Abstract. These are the rough ongoing lecture notes for the course 18.782 (Introduction to Arithmetic Geometry) taught at MIT in Fall 2019. The notes will be updated each weekend, following the lectures during the week. Be sure to clear your cache before reloading in order to get the most current version of the notes. If you find any typos/mistakes in the notes, please let me know!

1. Lecture 1

Note: This lecture is intended to be informal and motivational. It discusses several outside concepts (such as Riemann surfaces) which are not strictly needed in the remainder of the course. The lecture also briskly introduces several algebro-geometric concepts in an informal manner, to give intuition; all of these will be reintroduced more slowly and rigorously later in the semester.

1.1. Historical context. A Diophantine equation is a polynomial equation with integral (or rational) coefficients:

$$f(x_1,\ldots,x_n) = 0, \quad f \in \mathbb{Z}[x_1,\ldots,x_n].$$

Diophantus of Alexandria (c. 200-300 AD) wrote a series of texts called Arithmetica examining the problem of solving such equations in the integers or the rationals. Given $f$ as above and a commutative ring $R$, let us denote

$$V_f(R) = \{ P \in R^n : f(P) = 0 \}.$$

Diophantine analysis is the study of the relationship between Diophantine equations $f = 0$ and the sets $V_f(\mathbb{Z})$ or $V_f(\mathbb{Q})$. Below are some examples of Diophantine equations, which were mentioned by Fermat in his 1650 letter to Carcavi summarizing Fermat’s earlier works:

- For $p$ odd prime, $x^2 + y^2 = p$ is solvable in integers $(x,y) \in \mathbb{Z}^2$ if and only if $p \equiv 1 \mod 4$ (Fermat’s two-square theorem, proved by Euler around 1750).
- For every $k \in \mathbb{Z}_{\geq 1}$, the equation $x^2 + y^2 + z^2 + w^2 = k$ is solvable in integers (Lagrange’s four-square theorem, proved by Lagrange in 1770).
- For every positive nonsquare integer $N$, the equation $x^2 - Ny^2 = 1$ is solvable in integers (Pell’s equation; its solvability goes back to work of Brahmagupta c. 600-700 AD).
- $x^3 + y^3 = z^3$ has no solution in positive integers (the $n = 3$ case of the so-called Fermat’s Last Theorem; this case was proved by Euler in 1770).
- The only integer solutions of $y^2 = x^3 - 2$ are $(x,y) = (\pm 3,5)$.

(Fermat rarely gave proofs of his claims, and sometimes made incorrect ones.) The study of these and other Diophantine equations (especially involving binary quadratic forms) drove much of the research in number theory in the period following Fermat, and in particular formed the foundations for the development of algebraic number theory.

On the other hand, given an equation $f = 0$ as above, we can view the aggregate of its complex solutions $V_f(\mathbb{C})$ as a geometric object, namely a (complex) algebraic variety. Algebraic geometry is the field devoted to studying the geometric properties of algebraic varieties as well as their interactions with algebraic properties of the defining equations.

Date: December 1, 2019.
Arithmetic geometry is a relatively modern field of mathematics one of whose aims is to build a bridge between the geometric objects $V_f(\mathbb{C})$ and the arithmetic objects $V_f(\mathbb{Z})$ (or $V_f(\mathbb{Q})$). The fact that there should be a connection between the two aspects of the equation $f = 0$ is surprising and adds to the allure of the subject.

Below, we will briefly discuss some of the achievements in the twentieth century, to illustrate the role of geometry in the study of Diophantine equations. Suppose we are given a nonconstant polynomial $f = f(x,y) \in \mathbb{Q}[x,y]$. Since the set $V_f(\mathbb{R})$ of real points on the real affine plane $\mathbb{A}^2(\mathbb{R}) = \mathbb{R}^2$ typically, forms a one-dimensional object, we refer to $V_f$ as an affine plane curve defined over $\mathbb{Q}$. Let us make the two assumptions about $f$:

- $V_f$ is geometrically irreducible, i.e. $f$ is irreducible as a polynomial in $\mathbb{C}[x,y]$, and
- $V_f$ is nonsingular, i.e. $\nabla f(P) = \left(\frac{\partial f}{\partial x}(P), \frac{\partial f}{\partial y}(P)\right) \neq 0$ for every $P \in V_f(\mathbb{C})$.

Under these hypotheses, the set of complex points $V_f(\mathbb{C})$ admits the structure of a connected Riemann surface (i.e. complex manifold of dimension 1). In the informal discussion that follows, it will be useful for us to consider the "projective curve" $\overline{V}_f$ obtained by "compactifying" $V_f$ by adding a finite number of "points at infinity" (we will also assume nonsingularity at these points). Then $\overline{V}_f(\mathbb{C})$ is a closed Riemann surface, which is topologically characterized by its genus $g$ or number of "doughnut holes" on the surface. We have the following basic trichotomy, in differential geometric terms:

- $(g = 0)$. $\overline{V}_f(\mathbb{C})$ is topologically a sphere. It can be given a complete Riemannian metric of constant positive curvature, so as to be "positively curved." Its Euler characteristic is positive: $\chi > 0$.
- $(g = 1)$. $\overline{V}_f(\mathbb{C})$ is topologically a torus. It can be given a complete Riemannian metric of constant zero curvature, so as to be "flat" (like the world as imagined by Pacman.) Its Euler characteristic is zero: $\chi = 0$.
- $(g \geq 2)$. $\overline{V}_f(\mathbb{C})$ can be given a complete Riemannian metric of constant negative curvature, so as to be "negatively curved" (like on the surface of a saddle.) Its Euler characteristic is negative: $\chi < 0$.

Suppose now that $\overline{V}_f(\mathbb{Q})$ is nonempty. (This is a highly nontrivial assumption!) We then have the following trichotomy in the behavior of $\overline{V}_f(\mathbb{Q})$.

- $(g = 0)$. $\overline{V}_f(\mathbb{Q})$ is infinite.
- $(g = 1)$. $\overline{V}_f(\mathbb{Q})$ has the structure of an abelian group; and this group is in fact finitely generated. This latter statement was conjectured by Poincaré in 1901; it was proved by Mordell in 1921, with generalization by Weil in 1928.
- $(g \geq 2)$. $\overline{V}_f(\mathbb{Q})$ is finite. This was conjectured by Mordell in 1923, and proved by Faltings in 1983, which led to his 1986 Fields medal.

Thus, the trichotomy of curvature in geometry reverberates in arithmetic!

1.2. Rational points on conics. Let us consider the problem of finding rational solutions to Diophantine equations of degree 2 in two variables (i.e. equations of conics). Our analysis will serve to illustrate the utility of geometric intuition in solving Diophantine problems.

Example 1.1. Let us determine $V_f(\mathbb{Q})$ where

$$f(x,y) = x^2 + y^2 - 1. \quad (1)$$

Note that the set $V_f(\mathbb{R})$ of real solutions to equation (1) forms a unit circle in the real affine plane $\mathbb{A}^2(\mathbb{R}) = \mathbb{R}^2$. The rational solutions correspond to the points on the unit circle with rational $x$- and $y$-coordinates.
To determine $V_f(\mathbb{Q})$, first note that $(-1,0)$ lies in $V_f(\mathbb{Q})$. If $(x,y)$ is any other rational point in $V_f(\mathbb{Q})$, then the line joining $(-1,0)$ to $(x,y)$ has rational slope. Conversely, a line of rational slope $t \in \mathbb{Q}$ through $(-1,0)$ intersects $V_f$ at exactly at one other point $(x_t,y_t)$, and this point is rational. Indeed, note that $(x_t,y_t)$ solves simultaneously the equations
\[ x^2 + y^2 = 1 \quad \text{and} \quad y = t(x + 1). \]
Substitution gives
\[ x^2 + t^2(x + 1)^2 = 1 \implies (x + 1)((1 + t^2)x - (1 - t^2)) = 0. \]
If $x = -1$ then $y = 0$, which is excluded. If $x \neq 0$, we see that
\[ (x_t,y_t) = \left( \frac{1 - t^2}{1 + t^2}, \frac{2t}{1 + t^2} \right). \]
Therefore, we have
\[ V_f(\mathbb{Q}) = \left\{ \left( \frac{1 - t^2}{1 + t^2}, \frac{2t}{1 + t^2} \right) : t \in \mathbb{Q} \right\} \cup \{(-1,0)\}. \]
Note that we may informally view $(-1,0)$ as the point obtained by letting the parameter $t$ “tend to infinity.” Indeed, note that
\[ \lim_{|t| \to \infty} \left( \frac{1 - t^2}{1 + t^2}, \frac{2t}{1 + t^2} \right) = (-1,0). \]
To give a cleaner presentation of $V_f(\mathbb{Q})$, it will be useful to work with projective curves, which are introduced below.

