MATH 115, SUMMER 2012
LECTURE 30

JAMES MCIVOR

1. Properties of the Addition Law

Here we check that the addition of points defined above is associative and has inverses, and that $P+O = O+P = P$ for all $P \in E$ (so that E is actually a group), and also that it is commutative, i.e., $P+Q = Q+P$ (which means E is an abelian group).

Proof. (that O is the additive identity) What we have to show is that for any $P \in E$, $P + O = P$, which amounts to saying $O(OP) = P$. But in constructin $O(OP)$, we take the line from O to P; it meets E in 3 points: O, P, and some other point OP, and now we take the line from OP to O, which is the very same line from before. Therefore its third intersection with E is again P.

□

Proof. (of the existence of inverses) Fix $P \in E$. Let L be the tangent line to E at O. The line L meets E in other point[1]. Call this point Q[2]. The line from P to Q meets E in one other point. This is $-P$, the inverse of P.

Why? Because if you trace back trough the construction, you find $P+(-P) = O$.

- draw picture. Note - best to choose O strategically, so L doesn't meet E at infinity.

□

Proof. (of commutativity) This is the easiest of the three properties we have to check. It's just because when we talk about "the line between P and Q", this is symmetric in P and Q. So $PQ = QP$, and hence $P + Q = O(PQ) = O(QP) = Q + P$.

□

Proof. (of associativity - this is the harder part)

We pick three points $P, Q, R \in E$, and we must show

$$P + (Q + R) = (P + Q) + R$$

We make two simplifying assumptions. The first is that all nine points are not at infinity, so we may regard E as an affine (rather than projective) plane curve. If not, we can make a change of variables in the equation defining E to arrange that they are, so there is no loss of generality in this assumption.

Secondly, and this is a harder assumption to justify, let us assume that the three points are distinct. Having shown it in this case, if they're not we can deform

[1]because L is tangent to E, its intersection with E has "multiplicity two", so although most lines meet E in three points, L meets E in only one other point besides O.

[2]It could happen that in fact $Q = O$, in which case O is called an **inflection point** of E.

them a little bit so that they are distinct. Then the equality above holds, and since both expressions are continuous functions $E \times E \times E \to E$, the equality holds even as we deform them back to the original (non-distinct) points. The details of this continuity argument are omitted - see your book for more elaboration.

Now we take up the case when P, Q, R are distinct. We define nine points on E as follows:

$$\begin{aligned} P_1 &= Q \\ P_2 &= QR \\ P_3 &= R \\ P_4 &= AB \\ P_5 &= O \\ P_6 &= O(PQ) \\ P_7 &= P \\ P_8 &= O(QR) \\ P_9 &= (O(PQ))R \end{aligned}$$

If we can show that P_7, P_8, P_9 are collinear (all lie on one line), then the third intersection of the line containing P and $O(QR)$ with E must be $(O(PQ))R$, which means

$$P(O(QR)) = (O(PQ))R,$$

and therefore taking the line through these (identical) points and O gives the equality

$$O(P(O(QR))) = O((O(PQ))R),$$

which, if we stare at the definition of the addition, gives

$$P + (Q + R) = (P + Q) + R,$$

which is what we want.

So why do these three points lie on the same line?

First of all note that P_1, P_4, P_7 are collinear, as are P_2, P_5, P_8 and also P_3, P_6, P_9; this is all just from the definitions of the nine points.[3]

So we have these three lines, each of which passes through three of our nine points. Each line is given by a linear polynomial in x, y. Take the product of these three linear polynomials and you get a cubic polynomial, call it $g(x, y)$. This g defines a curve in $\mathbb{R}^2$, and all nine of our points lie on g, by the construction of g. Also all nine of our points live on E, so we have nine points which all lie on two cubic curves. Moreover, we have three of these points, say P_1, P_2, P_3 all on the same line.

Now it is a lemma that when nine points are common to two cubics, and three of them lie on a line, then the other six lie on a conic (see below for the proof of this lemma).

So let $q(x, y)$ be the quadratic polynomial defining this conic. Then $P_4, \ldots, P_9$ all lie on the conic curve C_q. But also P_4, P_5, P_6 are collinear. This means that $q(x, y)$ factors into two linear factors, so the conic C_q is the union of two lines (see lemma below).

