PACIFIC JOURNAL OF MATHEMATICS
Vol. 36, No. 1, 1971

AN ELEMENTARY PROOF OF THE RIEMANN HYPOTHESIS FOR AN ELLIPTIC CURVE OVER A FINITE FIELD

HORST G. ZIMMER

Let K be an elliptic function field over a finite field of constants k. This paper is aimed at presenting a valuation-theoretic proof of the analogue of Riemann's hypothesis for the zeta-function of K.

More precisely, K is regarded as the function field of the plane elliptic curve over k defined by the nonhomogeneous equation in two variables x, y

$$(1) \qquad \varphi(x, y) = y^2 + (a_0x + a_1)y + x^3 + b_1x^2 + b_2x + b_3 = 0$$

with coefficients $a_i, b_j \in k$ and nonzero discriminant D. Let p denote the characteristic of k, $q = p^r$ be the number of elements in k, and N the number of distinct solutions (ξ, η) in k of the equation (1). The analogue of the Riemann "hypothesis" for the elliptic curve (1) over k may be stated as the following (see [1], Chap. V, § 5)

THEOREM.

$$(2) \qquad |N - q| \leqq 2\sqrt{q} \ .$$

This theorem was, in several important cases, first proven by Hasse [3]. In the meantime various proofs and generalizations of it have been invented by the same author, A. Weil, P. Roquette and others. More recently, Manin [6], using ideas of Hasse's [4], gave an entirely elementary proof of the theorem under the supposition that the characteristic p of k is greater than 3. Elistratov [2] showed in a subsequent paper that Manin's argument carries over to the case of characteristic $p = 3$. Most of these proofs have in common that the characteristic p is presupposed to be $\neq 2, 3$ (or at least $\neq 2$) which permits one to assume that $a_0 = a_1 = b_1 = 0$ (or at least $a_0 = a_1 = 0$) in equation (1).

In the present paper, we give for all finite characteristics p a unified elementary proof of Riemann's hypothesis. Our method is closely related to that of Manin but, as opposed to it, brings valuation theory into play. This way our argumentation, on the one hand, avoids some of the computations which appear to be inevitable in Manin's proof and, on the other hand, circumvents a difficulty occurring in his reasoning (cf. MR [6], [2]). Altogether the valua-

tion-theoretic approach yields an explicit and perspicuous proof of the theorem.

Similar to Hasse's original argument, the truth of the inequality (2) will be inferred from the more general fact that a certain quadratic form is positive semi-definite. The setup of this paper has the advantage that it can be generalized to a proof of the positive semi-definiteness of a corresponding quadratic form in the case of an elliptic curve (1) defined over an arbitrary algebraic function field K in several variables over any field of constants k. This will be carried out in a different context in a subsequent paper.[1]

2. **A quadratic form.** First we build up the usual system $\mathfrak{S} = \{\mathfrak{p}\}$ of prime divisors $\mathfrak{p}$ of the function field K/k with respect to x. To this end we distinguish in K/k the rational function field $k(x)/k$. Let $\mathfrak{s} = \{\mathfrak{P}\}$ denote the system of prime divisors of $k(x)/k$ given by the prime polynomials and the "infinite" prime of $k(x)$. The system $\mathfrak{S} = \{\mathfrak{P}\}$ of K/k is then obtained by expanding the system $\mathfrak{s} = \{\mathfrak{P}\}$ of $k(x)/k$ to K/k in the familiar manner [1]. Observe that K is a finite algebraic extension of $k(x)$ of degree 2. We denote by $w_{\mathfrak{P}}$, $w_{\mathfrak{p}}$ the discrete valuations of $k(x)/k$, K/k respectively associated with $\mathfrak{P} \in \mathfrak{s}$, $\mathfrak{p} \in \mathfrak{S}$ and normalize each $w_{\mathfrak{p}}$ such that it attains the least positive value 1.

Each of the valuations $w_{\mathfrak{p}}$ of K/k with $\mathfrak{p} \in \mathfrak{S}$ satisfies the sharp inequality

$$w_{\mathfrak{p}}(z_1 + z_2) \leqq \min\{w_{\mathfrak{p}}(z_1), w_{\mathfrak{p}}(z_2)\} \qquad (z_1, z_2 \in K) \tag{3}$$

with the equality sign when $w_{\mathfrak{p}}(z_1) \neq w_{\mathfrak{p}}(z_2)$. Here the element $0 \in K$ is comprised by putting formally $w_{\mathfrak{p}}(0) = \infty$.

