ROCKY MOUNTAIN
JOURNAL OF MATHEMATICS
Volume 26, Number 3, Summer 1996

AN EFFECTIVE ROTH'S THEOREM FOR FUNCTION FIELDS

JULIE TZU-YUEH WANG

ABSTRACT. We will give a new proof of Roth's theorem for function fields which is motivated by Steinmetz's proof of Nevanlinna's second main theorem of slowly moving target functions. This method provides effective results.

0. Introduction. The correspondence between number theory and value distribution theory has been observed by Osgood [**3**] and Vojta [**6**]. Both these areas are related to function fields. For example, one can establish the analogue of Cartan's truncated second main theorem for function fields [**7**], which corresponds to the so-called abc conjecture for number fields. Usually for a corresponding result in function fields, one can also expect two proofs, one analogous to number theory and the other analogous to value distribution theory. For example, the Thue-Siegel-Roth theorem for function fields was proved by Uchiyama [**5**] with a line of proof similar to the one for number fields, and hence is ineffective. However, Roth's theorem for function fields should also correspond to some sort of second main theorem with moving target functions in value distribution theory (such as in Nevanlinna's conjecture with slowly moving target functions) which was proved by Steinmetz [**4**] in the case of functions. Indeed, since the ideas in [**4**] mainly involve Wronskians, one can expect an analogous proof for function fields. In this paper we will give a proof of Roth's theorem for function fields which is analogous to [**4**]. As we move on to the proof, it will become clear that the results are effective in the sense that the constants in the proof can be effectively determined from the method of the proof.

Let K be the function field of a smooth projective curve C over an algebraically closed field k of characteristic 0. For each $P \in C$ we have a valuation v_P on K. Each v_P can be extended to K^a, where K^a is the algebraic closure of K. For each $f \in K$ one can define the height $h_K(f) = \sum_{P\in C} -\min\{0, v_P(f)\} = \sum_{P\in C} \max\{0, v_P(f)\}$. We

Received by the editors on August 17, 1994, and in revised form on May 1, 1995.

will drop the index K when there is no confusion. Roth's theorem for function fields can be formulated as follows.

Roth's theorem. *Let S be a nonempty set of finitely many points in C. Assume that, for each $P \in S$, we have an element $a_P \in K^a$, and that v_P is extended to K^a in some way. Suppose that κ is a real number greater than 2. Then the elements $f \in K$ satisfying the approximation*

$$\sum_{P \in S} \max\{0, v_P(a_P - f)\} \geq \kappa h_K(f) \tag{1}$$

have bounded height.

Remark 1. If we enlarge the size of S, we get a stronger statement. So in the proof we can expand S as necessary.

Remark 2. Let L be a finite normal extension of K containing every a_P for $P \in S$, and let R be the smooth projective model of L. Then we have a surjective morphism $\pi : R \to C$. For every $Q \in \pi^{-1}(P)$, we can find an a_Q which is conjugate to a_P satisfying

$$v_P(a_P - f) = \frac{1}{[L:K]} \sum_{Q \in \pi^{-1}(P)} v_Q(a_Q - f).$$

Take $S' = \{Q : Q \in \pi^{-1}(S)\}$. Then since $h_K(f) = (1/[L:K])h_L(f)$ for all $f \in K$, (1) is equivalent to $\sum_{Q \in S'} \max\{0, v_Q(a_Q - f)\} \geq \kappa h_L(f)$. Therefore, without loss of generality, we can assume $a_P \in K$ for all $P \in S$.

1. The main theorem. Let $a_1, \ldots, a_q$ be distinct elements in K. Let $L(r)$ be the vector space over k spanned by $a_1^{n_1} \ldots a_q^{n_q}$ with $n_1, \ldots, n_q \geq 0$ and $n_1 + \cdots + n_q = r$. Let $\beta_1, \ldots, \beta_n$ be a base of $L(r)$, which can be assumed to consist of monomials. Let $b_1, \ldots, b_m$ be a base of $L(r+1)$ which can also be assumed to consist of monomials.

