\[ = \text{rank} \psi(W). \] Therefore the first lower bound for rank \( H(Q(\sqrt[3]{p})) \) in the corollary to Theorem 2.3 becomes \( 2 + \text{rank} H(k)^{3/2} - \text{rank} \psi(W) \); in particular this says that the \( l \)-class group \( H(Q(\sqrt[3]{p})) \) is not cyclic.

References

7. T. Honda, Pure cubic fields whose class numbers are multiples of 3, ibid. 3 (1971), pp. 7–12.

LeVeque's superelliptic equation over function fields

by

R. C. Mason (Cambridge) and B. Brindza (Debrecen)

I. Introduction. In a letter to Mordell written in 1925, later published, Siegel [9] proved that the hyperelliptic equation \( y^2 = g(x) \) has only finitely many solutions in integers \( x \) and \( y \); \( g \) denotes a polynomial with integer coefficients, possessing at least three simple zeros. Siegel's later investigations revealed his celebrated theorem [10] concerning the solutions of any polynomial equation \( F(x, y) = 0 \); he proved that there are only finitely many integer solutions, unless the curve associated with \( F \) has genus zero and no more than two infinite valuations. Siegel's proof was ineffective: he employed both the Mordell–Weil theorem and his own theorem on the approximation of algebraic numbers by rationals, which was a development of the pioneering work of Thue. In 1964 LeVeque [3] generalized Siegel's result on the hyperelliptic equation to prove that the superelliptic equation \( y^n = f(x) \) has only finitely many solutions in any ring of algebraic integers, unless of course it falls into the exceptional cases predicted by Siegel's general theorem. The conditions on \( f \) and \( m \) equivalent to the exceptional cases are given below \( (\ast) \). LeVeque's result was ineffective. In 1968 Baker proved the first general effective result on Diophantine equations by employing his celebrated lower bound for linear forms in logarithms: he effectively solved first the Thue equation, and then the hyperelliptic and certain superelliptic equations [1]. Baker's bounds were improved by Sprindzuk [11, 12]. LeVeque's theorem of 1964 was recently made completely effective by Brindza [2].

This paper is devoted to establishing a bound on the solutions of LeVeque's equation in the analogous case of function fields. Let \( k \) denote an algebraically closed field of characteristic zero, and \( k(z) \) the rational function field over \( k \). Let us consider the set of solutions \( X, Y \) in the ring of polynomials \( k[z] \) of the hyperelliptic equation \( Y^2 = G(X) \), where \( G \) is a polynomial with coefficients in \( k[z] \) and possessing at least three simple zeros. It is plainly possible for this equation to have infinitely many solutions, for example if the coefficients of \( G \) actually lie in \( k \). However, it is
possible to bound the degrees of the polynomial solutions $X$ and $Y$ of Schmid [9] succeeded in doing so in 1978. Schmidt's approach was developed from the work of Osgood, and his from Kolchin, on algebraic differential equations. In 1981 Mason discovered an entirely new approach to equations over function fields [4]. By means of a fundamental inequality (see Lemma 1 below) he was able to establish bounds on the degrees of the solutions which greatly improved on those of Schmid [6]. Moreover, the method actually led to an algorithm by which all the solutions of the hyperelliptic equation in $k[x]$ could be determined.

Now let $K$ denote an arbitrary finite extension of the rational function field $k(x)$, of genus $g_K$, say. Corresponding to $K$ is an algebraic curve $C$ defined over $k$, where the points of $C$ correspond to valuations on $K$. We shall blur the distinction between points of $C$ and the valuations; moreover, we shall assume that the valuations are additive with value group $Z$. If $S$ is a finite subset of $C$, then we can consider the ring $O_S$ of elements $x$ of $K$ with $v(x) > 0$ for $v$ outside $S$; that is, all the poles of $x$ lie in $S$. We shall be concerned with the solutions in $O_S$ of the equation