The system $\mathfrak{S}$ has the property that, for any given $0 \neq z \in K$,

$$w_{\mathfrak{p}}(z) \neq 0 \text{ only for finitely many } \mathfrak{p} \in \mathfrak{S}. \tag{4}$$

Furthermore, for $\mathfrak{S}$ there holds the product formula which we preferably write in the additive shape

$$\sum_{\mathfrak{p} \in \mathfrak{S}} f_{\mathfrak{p}} w_{\mathfrak{p}}(z) = 0 \qquad (0 \neq z \in K) \tag{5}$$

with the absolute residue class degrees $f_{\mathfrak{p}}$ of $\mathfrak{p}$ as multiplicities [1].

Now we form the algebraically independent *composite* over k of the elliptic function field K with itself, i.e., the elliptic function field $E = K(X, Y)$ over K as field of constants generated by the non-homogeneous equation in two variables X, Y over K

[1] This paper entitled "Die Néron-Tate'schen quadratischen Formen auf der rationalen Punktgruppe einer elliptischen Kurve" is to appear in the Journal of Number Theory.

$$(6)\qquad \varphi(X, Y) = Y^2 + (a_0X + a_1)Y + X^3 + b_1X^2 + b_2X + b_3 = 0$$

with the same coefficients a_i, b_j as in (1).

The rational points $P = (x_P, y_P)$ of the elliptic curve over K defined by (6), that is to say those points P with coordinates x_P, $y_P \in K$, together with the "zero point" $\mathcal{O} = (\infty, \infty)$ make up an (additive) abelian group $\mathfrak{E}$ under the following group operation [1].

For $P = (x_P, y_P)$, $Q = (x_Q, y_Q) \in \mathfrak{E}$ the sum $P + Q = (x_{P+Q}, y_{P+Q})$ is defined by (6) and

$$(7)\qquad x_{P+Q} = -(x_P + x_Q) - \left(\frac{y_P - y_Q}{x_P - x_Q}\right)^2 - a_0\left(\frac{y_P - y_Q}{x_P - x_Q}\right) - b_1$$

if $x_P \neq x_Q$, i.e., $P \neq \pm Q$, or

$$(8)\qquad x_{2P} = -2x_P - \left(\frac{\varphi_X(x_P, y_P)}{\varphi_Y(x_P, y_P)}\right)^2 + a_0\left(\frac{\varphi_X(x_P, y_P)}{\varphi_Y(x_P, y_P)}\right) - b_1$$

if $x_P = x_Q$ and $P = Q$, where φ_X, φ_Y stand for the partial derivatives of φ relative to X, Y respectively, so that

$$\varphi_X(x_P, y_P) = a_0y_P + 3x_P^2 + 2b_1x_P + b_2, \quad \varphi_Y(x_P, y_P) = 2y_P + a_0x_P + a_1 .$$

Observe that P and $-P$ have the same first coordinate $x_P = x_{-P}$.

We are now in a position to define a quadratic form d on the group of rational points $\mathfrak{E}$ of the curve (6) over K. Letting again $P = (x_P, y_P)$, we set

$$(9)\qquad d(P) = -\tfrac{1}{2}\sum_{P<0} f_{\mathfrak{p}} w_{\mathfrak{p}}(x_P) , \qquad (P \in \mathfrak{E})$$

where the shorthand notation "$P < 0$", to which we shall stick throughout in formulas involving (9), means that the summation is over all prime divisors $\mathfrak{p} \in \mathfrak{S}$ with $w_{\mathfrak{p}}(x_P) < 0$. For $P = \mathcal{O}$ we agree to put $d(P) = 0$. Notice that the condition (4) ensures that d is well-defined.

Our task consists first in showing that the function d defined by (9) is indeed a *quadratic form* on $\mathfrak{E}$, i.e., that d satisfies the condition

$$(10)\quad d(P + Q) + d(P - Q) = 2d(P) + 2d(Q) \quad \text{for any two } P, Q \in \mathfrak{E}$$

which, as one verifies by induction, is tantamount to

$$(11)\qquad d\left(\sum_{\lambda=1}^{s} n_\lambda P_\lambda\right) = \tfrac{1}{2}\sum_{\mu,\nu=1}^{s} n_\mu n_\nu \{d(P_\mu + P_\nu) - d(P_\mu) - d(P_\nu)\}$$

for any s points $P_1, \cdots, P_s \in \mathfrak{E}$ and rational integers $n_1, \cdots, n_s$.

We remark that, in order to prove (2), it would suffice to establish (10) for all integral multiples of two particular rational points. But our method will at once yield a proof of (10) for any two points of $\mathfrak{E}$.

For the proof of (10) we shall need two technical lemmas. The first one is due to E. Lutz [5]. It can in fact be enunciated for an elliptic curve (6) over any field K with an (additive) nonarchimedean valuation w provided that the coefficients of (6) enjoy the property $w(a_i) = 0$, $w(b_j) = 0$ whenever $a_i \neq 0$, $b_j \neq 0$ respectively. Let this be the case. To comprise the element $\mathcal{O} = (\infty, \infty)$ of $\mathcal{C}$ we put formally $w(\infty) = -\infty$.