Main theorem. *Suppose that t, $a_1, \ldots, a_q$ are S-units and that f is a nonzero element of K. If $f\beta_1, \ldots, f\beta_n, b_1, \ldots, b_m$ are linearly* *independent over k, then*

$$\sum_{i=1}^{q}\sum_{P\in S}\max\{0, v_P(f-a_i)\} \le \frac{m+n}{n}h(f) + \frac{(m+n-1)(m+n)}{n}(2g-2+2|S|+h(t)) + (q-1)^2\sum_{i=1}^{q}h(a_i).$$

Let t be a nonconstant function in K. Recall that the Wronskian matrix of $g_0,\dots,g_l\in K$ with respect to t is the matrix whose i,j-th entry is $(d^i/(dt)^i)(g_j)$, for $0\le i,\ j\le l$. We will denote the Wronskian matrix of $g_0,\dots,g_l\in K$ with respect to t as $W(g_0,\dots,g_l)$ and the Wronskian matrix of $g_0,\dots,g_l\in K$ with respect to t_P as $W_{t_P}(g_0,\dots,g_l)$, where t_P is a local parameter of a point P in C.

Definition. (1) $B[f]:=W(b_1,\dots,b_m,f\beta_1,\dots,f\beta_n)$
(2) $B_{t_P}[f]:=W_{t_P}(b_1,\dots,b_m,f\beta_1,\dots,f\beta_n)$.

Proposition 1. $\det B[f-a_i]=\det B[f]$.

Proof. This is because $a_i\beta_j\in L(r+1)$. □

Proof of the main theorem. Since $f\beta_1,\dots,f\beta_n,\ b_1,\dots,b_m$ are linearly independent over k, we have $\det B[f]\ne 0$. Let Q (respectively Q_j) be the matrix obtained by multiplying the $(m+1)$-st through $(m+n)$-th columns of $B[f]$, respectively $B[f-a_j]$, by f^{-1}, respectively $(f-a_j)^{-1}$. Then

$$\det B[f-a_j]=(f-a_j)^n\det Q_j.$$

Let $B^*[f]$, respectively Q^*, Q_j^*, be the determinant of the matrix obtained by multiplying the i-th $(1\le i\le m)$ column of $B[f]$ by b_i^{-1} and multiplying the l-th $(m+1\le l\le m+n)$ column of $B[f]$, respectively Q, Q_j, by β_l^{-1}, that is, $B^*[f]$, respectively Q^*, Q_j^*, is the logarithmic

Wronskian determinant of $B[f]$, respectively Q, Q_j. Together with Proposition 1, we have

$$(f-a_j)^n = \frac{\det B[f]}{\det Q_j} = \frac{B^*[f]}{Q_j^*}.$$

Therefore $v_P(f-a_i) = (1/n)v_P(B^*[f]) - (1/n)v_P(Q_j^*)$. We have

$$\begin{aligned}\max\{0, v_P(f-a_j)\} \le \frac{1}{n}&\max\{0, v_P(B^*[f])\} \\ &- \frac{1}{n}\min\{0, v_P(Q_j^*)\}.\end{aligned} \tag{2}$$

Since $a_i - a_j = (f-a_j) - (f-a_i)$, we have

$$v_P(a_i - a_j) \ge \min\{v_P(f-a_i), v_P(f-a_j)\}. \tag{3}$$

Let

$$\alpha_P = \max_{i \ne j}\{0, v_P(a_i - a_j)\}. \tag{4}$$

Then

$$\begin{aligned}\sum_{P\in S}\sum_{j=1}^{q}\max\{0, v_P(f-a_j)\} \le \frac{1}{n}&\sum_{P\in S}\max\{0, v_P(B^*[f])\} \\ &+ \frac{1}{n}\sum_{P\in S}\max_{1\le j\le q}\{-\min\{0, v_P(Q_j^*)\}\} \\ &+ (q-1)\sum_{P\in S}\alpha_P.\end{aligned} \tag{5}$$

The theorem then follows from the following lemmas:

Lemma 1. *If t is an S-integer, then $\sum_{P\in C}\max\{0, v_P(dt/dt_P)\} \le 2g - 2 + |S| + h(t)$.*

Lemma 2. *$h(B^*[f]) \le (m+n)h(f) + ((m+n-1)(m+n)/2)(2g - 2 + 2|S| + h(t))$.*

Lemma 3. $\sum_{P\in S}\max_{1\le j\le q}\{-\min\{0,v_P(Q_j^*)\}\}\le((m+n-1)(m+n)/2)(2g-2+2|S|+h(t))$.

