Math. Proc. Camb. Phil. Soc. (1983), **93**, 219–230
Printed in Great Britain

The hyperelliptic equation over function fields

By R. C. MASON†
St John's College, Cambridge

(*Received* 25 *March* 1982; *Revised* 22 *September* 1982)

1. *Introduction.* Siegel, in a letter to Mordell of 1925 (9), proved that the hyperelliptic equation $y^2 = g(x)$ has only finitely many solutions in integers x and y, where g denotes a square-free polynomial of degree at least three with integer coefficients. Siegel's method reduces the hyperelliptic equation to a finite set of Thue equations $f(x, y) = 1$, where f denotes a binary form with algebraic coefficients and at least three distinct linear factors; x and y are integral in a fixed algebraic number field. Siegel had already proved that the Thue equations so obtained have only finitely many solutions. However, as is well known, the work of Siegel is ineffective in that it fails to provide bounds on the integer solutions of $y^2 = g(x)$. In 1969 Baker (1), using the theory of linear forms in logarithms, employed Siegel's technique to establish explicit bounds on x and y; Baker's result thus reduced the problem of determining all integer solutions of the hyperelliptic equation to a finite amount of computation.

The present paper solves the analogous problem for function fields. Let k be an algebraically closed field of characteristic zero, and let $k(z)$ denote the rational function field over k. In 1978 Schmidt (8), developing work of Osgood and Kolchin on solutions of algebraic differential equations*, established explicit bounds on the degrees of all polynomials X and Y in $k[z]$ satisfying the hyperelliptic equation $Y^2 = G(X)$; here G denotes a square-free polynomial of degree at least three in X with coefficients in $k[z]$. However, since the ground field k is infinite, bounds on the degrees of X and Y do not in themselves enable one to compute the polynomial solutions of the hyperelliptic equation. Here we provide an algorithm for determining all solutions X and Y in $k[z]$ of $Y^2 = G(X)$; moreover we establish (Theorem 2) a simple criterion for the equation to have an infinite number of solutions. Our method differs completely from that of Schmidt, and is much simpler; in particular, it will be observed that no previous results on Thue equations are employed, and that somewhat better bounds are established on the degrees of X and Y without any appeal to the Riemann–Roch theorem. The crux of our proof is an inequality (see §3), whose motivation lies in the study of linear forms in the logarithms of algebraic functions; a form of this inequality was derived for the first time in our solution of the Thue equation over function fields (5).

Our results in fact refer to a rather more general situation than that described above: we shall deal with the hyperelliptic equation over an arbitrary finite extension K of $k(z)$. Let $\mathcal{O}$ denote the ring of elements of K integral over $k[z]$, and suppose that $\alpha_1, \ldots, \alpha_n$ are n ($\geqslant 3$) distinct elements of $\mathcal{O}$; we shall prove the following results.

† Supported by the Science Research Council and assisted by the Imperial Chemical Industries Educational Trust.

* See also Queen's Papers in Pure and Applied Maths., no. 54 (1980), 33–65.

https://doi.org/10.1017/S0305004100060497 Published online by Cambridge University Press

THEOREM 1. *All solutions X, Y in $\mathcal{O}$ of the hyperelliptic equation*

$$Y^2 = (X-\alpha_1)\ldots(X-\alpha_n), \tag{1}$$

can be effectively determined.

THEOREM 2. *The equation* (1) *has an infinite number of solutions X, Y in $\mathcal{O}$ if and only if there exists a substitution of the form*

$$X = \alpha x + \beta, \quad Y = \gamma y,$$

with α, β, γ in $\mathcal{O}$, $\alpha\gamma$ non-zero, which transforms (1) *into the equation $y^2 = g(x)$, where g has coefficients in the ground field k. Moreover, if such a substitution exists then, apart from the infinite family obtained by choosing x and y in k, the equation* (1) *has only finitely many solutions in $\mathcal{O}$.*

It should be mentioned here that the celebrated theorem of Manin and Grauert (see (7)) states that any curve defined over K, of genus at least two, has only finitely many points in K (not merely in $\mathcal{O}$), provided that it is not birationally equivalent over K to a curve defined over k; this is the analogue for function fields of the famous Mordell conjecture for number fields. It would be possible to derive Theorem 2 from this result, if n is at least five. However, the work of Grauert relies heavily on deep results from algebraic geometry, as developed by Weil and Zariski; moreover, no attempt appears to have been made to derive any effective results from the Manin–Grauert theorem, and without doubt such a task would be most formidable, even if it were to prove possible in principle.

For the proofs of Theorems 1 and 2 we shall consider, following Siegel, the field L generated over K by the square roots of $X-\alpha_1$, $X-\alpha_2$ and $X-\alpha_3$. In §6 we shall show that there are only a finite number of effectively determinable possibilities for L as X, Y run over the solutions in $\mathcal{O}$ of (1); here we shall appeal to a well-known result concerning torsion points on abelian varieties. It will thus suffice to suppose that $X-\alpha_1$, $X-\alpha_2$ and $X-\alpha_3$ are all squares in a fixed field L. Under this supposition we shall establish Theorems 1 and 2 in §4, using the fundamental inequality obtained in §3. Preliminary results on valuations will be recorded in §2. In §5 we shall calculate a bound on the heights of solutions X, Y in $\mathcal{O}$ of (1). Finally, in §7, we shall solve a particular elliptic equation in polynomials, thereby illustrating the algorithm described in §4. It will be observed that no machine computation is necessary here, in contrast with the analysis for the rational case (1); this is a consequence of the much sharper bounds appertaining in this context.

We note that, in order to speak meaningfully of the computation of solutions in Theorem 1, it is necessary to assume from the outset that the ground field k is presented explicitly; this means here that there is an algorithm to determine the zeros of any polynomial with coefficients in k.

