J. Austral. Math. Soc. (Series A) 25 (1978), 385-422

THUE'S EQUATION OVER FUNCTION FIELDS

WOLFGANG M. SCHMIDT

Dedicated to Professor K. Mahler on the occasion of his 75th birthday

(Received 14 November 1977)

Communicated by J. H. Coates

Abstract

Suppose we are given a "Thue equation" $f(x, y) = 1$, where f is a binary form with coefficients in a function field K of characteristic zero. A typical result is that if f is of degree at least 5 and has no multiple factors, then every solution $\mathbf{x} = (x, y)$ of the equation with components in K has $\mathrm{H}(\mathbf{x}) \leqslant 90\mathrm{H}(f) + 250g$. Here g is the genus of K and $\mathrm{H}(\mathbf{x})$, $\mathrm{H}(f)$ are suitably defined heights. No assumption is made that $\mathbf{x}$ be "integral" in some sense. As an application, bounds are derived for "integral" solutions of hyperelliptic equations over K.

Subject classification (*Amer. Math. Soc.* (*MOS*) 1970): primary 10 B 40; secondary 10 F 45, 10 M 05, 12 A 90.

1. Introduction

Thue (1909) proved that if $f(X, Y)$ is a form with rational coefficients and with at least 3 distinct linear factors, then the equation

$$f(x, y) = 1 \tag{1.1}$$

has only a finite number of solutions in rational integers x, y. In fact there are only finitely many rational solutions x, y whose denominators are composed of powers of primes belonging to an arbitrary but fixed finite set $\mathfrak{S}$ of primes. (Essentially Mahler (1933a, b).) This result continues to hold for number fields:

Partially supported by NSF-MCS 75-08233 A 02.

https://doi.org/10.1017/S1446788700021406 Published online by Cambridge University Press

If the coefficients of f lie in a number field K, then (1.1) has only finitely many solutions x, y in K having $v(x) \geqslant 0$, $v(y) \geqslant 0$ for every valuation v of K which does not belong to a given finite set $\mathfrak{S}$ of valuations. Now consider a hyperelliptic equation

$$y^2 = f(x) \tag{1.2}$$

where $f(X) = a(X-\alpha_1)^{e_1} \dots (X-\alpha_s)^{e_s}$ with at least 3 odd exponents e_i. It is a well-known consequence (Siegel (1929)) of the results on Thue equations that if $f(X)$ is a polynomial with coefficients in a number field K, then there are only finitely many solutions x, y in K with $v(x)$, $v(y) \geqslant 0$ for all valuation $v \notin \mathfrak{S}$.

All of the above results were originally derived from the non-effective method of Thue which does not allow one to find explicit bounds for the solutions. Baker (1968, 1969) was able to exhibit explicit bounds; more precisely, the "heights" of x, y are bounded in terms of the degree and the heights of the coefficients of f. These bounds enable one, at least in principle, to find all the solutions of (1.1) or (1.2) which are "$\mathfrak{S}$-integral" in the sense that $v(x) \geqslant 0$, $v(y) \geqslant 0$ for $v \notin \mathfrak{S}$.

Next, let k be an algebraically closed field of characteristic zero, and let $K = k(T)$ be the field of rational functions over k in the variable T. Consider a Thue equation (1.1) where now the coefficients of f lie in K, and consider possible solutions x, y which lie in the ring $k[T]$, that is, which are polynomials in T. Thue's method allows one to conclude (see, e.g., Uchiyama (1961) that these solutions x, y are polynomials of bounded degree. (One cannot assert the finiteness of the number of solutions, as is shown by the example $x^3 - 2y^3 = 1$, where k is the field of complex numbers and where there are infinitely many solutions x, y in k.) More generally, let K be a function field (of transcendence degree 1) over k, and let $\mathfrak{S}$ be a finite set of valuations of K/k. Then if $f(X, Y)$ has coefficients in K, the solutions x, y of (1.1) which lie in K and are $\mathfrak{S}$-integral have bounded *height*. Here the (additive) height $\mathbf{H}(x)$ is defined by

$$\mathbf{H}(x) = -\sum_{v} \min(0, v(x)),$$

where v runs through the valuations of K/k with value group $\mathbf{Z}$, the rational integers. (In the special case when $K = k(T)$ and $x \in k[T]$, we have $v(x) \geqslant 0$ except for the valuation $v_0(x) = -\deg x$, so that here $\mathbf{H}(x) = \deg x$.) Similar results pertain for the hyperelliptic equation (1.2). If these results are derived by the classical method of Thue, then no explicit bounds for $\mathbf{H}(x)$, $\mathbf{H}(y)$ can be given.

In (Schmidt (1976)) I used a more recent method of Osgood (1973, 1975) to derive bounds for the heights of Thue equations in the special case when $K = k[T]$ and $x, y \in k[T]$. This will now be extended to function fields. We first have to introduce some notation. Given a vector $\mathbf{x} = (x_1, \dots, x_n)$ with components in K, write

$$\mathbf{v}(\mathbf{x}) = \min(0, v(x_1), \dots, v(x_n)).$$

https://doi.org/10.1017/S1446788700021406 Published online by Cambridge University Press

If f is a polynomial with coefficients in K, let $\mathbf{v}(f)$ be defined in terms of the vector whose components are the coefficients of f. We define the height of $\mathbf{x}$ by

$$\mathrm{H}(\mathbf{x}) = -\sum_{v} \mathbf{v}(\mathbf{x}),$$

where v runs through the valuations of K/k with value group $\mathbf{Z}$, and we define the height $\mathrm{H}(f)$ of a polynomial in the obvious way. We note that

$$\mathbf{v}(x_1) + \ldots + \mathbf{v}(x_n) \leqslant \mathbf{v}(\mathbf{x}) \leqslant \mathbf{v}(x_i),$$

so that

$$\mathrm{H}(x_i) \leqslant \mathrm{H}(\mathbf{x}) \leqslant \mathrm{H}(x_1) + \ldots + \mathrm{H}(x_n) \quad (i = 1, \ldots, n).$$

Observe that if $K = k(T)$ and if $\mathbf{x} = (x_1/y, \ldots, x_n/y)$ with polynomials $x_1, \ldots, x_n$, y in $k[T]$ which are coprime though not necessarily coprime in pairs, then

$$\mathrm{H}(\mathbf{x}) = \max(\deg y, \deg x_1, \ldots, \deg x_n).$$

THEOREM 1. *Suppose K/k is a function field of genus g, and* (1.1) *is a Thue equation over K, where the form f is of degree d without multiple factors. Then*

(i) *If $d \geqslant 5$, every solution $\mathbf{x} = (x, y)$ with components in K has*

$$\mathrm{H}(\mathbf{x}) \leqslant 89\mathrm{H}(f) + 211g.$$

(ii) *If $d \geqslant 3$ and if $\mathfrak{S}$ is a finite set of valuations of K/k, then every $\mathfrak{S}$-integral solution has*

$$\mathrm{H}(\mathbf{x}) \leqslant 89\mathrm{H}(f) + 212g + |\mathfrak{S}| - 1,$$

where $|\mathfrak{S}|$ is the cardinality of $\mathfrak{S}$.

In the above theorem, the condition that f have no multiple factors can be relaxed.

No special importance is attached to the constants in our estimates, which could be improved with some extra effort. The estimate (ii) on $\mathfrak{S}$-integral solutions is not unexpected, but the estimate (i) is a surprise, since it is for all solutions with components in K. That the heights $\mathrm{H}(\mathbf{x})$ of all these solutions of (1.1) are bounded is known by the analogue of "Mordell's conjecture" for function fields, which was proved by Manin (1963) and by Grauert (1965). (See also Samuel (1966).) It is surprising that an estimate such as (i) comes out of the elementary arguments of the present paper.

COROLLARY 1.1. *Let a, b, c be non-zero polynomials lying in $k[T]$ and having degrees $\leqslant \delta$. Then if $d \geqslant 5$, the solutions of*

$$ax^d + by^d + cz^d = 0$$

in coprime non-zero polynomials x, y, z in $k[T]$ have degree at most 89δ.

https://doi.org/10.1017/S1446788700021406 Published online by Cambridge University Press

Namely, set $K = k(T)$, so that $g = 0$, and set $f(X, Y) = -(a/c)X^d - (b/c)Y^d$, so that $H(f) = \max(\deg a, \deg b, \deg c) \leqslant \delta$. By part (i) of the Theorem, the solutions of $f(x/z, y/z) = 1$ in coprime elements x, y, z of $k[T]$ have

$$\max(\deg x, \deg y, \deg z) = \mathbf{H}(x/z, y/z) \leqslant 89\mathbf{H}(f) \leqslant 89\delta.$$

COROLLARY 1.2. *Suppose that the form $f(X, Y)$ (of degree $d \geqslant 3$ and with distinct factors) has coefficients in $k[T]$ and that $m \neq 0$ lies in $k[T]$, and let $H(f, m)$ be the maximum of the degrees of m and of the coefficients of f. Then the solutions x, y in $k[T]$ of*

$$f(x, y) = m$$

have degrees not exceeding $89H(f, m)$.

In particular, if m and the coefficients of f lie in the ground field k, then so do x, y. The corollary follows from the second assertion of the theorem by the observation that if m and the coefficients of f are coprime (as polynomials in T), then by application of the sum formula (see (2.1)),

$$H(f, m) = -\sum_v \min(v(f_0), \ldots, v(f_d), v(m)) = \mathbf{H}(m^{-1}f)$$

where $f_0, \ldots, f_d$ are the coefficients of f and $m^{-1}f$ is the polynomial f divided by m. This corollary was already shown in Schmidt (1976).

THEOREM 2. *Let K/k be a function field of genus g, and let $\mathfrak{S}$ be a finite set of valuations of K/k. Let the polynomial $f(X)$ have its coefficients and its roots in K, and suppose that at least 3 of its roots have odd multiplicity. Then if x, y in K are solutions of the hyperelliptic equation (1.2) and if x is $\mathfrak{S}$-integral, then*

$$\mathbf{H}(x) \leqslant 10^6(\mathbf{H}(f) + g + |\mathfrak{S}|).$$

Clearly this implies an estimate also for $\mathbf{H}(y)$.

COROLLARY 2.1. *Suppose $m \neq 0$ and the coefficients as well as the roots of $f(X)$ lie in $k[T]$, and let $H(f, m)$ be the maximum of the degrees of m and of the coefficients of f. Then the solutions of*

$$my^2 = f(x) \tag{1.3}$$

with $x \in k[T]$, $y \in k(T)$ have

$$\deg x \leqslant 10^6 H(f, m). \tag{1.4}$$

We simply apply the theorem with $K = k(T)$, so that $g = 0$, and with $\mathfrak{S}$ consisting only of the valuation $v_0 = -\deg$, so that $|\mathfrak{S}| = 1$, to obtain

$$\deg x = \mathbf{H}(x) \leqslant 10^6(\mathbf{H}(m^{-1}f) + 1).$$

https://doi.org/10.1017/S1446788700021406 Published online by Cambridge University Press

Here $\mathbf{H}(m^{-1}f) = H(f,m)$ if, as we may suppose, m and the coefficients of f are coprime. Now replacing T by T^l, we have to replace $\deg x$ and $H(f,m)$ by l times themselves, so that we get $l\deg x \leqslant 10^6(lH(f,m)+1)$. Since this is true for arbitrarily large values of l, (1.4) follows.

COROLLARY 2.2. *Suppose $m \neq 0$ and the coefficients (but not necessarily the roots) of $f(X)$ lie in $k[T]$, and let $H(f,m)$ be as above. Then if d is the degree of f in X, every solution of* (1.3) *with $x \in k[T]$, $y \in k(T)$ has*

$$\deg x \leqslant 10^6 dH(f,m). \tag{1.5}$$

For the proof, set $K = k(T)$, and let $\mathfrak{S}_0$ consist of the valuation $v_0 = -\deg$ of K/k, so that $|\mathfrak{S}_0| = 1$. Further let K_1 be obtained from K by adjoining the roots of $f(X)$. The valuations of K_1/k extending v_0 form a set $\mathfrak{S}_1$ of cardinality $|\mathfrak{S}_1| \leqslant \Delta$, where Δ is the degree of K_1 over K. While $m^{-1}f$ had height $\mathbf{H}_K(m^{-1}f) \leqslant H(f,m)$ over K, its height over K_1 is $\mathbf{H}_{K_1}(m^{-1}f) = \Delta\mathbf{H}_K(m^{-1}f) \leqslant \Delta H(f,m)$ (see (2.11) below). Finally (Lemma H), the genus of K_1/k is $g_1 \leqslant (\Delta-1)dH(f,m)$. So by Theorem 2, the solutions x, y in K_1 with x $\mathfrak{S}_1$-integral have height

$$\mathbf{H}_{K_1}(x) \leqslant 10^6(\Delta H(f,m) + (\Delta-1)dH(f,m) + \Delta) \tag{1.6}$$

with respect to K_1. Now if x, y are in K then $\mathbf{H}_{K_1}(x) = \Delta H_K(x)$ (see (2.11)), so that upon dividing (1.6) by Δ we get

$$\deg x = \mathbf{H}_K(x) \leqslant 10^6(dH(f,m)+1).$$

Using again the trick of replacing T by T^l we obtain (1.5).

For the more special equation

$$my^2 = ax^d + b$$

the estimate for g_1 above can be improved to $g_1 < \Delta H(f,m)$ (see Lemma H), so that the final estimate may be improved to

$$\deg x \leqslant 10^6 \cdot 2H(f,m) = 2 \cdot 10^6 \max(\deg a, \deg b, \deg m).$$

But Davenport (1965) had obtained

$$\deg(ax^d - my^2) \geqslant \tfrac{1}{2}((d-2)\deg x - \deg a - \deg m) + 1$$

which holds unless $ax^d - my^2 = 0$ or $\deg x = 0$, which yields the much better estimate

$$(d-2)\deg x \leqslant 2\deg b + \deg a + \deg m - 2.$$

(The condition, $\deg x = 0$, is not stated by Davenport.)

The reader will be required to know the rudiments of the theory of function fields, including the Riemann-Roch Theorem.

2. Preliminaries

The additive version of the well-known product formula in K/k is the "sum formula"

$$\sum_v v(x) = 0 \tag{2.1}$$

for $x \neq 0$ in K. Here v runs through the valuations of K/k with value group $\mathbf{Z}$.

Besides the function $\mathbf{v}(\mathbf{x})$ introduced in the introduction, we shall need

$$v(\mathbf{x}) = \min(v(x_1), \ldots, v(x_m));$$

then $\mathbf{v}(\mathbf{x}) = \min(0, v(\mathbf{x}))$. Similarly define $v(f)$ for a polynomial f. Put

$$H(\mathbf{x}) = -\sum_v v(\mathbf{x}),$$

and define $H(f)$ in the obvious fashion. The sum formula shows that for $\mathbf{x} \neq \mathbf{0}$ and $f \neq 0$ we have $H(\mathbf{x}) \geqslant 0$, $H(f) \geqslant 0$. Furthermore, $H(\lambda\mathbf{x}) = H(\mathbf{x})$ if $\lambda \neq 0$ is in K. In particular, $H(\mathbf{x}) = 0$ if $\mathbf{x}$ is proportional to a vector in k^n. Conversely, if this is not the case, then some ratio $x_i/x_j \notin k$, so that for some v we have $x(x_i/x_j) \neq 0$ or $v(x_i) \neq v(x_j)$, which has the consequence that $H(\mathbf{x}) > 0$.

