8

DOES THE SET OF POINTS OF AN ELLIPTIC CURVE DETERMINE THE GROUP?

H. W. Lenstra, Jr.* and J. Pila**

* *Department of Mathematics # 3840*
University of California at Berkeley,
Berkeley, CA 94720–3840,
U. S. A.
hwl@math.berkeley.edu

** *Department of Mathematics*
The University of Melbourne
Parkville, Victoria 3052
Australia
pila@mundoe.maths.mu.oz.au

ABSTRACT

Let E be an elliptic curve over a field k, given in Weierstrass form. As is well known, the set $E(k)$ of points of E over k forms an abelian group in a natural way, the point at infinity being the zero element. It is often heard that the group structure on $E(k)$ is "determined" by the fact that three points of $E(k)$ add up to zero if they lie on a line. In this paper we investigate whether this statement is correct if taken literally. We find that it is not. In fact, we exhibit a field k and two elliptic curves E, E' over k for which $E(k)$ and $E'(k)$ are equal as subsets of the set of points of the projective plane over k, but have different group structures. Our main result states that this is a rare phenomenon: it occurs if and only if k has characteristic 2 and $E(k)$ has order 5. We also encounter an elliptic curve E for which $E(k)$ has a $\mathbf{Z}[i]$-module structure for many fields k, even though E does not have complex multiplication by $\mathbf{Z}[i]$.

1991 Mathematics Subject Classification: 14H52.

Acknowledgements

The authors are grateful to Ken Ribet and Ron Livné for clarifying discussions. They thank the NSF for support under Grants DMS 90–02939 and DMS 91–04316.

W. Bosma and A. van der Poorten (eds.), Computational Algebra and Number Theory, 111–118.

1 INTRODUCTION

Let k be a field, and let E be an elliptic curve over k, given by a Weierstrass equation

$$y^2 + a_1 xy + a_3 y = x^3 + a_2 x^2 + a_4 x + a_6$$

with coefficients a_1, a_2, a_3, a_4, a_6 in k and with a non-zero discriminant (see [6, Chapter III, Section 1]). The set $E(k)$ of points of E over k is the set of points (x, y) in $k \times k$ satisfying the equation, together with a "point at infinity", which is denoted by O.

It is well known that $E(k)$ forms, in a natural way, an abelian group which is written additively, and which has the following properties: O is the zero element; three points on $E(k)$ that lie on a straight line add up to O. Here it is understood that the lines passing through O are the vertical lines where x is constant.

In informal accounts of elliptic curves it is often implied that the properties just mentioned characterize the group addition on $E(k)$. In the present paper we investigate whether this statement is correct, if taken literally. It would imply that the group structure on $E(k)$ is determined by the set $E(k)$, as a subset of $(k \times k) \cup \{O\}$. The difficulty is that over general fields the intersection multiplicity of a line with the curve at a point cannot be computed from the set $E(k)$ alone, without reference to the equation defining the curve.

The following example shows that the group structure in general is not determined by the set of points. Let $k = \mathbf{F}_2$ be the field of two elements and let the elliptic curves E and E' over k be defined by

$$\begin{aligned} E: \quad & y^2 + y = x^3 + x^2 \\ E': \quad & y^2 + y = x^3 + x\,. \end{aligned}$$

In this case we have $E(k) = E'(k) = (k \times k) \cup \{O\}$. Each of $E(k)$ and $E'(k)$ is a group of order 5 but the group structures are different: in $E(k)$ one has $(0,0) + (0,0) = (1,1)$ whereas $(0,0) + (0,0) = (1,0)$ in $E'(k)$.

Our first result expresses that this example is typical.

1.1 Theorem *Let k be a field and let E be an elliptic curve over k, given by a Weierstrass equation. Then there exists an elliptic curve E' over k, given by a Weierstrass equation, such that $E(k)$ and $E'(k)$ are equal as sets but not as groups, if and only if $E(k)$ has order 5 and k has characteristic 2. If E' exists, then its Weierstrass equation is uniquely determined and E' is isomorphic to E.*

The proof of Theorem 1.1 is given in Section 3.