Fix a base field $k$ (say $\mathbb{Q}$). The projective $n$-space $\mathbb{P}^n$ over $k$ is defined by setting, for each field extension $L/k$,
\[ \mathbb{P}^n(L) = \frac{L^{n+1} \setminus \{(0, \ldots, 0)\}}{\sim} \]
where the equivalence relation $\sim$ on $L^{n+1} \setminus \{(0, \ldots, 0)\}$ is given by
\[ (a_0, \ldots, a_n) \sim (b_1, \ldots, b_n) \]
if and only if $(a_0, \ldots, a_n) = (\lambda \cdot b_1, \ldots, \lambda \cdot b_n)$ for some $\lambda \in L^\times$. Note that $\mathbb{P}^n(L)$ may be viewed as the space of lines through the origin (more precisely, one-dimensional vector subspaces) in the vector space $L^{n+1}$ over $L$. Given $a_0, \ldots, a_n \in L$ not all zero, we shall denote the class of $(a_0, \ldots, a_n)$ in $\mathbb{P}^n(L)$ by $(a_0 : \ldots : a_n)$.

**Example 1.2.** As a set,
\[ \mathbb{P}^1(L) = \{[1 : a] : a \in L\} \cup \{[0 : 1]\} = L \cup \{\infty\}. \]
For example, $\mathbb{P}^1(\mathbb{C})$ is the Riemann sphere giving a one-point compactification of the complex plane $\mathbb{C}$. In general, we have
\[ \mathbb{P}^n(L) = \{[1 : a_1 : \ldots : a_n] : a_i \in L\} \cup \{[0 : a_1 : \ldots : a_n] : a_i \in L\} = \mathbb{A}^n(L) \cup \mathbb{P}^{n-1}(L). \]
Thus, $\mathbb{P}^n$ may be viewed as a “compactification” of the affine space $\mathbb{A}^n$ obtained by adding $\mathbb{P}^{n-1}$ “at infinity.”

For a homogeneous polynomial $F(X_0, X_1, X_2) \in \mathbb{Q}[X_0, X_1, X_2]$ of degree $d \geq 1$ and field extension $L/\mathbb{Q}$, the zero set
\[ \mathbb{V}_F(L) = \{(a_0 : a_1 : a_2) \in \mathbb{P}^2(L) : F(a_0, a_1, a_2) = 0\} \]
is well-defined, since $F(\lambda X_0, \lambda X_1, \lambda X_2) = \lambda^d F(X_0, X_1, X_2)$ for every $\lambda \in L^\times$. We refer to $\mathbb{V}_F$ as a projective curve defined over $\mathbb{Q}$. 3
Example 1.3. Let \( f(x, y) = x^2 + y^2 - 1 \) as in Example 1. The zero set \( \mathbb{V}_F(C) \) in \( \mathbb{P}^2(\mathbb{C}) \) of the homogenization
\[
F(X_0, X_1, X_2) = X_0^2 f \left( \frac{X_1}{X_0}, \frac{X_2}{X_0} \right) = X_1^2 + X_2^2 - X_0^2
\]
is given by the union of two sets
\[
\{ [1 : a_1 : a_2] \in \mathbb{P}^2(\mathbb{C}) : F(1, a_1, a_2) = f(a_1, a_2) = 0 \} = V_f(\mathbb{C})
\]
and
\[
\{ [0 : a_1 : a_2] \in \mathbb{P}^2(\mathbb{C}) : F(0, a_1, a_2) = a_1^2 + a_2^2 = 0 \} = \{ (0 : 1 : i), (0 : 1 : -i) \}.
\]
Thus, we have
\[
\mathbb{V}_F(\mathbb{C}) = V_f(\mathbb{C}) \cup \{ (0 : 1 : i), (0 : 1 : -i) \}
\]
and we may view \((0 : 1 : i)\) and \((0 : 1 : -i)\) as points “at infinity” of \( V_f \) in \( \mathbb{P}^2 \). Note that we have a well-defined map
\[
\mathbb{P}^1(\mathbb{C}) \to \mathbb{V}_F(\mathbb{C}) \subset \mathbb{P}^2(\mathbb{C})
\]
given by
\[
(t_0 : t_1) \mapsto (t_0^2 + t_1^2 : t_0^2 - t_1^2 : 2t_0t_1),
\]
which in fact establishes a bijection between (the complex points and) rational points of \( \mathbb{P}^1 \) and \( \mathbb{V}_F \). Note that the restriction of the map above to the affine line \( \mathbb{A}^1 = \{ [1 : t] \} \subset \mathbb{P}^1 \) can be written
\[
t = (1 : t) \mapsto (1 + t^2 : 1 - t^2 : 2t) = \left( 1 : \frac{1 - t^2}{1 + t^2} : \frac{2t}{1 + t^2} \right) = \left( \frac{1 - t^2}{1 + t^2} : \frac{2t}{1 + t^2} \right)
\]
which agrees with the rational parametrization given in Example 1. We will later revisit this example and show that the map given above is an isomorphism of algebraic curves over \( \mathbb{Q} \), so we may write
\[
\mathbb{P}^1 \simeq V_F.
\]

It is easy to see that, although our discussion was limited to a particular conic section \( x^2 + y^2 = 1 \), the same argument can be used a to give a rational parametrization of rational points for an arbitrary nondegenerate conic section, provided that it has at least one rational point to begin with. This leads to the following result:

**Theorem 1.4.** Let \( C \) be a geometrically irreducible (smooth) projective curve of degree 2 in \( \mathbb{P}^2 \) defined over \( \mathbb{Q} \). Then the following are equivalent:

1. \( C(\mathbb{Q}) \) is nonempty.
2. \( C \simeq \mathbb{P}^1 \) over \( \mathbb{Q} \).

This naturally leads us to the problem of finding necessary and sufficient conditions for a curve \( C \) as above to have a rational point. This will be solved by the Hasse-Minkowski theorem, which will be covered in later lectures.

2. **Lecture 2**

2.1. **Absolute values.**

**Definition 2.1.** An absolute value on a field \( K \) is a function
\[
| \cdot | : K \to \mathbb{R}_{\geq 0}
\]
such that, for all \( x, y \in K \),

- \(|x| = 0 \iff x = 0\)
- \(|xy| = |x| \cdot |y|\)
- \(|x + y| \leq |x| + |y|\) (triangle inequality)
If $|\cdot|$ satisfies the strong triangle inequality
$$|x + y| \leq \max\{|x|, |y|\}$$
for all $x, y \in K$, then $|\cdot|$ is nonarchimedean; otherwise, it is archimedean.

**Example 2.2.** We have the following.

1. The usual absolute value on $\mathbb{R}$ or $\mathbb{C}$, and its restriction to subfields.
   
   We shall denote by $|\cdot|_\infty$ the restriction of this to $\mathbb{Q}$.