Gauss' Lemma says that

$$v(fg) = v(f) + f(g)$$

for polynomials f, g, and this implies that $H(fg) = H(f) + H(g)$. Repeated application of Gauss' Lemma shows that if $f(X) = f_0(X-\alpha_1)\ldots(X-\alpha_d)$, and if v is extended in some way to $K(\alpha_1, \ldots, \alpha_d)$, then

$$v(f) = v(f_0) + \sum_{i=1}^{d} \min(0, v(\alpha_i)) = v(f_0) + \sum_{i=1}^{d} \mathbf{v}(\alpha_i). \tag{2.2}$$

Thus if $\alpha_i \in K$, then

$$\mathrm{H}(\alpha_i) = -\sum_v \mathbf{v}(\alpha_i) \leqslant -\sum_v (v(f) - v(f_0)) = -\sum_v v(f) = H(f) \quad (1 \leqslant i \leqslant d). \tag{2.3}$$

It further follows from (2.2) that

$$v(f) \leqslant v(f_0\alpha_i) \quad (1 \leqslant i \leqslant d), \tag{2.4}$$

more generally that

$$v(f) \leqslant v(f_0\alpha_{i_1}\alpha_{i_2}\cdots\alpha_{i_s}) \quad (1 \leqslant i_1 < i_2 < \ldots < i_s \leqslant d). \tag{2.5}$$

The discriminant

$$D = f_0^{2d-2} \prod_{1 \leqslant i < j \leqslant d} (\alpha_i - \alpha_j)^2$$

is a polynomial of degree $2d-2$ in the coefficients of f, and hence

$$v(D) \geqslant (2d-2)\, v(f). \tag{2.6}$$

https://doi.org/10.1017/S1446788700021406 Published online by Cambridge University Press

Now D divided by $(\alpha_1-\alpha_2)^2$ is a polynomial of degree at most $2d-2$ in each α_i, so that by (2.5)

$$v(D/(\alpha_1-\alpha_2)^2)\geqslant(2d-2)\,v(f)$$

and

$$v(\alpha_1-\alpha_2)\leqslant\tfrac{1}{2}v(D)-(d-1)\,v(f). \tag{2.7}$$

In a similar way it is seen that

$$v\left(\frac{(\alpha_1-\alpha_2)(\alpha_1-\alpha_3)}{\alpha_2-\alpha_3}\right)\leqslant\tfrac{1}{2}v(D)-(d-1)\,v(f). \tag{2.8}$$

Finally, D divided by $(f_0(\alpha_1-\alpha_2)\dots(\alpha_1-\alpha_d))^2$ is a polynomial in each α_i of degree at most $2d-4$, so that by (2.5),

$$v(D/(f_0(\alpha_1-\alpha_2)\dots(\alpha_1-\alpha_d))^2)\geqslant(2d-4)\,v(f)$$

and

$$v(f_0(\alpha_1-\alpha_2)\dots(\alpha_1-\alpha_d))\leqslant\tfrac{1}{2}v(D)-(d-2)\,v(f). \tag{2.9}$$

Now let L be an extension of K of degree Δ. Write $V|v$ if the valuation V of L/k is an extension of the valuation v of K/k. If V has ramification index e_V over v, the value group of V is $e_V^{-1}\mathbf{Z}$. But we now want our valuations to have value group $\mathbf{Z}$, and hence we renormalize V to have value group $\mathbf{Z}$, so that now $V(x)=e_V\,v(x)$ for $x\in k$. It is well known that

$$\sum_{V|v}e_V=\Delta,$$

and therefore

$$\sum_{V|v}V(x)=\Delta v(x)\quad\text{and}\quad\sum_{V|v}\mathbf{V}(\mathbf{x})=\Delta\mathbf{v}(\mathbf{x}) \tag{2.10}$$

if $x\in K$ and $\mathbf{x}\in K^n$. If $\mathbf{x}\in K^n$ we may form both the height $\mathbf{H}_K(\mathbf{x})$ defined over K and the height $\mathbf{H}_L(\mathbf{x})$ defined over L. In view of (2.10) we have

$$\mathbf{H}_L(\mathbf{x})=\Delta\mathbf{H}_K(\mathbf{x}). \tag{2.11}$$

Suppose $f(X)=f_0x^\Delta+\dots+f_\Delta$ is irreducible over K, let α be a root of f, and let $L=K(\alpha)$, so that $[L:K]=\Delta$. We claim that

$$v(f)=v(f_0)+\sum_{V|v}\mathbf{V}(\alpha) \tag{2.12}$$

for any valuation v of K/k. For if L' is the splitting field of f over K and if f has roots $\alpha=\alpha_1,\alpha_2,\dots,\alpha_\Delta$ in L', then every valuation V' of L'/k has

$$V'(f)=V'(f_0)+\sum_{i=1}^{\Delta}\mathbf{V}'(\alpha_i)$$

by (2.2). Given a valuation v of K/k and summing over all its extensions V' to L', we obtain

$$\Delta' v(f) = \Delta' v(f_0) + \sum_{i=1}^{\Delta} \sum_{V'|v} \mathbf{V}'(\alpha_i), \tag{2.13}$$

where $\Delta' = [L' : K]$. Since $\alpha = \alpha_1, \ldots, \alpha_\Delta$ are conjugates,

$$\begin{aligned} \sum_{i=1}^{\Delta} \sum_{V'|v} \mathbf{V}'(\alpha_i) &= \Delta \sum_{V'|v} \mathbf{V}'(\alpha) = \Delta \sum_{\substack{V|v \\ \text{of } L}} \sum_{\substack{V'|V \\ \text{of } L'}} \mathbf{V}'(\alpha) \\ &= \Delta[L' : L] \sum_{V|v} \mathbf{V}(\alpha) = \Delta' \sum_{V|v} \mathbf{V}(\alpha). \end{aligned}$$

Dividing (2.13) and the last equation by Δ' we obtain (2.12).

3. Geometry of numbers in function fields

Let K be a function field (in one variable) over the ground field k. We allow more generality in this section than in the rest of the paper: we assume k to be algebraically closed in K, but k need not be algebraically closed or be of characteristic zero. Prime divisors of K/k will be denoted by $\mathfrak{P}$. With $\mathfrak{P}$ is associated a "place" of K/k with a residue class field which is a finite algebraic extension of k, of a degree $d(\mathfrak{P})$ called the *degree* of $\mathfrak{P}$. Also associated with $\mathfrak{P}$ is a valuation $v_{\mathfrak{P}}$ with value group $\mathbf{Z}$, the integers. The *group of divisors* is the free abelian multiplicative group generated by the prime divisors. The definition of the degree is extended from prime divisors to divisors in general by the rule that $d(\mathfrak{A}\mathfrak{B}) = d(\mathfrak{A}) + d(\mathfrak{B})$ for any divisors $\mathfrak{A}, \mathfrak{B}$. Further write $v_{\mathfrak{P}}(\mathfrak{A}) = n$ if $\mathfrak{P}$ occurs with exponent n in the representation of $\mathfrak{A}$ as a product of prime divisors.

Define $L(\mathfrak{A})$ as the set of $X \in K$ having $v_{\mathfrak{P}}(X) \geqslant v_{\mathfrak{P}}(\mathfrak{A})$ for every $\mathfrak{P}$. Then $L(\mathfrak{A})$ turns out to be a finite dimensional vector space over k; we denote its dimension *by* $l(\mathfrak{A})$. Riemann's Theorem (which is part of the Riemann–Roch Theorem) asserts that

$$l(\mathfrak{A}^{-1}) \geqslant d(\mathfrak{A}) + 1 - g$$

for a certain integer g depending only on K/k. The smallest integer g with this property turns out to be non-negative and is called the *genus* of K/k.

It is known (Eichler (1963); see also Armitage (1967)) that the Riemann–Roch Theorem is a consequence of the analogue of Minkowski's Geometry of Numbers where integers and reals are replaced by polynomials and power series, respectively; this new type of Geometry of Numbers was first studied by Mahler (1940). Our next theorem contains both Riemann's Theorem and Mahler's analogue of Minkowski's Linear Forms Theorem.

Let $\mathfrak{P}$ be a divisor and $A = (\alpha_{ij})$ an $(n \times n)$-matrix with entries in $K_{\mathfrak{P}}$, the completion of K with respect to $v_{\mathfrak{P}}$. Write $L_{\mathfrak{P}}(A)$ for the set of n-tuples $(x_1, \ldots, x_n) \in K^n$ having

$$\begin{aligned} &v_{\mathfrak{P}}(\alpha_{11}x_1 + \ldots + \alpha_{1n}x_n) \geqq 0, \\ (3.1) \qquad &\ldots\ldots \\ &v_{\mathfrak{P}}(\alpha_{n1}x_1 + \ldots + \{\alpha_{nn}x_n) \geqq 0. \end{aligned}$$

$L_{\mathfrak{P}}(A)$ is easily seen to be a vector space over k.

A *matrix repartition* $\mathscr{A}$ will be a mapping $\mathfrak{P} \to A_{\mathfrak{P}}$ from the set of prime divisors into matrices such that $A_{\mathfrak{P}} = I$, the identity matrix, for all but finitely many $\mathfrak{P}$. We put

$$d(\mathscr{A}) = \sum_{\mathfrak{P}} d(\mathfrak{P}) v_{\mathfrak{P}}(\det A_{\mathfrak{P}}).$$

With $\mathscr{A}$ we associate

$$L(\mathscr{A}) = \bigcap_{\mathfrak{P}} L_{\mathfrak{P}}(A_{\mathfrak{P}}),$$

which consists of $(x_1, \ldots, x_n)$ having (3.1) for each $\mathfrak{P}$, and which is obviously a vector space over k. Writing $l(\mathscr{A})$ for $\dim_k L(\mathscr{A})$ we have

THEOREM 3. *$l(\mathscr{A}) \geqq d(\mathscr{A}) + n - ng$.*

The case $n = 1$ is Riemann's Theorem. For if $A_{\mathfrak{P}} = (\alpha_{\mathfrak{P}})$ and if we put $\mathfrak{A} = \Pi_{\mathfrak{P}}{}^{v_{\mathfrak{P}}(\alpha_{\mathfrak{P}})}$, then $d(\mathfrak{A}) = d(\mathscr{A})$ and $L(\mathfrak{A}^{-1}) = L(\mathscr{A})$. Conversely, given $\mathfrak{A}$ there is an $\mathscr{A}$ with $d(\mathscr{A}) = d(\mathfrak{A})$ and $L(\mathscr{A}) = L(\mathfrak{A}^{-1})$. The constant $n - ng$ in the theorem is best possible as may be seen by choosing suitable diagonal matrices $A_{\mathfrak{P}}$. If some $A_{\mathfrak{P}}$ is singular, then $v(\det A_{\mathfrak{P}}) = \infty$, and the theorem asserts that $l(\mathscr{A}) = \infty$. Many proofs are possible; we choose here to deduce everything from Riemann's Theorem.

PROOF. The inductive step from $n-1$ to n is as follows. We first reduce the case where some $A_{\mathfrak{P}}$ is singular to the case where each $A_{\mathfrak{P}}$ is non-singular. If $\lambda_1 \mathbf{a}_1 + \ldots + \lambda_n \mathbf{a}_n = \mathbf{0}$ is a non-trivial relation of linear dependence of the rows of $A_{\mathfrak{P}}$, suppose without loss of generality that $v_{\mathfrak{P}}(\lambda_1) \geqq \ldots \geqq v_{\mathfrak{P}}(\lambda_n)$. Then

$$\mathbf{a}_n = \mu_1 \mathbf{a}_1 + \ldots + \mu_{n-1} \mathbf{a}_{n-1}$$

with $v(\mu_i) \geqq 0$, and the n inequalities (3.1) follow from the first $n-1$ inequalities. Continuing in this way we see that after reordering, the first rows of $A_{\mathfrak{P}}$, say $r < n$ rows, will be linearly independent, and the n inequalities (3.1) follow from the first r inequalities. We can choose a non-singular $A'_{\mathfrak{P}}$ whose first r rows are the same as those of $A_{\mathfrak{P}}$, and which has $v_{\mathfrak{P}}(\det A'_{\mathfrak{P}})$ arbitrarily large. Let $\mathscr{A}'$ be obtained from $\mathscr{A}$ by replacing each singular $A_{\mathfrak{P}}$ by $A'_{\mathfrak{P}}$. Then $L(\mathscr{A}') = L(\mathscr{A}')$ and $l(\mathscr{A}) \geqq d(\mathscr{A}') + n - ng$. Since we can make $d(\mathscr{A}')$ arbitrarily large, we get $l(\mathscr{A}) = \infty$.

https://doi.org/10.1017/S1446788700021406 Published online by Cambridge University Press

We may thus suppose that each $A_{\mathfrak{P}}$ is non-singular. Suppose without loss of generality that $v_{\mathfrak{P}}(\alpha_{11}^{(\mathfrak{P})}) \leqslant \ldots \leqslant v_{\mathfrak{P}}(\alpha_{n1}^{(\mathfrak{P})})$. Then for a given $\mathfrak{P}$, the system (3.1) of inequalities is equivalent to the one obtained by subtracting $\alpha_{21}^{(\mathfrak{P})}/\alpha_{11}^{(\mathfrak{P})}$ times $\alpha_{11}^{(\mathfrak{P})}x_1 + \ldots + \alpha_{n1}^{(\mathfrak{P})}x_n$ from $\alpha_{21}^{(\mathfrak{P})}x_1 + \ldots + \alpha_{2n}^{(\mathfrak{P})}x_n$. We thus may suppose without loss of generality that $\alpha_{21}^{(\mathfrak{P})} = 0$, and more generally that $\alpha_{21}^{(\mathfrak{P})} = \ldots = \alpha_{n1}^{(\mathfrak{P})} = 0$. Then $A_{\mathfrak{P}}$ is of the form

$$A_{\mathfrak{P}} = \begin{pmatrix} \alpha_1^{(\mathfrak{P})} & \alpha_2^{(\mathfrak{P})} & \ldots & \alpha_n^{(\mathfrak{P})} \\ 0 & \beta_{22}^{(\mathfrak{P})} & \ldots & \beta_{2n}^{(\mathfrak{P})} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & \beta_{n2}^{(\mathfrak{P})} & \ldots & \beta_{nn}^{(\mathfrak{P})} \end{pmatrix} = \begin{pmatrix} \alpha_1^{(\mathfrak{P})} & \mathbf{a}^{(\mathfrak{P})} \\ 0 & B_{\mathfrak{P}} \end{pmatrix},$$

say.