2. *Preliminaries.* We begin by recalling the definitions of the canonical valuations on the finite extension K of $k(z)$ (see (5)). Since $K/k(z)$ is separable, K is obtained from $k(z)$ by adjoining a single element y, of degree d over $k(z)$, say. By considering the Puiseux expansions of y as in (5), we may construct valuations v on K by defining $|f|_v = \exp(-\mathrm{ord}_v f)$ for f non-zero, where $\mathrm{ord}_v f$ denotes the order of vanishing of the Laurent expansion of f in powers of z_v. If $|z-a|_v < 1$ for some a in k, then the local parameter z_v is equal to $(z-a)^{1/e_v}$ and we write $v|a$, whilst if $|z|_v > 1$ then $z_v = z^{-1/e_v}$

https://doi.org/10.1017/S0305004100060497 Published online by Cambridge University Press

and we write $v|\infty$; e_v is termed the *ramification index* of v. Each valuation defined is non-archimedean and discrete; furthermore, the *product formula*

$$\prod_v |f|_v = 1 \tag{2}$$

is satisfied for any non-zero f. We also define the *height* $H(f)$ by

$$H(f) = \sum_v \max(0, \log |f|_v),$$

as in (8). (This differs slightly from previously; in (5) we took the height to be

$$\exp(H(f)/[K:k(f,z)]).)$$

We introduce the parameter

$$g = \tfrac{1}{2}\sum_v (e_v - 1) - d + 1; \tag{3}$$

g is termed the *genus* of K.

The mapping $f \mapsto f'$ on $k(z)$ extends uniquely to a derivation on K. We may also define a local derivation d/dv on the field of formal Laurent expansions in powers of z_v, namely differentiation with respect to z_v. If $v|a$ is a finite valuation on K, then $z = a + z_v^{e_v}$, so $\log |dz/dv|_v = 1 - e_v$; similarly $\log |dz/dv|_v = 1 + e_v$ if $v|\infty$. In view of (3) above, and the product formula (2), we obtain

$$\exp(2-2g) = \prod_v \left|\frac{df}{dv}\right|_v \tag{4}$$

for any f in $K\backslash k$; this follows since $df/dv = f'(dz/dv)$ for each v, and $\sum_{v|\infty} e_v = d$. We also record two inequalities on the order of vanishing of the local derivative of f at v, readily obtained by considering the Laurent expansion of f in powers of z_v. We have namely

$$\left|\frac{df}{dv}\right|_v \leqslant e|f|_v \quad \text{for all } f, \quad \text{and} \quad \left|\frac{df}{dv}\right|_v \leqslant 1 \quad \text{if} \quad |f|_v \leqslant 1; \tag{5}$$

these inequalities will form the basis of the proof of the fundamental lemma in the next section.

Finally we record a lemma which asserts that an element of K may be determined effectively by the set of values it assumes. The proof of this result may be found in ((5), §4).

LEMMA. *Suppose that for each valuation v on K we are given an integer m_v, such that $m_v = 0$ for all but a finite number of v. Then we can determine effectively whether there exists an element f in K such that $|f|_v = e^{m_v}$ for all v. Moreover, if such an f exists then it may be computed, and it is unique to within a non-zero factor in k.*

3. *The fundamental inequality.* As in §1, X and Y will denote a solution in $\mathcal{O}$ of (1), and L will denote the extension field of K generated by ξ_1, ξ_2 and ξ_3, such that $\xi_i^2 = X - \alpha_i$ for $i = 1, 2, 3$. In §6 we shall show that L may be chosen from a finite effectively determinable set; it suffices then to assume that L is fixed, as we shall do in this and the following section. We denote by $\mathcal{O}_L$ the ring of elements of L integral over $k[z]$, so ξ_1, ξ_2 and ξ_3 are elements of $\mathcal{O}_L$; we also define elements β_i, $\hat{\beta}_i$ of $\mathcal{O}_L$ for

https://doi.org/10.1017/S0305004100060497 Published online by Cambridge University Press

$i = 1, 2, 3$ by $\beta_1 = \xi_2 - \xi_3$, $\hat{\beta}_1 = \xi_2 + \xi_3$, with β_2, $\hat{\beta}_2$, β_3 and $\hat{\beta}_3$ defined by permutation of the suffixes. Then $\beta_1\hat{\beta}_1 = \alpha_3 - \alpha_2$, $\beta_2\hat{\beta}_2 = \alpha_1 - \alpha_3$, $\beta_3\hat{\beta}_3 = \alpha_2 - \alpha_1$, and

$$\beta_1 + \beta_2 + \beta_3 = 0; \tag{6}$$

further, three similar equations,

$$\beta_1 + \hat{\beta}_2 - \hat{\beta}_3 = -\hat{\beta}_1 + \beta_2 + \hat{\beta}_3 = \hat{\beta}_1 - \hat{\beta}_2 + \beta_3 = 0 \tag{7}$$

hold. It is these equations which form the crux of the analysis, both in the classical case of algebraic numbers (see (2)), and here. In the classical case each β_i is factorized as $\delta_i\epsilon_i^3$, where δ_i has bounded height and ϵ_i is a unit; this is possible since each β_i divides the fixed element $\mu = (\alpha_3 - \alpha_2)(\alpha_1 - \alpha_3)(\alpha_2 - \alpha_1)$. Thus equation (6) becomes a Thue equation in ϵ_2/ϵ_1 and ϵ_3/ϵ_1, namely

$$\delta_2(\epsilon_2/\epsilon_1)^3 + \delta_3(\epsilon_3/\epsilon_1)^3 = -\delta_1;$$

other Thue equations may be derived from (7). Using his bounds for the general Thue equation, Baker (1) obtained explicit bounds for the integer solutions of the hyperelliptic equation; these bounds are multiply exponential with respect to the coefficients. In the case of function fields, a modification of the procedure, involving the use of the Riemann–Roch theorem, is made by Schmidt (8); he transforms (6) into an infinite set of Thue equations, whose coefficients nevertheless have bounded heights. The inequality proved below enables us to solve (6) directly, without reduction to a system of Thue equations as in previous work; it will be noted, moreover, that the bounds thereby obtained in §5 vary only as a fixed multiple of the height, and are not exponential as in the classical case. The inequality established here in fact takes a simpler and sharper form than previously (5).

