negative real numbers $\mathbf{r}$,

$$\lim_{|\gamma|\to\infty} |a_\gamma| \mathbf{r}^\gamma = 0.$$

We recall that to each m-tuple $\mathbf{r} = (r_1, \ldots, r_m)$ of non-negative real numbers, we can associate a non-Archimedean absolute value $|\ |_{\mathbf{r}}$ on the ring $\mathcal{E}_m$ by defining

$$|f|_{\mathbf{r}} = \sup_{\gamma} |a_\gamma| \mathbf{r}^\gamma,$$

where as above f in $\mathcal{E}_m$ is given by the power series expansion

$$f(z) = \sum_{\gamma} a_\gamma z^\gamma.$$

The non-trivial thing that needs to be checked is that if f and g are two elements of $\mathcal{E}_m$, then $|fg|_{\mathbf{r}} = |f|_{\mathbf{r}} |g|_{\mathbf{r}}$. When all the r_j equal 1 this is worked out, for instance, in [6, §5.1.2]. In general, by extending the field $\mathbb{F}$ if necessary, we may assume the r_j are elements of $|\mathbb{F}^\times|$ and then reduce to the case when all the r_j are 1 by an affine rescaling of the variables.

For our purposes, we only need to consider those $\mathbf{r}$ for which all the r_j are equal, or in other words m-tuples of the form $\mathbf{r} = (r, \ldots, r)$. We will denote the associated absolute value on $\mathcal{E}_m$ by $|\ |_r$. Clearly, we have:

Proposition 2.1. *If f is in $\mathcal{E}_m$, then $|f|_r$ is a non-decreasing function of r.*

If f happens to be a polynomial of degree d, then we easily see that as $r \to \infty$,

$$\log |f|_r = d \log r + O(1).$$

Thus, in our ABC theorems for entire functions, $\log |f|_r$ will play the role played by the degree in the case of polynomials on the left-hand side of the inequalities.

We will make use of the following observation on several occasions:

Corollary 2.2. *If f, g and h are elements of $\mathcal{E}_m$ such that $f = gh$ and if $r_0 > 0$, then for all $r \geq r_0$,*

$$\log |g|_r \leq \log |f|_r + O(1).$$

Proof. By the multiplicativity of $|\ |_r$ and Proposition 2.1,

$$\log |f|_r = \log |g|_r + \log |h|_r \geq \log |g|_r + \log |h|_{r_0},$$

which gives the required inequality. ∎

We recall the elementary

Lemma 2.3 (Logarithmic derivative lemma). *Let f be an entire function in $\mathcal{E}_m$ and let γ be a multi-index. Then*

$$|\partial^\gamma f|_r \leq \frac{|f|_r}{r^{|\gamma|}}.$$

Proof. Differentiate the power series defining f and use the fact that $|k| \leq 1$ for any integer k. ∎

We need to make use of some ring-theoretic properties of $\mathcal{E}_m$. The reader can see [11] for a detailed treatment.

As usual, we will call an element P of $\mathcal{E}_m$ *irreducible* if whenever we write $P = fg$ with f and g in $\mathcal{E}_m$, then at least one of f or g is a unit in $\mathcal{E}_m$. As is well-known, the only units in $\mathcal{E}_m$ are the non-zero constant functions—see e.g. [12, Cor. 2.4].

PROPOSITION 2.4. *Let P be an irreducible element of $\mathcal{E}_m$ and let j be an integer between* 1 *and m. If P divides $\partial P/\partial z_j$, then $\partial P/\partial z_j \equiv 0$.*

Proof. Suppose
$$\frac{\partial P}{\partial z_j} = Pg$$
for some g in $\mathcal{E}_m$. Then
$$\left|\frac{\partial P}{\partial z_j}\right|_r = |P|_r |g|_r.$$
From Lemma 2.3, $|g|_r \leq 1/r$, and so by Proposition 2.1, $|g|_r \equiv 0$. ∎

Suppose f and g are non-constant elements of $\mathcal{E}_m$ such that g divides f in $\mathcal{E}_m$. Then f and g can also be considered as analytic functions on the closed ball of radius r, i.e.,
$$\mathbf{B}^m(r) = \{(z_1, \ldots, z_m) \in \mathbb{F}^m : \max |z_j| \leq r\}.$$
For large r, the function g will have zeros inside the ball (again see [12, Cor. 2.4]), and hence will not be a unit in the ring of analytic functions on the ball $\mathbf{B}^m(r)$. These rings are Tate algebras when $r \in |\mathbb{F}^\times|$, and hence factorial [6, §5.2.6, Th. 1]. Thus, $\mathcal{E}_m$ is a subring of a factorial ring in which g is not a unit, and hence some power of g will not divide f in $\mathcal{E}_m$. Therefore, one can speak of the multiplicity with which an entire function g divides another entire function f. Although $\mathcal{E}_m$ itself is not factorial, the notion of "greatest common divisor" does make sense in $\mathcal{E}_m$; see [11] or the appendix of [12]. Of course, greatest common divisors are only defined up to units, hence multiplicative constants. Given two entire functions f_1 and f_2, when we write something like $g = \gcd(f_1, f_2)$, we mean picking any function g which is a greatest common divisor of f_1 and f_2. Hence g is only well-defined up to a choice of multiplicative constant.

Since we have greatest common divisors, we can define a good notion of the "radical" or the "square free part" of an entire function, at least in characteristic zero. The definition we give here will not be the square free part in positive characteristic, but will be the suitable thing to put on the right-hand side in our basic ABC theorem. We will discuss the existence of the square-free part of an analytic function in positive characteristic in a later section.

Proposition 2.5. *Let f be an entire function in $\mathcal{E}_m$. For each j from* 1 *to m, define*

$$g_j = \gcd\left(f, \frac{\partial f}{\partial z_j}\right) \quad \text{and} \quad h_j = \frac{f}{g_j}.$$

Let $R(f)$ be the least common multiple of the h_j. Then:

(i) *$R(f)$ divides f;*
(ii) *for any non-constant g in $\mathcal{E}_m$, g^2 does not divide $R(f)$;*
(iii) *if P is an irreducible element of $\mathcal{E}_m$ that divides f, then P also divides $R(f)$ if and only if the multiplicity to which P divides f is not divisible by the characteristic of $\mathbb{F}$.*

We will call $R(f)$ as defined in Proposition 2.5 the *radical* of f.

Proof. For (i), note that because each of the h_j divides f, it is clear that their least common multiple $R(f)$ also divides f.

To show (ii), suppose g is a non-constant element of $\mathcal{E}_m$ such that g^2 divides $R(f)$. Then g^2 must also divide f since $R(f)$ divides f. Let $s \geq 2$ be the largest integer such that g^s divides f. Then $f = g^s \widetilde{f}$, where $\widetilde{f}$ is an element of $\mathcal{E}_m$ not divisible by g. Because

$$\frac{\partial f}{\partial z_j} = \widetilde{f} s g^{s-1} \frac{\partial g}{\partial z_j} + g^s \frac{\partial \widetilde{f}}{\partial z_j},$$

we see that g^{s-1} divides g_j for all j, and hence g^2 does not divide h_j for any j. Thus, g^2 cannot divide $R(f)$.

To show (iii), let P be a non-constant irreducible element of $\mathcal{E}_m$ that divides f. Let s be the largest integer such that P^s divides f. Then $f = P^s \widetilde{f}$ with $\widetilde{f}$ relatively prime to P. If $\partial P/\partial z_j \equiv 0$ for all j, then because we have assumed that P is non-constant, it follows that P is a pth power and $\mathbb{F}$ has positive characteristic p. But then P would not be irreducible, and so there must exist some j such that $\partial P/\partial z_j \not\equiv 0$. Because

$$\frac{\partial f}{\partial z_j} = s P^{s-1} \frac{\partial P}{\partial z_j} \widetilde{f} + P^s \frac{\partial \widetilde{f}}{\partial z_j},$$

we conclude from Proposition 2.4 that P^s divides $\partial f/\partial z_j$ if and only if $s = 0$ in $\mathbb{F}$. ∎

We will now state and prove the most basic version of an ABC theorem for non-Archimedean entire functions of several variables. We feel that discussing this basic case here will help the reader see the main ideas behind what we will do.

THEOREM 2.6 (Basic ABC theorem). *Let $f_2 = f_0 + f_1$ be entire functions such that f_0 and f_1 are relatively prime in $\mathcal{E}_m$. If $\mathbb{F}$ has characteristic zero, assume that at least one of f_0 or f_1 is non-constant. If $\mathbb{F}$ has positive characteristic p, then assume that at least one of f_0 or f_1 is not a pth power in $\mathcal{E}_m$. Let $r_0 > 0$. Then, for $r \geq r_0$,*

$$\max_{0 \leq i \leq 2} \log |f_i|_r \leq \log |R(f_0 f_1 f_2)|_r - \log r + O(1).$$

Proof. We follow the standard polynomial proof, as given for instance in [32], *mutatis mutandis.* Without loss of generality assume that f_0 is non-constant and if $\mathbb{F}$ has positive characteristic p that f_0 is not a pth power in $\mathcal{E}_m$. This implies there exists a non-constant irreducible element P_0 in $\mathcal{E}_m$ that divides f_0 to a multiplicity s_0 not divisible by the characteristic of $\mathbb{F}$. Without loss of generality, assume that $\partial P_0/\partial z_1 \not\equiv 0$. Consider the Wronskian determinant,

$$W = \det \begin{pmatrix} f_0 & f_1 \\ \dfrac{\partial f_0}{\partial z_1} & \dfrac{\partial f_1}{\partial z_1} \end{pmatrix} = \det \begin{pmatrix} f_0 & f_2 \\ \dfrac{\partial f_0}{\partial z_1} & \dfrac{\partial f_2}{\partial z_1} \end{pmatrix} = \det \begin{pmatrix} f_2 & f_1 \\ \dfrac{\partial f_2}{\partial z_1} & \dfrac{\partial f_1}{\partial z_1} \end{pmatrix},$$

where the first equality defines W in $\mathcal{E}_m$ and the second two equalities follow from $f_2 = f_0 + f_1$.

We first claim that $W \not\equiv 0$. Indeed, if $W \equiv 0$, then

$$f_0 \frac{\partial f_1}{\partial z_1} = f_1 \frac{\partial f_0}{\partial z_1}.$$

Because $P_0^{s_0}$ divides f_0 and does not divide $\partial f_0/\partial z_1$, this would imply that P_0 divides f_1. But f_0 and f_1 were assumed relatively prime, and hence $W \not\equiv 0$.