We know from our induction hypothesis that $(n-1)$-tuples $(x_2, \ldots, x_n)$ having

$$(3.2) \qquad v_{\mathfrak{P}}(\beta_{j2}^{(\mathfrak{P})}x_2 + \ldots + \beta_{jn}^{(\mathfrak{P})}x_n) \geqslant 0 \quad (j = 2, \ldots, n)$$

for all $\mathfrak{P}$ form a vector space $L(\mathscr{B}))$ over k of dimension

$$\begin{aligned} l(\mathscr{B}) &= d(\mathscr{B}) + (n-1)(1-g) \\ &= d(\mathscr{A}) + n - ng - \sum_{\mathfrak{P}} d_{\mathfrak{P}} v_{\mathfrak{P}}(\alpha_1^{(\mathfrak{P})}) + g - 1). \end{aligned}$$

Suppose the matrices $A_{\mathfrak{P}}$ are distinct from I for $\mathfrak{P}$ in the finite set Π. Choose a large positive integer c and write $L(\mathscr{C})$ for the set of $x_1 \in K$ having

$$\begin{aligned} &v_{\mathfrak{P}}(\alpha_1^{(\mathfrak{P})}x_1) \geqslant -c \quad &&\text{for } \mathfrak{P} \in \Pi \\ &v_{\mathfrak{P}}(x_1) \geqslant 0 &&\text{for } \mathfrak{P} \notin \Pi \end{aligned}$$

By Riemann's Theorem, $L(\mathscr{C})$ is a vector space of dimension

$$\geqslant c \sum_{\mathfrak{P} \in \Pi} d(\mathfrak{P}) + \sum_{\mathfrak{P}} d(\mathfrak{P}) v_{\mathfrak{B}}(\alpha_1^{(\mathfrak{P})}) + 1 - g.$$

So if $L(\mathscr{B}) + L(\mathscr{C})$ consists of n-tuples $(x_1, \ldots, x_n)$ with

$$x_1 \in L(\mathscr{C}) \quad \text{and} \quad (x_2, \ldots, x_n) \in L(\mathscr{B}),$$

then

$$\dim(L(\mathscr{B}) + L(\mathscr{C})) \geqslant d(\mathscr{A}) + n - ng + c \sum_{\mathfrak{P} \in \Pi} d(\mathfrak{P}).$$

If c is chosen sufficiently large, then

$$(3.3) \qquad v_{\mathfrak{P}}(\alpha_1^{(\mathfrak{P})}x_1 + \ldots + \alpha_n^{(\mathfrak{P})}x_n) \geqslant -c$$

for every $\mathfrak{P} \in \Pi$ and every n-tuple in $L(\mathscr{B}) + L(\mathscr{C})$: this follows from the fact that (since $B_{\mathfrak{P}}$ is non-singular) $\alpha_2^{(\mathfrak{P})}x_2 + \ldots + \alpha_n^{(\mathfrak{P})}x_n$ is a linear combination of $\beta_{i2}^{(\mathfrak{P})}x_2 + \ldots + \beta_{in}^{(\mathfrak{P})}x_n$ $(i = 2, \ldots, n)$, so that $v_{\mathfrak{P}}(\alpha_2^{(\mathfrak{P})}x_2 + \ldots + \alpha_n^{(\mathfrak{P})}x_n) \geqslant -c$ by (3.2).

https://doi.org/10.1017/S1446788700021406 Published online by Cambridge University Press

But of course what we *want* in $L(\mathscr{A})$ is that

$$v_{\mathfrak{P}}(\alpha_1^{(\mathfrak{P})}x_1+\ldots+\alpha_n^{(\mathfrak{P})}x_n)\geqslant 0. \tag{3.4}$$

Lemma A. *Let S be a vector space over k consisting of certain n-tuples $(x_1,\ldots,x_n)$ with components in K, and suppose that*

$$v_{\mathfrak{P}}(\alpha_1^{(\mathfrak{P})}x_1+\ldots+\alpha_n^{(\mathfrak{P})}x_n)\geqslant -l$$

for some particular $\mathfrak{P}$ and every $(x_1,\ldots,x_n)\in S$. Then if S' is the subspace of S where

$$v_{\mathfrak{P}}(\alpha_1^{(\mathfrak{P})}x_1+\ldots+\alpha_n^{(\mathfrak{P})}x_n)\geqslant -(l-1),$$

we have

$$\dim S'\geqslant \dim S-d(\mathfrak{P}).$$

Assuming the truth of the lemma, we may decrease c to $c-1$ in (3.3) for one particular $\mathfrak{P}\in\Pi$, to obtain a subspace S' of $L(\mathscr{B})+L(\mathscr{C})$ of dimension $\geqslant\dim(L(\mathscr{B})+L(\mathscr{C}))-d(\mathfrak{P})$. Gradually reducing c to 0 for this particular $\mathfrak{P}$, we get a subspace S'' of $L(\mathscr{B})+L(\mathscr{C})$ of dimension $\geqslant\dim(L(\mathscr{B})+L(\mathscr{C}))-cd(\mathfrak{P})$. If we reduce c to 0 for every $\mathfrak{P}\in\Pi$, we obtain a subspace S''' of dimension

$$\begin{aligned}\dim S'''&\geqslant\dim(L(\mathscr{B})+L(\mathscr{C}))-c\sum_{\mathfrak{P}\in\Pi}d(\mathfrak{P})\\&\geqslant d(\mathscr{A})+n-ng.\end{aligned}$$

Since S''' is contained in $L(\mathscr{A})$, the theorem will follow.

As for the lemma, it will suffice to show that any $d(\mathfrak{P})+1$ elements $\mathbf{x}_i=(x_{i1},\ldots,x_{in})$ of S have a linear combination which lies in S'. Let β be so that $v_{\mathfrak{P}}(\beta)=l$. The image of $\beta(\alpha_1^{(\mathfrak{P})}x_{i1}+\ldots+\alpha_n^{(\mathfrak{P})}x_{in})$ under the place $\mathfrak{P}$ is an element γ_i in the residue class field of $\mathfrak{P}$. This residue class field is of degree $d(\mathfrak{P})$ over k, so that we have a non-trivial relation

$$c_1\gamma_1+\ldots+c_{d(\mathfrak{P})+1}\gamma_{d(\mathfrak{P})+1}=0$$

with coefficients in k. Thus setting

$$\mathbf{x}=c_1\mathbf{x}_1+\ldots+c_{d(\mathfrak{P})+1}\mathbf{x}_{d(\mathfrak{P})+1}=(x_1,\ldots,x_n),$$

say, we have $v_{\mathfrak{P}}(\beta(\alpha_1^{(\mathfrak{P})}x_1+\ldots+\alpha_n^{(\mathfrak{P})}x_n))>0$, so that

$$v_{\mathfrak{P}}(\alpha_1^{(\mathfrak{P})}x_1+\ldots+\alpha_n^{(\mathfrak{P})}x_n)\geqslant -(l-1),$$

and $\mathbf{x}\in S'$.

The following is now obvious:

Corollary 3.1. *Suppose that for each $\mathfrak{P}$ the inequalities* (3.1) *are replaced by*

$$\begin{aligned}&v_{\mathfrak{P}}(\alpha_{11}^{(\mathfrak{P})}x_1+\ldots+\alpha_{1n}^{(\mathfrak{P})}x_n)\geqslant c_1^{(\mathfrak{P})},\\&\ldots\ldots\\&v_{\mathfrak{P}}(\alpha_{n1}^{(\mathfrak{P})}x_1+\ldots+\alpha_{nn}^{(\mathfrak{P})}x_n)\geqslant c_n^{(\mathfrak{P})},\end{aligned}\tag{3.5}$$

where the integers $c_i^{(\mathfrak{P})}$ are zero with finitely many exceptions. Put

$$d(\mathbf{c}) = \sum_{\mathfrak{P}} \sum_{i=1}^{n} d(\mathfrak{P})\, v_{\mathfrak{P}}(c_i^{(\mathfrak{P})}).$$

The solutions $(x_1, \dots, x_n) \in K^n$ of (3.5) *form a k-vector space of dimension*

$$\geqslant d(\mathscr{A}) - d(\mathbf{c}) + n - ng.$$

4. Linear equations

From here on k will again be of characteristic zero and algebraically closed. Let K/k be a function field of genus g.

THEOREM 4. *Let $L_1, \dots, L_s$ be extensions of K of respective degrees $\Delta(1), \dots, \Delta(s)$, put $d = \Delta(1) + \dots + \Delta(s)$, and suppose that $n > d$. For $i = 1, \dots, s$ let $\mathbf{y}_i = (y_{i1}, \dots, y_{in})$ be a non-zero vector in L_i^n of height $H_i(\mathbf{y}_i)$ over L_i. Then there is a non-zero $\mathbf{x} = (x_1, \dots, x_n) \in K^n$ with*

$$x_1 y_{i1} + \dots + x_n y_{in} = 0 \quad (i = 1, \dots, s) \tag{4.1}$$

having

$$H_K(\mathbf{x}) \leqslant \frac{1}{n-d}(H_1(\mathbf{y}_1) + \dots + H_s(\mathbf{y}_s) + ng).$$

PROOF. All but finitely many valuations v of K/k are unramified in L_i, and all but finitely many valuations V_i of L_i/k have

$$\min(V_i(y_{i1}), \dots, V_i(y_{in})) = 0. \tag{4.2}$$

Pick a valuation v_0 of K/k which is unramified in $L_1, \dots, L_s$ and such that for $i = 1, \dots, s$ each extension V_i to L_i satisfies (4.2). Then v_0 extends to $\Delta(i)$ distinct valuations $V_{i1}, \dots, V_{i\Delta(i)}$ in L_i. There are $\Delta(i)$ embeddings $\varphi_{i1}, \dots, \varphi_{i\Delta(i)}$ of L_i into the completion $\hat{K}$ of K with respect to v_0; write $\varphi_{ij}(z) = z^{(ij)}$. The valuations V_{ij} are then given by $V_{ij}(z) = v_0(z^{(ij)})$ $(i = 1, \dots, s; j = 1, \dots, \Delta(i))$.

Consider the linear form

$$\mathfrak{L}^{(ij)}(\mathbf{x}) = x_1 y_{i1}^{(ij)} + \dots + x_n y_{in}^{ij} \quad (i = 1, \dots, s; j = 1, \dots, \Delta(i)).$$

If $\mathfrak{L}^{(11)}, \dots, \mathfrak{L}^{(1\Delta(1))}$ are linearly dependent, say $c_1 \mathfrak{L}^{(11)} + \dots + c_{\Delta(1)} \mathfrak{L}^{(1\Delta(1))} = 0$ with $c_j \in \hat{K}$, we may suppose that $v_0(c_{\Delta(1)}) \leqslant \dots \leqslant v_0(c_1)$. Dividing by $c_{\Delta(1)}$ and changing the notation we obtain

$$\mathfrak{L}^{(1\Delta(1))} = a_1 \mathfrak{L}^{(11)} + \dots + a_{\Delta(1)-1} \mathfrak{L}^{(1(\Delta(1)-1))},$$

with $v_0(a_l) \geqslant 0$. Continuing in this fashion and reordering, we obtain linearly independent forms $\mathfrak{L}^{(11)}, \dots, \mathfrak{L}^{(1\Phi(1))}$ such that each $\mathfrak{L}^{(1j)}$ is a combination of these

https://doi.org/10.1017/S1446788700021406 Published online by Cambridge University Press

forms with coefficients $a_{jl} = a_{jl}^{(1)}$ having $v_0(a_{jl}^{(1)}) \geqslant 0$. Next, if $\mathfrak{L}^{(11)}, \dots, \mathfrak{L}^{(1\Phi(1))}$, $\mathfrak{L}^{(21)}, \dots, \mathfrak{L}^{(2\Delta(2))}$ are linearly dependent, say

$$u_1 \mathfrak{L}^{(11)} + \dots + u_{\Phi(1)} \mathfrak{L}'^{(1\Phi(1))} + u_1 \mathfrak{L}^{(21)} + \dots + u_{\Delta(2)} \mathfrak{L}^{(2\Delta(2))} = 0,$$

we may suppose that $v_0(\hat{u}_{\Delta(2)}) \leqslant \dots \leqslant v_0(\hat{u}_1)$, so that with a different notation we obtain

$$\mathfrak{L}^{(2\Delta(2))} = b_1^{(2)} \mathfrak{L}^{(11)} + \dots + b_{\Delta(1)}^{(2)} \mathfrak{L}^{(1\Delta(1))} + a_1^{(2)} \mathfrak{L}^{(21)} + \dots + a_{\Delta(2)-1}^{(2)} \mathfrak{L}^{(2(\Delta(2))-1)}$$

with $v_0(a_l^{(2)}) \geqslant 0$. Continuing and reordering we obtain linearly independent forms $\mathfrak{L}^{(11)}, \dots, \mathfrak{L}^{(1\Phi(1))}, \mathfrak{L}^{(21)}, \dots, \mathfrak{L}^{(2\Phi(2))}$ such that each $\mathfrak{L}^{(2j)}$ is a linear combination of these forms in such a way that the coefficient $a_{jl}^{(2)}$ of $\mathfrak{L}^{(2l)}$ has $v_0(a_{jl}^{(2)}) \geqslant 0$. It is possible that each $\mathfrak{L}^{(2j)}$ is a linear combination of $\mathfrak{L}^{(11)}, \dots, \mathfrak{L}^{(1\Delta(1))}$, in which case $\Delta(2) = 0$. Repeating this process with the forms $\mathfrak{L}^{(3j)}$, etc., we finally get linearly independent forms $\mathfrak{L}^{(11)}, \dots, \mathfrak{L}^{(1\Phi(1))}, \dots, \mathfrak{L}^{(s1)}, \dots, \mathfrak{L}^{(s\Phi(s))}$ such that each $\mathfrak{L}^{(ij)}$ is a linear combination of the forms $\mathfrak{L}^{(hl)}$ with $h \leqslant i$ and $1 \leqslant l \leqslant \Phi(h)$ where the coefficient $a_{jl}^{(i)}$ of $\mathfrak{L}^{(il)}$ has $v_0(a_{jl}^{(i)}) \geqslant 0$.

The equations $\mathfrak{L}^{(ij)}(\mathbf{x}) = 0$ $(1 \leqslant i \leqslant s,\ 1 \leqslant j \leqslant \Delta(i))$ define a subspace of n-space of dimension

$$n - \Phi(1) - \dots - \Phi(s) = n - \Phi = e,$$

say. After reordering of coordinates, the equations of this subspace will be

$$x_i = b_{i1} x_1 + \dots + b_{ie} x_e \quad (i = e+1, \dots, n), \tag{4.3}$$

with coefficients b_{ij} having $v_0(b_{ij}) \geqslant 0$.

Let P_i $(i = 1, \dots, s)$ be the smallest integer with $P_i \geqslant \Delta(i)^{-1} H_i(\mathbf{y}_i)$, and define P_{ij} $(i = 1, \dots, s;\ j = 1, \dots, \Phi(i))$ by

$$P_{ij} = \begin{cases} P_i + 1 & \text{if } j = 1, \\ P_i & \text{otherwise.} \end{cases}$$

Let Q be the smallest integer with

$$Q \geqslant e^{-1} \left(\sum_{i=1}^{s} \sum_{j=1}^{\Phi(i)} P_{ij} + ng + 1 - n \right). \tag{4.4}$$

The system of inequalities

$$v(x_1) \geqslant 0, \dots, v(x_n) \geqslant 0 \quad \text{for } v \neq v_0, \tag{4.5a}$$

$$v_0(\mathfrak{L}^{(ij)}(\mathbf{x})) \geqslant P_{ij} \quad (1 \leqslant i \leqslant s, 1 \leqslant j \leqslant \Phi(i)), \tag{4.5b}$$

$$v_0(x_1), \dots, v_0(x_e) \geqslant -Q \tag{4.5c}$$

is a system such as (3.5) in Corollary 3.1. We have $A_{\mathfrak{P}} = I$ unless $\mathfrak{P} = \mathfrak{P}_0$, the prime divisor belonging to v_0. Moreover $v_0(\det A_{\mathfrak{P}_0}) \geqslant 0$ since

$$\min(v_0(y_{i1}^{(ij)}), \dots, v_0(y_{in}^{(ij)})) = \min(V_{ij}(y_{i1}), \dots, V_{ij}(y_{in})) = 0$$

by our choice of v_0 below (4.2). Thus $d(\mathscr{A}) \geqslant 0$. Since each $d(\mathfrak{P}) = 1$, since $c_i^{(\mathfrak{P})} = 0$ unless $\mathfrak{P} = \mathfrak{P}_0$, and since

$$\sum_{i=1}^{n} v_0(c_i^{(\mathfrak{P}_0)}) = \sum_{i=1}^{s} \sum_{j=1}^{\Phi(i)} P_{ij} - eQ \leqslant -(ng+1-n),$$

we find that

$$d(\mathscr{A}) - d(\mathbf{c}) + n - ng > 0.$$

By Corollary 3.1 there is a non-zero $\mathbf{x} = (x_1, \dots, x_n) \in K^n$ with (4.5a), (4.5b), (4.5c).

