LECTURE 22, WEDNESDAY 27.04.04

FRANZ LEMMERMEYER

Last time, we defined the height of a rational number $x = \frac{p}{q}$ in lowest terms by $H(x) = \max\{|p|, |q|\}$ and proved

Proposition 1. *Let $f, g \in \mathbb{Z}[X]$ be coprime, and put $n = \max\{\deg f, \deg g\}$. Then there exist constants $C_1, C_2 > 0$ such that for all $x \in \mathbb{Q}$ we have*

$$C_1 H(x)^n \le H\Big(\frac{f(x)}{g(x)}\Big) \le C_2 H(x)^n.$$

Now we will apply this result to rational points on elliptic curves.

1. The Naive Height

Now assume that $P = (x, y)$ is a rational point on a curve. Then we put $H(P) = H(x)$; for elliptic curves, we think of $\mathcal{O}$ as $(\frac{1}{0}, \frac{1}{0})$ and put $H(\mathcal{O}) = 0$. The function $H : E(\mathbb{Q}) \longrightarrow \mathbb{N}$ is called the naive height on E. It will be used to define the "canonical" (or Néron-Tate) height below.

Lemma 2. *Let* E *be an elliptic curve defined over $\mathbb{Q}$. Then for all $\kappa > 0$, the set $\{P \in \mathrm{E}(\mathbb{Q}) : H(P) < \kappa\}$ of rational points with bounded height is finite.*

Proof. Clearly the set of all $x \in \mathbb{Q}$ with $H(x) < e^\kappa$ is finite, hence the same is true for the set of all x-coordinates of points $P \in \mathrm{E}(\mathbb{Q})$ with $H(P) < \kappa$. To each such x-coordinate, however, there are at most two values of y with $(x, y) \in \mathrm{E}(\mathbb{Q})$. □

The height function on $E(\mathbb{Q})$ satisfies certain relations that will be needed in our proof of the Mordell-Weil theorem. We'll start with

Lemma 3. *Let* E : $y^2 = x^3 + ax + b$ *an elliptic curve defined over $\mathbb{Q}$. Then there exist constants $C_1, C_2 > 0$ such that $C_1 H(P)^4 \le H(2P) \le C_2 H(P)^4$ for all $P \in \mathrm{E}(\mathbb{Q})$.*

Proof. Let us write $P = (x, y)$ and $P^* = 2P = (x^*, y^*)$; here we assume that $P^* \ne \mathcal{O}$ (when proving statements like the one above, we may always exclude finitely many points, because these can be taken care of afterwards by simply changing the constants C_1, C_2). By the duplication formula on elliptic curves $y^2 = f(x) = x^3 + ax + b$ we have

$$x^* = \frac{f'(x)^2 - (8x + 4a)f(x)}{4f(x)} = \frac{x^4 - 2ax^2 - 8bx + a^2}{4(x^3 + ax + b)}.$$

The first fraction shows that any common divisor of numerator and denominator divides f and f', which implies that f and f' have a common root in $\mathbb{Z}[X]$; but then f has a double root, hence E is singular and not an elliptic curve. Thus we can apply Proposition 1, and this proves the claim. □

In the proof of Proposition 1 we have used relations between the polynomials; the result above can be proved directly by writing down these relations. Consider the polynomials $f(x) = x^4 - 2ax^2 - 8bx + a^2$ and $g(x) = 4(x^3 + ax + b)$; typing

```
f=x^4-2*a*x^2-8*b*x+a^2: g=4*(x^3+a*x+ b):polresultant(f,g)
```

into `pari` shows that the resultant $R_{f,g}$ is

$$R_{f,g} = 2^8(4a^3 + 27b^2)^2.$$

Thus there exist polynomials u, v with $fu + gv = R_{f,g}$. We get a better relation by using the Euclidean algorithm and Bezout: the command `bezout(f,g)` gives polynomials

$$u_1 = \frac{1}{D}(3x^2 + 4a), \quad v_1 = \frac{1}{4D}(-3x^3 + 5ax + 27b),$$

where $D = 4a^3 + 27b^2$ and where $fu_1 + gv_1 = 1$. Multiplying through by $4D$ shows that