[3]Also you may observe that P_1, P_2, P_3 and also P_4, P_5, P_6 are collinear - we will use this shortly, but we don't need it just yet.

Now we're almost done. We know that P_7, P_8, P_9 lie on the conic C_q, which is a union of two lines. One of the lines, call it L_1, contains P_4, P_5, P_6, and the other is some line L_2 we don't know anything about. If one of P_7, P_8, P_9 lies on the first line L_1, then L_1 contains at least four points of E. By the lemma below, then, L_1 is a component of E, which is therefore reducible. But this contradicts the irreducibility of E (all elliptic curves are irreducible by definition). So none of P_7, P_8, P_9 lie on L_1, hence they all lie on L_2, hence they are collinear and we're done.

□

Lemma 1 (The "Too-Many-Points" Lemma)**.** *Suppose C is a degree d (affine or projective) plane curve, and $P_1, \ldots, P_k$ are points of C that are collinear, say all lying on some line L. If $k > d$, then C is reducible, and L is a component of C.*

Roughly speaking, it says that if a degree d curve intersects a line in too many (more than d) points, then the line must be a component of C. This means in fact that their intersection contains infinitely many points - namely every point of L!

Proof. Let C be given by the equation $f(x, y) = 0$, where f is a polynomial of two variables x, y of degree d. First assume the line L is not vertical (which corresponds to slope $= \infty$). Then L has an equation of the form $y = mx + b$ for some $m, b \in \mathbb{R}$. Points (x, y) satisfying this equation and the equation for C correspond to values of x satisfying

$$f(x, mx + b) = 0,$$

which is just a polynomial in x of degree d. The fundamental theorem of algebra[4] says that this equation 1) is identically zero, or 2) has exactly[5] d solutions if x is allowed to be complex. In case 2), some or all of these complex numbers may actually be real, but in any case, if we look at only the real number solutions, there are at most d of them. Thus since there are $k > d$ points $P_1, \ldots P_k$ satisfying this equation[6], then we must be in case 1). Then every point of the line is also a point of C, which means exactly that L is a component of C. Algebraically, it means that f factors into a linear polynomial times some degree $d - 1$ polynomial (which may or may not be irreducible).

□

2. Rational Points on an Elliptic Cruve

So far we've been working with all real points on E. But in number theory we only really care about the raional points, since these relate back to solving Diophantine equations, which is what we wanted to do in the first place.

Recall that a subgroup of a group G is a subset that is closed uner addition and taking inverses. The best thing about the group structure we just deifned is that the rational points form a subgroup, at least if we choose the identity element correctly.

---

[4]If you haven't seen this theorem, don't worry.

[5]Some of the solutions may be repeated, for example $x^2 = 0$ has the "double root" $x = 0$.

[6]Really it is just the x-coordinates of the points, and not the points themselves, that satisfy the equation.

Proposition 1. *If E is an elliptic curve, defined by a polynomial with* rational *coefficients, and the point O is a rational point on E, then the group structure defined above, using O as identity element, is such that the set of rational points, denoted $E(\mathbb{Q})$, forms a subgroup of E. Conversely, if the rational points form a subgroup, then the identity element we chose must also be a rational point.*

Proof. Clear - lines through rational points, give other rational points. Similar to arguments in Pythagorean triples lecture. □

3. What kinds of groups do we get?

Now we are interested in the subgroups $E(\mathbb{Q})$ of rational points. Some questions emerge:

(1) Our group depends on the choice of identity element O. How does the group change if we choose different points for O?
(2) If we look at all the possible elliptic curves E, what groups $E(\mathbb{Q})$ arise?

The first question is easier, the answer to the second is a very famous theorem of Mordell from 1922.

Let us look at the first question. Say we first choose a point O_1 as our identity - it gives us a way of adding points, which we write $P + Q$.

Now we choose a different point O_2, and instead use that as our identity element. Then we get a different way of adding points, let's write it as $P \oplus Q$ to distinguish it from the first addition law. How are $+$ and $\oplus$ related?

