Don't Stop Believin'
There are Group Laws on the Cubics

Josh Mollner

July 29, 2008

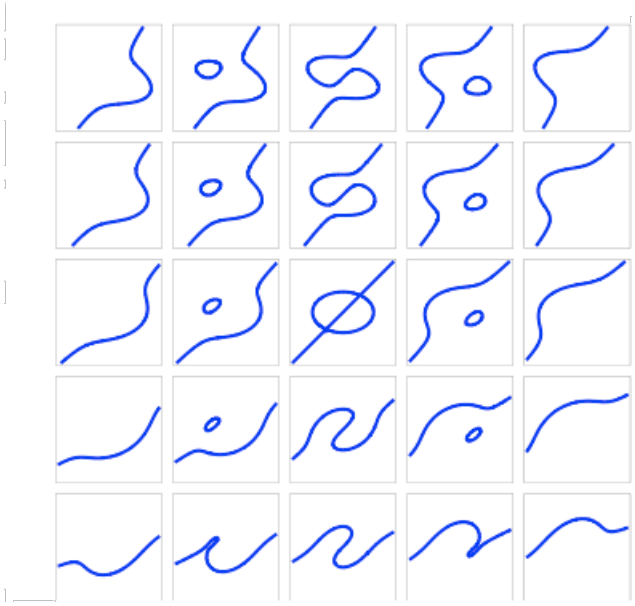
1 Introduction

The study of cubic curves is an important area of mathematics. For example, they were used in Andrew Wiles' proof of Fermat's Last Theorem. An important theorem regarding cubics is Mordell's Theorem, which states that the set of rational points on a rational, non-singular, irreducible cubic with at least one rational point is a finitely-generated abelian group. (These terms will be defined later in the introduction and illustrated with examples.) While we will not show that this group is finitely generated in this paper, we will be able to prove or at least sketch the proof of most of the other claims of Mordell's Theorem.

We will begin by defining a cubic plane curve. A cubic equation is any polynomial equation of the form

$$P(x, y) = ax^3 + bx^2y + cxy^2 + dy^3 + ex^2 + fxy + gy^2 + hx + iy + j = 0.$$

A cubic curve is the set of points in $\mathbb{R}^2$ satisfying a cubic equation. The following is a sample of the wide variety of curves which are cubics:



Note that the cubic depicted in the center of the graphic is the union of an ellipse and a line. It is in fact true that the union of any conic and any line is a cubic. To see that this is true, let $P_1(x, y) = 0$ be the equation of a conic, and $P_2(x, y) = 0$ be the equation of a line. P_1 is a polynomial of degree two and P_2 is a polynomial of degree one, and so $P(x, y) = P_1(x, y)P_2(x, y)$ is a polynomial of degree three. Consequently, $P(x, y) = 0$ is a cubic, but we note that this is simply the set of solutions to $P_1(x, y) = 0$ union the set of solutions to $P_2(x, y) = 0$. Similarly, the union of any three lines is also a cubic.

From a geometric perspective, any cubic that is the union of a line and a conic or the union of three lines is *reducible*. From an algebraic perspective, a cubic is reducible if its defining equation can be factored as a product of polynomials with lower degrees. All other cubics are said to be *irreducible*.

We will say that a cubic is *singular* at a point (a, b) if:

$$\frac{\partial P}{\partial x}(a, b) = 0 \text{ and } \frac{\partial P}{\partial y}(a, b) = 0.$$

Geometrically, a point of singularity is a cusp or a point at which the cubic loops back upon itself. A cubic that is not singular at any of its points is called *non-singular*.

From here, we will be primarily interested in non-singular, irreducible cubics. The crucial problem with reducible cubics is that it is possible to draw a line which intersects the cubic at infinitely many points (namely, the linear component of the reducible cubic). This is problemic because we will wish to use a theorem that any line intersecting a cubic at two points then intersects the cubic at exactly three points. The

crucial problem with non-singular cubics is that a tangent line cannot be constructed at a singular point.

In order to state Mordell's Theorem, we will need to define a rational point. We will say that a point (x, y) in $\mathbb{R}^2$ is ***rational*** if x and y are both rational numbers. In the paper, we will demonstrate that the set of rational points on a cubic curve can be given the structure of an abelian group. An ***abelian group*** is a set of elements, G, together with an operation, $+$, such that the following five properties hold:

1. Closure: For all x and y in G, $x + y$ is also in G.
2. Associativity: For all x, y, and z in G, $(x + y) + z = x + (y + z)$.
3. Identity Element: There exists a unique element i in G such that for all x in G, $x + i = x$.
4. Inverse Element: For each x in G, there exists an element y in G such that $x+y = i$.
5. Commutativity: For every x and y in G, $x + y = y + x$.