Theorem 1.1 makes us wonder whether it often occurs, over fields of characteristic 2, that the number of points of an elliptic curve equals 5. To address this question,

we first give a precise description of elliptic curves with a subgroup of order 5. We define a *Weierstrass transformation* to be a change of coordinates of the form $(x,y) \to (r^2x+u, r^3y+vx+w)$ where $r, u, v, w \in k$ with $r \neq 0$. Such a change of coordinates preserves the Weierstrass form of equations defining elliptic curves.

1.2 Theorem *Let k be a field and let $a, b, b', c, d, d' \in k$ be such that $a \neq c$, $b \neq b'$ and $d \neq d'$. Then the following two assertions are equivalent:*

(a) *there exists a Weierstrass equation defining an elliptic curve E over k with the property that $\{O, (a,b), (a,b'), (c,d), (c,d')\}$ is a subgroup of order 5 of $E(k)$ with $2 \cdot (a,b) + (c,d) = O$;*

(b) $(c-a)^3 = (d'-d)(b'-b)$ *and* $t = (d'-d)/(b'-b)$ *satisfies* $t - t^{-1} \neq 11$.

If a Weierstrass equation as in (a) *exists, then it is unique and there is a Weierstrass transformation transforming $(a,b), (a,b'), (c,d), (c,d')$ into $(0,0), (0,t), (t,0), (t,t^2)$ respectively; letting $t = (d'-d)/(b'-b)$ the equation transforms into*

$$y^2 + (1-t)xy - ty = x^3 - tx^2 .$$

This is proved by means of a well-known argument (see [6, Exercise 8.13]; cf. [2, Table 3, entry 13]) which is given in Section 2.

Next we discuss the situation that k has characteristic 2 and $E(k)$ has order 5. The following theorem summarizes what we know about this: the parameter t in the previous theorem is either equal to 1 or it is transcendental; both cases actually occur.

1.3 Theorem

(a) *Let k be a field of characteristic 2, and let E be an elliptic curve over k, given by a Weierstrass equation, such that $E(k)$ has order 5. Then there is a Weierstrass transformation that brings the equation for E in the form*

$$y^2 + (1+t)xy + ty = x^3 + tx^2$$

where $t \in k$ is equal to 1 or is transcendental over $\mathbf{F}_2$.

(b) *The Weierstrass equation $y^2 + y = x^3 + x^2$ defines an elliptic curve E over $\mathbf{F}_2$ for which $E(k)$ has order 5 if $k = \mathbf{F}_2, \mathbf{F}_4$ or $\mathbf{F}_8$ but not for any other algebraic extension of $\mathbf{F}_2$. Also $E(k)$ is of order 5 for any purely transcendental extension k of $\mathbf{F}_2, \mathbf{F}_4$ or $\mathbf{F}_8$.*

(c) *Let k be any field and let t be transcendental over k. Then the Weierstrass equation $y^2 + (1-t)xy - ty = x^3 - tx^2$ defines an elliptic curve E over $k(t)$ for which $E(k(t))$ has order 5.*

The proof of this theorem is given in Section 4.

The curves occurring in Theorem 1.2 have a curious property which is formulated in our final result. We denote by $\mathbf{Z}[i]$ the ring of Gaussian integers, with $i^2 = -1$.

1.4 Theorem *Let k be a field, σ an automorphism of k with $\sigma^2 = \mathrm{id}_k$, and t an element of k satisfying $t \cdot \sigma t = -1$ with $t + \sigma t \neq 11$. Then the equation*

$$y^2 + (1-t)\,xy - t\,y \;=\; x^3 - t\,x^2$$

defines an elliptic curve E over k for which the group $E(k)$ has a module structure over $\mathbf{Z}[i]$, but that does not have complex multiplication by $\mathbf{Z}[i]$ unless $t = \sigma t$ or $(t + \sigma t - 9)^2 = 5$.