Since each $\mathfrak{L}^{(1j)}$ is a linear combination of $\mathfrak{L}^{(11)}, \dots, \mathfrak{L}^{(1\Phi(1))}$ with coefficients $a_{jl}^{(1)}$ having $v_0(a_{jl}^{(1)}) \geqslant 0$, we have

$$(4.6) \qquad v_0(\mathfrak{L}^{(1j)}(\mathbf{x})) \geqslant P_1 \quad (j = 1, \dots, \Delta(1))$$

by (4.5b).

Now if V_1 ranges over the valuations of L_1/k,

$$\begin{aligned} &\sum_{V_1} V_1(y_{11}x_1 + \dots + y_{1n}y_n) \\ &\qquad = \sum_{V_1 | v_0} V_1(\dots) + \sum_{V_1 \nmid v_0} V_1(\dots) \\ &\qquad \geqslant \sum_{j=1}^{\Delta(1)} v_0(\mathfrak{L}^{(1j)}(\mathbf{x})) + \sum_{V_1 \nmid v_0} \min(V_1(y_{11}), \dots, V_1(y_{1n})) \end{aligned}$$

by (4.5a); and by (4.5b) and (4.6) this is

$$(4.7) \qquad \geqslant \Delta(1) P_1 + 1 + \sum_{V_1} \min(V_1(y_{11}), \dots, V_1(y_{1n})),$$

since by our choice of v_0 below (4.2) the minimum in (4.7) is zero if $V_1 | v_0$. Since $\Delta(1) P_1 + 1 > H_1(\mathbf{y}_1)$, and by the definition of $H_1(\mathbf{y}_1)$ we obtain

$$> H_1(\mathbf{y}_1) - H_1(\mathbf{y}_1) = 0.$$

An appeal to the sum formula in L_1/k yields

$$(4.8) \qquad x_1 y_{11} + \dots + x_n y_{1n} = 0.$$

Since each $\mathfrak{L}^{(2j)}$ is a linear combination of $\mathfrak{L}^{(11)}, \dots, \mathfrak{L}^{(1\Phi(1))}, \mathfrak{L}^{(21)}, \dots, \mathfrak{L}^{(2\Phi(2))}$, since $\mathfrak{L}^{(11)}(\mathbf{x}) = \dots = \mathfrak{L}^{(1\Phi(1))}(\mathbf{x}) = 0$ by (4.8), and since the coefficient $a_{jl}^{(2)}$ of $\mathfrak{L}^{(2l)}$ in the combination has $v_0(a_{jl}^{(2)}) \geqslant 0$, we have

$$v_0(\mathfrak{L}^{(2j)}(\mathbf{x})) \geqslant P_2 \quad (j = 1, \dots, \Delta(2))$$

by (4.5b). If $\Phi(2)=0$, then

$$x_1 y_{21}+\ldots+x_n y_{2n}=0, \tag{4.9}$$

more generally $\mathfrak{L}^{(2j)}(\mathbf{x})=0$, follows from (4.8). If $\Phi(2)>0$ we find that with V_2 ranging over the valuation of L_2/k,

$$\sum_{V_2} V_2(y_{21}x_1+\ldots+y_{2n}x_n)>0,$$

which again implies (4.9).

Continuing in this manner we see that (4.1) holds for each i, hence that $\mathfrak{L}^{(ij)}(\mathbf{x})=0$ for each i,j. Thus (4.3) holds, which in conjunction with (4.5c) yields $v_0(x_j)\geqslant -Q$ $(j=1,\ldots,n)$, so that by (4.5a) we get

$$H_K(\mathbf{x})=-\sum_v \min(v(x_1),\ldots,v(x_n))\leqslant -\min(v_0(x_1),\ldots,v_0(x_n))\leqslant Q. \tag{4.10}$$

Now when $\Phi(i)>0$ we have

$$\begin{aligned}\sum_{j=1}^{\Phi(i)} P_{ij} &= \Phi(i)\,P_i+1\\ &\leqslant \Phi(i)\,\Delta(i)^{-1}(H_i(\mathbf{y}_i)+\Delta(i)-1)+1\\ &\leqslant H_i(\mathbf{y}_i)+\Phi(i),\end{aligned}$$

and

$$\begin{aligned}&\sum_{i=1}^{s}\sum_{j=1}^{\Phi(i)} P_{ij}+ng+1-n\\ &\qquad\leqslant H_1(\mathbf{y}_1)+\ldots+H_s(\mathbf{y}_s)+ng-(n-\Phi-1)\\ &\qquad= H_1(\mathbf{y}_1)+\ldots+H_s(\mathbf{y}_s)+ng-(e-1).\end{aligned}$$

By the definition of Q involving (4.4) we have

$$Q\leqslant e^{-1}(H_1(\mathbf{y}_1)+\ldots+H_s(\mathbf{y}_s)+ng),$$

and the theorem follows from (4.10) and from $e=n-\Phi\geqslant n-d$.

5. Construction of a differential equation

Let K/k be a function field. Assume that we are given an element $T\in K$, $T\notin k$; then T is transcendental over k. Thus we have the function fields

$$k(T)\subseteq K.$$

Each valuation v of K/k is the extension of some valuation u of $k(T)/k$; denote the ramification index of v over u by ε_v. (v has value group $\mathbf{Z}$ and hence u has value group $\varepsilon_v\mathbf{Z}$.) The valuation $u_\infty=-\deg$ of $k(T)/k$ with $u_\infty(T)=-1$ will be called

https://doi.org/10.1017/S1446788700021406 Published online by Cambridge University Press

the "infinite valuation". Extensions of u_∞ to K will be called "infinite valuations"; the other valuations of K/k will be called "finite". The divisor

$$\mathfrak{D} = \prod_{v \text{ finite}} \mathfrak{P}_v^{\varepsilon_v - 1},$$

where $\mathfrak{P}_v$ is the prime divisor associated with v is the *different*. Its degree is

$$d(\mathfrak{D}) = \sum_{v \text{ finite}} (\varepsilon_v - 1).$$

The derivation with respect to T in $k(T)$ may be uniquely extended to a derivation in K, and indeed in any finite algebraic extension of K. Denote the derivation of x by x' or by dx/dT.

LEMMA B. *Suppose $x \in K$. Then*
(a) $v(x') \geqslant v(x) + \varepsilon_v$ *if v is infinite,*
(b) $v(x') \geqslant v(x) - \varepsilon_v$ *if v is finite,*
(c) $v(x') \geqslant 1 - \varepsilon_v$ *if v is finite and* $v(x) \geqslant 0$.

PROOF. If v is infinite, then there is an embedding of K into a field $k((T_v))$ of formal power series in a quantity T_v such that $v(c_a T_v^a + c_{a+1} T_v^{a+1} + \ldots) = a$ and $T = 1/T_v^{\varepsilon_v}$. Then $v(dx/dT_v) \geqslant v(x) - 1$ and $v(dT/dT_v) = -\varepsilon_v - 1$, so that

$$v(x') = v(dx/dT) = v(dx/dT_v) - v(dT/dT_v) \geqslant v(x) + \varepsilon_v.$$

If v is finite, then $v(T-c) = 0$ for a unique $c \in k$, and there is an embedding of K into a field $k((T_v))$ such that $v(c_a T_v^a + \ldots) = a$ and $T = c + T_v^{\varepsilon_v}$. We have $v(dx/dT_v) \geqslant v(x) - 1$ and $v(dT/dT_v) = \varepsilon_v - 1$, whence $v(x') \geqslant v(x) - \varepsilon_v$. But if $v(x) \geqslant 0$ then $v(dx/dT_v) \geqslant 0$, and $v(x') \geqslant 1 - \varepsilon_v$.

COROLLARY B.1. *If $\mathbf{x} = (x, y) \in K^2$, then $v(x'y - y'x) \geqslant 2v(\mathbf{x}) - v(\mathfrak{D})$.*

PROOF. Note that $v(z) \geqslant 0$ implies $v(z') \geqslant -v(\mathfrak{D})$. So if, say, $v(x) \geqslant v(y)$, then $v(x'y - y'x) = v(y^2(x/y)') \geqslant v(y^2) - v(\mathfrak{D}) = 2v(\mathbf{x}) - v(\mathfrak{D})$.

An element $x \in K$ will be called *integral* (over $k[T]$) if $v(x) \geqslant 0$ for every finite valuation. Given a polynomial $f(X)$ with integral coefficients put

$$H_\infty(f) = -\sum_{v \text{ infin}} v(f);$$

then $H_\infty(f) \geqslant H(f) \geqslant 0$. If $E(Z, Z')$ is a "differential polynomial", define $v(E)$ as with every other polynomial, i.e. by $v(E) = v(\mathbf{x})$ where the components of $\mathbf{x}$ are the coefficients of E, and define the height $H(E)$ in the obvious way.

https://doi.org/10.1017/S1446788700021406 Published online by Cambridge University Press

THEOREM 5. *Suppose* $f(X) = p_0 X^d + \ldots + p_d$ *is a polynomial whose coefficients lie in* K *and are integral, and which has no multiple factors. Suppose* $t > \frac{1}{2}d$. *Then there is a non-zero differential polynomial*

$$E(Z,Z') = (m_0 Z^{t-2} + \ldots + m_{t-2}) Z' - (n_0 Z^t + \ldots + n_t)$$

with coefficients m_i, n_i *in* K *and height*

$$H(E) \leqslant (2t-d)^{-1}((d+t-2)H(f) + dH_\infty(f) + dd(\mathfrak{D}) + 2tg)$$

such that every root α *of* f *(in some extension of* K*) satisfies the differential equation*

$$E(\alpha, \alpha') = 0. \tag{5.1}$$

(Note that in the above inequality, $dd(\mathfrak{D})$ is equal to d, the degree of f, times $d(\mathfrak{D})$, the degree of the different $\mathfrak{D}$.)

PROOF. Let $f(X) = f_1(X) f_2(X) \ldots f_s(X)$ be the factorization of f over K, and let α_i be a root of f_i $(i = 1, \ldots, s)$. It will be enough if (5.1) is satisfied by $\alpha_1, \ldots, \alpha_s$. The root α_i lies in a field L_i of degree $\Delta(i)$ over K, where $\Delta(i)$ is the degree of f_i.

The polynomial $f(X)$ is a function $f(X, T)$ of X and of T; denote its partial derivatives by f_X, f_T. In view of $f(\alpha_i) = f(\alpha_i, T) = 0$ we have $f_X(\alpha_i)\alpha_i' + f_T(\alpha_i) = 0$, and here $f_X(\alpha_i) \neq 0$ since α_i is not a double root of f. The desired equation $E(\alpha_i, \alpha_i') = 0$ may therefore be written in the form

$$\alpha_i^{t-2} f_T(\alpha_i) m_0 + \ldots + f_T(\alpha_i) m_{t-2} + \alpha_i^t f_X(\alpha_i) n_0 + \ldots + f_X(\alpha_i) n_t = 0. \tag{5.2}$$

This is a linear equation in the $n = 2t$ unknowns $m_0, \ldots, m_{t-2}, n_0, \ldots, n_t$. It is of the type (4.1) if $\mathbf{y}_i$ is the vector whose components are

$$\alpha_i^{t-2} f_T(\alpha_i), \ldots, f_T(\alpha_i), \alpha_i^t f_X(\alpha_i), \ldots, f_X(\alpha_i). \tag{5.3}$$

We need

LEMMA C. *The height* $H_i(\mathbf{y}_i)$ *of* $\mathbf{y}_i \in L_i^n$ *satisfies*

$$H_i(\mathbf{y}_i) \leqslant (d+t-2) H(f_i) + \Delta(i) H_\infty(f) + \Delta(i) d(\mathfrak{D}). \tag{5.4}$$

PROOF. Observe that α_i occurs at most to the exponent d in $f_T(\alpha_i)$ and at most $d-1$ in $f_X(\alpha_i)$, so that α_i occurs at most to the exponent $d+t-1$ in the components of $\mathbf{y}_i$. But the only term where the exponent $d+t-1$ occurs is $dp_0 \alpha_i^{d+t-1}$ in $\alpha_i^t f_X(\alpha_i)$. Thus if V_i is any valuation of L_i/k, then

$$\begin{aligned} V_i(\mathbf{y}_i) \geqslant \min(&V_i(f_T), V_i(f_T) + (d+t-2) V_i(\alpha_i), \\ &V_i(f_X), V_i(f_X) + (d+t-2) V_i(\alpha_i), V_i(p_0 \alpha_i^{d+t-1})). \end{aligned}$$

https://doi.org/10.1017/S1446788700021406 Published online by Cambridge University Press

Observe that $V_i(p_0\alpha_i) \geqslant V_i(f)$ by (2.4). If V_i is infinite (that is, if it is an extension of u_∞), $V_i(f_T) \geqslant V_i(f)$ by (a) of Lemma B, since V_i is an extension of an infinite valuation of K. We always have $V_i(f_X) \geqslant V_i(f)$, so that for V_i infinite,

$$V_i(\mathbf{y}_i) \geqslant V_i(f) + (d+t-2)\min(0, V_i(\alpha_i)).$$

If V_i is finite, it is an extension of a finite valuation v of K/k, and we have $v(f) \geqslant 0$, $v(f_X) \geqslant 0$ and $v(f_T) \geqslant 1-\varepsilon_v$ by (c) of Lemma B, so that $V_i(f)$, $V_i(f_X) \geqslant 0$ and $V_i(f_T) \geqslant (1-\varepsilon_v)e_{V_i}$. (Here V_i (like any valuation with cap V) has value group $\mathbf{Z}$, so that $V_i(x) = e_{V_i}v(x)$ for $x \in K$ where e_{V_i} is the ramification index.) Therefore if V_i is finite,

$$V_i(\mathbf{y}_i) \geqslant (d+t-2)\min(0, V_i(\alpha_i)) + (1-\varepsilon_v)e_{V_i}.$$

Now

$$\sum_{v \text{ finite}} \sum_{V_i|v} (1-\varepsilon_v)e_{V_i} = \Delta(i) \sum_{v \text{ finite}} (1-\varepsilon_v) = -\Delta(i)\,d(\mathfrak{D}),$$

and

$$\sum_{v} \sum_{V_i|v} \min(0, V_i(\alpha_i)) = \sum_{v} (-v(p_{i0}) + v(f_i)) = -H(f_i)$$

by (2.12), where p_{i0} is the leading coefficient of f_i. Finally

$$\sum_{v \text{ infin}} \sum_{V_1|v} V_i(f) = \Delta(i) \sum_{v \text{ infin}} v(f) = -\Delta(i)\,H_\infty(f).$$

Combining our estimates we obtain the desired (5.4).

Gauss' Lemma together with $\Delta(1)+\ldots+\Delta(s) = d$ yields

$$H_1(\mathbf{y}_1)+\ldots+H_s(\mathbf{y}_s) \leqslant (d+t-2)\,H(f) + dH_\infty(f) + d\cdot d(\mathfrak{D}).$$

Theorem 5 is now an immediate consequence of Theorem 4.