LEMMA. *Suppose that γ_1, γ_2 and γ_3 are non-zero elements of K such that $\gamma_1 + \gamma_2 + \gamma_3 = 0$, and such that $|\gamma_1|_v = |\gamma_2|_v = |\gamma_3|_v$ for all valuations v not in the finite set $\mathscr{V}$. Then either γ_2/γ_3 lies in k, or*

$$H(\gamma_2/\gamma_3) \leqslant 2g - 2 + |\mathscr{V}|, \tag{8}$$

where $|\mathscr{V}|$ denotes the number of elements of $\mathscr{V}$.

Proof. It will be convenient to write $f = \gamma_2/\gamma_3$, so in view of the equation

$$\gamma_1 + \gamma_2 + \gamma_3 = 0$$

we have

$$\gamma_1/\gamma_3 = -(f+1),$$

and we may assume that f is not in k. Let $\mathscr{V}_1$, $\mathscr{V}_2$, $\mathscr{V}_3$ denote the sets of valuations v on K such that $|f|_v > 1$, $|f|_v < 1$, $|f+1|_v < 1$ respectively, so that each of the disjoint sets $\mathscr{V}_1$, $\mathscr{V}_2$, $\mathscr{V}_3$ is contained in $\mathscr{V}$. Furthermore, for v in $\mathscr{V}_1$, $\mathscr{V}_2$ or $\mathscr{V}_3$, let m_v denote the positive integer $\log|f|_v$, $-\log|f|_v$ or $-\log|f+1|_v$ respectively. In view of the product formula (2), applied to f and to $f+1$, we have $H(f) = \Sigma m_v$ for each $i = 1, 2, 3$, where the sum is taken over v in $\mathscr{V}_i$.

We now employ the inequalities (5) to estimate the order of vanishing of df/dv at v. If v lies in $\mathscr{V}_1$, then we have $|df/dv|_v \leqslant e|f|_v = \exp(1+m_v)$, whilst if v lies in $\mathscr{V}_2$ or $\mathscr{V}_3$, then we have $|df/dv|_v \leqslant \exp(1-m_v)$. Since $|df/dv|_v \leqslant 1$ for all other valuations v, the relation (4), on taking logarithms, yields

$$2 - 2g \leqslant \Sigma(1+m_v) + \Sigma(1-m_v),$$

where the first sum is taken over v in $\mathscr{V}_1$, and the second over v in $\mathscr{V}_2 \cup \mathscr{V}_3$. The required inequality (8) now follows.

https://doi.org/10.1017/S0305004100060497 Published online by Cambridge University Press

The lemma just proved shows that the height of γ_2/γ_3 is bounded. We now show further that, unless γ_2/γ_3 is an element of k, it has only finitely many possibilities, which may be effectively determined. For it follows from (8) that, since $\mathscr{V}$ is finite and $|\gamma_2/\gamma_3|_v = 1$ for v not in $\mathscr{V}$, the complete set of values assumed by γ_2/γ_3 has only a finite number of possibilities. Using the lemma in § 2, we conclude that γ_2/γ_3 has, up to a factor in k, only finitely many possibilities, effectively determinable. Thus we may write $\gamma_2/\gamma_3 = a\alpha$, where a is some element of k and α belongs to a finite set of elements of K. Similarly γ_1/γ_3 has only finitely many possibilities up to a factor in k, so we may write $\gamma_1/\gamma_3 = b\beta$, with b in k and β in a finite effectively determinable set. However, from $\gamma_1+\gamma_2+\gamma_3 = 0$ we have $a\alpha+b\beta = -1$, which, on differentiating, gives

$$a\alpha' + b\beta' = 0.$$

Therefore a and b are determined unless α' and β' both vanish, that is unless α and β are elements of k. Hence we conclude that γ_2/γ_3 is either an element of k, or belongs to a finite effectively determinable set of elements of K, as required.

4. *Proofs of Theorems* 1 *and* 2. We shall continue to assume throughout this section that X and Y are elements of $\mathcal{O}$ satisfying (1), and such that $X-\alpha_1$, $X-\alpha_2$ and $X-\alpha_3$ are squares in the fixed field L. We shall prove in § 6 that L may be chosen from a finite effectively determinable set, and, in view of this result, our assumption suffices for the proofs of Theorems 1 and 2. As in § 3 we define elements β_i, $\hat{\beta}_i$ in $\mathcal{O}_L$ for $i = 1, 2, 3$, which satisfy

$$\beta_1+\beta_2+\beta_3 = 0 \tag{6}$$

with μ/β_1, μ/β_2, μ/β_3 also elements of $\mathcal{O}_L$, where $\mu = (\alpha_3-\alpha_2)(\alpha_1-\alpha_3)(\alpha_2-\alpha_1)$. If v is a finite valuation on L such that $|\mu|_v = 1$, then it follows that $|\beta_i|_v = 1$ for $i = 1, 2, 3$. Thus if $\mathscr{V}$ denotes the set of infinite valuations on L and those finite valuations v for which $|\mu|_v < 1$, then $|\beta_1|_v = |\beta_2|_v = |\beta_3|_v$ for v not in $\mathscr{V}$. From the last paragraph in § 3, we conclude that β_2/β_3 is either an element of k, or a member of a finite effectively determinable set. We now show that the ratio $\beta_2/\beta_3 = f$ determines X uniquely. For brevity we denote by λ the element $(\alpha_1-\alpha_3)/(\alpha_2-\alpha_1)$. We have $\hat{\beta}_2/\hat{\beta}_3 = \lambda/f$, and so, using equation (6) and the equation $\beta_1+\hat{\beta}_2-\hat{\beta}_3 = 0$ in (7), we obtain $\beta_1/\beta_3 = -(f+1)$ and $\beta_1/\hat{\beta}_3 = 1-\lambda/f$. Thus $\hat{\beta}_3/\beta_3 = f(f+1)/(\lambda-f)$; but $\beta_3\hat{\beta}_3 = \alpha_2-\alpha_1$ and

$$(X-\alpha_1)/(\alpha_2-\alpha_1) = \xi_1^2/\beta_3\hat{\beta}_3 = \tfrac{1}{4}(\beta_3+\hat{\beta}_3)^2/\beta_3\hat{\beta}_3, \tag{9}$$

so f determines X as required.