2. The trivial absolute value on any field $K$ is defined by
   $$|x| = \begin{cases} 1 & x \neq 0, \\ 0 & x = 0. \end{cases}$$

We will construct new absolute values $|\cdot|_p$ on $\mathbb{Q}$ for each prime number $p$. By the Fundamental Theorem of Arithmetic, for each $n \in \mathbb{Q}^\times$ there exist unique $u \in \{\pm 1\}$ and $n_2, n_3, n_4, \ldots \in \mathbb{Z}$ (indexed by primes, with all but finitely many $n_p$ zero) such that
$$n = u \prod_{p \text{ prime}} p^{n_p}.$$  

**Definition 2.3.** For each prime $p$, the $p$-adic valuation on $\mathbb{Q}$ is the function
$$v_p : \mathbb{Q}^\times \rightarrow \mathbb{Z}$$
given by $v_p(n) = n_p$. We extend $v_p$ to a function $v_p : \mathbb{Q} \rightarrow \mathbb{Z} \cup \{+\infty\}$ by setting $v_p(0) = +\infty$.

**Lemma 2.4.** For any prime $p$ and $x, y \in \mathbb{Q}$, we have

1. $v_p(x) = +\infty \iff x = 0$
2. $v_p(xy) = v_p(x) + v_p(y)$
3. $v_p(x + y) \geq \min\{v_p(x), v_p(y)\}$.

**Proof.** (1) and (2) are clear. To prove (3), we may assume $x \neq 0$ and $y \neq 0$ since otherwise the claim is clear (note that $+\infty \geq n$ for any $n \in \mathbb{Z}$) by convention. Let us write $x = p^{r_p}u$ and $y = p^{r_q}v$ with $u, v, r, s \in \mathbb{Z}$ not divisible by $p$. Without loss of generality, $v_p(x) = n = v_p(y).$ Then
$$\frac{r}{s} + p^{m-n}u = p^{n-r}v + p^{m-n}us.$$  

Since $p \nmid sv$, it follows that $v_p(x + y) \geq v_p(x) = \min\{v_p(x), v_p(y)\}$ as desired. \qed

**Corollary-Definition 2.5.** The function $|\cdot|_p : \mathbb{Q} \rightarrow \mathbb{R}_{\geq 0}$ given by $|x|_p = p^{-v_p(x)}$ is a nonarchimedean absolute value, called the $p$-adic absolute value on $\mathbb{Q}$.

**Remark.** Note the **product formula**: for every $x \in \mathbb{Q}^\times$, we have $\prod_{p \leq \infty} |x|_p = 1$.

**2.2. Completions.** Let $K$ be a field with absolute value $|\cdot|$.

**Definition 2.6.** A sequence $(a_i)$ in $K$

(a) **converges** to $\ell \in K$ if for all $\epsilon > 0$ there exists $N \geq 0$ such that $|a_i - \ell| < \epsilon$ for all $i \geq N$.

(b) **is Cauchy** if $\forall \epsilon > 0$ there exists $N$ such that $|a_i - a_j| < \epsilon$ for all $i, j \geq N$.

Note that every convergent sequence is Cauchy. We say that $K$ is complete with respect to $|\cdot|$ if every Cauchy sequence converges in $K$.

**Example 2.7.** $\mathbb{Q}$ is not complete with respect to the usual absolute value $|\cdot|_\infty$, but $\mathbb{R}$ is.

**Definition 2.8.** Two sequences $(a_i)$ and $(b_i)$ in $K$ are **equivalent** if $\lim_{i \to \infty} |a_i - b_i| = 0$.

**Definition 2.9.** The **completion** $\hat{K}$ of $K$ with respect to $|\cdot|$ is the set of equivalence classes of Cauchy sequences in $K$. 

5
Note that \( \hat{K} \) forms a field with respect to termwise addition and multiplication:

\[
[(a_i) + (b_i)] = [(a_i + b_i)], \\
[(a_i) \cdot (b_i)] = [(a_i \cdot b_i)].
\]

The additive identity in \( \hat{K} \) is the class \( 0 = [(0, 0, 0, \ldots)] \) and the multiplicative identity is the class \( 1 = [(1, 1, 1, \ldots)] \). If \((a_i)\) is a Cauchy sequence in \( K \) and \((a_i) \neq 0\), then there exists \( N \geq 0 \) such that \( a_i \neq 0 \) for all \( i \geq N \). So defining

\[
b_i = \begin{cases} 
0 & i \leq 0, \\
a_i^{-1} & i \geq N
\end{cases}
\]

we see that \([(a_i)] \cdot [(b_i)] = [(0, \ldots, 0, 1, 1, \ldots)] = 1\) so \([(a_i)]\) is invertible.

Let \( \| \cdot \| : \hat{K} \to \mathbb{R}_{\geq 0} \) be the function given by \( \|(a_i)\| = \lim_{i \to \infty} |a_i| \). Then \( \| \cdot \| \) is an absolute value on \( \hat{K} \), and \( \hat{K} \) is complete with respect to \( \| \cdot \| \). The restriction of \( \| \cdot \| \) to \( K \) under the embedding \( K \to \hat{K} \) given by \( a \mapsto [(a, a, \ldots)] \) recovers the original absolute value \( |\cdot| \) on \( K \). In the absence of confusion, we will use the same notation for the absolute value on \( \hat{K} \) and the absolute value on \( K \). Finally, note that if \((a_i)\) is a Cauchy sequence in \( K \) and \( a = [(a_i)] \) denotes its class in \( \hat{K} \), then

\[
\lim_{i \to \infty} a_i = a \text{ in } \hat{K}.
\]

**Proposition 2.10.** Let \( L \) be a field complete with respect to an absolute value \( |\cdot| \), and let \( K \subset L \) be an arbitrary subfield (so that \( K \) inherits an absolute value from \( L \)).

1. The inclusion \( K \subset L \) extends uniquely to a field embedding \( \hat{K} \to L \).
2. If every element of \( L \) is a limit of a sequence in \( K \), then \( \hat{K} \approx L \).

**Proof.** The embedding \( \hat{K} \to L \) is given by \([(a_i)] \mapsto \lim_{i \to \infty} a_i \). \( \square \)

**Definition 2.11.** For each prime \( p \), the field \( \mathbb{Q}_p \) of \( p \)-adic numbers is the completion of \( \mathbb{Q} \) with respect to the \( p \)-adic absolute value \( |\cdot|_p \).

2.3. Ostrowski’s theorem.

**Definition 2.12.** Two absolute values \( |\cdot|, |\cdot|' \) on a field \( K \) are equivalent if there is a real number \( \alpha > 0 \) such that

\[
|x|' = |x|^\alpha
\]

for every \( x \in K \).

**Remark.** If two absolute values \( |\cdot| \) and \( |\cdot|' \) on \( K \) are equivalent, then they have the same completions \((\hat{K}, |\cdot|) \approx (\hat{K}, |\cdot|')\).

**Theorem 2.13** (Ostrowski). Up to equivalence, the nontrivial absolute values on \( \mathbb{Q} \) are precisely the \( p \)-adic absolute values \( |\cdot|_p \) for \( p \) prime and \( |\cdot|_\infty \).

