PACIFIC JOURNAL OF MATHEMATICS
Vol. 135, No. 1, 1988

ON A THEOREM DUE TO CASSELS

JOSÉ M. SOUTO MENÉNDEZ

Using properties of one-dimensional formal groups, a proof is given of a theorem on the valuations of the torsion points of elliptic curves defined over p-adic fields.

1. Introduction. The aim of the present note is to give a proof of Theorem 5, due to Cassels, on the valuations of the torsion points of an elliptic curve defined over a local field K of characteristic zero. Cassels's proof relies on the addition formulas for the Weierstrass $\wp$ and $\wp'$ functions. The one given here follows from the properties of the torsion points of one-dimensional formal groups defined over the ring of integers of K.

The reader could also look at Oort [5] for another approach to Cassels' theorem.

2. Torsion points of formal groups. In the following we denote by K a local field, finite extension of the field Q_p of p-adic numbers, with ring of integers A; we assume that the normalized valuation v of K is extended to the algebraic closure $\overline{K}$ of K. We denote by $\mathfrak{p}_K$ (resp. $\mathfrak{p}_{\overline{K}}$) the maximal ideal of A (resp. of the valuation ring of $\overline{K}$), and by $e = v(p)$ the ramification index of K/Q_P.

Let F be a one-dimensional formal group of finite height $h \geq 1$, defined over A; as usual (see [3]), for each $a \in Z_p$ we denote by $[a](X) \in A[[X]]$ the unique endomorphism of F such that $[a](X) = aX + \cdots$. The group of points $F(\mathfrak{p}_{\overline{K}})$ of F with values in $\overline{K}$ has a structure of a module over Z_p, by means of the operation $a \cdot x = [a](x)$, $a \in Z_p$, $x \in F(\mathfrak{p}_{\overline{K}})$; $F(\mathfrak{p}_K)$ is a sub-Z_p-module of $F(\mathfrak{p}_{\overline{K}})$.

Let $[p](X) = \sum_{i=1}^{\infty} a_i X^i$ $(a_1 = p)$ be the "multiplication by p" in the formal group F; setting $q = p^h$, one has $a_i \in \mathfrak{p}_K$ if $i = 1, \ldots, q-1$ and $v(a_q) = 0$. We shall be interested in the valuations of the torsion points $x \in F(\mathfrak{p}_{\overline{K}})$. The most convenient thing is to consider the Newton polygon of the series $[p](X)$, that is the lower convex envelope of the points $(i, v(a_i)) \in R^2$ $(i \geq 1)$.

If $P_0 = (1, e)$, $P_1 = (q_1, e_1), \ldots, P_m = (q, 0)$ are the vertices of such a polygon (where $e_i = v(a_{q_i})$), the slopes are the negative of the numbers

$\alpha_1 = (e - e_1)/(q_1 - 1), \ldots, \alpha_m = e_{m-1}/(q_m - q_{m-1})$ $(\alpha_1 > \alpha_2 > \cdots > \alpha_m)$. If $q_i \leq r \leq q_{i+1}$ (for $i = 0, \ldots, m-1$), for any $x \in \mathfrak{p}_{\overline{K}}$ one has $v(a_r x^r) \geq \inf(v(a_{q_i} x^{q_i}), v(a_{q_{i+1}} x^{q_{i+1}}))$; moreover, if $r > q_m = q$, for $x \in \mathfrak{p}_{\overline{K}}$, $v(a_r x^r) > v(a_q x^q)$. Therefore, for any $x \in \mathfrak{p}_{\overline{K}}$ with $[p](x) = 0$, there exists $i = 0, \ldots, m-1$ such that $v(a_{q_i} x^{q_i}) = v(a_{q_{i+1}} x^{q_{i+1}})$, so that $v(x) = \alpha_{i+1}$. Moreover (see Koblitz [4]) the number of roots $x \in \mathfrak{p}_{\overline{K}}$ of the series $[p](X)$, of valuation α_{i+1}, is $q_{i+1} - q_i$.

LEMMA 1. *With the above notations, the q_i are powers of p.*

Proof. For each $i = 1, \ldots, m$, the set

$$\{x \in F(\mathfrak{p}_{\overline{K}}) \mid [p](x) = 0,\ v(x) \geq \alpha_i\}$$

is an elementary abelian p-group (with the operation given by the formal group law F); as its order is $(q_i - q_{i-1}) + \cdots + (q_1 - 1) + 1$, the lemma is obvious.