Corollary 5.1. *Whether $f(X)$ in Theorem 5 has integral coefficients or not, we can always find a differential polynomial $E(Z, Z')$ with the desired properties and with*

$$H(E) \leqslant (2t-d)^{-1}((2d+t-2)\,H(f) + d(\mathfrak{D}) + (d+2t)g). \tag{5.5}$$

Proof. Let v_∞ be an arbitrary infinite valuation of K/k. By Riemann's Theorem there is a $\lambda \neq 0$ in K having

$$v(\lambda) \geqslant -v(f) \quad \text{if } v \neq v_\infty,$$

$$v_\infty(\lambda) \geqslant \sum_{v \neq v_\infty} v(f) - g.$$

Putting $f_1 = \lambda f$ we have $v(f_1) \geqslant 0$ if v is finite, so that the coefficients of f_1 are integral.

https://doi.org/10.1017/S1446788700021406 Published online by Cambridge University Press

Moreover, $H(f_1) = H(f)$ and

$$H_\infty(f_1) = -\sum_{v \text{ infin}} (v(\lambda)+v(f)) \leqslant -v_\infty(\lambda) - v_\infty(f)$$
$$\leqslant -\sum_v v(f) + g = H(f)+g.$$

It will suffice to apply Theorem 5 to f_1.

Up to now T was fairly arbitrary. We now have

LEMMA D. *For a suitable choice of T,*

$$d(\mathfrak{D}) \leqslant 3g. \tag{5.6}$$

PROOF. Pick an arbitrary prime divisor and denote it by $\mathfrak{P}_\infty$. By Riemann's Theorem the elements $T \in K$ having

$$v_\infty(T) \geqslant -g-1,$$
$$v(T) \geqslant 0 \quad \text{if } v \neq v_\infty$$

form a vector space of dimension $\geqslant 2$, which therefore contains an element $T \notin k$. The divisor of T is

$$(T) = \mathfrak{M}/\mathfrak{P}_\infty^n$$

where $n \leqslant g+1$ and $\mathfrak{M}$ is integral. The degree $[K : k(T)] = n$ (Deuring (1972), § 11). Now (see Eichler (1963), p. 150)

$$g = \tfrac{1}{2}d(\mathfrak{D}') - n + 1,$$

where $\mathfrak{D}' = \prod_v \mathfrak{P}_v^{e_v - 1}$ is the "pseudo-different" of K over $k(T)$. Thus

$$d(\mathfrak{D}') = d(\mathfrak{D}) + \sum_{v \text{ infin}} (e_v - 1) = d(\mathfrak{D}) + n - 1,$$

since only v_∞ is infinite and $e_{v_\infty} = n$. We get $g = \frac{1}{2}(d(\mathfrak{D})+n-1)-n+1$, and therefore $d(\mathfrak{D}) = 2g+n-1 \leqslant 3g$.

6. Solutions of differential equations

Let $k \subseteq k(T) \subseteq K$ be as in the last section; again let $\mathfrak{D}$ be the different of K over $k[T]$, and $d(\mathfrak{D})$ its degree.

THEOREM 5. *Suppose $\alpha \in K$ is a solution of a differential equation*

$$E(\alpha, \alpha') = M(\alpha)\alpha' - N(\alpha) = 0, \tag{6.1}$$

where $M(X), N(X)$ are polynomials with coefficients in K and without common factor of positive degree. Suppose the equation is not linear or Ricatti, that is, not with

https://doi.org/10.1017/S1446788700021406 Published online by Cambridge University Press

$\deg M = 0$ *and* $\deg N \leqslant 2$. *Then*

$$\mathrm{H}(\alpha) \leqslant 8H(E) + d(\mathfrak{D}).$$

Proof. Write $M = m_0 X^{\mu} + \ldots + m_{\mu}$, $N = n_0 X^{\nu} + \ldots + n_{\nu}$ with $m_0, n_0 \neq 0$. Fix a valuation v of K/k at the moment, and suppose that

$$v(M(\alpha)) \geqslant v(E). \tag{6.2}$$

Since α satisfies the equation

$$m_0 \alpha^{\mu} + \ldots + m_{\mu-1} \alpha + (m_{\mu} - M(\alpha)) = 0$$

whose coefficients have valuation $\geqslant v(E)$, it follows from (2.4) that

$$v(\alpha) \geqslant v(E) - v(m_0).$$

If v is infinite (as defined in Section 5), we get $v(\alpha') > v(E) - v(m_0) \geqslant 2(v(E) - v(m_0))$ by part (a) of Lemma B. If v is finite, we obtain $v(\alpha') \geqslant v(E) - v(m_0) - \varepsilon_v$ by part (b), and if $v(E) < v(m_0)$ then this is $\geqslant 2(v(E) - v(m_0)) + 1 - \varepsilon_v$. If $v(E) = v(m_0)$, then $v(\alpha') \geqslant 1 - \varepsilon_v = 2(v(E) - v(m_0)) + 1 - \varepsilon_v$ by part (c) of Lemma B. Hence in all cases

$$v(\alpha') \geqslant 2(v(E) - v(m_0)) - v(\mathfrak{D}).$$

(Observe that $v(E) \leqslant v(m_0)$ by definition of $v(E)$.) The differential equation (6.1) implies that

$$v(N(\alpha)) \geqslant v(M(\alpha)) + 2(v(E) - v(m_0)) - v(\mathfrak{D}). \tag{6.3}$$

The resultant R of $M(X), N(X)$ may be written as

$$R = M(X)\, V(X) + N(X)\, W(X), \tag{6.4}$$

where $V(X), W(X)$ are certain polynomials defined in terms of determinants. In particular, V, W are of respective degrees $\leqslant \nu - 1$, $\mu - 1$, and

$$v(V), v(W) \geqslant (\mu + \nu - 1)\, v(E).$$

Now since $v(m_0 \alpha) \geqslant v(E)$, it follows that

$$v(m_0^{\nu-1} V(\alpha)) \geqslant (\mu + \nu - 1 + \nu - 1)\, v(E), \qquad v(m_0^{\mu-1} W(\alpha)) \geqslant (\mu + \nu - 1 + \mu - 1)\, v(E).$$

Thus with

$$\omega = \max(\nu - 1, \mu + 1),$$

(6.3), (6.4) yield

$$v(m_0^{\omega} R) \geqslant v(M(\alpha)) + (\mu + \nu - 1 + \omega)\, v(E) - v(\mathfrak{D}),$$

whence

$$v(M(\alpha)) \leqslant -(\mu + \nu - 1)\, v(E) + \omega(v(m_0) - v(E)) + v(R) + v(\mathfrak{D}). \tag{6.5}$$

https://doi.org/10.1017/S1446788700021406 Published online by Cambridge University Press

Since $v(m_0)\geqslant v(E)$, since $v(R)\geqslant(\mu+\nu)\,v(E)$ and since $v(\mathfrak{D})\geqslant 0$, (6.5) is always true, irrespective of (6.2).

Now either

$$v(m_0\,\alpha^{\mu})<\min\left(v(m_1\,\alpha^{\mu-1}),\ldots,v(m_{\mu})\right). \tag{6.6}$$

Then $v(m_0\,\alpha^{\mu})=v(M(\alpha))$ and $\mu v(\alpha)=v(M(\alpha))-v(m_0)$. Or if (6.6) is not true, then $v(m_0\,\alpha^{\mu})\geqslant v(m_i\,\alpha^{\mu-i})$ for some i $(1\leqslant i\leqslant\mu)$. Then $v(\alpha^i)\geqslant v(m_i)-v(m_0)\geqslant v(E)-v(m_0)$, whence $v(\alpha)\geqslant v(E)-v(m_0)$ and $\mu v(\alpha)\geqslant\mu(v(E)-v(m_0))$. So whether (6.6) is true or not,

$$\mu\min\left(0,v(\alpha)\right)\geqslant\min\left(\mu(v(E)-v(m_0)),\;v(M(\alpha))-v(m_0)\right).$$

Using the sum formula for $m_0^{\mu}\,M(\alpha)^{-1}$ we obtain

$$\begin{aligned}\mu\mathrm{H}(\alpha)&=-\mu\sum_{v}\min\left(0,v(\alpha)\right)\\&\leqslant-\sum_{v}\min\left(\mu v(E)-v(M(\alpha)),\;(\mu-1)\,v(m_0)\right).\end{aligned}$$

Involving (6.5) and once again the sum formula, we get

$$\begin{aligned}\mu\mathrm{H}(\alpha)&\leqslant\sum_{v}\max\left(-(2\mu+v-1)\,v(E)+\omega(v(m_0)-v(E))+v(R)+v(\mathfrak{D}),(1-\mu)\,v(m_0)\right)\\&=\sum_{v}\max\left(-(2\mu+\nu-1+\omega)\,v(E)+v(\mathfrak{D}),\;-v(R)+(1-\mu-\omega)\,v(m_0)\right).\end{aligned}$$

Recall that $\sum_v v(\mathfrak{D})=d(\mathfrak{D})$, and note that

$$\begin{aligned}-v(R)+(1-\mu-\omega)\,v(m_0)&\leqslant-(\mu+\nu)\,v(E)+(1-\mu-\omega)\,v(E)\\&=(1-2\mu-\nu-\omega)\,v(E).\end{aligned}$$

It then follows that

$$\begin{aligned}\mu\mathrm{H}(\alpha)&\leqslant(1-2\mu-\nu-\omega)\sum_{v}v(E)+d(\mathfrak{D})\\&=(2\mu+\nu+\omega-1)\,H(E)+d(\mathfrak{D}),\end{aligned}$$

and if $\mu>0$, then

$$\mathrm{H}(\alpha)\leqslant\mu^{-1}(2\mu+\nu+\omega-1)\,H(E)+d(\mathfrak{D}).$$

Now suppose that

$$\nu\leqslant2\mu+2.$$

By our assumption that the differential equation be neither linear nor Ricatti we get $\mu>0$. Further

$$\begin{aligned}2\mu+\nu+\omega-1&\leqslant2\mu+(2\mu+2)-1+\max\left(2\mu+1,\mu+1\right)\\&=6\mu+2\leqslant8\mu,\end{aligned}$$

https://doi.org/10.1017/S1446788700021406 Published online by Cambridge University Press

and

$$\mathrm{H}(\alpha) \leqslant 8H(E) + d(\mathfrak{D}),$$

as desired.

If $\nu > 2\mu + 2$, we use the method of Osgood (1975) and Schmidt (1976). We set $\hat{\alpha} = 1/\alpha$ and we obtain a differential equation $\hat{M}(\hat{\alpha})\hat{\alpha}' = \hat{N}(\hat{\alpha})$ for $\hat{\alpha}$, of the same type and same height: $H(\hat{E}) = H(E)$. But now the degrees $\hat{\mu}, \hat{\nu}$ have $\hat{\nu} \leqslant 2\hat{\mu} + 2$. We obtain $\mathrm{H}(\alpha) = \mathrm{H}(\hat{\alpha}) \leqslant 8H(E) + d(\mathfrak{D})$.

7. Solutions of the Thue Equation

We are interested in solutions $x, y \in K$ of

$$f(x,y) = 1,$$

where f is a form of degree $d \geqslant 3$, without multiple factors and with coefficients in K. We further suppose that $f(X, Y)$ is not divisible by X or by Y; this can always be achieved by a linear change of variables with coefficients in k, and such a change of variables will not affect heights. Then

$$f(X, Y) = f_0(X - \alpha_1 Y) \dots (X - \alpha_d Y) = g_0(\beta_1 X - Y) \dots (\beta_d X - Y)$$

with $\alpha_i \beta_i = 1$ $(i = 1, \dots, d)$.

Again let $T \in K$ be transcendental over k. Assuming $t > \frac{1}{2}d$ we can apply Corollary 5.1 to the polynomial $f(X, 1)$ to construct a certain differential polynomial

$$E(Z, Z') = M(Z)Z' - N(Z) = (m_0 Z^{t-2} + \dots + m_{t-2})Z' - (n_0 Z^t + \dots + n_t)$$

having $E(\alpha_i, \alpha_i') = 0$ $(i = 1, \dots, d)$. Putting

$$\begin{aligned} G(Z, Z', W, W') &= W^t E(Z/W, (Z/W)') \\ &= (m_0 Z^{t-2} + \dots + m_{t-2} W^{t-2})(Z'W - W'Z) - (n_0 Z^t + \dots + n_t W^t) \end{aligned}$$

we have $G(\alpha_i, \alpha_i', 1, 0) = 0$ and $G(1, 0, \beta_i, \beta_i') = 0$ $(i = 1, \dots, d)$.

Lemma E. *Suppose $x, y \in K$ with*

$$F(x,y) = 1 \quad \text{and} \quad G(x, x', y, y') \neq 0. \tag{7.1}$$

(i) *Then if $t < d - 1$, the point $\mathbf{x} = (x, y)$ has height*

$$\mathrm{H}(\mathbf{x}) \leqslant (d - t - 1)^{-1}(H(E) + (2d-2)\mathrm{H}(f) + d(\mathfrak{D})). \tag{7.2}$$

(ii) *If $t < d$ and if $\mathbf{x}$ is $\mathfrak{S}$-integral for a finite set $\mathfrak{S}$ of valuations, then*

$$\mathrm{H}(\mathbf{x}) \leqslant (d - t)^{-1}(H(E) + (2d-2)\mathrm{H}(f) + d(\mathfrak{D}) + |\mathfrak{S}|). \tag{7.3}$$

Proof. Write $\psi(v) = (d-2)v(f) - \frac{1}{2}v(D)$ where D is the discriminant of $f(X, 1)$, and also of $f(1, Y)$. We begin with the observation that by the sum formula and by (2.6),

$$\sum_{v} \min(0, \psi(v)) = \sum_{v} \min(\tfrac{1}{2}v(D), (d-2)\,v(f))$$

(7.4)
$$\geqslant \sum_{v} \min((d-1)\,v(f), (d-2)\,v(f))$$
$$\geqslant \sum_{v} (d-1)\,\mathbf{v}(f)$$
$$= -(d-1)\,\mathrm{H}(f).$$

Now let $\mathfrak{T}$ be the set of valuations having

$$dv(\mathbf{x}) \geqslant \psi(v),$$

and $\mathfrak{T}'$ the set with $dv(\mathbf{x}) < \psi(v)$. Then

(7.5)
$$\sum_{v \in \mathfrak{T}} \mathbf{v}(\mathbf{x}) = \sum_{v \in \mathfrak{T}} \min(0, v(\mathbf{x})) \geqslant \sum_{v \in \mathfrak{T}} \min(0, d^{-1}\psi(v)).$$

Next, let $v \in \mathfrak{T}'$. Suppose that

(7.6)
$$v(x) \geqslant v(y),$$

say. Let v be extended in some way to a valuation of the field $K(\alpha_1, \ldots, \alpha_d)$, and suppose without loss of generality that

(7.7)
$$v(x-\alpha_1 y) \geqslant \ldots \geqslant v(x-\alpha_d y).$$

Then

$$v((\alpha_1-\alpha_i)\,y) = v(x-\alpha_i y-(x-\alpha_1 y)) \geqslant v(x-\alpha_i y)$$

and therefore

$$0 = v(1) = v(f(x,y)) = v(f_0)+v(x-\alpha_1 y)+\ldots+v(x-\alpha_d y)$$
$$\leqslant v(f_0(\alpha_1-\alpha_2)\ldots(\alpha_1-\alpha_d))+v(y^d)+v(\alpha_1-(x/y)),$$

which by (2.9) is

$$\leqslant \tfrac{1}{2}v(D)-(d-2)\,v(f)+v(y^d)+v(\alpha_1-(x/y)).$$

We obtain

(7.8)
$$v(\alpha_1-(x/y)) \geqslant (d-2)\,v(f)-\tfrac{1}{2}v(D)-dv(y),$$

whence

$$v(\alpha_1-(x/y)) \geqslant \psi(v)-dv(\mathbf{x}) > 0$$

by (7.6) and since $v \in \mathfrak{T}'$. Using (7.6) again we find that $v(x/y) \geqslant 0$, $v(\alpha_1) \geqslant 0$. Taking the derivative we obtain

$$v(\alpha_1'-(x/y)') \geqslant \psi(v)-dv(\mathbf{x})-1-v(\mathfrak{D})$$

https://doi.org/10.1017/S1446788700021406 Published online by Cambridge University Press

by Lemma B, and therefore

$$v\left(E\left(\frac{x}{y'}\left(\frac{x}{y}\right)'\right)\right) = v\left(E\left(\frac{x}{y'}\left(\frac{x}{y}\right)'\right) - E(\alpha_1, \alpha_1')\right)$$

$$\geqslant v(E) + \psi(v) - dv(\mathbf{x}) - 1 - v(\mathfrak{D}),$$

which in turn yields

$$v(G(x, x', y, y')) \geqslant v(E) + \psi(v) - (d-t)\, v(\mathbf{x}) - 1 - v(\mathfrak{D}). \tag{7.9}$$