$$Y^n = F(X),$$

where $F$ is a polynomial with coefficients in $K$. The notion on $K$ corresponding to 'degree' on $k[x]$ is 'height': we define

$$H_K(x) = -\sum_{v \in C} \min(0, v(x)).$$

For a polynomial $f$ over $K$ and a valuation $v$, we define $v(f)$ to be the smallest of the $v(x)$ as $x$ ranges over the coefficients of $f$: the height of $F$ is then defined as above. The final ingredient needed before stating our theorem is some information on the factorisation of $F$ which corresponds to certain invariants on the curve associated with $(\ast)$. We suppose that $L$ is a finite extension of $K$ in which $F$ factorises completely, say

$$F(x) = \prod_{i=1}^n (X-x_i)^{s_i},$$

where $x_1, \ldots, x_n$ are distinct. Let us write $m_i = m(m, r_i)$ for $i = 1, \ldots, n$. We shall prove the following theorem.

**Theorem.** Provided that $(t_1, \ldots, t_n)$ is not a permutation of either $(t, 1, \ldots, 1)$ or $(2, 2, 1, \ldots, 1)$ (\$\lambda\$), all the solutions $X, Y$ in $O_S$ of $(\ast)$ satisfy

$$H_K(X) \leq 78H_K(F) + 12g_K + 6|S|,$$

where $|S|$ denotes the cardinality of $S$. Furthermore, in the exceptional cases (\$\lambda\$), we can find an extension field $K'$ of $K$ and a set of valuations $S'$ on $K'$ such that $(\ast)$ has a family of solutions in $O_{S'}$ of unbounded height.

It is easy to see that the exceptions covered by (\$\lambda\$) are in exact correspondence with those discovered by Siegel in [10]; the curve associated with $(\ast)$ then has genus zero, and the infinite valuations number one and two respectively. Our bound may be compared with those of Schmidt, $10^6H_K(F) + g_K + |S|$ and Mason, $26H_K(F) + 8g_K + 4|S|$, both for the hyperelliptic equation. It should also be mentioned that, by using the fundamental inequality in a different way, Mason was able to establish ([6], p. 120) bounds on the heights of all the solutions in $K$, not just those in $O_S$, of the superelliptic equation $(\ast)$, subject to certain degree restrictions.

In the next section we shall recall the fundamental inequality first proved by Mason in 1983, together with a genus estimate which will be of considerable value here. The proof of the theorem will then follow from an indirect application of the fundamental inequality. The exact form in which the inequality is applied depends on the integers $t_1, \ldots, t_n$; there are three cases to consider, the first two are disposed of in Section 3, and the third in Section 4.

### 2. Preliminaries.

The following inequality has formed the crucial step in the effective resolutions of the various families of equations over function fields. Proofs of the inequality may be found in [5] or [6], p. 14.

**Lemma 1.** Suppose that $\gamma_1, \gamma_2$ and $\gamma_3$ are non-zero elements of $K$ with $\gamma_1 + \gamma_2 + \gamma_3 = 0$. Let us further suppose that there is a finite set $\mathcal{V}'$ of valuations on $K$ such that $v(\gamma_1) = v(\gamma_2) = v(\gamma_3)$ for $v$ outside $\mathcal{V}'$. Then either $\gamma_1/\gamma_2$ lies in $k$, in which case $H_K(\gamma_1/\gamma_2) = 0$, or

$$H_K(\gamma_1/\gamma_2) \leq |\mathcal{V}'| + 2g_K - 2.$$

In the case of the Thue equation $f(X, Y) = \mu$, the inequality is applied directly to the identity

$$(X-a_1)(X-a_2) + (X-a_3)(X-a_4) + (X-a_5)(X-a_6)(X-a_7)(X-a_8) = 0,$$

where each $X-a_i$ is a factor of the binary form $f$. A direct application is also made in the case of the hyperelliptic equation (see § 4). However, in the case of the superelliptic equation an indirect approach yields much better bounds (see § 3). In the proof it will be necessary to factorise $F$ partially by adjoining some of the zeros $x_1, \ldots, x_n$ and in order to apply Lemma 1 we need a bound on the resulting growth in the genus. This is achieved in the following ([6], p. 67).