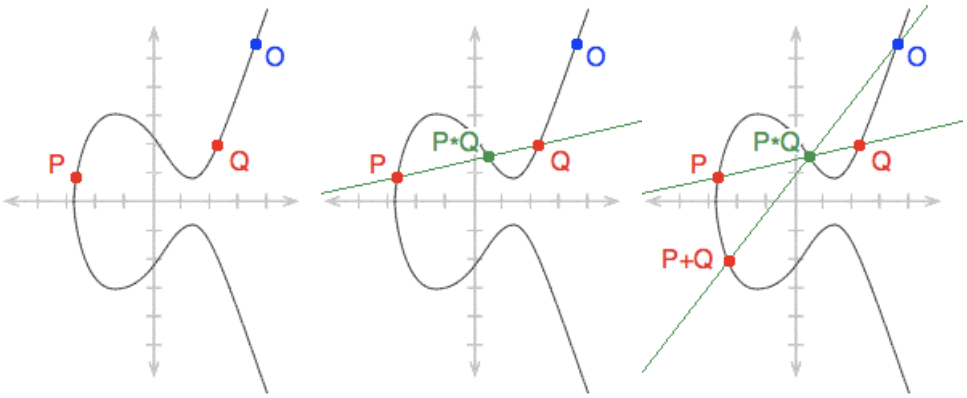
2 The Group Law on Points of a Cubic

We will first show that the set of all points on an irreducible, non-singular cubic can be formed into an abelian group. Later, we will show that if we restrict our attention to the set of rational points on such a curve, they also form a group, provided that we impose a few further conditions on the types of cubics under consideration.

We will begin by defining the operation of addition of points. Given an irreducible, non-singular cubic curve, C, we choose any point on the curve and label it O. We will later show that O is the identity element in the group of points on C. Then given two points P and Q on the curve, we define $P + Q$ as follows:

- Draw the line connecting P and Q. It will intersect the curve at a third point, which we denote $P * Q$.
- Draw the line connecting $P*Q$ and O. It will intersect the curve at a third point, which will be defined to be $P + Q$.

Note that according to this notation, we can write $P + Q = (P * Q) * O$ for any two points on the cubic curve P and Q.



A few concerns might arise after reading the above definition. First, it is not obvious that a line intersecting the curve at two points also intersects the curve in a third point. Hence, this process may not result in a well-defined operation. Even worse, it may be possible for a line to intersect a cubic at four or more points, and so there may not be a well-defined method for identifying the third point of intersection. Finally, it is not obvious that this operation will satisfy the five properties of an abelian group. Nevertheless, we must remind ourselves to not stop believin' and to hold on to that feelin'.

2.1 Well-Definedness

We will first consider the well-definedness of addition. Assuming that we are working in projective space, $\mathbb{P}^2_{\mathbb{R}}$, we will show that any line intersecting a cubic curve at two points in fact intersects it at exactly three points, provided that we count points of intersection correctly. That is, we say that a line which is tangent to the cubic intersects it twice at the point of tangency. We also say that a line which is tangent to the curve at an inflection point intersects the curve thrice at the point of inflection.

Upon proving the following theorem, it will be apparent that the operation of addition of points defined above will be well-defined, since we will always be able to find the third point of intersection which the process requires.

Theorem. *Let l be a line that intersects an irreducible cubic C at least twice, counting multiplicities. Then l intersects C exactly three times, counting multiplicities.*

Proof. There exists a linear transformation which will map l to $y = 0$. We apply this linear transformation to l and C. This will not change the fact that C is irreducible, nor will it change the intersection multiplicities of l with C. Thus proving this theorem in general is equivalent to proving it in the case where l is the line $y = 0$.

Let $P(X,Y,Z)$ be a homogeneous polynomial that defines C in $\mathbb{P}^2$. Since C is a cubic, we can write:

$$P(X,Y,Z) = aX^3+bX^2Y+CX^2Y+dY^3+eX^2Z+fXYZ+gY^2Z+hXZ^2+iYZ^2+jZ^3.$$

Moreover, since C is irreducible, y is not a factor of P, and so a, e, h, and j are not all zero. Consequently, $P(X,0,Z) = aX^3 + eX^2Z + hXZ^2 + jZ^3$ defines a cubic.

Intersections of C and $y = 0$ correspond to solutions of $P(X,0,Z) = 0$. Since we assumed that C intersects $y = 0$ twice, we know that there are two solutions to $P(X,0,Z) = 0$, say $[b_1 : 0 : a_1]$ and $[b_2 : 0 : a_2]$.

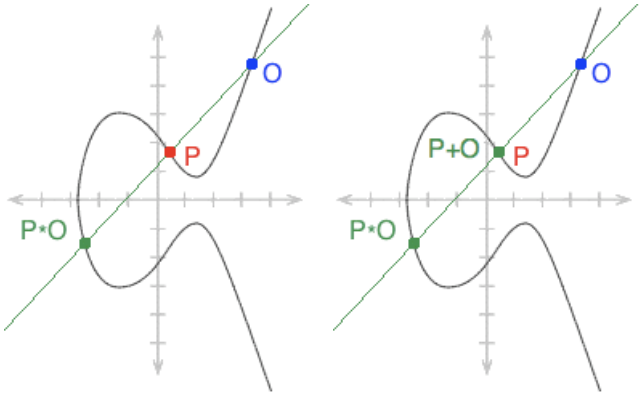
Thus $(a_1X - b_1Z)$ and $(a_2X - b_2Z)$ divide $P(X,0,Z)$. However, since $P(X,0,Z) = 0$ is a cubic equation, when two terms of degree one are factored out of it, we are left with another term of degree one, of the form $(a_3X - b_3Z)$. Thus, $[b_3 : 0 : a_3]$ is another point of intersection of C with $y = 0$. Therefore, C intersects the $y = 0$ exactly three times.

□

2.2 Closure, Identity, Inverses

Now that we have shown the operation, $+$, is well-defined, it only remains to demonstrate that it turns the points of the cubic into an abelian group. It is obvious that the property of closure is satisfied, since $P + Q$ was defined to be the third point of intersection of a line with the cubic.