Lemma 4. $\sum_{P\in S}\alpha_P\le(q-1)\sum_{i=1}^q h(a_i)$.

Proof of Lemma 1. If $v_P(t)<0$, then $v_P(dt/dt_P)=v_P(t)-1$. On the other hand, if $v_P(t)\ge 0$, then $v_P(dt/dt_P)\ge 0$. By the Riemann-Roch theorem,

$$\begin{aligned}2g-2&=\sum_{v_P(t)<0}v_P\left(\frac{dt}{dt_P}\right)+\sum_{v_P(t)\ge 0}v_P\left(\frac{dt}{dt_P}\right)\\&=\sum_{v_P(t)<0}(v_P(t)-1)+\sum_{P\in C}\max\left\{0,v_P\left(\frac{dt}{dt_P}\right)\right\}\\&\ge -h(t)-|S|+\sum_{P\in C}\max\left\{0,v_P\left(\frac{dt}{dt_P}\right)\right\}.\end{aligned}$$

Therefore, $\sum_{P\in C}\max\{0,v_P(dt/dt_P)\}\le 2g-2+|S|+h(t)$. □

Proof of Lemma 2. Let $B_{t_P}^*[f]$, respectively $Q_{t_P}^*$), be obtained in the same way as $B^*[f]$, respectively $Q_{t_P}^*$, but by taking derivatives with respect to t_P instead. By the basic properties of Wronskian determinants, we have

$$B^*[f]=B_{t_P}^*[f]\left(\frac{dt}{dt_P}\right)^{-(m+n-1)(m+n)/2}.$$

Therefore,

$$\begin{aligned}\min\{0,v_P(B^*[f])\}&\ge\min\{0,v_P(B_{t_P}^*[f])\}\\&\qquad-\frac{(m+n-1)(m+n)}{2}\max\left\{0,v_P\left(\frac{dt}{dt_P}\right)\right\}.\end{aligned}\tag{6}$$

For $P\in S$, we use the relation $B_{t_P}^*[f]=f^nQ_{t_P}^*$. Since $Q_{t_P}^*$ is a logarithmic Wronskian determinant, the order of its pole at P does not exceed $(m+n-1)(m+n)/2$. Therefore,

$$\min\{0,v_P(B_{t_P}^*[f])\}\ge n\min\{0,v_P(f)\}-\frac{(m+n-1)(m+n)}{2}.\tag{7}$$

If $P \notin S$ and $v_P(f) \geq 0$, then because $b_1, \dots, b_m, \beta_1, \dots, \beta_n$ are S-units

$$v_P(B^*_{t_P}[f]) \geq 0. \tag{8}$$

If $P \notin S$ and $v_P(f) < 0$, then

$$\begin{aligned}\det B_{t_P}[f] &= \det W_{t_P}(b_1, \dots, b_m, f\beta_1, \dots, f\beta_n)\\ &= f^{m+n} \det W_{t_P}\left(\frac{b_1}{f}, \cdots, \frac{b_m}{f}, \beta_1, \dots, \beta_n\right).\end{aligned}$$

Then, since $b_1, \dots, b_m, \beta_1, \dots, \beta_n$ are S-units,

$$\begin{aligned}\min\{0, v_P(B^*_{t_P}[f])\} &= \min\{0, v_P(\det B_{t_P}[f])\}\\ &\geq (m+n)\min\{0, v_P(f)\}.\end{aligned} \tag{9}$$

By (6), (7), (8), (9) and Lemma 1, we have

$$\begin{aligned}h(B^*[f]) &\leq (m+n)h(f)\\ &\quad + \frac{(m+n-1)(m+n)}{2}(2g-2+2|S|+h(t)). \qquad \square\end{aligned}$$