This is proved in Section 2.

2 POINTS OF ORDER FIVE

In this section we prove Theorem 1.2 and Theorem 1.4. We begin with the implication (a)⇒(b) of Theorem 1.2.

Proof Assume that (a) of Theorem 1.2 holds. Applying a Weierstrass transformation we may replace the points $(a,b), (a,b'), (c,d), (c,d')$ by $(0,0), (0,t), (t,0), (t,rt)$ respectively, where $t = (c-a)^3/(b'-b)^2$ and $r = (d'-d)/(b'-b)$. Let the Weierstrass equation be

$$y^2 + a_1\,xy + a_3\,y \;=\; x^3 + a_2\,x^2 + a_4\,x + a_6\,.$$

The line passing through $(0,0)$ and $(t,0)$ is given by $y = 0$. Now the condition $2 \cdot (0,0) + (t,0) = O$ expresses that this line is tangent to the curve in $(0,0)$ and that it also intersects the curve in $(t,0)$. This means that upon substitution of $y = 0$ the Weierstrass equation reduces to the equation $0 = x^2\,(x-t)$. Therefore we have $a_2 = -t$ and $a_4 = a_6 = 0$. Since $(0,t)$ lies on the curve we have $a_3 = -t$.

Because $(0,0)$ and $(0,t)$ have the same x-coordinate, we have $(0,t) = -(0,0)$ in the group $E(k)$. From (a) it thus follows that $2 \cdot (t,0) + (0,t) = -5 \cdot (0,0) = O$. This implies that the line $x + y = t$, which meets the curve in $(0,t)$ and $(t,0)$, is tangent to it in the latter point. Therefore the Weierstrass equation reduces to $x(x-t)^2 = 0$ upon substitution of $y = t - x$. This leads to $a_1 = 1 - t$, so that the Weierstrass equation is as in the statement of the theorem. Because (t,rt) lies on the curve, we have $t = r = (d'-d)/(b'-b)$, so $(c-a)^3 = (d'-d)(b'-b)$. The discriminant of the Weierstrass equation is found to be $t^6\,(t - t^{-1} - 11)$ (cf. [6, Chapter III, Section 1] and [2, Table 3, entry 13]). It is non-zero, so $t - t^{-1} \neq 11$. This proves (b) and the uniqueness statement of the theorem is also proved.

The proof that (b) implies (a) is now straightforward, since we know which equation we have to try. This proves Theorem 1.2. □

Proof Theorem 1.4 arises from the observation that multiplication by 2 is an automorphism of any additive group of order 5. In the situation of Theorem 1.2, this

automorphism maps $(a,b),(a,b'),(c,d),(c,d')$ to $(c,d'),(c,d),(a,b),(a,b')$ respectively, so (a) holds for a,b,b',c,d,d' if and only if it holds for c,d',d,a,b,b'. Then by the uniqueness statement of Theorem 1.2 the two elliptic curves are the same. Note that the permutation changes t into $-t^{-1}$. Therefore there is not only a Weierstrass transformation that transforms the equation for the curve into

$$y^2+(1-t)\,xy-t\,y \;=\; x^3-t\,x^2$$

and the points $(a,b),(a,b'),(c,d),(c,d')$ into $(0,0),(0,t),(t,0),(t,t^2)$ respectively, but there is also a Weierstrass transformation that transforms the equation for the curve into

$$y^2+(1+t^{-1})\,xy+t^{-1}y \;=\; x^3+t^{-1}\,x^2$$

and the points $(a,b),(a,b'),(c,d),(c,d')$ into $(-t^{-1},0),(-t^{-1},t^{-2}),(0,-t^{-1}),(0,0)$ respectively. Composing the inverse of the first Weierstrass transformation with the second we obtain a transformation of the first equation into the second. Let now σ be as in Theorem 1.4. Then σ interchanges t and $-t^{-1}$, so the solution sets of the two equations are mapped to each other by $\sigma(x,y)=(\sigma x,\sigma y)$. With $\sigma O=O$ this is actually a group isomorphism. Composing the Weierstrass transformation that links the two equations with σ one finds a group automorphism i of $E(k)$, where E is the curve given by the first equation. Explicitly, one has

$$i(x,y) \;=\; \left(\sigma(x)\,t^2+t\;,\;-\sigma(x)\,t^2-\sigma(y)\,t^3\right).$$