We now distinguish two cases, according to the values of the parameters $\alpha_1, \ldots, \alpha_n$ in (1). In the first case at least one of the ratios $\lambda_i = (\alpha_1-\alpha_i)/(\alpha_2-\alpha_1)$, $3 \leqslant i \leqslant n$, is not an element of k. Rearranging $\alpha_3, \ldots, \alpha_n$ if necessary, we may assume that in fact $\lambda_3 = \lambda$ is not an element of k. However, if $\beta_2/\beta_3 = f$, then $\hat{\beta}_2/\hat{\beta}_3 = \lambda/f$, so at least one of β_2/β_3 and $\hat{\beta}_2/\hat{\beta}_3$ is not in k. In view of the equations $\beta_1+\beta_2+\beta_3 = \beta_1+\hat{\beta}_2-\hat{\beta}_3 = 0$, the last paragraph in § 3 implies that at least one of f and λ/f belongs to a finite set of effectively determinable elements of L. Since f determines X uniquely, we conclude that Theorems 1 and 2 are established in this case.

We now assume that the alternative case holds, so that each λ_i, $3 \leqslant i \leqslant n$, lies in k. As before, X is uniquely determined by $f = \beta_2/\beta_3$, which is either an element of k, or a member of a finite effectively determinable set. We shall show that, unless Y

https://doi.org/10.1017/S0305004100060497 Published online by Cambridge University Press

vanishes, all the solutions X, Y in $\mathcal{O}$ of (1) such that f lies in k are derived from the solutions x, y in k of $y^2 = \Pi_{i=1}^{n}(x+\lambda_i)$ via a substitution of the form

$$X = \alpha x + \beta, \quad Y = \gamma y \quad (\alpha, \beta, \gamma \in \mathcal{O}, \quad \alpha\gamma \neq 0),$$

as specified in Theorem 2. For, if f lies in k, then since $\hat{\beta}_3/\beta_3 = f(f+1)/(\lambda - f)$, $\hat{\beta}_3/\beta_3$ is also an element of k and, in view of (9), so also is $x = (X-\alpha_1)/(\alpha_2-\alpha_1)$. This substitution transforms (1) into $Y^2 = (\alpha_2-\alpha_1)^n \Pi(x+\lambda_i)$, where Y is an element of $\mathcal{O}$ and $x, \lambda_1, \ldots, \lambda_n$ are elements of k. Assuming $Y \neq 0$, this equation possesses a solution if and only if $(\alpha_2-\alpha_1)^n$ is a square in $\mathcal{O}$, say $(\alpha_2-\alpha_1)^n = \gamma^2$. The substitution $Y = \gamma y$ transforms the above equation into $y^2 = \Pi(x+\lambda_i)$. Theorems 1 and 2 are therefore established in this case also.

To complete the proof of Theorem 2 we have to show that L has only finitely many possibilities; the proof of Theorem 1 is completed by showing that these possibilities may be determined effectively; both these aims will be accomplished in §6.

5. *Bounds.* In this section we shall establish the inequality

$$H(X) \leqslant 8g + 4r + 26H \tag{10}$$

for any solution X, Y in $\mathcal{O}$ of the hyperelliptic equation (1). Here g denotes the genus of K, r denotes the number of infinite valuations of K, and H is the height (as in (8)) of the polynomial on the right of (1), that is $H = \Sigma H(\alpha_i)$. Schmidt (8) establishes the inequality $H(X) \leqslant 10^6(g+r+H)$.

As in §4, we denote by $\mathscr{V}$ the set of infinite valuations on L and those finite valuations ν with $|\mu|_\nu < 1$. Furthermore, we denote by M the integer $\max(0, 2g_L - 2 + |\mathscr{V}|)$, where g_L is the genus of L. Applying the fundamental inequality (8) to the equation $\beta_1 + \beta_2 + \beta_3 = 0$, we have

$$H_L(\beta_2/\beta_3) \leqslant M,$$

where $H_L(f)$ is the height of an element f in L. Furthermore, in view of the equations (7), M is also an upper bound for each of $H_L(\hat{\beta}_2/\beta_3)$, $H_L(\beta_2/\hat{\beta}_3)$ and $H_L(\hat{\beta}_2/\hat{\beta}_3)$. Since $\beta_2\hat{\beta}_2 = \alpha_1 - \alpha_3$, $\beta_3\hat{\beta}_3 = \alpha_2 - \alpha_1$, and $X - \alpha_1 = \xi_1^2 = \frac{1}{4}(\hat{\beta}_2 - \beta_2)^2$, it follows that

$$2(2X - \alpha_1 - \alpha_3)/(\alpha_2 - \alpha_1) = (\hat{\beta}_2/\beta_3)(\hat{\beta}_2/\hat{\beta}_3) + (\beta_2/\beta_3)(\beta_2/\hat{\beta}_3).$$

However, if f and h are any two elements of L, the inequality

$$\max\{H_L(f+h), H_L(fh)\} \leqslant H_L(f) + H_L(h)$$

is readily established. We have, therefore,

$$H_L((2X - \alpha_1 - \alpha_3)/(\alpha_2 - \alpha_1)) \leqslant 4M,$$

so it will suffice to calculate a bound on the integer M.

We denote by δ the degree of the extension L/K, so in view of (3) we have

$$2g_L - 2 = \sum_{\nu} (e_\nu - 1) - 2d\delta.$$

Here e_ν denotes the ramification index of a valuation ν on L, as defined in §2. Since $\Sigma_{\nu|v} e_\nu = \delta e_v$ for each v on K, the genera of L and K are related by

$$(2g_L - 2) - \delta(2g - 2) = \sum_{v} (\delta - r_v),$$

https://doi.org/10.1017/S0305004100060497 Published online by Cambridge University Press

where r_v denotes the number of valuations ν on L such that $\nu|v$. It follows *since $g \geqslant 0$ that

$$M \leqslant \delta(2g+r+|\mathscr{W}|),$$

where $\mathscr{W}$ is the set of finite valuations v on K such that $r_v < \delta$ or $|\mu|_v < 1$. It is therefore sufficient to establish the inequality $|\mathscr{W}| \leqslant 6H$. For if f is an element of K then $H_L(f) = \delta H(f)$, and so

$$H((2X-\alpha_1-\alpha_3)/(\alpha_2-\alpha_1)) \leqslant 4(2g+r+|\mathscr{W}|).$$

Hence

$$H(X) \leqslant 2H+8g+4r+4|\mathscr{W}|,$$

and so the desired inequality (10) follows.