This holds for $v \in \mathfrak{T}'$. On the other hand, for every v,

$$v(G(x, x', y, y')) \geqslant v(E) + tv(\mathbf{x}) - v(\mathfrak{D}) \tag{7.10}$$

by Corollary B.1. For $v \in \mathfrak{T}'$ we combine (7.9) and (7.10) to obtain

$$\min\left((d-t)\, v(\mathbf{x}) + 1, -tv(\mathbf{x})\right) \tag{7.11}$$

$$\geqslant -v(G(x, x', y, y')) + v(E) - v(\mathfrak{D}) + \min(0, \psi(v)).$$

(i) Now if $d > t+1$, we observe that

$$\min((d-t-1)\, v, 0) \geqslant \min((d-t)\, v + 1, -tv)$$

for every integer v, so that (7.11) yields

$$(d-t-1) \sum_{v \in \mathfrak{T}'} \mathbf{v}(\mathbf{x}) \geqslant -\sum_{v \in \mathfrak{T}'} v(G(x, x', y, y')) + \sum_{v \in \mathfrak{T}'} v(E) - \sum_{v \in \mathfrak{T}'} v(\mathfrak{D}) + \sum_{v \in \mathfrak{T}'} \min(0, \psi(v)).$$

Using (7.5) and the sum formula for $G(x, x', y, y')$ we have

$$(d-t-1) \sum_{v} \mathbf{v}(\mathbf{x}) \geqslant \sum_{v \in \mathfrak{T}} v(G(x, x', y, y')) + \sum_{v \in \mathfrak{T}'} v(E) - \sum_{v \in \mathfrak{T}'} v(\mathfrak{D}) + \sum_{v} \min(0, \psi(v)).$$

Utilizing (7.10), (7.5) and (7.4) we obtain

$$\geqslant \sum_{v} v(E) + t \sum_{v \in \mathfrak{T}} v(\mathbf{x}) - \sum_{v} v(\mathfrak{D}) + \sum_{v} \min(0, \psi(v))$$

$$\geqslant -H(E) - d(\mathfrak{D}) + 2 \sum_{v} \min(0, \psi(v))$$

$$\geqslant -H(E) - (2d-2)\, \mathrm{H}(f) - d(\mathfrak{D}),$$

and (7.2) follows.

(ii) On the other hand if $d > t$, we observe that

$$\min((d-t)\, v, 0) \geqslant \min((d-t)\, v + 1, -tv) - \delta_v$$

for all integers v, where $\delta_v = 1$ if $v < 0$ and $\delta_v = 0$ otherwise. For $\mathfrak{S}$-integral $\mathbf{x}$ the number of v with $v(\mathbf{x}) < 0$ is at most $|\mathfrak{S}|$, so that (7.11) yields

$$(d-t) \sum_{v \in \mathfrak{T}'} \mathbf{v}(\mathbf{x}) \geqslant -\sum_{v \in \mathfrak{T}'} G(x, x', y, y') + \sum_{v \in \mathfrak{T}'} v(E) - \sum_{v \in \mathfrak{T}'} v(\mathfrak{D}) + \sum_{v \in \mathfrak{T}'} \min(0, \psi(v)) - |\mathfrak{S}|.$$

Continuing as in the case (i) we get

$$(d-t)\sum_v \mathbf{v}(\mathbf{x}) \geqslant -H(E)-(2d-2)\mathrm{H}(f)-d(\mathfrak{D})-|\mathfrak{S}|$$

and (7.3).

8. Proof of the theorem on Thue equations

Pick $T \in K$ according to Lemma D such that $d(\mathfrak{D}) \leqslant 3g$. Set

$$t = (2d+\varepsilon)/3$$

where $\varepsilon = 0$ if $d \equiv 0 \pmod 3$, $\varepsilon = 1$ if $d \equiv 1 \pmod 3$ and $\varepsilon = -1$ if $d \equiv 2 \pmod 3$. Then t is integral with $d/2 < t < d$ for $d \geqslant 3$ and $d/2 < t < d-1$ for $d \geqslant 5$. Construct a differential polynomial $E(Z, Z')$ according to Corollary 5.1, so that by (5.5),

$$H(E) \leqslant (d+2\varepsilon)^{-1}((8d+\varepsilon-6)H(f)+(16d+2\varepsilon)g).$$

If, say, $d \equiv 2 \pmod 3$, we have

$$H(E) \leqslant (d-2)^{-1}((8d-7)H(f)+(16d-2)g).$$

The right-hand side is a decreasing function of d, and substituting $d = 5$ we get

(8.1) $$H(E) \leqslant 11H(f)+26g.$$

This estimate is also valid if $d \equiv 0$ or $d \equiv 1 \pmod 3$. Define $G(Z, Z', W, W')$ as in Section 7.

It will suffice to consider solutions of the Thue equation $f(x, y) = 1$ with $y \neq 0$. At first we shall assume that

(8.2) $$E(x/y, (x/y)') \neq 0,$$

so that $G(x, x', y, y') \neq 0$. If $d \geqslant 5$, part (i) of Lemma E yields

$$\mathrm{H}(\mathbf{x}) \leqslant H(E)+3g+(d-\varepsilon-3)^{-1}(6d-6)\mathrm{H}(f).$$

It is easily seen that for $d \geqslant 5$ the factor $(d-\varepsilon-3)^{-1}(6d-6)$ is at most 12, so that by (8.1),

(8.3) $$\mathrm{H}(\mathbf{x}) \leqslant 23\mathrm{H}(f)+29g.$$

If $d \geqslant 3$ and $\mathbf{x}$ is $\mathfrak{S}$-integral, part (ii) of Lemma E yields

$$\mathrm{H}(\mathbf{x}) \leqslant H(E)+3g+|\mathfrak{S}|+(d-\varepsilon)^{-1}(6d-6)\mathrm{H}(f).$$

The factor in front of $\mathrm{H}(f)$ is at most 6, and an appeal to (8.1) gives

(8.4) $$\mathrm{H}(\mathbf{x}) \leqslant 17\mathrm{H}(f)+29g+|\mathfrak{S}|.$$

Next, assume that

$$E(x/y,(x/y)')=0. \tag{8.5}$$

We may write $E(Z,Z')=g(Z)\,E_1(Z,Z')$, where

$$E_1(Z)=M(Z)Z'-N(Z)$$

and where $M(Z),N(Z),g(Z)$ are polynomials such that $M(Z),N(Z)$ have no common factor. Clearly $H(E)=H(E_1)+H(g)$, and therefore $H(E_1),H(g)\leqslant H(E)$. Now if $g(x/y)=0$, then

$$\mathbf{H}(x/y)\leqslant H(g)\leqslant H(E) \tag{8.6}$$

by (2.3). The other possibility is that

$$E_1(x/y,(x/y)')=0. \tag{8.7}$$

If this differential equation in x/y is not linear or Ricatti, then by Theorem 5,

$$\mathbf{H}(x/y)\leqslant 8H(E)+3g,$$

which then in view of (8.6) holds whenever (8.5) is true. Now $f(x,y)=1$ yields $0=v(f(x,y))\geqslant v(f)+dv(\mathbf{x})$, or $-v(\mathbf{x})\geqslant d^{-1}v(f)$. Further since $H(\mathbf{x})=\mathbf{H}(x/y)$,

$$\begin{aligned}\mathbf{H}(\mathbf{x}) &= -\sum_v \min(0,v(\mathbf{x})) = -\sum_v v(\mathbf{x})-\sum_v \min(0,-v(\mathbf{x}))\\ &\leqslant H(\mathbf{x})-d^{-1}\sum_v \min(0,v(f))\\ &\leqslant 8H(E)+3g+\mathbf{H}(f),\end{aligned}$$

and substituting (8.1) we obtain

$$\mathbf{H}(\mathbf{x})\leqslant 89\mathbf{H}(f)+211g. \tag{8.8}$$

There remains the case when the differential equation (8.7) is linear or Ricatti. Let v be a valuation of K/k, extended in some way to a valuation of $K(\alpha_1,\ldots,\alpha_d)$, and suppose that (7.6) and (7.7) hold, so that (7.8) is true. If the differential equation is linear, then any three solutions z,z_1,z_2 have

$$\frac{z-z_1}{z_2-z_1}=c$$

with $c\in k$, whence $v(z-z_1)=v(z_2-z_1)$. In particular, with $z=x/y$, $z_1=\alpha_1$, $z_2=\alpha_2$, we have $v((x/y-\alpha_1)=v(\alpha_2-\alpha_1)$, hence by (2.7)

$$v((x/y)-\alpha_1)\leqslant \tfrac{1}{2}v(D)-(d-1)\,v(f). \tag{8.9}$$

https://doi.org/10.1017/S1446788700021406 Published online by Cambridge University Press

If the differential equation is Ricatti, then any four solutions z, z_1, z_2, z_3 have

$$\frac{z-z_1}{z-z_3}\Big/\frac{z_2-z_1}{z_2-z_3} = c$$

with $c \in k$, whence $v(z-z_1)+v(z_2-z_3) = v(z-z_3)+v(z_2-z_1)$. In particular,

$$v((x/y)-\alpha_1)+v(\alpha_2-\alpha_3) = v((x/y)-\alpha_3)+v(\alpha_2-\alpha_1).$$

If $v((x/y)-\alpha_1) \leqslant v(\alpha_3-\alpha_1)$, then again (8.9) by (2.7). Otherwise

$$v((x/y)-\alpha_1) > v(\alpha_3-\alpha_1) \quad \text{and} \quad v((x/y)-\alpha_3) = v(\alpha_3-\alpha_1),$$

so that $v((x/y)-\alpha_1) = v(\alpha_3-\alpha_1)+v(\alpha_2-\alpha_1)-v(\alpha_2-\alpha_3)$, and (2.8) yields (8.9) again. Combining (8.9) with (7.8) we obtain

$$dv(y) \geqslant (2d-3)\,v(f)-v(D).$$

This is true under the sole condition (7.6), that is $v(x) \geqslant v(y)$. A similar estimate holds under the condition that $v(x) \leqslant v(y)$. Therefore

$$d\min(0, v(x), v(y)) \geqslant \min(0, (2d-3)\,v(f)-v(D)),$$

and using the sum formula for D we obtain

$$\begin{aligned} d\mathbf{H}(\mathbf{x}) &= -d\sum_v \mathbf{v}(\mathbf{x}) \\ &\leqslant -\sum_v \min(v(D), (2d-3)\,v(f)) \\ &\leqslant -\sum_v \min((2d-2)\,v(f), (2d-3)\,v(f)) \quad \text{(by (2.6))} \\ &\leqslant -(2d-2)\sum_v \mathbf{v}(f) \\ &= (2d-2)\,\mathbf{H}(f). \end{aligned}$$

After division by d we have

$$\mathbf{H}(\mathbf{x}) \leqslant 2\mathbf{H}(f). \tag{8.10}$$

If $d \geqslant 5$ we may combine (8.3), (8.8) and (8.10) to obtain part (i) of Theorem 1. If $d \geqslant 3$ and $\mathbf{x}$ is $\mathfrak{S}$-integral we combine (8.4), (8.8) and (8.10) to obtain

$$\mathbf{H}(\mathbf{x}) \leqslant 89\mathbf{H}(f)+211g+|\mathfrak{S}|. \tag{8.11}$$

The desired part (ii) of Theorem 1 follows immediately if $g > 0$. If $g = 0$, then $K = k(T)$ for some T. We may suppose that $v_0 = -\deg$ lies in $\mathfrak{S}$. For if $c \in k$ and the valuation v with $v(T-c) = 1$ lies in $\mathfrak{S}$, replace T by $T_1 = (T-c)^{-1}$. The extension field $L = K(T^{1/n}) = k(T^{1/n})$ has again genus 0, but $\mathbf{H}_L(\mathbf{x}) = n\mathbf{H}(\mathbf{x})$ and

$\mathrm{H}_L(f) = n\mathrm{H}(f)$. Since v_0 is completely ramified in L, the set $\mathfrak{S}_L$ of extensions of elements of $\mathfrak{S}$ to L has cardinality $|\mathfrak{S}_L| \leqslant n(|\mathfrak{S}|-1)+1$. Thus by (8.11), applied in L,

$$n\mathrm{H}(\mathbf{x}) = \mathrm{H}_L(\mathbf{x}) \leqslant 89\mathrm{H}_L(f) + |\mathfrak{S}_L| \leqslant 89n\mathrm{H}(f) + n(|\mathfrak{S}|-1)+1.$$

Since n is arbitrary, $\mathrm{H}(\mathbf{x}) \leqslant 89\mathrm{H}(f) + |\mathfrak{S}|-1$.

9. Some genus estimates

Lemma F. *Let K/k be a function field of genus g. Let z be in K but not a square in K. Then $L = K(\sqrt{z})$ has genus*

$$g_L \leqslant 2g + \mathrm{H}_K(z) - 1.$$

Proof. We have (Eichler (1963), Ch. III.3, formula (10))

$$g_L = g(L/K) + [L:K]g = g(L/K) + 2g$$

where (Eichler (1963), formula (9))

$$g(L/K) = \tfrac{1}{2}g(\mathfrak{D}_{L/K}) - 1$$

and $g(\mathfrak{D}_{L/K})$ is the ramification index of L over K. It is (Eichler (1963), Ch. III.2, formula (36)) equal to $\sum_V (e_V - 1)$ where V runs through the valuations of L/k and e_V is the ramification index of V over K. Suppose that the principal divisor (z) in K is of the form

$$(z) = \mathfrak{P}_1^{a_1} \cdots \mathfrak{P}_r^{a_r} / (\mathfrak{Q}_1^{b_1} \cdots \mathfrak{Q}_t^{b_t})$$

with distinct prime divisors $\mathfrak{P}_1, \ldots, \mathfrak{P}_r$, $\mathfrak{Q}_1, \ldots, \mathfrak{Q}_t$. Then a prime divisor $\mathfrak{P}$ will ramify in the extension field L precisely if $\mathfrak{P}$ is equal to some $\mathfrak{P}_i$ with odd a_i, or to some $\mathfrak{Q}_j$ with odd b_j, and moreover the ramification index is 2 in this case. So

$$g(\mathfrak{D}_{L/K}) \leqslant r + t \leqslant a_1 + \ldots + a_r + b_1 + \ldots + b_t.$$

Now

$$\mathrm{H}(z) = \mathrm{H}_K(z) = -\sum_v \min(0, v(z)) = b_1 + \ldots + b_t$$
$$= \mathrm{H}(1/z) = a_1 + \ldots + a_r.$$

We obtain $g(\mathfrak{D}_{L/K}) \leqslant 2\mathrm{H}(z)$, whence $g(L/K) \leqslant \mathrm{H}(z) - 1$, whence the lemma.