We claim that $i^2=-1$. This can either be verified by an explicit computation, or one can argue as follows. From $\sigma^2=\mathrm{id}_k$ it follows that i^2 belongs to the automorphism group of the curve. Suppose first that t is transcendental over the prime field of k. Then the j-invariant of E, which is given by

$$j(E) \;=\; \frac{(t^4-12\,t^3+14\,t^2+12\,t+1)^3}{t^5\,(t^2-11\,t-1)}$$

is transcendental as well. Therefore E has no complex multiplication, and its only automorphisms are 1 and -1, so that $i^2=\pm 1$. Checking the action of i on the given points of order 5 one finds that $i^2=-1$. This is, by specialization, then also true if t is not transcendental. Hence $E(k)$ has a module structure over the ring of Gaussian integers.

If E has complex multiplication by $\mathbf{Z}[i]$ then one has $j(E)=1728$. From

$$j(E)-1728 \;=\; \frac{(t^2+1)^2\,(t^4-18\,t^3+74\,t^2+18\,t+1)^2}{t^5\,(t^2-11\,t-1)}$$

one finds that this is equivalent to $t=\sigma t$ or $(t+\sigma t-9)^2=5$. This proves Theorem 1.4. □

We remark that $t^4-18\,t^3+74\,t^2+18\,t+1=0$ defines, in characteristic zero, a unit in the ring $\mathbf{Z}[\zeta_{20}+\zeta_{20}^{-1}]$ and that $t^4-12\,t^3+14\,t^2+12\,t+1=0$ likewise defines a unit of $\mathbf{Z}[\zeta_{15}+\zeta_{15}^{-1}]$; here ζ_m denotes a primitive m-th root of unity.

2.1 Remark Theorem 1.2 shows that pairs consisting of an elliptic curve and a point of order 5 on it can be parameterized with a single parameter t. This reflects the fact that the modular curve $X_1(5)$ is rational. The cusps of $X_1(5)$ correspond to $t = 0$, $t = \infty$ and $t - 1/t = 11$. The transformation $t \to -1/t$ is the "diamond operator" $\langle 2 \rangle$ on $X_1(5)$.

3 EQUAL SETS OF POINTS WITH DIFFERENT GROUP STRUCTURES

In this section we prove Theorem 1.1.

Proof We begin with the "if" part. Suppose that E is an elliptic curve over a field k of characteristic 2, and that $E(k)$ has order 5. Then $E(k) = \{O, P, -P, Q, -Q\}$ for certain P, Q. Applying a Weierstrass transformation we may assume that $P = (0,0)$. Then $-P$ also has x-coordinate equal to 0 and the x-coordinate of Q is different from 0. Therefore there is a unique Weierstrass transformation of the type $(x, y) \to (x, y + tx)$ (with $t \in k$) that maps Q to $-Q$ and fixes P and $-P$. Since the characteristic of k equals 2 this Weierstrass transformation has order 2, so it maps $-Q$ to Q. Therefore it transforms the equation for E into the equation for an isomorphic elliptic curve E' for which $E'(k) = \{O, P, -P, Q, -Q\}$ as well. The isomorphism $E(k) \to E'(k)$ fixes P and $-P$ and interchanges Q and $-Q$. This is not a group automorphism of $E(k)$, so the group structures on $E(k)$ and $E'(k)$ are different.