Let $F = f_0 f_1 f_2$, $G = \gcd(F, \partial F/\partial z_1)$, and $H = F/G$. Then by definition H divides $R(f_0 f_1 f_2)$, and so

$$\log |H|_r \leq \log |R(f_0 f_1 f_2)|_r + O(1)$$

for $r \geq r_0$ by Corollary 2.2. We also claim that G divides W. Indeed, suppose that P is an irreducible element that divides G. Then P divides F and so it divides one of the f_i and hence exactly one of the f_i since the f_i are relatively prime. Thus, suppose that P divides f_i and hence F with multiplicity s. Then P^{s-1} divides $\partial f_i/\partial z_1$ and hence also W. If P^s also divides G and hence $\partial F/\partial z_1$, then either s is divisible by the characteristic of $\mathbb{F}$ or $\partial P/\partial z_1 = 0$. But in either of these cases, P^s also divides $\partial f_i/\partial z_1$, and so P^s also divides W. Thus, G divides W as claimed. Again applying Corollary 2.2, we see that for $r \geq r_0$,

$$\log |G|_r \leq \log |W|_r + O(1).$$

By Lemma 2.3,
$$\left| f_i \frac{\partial f_j}{\partial z_1} \right|_r \le \frac{|f_i f_j|_r}{r},$$
and hence using each of the three determinants defining W,
$$\log|W|_r \le \log\min\{|f_0 f_1|_r, |f_0 f_2|_r, |f_1 f_2|_r\} - \log r.$$
Hence,
$$\begin{aligned}\log\max|f_i|_r &= \log|f_0|_r + \log|f_1|_r + \log|f_2|_r - \log\min_{0\le i<j\le 2}|f_i f_j|_r\\ &= \log|F|_r - \log\min_{0\le i<j\le 2}|f_i f_j|_r\\ &= \log|H|_r + \log|G|_r - \log\min_{0\le i<j\le 2}|f_i f_j|_r\\ &\le \log|R(F)|_r + \log|W|_r - \log\min_{0\le i<j\le 2}|f_i f_j|_r + O(1)\\ &\le \log|R(F)|_r - \log r + O(1)\end{aligned}$$
for $r \ge r_0$. ∎

We conclude this section with a discussion of counting functions. For a polynomial in one variable, it is a simple matter to count the zeros, with or without multiplicity, because they are finite in number. For a one-variable convergent power series, the zeros are discrete, so one can create a counting function by counting them up to a certain size and then seeing how the number of zeros grows as the maximum size considered is allowed to grow. This is in complete analogy to Nevanlinna's notion of a counting function to count the number of zeros of a complex entire or meromorphic function.

For several-variable polynomials, one generally does not try to "count" zeros. Rather, one counts irreducible factors, usually weighted by the degree of the irreducible factor. For complex holomorphic functions of several variables, including the case of complex polynomials, one can define counting functions in a very geometric way by integrating certain differential forms over the irreducible components of the zero divisor of the function; see e.g. [24] or [28].

One approach to defining non-Archimedean counting functions in several variables is the approach initiated by Hà Huy Khoái [17] and used by An and Manh [1]–[3]. Although this approach is, in principle, aesthetically pleasing because of its definition in terms of the geometry of the Newton polytope associated to a several-variable power series, in practice, working with counting functions defined in this way seems to be rather difficult and not to produce particularly aesthetic proofs. For instance, the difficulty in working with this notion of counting function seems to have something to do with An and Manh's need for some of their restrictive hypotheses in [3]. Moreover, working with this definition seems to obscure connections to proofs of similar results for polynomials. In [12], Cherry and Ye preferred not to give an *a priori* natural definition of counting function, but rather first proved [12, Lem. 2.3] that starting with a power series of several variables, the counting functions of the one-variable power series obtained by restricting to a sufficiently generic line through the origin do not depend on the generic line chosen and can be expressed in terms of the power series coefficients. The pay-back for doing this work first before giving what may seem like an unnatural definition for the counting function is that one can then in a relatively straightforward manner connect Cherry and Ye's notion of counting function with $|f|_r$ through a Poisson–Jensen–Green type formula [12, Th. 3.1]. Then one can work with $|f|_r$ in a relatively straightforward manner and in close analogy with how one would naturally work with a several-variable polynomial.

Suppose

$$f = \sum_{\gamma} a_\gamma z^\gamma$$

is an entire function on $\mathbb{A}^m$. As earlier in this section, let $r > 0$ and let $\mathbf{r} = (r, \ldots, r)$. Cherry and Ye define the *unintegrated counting function* of zeros of f by

$$n_f(0, r) = \sup\{|\gamma| : |a_\gamma|\mathbf{r}^\gamma = |f|_r\}.$$

This is the number of zeros, counting multiplicity, that f has on a sufficiently generic line through the origin with $\max|z_j| \le r$. Also, define

$$n_f(0, 0) = \lim_{r \to 0} n_f(0, r) = \min\{|\gamma| : a_\gamma \neq 0\}.$$

As is typical in Nevanlinna theory, it is more convenient to work with the *integrated counting function* of zeros

$$N_f(0, r) = n_f(0, 0)\log r + \int_0^r (n_f(0, t) - n_f(0, 0))\,\frac{dt}{t}.$$

Immediately from the definition we see that if f is a non-constant entire function, then for $r \ge 1$,

$$\log r \le N_f(0, r) + O(1). \tag{1}$$

Cherry and Ye's Poisson–Jensen–Green formula [12, Th. 3.1] then says that there exists a constant C_f depending on f but not on r such that

$$N_f(0, r) = \log|f|_r + C_f \tag{2}$$

for all r. These counting functions count zeros of f with multiplicity.

The following proposition for counting functions corresponds to Corollary 2.2.

PROPOSITION 2.7. *Let $f = gh$ be entire functions. Then*:

(i) $n_f(0,r) = n_g(0,r) + n_h(0,r)$ *for all* $r \geq 0$;
(ii) $N_f(0,r) = N_g(0,r) + N_h(0,r)$ *for all* $r \geq 0$;
(iii) $N_f(0,r) \geq N_g(0,r)$ *for all* $r \geq 1$.

Proof. The equality in (ii) follows immediately from (i) and the definition of the integrated counting functions. The inequality in (iii) follows from (ii) and the fact that $N_h(0,r) \geq 0$ if $r \geq 1$. Thus, we need to show (i). To do so, let

$$f(z) = \sum_{\alpha} a_\alpha z^\alpha, \qquad g(z) = \sum_{\beta} b_\beta z^\beta, \qquad h(z) = \sum_{\gamma} c_\gamma z^\gamma.$$

We leave the case $r = 0$ for the reader. Let $r > 0$. Let β_0 and γ_0 be the largest multi-indices in the graded lexicographical ordering such that

$$|b_{\beta_0}|r^{|\beta_0|} = |g|_r \quad \text{and} \quad |c_{\gamma_0}|r^{|\gamma_0|} = |h|_r$$

respectively. By definition, $n_g(0,r) = |\beta_0|$ and $n_h(0,r) = |\gamma_0|$. Therefore for $|\alpha| > |\beta_0| + |\gamma_0|$, we have

$$|a_\alpha|r^{|\alpha|} \leq \max_{\beta+\gamma=\alpha} |b_\beta|r^{|\beta|}|c_\gamma|r^{|\gamma|} < |g|_r|h|_r = |f|_r,$$

where the second inequality follows from the fact that if

$$|\beta| + |\gamma| = |\alpha| > |\beta_0| + |\gamma_0|,$$

then we must have

$$|\beta| > |\beta_0| \quad \text{or} \quad |\gamma| > |\gamma_0|.$$

On the other hand, if we consider $\alpha_0 = \beta_0 + \gamma_0$, then

$$\alpha_0 = \sum_{\beta+\gamma=\alpha_0} b_\beta c_\gamma.$$

If $\beta \neq \beta_0$ (and so $\gamma \neq \gamma_0$), then either β comes after β_0 or γ comes after γ_0 in the graded lexicographical ordering, which means

$$|b_\beta|r^{|\beta|}|c_\gamma|r^{|\gamma|} < |b_{\beta_0}|r^{|\beta_0|}|c_{\gamma_0}|r^{|\gamma_0|},$$

and so

$$|a_{\alpha_0}|r^{|\alpha_0|} = |b_{\beta_0}|r^{|\beta_0|}|c_{\gamma_0}|r^{|\gamma_0|} = |g|_r|h|_r = |f|_r.$$

Thus,

$$n_f(0,r) = |\alpha_0| = |\beta_0| + |\gamma_0| = n_g(0,r) + n_h(0,r). \ \blacksquare$$

In [12], Cherry and Ye did not discuss truncated counting functions, where zeros are counted without multiplicity or with their multiplicities "truncated" to a certain level. In complex Nevanlinna theory, since one has a natural geometric definition for counting functions defined as integrals over irreducible components of an analytic divisor, it is straightforward to define truncated counting functions. Since Cherry and Ye's definition of counting functions is given in terms of power series coefficients, it is clear that there will be no obvious definition of truncated counting functions in terms of the power series coefficients. Instead, we use the nice ring-theoretic properties of $\mathcal{E}_m$ discussed above, and *in characteristic zero* simply define *truncated counting functions* by

$$n_f^{(1)}(0,r) = n_{R(f)}(0,r) \quad \text{and} \quad N_f^{(1)}(0,r) = N_{R(f)}(0,r),$$

where, as before, $R(f)$ denotes the radical of f. Note that although $R(f)$ is only defined up to a multiplicative constant, $n^{(1)}$ and $N^{(1)}$ are well-defined. In characteristic zero, Proposition 2.5 justifies calling the counting functions of the radical the "truncated" counting function for f because each irreducible factor of f appears with multiplicity one in $R(f)$. In positive characteristic, $N_{R(f)}$ might be called "overly truncated" because it completely ignores all irreducible factors of f which appear with multiplicity divisible by the characteristic. We will see in Section 4 how to define truncated counting functions in positive characteristic that include all irreducible factors. However, as we saw in Theorem 2.6, in positive characteristic we can sometimes give lower bounds on these overly truncated counting functions.

We complete this section by pointing out that Boutabaa and Escassut [7, 8] were the first to work out one-variable non-Archimedean Nevanlinna theory in positive characteristic. Their work also highlights that in working with Nevanlinna theory in positive characteristic, one may often ignore zeros whose multiplicity is divisible by the characteristic.

3. Hasse derivatives and generalized Wronskians. If $\mathbb{F}$ has characteristic zero, then a formal power series f in the variables $z = (z_1, \dots, z_m)$ is non-constant if and only if at least one of its formal partial derivatives $\partial f/\partial z_j$ is not identically zero. By contrast, if $\mathbb{F}$ has positive characteristic p, then any formal power series in $z^p = (z_1^p, \dots, z_m^p)$ is such that all its partial derivatives $\partial^\gamma f$ are identically zero for all $|\gamma| > 0$. Also, if $\mathbb{F}$ has positive characteristic p, then if γ is a multi-index such that $\gamma_i \geq p$ for some i, then $\partial^\gamma f = 0$ for all f. Therefore, we will introduce a modification of the standard derivative, known as the Hasse derivative, which is more useful in positive characteristic.