Proposition 2. *The group E with addition law $+$ is isomorphic to the group E with addition law $\oplus$. The isomorphism is given by $P \mapsto P + O_2$.*

- see book for proof.
- the same typew of result holds if we only look at subgroups of rational points.

Now we consider the second question. We know that E has the structure of an abelian group. What kinds of abelian groups are there? The most common ones are $\mathbb{Z}$ and $\mathbb{Z}/m$. But we can also combine these by the operation of direct sum, as follows:

$$\mathbb{Z} \oplus \mathbb{Z} = \{(a, b) \,|\, a, b \in \mathbb{Z}\}$$
$$\mathbb{Z} \oplus \mathbb{Z}/m = \{(a, b) \,|\, a \in \mathbb{Z}, b \in \mathbb{Z}/m\}$$
$$\mathbb{Z} \oplus \mathbb{Z}/m \oplus \mathbb{Z}/n = \{(a, b, c) \,|\, a \in \mathbb{Z}, b \in \mathbb{Z}/m, c \in \mathbb{Z}/n\}$$

in each case, the addition is defined "component by component", e.g., $(a, b) + (c, d) = (a + b, c + d)$.

But there are other, weirder, abelian groups, for example $\mathbb{Q}$ itself, if we forget about the multiplication, or $\mathbb{Q}/\mathbb{Z}$, which is like $\mathbb{Q}$ except with the extra rule[7] $1 = 0$, so for example $1/2 + 3/4 = 1/4$ in $\mathbb{Q}/\mathbb{Z}$.

An abelian group G is called **finitely generated** if there is a finite set $g_1, \ldots, g_r$ of generators, which means that any element of G can be obtained by adding and subtracting various copies of the g_i. Now it's a famous theorem of abstract algebra[8] that every finitely generated abelian group can be written in the form

$$\mathbb{Z}^r \oplus \mathbb{Z}/m_1 \oplus \ldots \oplus \mathbb{Z}/m_k$$

[7]This is just the factor group, if you've seen it before.
[8]The "Structure Theorem" for finitely generated abelian groups.

for some nonnegative integers $r, m_1, \ldots, m_k$. Thus it says that finitely generated abelian groups are all built out of finitely many copies of $\mathbb{Z}$ and $\mathbb{Z}/m$. We actually pretty much proved this when we proved the Smith Normal Form Theorem in lecture 23.

Thus if we know that a certain group is finitely generated, we understand that group very well. Mordell figured out that $E(\mathbb{Q})$ is a fintely generated abelian group, for any elliptic curve E:

Theorem 1 (Mordell-Weil Theorem)**.** *If E is an elliptic curve defined by a polynomial with rational coefficients, with rational identity element O, then the subgroup $E(\mathbb{Q})$ of rational points is a finitely generated abelian group.*

- the proof is too hard for us. But it's a very powerful theorem, because it says that all the rational points on an elliptic curve (if there are any) can be obtained by drawing lines with rational slopes, starting from finitely many fixed rational points (the generators).

- even more surprising, in the late 70s, Barry Mazur of Harvard showed that if $E(\mathbb{Q})$ is written as

$$\mathbb{Z}^r \oplus \mathbb{Z}/m_1 \oplus \ldots \oplus \mathbb{Z}/m_k,$$

which can be done by Mordell's Thm and the structure Thm for Fin Gen'd Ab Groups, then there are only two basic possibilities for the m_i: either there is only one $\mathbb{Z}/m_i$, i.e., $E(\mathbb{Q}) \cong \mathbb{Z}^r \oplus \mathbb{Z}/m$, and in this case m must be between 1 and 10, or else it's 12. Otherwise, $E(\mathbb{Q}) \cong \mathbb{Z}^r \oplus \mathbb{Z}/m \oplus \mathbb{Z}/2$, where m can only be $2, 4, 6, 8$. This means there are really not so many possibilities for the rational points on elliptic curves. It constitutes a huge breakthrough in understanding cubic diophantine equations.

4. Connections to Other Branches of Mathematics

-time permitting, mention elliptic integrals, arc length of ellipse, torus, etc.