Next we prove the "only if" part. Let it be supposed that E and E' are elliptic curves over a field k, given by Weierstrass equations

$$\begin{aligned} E: \quad & y^2 + a_1 xy + a_3 y &&= x^3 + a_2 x^2 + a_4 x + a_6 \\ E': \quad & y^2 + a_1' xy + a_3' y &&= x^3 + a_2' x^2 + a_4' x + a_6' . \end{aligned}$$

We assume that $E(k) = E'(k)$ and that the group structures on $E(k)$ and $E'(k)$ are different.

The map $E(k) \to E(k)$ sending P to $-P$ can be described as follows. If $P = O$ then $-P = O$. Next let $P = (x, y) \neq O$. Then there exists at most one other $Q \in E(k)$ with the same x-coordinate as P; if it exists then $-P = Q$ otherwise $-P = P$. This description is entirely in terms of the set $E(k)$. Therefore the map is the same as the map $E'(k) \to E'(k)$ sending P to $-P$. In particular, the two groups $E(k)$ and $E'(k)$ do not only have the same zero element but also the same elements of order two.

It is easy to see that a given set of cardinality at most four has at most one group structure for which the zero element and the elements of order two are prescribed. Hence what we just proved implies that the set $E(k) = E'(k)$ has at least 5 elements.

There are at most two points different from O on $E(k)$ with a given x-coordinate, so at least two different x-coordinates occur. Let $P = (a, b)$, $Q = (c, d) \in E(k)$ be chosen such that $a \neq c$. Then the point $-P$, which is the same for $E(k)$ and $E'(k)$, is of the form (a, b') and we have both $b+b' = -a_1 a - a_3$ and $b+b' = -a_1' a - a_3'$. Likewise, we have $-Q = (c, d')$, where $d + d' = -a_1 c - a_3 = -a_1' c - a_3'$. From

$$a_1 a + a_3 = a_1' a + a_3' \quad \text{and} \quad a_1 c + a_3 = a_1' c + a_3' \quad \text{with} \quad a \neq c$$

it clearly follows that $a_1 = a_1'$ and $a_3 = a_3'$. Since P and Q satisfy both the equation for E and the equation for E' we have now

$$a_2 a^2 + a_4 a + a_6 = a_2' a^2 + a_4' a + a_6' \quad \text{and} \quad a_2 c^2 + a_4 c + a_6 = a_2' c^2 + a_4' c + a_6' .$$

Hence the quadratic polynomial $(a_2 - a_2') x^2 + (a_4 - a_4') x + (a_6 - a_6')$ vanishes in $x = a$ and in $x = c$. If a third x-coordinate e would occur among the x-coordinates of the points in $E(k)$, then the polynomial would vanish in e as well, so it would be identically zero. In that case E and E' would be given by the same equation, contradicting that $E(k)$ and $E'(k)$ have different group structures. We conclude that there is no third coordinate, so $E(k) = E'(k) = \{O, P, -P, Q, -Q\}$. It follows that the order is equal to 5, and that $P \neq -P$ and $Q \neq -Q$.

On $E(k)$, we have $2P = Q$ or $2P = -Q$. Interchanging d and d' if necessary, we may assume that the latter alternative holds. Then we are in the situation of Theorem 1.2, so $(c-a)^3 = (d'-d)(b'-b)$. On $E'(k)$ we must have $2P = Q$ or the group structure would be the same. Thus from Theorem 1.2, with d and d' interchanged, we find that $(c-a)^3 = (d-d')(b'-b)$. Therefore $(d'-d)(b'-b) = (d-d')(b'-b)$, which implies that the characteristic of k equals 2. This proves the "only if" part of Theorem 1.1.

The uniqueness of E', given E, follows from Theorem 1.2. As we saw in the proof of the "if" part, the curve E' is isomorphic to E. This completes the proof of Theorem 1.1. □

4 FIVE POINTS IN CHARACTERISTIC TWO

In this section we prove Theorem 1.3.