If $\alpha = (\alpha_1, \dots, \alpha_m)$ and $\beta = (\beta_1, \dots, \beta_m)$ are multi-indices, we use $\alpha + \beta$ to denote the multi-index

$$\alpha + \beta = (\alpha_1 + \beta_1, \dots, \alpha_m + \beta_m).$$

We say that $\alpha \geq \beta$ if $\alpha_i \geq \beta_i$ for all i from 1 to m. Note that this notion of $\geq$ is not a total ordering on the set of multi-indices and is *not* the graded lexicographical ordering that was used in the proof of Proposition 2.7. If

$\alpha \geq \beta$, we use $\alpha - \beta$ to denote the multi-index

$$\alpha - \beta = (\alpha_1 - \beta_1, \ldots, \alpha_m - \beta_m).$$

Also, if $\alpha \geq \beta$, define the multinomial coefficient $\binom{\alpha}{\beta}$ by

$$\binom{\alpha}{\beta} = \binom{\alpha_1}{\beta_1} \cdots \binom{\alpha_m}{\beta_m},$$

where the $\binom{\alpha_i}{\beta_i}$ are the standard binomial coefficients. Given a formal power series

$$f = \sum_{\alpha} a_\alpha z^\alpha$$

and a multi-index γ, we define the *Hasse derivative* of multi-index γ of f, which we will denote $D^\gamma f$, to be the formal power series defined by

$$D^\gamma f = \sum_{\alpha \geq \gamma} \binom{\alpha}{\gamma} a_\alpha z^{\alpha - \gamma}.$$

Note that if $\gamma = (0, \ldots, 0)$, then $D^\gamma f = f$, and that if $|\gamma| = 1$, then $D^\gamma f = \partial^\gamma f$. Given a j from 1 to m and a positive integer k, we will use D_j^k as a short-hand notation for $D^\gamma f$ where $\gamma = (\gamma_0, \ldots, \gamma_m)$ with $\gamma_j = k$ and $\gamma_i = 0$ for $i \neq j$.

Because the multinomial coefficients $\binom{\alpha}{\gamma}$ are integers and hence have non-Archimedean absolute value at most 1, we see that if $\mathbf{r} = (r_1, \ldots, r_m)$ is an m-tuple of non-negative real numbers such that

$$\lim_{|\alpha| \to \infty} |a_\alpha| \mathbf{r}^\alpha = 0,$$

then

$$\lim_{|\alpha| \to \infty} \left| \binom{\alpha}{\gamma} a_\alpha \right| \mathbf{r}^{\alpha - \gamma} \leq \frac{1}{\mathbf{r}^\gamma} \lim_{|\alpha| \to \infty} |a_\alpha| \mathbf{r}^\alpha = 0,$$

and so we see that if f is in $\mathcal{E}_m$, then $D^\gamma f$ is also in $\mathcal{E}_m$.

Clearly,

$$\partial^\gamma f = \gamma! D^\gamma f, \qquad \text{where } \gamma! = \gamma_1! \cdots \gamma_m!.$$

Thus, in characteristic zero, the Hasse derivatives are just constant multiples of the ordinary derivatives, and so one sees immediately that they have similar properties to those of the ordinary partial derivative. In positive characteristic, one must check these.

PROPOSITION 3.1. *The Hasse derivatives have the following basic properties*:

(i) $D^\alpha[f + g] = D^\alpha f + D^\alpha g$;
(ii) $D^\alpha[fg] = \sum_{\beta + \gamma = \alpha} D^\beta f D^\gamma g$;
(iii) $D^\alpha D^\beta f = \binom{\alpha+\beta}{\beta} D^{\alpha+\beta} f$;

(iv) *if* $\mathbb{F}$ *has positive characteristic* p *and* $s \geq 0$ *is an integer, then*

$$D_i^{p^s} f^{p^s} = (D_i f)^{p^s}.$$

Proof. Property (i) is obvious. To check property (ii), write out both sides and compare like powers of z. What is needed for equality is that for multi-indices δ and ε with $\delta \geq \beta$ and $\varepsilon \geq \gamma$, one has

$$\sum_{\beta+\gamma=\alpha} \binom{\delta}{\beta}\binom{\varepsilon}{\gamma} = \binom{\delta+\varepsilon}{\alpha},$$

which is nothing other than Vandermonde's identity. To check property (iii), one uses the elementary observation that for $\gamma \geq \alpha + \beta$,

$$\binom{\alpha+\beta}{\beta}\binom{\gamma}{\alpha+\beta} = \binom{\gamma}{\alpha}\binom{\gamma-\alpha}{\beta}.$$

What one needs for (iv) is the fact that for any integer j,

$$\binom{jp^s}{p^s} \equiv j \bmod p,$$

which follows immediately from Lucas's theorem. ∎

We also want to point out that as with ordinary partial derivatives, the same proof as in Lemma 2.3 gives

Lemma 3.2 (Logarithmic derivative lemma). *Let* f *be an entire function in* $\mathcal{E}_m$ *and let* γ *be a multi-index. Then*

$$|D^\gamma f|_r \leq \frac{|f|_r}{r^{|\gamma|}}.$$

Corollary 3.3. *Let* f *be an entire function in* $\mathcal{E}_m$ *and let* γ *be a multi-index with* $|\gamma| > 0$. *If* f *divides* $D^\gamma f$, *then* $D^\gamma f \equiv 0$.

Proof. This follows from Lemma 3.2 and Corollary 2.2. ∎

We will denote the fraction field of $\mathcal{E}_m$ by $\mathcal{M}_m$ and call it the field of *meromorphic functions* on $\mathbb{A}^m$. One sees immediately that one can use Proposition 3.1(ii) to inductively extend the Hasse derivatives to the field $\mathcal{M}_m$ and that the four properties of Proposition 3.1 continue to hold for functions in $\mathcal{M}_m$.

For each integer $k \geq 2$, let

$$\mathcal{M}_m[k] = \{Q \in \mathcal{M}_m : D_j^i Q \equiv 0 \text{ for all } 0 < i < k \text{ and } 1 \leq j \leq m\}.$$

If $\mathbb{F}$ has positive characteristic p and if s is a positive integer, let

$$\mathcal{E}_m[p^s] = \{g^{p^s} : g \in \mathcal{E}_m\}.$$

Note that $\mathcal{E}_m[p^s]$ is a subring of $\mathcal{E}_m$ and that it consists of those elements f in $\mathcal{E}_m$ that can be written as convergent power series in

$$z^{p^s} = (z_1^{p^s}, \ldots, z_m^{p^s}).$$

PROPOSITION 3.4.

(A) *If* $\mathbb{F}$ *has characteristic* 0, *then for all* $k \geq 2$, *we have* $\mathcal{M}_m[k] = \mathbb{F}$.

(B) *If* $\mathbb{F}$ *has positive characteristic* p *and if* s *is an integer* ≥ 1, *then*

- (B1) $\mathcal{M}_m[p^{s-1}+1] = \mathcal{M}_m[p^s]$;
- (B2) $\mathcal{M}_m[p^s]$ *is the fraction field of* $\mathcal{E}_m[p^s]$ *and* $D_i^{p^s}$ *for* $i = 1, \ldots, m$ *are derivations on* $\mathcal{M}_m[p^s]$.

Proof. Clearly, we have (A). Proposition 3.1(iii) implies (B1). We show (B2) by induction on s. Let Q be an element of $\mathcal{M}_m[2] = \mathcal{M}_m[p]$, which implies $\partial Q/\partial z_i \equiv 0$ for all $i = 1, \ldots, m$. Write $Q = f/g$ with f and g relatively prime in $\mathcal{E}_m$. Suppose that f is not in $\mathcal{E}_m[p]$. Then there is an irreducible element P of $\mathcal{E}_m$ such that P divides f to a multiplicity not divisible by p and such that $\partial P/\partial z_i \not\equiv 0$ for some i. This implies that P divides f to a higher power than it divides $\partial f/\partial z_i$. Since $\partial Q/\partial z_i \equiv 0$, this would then imply P must divide g, contradicting the fact that f and g are relatively prime. Hence f must have been in $\mathcal{E}_m[p]$. Similarly, g must be in $\mathcal{E}_m[p]$. That D_i^p is a derivation on $\mathcal{M}_m[p]$ then follows from Proposition 3.1(ii) or (iv). By (B1), we have

$$\mathcal{M}_m[p^{s+1}] = \{Q \in \mathcal{M}_m[p^s] : D_i^{p^s} Q = 0 \text{ for all } i = 1, \ldots, m\},$$

and so the proof is completed by induction. Indeed, writing an element Q of $\mathcal{M}_m[p^{s+1}] \subset \mathcal{M}_m[p^s]$ as f/g with f and g relatively prime elements of $\mathcal{E}_m[p^s]$ and arguing as before using $D_i^{p^s}$ in place of the first partials, we see that every irreducible element P of $\mathcal{E}_m$ that divides either f or g must divide them with multiplicity divisible by p^{s+1}, and hence f and g must be in $\mathcal{E}_m[p^{s+1}]$. ■

THEOREM 3.5 (Hsia–Wang [18, Lem. 2]). *Let* $\mathbb{F}$ *have characteristic zero* (*resp. positive characteristic* p), *and let* $s \geq 1$ *be an integer. Let* $f = (f_0, \ldots, f_{n-1})$ *be an* n*-tuple of entire functions. For a multi-index* γ, *let*

$$D^\gamma f = (D^\gamma f_0, \ldots, D^\gamma f_{n-1}).$$

Let γ^0 *be the multi-index* $(0, \ldots, 0)$. *If* $f_0, \ldots, f_{n-1}$ *are linearly independent over* $\mathbb{F}$ (*resp.* $\mathcal{M}_m[p^s]$), *then there exist multi-indices* $\gamma^1, \ldots, \gamma^{n-1}$ *such that*

$$|\gamma^i| \leq |\gamma^{i-1}| + 1 \qquad (\textit{resp. } |\gamma^i| \leq |\gamma^{i-1}| + p^{s-1})$$

and such that

$$\det \begin{pmatrix} f_0 & \cdots & f_{n-1} \\ D^{\gamma^1} f_0 & \cdots & D^{\gamma^1} f_{n-1} \\ D^{\gamma^2} f_0 & \cdots & D^{\gamma^2} f_{n-1} \\ \vdots & \vdots & \vdots \\ D^{\gamma^{n-1}} f_0 & \cdots & D^{\gamma^{n-1}} f_{n-1} \end{pmatrix} \not\equiv 0.$$

REMARK. The determinant in Theorem 3.5 is called a *generalized Wronskian.* For polynomials in characteristic zero, this theorem, with a different proof, appears in [27]. In the case of complex entire functions of several variables, a similar theorem can be found in [15].

REMARK. Often (e.g. [18]), one tends to state this lemma with

$$|\gamma^i| \leq i \quad (\text{resp. } |\gamma^i| \leq ip^{s-1}),$$

but we will want to give a lower bound on $\sum |\gamma^i|$ in terms of $|\gamma^{n-1}|$, and so for us the observation that

$$|\gamma^{i-1}| \geq |\gamma^i| - 1 \quad (\text{resp. } |\gamma^{i-1}| \geq |\gamma^i| - p^{s-1})$$

is important.

REMARK. We also remark here that Theorem 3.5 can be used to derive a positive characteristic Cartan-type second main theorem for linearly non-degenerate non-Archimedean analytic curves in projective space. For instance, the proof given in [12, Th. 5.1] goes through once a non-vanishing generalized Wronskian exists.

Proof of Theorem 3.5. We write the proof in the case of positive characteristic. The same proof works in characteristic zero by Proposition 3.4(A) if all powers of p are replaced by 1.