Proof of Lemma 3. Let Q^*_{j,t_P} be obtained in the same way as Q^*_j but by taking derivatives with respect to t_P instead. We have

$$-\min\{0, v_P(Q^*_{j,t_P})\} \leq \frac{(m+n-1)(m+n)}{2}. \tag{10}$$

Again, by the basic properties of Wronskian determinants,

$$Q^*_j = Q^*_{j,t_P}(dt/dt_P)^{-(m+n-1)(m+n)/2}.$$

Therefore

$$\begin{aligned}&\sum_{P\in S} \max_{1\leq j\leq q}\{-\min\{0, v_P(Q^*_j)\}\}\\ &\qquad \leq \frac{(m+n-1)(m+n)}{2}\left(\sum_{P\in S}\max\left\{0, v_P\left(\frac{dt}{dt_P}\right)\right\} + |S|\right).\end{aligned}$$

By Lemma 1, we have

$$\sum_{P\in S}\max_{1\le j\le q}\{-\min\{0,v_P(Q_j^*)\}\\ \le \frac{(m+n-1)(m+n)}{2}(2g-2+2|S|+h(t)). \qquad \square$$

Proof of Lemma 4.

$$\begin{aligned}\sum_{P\in S}\alpha_P &= \sum_{P\in S}\max_{i\neq j}\{0,v_P(a_i-a_j)\}\\ &\le \sum_{P\in S}\sum_{i\neq j}\max\{0,v_P(a_i-a_j)\}\\ &\le \sum_{i\neq j}h(a_i-a_j)\\ &\le \sum_{i\neq j}(h(a_i)+h(a_j))\\ &\le (q-1)\sum_{i=1}^{q}h(a_i). \qquad \square\end{aligned}$$

Therefore we have completed the proof of the main theorem.

2. The proof of Roth's theorem. First we need to deal with the case when $\det B[f]=0$. We need the following properties:

Proposition 2. *If $g\in L(s)$, then $h(g)\le s\sum_{i=1}^{q}h(a_i)$.*

Proof. Since $g\in L(s)$, g is a linear combination of $a_1^{n_1}\dots a_q^{n_q}$ with $n_1,\dots,n_q\ge 0$ and $n_1+\dots+n_q=s$. Therefore,

$$v_P(g)\ge s\min\{v_P(a_1),\dots,v_P(a_q)\}$$

$$\min\{0,v_P(g)\}\ge s\sum_{i=1}^{q}\min\{0,v_P(a_i)\}.$$

Therefore, $h(g)\le s\sum_{i=1}^{q}h(a_i)$. $\square$

Lemma 5. *Let* $\beta_1, \dots, \beta_n \in L(r)$ *and* $b_1, \dots, b_m \in L(r+1)$. *If* $f\beta_1, \dots, f\beta_n, b_1, \dots, b_m$ *are linearly dependent over* k, *then* $h(f) \leq (2r+1)\sum_{i=1}^{q} h(a_i)$.

Proof. Since $f\beta_1, \dots, f\beta_n, b_1, \dots, b_m$ are linearly dependent over k,

$$f = \frac{\sum_{i=1}^{m} c_i b_i}{\sum_{j=1}^{m} d_j \beta_j}$$

for some $c_i, d_j \in k$. Since $\sum_{i=1}^{m} c_i b_i \in L(r+1)$ and $\sum_{j=1}^{m} d_j \beta_j \in L(r)$,

$$h(f) \leq h\bigg(\sum_{i=1}^{m} c_i b_i\bigg) + h\bigg(\sum_{j=1}^{m} d_j \beta_j\bigg) \leq (2r+1)\sum_{i=1}^{q} h(a_i)$$

by Proposition 2. $\square$

Let $l(r)$ be the dimension of $L(r)$. Since $l(r+1) \geq l(r)$, $\inf_r l(r+1)/l(r)$ is well defined.

Lemma 6. $\inf_r l(r+1)/l(r) = 1$.