PROPOSITION 2. *For any $x \in \mathfrak{p}_{\overline{K}}$, one has*
–if $v(x) < \alpha_m$, then $v([p](x)) = qv(x)$,
–if $\alpha_{i+1} < v(x) < \alpha_i$, then $v([p](x)) = e_i + q_i v(x)$,
–if $\alpha_1 < v(x)$, then $v([p](x)) = e + v(x)$.

Proof. For $x \in F(\mathfrak{p}_{\overline{K}})$ such that $v(x) < \alpha_m$, then for any $r \neq q_m = q$, $v(a_q x^q) < v(a_r x^r)$. In fact, when $r > q$ such a relation is obvious (since $v(a_q) = 0$); when $r < q$, one may write

$$\begin{aligned} v(x) < \alpha_m &= (v(a_{q_{m-1}}) - v(a_q))/(q - q_{m-1}) \\ &< (v(a_r) - v(a_q))/(q - r), \end{aligned}$$

hence $v(a_q x^q) < v(a_r x^r)$.

If $\alpha_{i+1} < v(x) < \alpha_i$ (with $i = 1, \ldots, m-1$), then for any $r \neq q_i$, one has $v(a_{q_i} x^{q_i}) < v(a_r x^r)$. In fact, for $r > q_i$ this relation comes from

$$v(x) > (v(a_{q_i}) - v(a_{q_{i+1}}))/(q_{i+1} - q_i) \geq (v(a_{q_i}) - v(a_r))/(r - q_i);$$

for $r < q_i$, it comes from

$$v(x) < (v(a_{q_{i-1}}) - v(a_{q_i}))/(q_i - q_{i-1}) \leq (v(a_r) - v(a_{q_i}))/(q_i - r).$$

The case $v(x) > \alpha_1$ is discussed similarly.

REMARKS. (1) If $v(x) = \alpha_i$, $[p](x) \neq 0$ (for $i = 1, \ldots, m$), arguing as above, one gets $v([p](x)) \geq e_i + q_i \alpha_i$.

(2) For $i > \alpha_1$, $x \to [p](x)$ induces an isomorphism $F(\mathfrak{p}_K^i) \to F(\mathfrak{p}_K^{i+e})$ (of course, we denote by $F(\mathfrak{p}_K^r)$ the set $\mathfrak{p}_K^r$ with the group

structure given by the group law F). The injectivity comes from the fact that in $\mathfrak{p}_{\overline{K}}$ the zeros of $[p](X)$ have valuation $\leq \alpha_1$. To show the surjectivity, let Π be a uniformizing parameter of K; we have to see that if $y \in \mathfrak{p}_K$ is such that $v(y) = i + e > \alpha_1 + e$, there is $x = \Pi^{\alpha_1} t$ $(t \in A)$ such that $[p](\Pi^{\alpha_1} t) = y$; now, the series

$$\frac{1}{\Pi^{\alpha_1+e}}(-y + p\Pi^{\alpha_1} T + a_2 \Pi^{2\alpha_1} T^2 + \cdots)$$

has coefficients in A and Weierstrass degree one, so the result follows from the Preparation Theorem for power series.

(3) If F is the multiplicative group, the Newton polygon of $[p](X)$ only has one slope. Proposition 2 gives then the well-known effect of "raising to the pth power" in the group of principal units of the local field K (or of any of its finite extensions).

PROPOSITION 3. *$F(\mathfrak{p}_K)$ is a Z_p-module of finite type, whose rank modulo torsion is $[K\colon Q_p]$. The torsion subgroup is a finite p-group.*

Proof. For each $i \geq 1$, let us denote, as above, by $F(\mathfrak{p}_K^i)$ the abelian group on the set $\mathfrak{p}_K^i$ with the operation given by $(x, y) \to F(x, y)$; of course, $F(\mathfrak{p}_K^i)$ is a Z_p-submodule of $F(\mathfrak{p}_K)$. It is trivial that $\mathfrak{p}_K^i/\mathfrak{p}_K^{i+1} \simeq F(\mathfrak{p}_K^i)/F(\mathfrak{p}_K^{i+1})$.

The filtration $F(\mathfrak{p}_K) \supset F(\mathfrak{p}_K^2) \supset \cdots$ is separated and produces in $F(\mathfrak{p}_K)$ the p-adic topology (if i is large enough, one of the remarks shows that $pF(\mathfrak{p}_K^i) = F(\mathfrak{p}_K^{i+e})$). According to a well-known lemma in commutative algebra, the finiteness of $F(\mathfrak{p}_K)$ as a module over Z_p, follows from the finiteness of $F(\mathfrak{p}_K)/pF(\mathfrak{p}_K)$, a quotient of $F(\mathfrak{p}_K)/F(\mathfrak{p}_K^{i+e}) = F(\mathfrak{p}_K)/pF(\mathfrak{p}_K^i)$ for i large enough.