Lemma G. *Let K/k be a function field of genus g, and let z_1, z_2, z_3 be in K with heights $\mathrm{H}(z_i) \leqslant H$ where $H \geqslant 1$. Then $L = K(\sqrt{z_1}, \sqrt{z_2}, \sqrt{z_3})$ has genus*

$$g_L \leqslant \Delta(\tfrac{3}{2}H + g) \tag{9.1}$$

where Δ is the degree $[L:K]$.

https://doi.org/10.1017/S1446788700021406 Published online by Cambridge University Press

PROOF. The field $L_1 = K(\sqrt{z_1})$ has by the preceding lemma genus

$$g_1 \leqslant 2g+H-1.$$

This holds even if z_1 is a square in K, for then $g_1 = g \leqslant 2g \leqslant 2g+H-1$. Now $\mathrm{H}_{L_1}(z_2) \leqslant 2\mathrm{H}_K(z_2) \leqslant 2H$, so that a second application of Lemma F yields $g_2 \leqslant 2g_1+2H-1$ for the genus g_2 of $K(\sqrt{z_1}, \sqrt{z_2})$, and therefore

$$g_2 \leqslant 4g+4H-3.$$

A third application of Lemma F yields

$$g_3 \leqslant 8g+12H-7.$$

If $\Delta = 1$ we have $g_L = g$, if $\Delta = 2$ we have $g_L \leqslant 2g+H-1$, if $\Delta = 4$ we have $g_L \leqslant 4g+4H-3$, and if $\Delta = 8$ then $g_L \leqslant 8g+12H-7$. The estimate (9.1) is true in each case.

LEMMA H. *Let $f(X) = f_0 X^d + \ldots + f_d$ be a polynomial whose coefficients are polynomials in T (they lie in $k[T]$) of degree $\leqslant m$. Let K be the splitting field of f over $k(T)$, and Δ the degree of K over $k(T)$. Then K/k has genus*

$$g \leqslant (\Delta-1)\, dm.$$

In the special case when $f(X) = f_0 X^d + f_d$,

$$g \leqslant (\Delta-1)(m-1).$$

PROOF. We may suppose that the roots of f are distinct. If $c \in k$ and $f_0(c) \neq 0$, $D(c) \neq 0$ where D is the discriminant, then we *claim* that the valuation with $v(T-c) = 1$ is unramified in K. For let $\alpha_1, \ldots, \alpha_d$ be the roots of $f(X) = f(X, T)$. (We recall that f is a polynomial in T also. The splitting field is $K = k(T, \alpha_1, \ldots, \alpha_d)$.) The roots of $f(X, c)$ are d distinct elements $a_1, \ldots, a_d$ in k. Pick $c_1, \ldots, c_d$ in k such that

$$\alpha = c_1 \alpha_1 + \ldots + c_d \alpha_d$$

generates K over $k(T)$ and that the $d!$ numbers

$$a_\sigma = c_1 a_{\sigma(1)} + \ldots + c_d a_{\sigma(d)},$$

where σ runs through the permutations of $1, \ldots, d$, are distinct. The field polynomial of α over $k(T)$ is a divisor of

$$\prod_\sigma (X - \alpha_\sigma)$$

where $\alpha_\sigma = c_1 \alpha_{\sigma(1)} + \ldots + c_d \alpha_{\sigma(d)}$. The image of the latter polynomial in the residue class field associated with v is $\prod_\sigma (X - a_\sigma)$, hence has distinct roots, and

https://doi.org/10.1017/S1446788700021406 Published online by Cambridge University Press

hence also the image of the field polynomial of α has distinct roots. Hence v is indeed unramified.

Now $\deg f_0 + \deg D \leqslant m + (2d-2)m = (2d-1)m$, and being overly generous and allowing for the valuation $v_\infty = -\deg$, we see that at most $(2d-1)m+1$ valuations v of $k(T)$ ramify in K. If v does ramify, the sum of $e_V - 1$ over the extensions V is at most $\Delta - 1$. So in the notation of Eichler ((1963), Ch. III.2 (36)),

$$g(\mathfrak{D}_{K/k(T)}) \leqslant (\Delta - 1)((2d-1)m+1).$$

Again (Eichler (1963), Ch. III.3 (5)),

$$\begin{aligned} g &= \tfrac{1}{2} g(\mathfrak{D}_{K/k(T)}) - \Delta + 1 \\ &\leqslant (\Delta - 1)((d - \tfrac{1}{2})m + \tfrac{1}{2} - 1) \\ &\leqslant (\Delta - 1)dm. \end{aligned}$$

In the more special case there are at most $2m$ valuations v which ramify in K. One obtains $g(\mathfrak{D}) \leqslant \Delta - 1) \cdot 2m$ and therefore

$$g = \tfrac{1}{2} g(\mathfrak{D}) - \Delta + 1 \leqslant (\Delta - 1)(m-1).$$

10. Solutions of the hyperelliptic equation

In the equation

$$y^2 = f(x) \tag{10.1}$$

set $f(X) = f_1(X) g^2(X)$ where $f_1(X)$ is square free and where the leading coefficient of g is 1. Then if v is any valuation, $v(g) \leqslant 0$, whence

$$\mathsf{v}(f) = \min(0, v(f)) = \min(0, v(f_1) + 2v(g)) \leqslant \min(0, v(f_1)) = \mathsf{v}(f_1),$$

and therefore $\mathrm{H}(f_1) \leqslant \mathrm{H}(f)$. Since $f(x)$ is a square in K precisely when $f_1(x)$ is, we may replace f by f_1. Since by hypothesis f contains at least three linear factors which occur to an odd power, $\deg f_1 \geqslant 3$. *Thus we may suppose without loss of generality that f is square free of degree $d \geqslant 3$.* Write

$$f = f_0(X - \alpha_1) \ldots (X - \alpha_d).$$

We remark that if $\mathfrak{T}$ is any set of valuations, then

$$\begin{aligned} -\sum_{v \in \mathfrak{T}} v(f/f_0) &= -\sum_{v \in \mathfrak{T}} v(f) + \sum_{v \in \mathfrak{T}} v(f_0) \\ &= -\sum_{v \in \mathfrak{T}} v(f) - \sum_{v \notin \mathfrak{T}} v(f_0) \leqslant -\sum_{v} \mathsf{v}(f) = \mathrm{H}(f). \end{aligned} \tag{10.2}$$

Given a solution of (10.1) where $x, y \in K$ and x is $\mathfrak{S}$-integral, we define divisors

https://doi.org/10.1017/S1446788700021406 Published online by Cambridge University Press

$\mathfrak{M}_i, \mathfrak{N}_i$ $(i = 1, \ldots, d)$ by the rule that the principal divisor

$$(x-\alpha_i) = \frac{\mathfrak{N}_i}{\mathfrak{M}_i},$$

and that

$$v(\mathfrak{M}_i) = \begin{cases} -v(x-\alpha_i) & \text{if } v(x-\alpha_i)<0 \text{ and } v \notin \mathfrak{S}, \\ 0 & \text{otherwise.} \end{cases}$$

We next define $\mathfrak{A}_i, \mathfrak{B}_i$ by the rule that $\mathfrak{N}_i = \mathfrak{A}_i^2 \mathfrak{B}_i$ and that $v(\mathfrak{B}_i) = 0$ or 1 for every valuation v. In particular this implies that

$$(10.3) \qquad (x-\alpha_i) = \frac{\mathfrak{A}_i^2 \mathfrak{B}_i}{\mathfrak{M}_i} \quad (i = 1, \ldots, d).$$

LEMMA I.

(i) *For* $v \notin \mathfrak{S}$,

$$\sum_{i=1}^{d} v(\mathfrak{M}_i) = -v(f/f_0).$$

(ii) *For* $v \notin \mathfrak{S}$ *and* $v(f) \geqslant 0$,

$$\sum_{i=1}^{d} v(\mathfrak{B}_i) \leqslant v(D)$$

where D *is the discriminant of* f.

PROOF. (i) Since $v(x) \geqslant 0$ for $v \notin \mathfrak{S}$, we have

$$\min(0, v(x-\alpha_i)) = \min(0, v(\alpha_i)),$$

whence by (2.2),

$$\sum_{i=1}^{d} v(\mathfrak{M}_i) = -\sum_{i=1}^{d} \min(0, v(x-\alpha_i)) = -\sum_{i=1}^{d} \min(0, v(\alpha_i)) = -v(f/f_0).$$

(ii) Suppose $\sum_{i=1}^{d} v(\mathfrak{B}_i) = z \geqslant 1$, so that $v(\mathfrak{B}_i) = 1$ for z values of i between 1 and d; say, $v(\mathfrak{B}_i) = z$ for $i \in \mathfrak{J}$ with $|\mathfrak{J}| = z$. Then $v(x-\alpha_i) \geqslant 1$ for $i \in \mathfrak{J}$ and $v(\alpha_i - \alpha_j) \geqslant 1$ for $i, j \in \mathfrak{J}$, so that

$$2 \sum_{\substack{(i,j) \in \mathfrak{J}^2 \\ i<j}} v(\alpha_i - \alpha_j) \geqslant z(z-1).$$

On the other hand by (2.5),

$$2v(f_0^{d-1}) + 2 \sum_{\substack{(i,j) \notin \mathfrak{J}^2 \\ i<j}} v(\alpha_i - \alpha_j) \geqslant (2d-2)\, v(f) \geqslant 0,$$

so that

$$v(D) \geqslant z(z-1)$$

and therefore $v(D) \geqslant z$ if $z \geqslant 2$.

https://doi.org/10.1017/S1446788700021406 Published online by Cambridge University Press

There remains the case when $z = 1$. By (10.3), (2.5), we have

$$v\left(\frac{(f_0)}{\mathfrak{M}_1 \dots \mathfrak{M}_d}\right) = v(f_0 \prod_{v(x-\alpha_i)<0} (x-\alpha_i)) \geqslant v(f) \geqslant 0,$$

and therefore the divisor

$$(f_0)\frac{\mathfrak{B}_1 \dots \mathfrak{B}_d}{\mathfrak{M}_1 \dots \mathfrak{M}_d}$$

has valuation $\geqslant z = 1$. But this divisor equals $(f(x))/(\mathfrak{A} \dots \mathfrak{A}_d)^2$, so is a complete square, and hence has valuation at least 2. Since $v(\mathfrak{B}_1 \dots \mathfrak{B}_d) = z = 1$, we find that

$$v(f_0) > \sum_{i=1}^{d} v(\mathfrak{M}_i).$$

Suppose that, say, $v(\alpha_1) \leqslant \dots \leqslant v(\alpha_a) < 0 \leqslant v(\alpha_{a+1}) \leqslant \dots \leqslant v(\alpha_d)$, where possibly $a = 0$. Then

$$\sum_{i=1}^{d} v(\mathfrak{M}_i) = -v(\alpha_1) - \dots - v(\alpha_\alpha)$$

and

$$\begin{aligned} v(D) &\geqslant (2d-2)\,v(f_0) + (2d-2)\,v(\alpha_1) + (2d-4)\,v(\alpha_2) + \dots + 2v(\alpha_{d-1}) \\ &\geqslant (2d-2)\,(v(f_0) + v(\alpha_1) + \dots + v(\alpha_a)) \geqslant 2d-2 > 1 = z. \end{aligned}$$

COROLLARY I.1.

(i)
$$\sum_v \sum_{i=1}^{d} v(\mathfrak{M}_i) \leqslant \mathbf{H}(f),$$

(ii)
$$\sum_v \sum_{i=1}^{d} v(\mathfrak{B}_i) \leqslant d\,|\mathfrak{S}| + (3d-2)\,\mathbf{H}(f).$$

PROOF. (i) By the construction of $\mathfrak{M}_i$ and by Lemma I, the double sum in question is

$$-\sum_{v \notin \mathfrak{S}} v(f/f_0),$$

which is at most $\mathbf{H}(f)$ by remark (10.2).

(ii) By the construction of $\mathfrak{B}_i$ and by Lemma I, the double sum in question is

$$\begin{aligned} &\leqslant \sum_{\substack{v \notin \mathfrak{S} \\ v(f) \geqslant 0}} v(D) + d \sum_{\substack{v \in \mathfrak{S} \\ \text{or } v(f)<0}} 1 \\ &= \sum_{\substack{v \in \mathfrak{S} \\ \text{or } v(f)<0}} (d - v(D)) \\ &\leqslant \sum_{\substack{v \in \mathfrak{S} \\ \text{or } v(f)<0}} (d + (2d-2)(-v(f)) \\ &\leqslant d\,|\mathfrak{S}| + (3d-2) \sum_{v(f)<0} -v(f) \\ &= d\,|\mathfrak{S}| + (3d-2)\,\mathbf{H}(f). \end{aligned}$$

https://doi.org/10.1017/S1446788700021406 Published online by Cambridge University Press

Write

$$a_i = \sum_v v(\mathfrak{A}_i). \tag{10.4}$$

By Riemann's Theorem there is a non-zero $y_i \in K$ with

$$v(y_i) \geqslant v(\mathfrak{A}_i) \quad \text{for } v \neq v_0,$$

$$v_0(y_i) \geqslant v_0(\mathfrak{A}_i) - a_i - g,$$

where v_0 is some arbitrary valuation picked in advance. Define z_i by the equation

$$x - \alpha_i = y_i^2 z_i. \tag{10.5}$$

Lemma J.

$$\sum_{i=1}^{d} \mathrm{H}(z_i) \leqslant 2gd + d|\mathfrak{S}| + 3d\mathrm{H}(f).$$

Proof. $\mathrm{H}(z_i) = \mathrm{H}(1/z_i) = \mathrm{H}(y_i^2/(x-\alpha_i))$, so that

$$\begin{aligned}
\mathrm{H}(z_i) &= -\sum_v \min(0, v(y_i^2/(x-\alpha_i)) \\
&= -\sum_v \min(0, 2v(y_i/\mathfrak{A}_i) + v(\mathfrak{M}_i) - v(\mathfrak{B}_i)) \\
&\leqslant \sum_v v(\mathfrak{B}_i) - 2\min(0, -a_i - g) \\
&\leqslant \sum_v v(\mathfrak{B}_i) + 2g + 2\max(0, a_i) \\
&= 2g + \max(\sum_v v(\mathfrak{B}_i), \sum_v v(\mathfrak{A}_i^2 \mathfrak{B}_i)) \\
&= 2g + \max(\sum_v v(\mathfrak{B}_i), \sum_v v(\mathfrak{M}_i)).
\end{aligned}$$

An appeal to Corollary I.1 yields

$$\sum_{i=1}^{d} \mathrm{H}(z_i) \leqslant 2gd + d|\mathfrak{S}| + (3d-2)\mathrm{H}(f) + \mathrm{H}(f)$$

and Lemma J.

It is a consequence of the lemma that there are three among $z_1, \ldots, z_d$ whose heights do not exceed

$$(d/(d-2))(2g + |\mathfrak{S}| + 3\mathrm{H}(f)) \leqslant 3(2g + |\mathfrak{S}| + 3\mathrm{H}(f)).$$

This yields the

Corollary J.1. *After suitable ordering of $z_1, \ldots, z_d$, we have*

$$\mathrm{H}(z_i) \leqslant 9\mathrm{H}(f) + 3|\mathfrak{S}| + 6g \quad (i = 1, 2, 3).$$

11. Proof of the theorem on hyperelliptic equations

Let ζ_i be quantities with $\zeta_i^2 = z_i$ $(i = 1, 2, 3)$, L the extension field $L = K(\zeta_1, \zeta_2, \zeta_3)$, Δ the degree $\Delta = [L : K]$ and g_L the genus of L/k. It follows from Corollary J.1 and Lemma G that

$$g_L \leqslant \Delta(14\mathrm{H}(f) + 5|\mathfrak{S}| + 10g). \tag{11.1}$$

Valuations v of K are extended to valuations V of L. We normalize these valuations V to have again value group $\mathbf{Z}$. Let $\mathfrak{S}_L$ be the set of valuations V which extend valuations v of $\mathfrak{S}$.

