AN ELEMENTARY PROOF OF THE GROUP LAW FOR ELLIPTIC CURVES

ABSTRACT. We give an elementary proof of the group law for elliptic curves using explicit formulas.

1. INTRODUCTION

In this short note we give an elementary proof of the well–known fact that the addition of points on an elliptic curve defines a group structure. We only use explicit and very well–known formulas for the coordinates of the addition of two points. Even though the arguments in the proof are elementary, making this approach work requires several intricate arguments and elaborate computer calculations. The approach of this note was used by Laurent Théry [Th07] to give a formal proof of the group law for elliptic curves in the formal proof management system Coq.

In the following K will denote an algebraically closed field with $\mathrm{char}(K) > 3$. An elliptic curve is defined as a pair (E, O) where E is a smooth algebraic curve of genus one and $O \in E$ a point.

Proposition 1.1. [Si86, Prop. 3.1] *Let E be an elliptic curve E. Then there exist $a, b \in K$ with $4a^3 + 27b^2 \neq 0$ and an isomorphism of curves*

$$\phi : E \to E_{a,b} := \{(x : y : z) \in \mathbf{P}(K)^2 | zy^2 = x^3 + axz^2 + bz^3\}$$

such that $\phi(O) = [0 : 1 : 0]$. Conversely, for any a, b with $4a^3 + 27b^2 \neq 0$ the variety $(E_{a,b}, [0 : 1 : 0])$ is an elliptic curve.

Corollary 1.2. *Let E be an elliptic curve E. Then there exist $a, b \in K$ with $4a^3 + 27b^2 \neq 0$ and a bijection*

$$\phi : E \to E_{a,b}^{affine} := \{(x, y) \in K^2 | y^2 = x^3 + ax + b\} \cup \{O\}$$

such that $\phi(O) = O$.

In the following we will take the affine point of view, i.e. an elliptic curve E will mean the set $E_{a,b}^{affine}$ for some $a, b \in K$ with $4a^3 + 27b^2 \neq 0$. The point O is called the 'point at infinity'. For points $A, B, C, \dots \in E \setminus \{O\}$ we will write $A = (x_A, y_A), B = (x_B, y_B), C = (x_C, y_C), \dots$ for the coordinates.

Date: June 17, 2009.

Definition. We define $+ : E \times E \to E, (A, B) \mapsto A + B$ as follows. We set $A + O = O + A := O$ for all A. If $(x_A, y_A) = (x_B, -y_B)$, then $A + B := O$. Otherwise $A + B := (x_{AB}, y_{AB})$ where

$$\begin{aligned} x_{AB} &:= \alpha(A,B)^2 - x_A - x_B \\ y_{AB} &:= -y_A + \alpha(A,B)(x_A - x_{AB}) \end{aligned} \tag{1}$$

with $\alpha(A, B) = \frac{y_A - y_B}{x_A - x_B}$ if $x_A \neq x_B$ and $\alpha(A, B) = \frac{3x_A^2 + a}{2y_B}$ if $x_A = x_B$.

Remark. This definition is motivated by geometry. If $A, B \neq O$, then we can take the line through A and B (respectively the tangent to A if $A = B$). This line intersects the curve E in three points (counted with multiplicities) A, B, C. Then define $A + B := -C$ where $-C$ is the reflection of C about the x-axis.

Theorem 1.3. *$(E, +)$ is an abelian group with neutral element O.*

This theorem is of course well known. For example in [Si86] it is shown that the above structure is isomorphic to the group structure of $\mathrm{Pic}^0(E)$, the Picard group of E. In particular $(E, +)$ forms a group. A more geometric argument, that $(E, +)$ defines a group is given in [Ku95, p. 87] or [Hu87]. Perhaps the most elementary proof can be found for $K = \mathbb{C}$ using the Weierstrass function (cf. [La78]). By the Lefschetz principle this shows the theorem for any algebraically closed field of characteristic zero.

We will give a completely elementary proof, just using the above explicit definition of the group structure through formulas. It was always clear that such a proof exists, but it turns out that this direct proof is more difficult than one might have imagined initially. Many special cases have to be dealt with separately and some are non–trivial. Furthermore it turns out that the explicit computations in the proof are very hard. The verification of some identities took several hours on a modern computer; this proof could not have been carried out before the 1980's.