**Proof.** Let \( \| \cdot \| \) be an absolute value on \( \mathbb{Q} \). By multiplicativity of \( \| \cdot \| \), it suffices that \( \| \cdot \| \) agrees on \( \mathbb{Z}_{\geq 1} \) with the trivial absolute value or \( |\cdot|_p^\alpha \) for some \( \alpha > 0 \) and \( p \leq \infty \). We have three cases:

**Case I.** Suppose there exists \( b \in \mathbb{Z}_{\geq 1} \) with \( \|b\| > 1 \).

Let \( b \) be the smallest such. Note that \( b \geq 2 \) since \( |1| = 1 \). Let \( \alpha > 0 \) be a real number such that \( |b| = b^\alpha = |b|^\alpha_\infty \). For any \( n \in \mathbb{Z}_{\geq 1} \), let us write

\[
n = a_0 + a_1 b + \cdots + a_s b^s, \quad 0 \leq a_i < b, \quad b^s \leq n < b^{s+1}.
\]
Then we have
\[
\|n\| \leq \|a_0\| + \|a_1\|\|b\| + \cdots + \|a_s\|\|b\|^s \\
\leq 1 + b^\alpha + \cdots + b^{\alpha s} \quad \text{(since } \|a_i\| \leq 1 \text{ by minimality of } b) \\
= (1 + b^{-\alpha} + \cdots + b^{-\alpha s})b^{\alpha s} \\
\leq cn^\alpha
\]
for some constant $c > 0$ independent of $n$. We have here used the observation that $1 + b^{-\alpha} + \cdots + b^{-\alpha s}$ is a partial sum of the geometric series $\sum_{i \geq 0} b^{-\alpha i}$ which is (absolutely) convergent whence bounded, since $0 < b^{-\alpha} < 1$. Thus, for each integer $N \geq 1$, if we apply the above argument to $n^N$ we obtain
\[
\|n^N\| \leq c(n^N)^\alpha \iff \|n\| \leq c^{1/N} n^\alpha.
\]
Letting $N \to \infty$, we thus deduce $\|n\| \leq n^\alpha$ for every $n \in \mathbb{Z}_{\geq 1}$. Next, noting that
\[
\|b^{s+1}\| \leq \|b^{s+1} - n\| + \|n\|
\]
we have
\[
\|n\| \geq \|b^{s+1}\| - \|b^{s+1} - n\| \\
\geq b^{(s+1)\alpha} - (b^{s+1} - n)^\alpha \quad \text{(by what we proved earlier)} \\
\geq b^{(2s+1)\alpha} - (b^{s+1} - b^\alpha)^\alpha = b^{(s+1)\alpha} \left(1 - \left(1 - \frac{1}{b}\right)^\alpha\right) \\
\geq cn^\alpha
\]
for some constant $c > 0$ independent of $n$. Applying the same argument as above, we find that $\|n\| \geq n^\alpha$ for every $n \in \mathbb{Z}_{\geq 1}$. Thus $\|n\| = n^\alpha$ for every $n \in \mathbb{Z}_{\geq 1}$.

**Case II.** Suppose $\|n\| = 1$ for every $n \in \mathbb{Z}_{\geq 1}$. Then $\|\cdot\|$ is the trivial absolute value.

**Case III.** Suppose $\|n\| \leq 1$ for every $n \in \mathbb{Z}_{\geq 1}$, and $\|b\| < 1$ for some $b \in \mathbb{Z}_{\geq 1}$.

Let $b$ be the smallest such. Then $b$ must be a prime number (say $p$), since if $b = rs$ with $r, s \in \mathbb{Z}_{\geq 1}$ then $1 > \|b\| = \|r\|\|s\|$ which implies $\|r\| < 1$ or $\|s\| < 1$, implying $b = r$ or $b = s$ by minimality of $b$. We now claim that $\|q\| = 1$ for every prime $q \neq p$. Indeed, if $\|q\| < 1$ then, for every $N \geq 1$, by coprimality of $p^N$ and $q^N$ there exist $u, v \in \mathbb{Z}$ such that
\[
1 = up^N + vq^N
\]
which implies
\[
1 \leq \|u\|\|p\|^N + \|v\|\|q\|^N \leq \|p\|^N + \|q\|^N.
\]
Taking $N$ to be sufficiently large, we obtain a contradiction; hence, $\|q\| = 1$ for any prime $q \neq p$.

Now, let $\alpha > 0$ be a real number such that $\|p\| = p^{-\alpha} = \|p\|_p^\alpha$. Then for any $n \in \mathbb{Z}_{\geq 1}$ we have
\[
\|n\| = \prod_q q^{v_p(n)} = \|p\|^{v_p(n)} = p^{-\alpha v_p(n)} = |n|^\alpha_p.
\]

3. Lecture 3

**Definition 3.1.** A discrete valuation on a field $K$ is a function
\[
v : K \to \mathbb{Z} \cup \{+\infty\}
\]
such that, for all $x, y \in K$,
- $v(x) = +\infty \iff x = 0$
- $v(xy) = v(x) + v(y)$
- $v(x + y) \geq \min\{v(x), v(y)\}$
Example 3.2. We have the following examples.

- The $p$-adic valuation $v_p$ on $\mathbb{Q}$ extends to a discrete valuation on $\mathbb{Q}_p$ by $v_p(\cdot) = -\log_p |\cdot|_p$.
- Let $\mathbb{C}(t)$ be the field of rational functions in one variable $t$ over $\mathbb{C}$. For every $a \in \mathbb{C}$, the function
  
  \[ v_a(f) = \text{order of vanishing of } f \text{ at } a \quad \text{for} \quad f \in \mathbb{C}(t)^x \]

  is a discrete valuation on $\mathbb{C}(t)$.
- The trivial valuation on $K$ is given by $v(x) = 0$ for any $x \in K^x$.

Definition-Lemma 3.3. Let $v : K \to \mathbb{Z}$ be a nontrivial discrete valuation.

(a) The valuation ring of $v$ is $\mathcal{O}_v = \{ x \in K : v(x) \geq 0 \}$.
(b) $\mathcal{O}_v$ has a unique maximal ideal $m_v = \{ x \in K : v(x) > 0 \}$ and $\mathcal{O}_v^x = \mathcal{O}_v \setminus m_v$.
(c) $m_v$ is principal. Any $\pi \in m_v$ such that $m_v = \pi \mathcal{O}_v$ is called a uniformizer of $\mathcal{O}_v$. Every nonzero ideal of $\mathcal{O}_v$ is of the form $\pi^k \mathcal{O}_v$ for some $k \geq 0$.
(d) We have $\text{Frac}\mathcal{O}_v = K$, and in fact $\mathcal{O}_v[\frac{1}{\pi}] = K$. The field $k_v = \mathcal{O}_v/m_v$ is called the residue field of the valuation.

Proof. (a) Since $v(0) = +\infty$ and $v(1) = 0$, we have $0, 1 \in \mathcal{O}_v$. If $x, y \in \mathcal{O}_v$ so $v(x), v(y) \geq 0$, then $v(xy) = v(x) + v(y) \geq 0$ and $v(x+y) \geq \min\{v(x), v(y)\} \geq 0$ so $xy \in \mathcal{O}_v$ and $x+y \in \mathcal{O}_v$. This shows that $\mathcal{O}_v$ is a subring of $K$.

(b) Arguing as in (a), we see that $m_v$ is an ideal of $\mathcal{O}_v$. If $x \in \mathcal{O}_v \setminus m_v$, then $v(x) = 0$ which implies that

\[ v\left(\frac{1}{x}\right) = v\left(\frac{1}{x}\right) + v(x) = v(1) = 0 \]

so that $1/x \in \mathcal{O}_v$ and $x \in \mathcal{O}_v^x$. It follows that $\mathcal{O}_v^x = \mathcal{O}_v \setminus m_v$, the inclusion $\mathcal{O}_v^x \subseteq \mathcal{O}_v \setminus m_v$ being obvious. This shows that $m_v$ is the unique maximal ideal of $\mathcal{O}_v$, as desired. (Recall that $R^x = R \setminus \bigcup m \in R$ maximal ideal $m$ for any commutative ring $R$.)

(c) Let $\pi \in m_v$ be any element such that $v(\pi) = \min\{v(x) : x \in m_v\}$. Then for any $x \in m_v$,

\[ v\left(\frac{x}{\pi}\right) = v(x) - v(\pi) \geq 0 \implies x/\pi \in \mathcal{O}_v \implies x \in \pi \mathcal{O}_v. \]

This shows that $m_v = \pi \mathcal{O}_v$. In fact, for any $x \in m_v$, we have $v(x/\pi) = 0$ or $x/\pi \in m$ (and in particular $v(x/\pi) \geq v(\pi)$ by minimality of $v(\pi)$). By an inductive argument, we conclude that $v(\pi) | v(x)$ for every $x \in m_v$. So if $I \subseteq \mathcal{O}_v$ is any nonzero ideal, then $\min\{v(x) : x \in I\} = k \cdot v(\pi)$ for some $k \in \mathbb{Z}_{\geq 0}$, and $I = \pi^k \mathcal{O}_v$ by arguing as above.