LEMMA 1. *For any two rational points $P = (x_P, y_P)$, $Q = (x_Q, y_Q)$ of the curve* (6) *over K we have*:

(a) *The inequalities $w(x_P) < 0$, $w(x_Q) < 0$ imply*

$$w(x_{P+Q}) < 0$$

and, moreover,

$$w(x_{P+Q}) \leqq \max\{w(x_P), w(x_Q)\}$$

with the equality sign when $w(x_P) \neq w(x_Q)$.

(b) *The inequalities $w(x_P) < 0$, $w(x_Q) \geqq 0$ entail*

$$w(x_{P+Q}) \geqq 0 .$$

We note first that the statements corresponding to (a), (b) with $P - Q$ instead of $P + Q$ are also valid since $x_{-Q} = x_Q$.

Lemma 1 is obviously true if P, Q or $P + Q$ is the zero point $\mathcal{O}$ of $\mathcal{C}$. Thus, we may assume that none of the points P, Q or $P + Q$ is $\mathcal{O}$. Using the sharp inequality (3) for w and the addition formula (7) for $P, Q \in \mathcal{C}$, one shows then that

(α) the relations $w(x_P) < w(x_Q) < 0$ imply $w(x_{P+Q}) = w(x_Q)$, while

(β) $w(x_P) < 0 \leqq w(x_Q)$ entail $w(x_{P+Q}) \geqq 0$. This proves (b) and part of (a).

It remains to verify that the relations $w(x_P) = w(x_Q) < 0$ imply $w(x_{P+Q}) \leqq w(x_P)$. We may assume $w(x_{P+Q}) \neq w(x_Q)$ since otherwise the assertion is true. Then we must have $w(x_{P+Q}) < 0$ because the assumption $w(x_{P+Q}) \geqq 0$ would, on grounds of the decomposition $P = (P + Q) - Q$, according to (β) lead to the inequality $w(x_P) \geqq 0$ contradicting the premise $w(x_P) < 0$. But then the same decomposition, because of the relations $w(x_{P+Q}) < 0$, $w(x_Q) < 0$, $w(x_{P+Q}) \neq w(x_Q)$, yields by statement (α) that

$$w(x_P) = \max\{w(x_{P+Q}), w(x_Q)\} \geqq w(x_{P+Q})$$

which proves the remaining portion of assertion (a).

To state the second lemma we return to our original situation of an elliptic function field K/k defined by an equation (1) with coef-

ficients a_i, b_j in the constant field k of K. We denote by w any of the valuations $w_\mathfrak{p}$ of K/k with $\mathfrak{p} \in \mathscr{S}$.

LEMMA 2. *Let $P = (x_P, y_P)$, $Q = (x_Q, y_Q)$ be any two points of $\mathscr{C}$ such that $P \neq \pm Q$.* i.e., $x_P \neq x_Q$.

(a) *If $w(x_P) \geqq 0$, $w(x_Q) \geqq 0$, then the inequalities $w(x_{P-Q}) < 0$, $w(x_{P+Q}) < 0$ imply*

$$-\tfrac{1}{2}w(x_{P-Q}) - \tfrac{1}{2}w(x_{P+Q}) = w(x_P - x_Q) ,$$

while $w(x_{P\mp Q}) < 0$, $w(x_{P\pm Q}) \geqq 0$ entail respectively

$$-\tfrac{1}{2}w(x_{P\mp Q}) = w(x_P - x_Q) ,$$

and for $w(x_{P-Q}) \geqq 0$, $w(x_{Q+Q}) \geqq 0$ there holds

$$w(x_P - x_Q) = 0 .$$

(b) *If $w(x_P) < 0$, $w(x_Q) < 0$, then we have also $w(x_{P-Q}) < 0$, $w(x_{P+Q}) < 0$, and the relation*

$$\tfrac{1}{2}w(x_{P-Q}) + \tfrac{1}{2}w(x_{P+Q}) = w(x_P) + w(x_Q) - w(x_P - x_Q)$$

is valid.