We proceed by induction on n. When $n = 1$, the theorem is trivial. Now assume that the theorem is true for $n-1$. By the induction hypothesis, there exist multi-indices $\gamma^0, \gamma^1, \ldots, \gamma^{n-2}$ with

$$|\gamma^i| \leq |\gamma^{i-1}| + p^{s-1}$$

and such that the $D^{\gamma^i} f$ for $i = 0, \ldots, n-2$ span an $n-1$-dimensional $\mathcal{M}_m$-vector subspace of $\mathcal{M}_m^n$. Let $k = |\gamma^{n-2}| + p^{s-1}$. Let V be the $\mathcal{M}_m$-vector subspace of $\mathcal{M}_m^n$ spanned by $D^\gamma f$ for all $|\gamma| \leq k$. If the theorem were not true, then V could not have dimension n, and so there exist $Q_0, \ldots, Q_{n-1}$ not all zero in $\mathcal{M}_m$ such that

$$Q_0 D^\gamma f_0 + \cdots + Q_{n-1} D^\gamma f_{n-1} \equiv 0 \tag{3}$$

for every γ with $|\gamma| \leq k$. Because the vectors

$$(D^{\gamma^0} f_j, \ldots, D^{\gamma^{n-2}} f_j) \quad \text{for } j = 0, \ldots, n-2$$

are linearly independent over $\mathcal{M}_m$ by the induction hypothesis, we can assume $Q_{n-1} \equiv 1$, and hence we have

$$Q_0 D^\gamma f_0 + \cdots + Q_{n-2} D^\gamma f_{n-2} + f_{n-1} \equiv 0 \quad \text{for all } |\gamma| \leq k.$$

Our goal is to show that the Q_j are in $\mathcal{M}_m[p^s]$. Note that

$$0 = D_l^1\Big[\sum_{j=0}^{n-2} Q_j D^{\gamma^i} f_j + D^{\gamma^i} f_{n-1}\Big] = \sum_{j=0}^{n-2} D_l^1 Q_j D^{\gamma^i} f_j + \sum_{j=0}^{n-1} Q_j D_l^1 D^{\gamma^i} f_j$$
$$= \sum_{j=0}^{n-2} D_l^1 Q_j D^{\gamma^i} f_j$$

for $i = 0, \ldots, n-2$ and $l = 1, \ldots, m$, where the last line follows from Proposition 3.1(iii) and equation (3). By the linear independence of the vectors $(D^{\gamma^0} f_j, \ldots, D^{\gamma^{n-2}} f_j)$, we conclude that $D_l^1 Q_j \equiv 0$ for all l and all $j = 1, \ldots, n-2$. Thus, the Q_j belong to $\mathcal{M}_m[2] = \mathcal{M}_m[p]$. Now assume the Q_j belong to $\mathcal{M}_m[p^t]$ for some $t \geq 1$. By Proposition 3.4(B), we can apply $D_l^{p^{t+1}}$ as if it were a derivation to get

$$0 = D_l^{p^{t+1}}\Big[\sum_{j=0}^{n-2} Q_j D^{\gamma^i} f_j + D^{\gamma^i} f_{n-1}\Big] \tag{4}$$
$$= \sum_{j=0}^{n-2} D_l^{p^{t+1}} Q_j D^{\gamma^i} f_j + \sum_{j=0}^{n-1} Q_j D_l^{p^{t+1}} D^{\gamma^i} f_j.$$

If $t < s$, we can use Proposition 3.1(iii), the fact that

$$|\gamma^i| \leq |\gamma^{n-2}| \leq (k-1)p^{s-1}$$

and equation (3) to conclude that the right-hand sum in (4) vanishes, and thus,

$$0 = \sum_{j=0}^{n-2} D_l^{p^{t+1}} Q_j D^{\gamma^i} f_j.$$

Again, by linear independence, we conclude $D_l^{p^{t+1}} Q_j \equiv 0$ and so the Q_j belong to $\mathcal{M}_m[p^{t+1}]$. Continuing in this manner, we find that the Q_j are in $\mathcal{M}_m[p^s]$, contradicting the assumption that the f_j are linearly independent over $\mathcal{M}_m[p^s]$. ∎

PROPOSITION 3.6. *Let f be an entire function in $\mathcal{E}_m$. Let $\gamma = (\gamma_1, \ldots, \gamma_m)$ be a multi-index. Let P be an irreducible element of $\mathcal{E}_m$ that divides f with exact multiplicity e. If $e > |\gamma|$, then $P^{e-|\gamma|}$ divides $D^\gamma f$. Moreover, if $\operatorname{char} \mathbb{F} = p > 0$ and if e is divisible by $p^s > \max\{\gamma_1, \ldots, \gamma_m\}$, then P^e divides $D^\gamma f$.*

Proof. Because the D_i commute, it suffices to show the proposition for $D^\gamma = D_i^k$.

In the special case that $\operatorname{char} \mathbb{F} = p > 0$ and e is divisible by

$$p^s > \max\{\gamma_1, \ldots, \gamma_m\},$$

the fact that P^e divides $D_i^k f$ follows easily from Proposition 3.1(ii) since for any non-negative integer l,

$$D_i^j P^{lp^s} = 0 \quad \text{for all } 0 < j < p^s$$

and since $p^s > k$ by assumption.

We show the general case by induction on e and k. The case $e = k$ is trivial. We now suppose that the proposition holds for all D_i^j with $j \le k$ and for e and then show that it also holds for k and $e+1$. Suppose $f = P^{e+1}g$ with g relatively prime to P. By Proposition 3.1(ii),

$$D_i^k f = D_i^k(P \cdot P^e g) = P D_i^k(P^e g) + \sum_{j=1}^{k} D_i^j P D_i^{k-j}(P^e g).$$

If $e > k \ge k - j$, then by induction, $P^{e-(k-j)}$ divides $D_i^{k-j}(P^e g)$ and hence P^{e-k} divides $D_i^{k-j}(P^e g)$ for all $j = 0, \dots, k$ and P^{e+1-k} divides $D_i^{k-j}(P^e g)$ for all $j > 0$. ∎

4. Higher radicals and truncated counting functions. If $\mathbb{F}$ has characteristic zero, if f is in $\mathcal{E}_m$, and if $l \ge 1$ is an integer, then clearly if P is an irreducible element of $\mathcal{E}_m$ that divides f with multiplicity e, then P divides $\gcd(f, R(f)^l)$ with multiplicity $\min\{l, e\}$. Thus, *in characteristic zero*, we can define the *lth truncated counting function* by

$$N_f^{(l)}(0, r) = N_{\gcd(f, R(f)^l)}(0, r).$$

We saw at the end of Section 2 that in positive characteristic p, the radical $R(f)$ does not contain those irreducible factors of f that divide f with multiplicity divisible by p. Although that was exactly what was appropriate in Theorem 2.6, when we consider

$$f_n = f_0 + \dots + f_{n-1}$$

with $n > 2$, we will not be able to ignore all irreducible factors with multiplicity divisible by p. Thus, we want to define truncated counting functions in positive characteristic that include all irreducible factors.

For the rest of this section, let $\mathbb{F}$ have positive characteristic p. We will use the following proposition to inductively define *higher p^s-radicals* for integers $s \ge 1$.

PROPOSITION 4.1. *Let f be an entire function in $\mathcal{E}_m$ and let $s \ge 1$ be an integer. Assume that we have defined a p^{s-1}-radical $R_{p^{s-1}}(f)$ that has the property that $R_{p^{s-1}}(f)$ is square-free and has the property that an irreducible element P of $\mathcal{E}_m$ divides $R_{p^{s-1}}(f)$ if and only if P divides f with multiplicity not divisible by p^s. Let*

$$\bar{f} = \frac{f}{\gcd(f, R_{p^{s-1}}(f)^{p^s})}.$$

For $i = 1, \ldots, m$, let $g_i = \gcd(\bar{f}, D_i^{p^s} \bar{f})$, let $h_i = \bar{f}/g_i$, and let H be the least common multiple of the h_i. Let

$$G = \frac{H}{\gcd(H, R_{p^{s-1}}(H)^{p^s-1})}.$$

Then:

(i) *if P is an irreducible element of $\mathcal{E}_m$ that divides G, then it divides G with multiplicity exactly p^s;*
(ii) *if P is an irreducible element of $\mathcal{E}_m$, then P divides G if and only if P divides f with multiplicity a multiple of p^s but not a multiple of p^{s+1}.*

It follows that G is a p^sth power, so we can let R be a p^sth root of G and let $R_{p^s}(f)$ be the least common multiple of $R_{p^{s-1}}(f)$ and R.

Proof. Our induction begins with the radical as defined in Section 2, so we let $R_{p^0}(f) = R(f)$. To show the inductive step, let P be an irreducible element of $\mathcal{E}_m$ that divides $\bar{f}$. Note that if P divides $\bar{f}$, then it divides it with multiplicity at least p^s. Write

$$\bar{f} = P^{jp^s+e}\widetilde{f},$$

where $j \geq 1$ and $0 \leq e < p^s$ are integers and $\widetilde{f}$ is relatively prime to P. Then, by Proposition 3.1(ii) and (iv) and Proposition 3.4(B), for $i = 1, \ldots, m$,

$$D_i^{p^s} \bar{f} = D_i^{p^s}[P^{jp^s} P^e \widetilde{f}] = P^e \widetilde{f}(D_i[P^j])^{p^s} + P^{jp^s} D_i^{p^s}[P^e \widetilde{f}]. \tag{5}$$

We first consider the case that j is not divisible by p. Because P is irreducible and hence not a pth power, there exists an i such that $D_i P \not\equiv 0$. Because

$$D_i(P^j) = jP^{j-1}D_i P,$$

we see from the assumption that j is not divisible by p, that P divides $D_i^{p^s} f$ with exact multiplicity $(j-1)p^s + e$, and so by Proposition 3.6, P divides H with exact multiplicity p^s.

In case j is divisible by p, we see from equation (5) that P divides H with multiplicity at most $e < p^s$. Thus, P does not divide R. ∎

We now show the existence of the square-free part of an entire function, which is square-free and contains all the irreducible factors dividing the function.

THEOREM 4.2. *Let f be an entire function in $\mathcal{E}_m$. There exists an entire function $S(f)$ in $\mathcal{E}_m$ such that $S(f)$ is square-free and such that an irreducible element P in $\mathcal{E}_m$ divides $S(f)$ if and only if it divides f.*

We will call $S(f)$ the *square-free part* of f. We define the *lth truncated counting function* by

$$N_f^{(l)}(0,r) = N_{\gcd(f,S(f)^l)}(0,r).$$

As in characteristic zero, the lth truncated counting function truncates to l all multiplicities higher than l.

Proof of Theorem 4.2. The proof is similar to the proof of the existence of greatest common divisors given in [11]. The case that f is identically zero is trivial, so assume that f is not identically zero. For $i = 1, 2, \ldots$, let r_i be an increasing sequence of elements of $|\mathbb{F}^\times|$ such that $r_i \to \infty$. Consider f as an element of the ring $\mathcal{A}_m(r_i)$ of analytic functions on $\mathbf{B}^m(r_i)$, the closed ball of radius r_i. Let z_0 be a point in $\mathbf{B}^m(r_1)$ such that $f(z_0) \neq 0$. Let $R_{p^s}(f)$ be the higher radicals of f defined as in Proposition 4.1 normalized so that for each s, we have $R_{p^s}(f)(z_0) = 1$. Because the ring $\mathcal{A}_m(r_i)$ is factorial [6, §5.2.6, Th. 1], only finitely many of the irreducible elements in $\mathcal{E}_m$ that divide f are non-units in $\mathcal{A}_m(r_i)$. Each of these divides f to some finite multiplicity, and so there exists some s_i such that every irreducible element in $\mathcal{E}_m$ that divides f and is not a unit in $\mathcal{A}_m(r_i)$ also divides $R_{p^s}(f)$ for all $s \geq s_i$. This means that for $s, t \geq s_i$, $R_{p^s}(f)$ and $R_{p^t}(f)$ differ by a unit in $\mathcal{A}_m(r_n)$.