Taking again i large enough so that $F(\mathfrak{p}_K^i)$ is torsion free, hence free, its rank is the same as $\dim_{F_p}(F(\mathfrak{p}_K^i)/pF(\mathfrak{p}_K^i)) = \dim_{F_p}(F(\mathfrak{p}_K^i)/F(\mathfrak{p}_K^{i+e}))$; since $(F(\mathfrak{p}_K^i)\colon F(\mathfrak{p}_K^{i+e})) = p^{[K:Q_p]}$, the proposition is clear.

PROPOSITION 4. *Let $x \in F(\mathfrak{p}_{\overline{K}})$ be a torsion point of order p^r. Then $v(x) \leq e/\varphi(p^r) = e/p^{r-1}(p-1)$.*

Proof. From Proposition 2, it is obvious that for any $x \in F(\mathfrak{p}_{\overline{K}})$, $v([p](x)) \geq v(x)$; therefore, if $v(x) > \alpha_1$, x is not a torsion point.

One proves the proposition by induction on r. If x is of order p, $v(x) \leq \alpha_1 = (e - e_1)/(q_1 - 1) \leq e/(p-1)$ (by Lemma 1). If x is of order p^r $(r > 1)$, then $v(x) \leq \alpha_1$ and $v([p](x)) \leq e/p^{r-2}(p-1)$, by the induction hypothesis; again by Proposition 2,

$$v(x) \leq \alpha_1 \Rightarrow v([p](x)) \geq pv(x),$$

so $v(x) \leq v([p](x))/p \leq e/p^{r-1}(p-1)$.

REMARK. Sometimes, one can be more precise about $v(x)$. If the height of F is $h = 1$, the Newton polygon has only one slope and all the points of order p have valuation $e/(p-1)$; in this case, if $x \in F(\mathfrak{p}_{\overline{K}})$ is of order p^r $(r \geq 1)$, $v(x) = e/p^{r-1}(p-1)$.

If the height of F is $h = 2$, there are two possibilities for the Newton polygon. If there is only one slope, the points $x \in F(\mathfrak{p}_{\overline{K}})$ of order p^r have exact valuation $v(x) = e/p^{2(r-1)}(p^2-1)$. If there are two slopes, one cannot say more than in Proposition 4.

3. Cassels's theorem. In the following theorem, E denotes an elliptic curve defined over the local field K, given by a minimal Weierstrass equation

$$\text{(X)} \qquad y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$$

$(a_i \in A)$. We write $E(K)$ for the group of points of E with values in K; $E(K)$ is an abelian group in the usual way, taking the point at infinity (0,1,0) of E as zero element. Notations are the same as in Tate **[6]**.

THEOREM 5. *Let $(x, y) \in E(K)$ a torsion point of E. If the order of (x, y) is not a power of p, then $x, y \in A$. If the order of (x, y) is p^r $(r \geq 1)$, then*

$$v(x) \geq -2e/p^{r-1}(p-1), \quad v(y) \geq -3e/p^{r-1}(p-1).$$

Proof. By reducing the equation (X) of E modulo the maximal ideal of A, we get the equation of a cubic $\tilde{E}$ defined over the residue field k of K. The set $\tilde{E}_{ns}(k)$ of nonsingular points of $\tilde{E}$ with values in k is a group, and one has the exact sequence

$$0 \to E_1(K) \to E_0(K) \to \tilde{E}_{ns}(k) \to 0;$$

here $E_0(K)$ denotes the subgroup of the elements of $E(K)$ that reduce to the nonsingular points of E, and

$$E_1(K) = 0 \cup \{(x, y) \in E(K) \mid v(x) \leq -2, \ v(y) \leq -3\}$$

is the kernel of the reduction map.