For the proof of Lemma 2 we will need the following four identities which are immediate consequences of the equation (1) for the coordinates of P, Q.

$$y_P - y_{-Q} = (y_{-P} - y_{-Q}) + \varphi_Y(x_P, y_P) , \tag{12}$$

$$y_P - y_Q = -(y_{-P} - y_{-Q}) - a_0(x_P - x_Q) , \tag{13}$$

$$\begin{aligned}&(y_P - y_Q)(y_P - y_{-Q})\\ &= -(x_P - x_Q)(x_P^2 + x_P x_Q + x_Q^2 + b_1(x_P + x_Q) + b_2 + a_0 y_P) ,\end{aligned} \tag{14}$$

$$\begin{aligned}&x_P^2 + x_P x_Q + x_Q^2 + b_1(x_P + x_Q) + b_2 + a_0 y_P\\ &= \varphi_X(x_P, y_P) + (x_P - x_Q)\{(x_P - x_Q) - (3x_P + b_1)\} .\end{aligned} \tag{15}$$

Also we shall make continual use of the property (3) of w.

Now, employing the addition formula (7) and the relations (12)–(15) for $P, Q \in \mathscr{C}$, we prove assertion (a) by showing first that, under the assumptions $w(x_P) \geqq 0$, $w(x_Q) \geqq 0$, the inequalities $w(x_{P-Q}) < 0$, $w(x_{P+Q}) < 0$ imply

$$-\tfrac{1}{2}w(x_{P-Q}) - \tfrac{1}{2}w(x_{P+Q}) = w(x_P - x_Q) - w(\varphi_X(x_P, y_P)) , \tag{16}$$

while $w(x_{P\mp Q}) < 0$, $w(x_{P\pm Q}) \geqq 0$ entail respectively

$$-\tfrac{1}{2}w(x_{P\mp Q}) = w(x_P - x_Q) - w(\varphi_Y(x_P, y_P)) ,$$

and for arbitrary $w(x_{P-Q})$, $w(x_{P+Q})$ there holds obviously $w(x_P - x_Q) \geqq 0$.

Then one verifies by means of Lemma 1 and the addition formulas (7), (8) that in the first case

$$0 \leqq w(\varphi_X(x_P, y_P)) < w(\varphi_Y(x_P, y_P)) , \tag{17}$$

while in the second case

$$w(\varphi_X(x_P, y_P)) \geqq w(\varphi_Y(x_P, y_P)) \geqq 0 , \tag{18}$$

and thirdly, again by virtue of (12)-(15), for $w(x_{P-Q}) \geqq 0$, $w(x_{P+Q}) \geqq 0$ the inequalities

$$w(\varphi_X(x_P, y_P)), \quad w(\varphi_Y(x_P, y_P)) \geqq w(x_P - x_Q) \geqq 0 \tag{19}$$

are true.

Further, we observe that the discriminant D of the curve (6), since it is the resultant of φ_X, φ_Y, admits a representation in the shape

$$D = \chi(a_i, b_j, X, Y)\varphi_X(X, Y) + \psi(a_i, b_j, X, Y)\varphi_Y(X, Y) , \tag{20}$$

where χ and ψ are polynomials in a_i, b_j, X, Y with rational integral coefficients mod p. But the relation (20) remaining true upon replacing (X, Y) by the coordinates (x_P, y_P) of the point $P \in \mathscr{C}$ it follows that in the first case $w(\varphi_X(x_P, y_P)) = 0$, in the second case $w(\varphi_Y(x_P, y_P)) = 0$, and in the third case $w(x_P - x_Q) = 0$ since otherwise (20) with (x_P, y_P) in place of (X, Y) and, respectively, one of the identities (17), (18) or (19) would lead to the inequality $w(D) > 0$ which contradicts the fact that D is a nonzero element of the constant field k of K, i.e., that $w(D) = 0$. We have thereby utilized the assumption that the elliptic curve (6) is already defined over k which means that its coefficients a_i, b_j lie in k and a fortiori $0 \neq D \in k$. This proves (a).

The first part of assertion (b) and the second part of (b) under either additional assumption $w(x_P) < w(x_Q) < 0$ or $w(x_Q) < w(x_P) < 0$ are immediate consequences of Lemma 1, (a).

However, if $w(x_P) = w(x_Q) < 0$ one has to discuss the three distinct possibilities $w(x_{P-Q}) < w(x_Q)$, $w(x_{P+Q}) < w(x_Q)$ and $w(x_{P-Q}) = w(x_{P+Q}) = w(x_Q)$ separately.

If $w(x_{P-Q}) < w(x_Q)$ the assertion (b) can be proven by applying Lemma 1, (a), the addition formulas (7), (8) and either the identities (12), (13), when K has a characteristic $\neq 2$, or the relation (16), when the characteristic of K is 2. More precisely, one shows in the former case that

$$w(x_{P+Q}) = w(x_Q),\ w(x_{P-Q}) = 2w(y_P - y_{-Q}) - 2w(x_P - x_Q) \text{ and}$$
$$2w(y_P - y_{-Q}) = 3w(x_Q),$$

while in the latter case $w(x_{P+Q}) < w(x_Q)$ is valid such that relation (16) remains true here and