In order to establish the inequality $|\mathscr{W}| \leqslant 6H$, we shall show first that every v in $\mathscr{W}$ satisfies $|\alpha_i-\alpha_j|_v < 1$ for some $1 \leqslant i < j \leqslant n$, $i \leqslant 3$; this fact will be of crucial importance in the next section. If $v \nmid \infty$ satisfies $|\alpha_i-\alpha_j|_v = 1$ for all i,j with $1 \leqslant i < j \leqslant n$, $i \leqslant 3$, then, for any $i = 1, 2, 3$, we have $|X-\alpha_i|_v = 1$ or $|Y|_v^2$. Thus the equation $X-\alpha_i = \xi_i^2$ is soluble for ξ_i in a Laurent series in powers of z_v. We conclude that, if ν is a valuation on L such that $\nu|v$, then $e_\nu = e_v$. Since $\Sigma_{\nu|v} e_\nu = \delta e_v$, we have $r_v = \delta$, and so v is not an element of $\mathscr{W}$, as required.

Let us denote by $\mathscr{W}_i$ the set of finite valuations v on K such that $|\alpha_i-\alpha_j|_v < 1$ for some $j \neq i$; in view of the last paragraph we have $\mathscr{W} \subset \mathscr{W}_1 \cup \mathscr{W}_2 \cup \mathscr{W}_3$. However, if v is finite then $-\log|\alpha_i-\alpha_j|_v$ is a non-negative integer, and so

$$|\mathscr{W}_i| \leqslant -\sum_{j \neq i} \sum_{v \nmid \infty} \log|\alpha_i-\alpha_j|_v = \sum_{j \neq i} \sum_{v \nmid \infty} \log|\alpha_i-\alpha_j|_v \leqslant \sum_{j \neq i} H(\alpha_i-\alpha_j).$$

Since $H = \Sigma H(\alpha_i)$, the sum of the right is at most $(n-2)H(\alpha_i)+H$. Assuming that $\alpha_1, \ldots, \alpha_n$ are chosen in such an order that $H(\alpha_1) \leqslant \ldots \leqslant H(\alpha_n)$, it follows that $H(\alpha_1)+H(\alpha_2)+H(\alpha_3) \leqslant 3H/n$, and so $|\mathscr{W}| \leqslant 6H$ as required.

6. *The determination of the field L.* In this section we shall complete the proofs of Theorems 1 and 2 by showing that there are only a finite number of effectively determinable possibilities for the field L, generated over K by the square roots of $X-\alpha_1$, $X-\alpha_2$ and $X-\alpha_3$. Clearly it suffices to show that a finite set $f_1, \ldots, f_m$ of non-zero elements of K may be computed such that, for each solution X, Y in $\mathcal{O}$ of (1), some ratio $(X-\alpha_1)/f_j$ is a square in K.

Let us first show that such a finite set $f_1, \ldots, f_m$ exists; here we shall use a familiar result from the theory of abelian varieties. We shall assume a knowledge of the concepts of a *divisor* on K and the *Jacobian variety* $\mathscr{J}$ as the quotient of the group of divisors of degree zero by the subgroup of principal divisors. We shall denote by $\mathscr{S}$ the set of infinite valuations plus the set $\mathscr{W}_1$ of finite valuations defined in § 5. It follows that the integer $\log|X-\alpha_1|_v$ is *even* for each v not in the finite set $\mathscr{S}$. Hence the principal divisor $(X-\alpha_1)$, with components $-\log|X-\alpha_1|_v$, may be expressed in the form $2\mathfrak{a}+\mathfrak{b}$, where $\mathfrak{a}$ and $\mathfrak{b}$ have degree zero, and $\mathfrak{b}$ has only finitely many possibilities which are effectively determinable. Indeed, $\mathfrak{b}$ may be assumed to have components 0 outside $\mathscr{S}$, components 0 or 1 on $\mathscr{S}$, with the exception of a single pole at some fixed infinite valuation v_0.

* In fact we obtain $H(X) \leqslant \max(2H, 26H+8g+4r-8)$ without using $g \geqslant 0$.

https://doi.org/10.1017/S0305004100060497 Published online by Cambridge University Press

We shall suppose henceforth that $\mathfrak{b}$ is fixed. In the quotient group $\mathscr{J}$ we have $2[\mathfrak{a}] = -[\mathfrak{b}]$. However, since (see (4)) $\mathscr{J}$ may be regarded as an abelian variety of dimension g, it follows [(6), page 64] that this equation has 4^g solutions in $\mathscr{J}$, given by $[\mathfrak{a}_1], \ldots, [\mathfrak{a}_p]$ say, where $p = 4^g$. Choosing, for each $j = 1, \ldots, p$, an element F_j in K such that $(F_j) = 2\mathfrak{a}_j + \mathfrak{b}$, it now follows that, if $(F) = 2\mathfrak{a} + \mathfrak{b}$ for some $\mathfrak{a}$, then $[\mathfrak{a}] = [\mathfrak{a}_j]$ for some j, and so F/F_j is a square in K. Taking the union of the sets $F_1, \ldots, F_p$ over the finitely many possibilities for $\mathfrak{b}$, we achieve a finite set $f_1, \ldots, f_m$ as required. We conclude that L has only finitely many possibilities, and so the proof of Theorem 2 is complete.

To complete the proof of Theorem 1, the following lemma is required.

LEMMA. *For any divisor $\mathfrak{b}$ of degree zero, a set of representatives for the p solutions of $2[\mathfrak{a}] = -[\mathfrak{b}]$ may be determined effectively.*

To obtain Theorem 1, we observe that, if $\mathfrak{a}_1, \ldots, \mathfrak{a}_p$ are the divisors determined by this lemma, then each $2\mathfrak{a}_j + \mathfrak{b}$ is principal. Using the lemma in § 2 we can compute, for each $j = 1, \ldots, p$, an element F_j satisfying $(F_j) = 2\mathfrak{a}_j + \mathfrak{b}$. Hence a finite set $f_1, \ldots, f_m$ may be computed, and so all the possibilities for L may be determined effectively, as required.