$$4(x^4-2ax^2-8bx+a^2)(3x^2+4a)+4(x^3+ax+b)(-3x^3+5ax+27b) = 4(4a^3+27b^2).$$

Plugging in $x = p/q$ and clearing denominators now gives

$$\begin{aligned}4(p^4 - 2ap^2q^2 - 8bpq^3 + a^2q^4)(3p^2 + 4aq^2)+\\ 4(p^3 + apq^2 + bq^3)(-3p^3 + 5apq^2 + 27bq^3) = 4Dq^6.\end{aligned}$$

Multiplying through by q then gives the desired relation between $F(p, q)$ and $G(p, q)$.

The command `bezout(polrecip(f),polrecip(g))` now gives a similar relation between $f_2(x) = a^2x^4 - 8bx^3 - 2ax^2 + 1$ and $g_2(x) = 4(bx^3 + ax^2 + 1)$, namely $f_2u_2 + g_2u_2 = 4D$ with

$$\begin{aligned}u_2(x) &= -4(2ab^2x^2 + a^2bx - a^3 - 3b^2),\\ v_2(x) &= 2a^3bx^3 - (a^4 + 16ab^2)x^2 + a^2bx + 3a^3 + 24b^2.\end{aligned}$$

Plugging in $x = \frac{q}{p}$ and clearing denominators gives

$$\begin{aligned}&4(bq^3 + apq^2 + p^3)(2a^3bq^3 - (a^4 + 16ab^2)pq^2 + a^2bp^2q + 3(a^3 + 8b^2)p^3)\\ &-4(a^2q^4 - 8bpq^3 - 2ap^2q^2 + p^4)(2ab^2p^2q^2 + a^2bp^3q - (a^3 + 3b^2)p^2) \;=\; 4Dp^6.\end{aligned}$$

Multiplying through by p now gives the second relation between $F(p, q)$ and $G(p, q)$.

2. The Canonical Height

We have seen above that the "naive" height has the property that $H(2P) \sim H(P)^4$. Taking logs and setting $h_0(P) = \log H(P)$, we find that $h_0 : E(\mathbb{Q}) \longrightarrow \mathbb{R}_{\ge 0}$ is a function with the property $h_0(2P) \sim 4h_0(P)$. The canonical height h is a function $h : E(\mathbb{Q}) \longrightarrow \mathbb{R}_{\ge 0}$ differing only slightly from h_0 and with the property $h(2P) = 4h(P)$. Since the properties of the canonical height are much nicer than those of the naive height, using the canonical height will in general allow us to give cleaner proofs. The main reason for introducing it, however, is that it is the canonical and not the naive height which is involved in the definition of the regulator of an elliptic curve: this is a matrix whose entries are formed using the canonical heights of a basis of the free part of $E(\mathbb{Q})$, and it occurs in the Birch–Swinnerton-Dyer Conjecture.

Proposition 4. *There is a unique function $h : E(\mathbb{Q}) \longrightarrow \mathbb{R}_{\geq 0}$ with the properties*

(1) *$|h(P) - h_0(P)| < C$ for some constant $C > 0$ and all $P \in E(\mathbb{Q})$,*
(2) *$h(2P) = 4h(P)$ for all $P \in E(\mathbb{Q})$;*

it is given by

$$h(P) = \lim_{n\to\infty} \frac{h_0(2^n P)}{4^n}. \tag{1}$$

Proof. Assume that h is a function with the properties (1) and (2). Then

$$|4^n h(P) - h_0(2^n P)| = |h(2^n P) - h_0(2^n P)| < C,$$

hence dividing through by 4^n and letting $n \to \infty$ we find that h must be given by (1), hence is unique.