$$w(\varphi_X(x_P, y_P)) = 2w(x_P)\ .$$

The case $w(x_{P+Q}) < w(x_Q)$ can be treated similarly since $x_{-Q} = x_Q$. If $w(x_{P-Q}) = w(x_{P+Q}) = w(x_Q)$ it is enough to establish

$$w(x_P - x_Q) = w(x_Q)\ .$$

To this end one applies in succession the addition formula (7) and, according as the characteristic is 2 or $\neq 2$, the relations (14), (15) or (12), (13) to show that the supposition $w(x_P - x_Q) > w(x_Q)$ would lead to a contradiction. Notice that in the former case $w(\varphi_X(x_P, y_P) = 2w(x_P)$ while in the latter $2w(\varphi_Y(x_P, y_P)) = 3w(x_P)$.

We are now ready to prove the relation (10) for d with regard to any two points $P, Q \in \mathscr{C}$ subject to the restriction $P \neq \pm Q$. Applying in succession part (b) of Lemma 1, part (a) of Lemma 2, the product formula (5) for $\mathscr{S}$ and the inequality (3) for $w_\mathfrak{p}$ with $\mathfrak{p} \in \mathscr{S}$, we obtain according to the definition (9) of d (in the notation introduced by (9)):

$$\begin{aligned}
&d(P - Q) + d(P + Q) + \tfrac{1}{2}\sum_{P,Q,P-Q<0} f_\mathfrak{p} w_\mathfrak{p}(x_{P-Q}) + \tfrac{1}{2}\sum_{P,Q,P+Q<0} f_\mathfrak{p} w_\mathfrak{p}(x_{P+Q}) \\
&= -\tfrac{1}{2}\sum_{P-Q<0\leqq P,Q} f_\mathfrak{p} w_\mathfrak{p}(x_{P-Q}) - \tfrac{1}{2}\sum_{P+Q<0\leqq P,Q} f_\mathfrak{p} w_\mathfrak{p}(x_{P+Q}) \\
&= \sum_{P,Q\geqq 0} f_\mathfrak{p} w_\mathfrak{p}(x_P - x_Q) \\
&= -\sum_{P<0\leqq Q} f_\mathfrak{p} w_\mathfrak{p}(x_P - x_Q) - \sum_{Q<0\leqq P} f_\mathfrak{p} w_\mathfrak{p}(x_P - x_Q) - \sum_{P,Q<0} f_\mathfrak{p} w_\mathfrak{p}(x_P - x_Q) \\
&= 2d(P) + 2d(Q) + \sum_{P,Q<0} f_\mathfrak{p} w_\mathfrak{p}(x_P) + \sum_{P,Q<0} f_\mathfrak{p} w_\mathfrak{p}(x_Q) - \sum_{P,Q<0} f_\mathfrak{p} w_\mathfrak{p}(x_P - x_Q)\ .
\end{aligned}$$

But in comparing the first with the last portion of this sequence of identities one recognizes by means of part (b) of Lemma 2 that (10) is valid for $P \neq \pm Q$.

If $P = Q$ or $-Q$, (10) can be established in a similar way, applying the addition formula (8) in place of (7).

As already pointed out at the beginning, the asserted inequality (2) turns out to be a consequence of the positive semi-definiteness of the quadratic form d on $\mathscr{C}$. We say that d is *positive semidefinite* on $\mathscr{C}$ if, for any two rational points P, Q of the curve (6) over K, the quadratic form $\delta_{P,Q}$, defined for rational integers m, n by setting according to (11)

$$\begin{aligned}\delta_{P,Q}(m, n) &= d(mP + nQ) \\ &= d(P)m^2 + \{d(P + Q) - d(P) - d(Q)\}mn + d(Q)n^2 ,\end{aligned} \tag{21}$$

has a discriminant

$$\Delta_{P,Q} = \{d(P + Q) - d(P) - d(Q)\}^2 - 4d(P)d(Q) \leqq 0 . \tag{22}$$

In fact the following general lemma can be proved.

LEMMA 3. *Let $\delta(u, v) = \alpha u^2 + \beta uv + \gamma v^2$ be a quadratic form in two real variables u, v with real coefficients α, β, γ such that δ satisfies the inequalities*

$$\delta(m\sigma, n\tau) \geqq 0 \tag{23}$$

for two fixed real numbers $\sigma, \tau \neq 0$ and all rational integers m, n. Then the discriminant $\Delta = \beta^2 - 4\alpha\gamma$ of δ is less than or equal to zero.

Proof. If $\alpha = 0$ or $\gamma = 0$, then (23) implies $\beta = 0$, i.e., $\Delta = 0$.