We remark that, by the argument above, each nonzero $x \in \mathcal{O}_v$ can be written in the form $\pi^k u$ where $k = v(x) / v(\pi) \in \mathbb{Z}_{\geq 0}$ and $u \in \mathcal{O}_v^x$.

(d) Let $\pi \in \mathcal{O}_v$ be a uniformizer. For any $x \in K$, we have $v(\pi^N x) = N v(\pi) + v(x) \geq 0$ for $N \gg 0$ and hence $\pi^N x \in \mathcal{O}_v$ and $x \in \mathcal{O}_v[1/\pi]$. \qed

3.2. $p$-adic integers.

Definition 3.4. The ring $\mathbb{Z}_p$ of $p$-adic integers is the valuation ring of $\mathbb{Q}_p$ with respect to $v_p$:

\[ \mathbb{Z}_p = \{ x \in \mathbb{Q}_p : v_p(x) \geq 0 \} = \{ x \in \mathbb{Q}_p : |x|_p \leq 1 \}. \]

Proposition 3.5. The following is immediate from the previous subsection.

1. The unique maximal ideal of $\mathbb{Z}_p$ is $p\mathbb{Z}_p$.
2. The group of $p$-adic units of $\mathbb{Z}_p^x = \mathbb{Z}_p \setminus p\mathbb{Z}_p$.
3. Every $x \in \mathbb{Z}_p$ is of the form $p^nu$ with $n = v_p(x) \in \mathbb{Z}_{\geq 0}$ and $u \in \mathbb{Z}_p^x$.

Proposition 3.6. The ring $\mathbb{Z}_p$ is the $p$-adic closure of $\mathbb{Z}$ in $\mathbb{Q}_p$. In other words:

1. If $a = \lim_{i \to \infty} a_i$ in $\mathbb{Q}_p$ with $a_i \in \mathbb{Z}$, then $a \in \mathbb{Z}_p$. \(8\)}
(2) For every \( a \in \mathbb{Z}_p \), there is a sequence \( (a_i) \) with \( a_i \in \mathbb{Z} \) such that \( a = \lim_{i \to \infty} a_i \) in \( \mathbb{Q}_p \).

**Proof.** (1) If \( \lim_{i \to \infty} a_i = a \) with \( a_i \in \mathbb{Z} \), then \( |a_i|_p = \lim_{i \to \infty} |a_i|_p \leq 1 \).

(2) Let \( a \in \mathbb{Z}_p \) be given, and write \( a = \lim_{i \to \infty} a_i \) with \( a_i \in \mathbb{Q} \). Since \( \lim_{i \to \infty} |a_i| = |a|_p \leq 1 \) and \( \cdot |_p \) takes values in \( \mathbb{P}_\mathbb{Q} \), we see that \( a = 0 \) or \( |a_i|_p \) must be eventually constant with \( |a_i|_p \leq 1 \) for \( i \gg 0 \). Thus, we may assume without loss of generality that \( |a_i|_p \leq 1 \) for all \( i \). Let us write

\[
\frac{a_i}{c_i} = \frac{b_i}{c_i} \quad \text{with} \quad b_i, c_i \in \mathbb{Z} \text{ such that } \gcd(b_i, c_i) = 1 \quad \text{and} \quad p \nmid c_i.
\]

For each \( i \geq 1 \), by coprimality of \( p^i \) and \( c_i \) there exist \( u_i, v_i \in \mathbb{Z} \) such that \( b_i = u_i p^i + v_i c_i \). Then

\[
\frac{a_i}{c_i} = \frac{u_i p^i + v_i c_i}{c_i} = v_i + p^i \frac{u_i}{c_i} \implies |a_i - v_i|_p = \left| p^i \frac{u_i}{c_i} \right|_p \leq p^{-i}.
\]

This shows that \( (v_i) \sim (a_i) \) and \( \lim_{i \to \infty} v_i = \lim_{i \to \infty} a_i = a \). \( \square \)

**Proposition 3.7 (p-adic expansion).** We have the following.

(a) Every \( a \in \mathbb{Z}_p \) can be written uniquely in the form

\[
a = b_0 + b_1p + b_2p^2 + \ldots, \quad b_i \in \{0, \ldots, p-1\}.
\]

Conversely, any series of the form \( \sum_{i=0}^{\infty} b_ip^i \) with \( b_i \in \{0, \ldots, p-1\} \) converges in \( \mathbb{Z}_p \).

(b) Each \( a \in \mathbb{Q}_p \) can be written in the form

\[
a = \sum_{i=k}^{\infty} b_ip^i, \quad b_i \in \{0, \ldots, p-1\}
\]

for some \( k \in \mathbb{Z} \) such that \( v_p(a) = \min\{i : b_i \neq 0\} \).

**Proof.** (a) Let \( a \in \mathbb{Z}_p \) be given. By Proposition 3.6, there exists \( b'_0 \in \mathbb{Z} \) such that \( |a - b'_0|_p < 1 \). Then

\[
a = b'_0 + pa_1 \quad \text{for some} \quad a'_1 \in \mathbb{Z}_p.
\]

Moreover, writing \( b'_0 = pq + b_0 \) with \( q \in \mathbb{Z} \), \( b_0 \in \{0, \ldots, p-1\} \) and \( a_1 = a'_1 + q \), we thus can write

\[
a = b_0 + pa_1, \quad b_0 \in \{0, \ldots, p-1\}, a_1 \in \mathbb{Z}_p.
\]

Moreover, \( b_0 \) and \( a_1 \) are uniquely determined by \( a \). Indeed, if we have \( b_0 + pa_1 = b'_0 + pa'_1 \) with \( b'_0 \in \{0, \ldots, p-1\} \) and \( a'_1 \in \mathbb{Z}_p \), then \( b_0 - b'_0 = p(a'_1 - a_1) \) showing that \( b_0 - b'_0 \) is an integer with \( |b_0 - b'_0|_p < p \) satisfying \( v_p(b_0 - b'_0) > 0 \), and therefore \( b_0 = b'_0 \) and \( a_1 = a'_1 \). Similarly, we can write \( a_1 \) uniquely as

\[
a_1 = b_1 + pa_2, \quad b_1 \in \{0, \ldots, p-1\}, a_2 \in \mathbb{Z}_p.
\]

Note then that \( a = b_0 + b_1p + p^2a_2 \) with \( a_2 \in \mathbb{Z}_p \). By recursion we obtain a sequence \( \{b_0, b_1, b_2, \ldots\} \) of elements in \( \{0, \ldots, p-1\} \) such that the sequence of partial sums \( \{b_0, b_0 + b_1p, b_0 + b_1p + b_2p^2, \ldots\} \) is Cauchy and converges to \( a \).

(b) This follows from the fact that \( \mathbb{Q}_p = \mathbb{Z}_p[\mathbb{Z}_p]/[1/p] \). \( \square \)

**Lemma 3.8.** For each \( n \geq 1 \), the inclusion \( \mathbb{Z} \to \mathbb{Z}_p \) induces a natural isomorphism

\[
\mathbb{Z}/p^n\mathbb{Z} \sim \mathbb{Z}_p/p^n\mathbb{Z}_p
\]

compatible with the projections

\[
\begin{array}{ccc}
\mathbb{Z}/p^{n+1}\mathbb{Z} & \longrightarrow & \mathbb{Z}/p^n\mathbb{Z} \\
\downarrow & & \downarrow \\
\mathbb{Z}_p/p^{n+1}\mathbb{Z}_p & \longrightarrow & \mathbb{Z}_p/p^n\mathbb{Z}_p
\end{array}
\]

**Proof.** The class of \( a = \sum_{i=0}^{\infty} b_ip^i \in \mathbb{Z}_p \) in \( \mathbb{Z}_p/p^n\mathbb{Z}_p \) is the same as the class of \( a' = \sum_{i=0}^{n} b_ip^i \in \mathbb{Z} \), so the homomorphism \( \mathbb{Z} \to \mathbb{Z}_p/p^n\mathbb{Z}_p \) is surjective. The kernel of this surjection is \( \mathbb{Z} \cap p^n \mathbb{Z}_p = p^n \mathbb{Z} \). \( \square \)
3.3. $\mathbb{Z}_p$ as an inverse limit.