Let $u_{i,i+1}$ be the unit in $\mathcal{A}^m(r_i)$ such that

$$R_{p^{s_i}}(f) = u_{i,i+1} R_{p^{s_{i+1}}}(f),$$

and note that $u_{i,i+1}(z_0) = 1$. Then, writing $u_{i,i+1}$ as a power series about z_0, we have

$$u_{i,i+1}(z) = 1 + \sum_{|\gamma|\geq 1} a_\gamma (z - z_0)^\gamma \qquad \text{with } |a_\gamma| r_i^{|\gamma|} < 1 \text{ for all } |\gamma| \geq 1.$$

Thus, for $j > i$,

$$|u_{j,j+1} - 1|_{r_i} < \frac{r_i}{r_j}.$$

Since $r_i/r_j \to 0$ as $j \to \infty$ with i fixed, we can define units v_i in $\mathcal{A}_m(r_i)$ by

$$v_i = \prod_{j=i}^{\infty} u_{j,j+1}.$$

For $j > i$, we have

$$\begin{aligned} R_{p^{s_j}}(f) v_i = R_{p^{s_j}}(f) \prod_{k=i}^{\infty} u_{k,k+1} &= R_{p^{s_j}}(f) \Big(\prod_{k=i}^{j-1} u_{k,k+1}\Big)\Big(\prod_{k=j}^{\infty} u_{k,k+1}\Big) \\ &= R_{p^{s_i}}(f) v_j. \end{aligned}$$

This precisely means that $R_{p^{s^i}}(f)v_i^{-1}$ converges to an entire function F as $i \to \infty$ because the difference between $R_{p^{s^i}}(f)v_i^{-1}$ and $R_{p^{s^j}}(f)v_j^{-1}$ is identically zero on $\mathbf{B}^m(r_i)$. Note also that $v_i F = R_{p^{s_i}}(f)$ in $\mathcal{A}_m(r_i)$.

We claim that F is square free and that an irreducible element P of $\mathcal{E}_m$ divides F if and only if it divides f.

To show that F is square free, suppose that P is an irreducible element of $\mathcal{E}_m$ such that P^2 divides F. This means P^2 divides $R_{p^{s_i}}(f)$ in $\mathcal{A}_m(r_i)$. For i sufficiently large, P is not a unit in $\mathcal{A}_m(r_i)$, and so it would be the case that $R_{p^{s_i}}(f)$ is not square free in $\mathcal{A}_m(r_i)$. However, the proof of Proposition 4.1 works equally well in the ring $\mathcal{A}_m(r_i)$, and thus $R_{p^{s_i}}(f)$ is also square free in $\mathcal{A}_m(r_i)$.

Finally, let P be an irreducible element of $\mathcal{E}_m$. Suppose P divides f. Then P divides $R_{p^{s_i}}(f)$ for all i sufficiently large. In other words, there exist analytic functions h_i in $\mathcal{A}_m(r_i)$ such that

$$Ph_i = R_{p^{s_i}}(f).$$

Because $Ph_i v_i^{-1}$ converges to F as $i \to \infty$, for $j > i$, we have

$$P(h_i v_i^{-1} - h_j v_j^{-1}) = 0 \quad \text{in } \mathcal{A}_m(r_i).$$

Thus $h_i v_i^{-1}$ converges to an entire function H such that $PH = F$, and so P divides F. For the other direction, suppose that P divides F. Then P divides $R_{p^{s_i}}(f)$ in $\mathcal{A}_m(r_i)$, and so again noticing that the proof of Proposition 4.1 also works for $\mathcal{A}_m(r_i)$, we deduce that P divides f in $\mathcal{A}_m(r_i)$. In other words, there exist analytic functions g_i in $\mathcal{A}_m(r_i)$ such that $Pg_i = f$. This implies that the g_i converge to an analytic function G such that $PG = f$, and hence P divides f in $\mathcal{E}_m$. ∎

5. Linear algebra. Let V be a vector space over a field $\mathbb{E}$. Let $v_0, \ldots, v_n$ be $n+1$ linearly dependent vectors in V. Call an index set $I \subset \{0, \ldots, n\}$ *minimal* if the set of vectors $\{v_i : i \in I\}$ is linearly dependent, but such that for every proper subset $I' \subsetneq I$, the sets of vectors $\{v_i : i \in I'\}$ are linearly independent.

LEMMA 5.1 (Brownawell–Masser). *Let $v_0, \ldots, v_n$ be $n+1$ vectors in a vector space V over a field $\mathbb{E}$ such that $\sum v_i = 0$. Assume that no proper subsum vanishes*, *i.e.*,

$$\sum_{i \in I} v_i \neq 0 \quad \textit{for all } I \subsetneq \{1, \ldots, n\}.$$

Then there exists an integer $u \geq 1$, a partition

$$\{0, \ldots, n\} = I_0 \cup \cdots \cup I_{u-1},$$

and non-empty subsets

$$J_l \subset \bigcup_{j=1}^{l} I_j, \qquad l = 0, \ldots, u-2,$$

such that I_0 *and* $I_j \cup J_{j-1}$ *for* $j = 1, \ldots, u-1$ *are minimal.*

Proof. If $\{0, \ldots, n\}$ is minimal, set $I_0 = \{0, \ldots, n\}$. If $\{0, \ldots, n\}$ is not minimal, see [10, Lem. 6]. ∎

In positive characteristic p, we want to apply Theorem 3.5 to entire functions linearly independent over $\mathbb{F}$, so we complete this section by proving that if a collection of functions are linearly independent over $\mathbb{F}$, then they are also linearly independent over $\mathcal{M}_m[p^s]$ for some integer $s \geq 1$.

LEMMA 5.2. *Let* $f_1, \ldots, f_n$ *be meromorphic functions in* $\mathcal{M}_m$ *linearly independent over* $\mathbb{F}$, *with* $\operatorname{char} \mathbb{F} = p > 0$. *Then there exists an integer* $s \geq 1$ *such that* $f_1, \ldots, f_n$ *are linearly independent over* $\mathcal{M}_m[p^s]$.

Proof. Suppose the lemma is not true. Then, $f_1, \ldots, f_n$ are linearly dependent over $\mathcal{M}_m[p^s]$ for every $s \geq 1$. For each $s \geq 1$, let $I_s \subset \{1, \ldots, n\}$ be minimal. Note that each I_s contains at least two indices, otherwise one of the functions f_j would be identically zero, and hence the f_j could not be linearly independent over $\mathbb{F}$. Because there are only finitely many possible subsets I_s, we may assume without loss of generality that $I_s = \{1, \ldots, t\}$ for infinitely many s. Thus, for infinitely many s, the functions $f_2, \ldots, f_t$ are linearly independent over $\mathcal{M}_m[p^s]$ and f_1 is in the $\mathcal{M}_m[p^s]$-linear span of $f_2, \ldots, f_t$. In other words, for infinitely many s there exist unique $Q_{s,2}, \ldots, Q_{s,t}$ in $\mathcal{M}_m[p^s]$ such that

$$f_1 = Q_{s,2} f_2 + \cdots + Q_{s,t} f_t.$$

On the other hand, $\mathcal{M}_m[p^{s'}] \subset \mathcal{M}_m[p^s]$ if $s' \geq s$, so by the linear independence of $f_2, \ldots, f_t$, the $Q_{s,j}$ do not depend on s. Hence, $Q_{s,j}$ is in $\mathcal{M}_m[p^s]$ for infinitely many s, and must therefore be in $\mathbb{F}$. This contradicts the linear independence of the f_j over $\mathbb{F}$. ∎

If $\mathbb{F}$ has positive characteristic p, then we define the *index of independence* of a collection $\mathcal{F}$ of entire or meromorphic functions to be the smallest integer s such that any subset of functions in $\mathcal{F}$ linearly independent over $\mathbb{F}$ remains linearly independent over $\mathcal{M}_m[p^s]$, provided such an integer exists. Lemma 5.2 shows that such an integer always exists if $\mathcal{F}$ is finite.

6. ABC theorems. In the three-function ABC theorem, one begins with $f_2 = f_0 + f_1$ with the f_i relatively prime. Note that here $\gcd(f_0, f_1, f_2) = 1$ implies that the f_j are also *pairwise* relatively prime because, by the linear dependence, if two of the functions have a common factor, it divides the third as well. To generalize to $n+1$ functions, one obviously wants to consider

$$f_n = f_0 + \cdots + f_{n-1},$$

or more symmetrically,

$$0 = f_0 + \cdots + f_n.$$

Some work on such generalizations, e.g. [29], assumes the rather strong hypothesis that the f_j are pairwise relatively prime. Other work, e.g. [10], assumes $\gcd(f_0, \dots, f_n) = 1$ and that

$$\sum_{i \in I} f_i \neq 0 \quad \text{for each proper index subset } I \subsetneq \{0, \dots, n\}.$$

This hypothesis is referred to as "no vanishing subsums". The recent work of De Bondt [5] generalizes these two hypotheses to the following: Suppose $0 = f_0 + \cdots + f_n$ and assume that for each index set $I \subseteq \{0, \dots, n\}$,

$$\text{if} \quad \sum_{i \in I} f_i = 0, \quad \text{then} \quad \gcd(\{f_i : i \in I\}) = 1.$$

In the three-function ABC theorem, the right-hand side of the ABC inequality involves the radical $R(f_0 f_1 f_2)$ of the product. But because the functions are pairwise relatively prime, this is the same as the product of the radicals: $R(f_0)R(f_1)R(f_2)$. When one begins with $n+1$ functions that are not necessarily pairwise relatively prime, then the square free part of the product, $S(f_0 \cdots f_n)$, will not in general be the same as the product of the square free parts, $S(f_0) \cdots S(f_n)$. Again, we follow De Bondt's lead by presenting generalized ABC inequalities of both types.

The following two theorems are our generalized ABC theorems for non-Archimedean entire functions of several variables.

THEOREM 6.1 (Generalized ABC theorem (first version)). *Let $f_0, \dots, f_n$ be $n+1 \geq 3$ entire functions in $\mathcal{E}_m$, not all of which are constant and none of which is identically zero. Assume*

$$0 = f_0 + \cdots + f_n \tag{6}$$

and assume that for each index set $I \subseteq \{0, \dots, n\}$,

$$\textit{if} \quad \sum_{i \in I} f_i = 0 \quad \textit{then} \quad \gcd(\{f_i : i \in I\}) = 1. \tag{7}$$

Let $2 \leq d \leq n$ be the dimension of the $\mathbb{F}$-vector space spanned by the f_i. If $\mathbb{F}$ has characteristic zero, let $c = 1$, and if $\mathbb{F}$ has positive characteristic p, let $c = p^{s-1}$, where s is the index of independence for the f_i. Then there exist integers a and b with

$$1 \leq a \leq c(d-1) \quad \textit{and} \quad b \geq a\left\lceil \frac{a}{c} \right\rceil - \frac{\left\lceil \frac{a}{c} \right\rceil \left(\left\lceil \frac{a}{c} \right\rceil - 1 \right)}{2} c \geq a$$

such that for $r \geq 1$,

$$\max_{0\leq j\leq n} \log |f_j|_r \leq \sum_{j=0}^{n} N_{f_j}^{(a)}(0,r) - b\log r + O(1). \tag{8}$$

Moreover, *if* $\operatorname{char}\mathbb{F} = p > 0$ *then there further exists a non-negative integer* σ *with* $p^\sigma \leq a$ *such that for* $r \geq 1$,

$$\max_{0\leq j\leq n} \log |f_j|_r \leq \sum_{j=0}^{n} N_{G_j}(0,r) - b\log r + O(1), \tag{9}$$

where $G_j = \gcd(f_j, R_{p^\sigma}(f_j)^a)$.