Thus, let $\alpha \neq 0$. The discriminant of the polynomial in u $\delta_n(u) = \delta(u, n\tau)$ is

$$\Delta_n = n^2\tau^2\Delta = \alpha^2(\rho_n - \omega_n)^2 , \tag{24}$$

where ρ_n, ω_n are the roots of $\delta_n(u)$ in the complex number field.

Suppose, by way of contradiction, that $\Delta > 0$. Then, because of (24), also the inequalities $\Delta_n > 0$ are fulfilled for all $n \neq 0$, whence $\rho_n \neq \omega_n$ are real roots of $\delta_n(u)$.

Let $\rho_n < \omega_n$, say. The assumption (23) implies for $n \neq 0$ that there are rational integers κ_n such that the estimates

$$\kappa_n\sigma \leqq \rho_n < \omega_n \leqq (\kappa_n + 1)\sigma \text{ or } (\kappa_n - 1)\sigma ,$$

i.e.,

$$0 < (\rho_n - \omega_n)^2 \leqq \sigma^2 \tag{25}$$

hold. However, according to (24) and on account of the assumption $\Delta > 0$, the discriminants Δ_n become arbitrarily large, as $n \to \infty$, which contradicts the estimates (25). Thus, we have proved $\Delta \leqq 0$.

Application of Lemma 3 with $\sigma = \tau = 1$ to the quadratic form (21) yields the positive semi-definiteness of d on $\mathscr{C}$. Note that the conditions (23) with $\sigma = \tau = 1$ are fulfilled for $\delta_{P,Q}$ since, by definition (9) of d, $d(mP + nQ) \geqq 0$.

We summarize our results in a

PROPOSITION. *The function d, defined by* (9), *on the group $\mathscr{C}$*

of rational points of the elliptic curve (6) *over* K *is a positive semi-definite quadratic form on* $\mathscr{C}$.

3. Proof of the theorm. Now we are in a position to prove the theorem.

We pick two distinct rational points of $\mathscr{C}$, namely $P_0 = (x, y)$ and its image $Q_0 = (x^q, y^q)$ under the so-called Frobenius endomorphism of $\mathscr{C}$ which consists of raising the coordinates of P_0 to the q-th power. The proof of the theorem then amounts to verifying the following

LEMMA 4.

$$d(P_0) = 1, \quad d(Q_0) = q, \quad d(P_0 + Q_0) - d(P_0) - d(Q_0) = q - N .$$

For assuming Lemma 4 to be true one realizes immediately that the inequality (22) with $P = P_0, Q = Q_0$ is equivalent to the assertion (2).

It remains to prove Lemma 4.

We first observe that the "infinite" prime divisor $\wp_\infty$ of $k(x)/k$ is the only one with $w_{\mathfrak{p}_\infty}(x) < 0$, such that the extensions $\mathfrak{p}_\infty$ of $\wp_\infty$ to the elliptic function field $K = k(x, y)/k$ are the only prime divisors of K/k with $w_{\mathfrak{p}_\infty}(x) < 0$. But p_∞ admits exactly one extension $\mathfrak{p}_\infty$ to K/k since it is ramified. This is because the equation (1) shows according to property (3) that

$$2w_{\mathfrak{p}_\infty}(y) = e_{\mathfrak{p}_\infty} w_{\wp_\infty}(x^3 + b_1x^2 + b_2x + b_3) = -3e_{\mathfrak{p}_\infty}$$

so that the relative ramification index $e_{\mathfrak{p}_\infty}$ of $\mathfrak{p}_\infty$ over $\wp_\infty$ is an even number less than or equal to the field degree $[K: k(x)] = 2$, i.e., $e_{\mathfrak{p}_\infty} = 2$. The absolute degree of $\mathfrak{p}_\infty$ is $f_{\mathfrak{p}_\infty} = 1$. By definition (9) of d we have therefore

$$d(P_0) = -\tfrac{1}{2} f_{\mathfrak{p}_\infty} w_{\mathfrak{p}_\infty}(x) = 1$$

and, similarly, $d(Q_0) = q$.

In order to compute $d(P_0 + Q_0)$ we shall consider the group operation formula (7) on $\mathscr{C}$ for $P = P_0, Q = Q_0$. Let $P_0 + Q_0 = (x_0, y_0)$. Formula (7) shows that $w_{\mathfrak{p}}(x_0) < 0$ for $\mathfrak{p} \neq \mathfrak{p}_\infty$ can happen if and only if $w_{\mathfrak{p}}(x - x^q) > 0$.