**Definition 3.9.** An inverse system of sets is a sequence $(A_n)$ of sets and maps $(f_n)$ of the form

$$\cdots \rightarrow A_{n+1} \xrightarrow{f_n} A_n \rightarrow \cdots \rightarrow A_2 \xrightarrow{f_2} A_1 \xrightarrow{f_1} A_0.$$ 

The inverse limit $A = \lim_{\leftarrow} A_n$ of such an inverse system is given by

$$A = \left\{ (a_n) \in \prod_{n=0}^{\infty} A_n : f_n(a_{n+1}) = a_n \text{ for all } n \geq 0 \right\}$$

equipped with projections $\epsilon_n : A \rightarrow A_n$ such that $\epsilon_n = f_n \circ \epsilon_{n+1}$ for all $n \geq 0$.

**Remark.** If $\{(A_n), (f_n)\}$ is an inverse system of groups (resp. rings) with group homomorphisms (resp. ring homomorphisms), then $\lim_{\leftarrow} A_n$ naturally has the structure of a group (resp. ring).

**Example 3.10.** The sequence of projections $\cdots \rightarrow \mathbb{Z}/p^{n+1}\mathbb{Z} \rightarrow \mathbb{Z}/p^n\mathbb{Z} \rightarrow \cdots \rightarrow \mathbb{Z}/p^2\mathbb{Z} \rightarrow \mathbb{Z}/p\mathbb{Z}$ forms an inverse system of rings.

**Proposition 3.11.** The projections $\mathbb{Z}_p \rightarrow \mathbb{Z}/p^n\mathbb{Z}_p$ induce an isomorphism

$$\mathbb{Z}_p \simeq \lim_{\leftarrow} \mathbb{Z}/p^n\mathbb{Z}_p \simeq \lim_{\leftarrow} \mathbb{Z}/p^n\mathbb{Z}.$$

**Proof.** Under the natural map $\mathbb{Z}_p \rightarrow \lim_{\leftarrow} \mathbb{Z}/p^n\mathbb{Z}_p$, each element $([b_0], [b_0+b_1p], [b_0+b_1p+b_2p^2], \ldots) \in \lim_{\leftarrow} \mathbb{Z}/p^n\mathbb{Z}_p$ is in the image of $\sum_{i=0}^{\infty} b_i p^i \in \mathbb{Z}_p$. The kernel of this map is $\bigcap_n p^n\mathbb{Z}_p = \{0\}$. □

**Remark.** An alternative but equivalent way to build up the $p$-adic numbers $\mathbb{Q}_p$ is to first define $\mathbb{Z}_p$ as the inverse limit $\lim_{\leftarrow} \mathbb{Z}/p^n\mathbb{Z}$ and define $\text{Frac}(\mathbb{Z}_p) = \mathbb{Q}_p$. This is the approach taken e.g. in the notes of Poonen in the 2009 version of this course.

4. Lecture 4

4.1. Solving equations in $\mathbb{Z}_p$. In the previous lecture, we defined $\mathbb{Z}_p = \{ x \in \mathbb{Q}_p : |x|_p \leq 1 \}$ and showed

$$\mathbb{Z}_p \simeq \lim_{\leftarrow} \mathbb{Z}/p^n\mathbb{Z}.$$ 

We will use this to our advantage in solving equations over $\mathbb{Z}_p$. Given $f \in \mathbb{Z}_p[x_1, \ldots, x_N]$, note that reduction modulo $p^n$ gives us an inverse system

$$\cdots \rightarrow V_f(\mathbb{Z}/p^{n+1}\mathbb{Z}) \xrightarrow{\text{mod } p^n} V_f(\mathbb{Z}/p^n\mathbb{Z}) \rightarrow \cdots \rightarrow V_f(\mathbb{Z}/p\mathbb{Z}).$$

**Lemma 4.1.** For any $f \in \mathbb{Z}_p[x_1, \ldots, x_N]$, we have an isomorphism

$$V_f(\mathbb{Z}_p) \xrightarrow{\sim} \lim_{\leftarrow} V_f(\mathbb{Z}/p^n\mathbb{Z})$$

given by $P \mapsto (P \mod p^n)$.

**Proof.** Note that

$$\lim_{\leftarrow} V_f(\mathbb{Z}/p^n\mathbb{Z}) \subseteq \lim_{\leftarrow} ((\mathbb{Z}/p^n\mathbb{Z})^N) \simeq (\lim_{\leftarrow} \mathbb{Z}/p^n\mathbb{Z})^N \cong \mathbb{Z}_p^N.$$ 

So for any $(P_n) \in \lim_{\leftarrow} V_f(\mathbb{Z}/p^n\mathbb{Z})$ there exists a unique $P \in \mathbb{Z}_p^N$ such that $P \equiv P_n \mod p^n$ for every $n \geq 0$. Since $f(P) \equiv f(P_n) \equiv 0 \mod p^n$ for every $n \geq 0$, it follows that $f(P) = 0$ and thus $P \in V_f(\mathbb{Z}_p)$. This shows that the map in the lemma is a bijection. □

**Note.** Each $V_f(\mathbb{Z}/p^n\mathbb{Z})$ is a finite set, allowing us to utilize the following observation.

**Lemma 4.2.** If $\{(A_n), (f_n)\}$ is an inverse system of sets where each $A_n$ is finite and nonempty, then the inverse limit $\lim_{\leftarrow} A_n$ is nonempty.
Proof. For each \( n \) and \( k \geq 0 \), let us write \( T_{k,n} = \text{Im}(A_{n+k} \to \cdots \to A_n) \subset A_n \). Then
\[
A_n = T_{0,n} \supseteq T_{1,n} \supseteq \cdots
\]
is a nonincreasing sequence of nonempty finite sets, and hence must eventually stabilize, i.e. there is \( N(n) \) such that \( T_{k,n} = T_{\ell,n} \) for all \( k, \ell \geq N(n) \). Let \( E_n \subseteq A_n \) denote this set \( T_{k,n} \) for \( k \geq N(n) \); note that \( E_n \) is finite and nonempty. Note finally that \((E_n), (f_n|_{E_{n+1}})\) is an inverse system and moreover each map \( E_{n+1} \to E_n \) is surjective (as seen by unwinding the definitions). Thus, choosing any \( e_1 \in E_1 \), an element \( e_2 \in E_2 \) such that \( f_1(e_2) = e_1 \) and so on, we obtain a sequence 
\[
(e_n) \in \lim_{\longleftarrow} E_n \subseteq \lim_{\longleftarrow} A_n
\]
so in particular \( \lim A_n \) is nonempty. \( \square \)

Corollary 4.3. For any \( f \in \mathbb{Z}_p[x_1, \ldots, x_n] \), the following are equivalent:

(a) \( V_f(\mathbb{Z}_p) \) is nonempty.

(b) \( V_f(\mathbb{Z}/p^n\mathbb{Z}) \) is nonempty for all \( n \geq 0 \).

We now establish an analogue of the above for results for zero sets of homogeneous polynomials in projective space.

Definition 4.4. A point \( P \in \mathbb{Z}_p^N \) or \((\mathbb{Z}/p^n\mathbb{Z})^N \) is primitive if \( P \not\equiv 0 \mod p \), i.e. not all coordinates of \( P = (a_1, \ldots, a_N) \) are multiples of \( p \) in \( \mathbb{Z}_p \).

Lemma 4.5. Let \( F \in \mathbb{Z}_l[X_0, \ldots, X_N] \) be homogeneous. The following are equivalent:

(a) \( V_F(\mathbb{Q}_p) \subseteq \mathbb{P}^N(\mathbb{Q}_p) \) is nonempty.