THEOREM 6.2 (Generalized ABC theorem (second version)). *Let* $f_0, \ldots$ $\ldots, f_n$ *be* $n+1 \geq 3$ *entire functions in* $\mathcal{E}_m$, *not all of which are constant and none of which is identically zero. Assume*

$$0 = f_0 + \cdots + f_n.$$

Let $2 \leq d \leq n$ *be the dimension of the* $\mathbb{F}$*-vector space spanned by the* f_i. *Let* $2 \leq k \leq n$ *and assume that*

$$\gcd(f_{i_1}, \ldots, f_{i_k}) = 1 \quad \textit{for any } 0 \leq i_1 < \cdots < i_k \leq n. \tag{10}$$

Let $\overline{k} = \min\{k, d\}$. *If* $\overline{k} > 2$, *further assume*

$$\sum_{i\in I} f_i \neq 0 \quad \textit{for each } I \subsetneq \{0,\ldots,n\}. \tag{11}$$

If $\mathbb{F}$ *has characteristic zero*, *let* $c = 1$, *and if* $\mathbb{F}$ *has positive characteristic* p, *let* $c = p^{s-1}$, *where* s *is the index of independence for the* f_i. *Then there exist integers* $\overline{a}$ *and* b *with*

$$1 \leq \overline{a} \leq c\sum_{i=1}^{\overline{k}-1}(d-i) \quad \textit{and} \quad \overline{a} \leq b$$

such that for $r \geq 1$,

$$\max_{0\leq j\leq n} \log |f_j|_r \leq N_F^{(\overline{a})}(0,r) - b\log r + O(1), \tag{12}$$

where $F = f_0 \cdots f_n$.

We remark that given explicit f_i, the constants a, b, σ, and $\overline{a}$ in Theorems 6.1 and 6.2 can be determined explicitly in terms of non-vanishing generalized Wronskians of subsets of the f_i, as will be evident from the proof; see (16), (17), (18), and (20).

Trivial examples in positive characteristic show that the dependence on the index of independence s cannot be removed.

In characteristic zero and one variable, Theorems 6.1 and 6.2 are due to Hu and Yang [20]. Below, we will give their proof, which in turn closely follows Brownawell and Masser [10], and simply observe that it works, given the proper set-up, just as well for several variables and in positive characteristic. That it is natural to express the upper bounds on a and $\overline{a}$ in terms of d is an observation of Zannier [35]. For polynomials of several variables, De Bondt [5] gave an alternative Wronskian-based proof that first reduces to the case of one-variable polynomials by generic specialization but then proves the one-variable case by introducing extra variables to force linear independence, thereby avoiding Lemma 5.1.

Before giving the proof of Theorems 6.1 and 6.2, we discuss how to derive from it various other existing results in the literature. First observe that in characteristic zero, $b \geq a(a+1)/2$ and so (8) implies

$$\max_{0\leq j\leq n} \log|f_j|_r \leq \sum_{j=0}^{n} N_{f_j}^{(a)}(0,r) - \frac{a(a+1)}{2}\log r + O(1),$$

which specializing to complex polynomials gives us

Corollary 6.3 ([5, Th. 2.1(4)]). *Let $f_0, \ldots, f_n$ be polynomials of several complex variables satisfying the hypothesis of the theorem. Then*

$$\max \deg f_j \leq \sum_{j=0}^{n} r_a(f_j) - \frac{a(a+1)}{2},$$

where $r_a(f_j) = \deg\gcd(f_j, R(f_j)^a)$ and a is as in the theorem.

Incorporating an idea of Bayat and Teimoori [4] as in De Bondt [5], we get

Corollary 6.4. *Assume $\mathbb{F}$ has characteristic zero, let $f_0 + \cdots + f_n = 0$ be as in Theorem 6.1 and let d be the dimension of the $\mathbb{F}$-vector space spanned by the f_j. Let C be the number of f_j which are constant functions. For any A with $d \leq A \leq n - C$,*

$$\max_{0\leq j\leq n} \log|f_j|_r \leq A\bigg(\sum_{j=0}^{n} N_{f_j}^{(1)}(0,r) - \frac{A+1}{2}\log r\bigg) + O(1)$$

for $r \geq 1$.

De Bondt gives examples that show that A cannot be made smaller than d in Corollary 6.4.

In the case of polynomials, Corollary 6.4 gives [5, Th. 2.1(5)], which implies [4, Th. 5] as explained in [5]. As remarked by De Bondt, the proof of [4, Th. 5] given by Bayat and Teimoori in [4] is not correct for polynomials of several variables because their Lemma 4 is easily seen to be false for several-variable polynomials. However, arguing as in [5], their Theorem 5 is correct, even for several variables. Of course, this also recovers the result of Shapiro and Sparer [29].

Proof of Corollary 6.4. Clearly (by Proposition 2.7),

$$N_{f_j}^{(a)}(0,r) \le aN_{f_j}^{(1)}(0,r),$$

so from (8), we get

$$\max_{0\le j\le n} \log|f_j|_r \le a\left(\sum_{j=0}^{n} N_{f_j}^{(1)}(0,r) - \frac{a+1}{2}\log r\right) + O(1)$$

for $r \ge 1$. The observation of Bayat and Teimoori is that the expression on the right is quadratic in a and hence increasing in a provided

$$a + \frac{1}{2} \le \frac{1}{\log r}\sum_{j=0}^{n} N_{f_j}^{(1)}(0,r).$$

Because

$$N_{f_j}^{(1)}(0,r) \ge \log r + O(1)$$

for any non-constant f_j and because there are $n+1-C$ non-constant f_j, we have

$$\sum_{j=0}^{n} N_{f_j}^{(1)}(0,r) \ge (n+1-C)\log r + O(1) \ge (n-C)\log r$$

for r sufficiently large, and so we can increase a up to $n-C$ and the inequality will hold for all sufficiently large r. We can then adjust the $O(1)$ term to make the inequality hold for $r \ge 1$. ■

We now digress a little bit to discuss one slightly subtle difference between entire functions and polynomials. The astute reader will notice that we assumed no vanishing subsums, i.e. (11), in Theorem 6.2, whereas De Bondt assumed the weaker hypothesis (7) in both versions of his ABC theorems. In the case of complex polynomials, the inequality in De Bondt's work that corresponds to our inequality (12) is

$$\max \deg f_j \le \deg\gcd(F, R(F)^{\overline{a}}) - b \le \deg\gcd(F, R(F)^{\overline{a}}) - \overline{a},$$

and this holds even if the hypothesis (11) of no vanishing subsums is weakened to (7). The reason for this is that in the case of polynomials, one of the polynomials f_{j_0} has maximal degree. Thus, one need only consider a minimal index set I such that j_0 is contained in I and such that

$$\sum_{j\in I} f_j = 0.$$

However, in the case of entire functions, it need not be the case that there is a fixed index j_0 such that for all r sufficiently large

$$|f_{j_0}|_r = \max_{0\le j\le n} |f_j|_r.$$

Thus, if one replaces the hypothesis (11) in Theorem 6.2 with the weaker hypothesis (7), it follows from (12) and the fact that for a positive integer l, for $r \geq 1$ and F and G entire functions we have

$$\max\{N_F^{(l)}(0,r), N_G^{(l)}(0,r)\} \leq N_{FG}^{(l)}(0,r),$$

that for each index j in $\{0,\ldots,n\}$, there exist integers $\overline{a}_j$ and b_j with

$$1 \leq \overline{a}_j \leq c\sum_{i=1}^{k}(d-i) \quad \text{and} \quad \overline{a}_j \leq b_j$$

such that for $r \geq 1$,

$$\log|f_j|_r \leq N_F^{(\overline{a}_j)}(0,r) - b_j \log r + O(1),$$

from which it follows that

$$\max_{0\leq j\leq n} \log|f_j|_r \leq N_F^{(\max \overline{a}_j)}(0,r) - (\min b_j)\log r + O(1).$$

Of course, it need not be that $\min b_j \geq \max \overline{a}_j$, and thus when subsums of the f_j may vanish, it is not clear for entire functions whether one can choose the same constant at which multiplicities are truncated when counting the zeros of F as the coefficient in front of $-\log r$. Whether that can be done is a somewhat interesting question, because if it can be done, then the proof cannot be a straightforward generalization of the existing polynomial proof. If it cannot be done, then this would be an example where a polynomial inequality does not completely generalize to an analogous inequality for entire functions.

Corollary 6.5. *With hypotheses and notation as in Theorem* 6.2, *we have*, *for* $r \geq 1$ *and any*

$$\overline{A} \geq c\sum_{i=1}^{\overline{k}-1}(d-i),$$

the inequality

$$\max_{0\leq j\leq n} \log|f_j|_r \leq \overline{A}(N_F^{(1)}(0,r) - \log r) + O(1).$$

Moreover, *the above inequality remains valid if the hypothesis* (11) *is weakened to hypothesis* (7).

Remark. The $O(1)$ term may depend on $\overline{A}$.

Proof of Corollary 6.5. Under the hypothesis (11), we have

$$\max_{0\leq j\leq n} \log|f_j|_r \leq N_F^{(\overline{a})}(0,r) - \overline{a}\log r + O(1),$$

which follows immediately from (12) because $\overline{a} \leq b$. Because

$$N_F^{(\overline{a})}(0,r) \leq \overline{a}N_F^{(1)}(0,r)$$

and because

$$N_F^{(1)}(0,r) - \log r$$

is bounded below for $r \geq 1$ since F is non-constant, we have

$$N_F^{(\overline{a})}(0,r) - \overline{a}\log r \leq \overline{a}(N_F^{(1)}(0,r) - \log r) \leq \overline{A}(N_F^{(1)}(0,r) - \log r) + O(1),$$

which gives the corollary when there are no vanishing subsums. However, even if there are vanishing subsums, the functions can be grouped into vanishing subsums

$$\sum_{i\in I} f_i = 0$$

with no vanishing sub-subsums. Any vanishing subsum consisting of all constants can be thrown out. Letting

$$F_I = \prod_{i\in I} f_i,$$

we have

$$\max_{j\in I} \log |f_j|_r \leq \overline{A}(N_{F_I}^{(1)}(0,r) - \log r) + O(1).$$

Because $\overline{A}$ was chosen independent of I and because

$$\max_I N_{F_I}^{(1)}(0,r) \leq N_F^{(1)}(0,r) \quad \text{for } r \geq 1,$$

the corollary follows in general. ∎

If $\mathbb{F}$ has characteristic zero and in the case of polynomials when $k = n$, Corollary 6.5 is [5, Th. 2.2(7)]. When $k = 3$, then we can take $\overline{A} = 2n - 3$ and so we recover the main result of Quang and Tuan in [26]:

$$\max_{0\leq j\leq n} \deg f_j \leq (2n-3)[\deg R(F) - 1] \tag{13}$$

if $\gcd(f_{i_1}, f_{i_2}, f_{i_3}) = 1$ for all triples of indices $i_1 < i_2 < i_3$. Note that Quang and Tuan neglected the necessary hypothesis that the functions in any vanishing subsum be relatively prime, i.e., hypothesis (7). We also remark that Browkin and Brzeziński [9] conjectured that (13) remains true for one-variable polynomials in characteristic zero if the gcd hypothesis is relaxed to $\gcd(f_0, \ldots, f_n) = 1$ and no vanishing subsums. This conjecture seems to be out of reach of the current Wronskian-based proofs.

