The group law on a nonsingular plane cubic curve

Enrique Acosta

Fall 2010

Let C be a nonsingular plane cubic curve. In this short paper we show how to define a group law on the set of points of C. Parts of this proof are in most books on elliptic curves, but they are usually incomplete because with elementary arguments there are lots of special cases which have to be analyzed separately and the authors probably don't have the time or space to do them all. Other authors rely on the concept of continuity to argue why their arguments also work in degenerate cases. The proof given here relies on a deep theorem from the theory on algebraic curves (Noether's theorem) and deals with all the cases simultaneously, thus giving a cleaner and complete proof, and at the same time exposing the reader to beautiful concepts and theorems from the theory of algebraic curves.

Definition 1. Let $a, b \in C$ be any two points on C. We define $a \cdot b \in C$ by the following (see Figure 1):

- $a \cdot b$ is the third intersection point of the line joining a and b with C if a and b are distinct.
- $a \cdot a$ is the third intersection point of the tangent line of C at a.

Note that these definitions take into account intersection multiplicities[1] and that the intersection of a tangent line to C at its intersection point is either 2 or 3 since C is a cubic. For example, if $a \neq b$ and the line through a and b is tangent to C at a, then $a \cdot b$ is a. Similarly, if a is an inflection point of C, then $a \cdot a$ is a again.

It is direct from the definition that $a \cdot b = b \cdot a$ for any a, b. Another simple property we will be using is the following (see Figure 1).

[1] For an elementary introduction to intersection multiplicities of plane curves including elementary ways to compute them and an elementary proof of Bezout's Theorem see Hilmar and Smyth's article *Euclid meets Bezout: Intersecting algebraic plane curves with the Euclidean algorithm* [4].

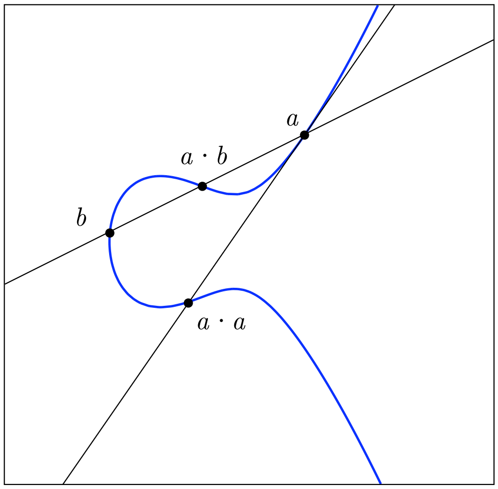

Figure 1: Cubic $y^2 = x(x^2 - 2.7x + 2)$

Lemma 2. *$a \cdot (a \cdot b) = b$ for any $a, b \in C$.*

One last simple property which will be useful in what follows is

Lemma 3. *$\cdot$ satisfies the cancelation law. That is, if $a \cdot b = a \cdot c$ then $b = c$.*

Proof. If $a \cdot b = a \cdot c$ then $a \cdot (a \cdot b) = a \cdot (a \cdot c)$. Now use Lemma 2 to conclude that $b = c$. □

Even though $\cdot$ seems like a nice operation, it has various flaws among which are the fact that it has no identity element and even worse, it is not associative (eg. $a \cdot (a \cdot b) \neq (a \cdot a) \cdot b$, see Lemma 2 and Figure 1). To define the group operation on C we will use $\cdot$ and we will need to fix a point $o \in C$. The group law $+$ on the cubic will depend on the choice of o, and is defined by the following rule:

Definition 4. Let $a, b \in C$ be any two points on C, we define $a + b \in C$ as

$$a + b := o \cdot (a \cdot b).$$

See Figure 2. We will now prove that $+$ defines an operation which makes C into a group with identity element o.

$+$ is commutative: $a + b = o \cdot (a \cdot b) = o \cdot (b \cdot a) = b + a$

o is the identity element: $o + a = o \cdot (o \cdot a) = a$ by Lemma 2.