Proof. Suppose this is not true. Then there exists $\delta > 0$ such that $l(r+1) \geq (1+\delta)l(r)$ for $r \geq 0$. Therefore

$$l(r) \geq (1+\delta)^r. \tag{11}$$

On the other hand, from the construction of $L(r)$

$$l(r) \leq \binom{q+r-1}{r} = O(r^{q-1}). \tag{12}$$

This contradicts (11). $\square$

Proof of Roth's theorem. It suffices to show the following:

$(*)$ *Let* $a_1, \dots, a_q$ *be distinct elements in* K, *and let* t *be a non-constant element of* K. *Let* S *be a finite set of points of* C *such that*

$t, a_1, \ldots, a_q$ are S-units. Suppose $\kappa > 2$. Then the elements $f \in K$ satisfying the approximation

$$\sum_{P \in S} \sum_{i=1}^{q} \max\{0, v_P(f - a_i)\} \geq \kappa h(f) \tag{13}$$

have bounded height.

Suppose for $\kappa > 2$ that

$$\sum_{P \in S} \sum_{i=1}^{q} \max\{0, v_P(f - a_i)\} \geq \kappa h(f).$$

Given an ε such that $\kappa - 2 - \varepsilon > 0$, we can find an integer r such that

$$\frac{l(r+1)}{l(r)} \leq 1 + \varepsilon. \tag{14}$$

Suppose $m = l(r+1)$ and $n = l(r)$. Then

$$\frac{m}{n} \leq 1 + \varepsilon. \tag{15}$$

Let $\beta_1, \ldots, \beta_n$ be a base of $L(r)$ and $b_1, \ldots, b_m$ be a base of $L(r+1)$. If $f\beta_1, \ldots, f\beta_n$, $b_1, \ldots, b_m$ are linearly dependent over k, then by Lemma 5

$$h(f) \leq (2r+1) \sum_{i=1}^{q} h(a_i).$$

Therefore, we only have to consider the case when $f\beta_1, \ldots, f\beta_n, b_1, \ldots,$ b_m are linearly independent over k. By the main theorem and (13)

$$\left(\kappa - \frac{m+n}{n}\right) h(f) \leq \frac{(m+n-1)(m+n)}{n}(2g - 2 + 2|S| + h(t)) + (q-1)^2 \sum_{i=1}^{q} h(a_i).$$

Since $m/n \leq 1+\varepsilon$, we have $\kappa - (m+n)/n \geq \kappa - 2 - \varepsilon$. Hence,

$$\begin{aligned}(\kappa - 2 - \varepsilon)h(f) &\leq (2+\varepsilon)(m+n-1)2(2g-2+2|S|+h(t)) \\ &\quad + (q-1)^2 \sum_{i=1}^{q} h(a_i) \\ &\leq 2(2+\varepsilon)\binom{q+r}{r+1}(2g-2+2|S|+h(t)) \\ &\quad + (q-1)^2 \sum_{i=1}^{q} h(a_i).\end{aligned}$$

Therefore $h(f)$ is bounded. This completes the proof of Roth's theorem. $\square$

REFERENCES

1. S. Lang, *Fundamentals of diophantine geometry*, Springer Verlag, New York, 1983.

2. ———, *Introduction to complex Hyperbolic spaces*, Springer Verlag, New York, 1987.

3. C.F. Osgood, *Sometimes effective Thue-Siegel-Roth-Schmidt-Nevanlinna bounds, or better*, J. Number Theory **21** (1985), 347–389.

4. N. Steinmetz, *Eine Verallgemeinerung des zweiten Nevanlinnaschen Hauptsatzes*, J. für Mathematik **368** (1985), 134–141.

5. S. Uchiyama, *Rational approximation to algebraic functions*, J. Fac. Sci. Hokkaido Univ. **15** (1961), 173–192.

6. P. Vojta, *Diophantine approximations and value distribution theory*, Springer Verlag, New York, 1987.

7. J.T.-Y. Wang, *The truncated second main theorem of function fields*, J. Number Theory, to appear.

Department of Mathematics, University of Notre Dame, Notre Dame, IN 46556

Current address: Institute of Mathematics, Academia Sinica, Nankang, Taipei 11529, Taiwan
E-mail address: jwang@math.sinica.edu.tw