One knows that there is a formal group law F defined over A, and an isomorphism

$$E_1(K) \xrightarrow{\sim} F(\mathfrak{p}_K),$$

$$(x, y) \to -x/y.$$

Such a formal group F is isomorphic to the additive one if E has bad reduction and the singularity of $\tilde{E}$ is a cusp, and of height one or two in the other cases; in the first case, $F(\mathfrak{p}_K)$ is of course torsion free, and in the other ones, the only possible torsion is p-torsion. In these cases, if $z \in F(\mathfrak{p}_K)$ has order p^r, $v(z) \leq e/p^{r-1}(p-1)$ (Proposition 4); since we have for the corresponding point $(x, y) \in E_1(K)$

$$x = z^{-2} - a_1 z^{-1} - a_2 - \cdots,$$
$$y = -z^{-1}x,$$

we get $v(x) = -2v(z) \geq -2e/p^{r-1}(p-1)$, $v(y) \geq -3e/p^{r-1}(p-1)$.

The theorem is proved taking account of the fact that

$$E(K) - E_1(K) = \{(x, y) \in E(K) \mid x, y \in A\}.$$

COROLLARY 6 (*Nagell-Lutz*). *Let E be an elliptic curve defined over Q, given by a minimal global Weierstrass equation of the form* (X) *with the a_i rational integers. Then the torsion points of $E(Q)$ have integer coordinates, with one possible exception: there could be a unique point of order two of the form $(a/4, b/8)$, with $a, b, \in Z$.*

Proof. For each prime number p, we denote by v_p the p-adic valuation of Q (extended to Q_p). Since we have $E(Q) \subset E(Q_p)$, we can apply the last theorem.

If $(x, y) \in E(Q)$ is a torsion point whose order is not a power of any prime number p, then $x, y \in Z_p$ for each p, so $x, y \in Z$.

If the order of $(x, y) \in E(Q)$ is p^r (p prime, $r \geq 1$), then for each prime $l \neq p$, $x, y \in Z_l$; moreover

$$v_p(x) \geq -2/p^{r-1}(p-1), \quad v_p(y) \geq -3/p^{r-1}(p-1),$$

so $x, y \in Z_p$ unless, perhaps, $p^r = 2, 3, 4$. If $p^r = 3$ or 4, again $x, y \in Z_p$, since $x, y \notin Z_p$ implies $v_p(x) \leq -2$, $v_p(y) \leq -3$.

So we are only left with the possibility of points of order $p^r = 2$; if $(x, y) \in E(Q)$ is one of those points, $x, y \in Z_l$ for each $l \neq 2$ and $\nu_2(x) = -2$, $v_2(y) = -3$; then (x, y) should belong to the kernel $E_l(Q_2)$ of the reduction of E modulo 2. Looking at the power series $[2](X) = 2X - a_1X^2 - 2a_2X^3 + \cdots$, we find that, in fact, if the formal group F associated to the model (X) of the curve E is of height one in Z_2 ($\Longleftrightarrow a_1 \notin 2Z$), there exists in $E(Q_2)$ a unique point of order two whose coordinates are $(a/4, b/8)$, $a, b \in Z_2$; such a point may or may not be in $E(Q)$.

REMARK. As shown in the proof, one has to study the possibility of a torsion point of order two in $E(Q)$ only when E has ordinary good reduction or split multiplicative reduction at 2.

4. Appendix. If $P = (x(P), y(P)) \in E(Q)$ is a torsion point of order different from two of the curve E given by the equation (X)—where the $a_i \in Z$—we know that $x(P),\ y(P) \in Z$, and so P verifies the hypothesis of the following proposition.

PROPOSITION 7. *Let Δ be the discriminant of the curve E. If $P = (x(P), y(P)) \in E(Q)$ is a point with integer coordinates such that $2P = (x(2P), y(2P))$ has also integer coordinates, then* $(2y(P) + a_1x(P) + a_3)^2 | \Delta$.

Proof. We only sketch it. We write, as in **[6]**,

$$\begin{aligned}
&b_2 \doteq a_1^2 + 4a_2, \quad b_4 = a_1a_3 + 2a_4, \quad b_6 = a_3^2 + 4a_6,\\
&b_8 = a_1^2a_6 - a_1a_3a_4 + 4a_2a_6 + a_2a_3^2 - a_4^2,\\
&\Delta = -b_2^2b_8 - 8b_4^3 - 27b_6^2 + 9b_2b_4b_6.
\end{aligned}$$