Next we show that the given point O is the identity. Let $P \in C$. We now compute $P + O$. We first draw the line connecting P and O. It will intersect the curve at a third point, which we label $P * O$. Notice that P, $P * O$, and O are collinear, so the line connecting O and $P * O$ is identical to the line connecting P and O. Thus, P is the third point of intersection on the line connecting O and $P * O$. Therefore, $P = P + O$, and O is the identity element.

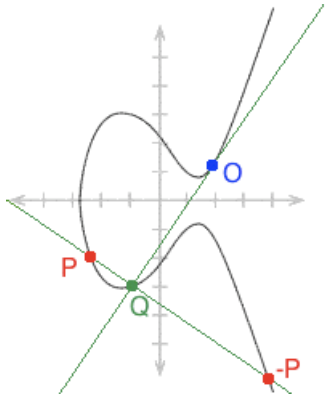


In order to show that O is unique, we assume that there exists another element, say S, which is also the identity. Then $P + S = (P * S) * O = P$. Consequently, P is the third point of intersection of the line joining $P * S$ and O, and so P, O, and $P * S$ are collinear and on the cubic. However, P, S, and $P * S$ are by construction also collinear points on the cubic. So $S = (P * S) * P$ and $O = (P * S) * P$. We conclude that $S = O$. Therefore, O is the unique identity element of our group.

Given a point P, we must show how to construct its inverse, that is, the point $-P$, such that $P + (-P) = O$. Since C is non-singular, there exists a tangent line at every point. We construct the tangent line to C at O. This line intersects the curve twice at O, and so it will have a third point of intersection with C, which we label Q. We then draw the line connecting Q to P, and we claim that the third point of intersection of this line with C, $Q * P$, is $-P$. It is easy to see that $P + (-P) = O$, since Q is by construction $P * (-P)$. Then by construction, the line joining Q and O is tangent to C at O, and so it intersects the curve twice at O. Consequently, $Q * O = O$. Thus, we have

$$P + (-P) = (P * (-P)) * O = Q * O = O.$$

In conclusion, given a point P on the cubic, $-P$ is the point $(O * O) * P$.



2.3 Associativity

The property of associativity is the most difficult to prove. A rigorous proof of this property is beyond the scope of this paper. However, we will include a proof that is valid in all but a few special cases and sketch a proof which is completely general. The former proof will need to make use of the following theorem, the Cayley-Bacharach Theorem, which is a classical result in Algebraic Geometry.

Theorem. *Let C_1 and C_2 be two cubics which intersect in exactly nine points. Suppose C is a third cubic which passes through eight of these nine points. Then C also passes through the ninth point.*

Proof. The general equation for a cubic is:

$$P(x, y) = ax^3 + bx^2y + cxy^2 + dy^3 + ex^2 + fxy + gy^2 + hx + iy + j = 0.$$

A cubic is determined by the ten coefficients $a, b, c, d, e, f, g, h, i$, and j. Since we can multiply the entire equation through by a scalar and still get the same curve, the set of all possible cubics is a nine-dimensional vector space. When we restrict the set of all cubics to the set of cubics which pass through a particular point, we impose one linear condition on the ten coefficients, and we restrict the set of cubics by one dimension. Consequently, given eight sufficiently general points, there is a one-dimensional family of cubics which pass through them. Since the nine points under consideration are the intersection of two cubics, it is an immediate consequence of Bezout's Theorem that no four of them are collinear and no conic can be drawn which contains seven of them. It can also be shown that under these conditions, the points meet the requirements of being sufficiently general.

The set of all cubics is nine-dimensional, and the set of cubics passing through these eight given points is one-dimensional. Hence, we can think of this subset of the cubics as a line in $\mathbb{R}^9$. If this line passes through the origin, then it is a one-dimensional vector space, and is spanned by any of its elements. If this line does not pass through the origin, then there exists a plane in $\mathbb{R}^9$ which contains both this line and the origin. This plane is a two-dimensional vector space, and is consequently spanned by any two linearly-independent elements of the space. Since the line lies in the plane, any element of the line can also be written as a linear combination of these two elements. Since the line does not pass through the origin, any two distinct elements of the line are linearly independent, and so span the line. In either case, given two distinct elements on the line, any point on the line can be written as a linear combination of these two elements.

Let $P_1(x, y) = 0$ and $P_2(x, y) = 0$ be the defining cubic equations for C_1 and C_2. Since C_1 and C_2 are distinct curves, from the above argument, the entire family of cubics passing through the eight points consists of cubic equations which are linear combinations of P_1 and P_2. Thus, the cubic C has the equation $P(x, y) = \lambda_1 P_(x, y) + \lambda_2 P_2(x, y) = 0$ for some λ_1 and λ_2.

At the ninth point of intersection of C_1 and C_2, $P_1(x, y) = 0$ and $P_2(x, y) = 0$, and so $P(x, y) = 0$ at that point as well. Since $P(x, y)$ vanishes there, we also have that C contains that point.

□

Having established the Cayley-Bacharach Theorem, we can use it to prove that addition of points on cubics is associative in all but a few special cases. That is, for any three points, P, Q, and R, on a cubic curve, C, that $(P + Q) + R = P + (Q + R)$.