It is clear that "+" is commutative, that O is a neutral element and that the inverse element for $A = (x_A, y_A)$ is given by $-A := (x_A, -y_A)$. The only difficult part is to show that "+" is in fact associative. This proof will require the remainder of this paper.

2. Proof of the associativity law for elliptic curves

In the following let E be a fixed elliptic curve.

We will use the following facts which follow immediately from the definition.

(1) "+" is commutative.
(2) For $A = (x_A, y_A) \in E \setminus \{O\}$ we have $A + A = O$ if and only if $y = 0$.
(3) If $A, B \in E \setminus \{O\}$ and $x_A = x_B$, then $A = B$ or $A = -B$.

Except for three special cases the operation "+" is given by Formula (1). In Section 2.1 we will show the associativity in 3 out of 4 cases in which addition is given by either of the two formulas. This will be done using explicit calculations.

In Section 2.2 we will prove several lemmas, which we will use in Section 2.3 to give the proof in the general case.

2.1. **Proof for the generic cases.** In this section we consider the cases in which only Formula (1) is being used in the definitions of $(A+B)+C$ and $A+(B+C)$.

Lemma 2.1. *Let $A, B, C \in E \setminus \{O\}$. If $A \neq \pm B$, $B \neq \pm C$, $A+B \neq \pm C$ and $B+C \neq \pm A$, then*

$$(A+B)+C = A+(B+C).$$

Proof. Write $(x_1, y_1) := (A+B)+C$ and $(x_2, y_2) := A+(B+C)$. Let

$$\begin{array}{rclrcl} \alpha & := & \frac{y_B - y_A}{x_B - x_A}, & \beta & := & \frac{y_A + y_C - \alpha(2x_A + x_B - \alpha^2)}{x_A + x_B + x_C - \alpha^2}, \\ \gamma & := & \frac{y_B - y_C}{x_B - x_C}, & \tau & := & \frac{y_A + y_B - \gamma(2x_B + x_C - \gamma^2)}{x_A + x_B + x_C - \gamma^2}. \end{array}$$

We get using Formula (1)

$$\begin{array}{rclrcl} x_1 & = & \beta^2 + x_A + x_B - x_C - \alpha^2, & y_1 & = & -y_C + \beta(2x_C - x_A - x_B - \beta^2 + \alpha^2), \\ x_2 & = & \tau^2 + x_B + x_C - x_A - \gamma^2, & y_2 & = & -y_A + \tau(2x_A - x_B - x_C - \tau^2 + \gamma^2). \end{array}$$

Setting

$$\begin{array}{rclrcl} \tilde\alpha & := & y_B - x_A, & \tilde\beta & := & (y_A + y_C)(x_B - x_A)^3 - \tilde\alpha((2x_A + x_B)(x_B - x_A)^2 - \tilde\alpha^2), \\ \tilde\gamma & := & y_B - y_C, & \tilde\tau & := & (y_A + y_B)(x_B - x_C)^3 - \tilde\gamma((2x_B + x_C)(x_B - x_C)^2 - \tilde\gamma^2), \\ \tilde\eta & := & x_B - x_A, & \tilde\mu & := & x_B - x_C. \end{array}$$

one can show that $x_1 = x_2$ is equivalent to

$$\begin{array}{l} \quad (\tilde\beta^2(x_B - x_C)^2 + (((2x_A - 2x_C)(x_B - x_C)^2 + \tilde\gamma^2)(x_B - x_A)^2 - \tilde\alpha^2(x_B - x_C)^2) \\ \quad ((x_A + x_B + x_C)(x_B - x_A)^2 - \tilde\alpha^2)^2)((x_A + x_B + x_C)(x_B - x_A)^2 - \tilde\gamma^2)^2 \\ -\tilde\tau^2((x_A + x_B + x_C)(x_B - x_A)^2 - \tilde\alpha^2)^2(x_B - x_A)^2 = 0 \end{array}$$

and $y_1 = y_2$ is equivalent to

$$\begin{array}{l} \quad (y_A - y_C)((x_A + x_B + x_C)\tilde\eta^2 - \tilde\alpha^2)^3((x_A + x_B + x_C)\tilde\mu^2 - \tilde\gamma^2)^3\tilde\eta^3\tilde\mu^3 \\ +\tilde\beta(((2x_C - x_A - x_B)\tilde\eta^2 + \tilde\alpha^2)((x_A + x_B + x_C)\tilde\eta^2 - \tilde\eta^2)^2 - \tilde\beta^2) \\ \qquad ((x_A + x_B + x_C)\tilde\mu^2 - \tilde\gamma^2)^3\tilde\mu^3 \\ \; -\tilde\tau(((2x_A - x_B - x_C)\tilde\mu^2 + \tilde\gamma^2)((x_A + x_B + x_C)\tilde\eta^2 - \tilde\gamma^2)^2 - \tilde\tau^2) \\ \qquad ((x_A + x_B + x_C)\tilde\eta^2 - \tilde\alpha^2)^3\tilde\eta^3 = 0. \end{array}$$