**Lemma 2.** Let $P(X)$ denote a non-zero polynomial of degree $n$ with coefficients in $K$. If $M$ is the field obtained by adjoining some zero $a$ of $P$ to $K$, then

$$g_M \leq 1 + [M : K](g_K - 1 + \frac{1}{2}(1 - 1/n)H_K(P)).$$

We observe that if $L$ is any finite extension of $K$, and $w$ is a valuation on $L$, then there is a valuation $v$ on $K$ and an integer $e_w$ such that $w(a)$
Lemma 2 with \( P = f \) we obtain

\[ g_M \leq d(g_k + 3(1 - 1/n)H_K(f)), \]

where \( d = [M : K] \). Let \( \mathcal{V} \) denote the set of valuations \( w \) on \( M \) at which at least one of the following occurs: \( w(F) < 0 \), \( w(\alpha) > 0 \), \( w(f'(\alpha_1)) > 0 \), \( w(f'(\alpha_2)) > 0 \), or \( w(\alpha) \) for some \( \alpha \) in \( S \): thus

\[ |\mathcal{V}| \leq H_M(F) + H_M(\alpha) + \{ (n-1)H_M(\alpha_1) + H_M(f) \} \]

\[ + \{ (n-1)H_M(\alpha_2) + H_M(f) \} + d |S|. \]

However, from the ordering of \( \alpha_1, \ldots, \alpha_n \) we have

\[ H_M(\alpha_1) + H_M(\alpha_2) \leq 2H_M(F)/n, \]

and hence

\[ |\mathcal{V}| \leq d (|S| + (6/(2n))H_K(F)). \]

Let \( \mathcal{W}_i, i = 1, 2 \), denote the set of valuations \( w \) outside \( \mathcal{V} \) such that \( w(X - \alpha_i) > 0 \). Now whenever \( w \) lies outside \( \mathcal{V} \) we obtain \( w(\alpha_i) \geq 0 \) for \( i = 1, 2 \), \( w(X(\alpha_i - \alpha_j)) = 0 \), \( w(\alpha_i) = 0 \) and \( w(f'(\alpha_i)) = 0 \). If in fact \( w \) lies in \( \mathcal{W}_i \) then \( w(X - \alpha_j) = 0 \) for \( j \neq i \), and so

\[ r_i w(X - \alpha_i) = mw(Y), \]

and in particular \( t_i \) divides \( w(X - \alpha_i) \). We conclude that

\[ t_i |\mathcal{W}_i| \leq H_M(X - \alpha_i) \quad (i = 1, 2). \]

If \( w \) lies outside \( \mathcal{W}_1, \mathcal{W}_2 \) and \( \mathcal{W} \) then \( w(X - \alpha_1) = w(X - \alpha_2) = w(\alpha_1 - \alpha_2) = 0 \), so we may apply Lemma 1 to obtain

\[ H_M \left( \frac{X - \alpha_1}{\alpha_1 - \alpha_2} \right) \leq |\mathcal{W}_1| + |\mathcal{W}_2| + |\mathcal{V}| + 2g_M. \]

A combination of this inequality with the bounds above on each of the terms on the right-hand side, together with the inequality

\[ H_M(X) \leq H_M \left( \frac{X - \alpha_1}{\alpha_1 - \alpha_2} \right) + 2H_M(\alpha_1) + H_M(\alpha_2), \]

derived from the height inequalities in Section 2, yields the truth of the theorem in the first case.