To show that (1) defines a function h, we need to prove that the sequence $a_n = 4^{-n} h_0(2^n P)$ is Cauchy. To this end we first observe that taking logs of the inequalities $c_1 H(P)^4 \leq H(2P) \leq c_2 H(P)^4$ gives

$$|h_0(2P) - 4h_0(P)| \leq c$$

for some constant $c > 0$ not depending on P. Now let $n > m \geq 0$ be integers; then

$$\begin{aligned}
&|4^{-n} h_0(2^n P) - 4^{-m} h_0(2^m P)| \\
&\qquad = \left| \sum_{i=m}^{n-1} (4^{-i-1} h_0(2^{i+1} P) - 4^{-i} h_0(2^i P)) \right| \\
&\qquad \leq \sum_{i=m}^{n-1} 4^{-i-1} |h_0(2^{i+1} P - 4h_0(2^i P)| \\
&\qquad \leq \sum_{i=m}^{n-1} 4^{-i-1} c \leq \frac{1}{3} 4^{-m} c.
\end{aligned}$$

But if m (and therefore n) go to infinity, then the right hand side converges to 0, which shows that (a_i) is Cauchy. Taking $m = 0$ and letting $n \to \infty$, on the other hand, shows that

$$|h(P) - h_0(P)| \leq \frac{c}{3},$$

which is property (1).

Property (2) is immediate from (1):

$$h(2P) = \lim_{n\to\infty} \frac{h_0(2^{n+1} P)}{4^n} = 4 \lim_{n\to\infty} \frac{h_0(2^{n+1} P)}{4^{n+1}} = 4h(P).$$

□

Here are some more properties of the canonical height:

Theorem 5. *The canonical height h has the following properties:*

(1) *For any $c' > 0$, there are only finitely many $P \in E(\mathbb{Q})$ with $h(P) < c'$.*
(2) *For all $m \in \mathbb{Z}$ and all $P \in E(\mathbb{Q})$ we have $h(mP) = m^2 h(P)$.*
(3) *We have $h(P) \geq 0$ for all $P \in E(\mathbb{Q})$, with equality if and only if P is a torsion point.*
(4) *The parallelogram law: for all $P, Q \in E(\mathbb{Q})$ we have*

$$h(P + Q) + h(P - Q) = 2h(P) + 2h(Q).$$

The only property of the canonical height that we will need for the proof of the Mordell-Weil Theorem is the parallelogram law, which will follow from

Lemma 6. *There is a constant $c > 0$ such that*

$$H(P+Q)H(P-Q) \le cH(P)^2H(Q)^2$$

for all $P, Q \in E(\mathbb{Q})$.

Its proof still requires some work, as we will see.

Proof of Theorem 5. Assume that $h(P) < c$. Then $|h(P) - h_0(P)| < c$ for all $P \in E(\mathbb{Q})$, hence $h_0(P) < c + c'$, and now the claim follows from Lemma 2.

For the proof of the parallelogram law (4), we observe that Lemma 6 implies

$$4^{-n}(h_0(2^nP + 2^nQ) + h_0(2^nP - 2^nQ) - 2h_0(2^nP) - 2h_0(2^nQ)) \le 4^{-n}c,$$

and letting $n \to \infty$ yields

$$h(P+Q) + h(P-Q) \le 2h(P) + 2h(Q).$$

Applying this inequality to $P' = P + Q$ and $Q' = P - Q$ shows

$$h(2P) + h(2Q) \le 2h(P+Q) + 2h(P-Q).$$

Using $h(2P) = 4h(P)$ and $h(2Q) = 4h(Q)$ then implies

$$2h(P) + 2h(Q) \le h(P+Q) + h(P-Q),$$

and the parallelogram law follows.

Now let us prove (2); since $h(-P) = h(P)$, we may assume that $m \ge 0$. The claim is trivial for $m = 0, 1$, and we already have proved the case $m = 2$ (it also follows from (5) by setting $P = Q$). Assume the claim holds for $m-1$ and m; then

$$\begin{aligned} h((m+1)P) &= -h((m-1)P) + 2h(mP) + 2h(P) \\ &= (-(m-1)^2 + 2m^2 + 2)h(P) = (m+1)^2h(P). \end{aligned}$$