By abuse of notation we now consider the equations over the polynomial ring $P := \mathbb{Z}[x_A, x_B, x_C, y_A, y_B, y_C, a, b]$. It suffices to show that the equalities hold in P/I where $I := (y_A^2 - x_A^3 - ax_A - b, y_B^2 - x_B^3 - ax_B - b, y_C^2 - x_C^3 - ax_C - b)$. This is equivalent to showing that both left hand sides lie in I. This was shown using the commutative algebra package 'CoCoA'. □

In a very similar way one can show the following two lemmas.

Lemma 2.2. *If $A, B \neq O, A \neq -A, A \neq \pm B, A + A \neq \pm B$ and $A + B \neq \pm A$, then*

$$(A + A) + B = A + (A + B).$$

.

Lemma 2.3. *If $A \neq O, A \neq -A$, $A + A \neq -(A + A)$, $(A + A) + A \neq \pm A$ and $A + A \neq \pm A$, then*

$$(A + A) + (A + A) = A + (A + (A + A)).$$

The next step would be to show that under the usual restrictions, $(A + B) + (A + B) = A + (B + (A + B))$. We will show this without reverting to explicit computations in the proof of Theorem 2.13.

2.2. **Proof of basic properties.**

Lemma 2.4. *For $A, B \in E$ we have*

$$-A - B = -(A + B).$$

Proof. The cases $A = O, B = O$ and $A = -B$ are trivial. In the other cases the lemma follows from an easy calculation using Formula (1). □

Lemma 2.5. *Let $A, B \in E$. If $A + B = A - B$ and $A \neq -A$, then $B = -B$.*

Proof. The cases $A = O$ respectively $B = O$ are trivial. If $A = \pm B$, then $B = -B$ follows easily from the uniqueness of the inverse element. So assume that $A, B \neq O, A \neq \pm B$. Using Formula (1) we get

$$\left(\frac{y_B - y_A}{x_B - x_A}\right)^2 - x_A - x_B = \left(\frac{-y_B - y_A}{x_B - x_A}\right)^2 - x_A - x_B.$$

This simplifies to $-2y_Ay_B = 2y_Ay_B$. Since $A \neq -A$ it follows that $y_A \neq 0$. We get $y_B = 0$ since $\mathrm{char}(K) > 3$, hence $B = -B$. □

Lemma 2.6 (Uniqueness of the neutral element)**.** *Let $A, B \in E$. If $A + B = A$, then $B = O$.*

Proof. The cases $A = O$ and $A = -B$ are trivial. Now assume that $A \neq O, A \neq -B$. Assume that $B \neq O$. Write $(x_C, y_C) := A + B = A = (x_A, y_A)$. It follows from Formula (1) that

$$y_A = y_C = -y_A + \alpha(P, Q)\underbrace{(x_A - x_C)}_{=0} = -y_A$$

i.e. $y_A = 0$, therefore $A = -A$. It follows that

$$A + B = A = -A = -A - B = A - B$$

According to Lemma 2.5 this means that $B = -B$, i.e. $y_B = 0$. In particular $A \neq B$, because otherwise we would get $B = A = A + B = A + A = A - A = O$. According to Formula (1) we get

$$x_A = x_C = \left(\frac{y_B - y_A}{x_B - x_A}\right)^2 - x_A - x_B = -x_A - x_B$$

since $y_A = y_B = 0$. Therefore x_A and $x_B = -x_A - x_A$ are zeros of the polynomial $P := X^3 + aX + b$. It follows that $x_0 = -x_A - x_B = x_A$ is the third zero since the second highest coefficient of P is zero. In particular x_A is a zero of degree 2. This leads to a contradiction, since we assumed that the discriminant $4a^3 + 27b^2$ of the polynomial $X^3 + aX + b$ is non–zero, i.e. the polynomial has distinct zeros. □