Fundamental to the proof of Theorems 6.1 and 6.2 are the following lemmas about generalized Wronskians.

LEMMA 6.6. *Let $f_0, \ldots, f_{n-1}$ be entire functions in $\mathcal{E}_m$. Let*

$$\gamma^1 = (\gamma_1^1, \ldots, \gamma_m^1), \quad \ldots, \quad \gamma^{n-1} = (\gamma_1^{n-1}, \ldots, \gamma_m^{n-1})$$

be multi-indices with $|\gamma^1| \leq \cdots \leq |\gamma^{n-1}|$ such that the associated generalized Wronskian W does not vanish identically. Let $\gamma^0 = (0, \ldots, 0)$. If P is an irreducible element which divides f_i with multiplicity $e > |\gamma^{n-1}|$, then P

divides W with multiplicity at least $e-|\gamma^{n-1}|$. Moreover, if $\operatorname{char}\mathbb{F}=p>0$ *and if p^t divides e and*

$$p^t > \max\{\gamma_i^j : 1\le j\le n-1 \text{ and } 1\le i\le m\},$$

then P^e divides W.

Remark. Note that $\max\{\gamma_i^j : 1\le j\le n-1 \text{ and } 1\le i\le m\}\le|\gamma^{n-1}|$.

Proof of Lemma 6.6. This follows immediately from Proposition 3.6. ∎

Lemma 6.7. *Let $f_0,\dots,f_{n-1}$ be entire functions in $\mathcal{E}_m$. Let $\gamma^1,\dots,\gamma^{n-1}$ be multi-indices with $|\gamma^1|\le\cdots\le|\gamma^{n-1}|$ such that the associated generalized Wronskian W does not vanish identically. Let $\gamma^0=(0,\dots,0)$ and let $F=f_0\cdots f_{n-1}$. Let $k\ge 2$ be the smallest integer such that for any k distinct indices $i_1,\dots,i_k$ in $\{0,\dots,n-1\}$ one has* $\gcd(f_{i_1},\dots,f_{i_k})=1$, *or if no such k exists, let $k=n+1$. Let*

$$\ell=\sum_{i=1}^{k-1}|\gamma^{n-i}|.$$

If P is an irreducible element which divides F with multiplicity $e>\ell$, then P divides W with multiplicity at least $e-\ell$.

Remark. Note that in positive characteristic p, even if P divides F with multiplicity a large multiple of p, this does not necessarily mean that P divides W with multiplicity divisible by p. This is because the powers of P may be split among the different f_i, so P need not divide any of the f_i with multiplicity divisible by p.

Proof of Lemma 6.7. By the hypothesis of the lemma, we may assume without loss of generality, by re-ordering the indices if necessary, that P does not divide f_j for $j\ge k-1$. Let $e_i\ge 0$ for $i=0,\dots,k-2$ be the multiplicities with which P divides the f_i, and assume $e_0\ge\cdots\ge e_{k-2}$. Then, by Proposition 3.6, P divides W with multiplicity at least

$$\sum_{i=0}^{k-2}\max\{0,e_i-|\gamma^{n-i-1}|\}\ge\sum_{i=0}^{k-2}(e_i-|\gamma^{n-i-1}|)=e-\ell. \ \blacksquare$$

Proof of Theorems 6.1 and 6.2. If $\mathbb{F}$ has positive characteristic, let G_i be as in the statement of Theorem 6.1. If $\mathbb{F}$ has characteristic zero, let

$$G_i=\gcd(f_i,S(f_i)^a).$$

Then G_i divides $\gcd(f_i,S(f_i)^a)$, and thus

$$N_{G_i}(0,r)\le N_{f_i}^{(a)}(0,r)$$

for $r\ge 1$ by Proposition 2.7 and the definition of truncated counting functions. Therefore, inequality (8) follows from inequality (9). Hence, to prove

Theorem 6.1, it suffices to prove (9), where in characteristic zero we interpret (9) with G_i as defined here in the proof.

We next observe that it suffices to prove Theorem 6.1 assuming there are no vanishing subsums. Indeed, if there are vanishing subsums, simply group the f_i into vanishing subsums with no vanishing sub-subsums, and note that these subsums still satisfy all the hypotheses of the theorem. Summing (9) over each minimal vanishing subsum clearly results in (9) for the general case.

Note also that if $\overline{k} = 2$ in Theorem 6.2, then the f_j are pairwise relatively prime, in which case

$$N_F^{(\ell)}(0,r) = \sum_{j=0}^{n} N_{f_j}^{(\ell)}(0,r),$$

and so Theorem 6.2 follows from Theorem 6.1. Thus, we henceforth assume (11) as we prove both theorems.

Consider the $\mathbb{F}$-linear span of $f_0, \ldots, f_n$ as an $\mathbb{F}$-vector space, and partition

$$\{0, \ldots, n\} = I_0 \cup \cdots \cup I_{u-1} \quad \text{with } J_j \in I_j \text{ for } j = 0, \ldots, u-2$$

as in Lemma 5.1. Let $J_{-1} = \emptyset$. Also, without loss of generality, assume $0 \in I_0$. Let n_j be the cardinality of I_j for $j = 0, \ldots, u-1$. Note that $n_0 \leq d+1$ by the minimality of I_0 and that $n_j \leq d$ for $j \geq 1$ by the minimality of $I_j \cup J_{j-1}$. Set

$$\gamma^{0,0} = \gamma^{1,0} = \cdots = \gamma^{u-1,0} = \gamma^0 = \gamma^1 = \cdots = \gamma^{u-1} = (0, \ldots, 0).$$

The f_i for i in $I_0 \setminus \{0\}$ are linearly independent over $\mathbb{F}$. Therefore by Lemma 5.2 and Theorem 3.5, there exist multi-indices $\gamma^{0,1}, \ldots, \gamma^{0,n_0-2}$ such that the generalized Wronskian W_0 formed by the f_i with respect to these multi-indices for i in $I_0 \setminus \{0\}$ is not identically zero, and moreover,

$$|\gamma^{0,i}| \leq |\gamma^{0,i-1}| + c \leq ci.$$

Note that if $n_0 = 2$, we simply let W_0 be the f_i for the unique index i in I_0 different from 0. Similarly, for $j = 1, \ldots, u-1$, there exist multi-indices $\gamma^{j,1}, \ldots, \gamma^{j,n_j-1}$ such that the generalized Wronskian W_j formed by the f_i with respect to these multi-indices for i in I_j is not identically zero, and

$$|\gamma^{j,i}| \leq |\gamma^{j,i-1}| + c \leq ci.$$

The total number of multi-indices we get this way is

$$n_0 - 2 + n_1 - 1 + \cdots + n_{u-1} - 1 = n - u.$$

Write these multi-indices as $\gamma^u, \ldots, \gamma^{n-1}$ with

$$|\gamma^u| \leq \cdots \leq |\gamma^{n-1}| \leq c(d-1),$$

and note that

$$|\gamma^i| \leq |\gamma^{i-1}| + c. \tag{14}$$

For each minimal index set $I_0, J_0 \cup I_1, \ldots, J_{u-2} \cup I_{u-1}$, there is a linear dependence relation

$$\sum_{i=0}^{n} c_{j,i} f_i = 0$$

with $c_{j,i}$ non-zero elements of $\mathbb{F}$ when i is in $J_{j-1} \cup I_j$ and 0 otherwise. Of course, this also gives rise to the linear equations

$$\sum_{i=0}^{n} c_{0,i} D^{\gamma^{0,q}} f_i = 0, \qquad q = 0, \ldots, n_0 - 2,$$

and

$$\sum_{i=0}^{n} c_{j,i} D^{\gamma^{j,q}} f_i = 0, \qquad q = 0, \ldots, n_j - 1,\ j = 1, \ldots, u-1.$$

Let M be the $n \times (n+1)$-matrix whose entries are $c_{j,i} D^{\gamma^{j,q}} f_i$, where the columns are indexed by i and the rows are indexed by j and q. Note that the sum of each row of M is zero. Let Δ_i denote the determinant of the matrix M with the ith column deleted. Because the ith column is the negative of the sum of the other columns, $\Delta_i = \pm \Delta_j$. From the block nature of M,

$$\Delta_0 = C_0 W_0 \cdots W_{u-1}, \tag{15}$$

where C_0 is a constant obtained by multiplying the appropriate $c_{j,i}$'s, and hence is non-zero. Thus, Δ_i is non-zero for all i, and up to a constant is the product of the generalized Wronskians W_j.

Define

$$a = |\gamma^{n-1}| \leq c(d-1), \tag{16}$$

$$\begin{aligned} b = \sum_{i=u}^{n-1} |\gamma^i| &\geq \sum_{i=u}^{n-2} \max\{1, |\gamma^{n-1}| - ci\} \\ &\geq a \left\lceil \frac{a}{c} \right\rceil - \frac{\left\lceil \frac{a}{c} \right\rceil \left(\left\lceil \frac{a}{c} \right\rceil - 1 \right)}{2} c \geq a \end{aligned} \tag{17}$$

and

$$\bar{a} = \sum_{i=1}^{\bar{k}-1} |\gamma^{n-i}| \leq c \sum_{i=1}^{\bar{k}-1} (d-i), \tag{18}$$

where the first inequality in (17) follows from (14). Note also that

$$b \geq \bar{a} \geq a \geq 1, \tag{19}$$

where $a \geq 1$ follows from the fact that not all the Wronskians W_j can be 1×1, for if they were, we would have either all the f_i constant or $\gcd(f_0,\dots,f_n) \neq 1$.

If $\mathbb{F}$ has characteristic $p > 0$, let σ be the largest integer such that

$$p^\sigma \leq \max\{\gamma_j^i : u \leq i \leq n-1 \text{ and } 1 \leq j \leq m\} \leq a, \tag{20}$$

where $\gamma^i = (\gamma_1^i,\dots,\gamma_m^i)$.