It remains to prove that points of canonical height 0 are exactly the torsion points. If $mP = \mathcal{O}$, then $0 = h(\mathcal{O}) = h(mP) = m^2h(P)$, hence $h(P) = 0$: torsion points have height 0. Conversely, assume that P is a rational point of height 0. Then $h(mP) = m^2h(P) = 0$ for all $m \ge 0$, and since there are only finitely many points with height < 1, say, we must have $mP = nP$ for some integers $m < n$; but then $(n-m)P = \mathcal{O}$, and P is torsion. □

3. Proof of Mordell-Weil

We now can prove the Theorem of Mordell Weil for elliptic curves defined over $\mathbb{Q}$ with three rational points of order 2:

Theorem 7. *Let $E : y^2 = (x - e_1)(x - e_2)(x - e_3)$ be an elliptic curve defined over $\mathbb{Q}$ with $e_1, e_2, e_3 \in \mathbb{Z}$. Then $E(\mathbb{Q})$ is finitely generated, and thus has the form $E(\mathbb{Q}) = E(\mathbb{Q})_{\text{tors}} \oplus \mathbb{Z}^r$ for some integer $r \ge 0$ called the Mordell-Weil rank of E.*

Proof. By the weak theorem of Mordell-Weil we know that $E(\mathbb{Q})/2E(\mathbb{Q})$ is finite; let $R_1, \dots, R_n$ be representatives of the finitely many cosets. Put

$$c = \max\{h(R_1), \dots, h(R_n)\}$$

and let $Q_1, \dots, Q_m$ denote the finitely many points with height $h(Q_i) \le c$. Let G be the subgroup of $E(\mathbb{Q})$ generated by $R_1, \dots, R_n, Q_1, \dots, Q_m$. We claim that $G = E(\mathbb{Q})$.

If not, then $E(\mathbb{Q}) \setminus G$ is not empty, and we choose a rational point $P \in E(\mathbb{Q}) \setminus G$ with minimal height. Since P is contained in one the cosets $R_i + 2E(\mathbb{Q})$, we can write $P - R_i = 2P_1$ for some $P_1 \in E(\mathbb{Q})$. Now

$$\begin{aligned} 4h(P_1) &= h(2P_1) = h(P - R_i) \\ &= 2h(P) + 2h(R_i) - h(P + R_i) \\ &\leq 2h(P) + 2c < 2h(P) + 2h(P) = 4h(P), \end{aligned}$$

where we have used that $c < h(P)$ since P is not among the Q_i. But $h(P_1) < h(P)$ and the minimality of $h(P)$ shows that $P_1 \in G$, hence $P = R_i + 2P_1 \in G$ as well. This contradiction finishes the proof. □

Note that the proof is constructive: given representatives R_i of the cosets of $E(\mathbb{Q})/2E(\mathbb{Q})$ (these can be read off the image of $\alpha : E(\mathbb{Q}) \longrightarrow (\mathbb{Q}^\times/\mathbb{Q}^{\times 2})^3$, if it can be determined), compute the constant c, find all rational points Q_i on E with height $< c$ by an exhaustive search, and then the R_i and Q_i generate $E(\mathbb{Q})$. Finding a basis from a set of generators essentially is linear algebra.

The problems, however, are these:

- Determining the image of α is difficult, in particular if $\mathbf{Ш}(E)[2]$ is not trivial.
- Finding all points of height $< c$ will in general be impossible due to the size of c.

4. Proof of the Parallelogram Law

In this section we will give the proof of the parallelogram law in the form of Lemma 6.

Recall that, for rational numbers x, y, we have seen that $H(x+y) \leq 2H(x)H(y)$; clearly, we cannot expect a nontrivial lower bound since the choice $y = -x$ gives $H(x - y) = 1$. Similarly, we have $H(xy) \leq H(x)H(y)$, and again there is no nontrivial lower bound since we are free to pick $y = \frac{1}{x}$. There is, however, a way to get a lower bound if we look at $H(x + y)$ and $H(xy)$ simultaneously.