In the second case, when \( t_1 = t_2 = 2 < t_3 \), we use the same method as above, but applied to the equation

\[ (X - \alpha_1)(\alpha_2 - X) + (\alpha_1 - \alpha_2) = 0. \]

Here we denote by \( M \) the subfield of \( L \) obtained by adjoining \( \alpha_1 \) and \( \alpha_3 \) to \( K \). The details of the proof are similar to those above, only the slight modification to \( H_M(\alpha_1) + H_M(\alpha_3) \leq 3H_M(f)/n \) is required.
4. Final case. Here we suppose that \( t_1 = t_2 = t_3 = 2 \). Now we denote by \( M \) the subfield of \( L \) obtained by adjoining \( a_1, a_2 \) and \( a_3 \) to \( K \); as before we write \( d = [M : K] \). A triple application of Lemma 2 yields

\[
q_M \leq d \left[ 2q_k + \frac{2}{3} (1 - 1/n) H_K(f) \right].
\]

Let \( \mathcal{V} \) denote the set of valuations \( w \) on \( M \) at which \( w(F) < 0 \), \( w(x) > 0 \), \( w(x') > 0 \), \( i = 1, 2 \) or 3, or \( w|v \) for some \( v \) in \( S \); we obtain the bound

\[
|\mathcal{V}| \leq d \left[ |S| + (8 - (3/n)) H_K(F) \right].
\]

As usual \( X, Y \) denotes a fixed solution of (**) with \( X \) in \( \mathcal{O}_S \), \( Y \) in \( K^* \). Now write \( N \) for the field obtained from \( M \) by adjoining the square roots of \( X - a_1 \), \( X - a_2 \) and \( X - a_3 \). From (**) we deduce that \( w(X-a_i) \) is even for \( w \) outside \( \mathcal{V} \), \( i = 1, 2, 3 \), and so ramification from \( M \) to \( N \) only occurs inside \( \mathcal{V} \). From [6], p. 34, we see that \( 2q_N - 2 + |\mathcal{X}| = [N : M] (2q_M - 2 + |\mathcal{V}|) \), where \( \mathcal{X} \) denotes the set of valuations \( n \) on \( N \) such that \( n|w \) for some \( w \) in \( \mathcal{V} \). The fundamental inequality can now be applied directly with \( \mathcal{V} = \mathcal{X} \) and

\[
\gamma_1 = \pm \sqrt{X - a_2} \pm \sqrt{X - a_3},
\]

and \( \gamma_2, \gamma_3 \) determined by permutation of the suffixes 1, 2, 3; the signs being chosen so that \( \gamma_1 + \gamma_2 + \gamma_3 = 0 \). Lemma 1 then yields ([6], p. 34)

\[
H_N \left( \frac{2X - a_1 - a_3}{a_2 - a_3} \right) \leq 4(2q_N + |\mathcal{X}|).
\]

The proof of the theorem is then completed by combining this last with the inequalities above. It will be observed that the difference in the use of the fundamental inequality between Sections 3 and 4 is that in the latter the set \( \mathcal{V} \) is independent of the solution \( X, Y \) of (**) but in the former it is not.

Finally, we deal with the necessity of the condition (**). If \( t_1 = t = t_3 = t_4 = 1 \), then we choose \( K' \) to be a field such that \( x \) is an \( m \)th power, and \( S' \) to contain the poles of \( a_1, \ldots, a_n \). In this case writing \( X = a_1 + T^{m'} \) for any \( T \) in \( \mathcal{O}_S \), forms a set of solutions of unbounded height. Similarly, if \( t_1 = t_2 = 2 \) and \( t_3 = \ldots = t_4 = 1 \), then the equation (**) may be transformed into \( R^2 - \alpha T^n = \beta \), where \( X \) is a polynomial in \( T \) and \( Y \) is a polynomial in \( R \). Choosing \( K' \) to contain the square root of \( \alpha \), and \( S' \) to contain the poles of \( a_1, \ldots, a_n \), together with the zeros and poles of some non-constant \( \eta \) in \( K' \), writing \( R + \sqrt{\alpha} T = \eta^n \) provides a set of solutions of unbounded height as \( m \to \infty \).

References