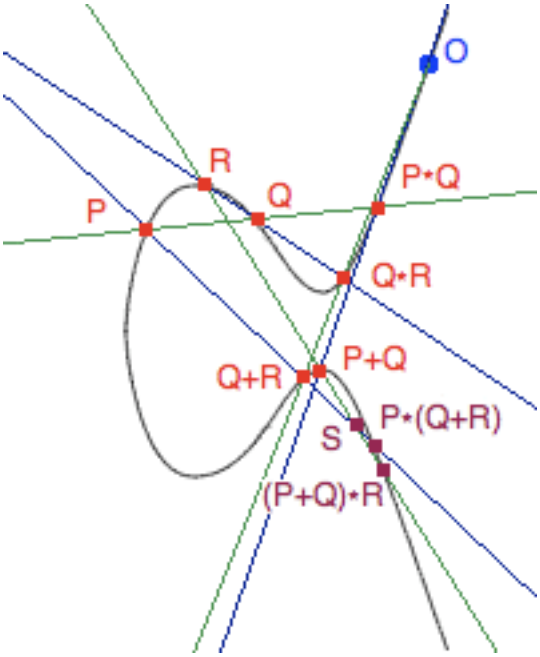
Proof. We note it is enough to show that $(P+Q)*R = P*(Q+R)$. Since $(P+Q)+R$ is the third point of intersection of the line joining $(P+Q)*R$ and O and $P+(Q+R)$ is the third point of intersection of the line joining $P*(Q+R)$ and O, if $(P+Q)*R = P*(Q+R)$, then we will also have $(P+Q)+R = P+(Q+R)$.

We know that $(P + Q) * R$ is the third point of intersection of the line joining $P + Q$ and R with the cubic. Similarly, we know that $P * (Q + R)$ is the third point of intersection of the line joining P and $Q+R$ with the cubic. Let S be the point at which these two lines intersect. If S lies on the cubic, then S must in fact be the third point of intersection of the line joining P and $Q + R$ with the cubic, and so $S = (P+Q)*R$. Similarly, if S lies on the cubic, $S = P * (Q + R)$. Thus, if these two lines intersect at a point on the cubic, we have $(P + Q) * R = P * (Q + R)$. Consequently, to prove associativity, it will be enough to show that S lies on the cubic, C.

We know that R, $P+Q$, and S are collinear, that P, Q, and $P*Q$ are collinear, and that O, $Q * R$, and $Q + R$ are collinear. Consequently, there exists a degenerate cubic which is the the union of these three lines. We call this cubic C_1.

We also know that P, $Q+R$, and S are collinear, that R, Q, and $Q*R$ are collinear, and that O, $P * Q$ and $P + Q$ are collinear. Thus there also exists a degenerate cubic curve consisting of the union of these three lines, which we call C_2.

Then the two cubics C_1 and C_2 intersect in the nine points $O, P, Q, R, P * Q, Q * R, P + Q, Q + R$, and S. By the way in which these points were constructed, we know that our original curve, C, passes through the first eight. Consequently, by the above theorem, we have that C also passes through S, as required. This proves associativity. □

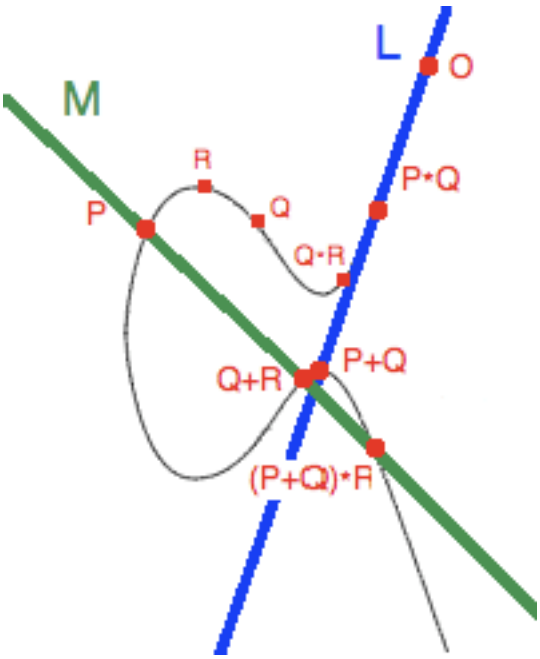


We must note that this proof is not completely general, because it may be possible that $O, P, Q, R, P * Q, Q * R, P + Q, Q + R$, and S may not all be distinct points. In this case, we would not be able to appeal to Cayley-Bacharach, and hence we would not be able to prove associativity in this manner.

This might lead you to stop believin', but there is a method to prove associativity in general. One way to do it is to use the explicit formulas for addition of a cubic in Weierstrass normal form that will be developed later in this section to check all the possible special cases.

Another method, a proof sketch of which is included below, uses intersection theory to show that it is possible to construct a curve, K, of degree two which passes through the six points O, $P * Q$, $P + Q$, $Q + R$, $(P + Q) * R$, and P, which may not all be distinct, but are listed by multiplicity (that is, if $P + Q = O$, for example, then K is tangent to the cubic, C at that point). We then consider the line through O, $P * Q$, and $P + Q$, which we denote L (these points are collinear by construction). Clearly, this line intersects K three times, counting multiplicities. Since K has degree two and L has degree one, this would contradict Bezout's Theorem unless K is reducible and L is a factor of K. Then, we can write $K = LM$, where $M = 0$ is another line.