REMARK. To every solution (ξ, η) in k of the elliptic equation (1) there corresponds exactly one prime divisor $\mathfrak{p} \neq \mathfrak{p}_\infty$ of degree $f_{\mathfrak{p}} = 1$ of K/k with $\xi, \eta =. \ x, y \bmod \mathfrak{p}$, and vice versa. Furthermore, for each of those (finitely many) $\mathfrak{p}$'s there holds

$$w_{\mathfrak{p}}(x - x^q) = e_{\mathfrak{p}} > 0 , \tag{26}$$

where $e_{\mathfrak{p}}$ denotes the ramification index of $\mathfrak{p}$ relative to $K/k(x)$ (cf. [1], Chap. IV, § 2).

The statement (26) is true on grounds of the decomposition

$$x^q - x = \prod_{\zeta \in k} (x - \zeta) \tag{27}$$

because of which we have $w_{\mathfrak{p}}(x^q - x) = e_{\mathfrak{p}} w_{\wp}(x^q - x) = e_{\mathfrak{p}}$ where $\wp$, as before, denotes the restriction of $\mathfrak{p}$ on $k(x)/k$.

At first we discuss the prime divisors $\mathfrak{p} \neq \mathfrak{p}_\infty$ of K/k with the property $w_{\mathfrak{p}}(x - x^q) > 0$. According to (27), these $\mathfrak{p}$'s have at most the degree $f_{\mathfrak{p}} = 2$, and their ramification index $e_{\mathfrak{p}}$ relative to $\wp$ does not exceed $[K: k(x)] = 2$.

We distinguish accordingly between three types of such prime divisors $\mathfrak{p} \in \mathscr{S}$.

(i) $e_{\mathfrak{p}} = 2, f_{\mathfrak{p}} = 1$. Then the restriction $\wp$ of $\mathfrak{p}$ on $k(x)$ is *ramified* in K.

In this case we have $w_{\mathfrak{p}}(y - y^q) = 1$. For $\mathfrak{p}$ necessarily divides the different $\varphi_Y(x, y)$ of the primitive element y for $K/k(x)$, and therefore $\mathfrak{p}$ is not a divisor of $\varphi_X(x, y)$ since, according to (20), $w_{\mathfrak{p}}(\varphi_Y(x, y)) > 0$ implies.

$$w_{\mathfrak{p}}(\varphi_X(x, y)) = 0 \ .$$

Hence, the two relations (14), (15) with $P = P_0, Q = Q_0$, i.e.,

$$\begin{aligned} &(y - y^q)\{\varphi_Y(x, y)^q + (y - y^q)\} \\ &= -(x - x^q)\{x^2 + xx^q + x^{2q} + b_1(x + x^q) + b_2 + a_0 y\} \ , \end{aligned}$$

$$\begin{aligned} &x^2 + xx^q + x^{2q} + b_1(x + x^q) + b_2 + a_0 y \\ &= \varphi_X(x, y) + (x - x^q)\{(x - x^q) - (3x + b_1)\} \ , \end{aligned}$$

together with (26) yield by the aid of (3) $w_{\mathfrak{p}}(y - y^q) = 1$.

We remark that $w_{\mathfrak{p}}(1 - y^{q-1}) = 1$ is impossible for ramification prime divisors $\mathfrak{p}$ of degree 1 of K/k since otherwise the corresponding restrictions $\wp$ on $k(x)/k$ would possess two distinct extensions to K/k. Hence $w_{\mathfrak{p}}(1 - y^{q-1}) = 0$ and $w_{\mathfrak{p}}(y) = 1$.

The formula (7) for the group addition in $\mathscr{E}$ reveals now that these ramification divisors $\mathfrak{p}$, because of (26), contribute the values

$$w_{\mathfrak{p}}(x_0) = 2w_{\mathfrak{p}}(y - y^q) - 2w_{\mathfrak{p}}(x - x^q) = 2 - 4 = -2$$

to $d(P_0 + Q_0)$.

Let M denote the number and $\mathscr{R} \subset \mathscr{S}$ the set of ramification prime divisors $\mathfrak{p} \neq \mathfrak{p}_\infty$ of degree 1 of K/k. In other words, M is the number of solutions in k of the elliptic equation (1) corresponding to those $\mathfrak{p} \in \mathscr{R}$.

(ii) $e_\mathfrak{p} = 1, f_\mathfrak{p} = 1$. Then the restriction $\wp$ of $\mathfrak{p}$ on $k(x)$ is *decomposed* in K.

In this case we have $w_\mathfrak{p}(y) = 0$ because otherwise $\wp$ would admit but one extension $\mathfrak{p}$ to K. Also the inequality $w_\mathfrak{p}(1 - y^{q-1}) > 0$ holds since $f_\mathfrak{p} = 1$ means that the residue class field of $\mathfrak{p}$ is k itself so that the element $\eta = .\ y \bmod \mathfrak{p}$ lies already in k, i.e., η satisfies $\eta \neq 0$, $\eta^{q-1} = 1$.