(b) \( V_F(\mathbb{Z}_p) \subseteq \mathbb{A}^{N+1}(\mathbb{Z}_p) \) contains a primitive point.

(c) \( V_F(\mathbb{Z}/p^n\mathbb{Z}) \) has a primitive point for all \( n \geq 1 \).

Proof. (a) \( \iff \) (b). If \( (a_0, \ldots, a_N) \in V_F(\mathbb{Z}_p) \) is primitive then \( (a_0, \ldots, a_N) \not\equiv 0 \) so that we have \( (a_0 : \cdots : a_N) \in V_F(\mathbb{Q}_p) \). Conversely, if \( P \in V_F(\mathbb{Q}_p) \) is represented by \( (a_0, \ldots, a_N) \in \mathbb{Z}_p^{N+1} \) and \( m = \min\{v_p(a_0), \ldots, v_p(a_N)\} \), then we have \( (p^{-m}a_0, \ldots, p^{-m}a_N) \in \mathbb{Z}_p^N \) and moreover this point is primitive.

(b) \( \implies \) (c). This is clear.

(c) \( \implies \) (b). Let \( V_f(\mathbb{Z}_p)^\text{prim} \) denote the set of points in \( V_f(\mathbb{Z}_p) \) that are primitive, and similarly define \( V_f(\mathbb{Z}/p^n\mathbb{Z})^\text{prim} \). The result then follows from Lemma 4.2 and the observation that
\[
V_f(\mathbb{Z}_p)^\text{prim} = \lim_{\longleftarrow} V_f(\mathbb{Z}/p^n\mathbb{Z})^\text{prim}. \quad \square
\]

4.2. Hensel’s lemma. Hensel’s lemma shows that, in favorable situations, one can determine if an equation \( f = 0 \) with \( f \in \mathbb{Z}_p[x_1, \ldots, x_n] \) is solvable in \( \mathbb{Z}_p \) by checking the finite set \( V_f(\mathbb{Z}/p\mathbb{Z}) \).

Lemma 4.6 (Hensel). Let \( f \in \mathbb{Z}_p[x] \) and \( z \in \mathbb{Z}/p\mathbb{Z} \) such that \( f(a) \equiv 0 \) and \( f'(a) \not\equiv 0 \mod p \). Then there exists unique \( \bar{a} \in \mathbb{Z}_p \) such that \( f(\bar{a}) = 0 \) in \( \mathbb{Z}_p \), and \( \bar{a} \equiv a \mod p \).

Proof. We will construct uniquely a sequence \( (a_n) \in \lim \mathbb{Z}/p^n\mathbb{Z} \) such that \( f(a_n) = 0 \) in \( \mathbb{Z}/p^n\mathbb{Z} \) for each \( n \geq 1 \) and \( a_1 = a \in \mathbb{Z}/p\mathbb{Z} \). We proceed by recursion on \( n \). Suppose \( a_n \in \mathbb{Z}/p^n\mathbb{Z} \) is constructed with \( f(a_n) = 0 \) and \( a_n \equiv a \mod p \). Let \( \bar{a}_{n+1} \in \mathbb{Z}/p^{n+1}\mathbb{Z} \) be any lift of \( a_n \) (in other words, \( \bar{a}_{n+1} \equiv a_n \mod p^n \)). Note that any other lift \( a_{n+1} \) of \( a_n \) in \( \mathbb{Z}/p^{n+1}\mathbb{Z} \) is of the form
\[
a_{n+1} = \bar{a}_{n+1} + p^nz
\]
for some \( z \in \mathbb{Z}_p \) determined uniquely up to \( p\mathbb{Z}_p \) (i.e. \( z \in \mathbb{Z}_p/p\mathbb{Z}_p \)) by \( a_{n+1} \). Our goal is to show that there is a unique choice of \( z \in \mathbb{Z}/p\mathbb{Z} \) so that \( f(a_{n+1}) = 0 \) in \( \mathbb{Z}/p^{n+1}\mathbb{Z} \).

Note that, by Taylor expansion, we have
\[
f(a_{n+1}) = f(\bar{a}_{n+1} + p^nz) = f(\bar{a}_{n+1}) + f'(\bar{a}_{n+1})p^nz + \cdots
\]
where the remaining terms (⋯) in the expansion are divisible by $p^{2n}$ and hence zero modulo $p^{n+1}$. We have $f(\tilde{a}_{n+1}) = p^n$ for some $c \in \mathbb{Z}/p\mathbb{Z}$ since $f(\tilde{a}_{n+1}) \equiv f(a_n) \equiv 0 \mod p^n$. Thus,

$$f(a_{n+1}) = p^n c f(\tilde{a}_{n+1}) p^n z = p^n (c + f'(\tilde{a}_{n+1}) z).$$

Since $f'(\tilde{a}_{n+1}) \equiv f'(a) \not\equiv 0 \mod p$, it follows that $c + f'(\tilde{a}_{n+1}) z$ is uniquely solvable in $z \in \mathbb{Z}/p\mathbb{Z}$. For this unique choice of $z$, we have $f(a_{n+1}) = 0$ in $\mathbb{Z}/p^{n+1}\mathbb{Z}$ and $a_{n+1} \equiv a_n \equiv a \mod p$ as desired. This completes the induction. The element $\tilde{a} \in \mathbb{Z}_p$ uniquely corresponding to $(a_n) \in \lim \mathbb{Z}/p^n\mathbb{Z}$ satisfies $f(\tilde{a}) \equiv f(a_n) \equiv 0 \mod p^n$ for every $n \geq 1$, showing that $f(\tilde{a}) = 0$. Finally, by construction we have $\tilde{a} \equiv a \mod p$.

\textbf{Corollary 4.7.} Let $f \in \mathbb{Z}_p[x_1, \ldots, x_n]$. If $P \in V_f(\mathbb{Z}/p\mathbb{Z})$ and $\nabla f(P) = \left( \frac{\partial f}{\partial x_1}(P), \ldots, \frac{\partial f}{\partial x_n}(P) \right) \neq 0$ in $\mathbb{Z}/p\mathbb{Z}$, then there exists $\tilde{P} \in V_f(\mathbb{Z}_p)$ such that $\tilde{P} \equiv P \mod p$. In particular, $V_f(\mathbb{Z}_p) \neq \emptyset$.

\textbf{Proof.} Say $\frac{\partial f}{\partial x_1}(P) \neq 0$ (the other cases being similar). Write $P = (a_1, \ldots, a_n)$, and choose $\tilde{a}_j \in \mathbb{Z}_p$ such that $\tilde{a}_j \equiv a_j \mod p$ for $j = 2, \ldots, n$. Apply Hensel’s lemma to $f(x, \tilde{a}_2, \ldots, \tilde{a}_n) \in \mathbb{Z}_p[x]$.

We give one application of Hensel’s lemma to study of homogeneous quadratic equations.

\textbf{Proposition 4.8.} Let

$$F(X_0, X_1, X_2) = aX_0^2 + bX_1^2 + cX_2^2 \in \mathbb{Z}[X_0, X_1, X_2]$$

with $abc \neq 0$ squarefree. If $p$ is a prime such that $p \nmid 2abc$, then $\nabla F(\mathbb{Q}_p) \neq \emptyset$.

\textbf{Proof.} By the Chevalley-Warning theorem, $V_F(\mathbb{Z}/p\mathbb{Z})$ has a primitive solution, say $P$. We have

$$\nabla F = (2aX_0, 2bX_1, 2cX_2)$$

which cannot vanish at $P$ since $P$ is primitive and $p \nmid 2abc$. By (the corollary to) Hensel’s lemma, there exists $\tilde{P} \in V_F(\mathbb{Z}_p)$ such that $\tilde{P} \equiv P \mod p$. Finally, note that $\tilde{P}$ is primitive since $P$ is primitive. Thus $\nabla F(\mathbb{Q}_p)$ is nonempty by Lemma 4.5.