Lemma 2.7. *Let $A \in E$. If $A \neq -A$ and $A + A \neq -A$, then $(A + A) - A = A$.*

Proof. The cases $A = O$ and $A + A = O$ are trivial. The general case follows from an easy computation using Formula (1). □

Lemma 2.8. *Let $A, B \in E$. If $A + B = -A$, then $B = -A - A$.*

Proof. The cases $A = O$, $B = O$, $A = B, A = -B$ are trivial. If $A = -A$, then $-A + B = -A$. Using Lemma 2.6 it follows that $B = O$. Hence $B = O = A - A = -A - A$. Now assume that $A \neq \pm B$ and $A \neq -A$, $A, B \neq O$. From $-A = A + B$ it follows that

$$x_A = \left(\frac{y_A - y_B}{x_A - x_B}\right)^2 - x_A - x_B$$

which is equivalent to $2y_Ay_B = y_A^2 + ax_B + b - 2x_A^3 + 3x_A^2x_B$, squaring we get

$$4x_B^3y_A^2 - x_B^2(3x_A^2 + a)^2 + x_B(2a^2x_A + 6x_A^5 - 12bx_A^2) - (y_A^2 - b)^2 + 4ax_A^4 + 8bx_A^3 = 0$$

which in turn is equivalent to

$$\left(x_B - \left(\left(\frac{3x_A^2 + a}{2y_A}\right)^2 - 2x_A\right)\right)(x_B - x_A)^2 = 0.$$

Since we excluded the case $x_A = x_B$ we get

$$x_B = \left(\frac{3x_A^2 + a}{2y_A}\right)^2 - 2x_A$$

i.e. $B = A+A$ or $B = -(A+A) = -A-A$. If $A+A = -A$, then $B = \pm(A+A) = \pm A$. Hence $A + A \neq -A$. By Lemma 2.7 it follows that $B = -A - A$ is a solution for the equation $A + B = -A$. If $B = A + A$ is also a solution, then $A + B = A + (A + A) = -A = A-(A+A) = A-B$. Since $A \neq -A$ it follows from Lemma 2.5, that $B = -B$. Therefore $B = -B = -A - A$. □

Lemma 2.9 (Cancelation rule)**.** *Let $A, B, \tilde{B} \in E$. If $A + B = A + \tilde{B}$, then $B = \tilde{B}$.*

Proof. If $A = O$, then immediately $B = \tilde{B}$. The cases $B = O$ and $A + B = O$ follow immediately from the uniqueness of the neutral element (Lemma 2.6) and the uniqueness of the inverse element. If $A + B = A + \tilde{B} = -A$, then using 2.8 we see that $B = -A - A$ and $\tilde{B} = -A - A$.

We therefore can assume that $A, B, \tilde{B} \neq O$ and $A + B = A + \tilde{B} \neq O, A + B \neq -A$. Writing $A + B = A + \tilde{B} =: (x_C, y_C)$ we get

$$\begin{array}{rcccl} x_C & = & \alpha(A,B)^2 - x_A - x_B & = & \alpha(A,\tilde{B})^2 - x_A - \tilde{x}_B \\ y_C & = & -y_A + \alpha(A,B)(x_A - x_C) & = & -y_A + \alpha(A,\tilde{B})(x_A - x_C). \end{array}$$

From $A + B \neq \pm A$ it follows that $x_A \neq x_C$, from the second equation we get $\alpha(A, B) = \alpha(A, \tilde{B})$. Using the first equation we get $x_B = \tilde{x}_B$, i.e. $B = -\tilde{B}$, or $B = \tilde{B}$. We consider the following two cases:

(1) If $A = -A$, then $B, \tilde{B} \neq -A = A$, hence

$$\frac{y_B - y_A}{x_B - x_A} = \alpha(A, B) = \alpha(A, \tilde{B}) = \frac{\tilde{y}_B - y_A}{\tilde{x}_B - x_A}.$$

Since $x_B = \tilde{x}_B$ we get $y_B = \tilde{y}_B$, therefore $B = \tilde{B}$.

(2) If $A \neq -A$, then assume that $B = -\tilde{B}$. It follows that $A + B = A + \tilde{B} = A - B$. By Lemma 2.5 $B = -B$, since $A \neq -A$. Therefore $B = \tilde{B}$.