Multiplication by two, $E \xrightarrow{2} E$, is given by a formula

$$2(x, y) = (x_2, y_2)$$

where $x_2 = u(x)/f(x)$, with $u(T) = T^4 - b_4T^2 - 2b_6T - b_8$ and $f(T) = 4T^3 + b_2T^2 + 2b_4T + b_6$; one has $\operatorname{disc}_3(f(T)) = 16\Delta$, and the relation

$$16u(T) - f'(T)^2 + 4(8T + b_2)f(T) = 0.$$

One verifies, with the notations of Bourbaki **[1]** (Ch. IV, §6),

$$\begin{aligned}
\operatorname{Res}_{4,3}(16u(T), f(T)) &= \operatorname{Res}_{4,3}(f'(T)^2 - 4(8T + b_2)f(T), f(T))\\
&= \operatorname{Res}_{4,3}(f'(T)^2, f(T)) = [\operatorname{Res}_{2,3}(f'(T), f(T))]^2\\
&= 16(\operatorname{disc}_3 f(T))^2 = 2^{12}\Delta^2;
\end{aligned}$$

therefore, $\operatorname{Res}_{4,3}(u(T), f(T)) = \Delta^2$.

On the other hand

$$\mathrm{Res}_{4,3}(u(T), f(T))$$

$$= \begin{vmatrix} 1 & 0 & 0 & 4 & 0 & 0 & 0 \\ 0 & 1 & 0 & b_2 & 4 & 0 & 0 \\ -b_4 & 0 & 1 & 2b_4 & b_2 & 4 & 0 \\ -2b_6 & -b_4 & 0 & b_6 & 2b_4 & b_2 & 4 \\ -b_8 & -2b_6 & -b_4 & 0 & b_6 & 2b_4 & b_2 \\ 0 & -b_8 & -2b_6 & 0 & 0 & b_6 & 2b_4 \\ 0 & 0 & -b_8 & 0 & 0 & 0 & b_6 \end{vmatrix}$$

$$= \begin{vmatrix} 1 & 0 & 0 & 4 & 0 & 0 & 0 \\ 0 & 1 & 0 & b_2 & 4 & 0 & 0 \\ -b_4 & 0 & 1 & 2b_4 & b_2 & 4 & 0 \\ -2b_6 & -b_4 & 0 & b_6 & 2b_4 & b_2 & 4 \\ -b_8 & -2b_6 & -b_4 & 0 & b_6 & 2b_4 & b_2 \\ 0 & -b_8 & -2b_6 & 0 & 0 & b_6 & 2b_4 \\ T^2u(T) & Tu(T) & u(T) & T^3f(T) & T^2f(T) & Tf(T) & f(T) \end{vmatrix}$$

$$\begin{aligned} &= -48\Delta T^2u(T) - 8b_2\Delta Tu(T) + (b_2^2 - 32b_4)\Delta u(T) + 12\Delta T^3 f(T) \\ &\quad - b_2\Delta T^2 f(T) - 10b_4\Delta Tf(T) + (b_2b_4 - 27b_6)\Delta f(T); \end{aligned}$$

here we have developed the last determinant by the last row, and made systematic use of the relation $4b_8 = b_2b_6 - b_4^2$.

Therefore,

$$\begin{aligned} \Delta = &(-48T^2 - 8b_2T + (b_2^2 - 32b_4))u(T) \\ &+ (12T^3 - b_2T^2 - 10b_4T + (b_2b_4 - 27b_6))f(T). \end{aligned}$$

Now, if $P = (x(P), y(P)) \in E(Q)$ and $2P = (x(2P), y(2P))$ have integer coordinates, as $u(x(P)) = x(2P)f(x(P))$, we get

$$f(x(P)) \mid \Delta.$$

Since $y^2 + a_1xy + a_3y = x^3 + a_2x^2 + a_4x + a_6$ implies

$$(2y + a_1x + a_3)^2 = 4x^3 + b_2x^2 + 2b_4x + b_6 = f(x),$$

the proposition is proved.

References

[1] N. Bourbaki, *Algèbre* , Chapitre 4 á 7. Ed. Masson, 1981.

[2] J. W. S. Cassels, *Diophantine equations with special reference to elliptic curves*, J. London Math. Soc., **41** (1966), 193–291.

[3] A. Frohlich, *Formal Groups*, Lecture Notes in Math. 74, Springer 1968.

[4] N. Koblitz, *p-adic Numbers, p-adic Analysis and Zeta Functions*, Graduate Texts in Mathematics 58, Springer 1977.

[5] F. Oort, *Elliptic curves: Diophantine torsion solutions and singular j-invariants*, Math. Ann., **207**, 139–162.

[6] J. Tate, *The arithmetic of elliptic curves*, Inv. Math., **23** (1974), 179–206.

Received May 16, 1985.

Perlora
Carreño
Asturias, Spain