The above lemma is relatively straightforward and it will suffice to sketch a proof. First we choose a unique representative divisor for each class in $\mathscr{J}$, as in the construction of the Jacobian variety (4). For, if $\mathfrak{A}$ is any divisor class in $\mathscr{J}$, then by the Riemann–Roch theorem there is a unique divisor $\mathfrak{a}$ in $\mathfrak{A}$ such that $\mathfrak{a}$ has a single pole, at v_0, of minimum order; furthermore, this order, s say, is at most g. If $\mathfrak{c}$ is any divisor on K, let $l(\mathfrak{c})$ denote the dimension of the space $V(\mathfrak{c})$ of elements f of K with $f = 0$ or $(f) + \mathfrak{c} \geqslant 0$. We may now express the conditions on the divisor $\mathfrak{a}$ of degree zero by

$$\mathfrak{a} \geqslant -s\mathfrak{v}_0, \quad l(s\mathfrak{v}_0 - \mathfrak{v}_0 - \mathfrak{b} - \mathfrak{a}) = 0, \quad l(-\mathfrak{b} - 2\mathfrak{a}) = 1, \tag{11}$$

where s is some integer, $0 \leqslant s \leqslant g$, and $\mathfrak{v}_0$ denotes the divisor of degree 1 with a single non-zero component, at v_0. Our object is to determine effectively the p solutions of (11). We proceed to show that the conditions in (11) may be expressed as a system of polynomial equations and inequalities over k. Such a system may be converted into a system of equations alone by the introduction of an extra variable y_0 and an extra equation $y_0 Q_1 \ldots Q_u = 1$, where $Q_1 \neq 0, \ldots, Q_u \neq 0$ are the inequalities to be satisfied. Since we know that this augmented system of equations possesses only finitely many solutions, it may be solved effectively, using the method of successive elimination by the resultant polynomial, as described in (10). The proof of the lemma will then be complete.

We shall suppose that $s \geqslant 1$ as the case $s = 0$ is trivial. Let us first assume that the components of $\mathfrak{a}$ are equal to 1 at valuations $v_1, \ldots, v_s$, $-s$ at v_0, and 0 otherwise; we wish to determine $v_1, \ldots, v_s$. As in § 2, we denote by y a primitive element of $K/k(z)$, and we may in fact choose y in $\mathscr{O}$. Let P then denote the minimal polynomial of y over $k[z]$, and let D denote the discriminant of P. From the proof (5) of the lemma in § 2, we note that, whenever f lies in $\mathscr{O}$, then Df lies in $k[y, z]$. We shall further assume that each v_i, $1 \leqslant i \leqslant s$, is not in $\mathscr{S}$, and satisfies $|D|_{v_i} = 1$. In this case each v_i is specified uniquely by y_i and z_i in k such that $|y - y_i|_{v_i} < 1$ and $|z - z_i|_{v_i} < 1$. It is the $2s$ unknowns $y_1, \ldots, y_s, z_1, \ldots, z_s$ which are determined by a system of polynomial equations and

https://doi.org/10.1017/S0305004100060497 Published online by Cambridge University Press

inequalities, which we now exhibit. The assumptions on $v_1, \ldots, v_s$ may be expressed as

$$P(y_i, z_i) = 0, \quad D\theta(y_i, z_i) \neq 0 \quad (1 \leqslant i \leqslant s),$$

where $\theta = \Pi_{j=2}^{n}(\alpha_1 - \alpha_j)$. Furthermore, $v_1, \ldots, v_s$ are distinct, and so

$$z_i \neq z_j \quad \text{or} \quad y_i \neq y_j \quad (1 \leqslant i < j \leqslant s).$$

It remains to express the two equations in (11) as polynomial equations and inequalities in $y_1, \ldots, z_s$. Using the same method as that of the proof (5) of the lemma in § 2, a basis $H_1, \ldots, H_t$ of $V(2s\mathfrak{v}_0 - \mathfrak{b})$ may be computed; since $\mathfrak{b}$ has only one pole, at the infinite valuation v_0, it follows that each H_i lies in $\mathcal{O}$. Since Dy' is a polynomial in y and z, we conclude that both DH_i and D^2H_i' lie in $k[y, z]$ for $i = 1, \ldots, t$. If we denote by $f(v)$ the constant coefficient of the Puiseux expansion of f at v, then f lies in $V(-\mathfrak{b} - 2\mathfrak{a})$ if and only if $f = \Sigma_{i=1}^{t} \pi_i H_i$ for some $\pi_1, \ldots, \pi_t$ in k with $f(v_j) = f'(v_j) = 0, 1 \leqslant j \leqslant s$. Since the polynomial D does not vanish at any v_j, these conditions are equivalent to the equations $\Sigma_{i=1}^{t} \pi_i (DH_i)(y_j, z_j) = \Sigma_{i=1}^{t} \pi_i (D^2H_i')(y_j, z_j) = 0$ for $1 \leqslant j \leqslant s$. The condition $l(-\mathfrak{b} - 2\mathfrak{a}) = 1$ is therefore equivalent to the condition that the $t \times 2s$ matrix with entries $DH_i(y_j, z_j)$ and $D^2H_i'(y_j, z_j)$, $1 \leqslant i \leqslant t$, $1 \leqslant j \leqslant s$, has rank $t - 1$. Since each entry in this matrix is a polynomial in y_j and z_j, the condition $l(-\mathfrak{b} - 2\mathfrak{a}) = 1$ may be expressed as a system of polynomial equations and inequalities in $y_1, \ldots, z_s$. Similarly the condition $l(s\mathfrak{v}_0 - \mathfrak{v}_0 - \mathfrak{b} - \mathfrak{a}) = 0$ may be so expressed; hence the proof of the lemma is complete in this case.