It can also be shown that M passes through P, $(P + Q) * R$, and $Q + R$. But then $(P + Q) * R$ must be the third point of intersection of the line through P and $Q + R$, and so $(P + Q) * R = P * (Q + R)$.



2.4 Commutativity

We have shown that addition of points forms the points of a cubic into a group. In order to show that it is actually an abelian group, we must also show that for any two points on the curve, P and Q, that $P+Q = Q+P$. This is clear, because the line joining P and Q is the same as the line joining Q and P, so the third points of intersection will be the same. Thus $P * Q = Q * P$, and so $P + Q = Q + P$.

2.5 If O is an Inflection Point

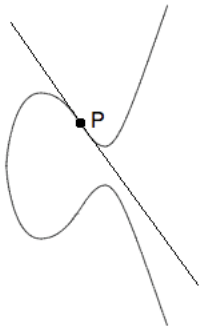
While this procedure turns the points of an irreducible, non-singular cubic into an abelian group for any choice of O, the case in which O is an inflection point is particularly interesting. By *inflection point*, we mean a point at which the tangent line to the cubic intersects the cubic with multiplicity three.

An example will make this clear. The tangent line of $y = x^3$ at $P = (0, 0)$ is $y = 0$. Plugging $y = 0$ into the equation of the cubic yields $x^3 = 0$, which gives the solution $x = 0$ with multiplicity three. Consequently, $y = 0$ intersects $y = x^3$ at $(0, 0)$ with multiplicity three, and so $P = (0, 0)$ is an inflection point of $y = x^3$.

Theorem. *A point P is an inflection point if and only if $P * P = P$.*

Proof. As we previously defined, $P * P$ is the third point of intersection with the cubic of the line through P and P (or the tangent line to the cubic at P). If P is a point of inflection, then the tangent line to the cubic at P intersects the cubic at P with multiplicity three, and so $P * P = P$.

If $P * P = P$, then P is the third point of intersection with the cubic of the tangent line to the cubic at P. Consequently, the tangent line to the cubic at P intersects the cubic at P with multiplicity three. Thus P is an inflection point. □



As a result of this theorem, we obtain several interesting properties, two of which are mentioned below.

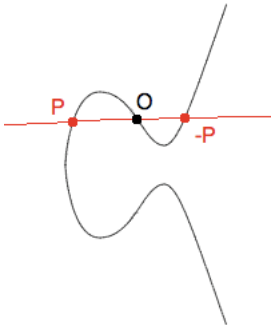
Lemma. *Let C be an irreducible, non-singular cubic curve, and let O be a point on C. Then O is an inflection point if and only if $P + Q + R = O$ whenever P, Q, and R are collinear.*

Proof. If P, Q, and R are collinear, then $P + Q + R = (P + Q) + R$. But $P * Q$ is simply R, since P, Q, and R are collinear. Then $P + Q$ is $R * O$, which we will call S. But then $S * R$ must be O, since by construction R, S, and O are collinear. And so $P + Q + R = S + R = O * O$.

But by the above theorem, $O * O$ is O if and only if O is an inflection point. Thus, we conclude that $P + Q + R = O$ if and only if O is an inflection point. □

Lemma. *$-P = P * O$ if and only if O is an inflection point.*

Proof. As we discussed earlier when we showed how to construct inverses, $-P$ is the point $(O * O) * P$. By the above theorem, $O * O = O$ if and only if O is an inflection point. Thus $-P = O * P$ if and only if O is an inflection point. □



3 The Group of Rational Points

In the previous section, we proved that addition of points as we defined it, is an operation which gives the points of an irreducible, non-singular cubic the structure of an abelian group. In this section, we will prove that addition of points also turns the set of rational points of an irreducible, non-singular cubic an abelian group, provided that we further restrict the set of cubics under consideration to be those which are also rational and which contain at least one rational point.

A *rational* cubic is a cubic that can be written in the form

$$P(x, y) = ax^3 + bx^2y + cxy^2 + dy^3 + ex^2 + fxy + gy^2 + hx + iy + j = 0,$$

where each of the coefficients, $a, b, c, d, e, f, g, h, i,$ and j are rational numbers. An important note is that if we multiply a cubic equation through by a scalar, we obtain the same cubic. Consequently, a cubic with an equation featuring irrational coefficients may actually be a rational cubic. For example, although we may be tempted to say that the cubic defined by

$$\sqrt{2}x^3 - \sqrt{2}y = 0$$

s irrational, it is not, since it is equivalent to the cubic defined by $x^3 - y = 0$, which has rational coefficients.

We must impose the condition that the cubic has at least one rational point, because otherwise the set of rational points on the cubic would be the empty set, and the empty set is not a group because it does not contain an identity element.

Given a rational, irreducible, non-singular cubic curve, C, with at least one rational point, we can choose any rational point to be the identity, O, and give C the structure defined earlier. However, we now only consider the subset of rational points. To show that the rational points are also a group, it remains to show that they are closed under addition and that inverses exist.

3.1 Closure

We will first show that the property of closure is satisfied. We proved in the previous section that if a line intersects the cubic in two points (counting multiplicities), then it intersects it in exactly three points (counting multiplicities). The proof that the rational points are closed under addition is similar. We will show that if a line intersects a rational cubic in two rational points (counting multiplicities), then it intersects the curve at three rational points. If we can show this, then we will know that if P and Q are rational, then $P * Q$ is rational as well. This is sufficient to prove closure, since if $P * Q$ is rational, then the same principle forces $P + Q$ to be rational as well, since O is defined to be a rational point.