In fact, let us consider the map

$$* : \mathbb{P}^1\mathbb{Q} \times \mathbb{P}^1\mathbb{Q} \longrightarrow \mathbb{P}^2\mathbb{Q} : [a : b] * [c : d] = [ac : ad + bc : bd]$$

from the direct product of two projective lines to the projective plane. Define the height of an element $x = [x_0 : x_1 : \cdots : x_n] \in \mathbb{P}^n\mathbb{Q}$ with $x_i \in \mathbb{Z}$ and $\gcd(x_0, \ldots, x_n) = 1$ to be $H(x) = \max\{|x_0|, \ldots, |x_n|\}$. Then

Lemma 8. *Let $u = [a : b]$ and $v = [c : d]$ be points on the rational projective line. Then*

$$\frac{1}{2}H(u)H(v) \leq H(u * v) \leq 2H(u)H(v).$$

Proof. We may assume that $a, b, c, d \in \mathbb{Z}$ and that $\gcd(a, b) = \gcd(c, d) = 1$. The upper bound is clear. Now

$$\begin{aligned} \gcd(ac, ad + bc, b) &= \gcd(ac, ad, b) \mid \gcd(a, b)\gcd(c, d) = 1, \\ \gcd(ac, ad + bc, d) &= \gcd(ac, bc, d) \mid \gcd(c, d)\gcd(a, b) = 1. \end{aligned}$$

Thus $\gcd(ac, ad + bc, bd) = 1$, hence $H(u * v) = \max\{|ac|, |ad + bc|, |bd|\}$, and we have to prove that

$$\max\{|a|, |b|\} \cdot \max\{|c|, |d|\} \leq 2\max\{|ac|, |ad + bc|, |bd|\},$$

which is equivalent to the inequalities

$$\begin{aligned}
|ac| &\le 2\max\{|ac|,|ad+bc|,|bd|\},\\
|ad| &\le 2\max\{|ac|,|ad+bc|,|bd|\},\\
|bc| &\le 2\max\{|ac|,|ad+bc|,|bd|\},\\
|bd| &\le 2\max\{|ac|,|ad+bc|,|bd|\}.
\end{aligned}$$

The first and the last are trivial, so it remains to show the second and third, and by symmetry it suffices to prove the second. This inequality holds trivially if $|d| \le 2|c|$ or $|a| \le 2|b|$, so assume that $|a| > 2|b|$ and $|d| > 2|c|$; this implies $|ad| > 4|bc|$. But now $|ad+bc| \ge |ad| - |bc| > \frac{3}{4}|ad|$, hence $2|ad+bc| > \frac{3}{2}|ad| > |ad|$. □

And now, finally, the parallelogram law. Write $P = (x_1, y_1)$, $Q = (x_2, y_2)$, $P + Q = (x_3, y_3)$ and $P - Q = (x_4, y_4)$. Then a mindless calculation using the addition formulas on $E : y^2 = x^3 + ax + b$ (note that we may assume $Q \ne \mathcal{O}$ since the claim in this case is trivial) shows that

$$\begin{aligned}
x_3 + x_4 &= 2\frac{(x_1x_2 + a)(x_1 + x_2) + 2b}{(x_2 - x_1)^2},\\
x_3x_4 &= \frac{(x_1x_2 - a)^2 - 4b(x_1 + x_2)}{(x_2 - x_1)^2}
\end{aligned}$$

Thus $[1 : x_3 + x_4 : x_3x_4] = [w_0 : w_1 : w_2] = w$ with

$$\begin{aligned}
w_0 &= 2(x_1x_2 + a)(x_1 + x_2) + 4b,\\
w_1 &= (x_1x_2 - a)^2 - 4b(x_1 + x_2)\\
w_2 &= (x_2 - x_1)^2.
\end{aligned}$$

Writing $x_1 = p/q$ and $x_2 = p'/q'$ as fractions in lowest terms it is easily seen that $H(w) \le cH(x_1)^2H(x_2^2)$. On the other hand, we have $u * v = w$ for $u = [1 : x_3]$, $v = [1 : x_4]$, hence $H(w) \ge \frac{1}{2}H(u)H(v)$. But $H(u) = H(x_3) = H(P + Q)$ and $H(v) = H(x_4) = H(P - Q)$. Putting $c' = 2c$, this gives

$$H(P + Q)H(P - Q) \le c' \cdot H(P)^2H(Q)^2.$$