We claim that F divides

$$\Delta_0 \prod_{i=0}^{n} G_i.$$

Indeed, suppose that P is an irreducible element of $\mathcal{E}_m$ which divides f_i with exact multiplicity e. If $\mathbb{F}$ has characteristic zero, then $P^{\min\{e,a\}}$ divides G_i by Proposition 2.5, and if $e > a$, then P^{e-a} divides W_j, where $i \in I_j$, by Lemma 6.6 and (16). Now suppose $\operatorname{char}\mathbb{F} = p$ and p^v is the largest power of p dividing e. If $p^v \leq p^\sigma \leq a$, then $P^{\min\{e,a\}}$ divides G_i by Propositon 4.1 and in the case $e > a$, we find that P^{e-a} divides W_j, where $i \in I_j$, again by Lemma 6.6 and (16). If $p^v > p^\sigma$, then P^e divides Δ_0 by Lemma 6.6 and (20). Thus in all cases, P^e divides $W_j G_i$, which divides $\Delta_0 G_i$ by (15). Therefore,

$$\log|F|_r \leq \log|\Delta_0|_r + \sum_{i=0}^{n} \log|G_i|_r + O(1) \tag{21}$$

by Corollary 2.2.

Let

$$F_0 = \prod_{i\in I_0\setminus\{0\}} f_i, \qquad F_j = \prod_{i\in I_j} f_i, \quad j = 1,\dots,u-1.$$

For each $j = 0,\dots,u-1$, the quotient W_j/F_j is the determinant of a matrix consisting of logarithmic derivatives, and so by Lemma 3.2,

$$\left|\frac{W_0}{F_0}\right|_r \leq -\Big(\sum_{q=0}^{n_0-2} |\gamma^{0,q}|\Big)\log r,$$

$$\left|\frac{W_j}{F_j}\right|_r \leq -\Big(\sum_{q=0}^{n_j-1} |\gamma^{0,q}|\Big)\log r \qquad \text{for } j = 1,\dots,u-1.$$

Then

$$\frac{f_0\Delta_0}{F} = C_0 \cdot \frac{W_0}{F_0}\cdots\frac{W_{u-1}}{F_{u-1}}$$

and hence

$$\log|f_0|_r + \log|\Delta_0|_r - \log|F|_r \leq -b\log r + O(1)$$

by Lemma 3.2 and (17). Similarly, if i is any index in I_0, then we can write W_0 as a determinant involving $D^{\gamma^{0,q}} f_l$ for $l \neq i$, and so $f_i W_0/f_0 F_0$ is also a sum of products of logarithmic derivatives, and hence

$$\log\left|\frac{f_iW_0}{f_0F_0}\right|_r \le -\Big(\sum_{q=0}^{n_0-2}|\gamma^{0,q}|\Big)\log r$$

as well. Thus

$$\log|f_i|_r + \log|\Delta_0|_r - \log|F|_r \le -b\log r + O(1)$$

for all i in I_0. Now let i be in I_1 and let j be in J_0. Then, by a similar argument,

$$\frac{f_i\Delta_0}{F} = C_0\cdot\frac{f_jW_0}{f_0F_0}\cdot\frac{f_iW_1}{f_jF_1}\cdot\frac{W_2}{F_2}\cdots\frac{W_{u-1}}{F_{u-1}},$$

and f_jW_0/f_0F_0 and f_iW_1/f_jF_1 are both sums of products of logarithmic derivatives. Hence

$$\log|f_i|_r + \log|\Delta_0|_r - \log|F|_r \le -b\log r + O(1).$$

Continuing, we find that for all i,

$$\log|f_i|_r + \log|\Delta_0|_r - \log|F|_r \le -b\log r + O(1),$$

whence

$$\max_{0\le i\le n}\log|f_i|_r \le \log|F|_r - \log|\Delta_0|_r - b\log r + O(1). \tag{22}$$

Combining this with (21), we get

$$\max_{0\le i\le n}\log|f_i|_r \le \sum_{i=0}^{n}\log|G_i|_r - b\log r + O(1).$$

If we use the Poisson–Jensen–Green type formula (2) and the definition of counting functions, this can also be written, for $r\ge 1$, as

$$\max_{0\le i\le n}\log|f_i|_r \le \sum_{i=0}^{n}N_{G_i}(0,r) - b\log r + O(1),$$

which is precisely (9).

We now show (12). Let P be an irreducible element of $\mathcal{E}_m$ that divides F. By the hypotheses of the theorem, there is at least one f_i such that P does not divide f_i. Because, as above, we can write Δ_0 as a product of Wronskians not involving f_i, we can use Lemma 6.7 to conclude by (18) that P divides Δ_0 with multiplicity at least $e-\overline{a}$. Thus, $F/\gcd(F,\Delta_0)$ divides $\gcd(F,S(F)^{\overline{a}})$. Hence, by Proposition 2.2, the Poisson–Jensen–Green type formula (2), and the definition of truncated counting functions, for $r\ge 1$,

$$\log|F|_r - \log|\Delta_0|_r \le |\gcd(F,S(F)^{\overline{a}})|_r + O(1) = N_F^{(\overline{a})}(0,r) + O(1).$$

Combining this with (22), we get for $r\ge 1$

$$\max_{0\le i\le n}\log|f_i|_r \le N_F^{(\overline{a})}(0,r) - b\log r + O(1),$$

which is (12). ■

References

[1] V. H. An, *Height of p-adic holomorphic maps in several variables and applications*, Acta Math. Vietnam. 27 (2002), 257–269.

[2] V. H. An and D. Q. Manh, *p-Adic Nevanlinna–Cartan theorem in several variables for Fermat type hypersurfaces*, East-West J. Math. 4 (2002), 87–99.

[3] —, —, *The "ABC" conjecture for p-adic entire functions of several variables*, Southeast Asian Bull. Math. 27 (2004), 959–972.

[4] M. Bayat and H. Teimoori, *A new bound for an extension of Mason's theorem for functions of several variables*, Arch. Math. (Basel) 82 (2004), 230–239.

[5] M. de Bondt, *Another generalization of Mason's ABC-theorem*, arXiv: 0707.0434v1.

[6] S. Bosch, U. Güntzer and R. Remmert, *Non-Archimedean Analysis*, Grundlehren Math. Wiss. 261, Springer, Berlin, 1984.

[7] A. Boutabaa and A. Escassut, *Nevanlinna theory in characteristic p and applications*, in: Analysis and Applications—ISAAC 2001 (Berlin), H. G. W. Begehr *et al.* (eds.), Int. Soc. Anal. Appl. Comput. 10, Kluwer, Dordrecht, 2003, 97–107.

[8] —, —, *Nevanlinna theory in positive characteristic and applications*, Ital. J. Pure Appl. Math. 23 (2008), 45–66.

[9] J. Browkin and J. Brzeziński, *Some remarks on the abc-conjecture*, Math. Comp. 62 (1994), 931–939.

[10] W. D. Brownawell and D. Masser, *Vanishing sums in function fields*, Math. Proc. Cambridge Philos. Soc. 100 (1986), 427–434.

[11] W. Cherry, *Existence of GCD's in rings of non-Archimedean entire functions*, preprint.

[12] W. Cherry and Z. Ye, *Non-Archimedean Nevanlinna theory in several variables and the non-Archimedean Nevanlinna inverse problem*, Trans. Amer. Math. Soc. 349 (1997), 5043–5071.

[13] M. van Frankenhuysen, *Hyperbolic spaces and the abc conjecture*, Ph.D. thesis, Katholieke Universiteit Nijmegen, 1995.

[14] —, *The ABC theorem for meromorphic functions*, arXiv: 0805.1729v1.

[15] H. Fujimoto, *Nonintegrated defect relation for meromorphic maps of complete Kähler manifolds into $P^{N_1}(C)\times\cdots\times P^{N_k}(C)$*, Japan. J. Math. (N.S.) 11 (1985), 233–264.

[16] D. Goldfeld, *Modular forms, elliptic curves and the ABC-conjecture*, in: A Panorama of Number Theory or the View from Baker's Garden (Zürich, 1999), G. Wüstholz (ed.), Cambridge Univ. Press, Cambridge, 2002, 128–147.

[17] Hà Huy Khoái, *La hauteur des fonctions holomorphes p-adiques de plusieurs variables*, C. R. Acad. Sci. Paris Sér. I Math. 312 (1991), 751–754.

[18] L.-C. Hsia and J. T.-Y. Wang, *The ABC theorem for higher-dimensional function fields*, Trans. Amer. Math. Soc. 356 (2004), 2871–2887.

[19] P.-C. Hu and C.-C. Yang, *The "abc" conjecture over function fields*, Proc. Japan Acad. Ser. A Math. Sci. 76 (2000), 118–120.

[20] —, —, *A note on Browkin–Brzeziński conjecture*, in: Ultrametric Functional Analysis, B. Diarra *et al.* (eds.), Contemp. Math. 384, Amer. Math. Soc., Providence, RI, 2005, 101–109.

[21] S. Lang, *Old and new conjectured Diophantine inequalities*, Bull. Amer. Math. Soc. (N.S.) 23 (1990), 37–75.

[22] R. C. Mason, *Diophantine Equations over Function Fields*, London Math. Soc. Lecture Note Ser. 96, Cambridge Univ. Press, Cambridge, 1984.

[23] —, *Norm form equations. I*, J. Number Theory 22 (1986), 190–207.

[24] J. Noguchi and T. Ochiai, *Geometric Function Theory in Several Complex Variables*, Transl. Math. Monogr. 80, Amer. Math. Soc., Providence, RI, 1990.

[25] J. Oesterlé, *Nouvelles approches du "théorème" de Fermat*, Séminaire Bourbaki 1987/88. Astérisque 161–162 (1988), Exp. No. 694, 4, 165–186.

[26] N. T. Quang and P. D. Tuan, *A note on Browkin–Brzeziński's conjecture*, Int. J. Contemp. Math. Sci. 2 (2007), 1335–1340.

[27] K. F. Roth, *Rational approximations to algebraic numbers*, Mathematika 2 (1955), 1–20.

[28] B. V. Shabat, *Distribution of Values of Holomorphic Mappings*, Transl. Math. Monogr. 61, Amer. Math. Soc., Providence, RI, 1985.

[29] H. N. Shapiro and G. H. Sparer, *Extension of a theorem of Mason*, Comm. Pure Appl. Math. 47 (1994), 711–718.

[30] C. L. Stewart and K. Yu, *On the abc conjecture*, *II*, Duke Math. J. 108 (2001), 169–181.

[31] W. W. Stothers, *Polynomial identities and Hauptmoduln*, Quart. J. Math. Oxford Ser. (2) 32 (1981), 349–370.

[32] L. N. Vaserstein, *Quantum* (abc)*-theorems*, J. Number Theory 81 (2000), 351–358.

[33] P. Vojta, *Diophantine Approximations and Value Distribution Theory*, Lecture Notes in Math. 1239, Springer, Berlin, 1987.

[34] J. F. Voloch, *Diagonal equations over function fields*, Bol. Soc. Brasil. Mat. 16 (1985), no. 2, 29–39.

[35] U. Zannier, *Some remarks on the S-unit equation in function fields*, Acta Arith. 64 (1993), 87–98.

Department of Mathematics
University of North Texas
1155 Union Circle #311430
Denton, TX 76203-5017, U.S.A.
E-mail: wcherry@unt.edu

Department of Mathematics and Statistics
MSC03 2150
1 University of New Mexico
Albuquerque, NM 87131-0001, U.S.A.
E-mail: ctoropu@hotmail.com

Received on 23.7.2008
and in revised form on 29.8.2008 (5768)