It remains to deal with a general divisor $\mathfrak{a}$ satisfying (11). It follows directly from (11) that $\mathfrak{a}$ has a single pole, at v_0, of order s, so we may suppose that $\mathfrak{a}$ has positive components $s_1, \ldots, s_q$ at valuations $v_1, \ldots, v_q$ respectively, with all other positive components occurring at valuations v either in $\mathcal{S}$, or such that $|D|_v < 1$. Since the sum of the positive components is s, there are only a finite number of possibilities for the integers $s_1, \ldots, s_q$ and for the other positive components, so we may sppose that these are fixed, and it only remains to determine the valuations $v_1, \ldots, v_q$. As before, our object is to express the conditions on $v_1, \ldots, v_q$ as a system of polynomial equations and inequalities in the $2q$ unknowns $y_1, \ldots, y_q, z_1, \ldots, z_q$. The condition $l(-\mathfrak{b} - 2\mathfrak{a}) = 1$ is again equivalent to a condition on the rank of a certain matrix, having entries of the form $D^{r_j+1} G_i^{(r_j)}(y_j, z_j)$, $0 \leqslant r_j < 2s_j$, $1 \leqslant j \leqslant q$, where $f^{(n)}$ denotes the nth derivative of f. Since these entries are polynomials in y_j and z_j, the condition is equivalent to a system of polynomial equations and inequalities, as required. Similarly the condition $l(s\mathfrak{v}_0 - \mathfrak{v}_0 - \mathfrak{b} - \mathfrak{a}) = 0$ may be so treated, and hence the proof of the lemma is complete.

It should be remarked that the above lemma is in principle unnecessary, since it is possible to deduce Theorem 1 directly from Theorem 2 and the bound for $H(X)$ in § 5. For, if $\omega_1, \ldots, \omega_d$ is a $k[z]$-basis of $\mathcal{O}$, then we may express any solution X, Y in $\mathcal{O}$ of (1) in the form

$$X = \sum_{j=0}^{J} \sum_{i=1}^{d} a_{ij} z^j \omega_i, \quad Y = \sum_{j=0}^{J} \sum_{i=1}^{d} b_{ij} z^j \omega_i,$$

where the a_{ij}, b_{ij} are elements of k. In view of the bound for $H(X)$ in § 5, J is bounded. Furthermore, the equation (1) is equivalent to a system of polynomial equations in the a_{ij} and b_{ij}. Provided that (1) possesses only finitely many solutions in $\mathcal{O}$, these may now be determined by successive elimination; when (1) possesses infinitely many solutions, according to Theorem 2, the exceptional solutions may be found by the

https://doi.org/10.1017/S0305004100060497 Published online by Cambridge University Press

imposition of the inequality $((X-\alpha_1)/(\alpha_2-\alpha_1))' \neq 0$. However, this algorithm for the solution of the hyperelliptic equation is far less efficient than that described earlier. The complexity of the method of successive elimination varies exponentially with the number of variables involved. In our earlier argument the number of variables is at most $2g$, which is much smaller than the $2d(J+1)$ variables a_{ij}, b_{ij}. As an illustration of the complexity of the method just advanced, the solution of the example in the next section would involve some 169 polynomial equations in 142 variables.

7. *Example.* In this section we shall solve the elliptic equation

$$Y^2 = X(X-z)(X-2z+1)$$

for X, Y in $k[z]$. This will give a simple illustration of the general method of analysis described in §4. It will be seen that an exactly similar argument would apply to any equation of the form $Y^2 = G(X)$, where G is square-free and has at least three roots in $k(z)$; the number of steps involved in the calculation varies exponentially with the height of G.

We shall assume throughout that Y is non-zero. First we shall determine all the possibilities for the field L, generated over $k(z)$ by the square roots of X, $X-z$ and $X-2z+1$. If a is any finite point of k such that $|X|_a < 1$, then either $a = 0, \frac{1}{2}$, or $|X|_a = |Y|_a^2$. We conclude that one of X, X/z, $X/(2z-1)$, $X/z(2z-1)$ is a square in $k[z]$. A similar conclusion obtains for the factors $X-z$ and $X-2z+1$; we deduce that L has at most 15 possibilities.

Let us first suppose that $X = h_1^2 z(2z-1)$, for some h_1 in $k[z]$. From degree parity arguments it follows that there exist h_2 and h_3 in $k[z]$ such that $X-z = h_2^2 z(z-1)$ and $X-2z+1 = h_3^2(2z-1)(z-1)$. In this case L is an extension of $k(z)$ of degree 4, generated by the square roots of $z(2z-1)$ and $z(z-1)$. The only valuations on L which are ramified are those $v|a$ with $a = 0$, $\frac{1}{2}$ or 1, and all these have $e_v = 2$; it follows from (3) that L has genus 0. We wish to determine, as in §4, all the possibilities for β_1, β_2 and β_3, where $\beta_1 = h_2\sqrt{[z(z-1)]} - h_3\sqrt{[(2z-1)(z-1)]}$, and β_2, β_3 are defined similarly. Since $\beta_1\hat{\beta}_1 = z-1$, and h_2, h_3 lie in $k[z]$, we deduce that, if v is a finite valuation on L, then $|\beta_1|_v = |\hat{\beta}_1|_v = 1$, or $1/e$, according as $|z-1|_v = 1$ or $|z-1|_v < 1$. Similar deductions may be made for the values of β_2 and β_3 at the finite valuations on L. It remains to investigate the four infinite valuations, which we denote by $|x|_i$, $1 \leqslant i \leqslant 4$. We define integers $b_{ij} = \log|\beta_i|_j$ for $1 \leqslant i \leqslant 3$, $1 \leqslant j \leqslant 4$; we wish to determine all the possibilities for the 3×4 matrix $\mathbf{B}$ with entries b_{ij}. Applying the fundamental inequality (8) to $\beta_1+\beta_2+\beta_3 = 0$, we obtain

$$\sum_{j=1}^{4} \max(0, b_{pj}-b_{qj}) \leqslant 6 \quad (1 \leqslant p \neq q \leqslant 3), \tag{12}$$

and so $|b_{pj}-b_{qj}| \leqslant 6$. Furthermore, $\log|\hat{\beta}_i|_j = 1-b_{ij}$, so, applying (8) to each of the equations in (7), we obtain $|b_{pj}+b_{qj}| \leqslant 7$, and so $|b_{ij}| \leqslant 6$ for each i, j. Hence there are only finitely many possibilities for $\mathbf{B}$. However, the number of possibilities may always be reduced by some further elementary considerations. For each j, choose indices p, q, r such that $b_{pj} \leqslant b_{qj} \leqslant b_{rj}$; it follows from $\beta_1+\beta_2+\beta_3 = 0$ that

$$|\beta_r|_j \leqslant \max(|\beta_p|_j, |\beta_q|_j), \quad \text{and so} \quad b_{qj} = b_{rj}.$$