Proof (a) Let k be a field of characteristic 2, and let E be an elliptic curve over k such that $E(k)$ has order 5. By Theorem 1.2, we can bring the equation for E in the form

$$y^2 + (1+t)\, xy + ty = x^3 + t x^2$$

for some $t \in k$. Suppose that $t \neq 1$ and that t is algebraic over $\mathbf{F}_2$. Then the field $\mathbf{F}_2(t)$, being finite, is perfect so t has a squareroot in $\mathbf{F}_2(t)$. This gives rise to a point $\left(t/(1+t), t^2/(1+\sqrt{t})^3\right)$ of order two in the group $E(k)$, which is impossible since $E(k)$ has order 5. This contradiction proves (a).

(b) By Theorem 1.2, the Weierstrass equation $y^2 + y = x^3 + x^2$ defines an elliptic curve E over $\mathbf{F}_2$. Since all $(x, y) \in \mathbf{F}_2 \times \mathbf{F}_2$ satisfy the equation, we have $E(\mathbf{F}_2) = (\mathbf{F}_2 \times \mathbf{F}_2) \cup \{O\}$, which is of order 5. From the fact that the coefficient a_1 in the Weierstrass equation vanishes one deduces that $E(L)$ does not have a point of order 2 for any extension field L of $\mathbf{F}_2$. Hence if $E(\mathbf{F}_4)$ or $E(\mathbf{F}_8)$ would have order greater than 5, it would have order at least 15, which contradicts Hasse's estimate

$$(\sqrt{q} - 1)^2 \leq \#E(\mathbf{F}_q) \leq (\sqrt{q} + 1)^2$$

(see [6, Chapter V, Theorem 1.1]) since $15 > (\sqrt{8} + 1)^2$. This proves that $E(\mathbf{F}_4)$ and $E(\mathbf{F}_8)$ have order 5. For $q \geq 16$ one has $(\sqrt{q} - 1)^2 > 5$, so Hasse's estimate implies that $E(k)$ has order greater than 5 for any algebraic extension k of $\mathbf{F}_2$ of degree at least 4.

To prove the last statement of (b), it suffices to remark that for any elliptic curve E over any field k one has $E(k(u)) = E(k)$ when u is transcendental over k. In fact, any point in $E(k(u))$ that is not in $E(k)$ would give an embedding of the function field of E over k into $k(u)$, contradicting Lüroth's theorem. This proves (b).

(c) Let t be transcendental over a field k. Then by Theorem 1.2 the equation $y^2 + (1 - t)xy - ty = x^3 - tx^2$ defines an elliptic curve over $k(t)$ for which $E(k(t))$ has a subgroup of order 5. To prove that there are no other points in $E(k(t))$ we may clearly assume that k is algebraically closed. We can form the elliptic surface S over k associated with E, as described in [5, Section 1]. The function field of S is the field of fractions of the ring $k[x, y, t]/(y^2 + (1 - t)xy - ty - x^3 + tx^2)$, which is just the rational function field $k(x, y)$. Therefore S is a rational elliptic surface, as defined in [5, Section 10]. The Mordell–Weil groups of rational elliptic surfaces over algebraically closed fields have been completely determined by Oguiso and Shioda [4]. Inspecting their list of the possible groups ([4, Corollary 2.1]) we see that if there is a subgroup of order 5, then that is the whole group.

This completes the proof of Theorem 1.3. □

REFERENCES

[1] R. Hartshorne, *Algebraic geometry*, New York: Springer–Verlag, 1977.

[2] D. S. Kubert, *Universal bounds on the torsion of elliptic curves*, Proc. London Math. Soc. **33** (1976), 193–237.

[3] S. Lang, *Fundamentals of diophantine geometry*, New York: Springer–Verlag, 1983.

[4] K. Oguiso and T. Shioda, *The Mordell–Weil lattice of a rational elliptic surface*, Comment. Math. Univ. St. Paul. **40** (1991), 83–99.

[5] T. Shioda, *On the Mordell–Weil lattices*, Comment. Math. Univ. St. Paul. **39** (1990), 211–240.

[6] J. H. Silverman, *The arithmetic of elliptic curves*, New York: Springer–Verlag, 1986.