Theorem. *If a line intersects a rational, irreducible, non-singular cubic, C, at two rational points, then its third point of intersection with C must also be rational.*

Proof. Given two rational points on a cubic curve, C, we consider the line l joining them. Since it is determined by two rational points, l must be a rational line. Thus, there exists a rational linear transformation which will map l to the x-axis, which is also a rational line. By *rational linear transformation*, we mean one which will map rational points to rational points. We apply this linear transformation to l and C. This will not change the fact that C is rational, irreducible, and non-singular. Moreover, since this linear transformation will map rational points to rational points, it will not change whether the points of intersection of l with C are rational or not rational, and so proving this theorem in general is equivalent to proving it in the case where l is the x-axis.

We let $P(X, Y, Z)$ be the homogeneous polynomial which defines C in $\mathbb{P}^2$. Since C is a cubic, we can write:

$$P(X,Y,Z) = aX^3+bX^2Y+CXY^2+dY^3+eX^2Z+fXYZ+gY^2Z+hXZ^2+iYZ^2+jZ^3.$$

So we also have that:

$$P(X,0,Z) = aX^3 + eX^2Z + hXZ^2 + jZ^3.$$

Since P is a rational cubic, we can write P in a form such that a, e, h, and j are rational.

While it is clear what it means for a point in $\mathbb{R}^2$ to be rational, we must define what we mean by a rational point in projective space. If the point lies in U_Z, then we say that it is rational if it can be written $[x : y : 1]$, where x and y are rational. If the point does not lie in U_Z, then we say that it is rational if it can be written $[x : y : 0]$, where x and y are rational.

Intersections of C and $y = 0$ correspond to solutions of $P(X, 0, Z) = 0$. We know that l intersects C at two rational points, which we denote P_1 and P_2. We also know that l intersects C at a third point, P_3, which we wish to prove is rational. We can write P_1 as $[b_1 : 0 : a_1]$ and P_2 as $[b_2 : 0 : a_2]$, where a_1, a_2, b_1, and b_2 are rational. We also write P_3 as $[b_3 : 0 : a_3]$. Consequently, P can be expressed in the following form:

$$P(X, 0, Z) = \lambda(a_1X - b_1Z)(a_2X - b_2Z)(a_3X - b_3Z)$$

for some non-zero value of λ.

If P_3 does not lie in the affine patch U_Z, then it can be written $[1 : 0 : 0]$, since it lies on the x-axis. Clearly, P_3 will be a rational point in this case. Thus, we must only show that P_3 is rational for the case in which it lies in U_Z. If P_3 lies in U_Z, we can write it as $[b_3 : 0 : 1]$. In order to show that P_3 is rational, we need only show that b_3 is rational.

There are three cases to consider, first, the case in which P_1 does not lie in U_Z, but P_2 does (this also covers the case in which P_1 lies in U_Z but P_2 does not). In this case, we can write P_1 as $[1 : 0 : 0]$ and P_2 as $[b_1 : 0 : 1]$, where b_1 is rational. In this case,

$$\begin{aligned} P(X, 0, Z) &= \lambda(-Z)(X - b_2Z)(X - b_3Z) \\ &= -\lambda X^2Z + \lambda(b_2 + b_3)XZ^2 - \lambda b_2b_3Z^3 \end{aligned}$$

By equating the coefficients in the two formulations which we have for $P(X, 0, Z)$ we find, $-\lambda = e$. Since P is a rational cubic, e is rational, and thus λ must be rational. The coefficient of the XZ^2 term must also be rational. Since λ is non-zero and rational, and since b_2 is rational, b_3 must also be rational, since otherwise, $\lambda(b_2 + b_3)$ would be irrational, which would contradict the fact that it must equal the rational number h. Thus, we have shown that P_3 is rational in this case.

The second case is that in which neither P_1 nor P_2 lie in U_Z. In this case, we can write both P_1 and P_2 as $[1 : 0 : 0]$, and we find:

$$\begin{aligned} P(X, 0, Z) &= \lambda(-Z)(-Z)(X - b_3Z) \\ &= -\lambda X^2Z - \lambda b_3Z^3 \end{aligned}$$

By a similar argument, λ must be rational in this case as well. Also, since λb_3 must be rational, and λ is non-zero and rational, we find that b_3 must be rational. Thus, we have shown that P_3 must be rational in this case as well.

The third case is that in which both P_1 and P_2 lie in U_Z. In this case, we can write P_1 as $[b_1 : 0 : 1]$ and P_2 as $[b_2 : 0 : 1]$, where both b_1 and b_2 are rational. In this case,

$$\begin{aligned} P(X, 0, Z) &= \lambda(X - b_1Z)(X - b_2Z)(X - b_3Z) \\ &= -\lambda X^3 - \lambda(b_1 + b_2 + b_3)X^2Z + \lambda(b_1b_2 + b_1b_3 + b_2b_3)XZ^2 - \lambda b_1b_2b_3Z^3 \end{aligned}$$

By a similar argument, λ must be rational here as well. Moreover, the coefficient to the X^2Z term must be rational. Since λ is rational and non-zero, and b_1 and b_2 are rational, b_3 must also be rational, and we have shown that P_3 must be rational in all cases.