Put

$$\eta_1 = y_2\zeta_2 - y_3\zeta_3, \quad \eta_2 = y_3\zeta_3 - y_1\zeta_1, \quad \eta_3 = y_1\zeta_1 - y_2\zeta_2,$$

$$\hat{\eta}_1 = y_2\zeta_2 + y_3\zeta_3, \quad \hat{\eta}_2 = y_3\zeta_3 + y_1\zeta_1, \quad \hat{\eta}_3 = y_1\zeta_1 + y_2\zeta_2,$$

so that by (10.5),

$$\eta_i\hat{\eta}_i = \alpha_h - \alpha_j \tag{11.2}$$

for every cyclic permutation (i, j, h) of $1, 2, 3)$.

Lemma K.

(i) $$\sum_{V\notin\mathfrak{S}_L} \mathrm{V}(\eta_i) \geqslant -\tfrac{1}{2}\Delta\mathrm{H}(f),$$

(ii) $$\sum_{V\notin\mathfrak{S}_L} V(\eta_i) \leqslant \tfrac{3}{2}\Delta\mathrm{H}(f).$$

Proof.

(i) $$V(\eta_i) \geqslant \min(V(y_1\zeta_1), V(y_2\zeta_2), V(y_3\zeta_3))$$
$$= \tfrac{1}{2}\min(V(x-\alpha_1), V(x-\alpha_2), V(x-\alpha_3)),$$

and for $V \notin \mathfrak{S}_L$ this is greater than or equal to

$$-\tfrac{1}{2}\max(V(\mathfrak{M}_1), V(\mathfrak{M}_2), V(\mathfrak{M}_3))$$
$$\geqslant -\frac{1}{2}\sum_{l=1}^{d} V(\mathfrak{M}_l).$$

So

$$\sum_{V\notin\mathfrak{S}_L} \mathrm{V}(\eta_i) \geqslant -\frac{1}{2}\sum_{V\notin\mathfrak{S}_L}\sum_{l=1}^{d} V(\mathfrak{M}_l) = -\frac{\Delta}{2}\sum_{v\notin\mathfrak{S}}\sum_{l=1}^{d} v(\mathfrak{M}_l)$$
$$\geqslant -\frac{\Delta}{2}\mathrm{H}(f)$$

https://doi.org/10.1017/S1446788700021406 Published online by Cambridge University Press

by Corollary I.1.

(ii)
$$\sum_{V\notin\mathfrak{S}_L} V(\eta_i) = \sum_{V\notin\mathfrak{S}_L} V(\eta_i\hat{\eta}_i) - \sum_{V\notin\mathfrak{S}_L} V(\hat{\eta}_i)$$
$$\leqslant \sum_{V\notin\mathfrak{S}_L} V(\alpha_h-\alpha_j)+\frac{\Delta}{2}\mathrm{H}(f)$$

by (11.2) and part (i), applied to $\hat{\eta}_i$ in place of η_i. Now

$$\sum_{V\notin\mathfrak{S}_L} V(\alpha_h-\alpha_j) = \Delta\sum_{v\notin\mathfrak{S}} v(\alpha_h-\alpha_j) = -\Delta\sum_{v\in\mathfrak{S}} v(\alpha_h-\alpha_j)$$
$$\leqslant -\Delta\sum_{v\in\mathfrak{S}} v(f/f_0) \leqslant \Delta\mathrm{H}(f)$$

by (2.4) and by remark (10.2), and (ii) follows.

LEMMA L. *There are* ε_i, η_i' $(i=1,2,3)$ *in* L *with*

(11.3) $$\eta_i = \varepsilon_i^5\eta_i' \quad (i=1,2,3)$$

and with

(11.4) $$\mathrm{H}_L(\eta_i') \leqslant 10\Delta\mathrm{H}(f)+5\Delta|\mathfrak{S}|+5g_L.$$

PROOF. Pick some particular valuation $V_0\notin\mathfrak{S}_L$. We will find ε_i with

(11.5) $$V(\varepsilon_i)\geqslant\tfrac{1}{5}V(\eta_i) \quad \text{for } V\neq V_0,$$

(11.6) $$V_0(\varepsilon_i)\geqslant\tfrac{1}{5}V_0(\eta_i)-2\Delta\mathrm{H}(f)-\Delta|\mathfrak{S}|-g_L.$$

Namely, the number of $V\notin\mathfrak{S}_L$ with $V(\eta_i)\neq 0$ is $\leqslant 2\Delta\mathrm{H}(f)$ by the preceding lemma, hence the total number of V with $V(\eta_i)\neq 0$ is $\leqslant 2\Delta\mathrm{H}(f)+|\mathfrak{S}_L|\leqslant 2\Delta\mathrm{H}(f)+\Delta|\mathfrak{S}|$. Hence if we replace the righthand sides of (11.5), (11.6) by the next largest integers, the sum will be $\leqslant -g_L$, so that by Riemann's Theorem in L there is indeed a non-zero ε_i in L having (11.5) and (11.6). If now η_i' is defined by (11.3), we obtain

$$\mathrm{H}_L(\eta_i') = \mathrm{H}_L(1/\eta_i') = \mathrm{H}_L(\varepsilon_i^5/\eta_i)$$
$$= -\sum_V \min(0, 5V(\varepsilon_i)-V(\eta_i))$$
$$= -\min(0, 5V_0(\varepsilon_i)-V_0(\eta_i))$$
$$\leqslant 5(2\Delta\mathrm{H}(f)+\Delta|\mathfrak{S}|+g_L).$$

LEMMA M. *Putting*

$$A = 30\Delta\mathrm{H}(f)+15\Delta|\mathfrak{S}|+15g_L,$$
$$B = 446\mathrm{A}+1055g_L+\Delta\mathrm{H}(f),$$

https://doi.org/10.1017/S1446788700021406 Published online by Cambridge University Press

we have

$$\mathrm{H}_L(x) \leqslant 8B + 2\Delta \mathrm{H}(f).$$

Proof. Let $\mathfrak{T}_L$ be the subset of $\mathfrak{S}_L$ where

$$V(\eta_3) \geqslant \tfrac{1}{2} V(\alpha_2 - \alpha_1),$$

and $\mathfrak{T}_L^-$ the subset of $\mathfrak{T}_L$ where $\min(V(\varepsilon_1), V(\varepsilon_2), V(\varepsilon_3)) < 0$. For $V \in \mathfrak{T}_L$ we have (using 2.4))

$$(11.7) \qquad V(\varepsilon_3) \geqslant \tfrac{1}{5} V(\eta_3) \geqslant \tfrac{1}{10} V(\alpha_2 - \alpha_1) \geqslant \tfrac{1}{10} V(f/f_0).$$

From the definition of η_1, η_2, η_3,

$$\eta_1 + \eta_2 + \eta_3 = 0,$$

and therefore

$$\eta_1'(\varepsilon_1/\varepsilon_3)^5 + \eta_2'(\varepsilon_2/\varepsilon_3)^5 + \eta_3' = 0.$$

This is a Thue equation over L with coefficients $\eta_1', \eta_2', \eta_3'$ in the variables $\varepsilon_1/\varepsilon_3, \varepsilon_2/\varepsilon_3$. The height of this Thue equation is

$$-\sum_V \min(V(\eta_1'), V(\eta_2'), V(\eta_3')) \leqslant \mathrm{H}_L(\eta_1') + \mathrm{H}_L(\eta_2') + \mathrm{H}_L(\eta_3') \leqslant A$$

by Lemma L. Theorem 1 yields

$$\sum_V \min(V(\varepsilon_1/\varepsilon_3), V(\varepsilon_2/\varepsilon_3), 0) \geqslant -89A - 211 g_L.$$

We obtain

$$\begin{aligned}\sum_{V \in \mathfrak{T}_L} \min(V(\varepsilon_1), V(\varepsilon_2), V(\varepsilon_3), 0) &= \sum_{V \in \mathfrak{T}_L^-} \min(V(\varepsilon_1), V(\varepsilon_2), V(\varepsilon_3)) \\ &= \sum_{V \in \mathfrak{T}_L^-} \min(V(\varepsilon_1/\varepsilon_3), V(\varepsilon_2/\varepsilon_3), 0) + \sum_{V \in \mathfrak{T}_L^-} V(\varepsilon_3) \\ &\geqslant -89A - 211 g_L + \tfrac{1}{10} \sum_{V \in \mathfrak{T}_L^-} V(f/f_0),\end{aligned}$$

and here

$$\sum_{V \in \mathfrak{T}_L^-} V(f/f_0) \geqslant \sum_V \min(0, V(f/f_0)) = \Delta \sum_v \min(0, v(f/f_0)) \geqslant -\Delta \mathrm{H}(f).$$

Since

$$\sum_V \min(V(\eta_1'), V(\eta_2'), V(\eta_3'), 0) \geqslant -A,$$

we get

$$\begin{aligned}\sum_{V \in \mathfrak{T}_L} \min(V(\eta_1), V(\eta_2), V(\eta_3), 0) & \\ &\geqslant -A - 5(89A + 211 g_L + \tfrac{1}{10}\Delta \mathrm{H}(f)) \\ &\geqslant -446A - 1055 g_L - \Delta \mathrm{H}(f) \\ &= -B.\end{aligned}$$

https://doi.org/10.1017/S1446788700021406 Published online by Cambridge University Press

All our arguments so far go through with ζ_3 replaced by $-\zeta_3$, whence η_1, η_2, η_3 replaced by $y_2\zeta_2+y_3\zeta_3=\hat{\eta}_1$, $-y_3\zeta_3-y_1\zeta_1$, $y_1\zeta_1-y_2\zeta_2=\eta_3$, respectively. This change does not affect $\mathfrak{T}_L$; but since η_1 is changed into $\hat{\eta}_1$, we get

$$\sum_{V\in\mathfrak{T}_L}\min(0, V(\eta_1), V(\hat{\eta}_1)) \geqslant \sum_{V\in\mathfrak{T}_L}\min(0, V(\eta_1)) + \sum_{V\in\mathfrak{T}_L}\min(0, V(\hat{\eta}_1)) \geqslant -2B,$$

and therefore

$$\sum_{V\in\mathfrak{T}_L}\min(0, V(x-\alpha_2), V(x-\alpha_3)) = 2\sum_{V\in\mathfrak{T}_L}\min(0, V(y_2\zeta_2), V(y_3\zeta_3)) \geqslant -4B.$$

Now $\mathrm{V}(x) \geqslant \mathrm{V}(x-\alpha_2)+\mathrm{V}(\alpha_2)$, and taking the sum over valuations V in $\mathfrak{T}_L$ and observing (2.3), (2.11) we obtain

$$\sum_{V\in\mathfrak{T}_L}\mathrm{V}(x) \geqslant -4B-\Delta H(f). \tag{11.8}$$

Recall the definition of $\mathfrak{T}_L$ and define $\mathfrak{R}_L$ to be the set of valuations $V\in\mathfrak{S}_L$ with $V(\hat{\eta}_3) \geqslant \frac{1}{2}V(\alpha_2-\alpha_1)$. By changing ζ_2 into $-\zeta_2$ every statement about $\mathfrak{T}_L$ goes into a statement about $\mathfrak{R}_L$, so that in particular (11.8) is true with $\mathfrak{T}_L$ replaced by $\mathfrak{R}_L$. In view of $V(\eta_3\hat{\eta}_3) = V(\alpha_2-\alpha_1)$, every $V\in\mathfrak{S}_L$ belongs to either $\mathfrak{T}_L$ or $\mathfrak{R}_L$, and we obtain

$$\sum_{V\in\mathfrak{S}_L}\mathrm{V}(x) \geqslant -8B-2\Delta H(f).$$

So

$$\mathrm{H}_L(x) = -\sum_{V\in\mathfrak{S}_L}\mathrm{V}(x) \leqslant 8B+2\Delta H(f).$$

The proof of Theorem 2 is now immediate. By (11.1) and by Lemma M we have $\mathrm{H}_L(x) \leqslant 10^6\Delta(\mathrm{H}(f)+g+|\mathfrak{S}|)$, whence the desired result as a consequence of (2.11).

Added in Proof. Recently B. Dwork (private communication) has obtained Corollary 1.1 for $d \geqslant 4$, and with a better bound on the degree. In fact he deals with more general equations $ax^{d_1}+by^{d_2}+cz^{d_3}=0$. On the other hand it may be seen that the method of the present paper in the special case of the equations of Corollary 1.1 requires only $d \geqslant 4$ and yields better estimates than stated.

References

J. V. Armitage (1967), "Algebraic functions and an analogue of the geometry of numbers: The Riemann–Roch Theorem", *Arch. Math.* **18**, 383–393.

A. Baker (1968), "Contributions to the theory of diophantine equations (I), on the representation of integers by binary forms", *Phil. Trans. Royal Soc.*, London, A **263**, 173–191.

A. Baker (1969), "Bounds for the solutions of the hyperelliptic equation", *Proc. Camb. Phil. Soc.* **65**, 439–444.

H. Davenport (1965), "On $f^3(t)-g^2(t)$", *Norske Vid. Selsk. Forh.* (*Trondheim*), **38**, 86–87.

https://doi.org/10.1017/S1446788700021406 Published online by Cambridge University Press

M. Deuring (1972), *Lectures on the Theory of Algebraic Functions of One Variable* (Springer Lecture Notes **314**)

M. Eichler (1963), *Einführung in die Theorie der algebraischen Zahlen und Funktionen* (Birkhäuser Verlag, Basel and Stuttgart).

H. Grauert (1965), "Mordell's Vermutung über rationale Punkte auf algebraischen Kurven and Funktionenkörper", *Inst. Haut. Études* **25**, 131–149.

K. Mahler (1933a), "Zur Approximation algebraischer Zahlen (I). Über den grössten Primteiler binärer Formen", *Math. Ann.* **107**, 691–730.

K. Mahler (1933b), Über die rationalen Punkte auf Kurven vom Geschlecht Eins", *J. Reine Angew, Math.* **170**, 168–178.

K. Mahler (1940), "An analogue to Minkowski's geometry of numbers in a field of series", *Ann. of Math.* (2) **42**, 305–320.

Ju. I. Manin (1963), "Rational points on algebraic curves over function fields", *Izv. Akad. Nauk Mat.* **27**, 1395–1440.

C. F. Osgood (1973), "An effective lower bound on the 'diophantine approximation' of algebraic functions by rational functions", *Mathematika*, **20**, 4–15.

C. F. Osgood (1975), "Effective bounds on the 'diophantine approximation' of algebraic functions over fields of arbitrary characteristic and applications to differential equations", *Indag Math.* **37**, 105–119.

P. Samuel (1966), *Lectures on Old and New Results on Algebraic Curves*, Tata Inst., **36**.

W. M. Schmidt (1976), "On Osgood's effective Thue theorem for algebraic functions", *Commun. on Pure and Applied Math.* **29**, 759–773.

C. L. Siegel (1929), "Über einige Anwendungen diophantischer Approximationen", *Abh. d. Preuss. Akad. d. Wiss., Math. Phys. Kl.* **1**.

A. Thue (1909), "Über Annäherungswerte algebraischer Zahlen", *Journal f. Math.* **135**, 284–305.

S. Uchiyama (1961), "Rational approximations to algebraic functions", *J. Fac. Sci. Hokkaido Univ. Ser.* 1, **15**, 173–192.

University of Colorado
Boulder, Colorado 80309, USA

https://doi.org/10.1017/S1446788700021406 Published online by Cambridge University Press