Applying this argument to (7), we deduce that either

$$b_{pj} = b_{qj} = b_{rj} \leqslant 0, \quad \text{or} \quad b_{qj} = b_{rj} = 1-b_{pj} \geqslant 1.$$

https://doi.org/10.1017/S0305004100060497 Published online by Cambridge University Press

Furthermore, since $\Pi_{v\nmid\infty}|\beta_i|_v = 1/e^2$ for $i = 1, 2, 3$, the product formula (2) yields $\Sigma_{j=1}^4 b_{ij} = 2$. It now follows that the only possibilities for $\mathbf{B}$ are the matrices

$$\mathbf{A}(n) = \begin{pmatrix} -n & -n & 1+n & 1+n \\ -n & 1+n & -n & 1+n \\ -n & 1+n & 1+n & -n \end{pmatrix} \quad (n \geqslant 0),$$

or one of the matrices obtained from $\mathbf{A}(n)$ by permuting the columns. Taking into account the choice made in the labelling of the infinite valuations on L, and the fact that h_1, h_2, h_3 lie in $k[z]$, we deduce that only the matrices $\mathbf{A}(n)$ need be considered. In view of the inequality (12) above, we have $n \leqslant 2$. Now $\deg h_3 = \log|\beta_1 - \beta_1|_j - 1 = n$, so X has degree $2n+2$. If $n = 0$, then h_1, h_2 and h_3 are elements of k, and from the conditions $|\beta_1|_1 = |\beta_2|_1 = 1$, we obtain the linear equations $h_2 - \sqrt{2}\, h_3 = h_3 - h_1 = 0$; from (7) we then obtain $h_3^2 = 1$, and this yields the solutions $X = z(2z-1)$. If $n = 1$, then $h_3 = a + bz$ for some a, b in k with b non-zero. From the condition $|\beta_1|_1 = 1/e$ we obtain $8a + b = 0$, and from $|\beta_2|_1 = 1/e$ we obtain $4a + b = 0$, which are incompatible. Similarly the case $n = 2$ may be excluded.

Now let us assume that one of X, X/z, $X/(2z-1)$ is a square in $k[z]$. In the first case, since $(X-z)(X-2z+1) = Y^2/X$, it is a square in $k[z]$, T^2 say. We have

$$\left(X - \frac{3z-1}{2} + T\right)\left(X - \frac{3z-1}{2} - T\right) = \frac{(z-1)^2}{4}. \tag{13}$$

Both factors on the left are polynomials, so X has degree at most 2. If X lies in k, then, since $(z-X)(2z-1-X)$ is a square, $X = 1$. If X has degree 2, then, for some κ in k, the two factors on the left of (13) are equal to $\frac{1}{4}\kappa$ and $(z-1)^2/\kappa$; since X is a square we deduce $\kappa = 1$ and $X = \frac{1}{2}(z+\frac{1}{2})^2$. In the other cases, when X/z or $X/(2z-1)$ is a square in $k[z]$, then (13) holds with T in $\mathcal{O}_K$, where $K = k(\sqrt{z})$ or $k(\sqrt{(2z-1)})$ respectively. In both cases K has only one infinite valuation, and so we deduce from (13) that X has degree at most 2. Thus if X/z is a square, $X = \kappa z$, for some κ in k, whence $\kappa = 2$, whilst, if $X/(2z-1)$ is a square, $X = \kappa(2z-1)$ for some κ in k, whence $\kappa = \frac{1}{2}$.

We conclude that the solutions X, Y in $k[z]$ of $Y^2 = X(X-z)(X-2z+1)$ with $Y \neq 0$ are

$$X = z(2z-1),\ Y = \pm\sqrt{2}\, z(z-1)(2z-1); \quad X = 1,\ Y = \pm\sqrt{2}\,(z-1);$$

$$X = \tfrac{1}{2}(z+\tfrac{1}{2})^2,\ Y = \pm\tfrac{1}{4}\sqrt{2}\,(z^2-\tfrac{1}{4})(z-\tfrac{3}{2}); \quad X = 2z,\ Y = \pm\sqrt{2}\, z;$$

$$X = z - \tfrac{1}{2},\ Y = \pm\tfrac{1}{2}\sqrt{2}\,(z-\tfrac{1}{2}).$$

I would like to express my thanks, to Professor Baker for his helpful advice on the presentation of this paper, and to the referee for several useful comments.

REFERENCES

(1) Baker, A. Bounds for the solutions of the hyperelliptic equation. *Proc. Cambridge Philos. Soc.* **65** (1969), 439–444.

(2) Baker, A. *Transcendental number theory* (Cambridge, 1975).

(3) Eichler, M. *Introduction to the theory of algebraic numbers and functions* (New York, Academic Press, 1966).

(4) Lang, S. *Abelian varieties* (New York, Interscience, 1959).

https://doi.org/10.1017/S0305004100060497 Published online by Cambridge University Press

(5) MASON, R. C. On Thue's equation over function fields. *J. London Math. Soc. Ser.* 2, **24** (1981), 414–426.
(6) MUMFORD, D. *Abelian varieties* (India, Oxford University Press, 1970).
(7) SAMUEL, P. *Lectures on old and new results on algebraic curves* (Bombay, 1966).
(8) SCHMIDT, W. M. Thue's equation over function fields. *J. Austral. Math. Soc. Ser. A* **25** (1978), 385–422.
(9) SIEGEL, C. L. The integer solutions of the equation $y^2 = ax^n + bx^{n-1} + \ldots + k$. *J. London Math. Soc.* **1** (1926), 66–68 (under the pseudonym X).
(10) VAN DER WAERDEN, B. L. *Moderne Algebra*, vol. II (Berlin, Springer, 1931).

https://doi.org/10.1017/S0305004100060497 Published online by Cambridge University Press