□

3.2 Existence of Inverses

We have shown that the set of rational points on a rational, irreducible, non-singular cubic are closed under addition. In order to show that addition of points gives these rational points the structure of a group, it only remains to show that inverses exist.

Previously, we have shown that given a point P on the cubic, $-P = (O*O)*P$. We must show that if we are given a rational point on the cubic, this process will produce another rational point on the cubic which is its inverse.

In the previous section, we showed that the line between any two rational points also intersects a cubic at a third rational point. Since O is defined to be a rational point, $O * O$ is therefore a rational point. And given a rational point P, $(O * O) * P$ is then a rational point as well. By the same argument used previously, this point will be the inverse of P. Thus inverses exist, and the set of rational points on a rational, irreducible, non-singular cubic with at least one rational point is an abelian group.

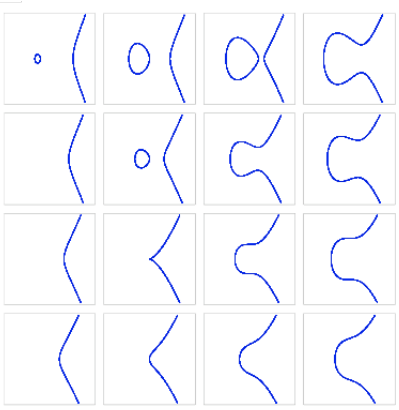
4 Explicit Formulas

4.1 Weierstrass Normal Form

It can be shown that any non-singular, irreducible cubic is birationally equivalent to a cubic in Weierstrass normal form with O as the point at infinity. A cubic in Weierstrass normal form given by:

$$y^2 = x^3 + ax + b.$$

Cubics in Weierstrass normal form are known as elliptic curves. A sampling of elliptic curves are depicted below. By *birationally equivalent*, we mean that there exist mutually inverse rational maps between the given cubic and a cubic in Weierstrass form. By a *rational map*, we mean a morphism defined only on a dense open set. Because of this, the group law on a cubic in Weierstrass normal form is an especially important case, and as such is worthy of further investigation.



4.2 Explicit Formulas

In this section, we will calculate explicit formulas for addition of points on a cubic in Weierstrass normal form where O is the point at infinity. To determine the point at infinity, we will homogenize the above equation of a cubic in Weierstrass normal form by setting $x = X/Z$ and $y = Y/Z$. This gives us the equation:

$$Y^2Z = X^3 + aXZ^2 + bZ^3.$$

To determine the point at infinity, we set $Z = 0$. This gives us $X^3 = 0$, which has the root $X = 0$ with multiplicity three. So the point at infinity is $[0 : 1 : 0]$, and since we have a triple root, this is an inflection point of the cubic.

Returning from projective space, we can interpret this result as follows: the cubic C consists of its graph in the xy plane, plus O, the point at infinity $[0 : 1 : 0]$, which can be interpreted as the point at which all vertical lines intersect. We know any line through two points on the cubic intersects it at a third point. If this line is vertical, then it intersects the cubic at O. If the line is non-vertical, then it must intersect the cubic at another point in the xy plane.

Before we discuss explicit formulas for addition of points, we will first observe that any curve in Weierstrass normal form is symmetric about the x-axis. If a particular (x_0, y_0) satisfies

$$(y_0)^2 = (x_0)^3 + ax_0 + b$$

then $(x_0, -y_0)$ also satisfies it, since

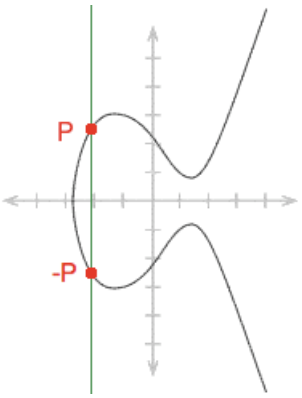
$$(-y_0)^2 = (y_0)^2 = (x_0)^3 + ax_0 + b.$$

Armed with this knowledge, we can examine the group structure of a cubic curve in Weierstrass normal form more closely. Given two points P and Q on the curve, we

find $P + Q$ by first drawing the line joining P and Q. The third point of intersection of that line with the cubic is $P * Q$. We next draw the line joining $P * Q$ and O, which is simply the vertical line through $P * Q$. Its third point of intersection with the cubic must be the reflection of $P * Q$ about the x-axis, since the curve is symmetric about the x-axis.

Moreover, given a point P on the curve, we claim that its inverse, $-P$, is its reflection about the x-axis. To demonstrate that this is true, we find $P * (-P)$, which is the third point of intersection with the curve of the line between P and $-P$. Since $-P$ is the reflection of P about the x-axis, the line between them is vertical, and the third point of intersection is O. Since O is an inflection point, $O * O$ is O, and so $P + (-P)$ is O, and the claim is proved. If P is O, however, it has no reflection about the x-axis. In this case, $-O$ is O, since $O + O = (O * O) * O = O * O = O$, since O is an inflection point. Thus, if $P = (x, y)$ is a point on the cubic, the explicit formula for calculating $-P$ is simply

$$-P = (x, -y).$$



Using this geometric intuition, we can also develop explicit algebraic formulas for the addition of points.

Let us use the following notation: $P = (x_1, y_1)$, $Q = (x_2, y_2)$, $P * Q = (x_3, y_3)$. Then $P + Q = (x_3, -y_3)$. The problem is to compute x_3 and y_3, given points values for the points P and Q.