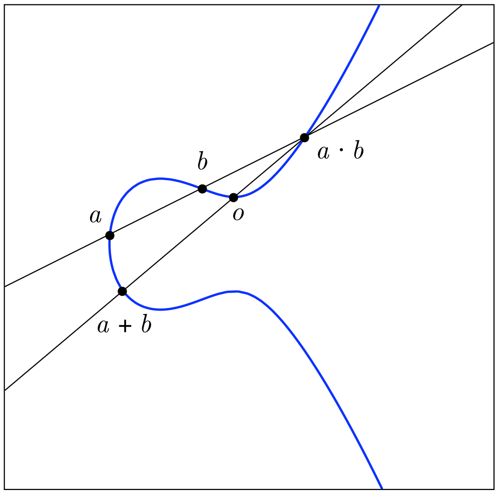

Figure 2: Definition of the group law

Inverses exist: The inverse of a is $a \cdot (o \cdot o)$ as the following shows:

$$\begin{aligned} a + a \cdot (o \cdot o) &= o \cdot (a \cdot (a \cdot (o \cdot o))) \\ &= o \cdot (o \cdot o) \\ &= o \end{aligned}$$

where we have used Lemma 2 twice.

Associativity: Using Lemma 3 one easily sees that proving that $(a+b)+c = a + (b+c)$ is equivalent to proving the following identity:

$$a \cdot (o \cdot (b \cdot c)) = c \cdot (o \cdot (a \cdot b)). \tag{1}$$

To prove this identity we will need the following result from Fulton's book [1] (Proposition 3, Chapter 5 section 6 p. 124) (see also [5](Theorem 7.3, p.122))

Theorem 5. *If the nine points of intersection of two cubics F and G are simple points of F, then every cubic though eight of this points also goes through the ninth.*

Note that the statement allows the points to have multiplicities, i.e, there could be points of intersection between F and G with multiplicity greater than one, and the conclusion of the theorem holds *counting multiplicities* as long as the points of intersection are simple points of F. For more details see the comments following the statement of the theorem in the book. We will use this theorem in the case when F is smooth when the hypothesis on the simplicity of the points is automatically satisfied.

The proof of theorem 5 in Walker [5] uses the language of places which is somewhat outdated. The presentation in [1] is more modern. A modern and relatively elementary presentation of Noether's Theorem in its maximum generality which is the fundamental tool in the proof of theorem 5 is given in [2], theorem 4.1. See also Noether's $AF + BG$ Theorem in Griffiths and Harris [3].

Before we go on, let us state a beautiful corollary of theorem 5. Recall that a flex of an algebraic curve is a non-singular point at which the tangent line to the curve intersects with multiplicity 3 or more (the tangent line being by definition the line that intersects the curve multiplicity 2 or more at that point). In particular, flexes to cubics are smooth points where the tangent line intersects with multiplicity 3 (by Bezout), and don't intersect the curve at any other point.

Corollary. *A line joining two flexes of a cubic passes through a third flex.*

Now let's get back to the proof of associativity of the + operation on a cubic. Remember we need to prove the equality (1) no matter what $a, b, c \in C$ are. Define the following lines and cubics:

$D1$ is the union of the following three lines:

- The line through a and $o \cdot (b \cdot c)$
- The line through b and c
- The line through $a \cdot b$ and o

$D2$ is the union of the following three lines:

- The line through c and $o \cdot (a \cdot b)$
- The line through a and b
- The line through $b \cdot c$ and o

If the points defining any of the lines agree, then we take the tangent line to C at the given point. With these definitions it follows that both $D1$ and $D2$ go through the following eight points of C counting multiplicities:

$$\begin{array}{cccc} a & c & a \cdot b & o \cdot (a \cdot b) \\ b & o & b \cdot c & o \cdot (b \cdot c) \end{array}$$