□

Lemma 2.10. *For any $A, B \in E$ we have*

$$(A + B) - B = A.$$

Proof. The cases $A = O, B = O$ respectively $A = -B$ are trivial. The case $A = B$ follows from Lemma 2.7. If $A + B = -B$ and $A \neq -B$, then by Lemma 2.8 $A = -B - B$, hence $(A + B) - B = -B - B = A$.

Now assume that $A, B \neq O, A \neq \pm B, A + B \neq -B$. This case follows from an explicit computation using Formula (1). □

Corollary 2.11. *Let $A, B, C \in E$. If $A + B = C$, then $A = C - B$.*

Proof. From Lemma 2.10 we get $A + B = A + (C - A)$, the corollary now follows from Lemma 2.9. □

2.3. Completion of the proof.

Lemma 2.12. *Let $A, B, C \in E$. Assume that*

(1) *$(A + B) \neq C$ and $A \neq (B + C)$, or*
(2) *$A = B$, or $B = C$, or $A = C$, or*
(3) *$O \in \{A, B, C, A + B, B + C, (A + B) + C, A + (B + C)\}$,*

then

$$(A+B)+C = A+(B+C).$$

Proof. The cases $A=O$, $B=O$, $C=O$ and $A=C$ are trivial. The cases $A=-B$ and $C=-B$ follow immediately from Lemma 2.10. If $A+B=-C$, then by Lemma 2.10

$$(A+B)+C = O = A-A = A+(B+(-B-A)) = A+(B+C).$$

The case $B+C=-A$ works the same. We thus established part (3) of the lemma.

We can therefore assume that $A,B,C \neq O$, $A \neq C, B \neq -A, -C, A+B \neq -C$ and $B+C \neq -A$.

If $A=B$, then we have to show that $(A+A)+C = A+(A+C)$. This follows from Lemmas 2.2 ($C \neq A+A$) and 2.3 ($C = A+A$). The case $B=C$ again works the same. This shows part (2) of the lemma.

The remaining cases of part (1) now follow immediately from Lemma 2.1. □

Theorem 2.13. *Let $A,B,C \in E(K)$, then*

$$(A+B)+C = A+(B+C)$$

Proof. By Lemma 2.12. we only have to prove the theorem for A,B,C with $A+B=C$ or $B+C=A$. It is clearly enough to consider only the case $A+B=C$. We therefore have to show that

$$(A+B)+(A+B) = A+(B+(A+B)).$$

By Lemma 2.12 we can assume that $A,B,C,A+B,B+C,(A+B)+C,A+(B+C) \neq O$ and A,B,C are pairwise different.

If $(A+B)+(A+B) = -A$, then $A+B = (-B-A)-A$ by Corollary 2.11. Furthermore $(-B-A)-A = -B+(-A-A)$ by the second part of Lemma 2.12, hence $A+B = -B+(-A-A)$. We get

$$\begin{aligned} A+(B+(A+B)) &= A+(B+(-B+(-A-A))) = A+(-A-A) = \\ &= -A = (A+B)+(A+B). \end{aligned}$$

If $(A+B)+(A+B) \neq -A$, then $((A+B)+(A+B))-A = (A+B)+((A+B)-A)$ by the second part of Lemma 2.12. Hence

$$\begin{aligned} ((A+B)+(A+B))-A &= (A+B)+((A+B)-A) = (A+B)+B = \\ &= (A+(B+(A+B)))-A. \end{aligned}$$

From Lemma 2.9 it follows, that $(A+B)+(A+B) = A+(B+(A+B))$. □

References

[Hu87] D. Husemoller, *Elliptic curves*, Graduate Texts in Mathematics, 111. Springer-Verlag, New York (1987)

[Ku95] E. Kunz, *Ebene algebraische Kurven*, Der Regensburger Trichter, Band 23, Regensburg (1995)

[La78] S. Lang, *Elliptic curves: Diophantine analysis*, Grundlehren der Mathematischen Wissenschaften, 231. Springer-Verlag, Berlin-New York (1978)

[Si86] J. H. Silverman, *The Arithmetic of Elliptic Curves*, Springer Verlag, Berlin–Heidelberg–New York (1986)

[Th07] L. Théry, *Proving the group law for elliptic curves formally*, INRIA Rapport technique n. 0311, available from
`http://hal.inria.fr/inria-00129237/en/`