We first observe that the line joining P and Q has equation $y = mx + n$, where

$$m = \frac{y_2 - y_1}{x_2 - x_1} \text{ and } n = y_1 - mx_1$$

We can find the points of intersection of the line with the cubic by solving the system of equations:

$$\begin{cases} y = mx + n \\ y^2 = x^3 + ax + b \end{cases}$$

Plugging the equation for y from the first equation into the second equation yields:

$$(mx+n)^2 = x^3 + ax + b$$

Shifting everything to one side yields:

$$0 = x^3 - m^2x^2 + (a-2mn)x + (b-n^2)$$

This is a cubic equation in x, and we know that its roots correspond to the x coordinates of the three points of intersection of the line with the cubic, which we know are x_1, x_2, and x_3. Thus,

$$x^3 - m^2x^2 + (a-2mn)x + (b-n^2) = (x-x_1)(x-x_2)(x-x_3)$$

Since the coefficients of the x^2 term must be equal, we obtain

$$m^2 = x_1 + x_2 + x_3$$

Thus, we find that

$x_3 = m^2 - x_1 - x_2$ and $y_3 = mx_3 + n = m(m^2 - x_1 - x_2) + n = m^3 - mx_1 - mx_2 + y_1 - mx_1$

Consequently,

$$P + Q = (m^2 - x_1 - x_2, -m^3 + 2mx_1 + mx_2 - y_1)$$

This equation is well-defined if P and Q are distinct points, however, if we want to calculate $P + P$, we cannot calculate m in the manner described above. In place of m, we need its analogue. As we described before, the line joining P and P is defined to be the tangent line to the curve at C. Thus, in place of m, we want to use the slope of the tangent line to the curve at P. By using implicit differentiation with respect to x, we find:

$$2y\frac{dy}{dx} = 3x^2 + a$$

And so

$$m = \frac{dy}{dx} = \frac{3x^2 + a}{2y}$$

We now have defined explicit formulas for adding points in this special case:

$$P + Q = (m^2 - x_1 - x_2, -m^3 + 2mx_1 + mx_2 - y_1)$$

where

$$m = \begin{cases} \dfrac{y_2 - y_1}{x_2 - x_1} & \text{if } P \neq Q \\ \dfrac{3x_1^2 + a}{2y_1} & \text{if } P = Q \end{cases}$$

4.2.1 Example One

We will illustrate this with an example. We look at following cubic curve, which is in Weierstrass form:

$$y^2 = x^3 - 4x + 1$$

We will add the points $P = (3, 4)$ and $Q = (4, 7)$. We obtain

$$m = 3 \text{ and } n = -5$$

Thus

$$P + Q = (m^2 - x_1 - x_2, -m^3 + 2mx_1 + mx_2 - y_1) = (2, -1)$$

4.2.2 Example Two

If we had wanted to calculate $P + P$, we would have:

$$m = 23/8 \text{ and } n = -37/8$$

Thus,

$$P + P = (m^2 - x_1 - x_2, -m^3 + 2mx_1 + mx_2 - y_1) = (145/64, -967/512)$$

5 Conclusion

5.1 Mordell's Theorem

Mordell's Theorem states that if an irreducible, non-singular cubic plane curve has a rational point, then the group of rational points is finitely generated. The proof is quite lengthy and complex, and it will not be reproduced here. However, we can begin to get an idea of what it means.

Given a rational, non-singular, irreducible cubic with a rational point, we have shown that the set of rational points on the cubic is an abelian group under the operation of addition of points. From a geometric perspective, what it means for this group to be finitely-generated is that there exists a finite number of rational points, P_1, ..., P_n, such that any rational point on the curve can be obtained through the addition of some combination of these points. Since addition of points is associative and commutative, the order in which these points are added does not matter. Consequently, if Q is a rational point on the cubic, we can write

$$Q = \sum_{i=1}^{n} n_i P_i, \text{ where } n_i \in \mathbb{Z}$$

The proof of this theorem uses the fact that any non-singular cubic with a rational point is birationally equivalent to a cubic in Weierstrass normal form, and it also uses the explicit formulas for the addition of points on a cubic in Weierstrass normal form which we developed in the previous section. Then by examining various properties of the height function of rational numbers, Mordell was able to show that the set of rational points on a rational, irreducible, non-singular cubic with at least one rational point is finitely generated.

5.2 Open Questions

While Mordell's Theorem is nice, it does leave us with some important questions which are as yet still unanswered. Given a rational, irreducible, non-singular cubic with a rational point, we know that there exists a finite generating set for the group of rational points. An important open question is, given a cubic, which are the rational points which compose this finite generating set?

In addition, it is not yet known how to determine, given a cubic, the minimum number of points needed in the finite generating set.

A third open problem is how to determine in a finite number of steps whether a given cubic has a rational point at all, and so given an arbitrary cubic, it may not even be possible to know whether Mordell's Theorem applies to it.

References

[1] Bix, Robert, *Cubics and Conics.* Springer-Verlag, New York, 1998.

[2] Husemoller, Dale, *Elliptic Curves.* Springer-Verlag, New York, 1987.

[3] Smith, Karen, *An Invitation to Algebraic Geometry.* Springer-Verlag, New York, 2000.

[4] Silverman, Joseph; Tate, Johan *Rational Points on Elliptic Curves.* Springer-Verlag, New York, 1992.