But then the formula (7), in virtue of (26) and $w_\mathfrak{p}(y - y^q) \geqq 1$, shows

$$w_\mathfrak{p}(x_0) \geqq 0 .$$

Let $2L$ denote the number and $\mathscr{D} \subset \mathscr{S}$ the set of decomposition prime divisors $\mathfrak{p}$ of degree 1 of K/k. In other words, $2L$ is the number of solutions in k of the equation (1) corresponding to those $\mathfrak{p} \in \mathscr{D}$.

(iii) $e_\mathfrak{p} = 1, f_\mathfrak{p} = 2$. Then the restriction $\wp$ of $\mathfrak{p}$ on $k(x)$ is *inert* in K.

In this case we have $w_\mathfrak{p}(y) = 0$ and $w_\mathfrak{p}(1 - y^{q-1}) = 0$ since the residue class field of $\mathfrak{p}$ is a proper extension of k of degree $f_\mathfrak{p} = 2$ such that $\xi = .x \bmod \mathfrak{p}$ lies in k while $\eta = .y \bmod \mathfrak{p}$ does not lie in k, which means $\eta \neq 0$, $\eta^{q-1} \neq 1$.

Thus the inertia prime divisors $\mathfrak{p}$ of degree 2 of K/k do not yield any solutions in k of (1), but they contribute, on account of (26), the values

$$w_\mathfrak{p}(x_0) = 2w_\mathfrak{p}(y - y^q) - 2w_\mathfrak{p}(x - x^q) = 0 - 2 = -2$$

to $d(P_0 + Q_0)$.

The number of inertia prime divisors $\mathfrak{p}$ of degree 2 of K/k is $q - L - M$. Denote the set of those $\mathfrak{p}$'s by $\mathscr{J} \subset \mathscr{S}$.

According to what we have said under (i), (ii), (iii), the number N of solutions in k of the elliptic equation (1) is

$$N = 2L + M .$$

In order to compute $d(P_0 + Q_0)$ it remains to consider the infinite prime divisor $\mathfrak{p}_\infty$ of K/k. We have $w_{\mathfrak{p}_\infty}(x) = -2 < 0$, $w_{\mathfrak{p}_\infty}(x^q) = -2q < -2 < 0$. Lemma 1, (a) therefore yields

$$w_{\mathfrak{p}_\infty}(x_0) = w_{\mathfrak{p}_\infty}(x) = -2 .$$

Hence, by (9), we obtain altogether

$$\begin{aligned} d(P_0 + Q_0) &= -\tfrac{1}{2}\Big\{f_{\mathfrak{p}_\infty} w_{\mathfrak{p}_\infty}(x_0) + \sum_{\mathfrak{p} \in \mathscr{S}} f_\mathfrak{p} w_\mathfrak{p}(x_0) + \sum_{\mathfrak{p} \in \mathscr{J}} f_\mathfrak{p} w_\mathfrak{p}(x_0)\Big\} \\ &= -\tfrac{1}{2}\{-2 - 2M - 4(q - L - M)\} , \end{aligned}$$

whence

$$d(P_0 + Q_0) = 1 + 2(q - L) - M$$

such that

$$d(P_0 + Q_0) - d(P_0) - d(Q_0) = q - N .$$

This proves Lemma 4.

REFERENCES

1. M. Eichler, *Introduction to the Theory of Algebraic Numbers and Functions*, Academic Press, New York and London, 1966.
2. I. V. Elistratov, *Elementary proof of Hasse's theorem*, Izdat. Saratov Univ., Saratov 1966, 21-26; MR **34** #4253 (1967).
3. H. Hasse, *Beweis des Analogons der Riemannschen Vermutung für die Artinschen und F. K. Schmidtschen Kongruenzzetafunktionen in gewissen elliptischen Fällen*, Nachr. Ges. Wiss. Göttingen, Math.-Phys. Kl. I, **42** (1933), 253-262.
4. H. Hasse, *Zur Theorie der abstrakten elliptischen Funktionenkörper*, J. Reine Angew. Math. **175** (1936), 55-62, 69-88, 193-208.
5. E. Lutz, *Sur l'equation* $y^2 = x^3 - Ax - B$ *dans les corps p-adiques*, J. Reine Angew. Math. **177** (1937), 238-244.
6. Ju. I. Manin, *On cubic congruences to a prime modulus*, Izv. Akad. Nauk SSSR, Mat. Ser. **20** (1956), 673-678; or Amer. Math. Soc. Trans. (2) **13** (1960), 1-7; MR **18** 380 (1957).

Received January 19, 1970. This research was sponsored in part by NSF Grant GP-9661. Currently at Universität Katlsruhe-Germany.

UNIVERSITY OF CALIFORNIA, LOS ANGELES
OHIO STATE UNIVERSITY, COLUMBUS