Moreover, the nine points of intersection between C and $D1$ are the eight points above and the extra $a \cdot (o \cdot (b \cdot c))$, and the nine points of intersection between C and $D2$ are the same eight points and the extra $c \cdot (o \cdot (a \cdot b))$. Now use theorem 5 with C as F, which by hypothesis is smooth, $D1$ as G, and $D2$ as the third cubic. The conclusion of the theorem tells us that $D2$ intersects C at $a \cdot (o \cdot (b \cdot c))$ with the same multiplicity as $D1$ does. This implies that $a \cdot (o \cdot (b \cdot c)) = c \cdot (o \cdot (a \cdot b))$ which is what we wanted to prove. This concludes the proof that + is associative.

Simplifications Remember that the inverse of an arbitrary point a is $a\cdot(o\cdot o)$, but if o is a flex then $o\cdot o = o$, and so the inverse of a is just $a\cdot o$. Thus, the group law takes on a simpler form if we take o to be a flex of C.

For example, if the cubic is in Weierstrass form with affine equation $y^2 = x^3 + Ax + B$, then the point at infinity O is a flex and the usual description of the group law for elliptic curves with its "reflection with respect to the x-axis" is a special case of the construction above where we take O as o. The "reflection" of $a\cdot b$ in the computation of $a+b = o\cdot(a\cdot b)$ is just the computation of the third intersection point of the line joining $a\cdot b$ and the point O at infinity, since this line is just the vertical line through $a\cdot b$. The fact that O is a flex explains why the inverse of a is just the reflection of a along the x-axis since this in just $a\cdot O$.

A Note About the Field Throughout the paper we assumed that the field was algebraically closed to be able to use Bezout's theorem. It is important however to note that the group operation is closed when you restrict it to points lying on a field as long as the coefficients of the equation also lie on the field. For example, if an elliptic curve is given by a Weierstrass affine equation $y^2 = x^3 + Ax + B$ with $A, B \in \mathbb{Q}$, then the set of points with rational coordinates is a subgroup of the whole group over the complex numbers. The explanation of why this is the case is as follows:

If you take two points p, q with rational coordinates, then you can parametrize the line joining them as $\vec{r}(t) = p + t\overrightarrow{pq}$, and then to find the x and y coordinates of the third intersection point you plug in $\vec{r}(t)$ into the equation of the curve to find the other root besides $t = 0$ and $t = 1$. This is an equation in t degree 3 with rational coefficients because $\overrightarrow{pq}$ has rational components. Two of the roots of the equation are rational, and since the product of the roots is the constant coefficient of the equation, then the third root t_0 is rational. Therefore, the third point of intersection is given by $\vec{r}(t_0) = p + t_0\overrightarrow{pq}$ which therefore has rational coordinates.

This argument obviously works for any field, so this shows that the group law is closed when you restrict to points that belong to a field k as long as the equation has coefficients that belong to k.

Final Remarks As an illustration to see why the proof of associativity actually requires this machinery and is not so straight forward, the reader should have a look at Figure 1, and using as o the point at infinity, draw the pictures to find the point corresponding to $a + a + b$ in the two ways $2a + b$ and $a + (a + b)$. This would be one of the degenerate cases that would have to be studied separately is we require all the points to be distinct, which is what is sometimes done in the presentations of this proof.

References

[1] W. Fulton. *Algebraic curves.* WA Benjamin, Inc., 1969.

[2] W. Fulton. Adjoints and Max Noether's Fundamentalsatz. In *Algebra, arithmetic and geometry with applications: papers from Shreeram S. Abhyankar's 70th birthday conference*, page 301. Springer Verlag, 2004.

[3] Phillip A Griffiths and Joseph Harris. *Principles of algebraic geometry.* Wiley classics library, 1994.

[4] J. Hilmar and C. Smyth. Euclid meets Bezout: Intersecting algebraic plane curves with the Euclidean algorithm. *American Mathematical Monthly*, 117(3):250–260, 2010.

[5] R.J. Walker. *Algebraic curves.* Princeton University Press New Jersey, 1950.