4.3. \textbf{Structure of $\mathbb{Z}_p^\times$.} We close this lecture by giving an application of Hensel’s lemma to the structure of $p$-adic units. For each $n \geq 1$, let $U_n = 1 + p^n\mathbb{Z}_p \leq \mathbb{Z}_p^\times$ be the kernel of the surjective homomorphism

$$\mathbb{Z}_p^\times \to (\mathbb{Z}/p^n\mathbb{Z})^\times.$$ 

We have a filtration $\mathbb{Z}_p^\times \supset U_1 \supset U_2 \supset \cdots$.

\textbf{Lemma 4.9.} We have the following.

1. $\mathbb{Z}_p^\times / U_1 \cong \mathbb{F}_p^\times$.
2. $U_n / U_{n+1} \cong \mathbb{Z}/p\mathbb{Z}$ for all $n \geq 1$.

\textbf{Proof.} (a) is clear. For (b), note that the map $U_n \to \mathbb{Z}/p\mathbb{Z}$ given by $1 + p^n z \mapsto z \mod p$ is surjective with kernel $U_{n+1}$.

\textbf{Corollary 4.10.} For each $n \geq 1$, we have $\# U_1 / U_n = p^{-n}$.

\textbf{Proposition 4.11.} Let $\mu_{p-1}$ be the set of solutions to $x^{p-1} = 1$ in $\mathbb{Z}_p^\times$. Then $\mu_{p-1}$ is a group mapping isomorphically onto $\mathbb{F}_p^\times$, and $\mathbb{Z}_p^\times = U_1 \times \mu_{p-1}$.

\textbf{Proof.} We have $\mu_{p-1} = \ker((\mathbb{Z}_p^\times \xrightarrow{x^{p-1}-1} \mathbb{Z}_p^\times) \leq \mathbb{Z}_p^\times$. Note that $f(x) = x^{p-1} - 1 = 0$ has $p-1$ distinct roots $1, \ldots, p-1 \in \mathbb{F}_p^\times$. By Hensel’s lemma, for each $a \in \mathbb{F}_p^\times$ there exists a unique $\tilde{a} \in \mu_{p-1}$ such that $\tilde{a} \equiv a \mod p$, so the composition $\mu_{p-1} \to \mathbb{Z}_p^\times \to \mathbb{F}_p^\times$ is an isomorphism. Since $\mu_{p-1} \cap U_1 = \{1\}$, we deduce that $\mathbb{Z}_p^\times = U_1 \times \mu_{p-1}$.

\textbf{Remark.} We have $\mu_{p-1} \cong \mathbb{F}_p^\times \cong \mathbb{Z}/(p-1)\mathbb{Z}$.
Proposition 4.12. Suppose that $p$ is an odd prime. Then we have an isomorphism $\mathbb{Z}_p \xrightarrow{\sim} U_1$ given by $z \mapsto (1 + p)^z$.

Proof. First, note that if $\alpha \in U_n \setminus U_{n+1}$ then $\alpha^p \in U_{n+1} \setminus U_{n+2}$. Indeed, if $\alpha = 1 + kp^n$ with $k \in \mathbb{Z}_p^\times$, then we have

$$\alpha^p = (1 + kp^n)^p = 1 + \binom{p}{1} kp^n + \binom{p}{2} k^2 p^{2n} + \cdots + k^p p^{pn} \equiv 1 + kp^{n+1} \mod p^{n+2}.$$

Now, let $u = 1 + p \in U_1 \setminus U_2$. Then, by the computations above, the class of $u$ in $U_1/U_n$ satisfies $u^{p^{n-2}} \equiv 1 \mod U_n$ and $u^{p^{n-1}} \equiv 1 \mod U_n$, so in particular $u \mod U_n$ has exact order $p^{n-1}$ in $U_1/U_n$. On the other hand $\#U_1/U_n = p^{n-1}$. This shows that $U_1/U_n$ is cyclic and generated by $u$, so we have an isomorphism

$$\mathbb{Z}/p^{n-1}\mathbb{Z} \xrightarrow{\sim} U_1/U_n$$

given by $z \mapsto u^z$. Compatibility of these isomorphisms for varying $n$ shows that

$$\mathbb{Z}_p \simeq \lim_{n \to \infty} \mathbb{Z}/p^n\mathbb{Z} \simeq \lim_{n \to \infty} U_1/U_n \simeq U_1.$$

Remark. The analysis of $U_1$ in the case where $p = 2$ is left as an exercise.

Corollary 4.13. If $p$ is an odd prime, then $\mathbb{Z}_p^\times \simeq \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}_p$.

5. Lecture 5

5.1. Quadratic forms. Let $k$ be a field of characteristic different from 2.

Definition 5.1. A quadratic form over $k$ is a homogeneous polynomial $q \in k[x_1, \ldots, x_n]$ of degree 2.

A quadratic form $q \in k[x_1, \ldots, x_n]$ determines a function $q : k^n \to k$ by evaluation, and is in fact determined by it.

Definition 5.2. A quadratic form on a finite-dimensional vector space $V/k$ is a function $q : V \to k$ such that there is an isomorphism $k^n \xrightarrow{\sim} V$ of vector spaces ($n = \dim V$) so that the composition

$$k^n \xrightarrow{\sim} V \to k$$

is a quadratic form in the previous sense. (In other words, there is a basis $e_1, \ldots, e_n$ of $V$ such that the function $q(x_1e_1 + \cdots + x_ne_n)$ is a homogeneous polynomial of degree 2 in $x_1, \ldots, x_n$.)

Definition 5.3. Let $V$ be a (finite-dimensional) vector space over $k$.

(a) A bilinear form on a $k$-vector space $V$ is a function

$$B : V \times V \to k$$

which is $k$-linear in each variables, i.e. for every $u, v, w \in V$ and $\lambda \in k$ we have $B(u + v, w) = B(u, w) + B(v, w)$ and $B(\lambda u, v) = \lambda B(u, v)$, and similarly $B(u, v + w) = B(u, v) + B(u, w)$ and $B(u, \lambda v) = \lambda B(u, v)$.

(b) A bilinear form $B : V \times V \to k$ is symmetric if $B(v, w) = B(w, v)$ for every $v \in V$.

Proposition 5.4. For each finite-dimensional vector space $V/k$, we have a bijection

$$\{\text{quadratic forms on } V\} \leftrightarrow \{\text{symmetric bilinear forms on } V\}$$

$$(q : V \to k) \mapsto B(u, v) := \frac{q(u+v) - q(u) - q(v)}{2}$$

$q(v) := B(v, v) \leftrightarrow (B : V \times V \to k)$. 

13
Example 5.5. If $q(x, y) = x^2 + y^2$ is a quadratic form on $k^2$, the corresponding bilinear form is
\[ B((x_1, y_1), (x_2, y_2)) = \frac{1}{2}((x_1 + x_2)^2 + (y_1 + y_2)^2) - (x_1^2 + y_1^2) - (x_2^2 + y_2^2) = x_1 x_2 + y_1 y_2. \]
Conversely, the quadratic form associated to the bilinear form $B((x_1, y_1), (x_2, y_2)) = x_1 x_2 + y_1 y_2$ is $q(x, y) = B((x, y), (x, y)) = x^2 + y^2$.

Note. Each quadratic form $q: k^n \to k$ is of the form $q(v) = v^t A v$ ($v \in k^n$ column vector) for some unique symmetric $n \times n$ matrix $A \in M_{n \times n}(k)$. The associated bilinear form is then $B(u, v) = u^t A v$.

Example 5.6. We have
\[ q(x, y) = 5x^2 - 2xy + 4y^2 = \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} 5 & -1 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}. \]

Definition 5.7. We define the following.
(a) The rank of a quadratic form $q$ is the rank of its associated symmetric matrix $A$.
(b) A quadratic form $q(x_1, \ldots, x_n)$ is nondegenerate if one of the following equivalent conditions holds:
- Rank of $q$ is $n$.
- The associated matrix $A$ is invertible.
- For every $v \in V$, the linear map $B(\cdot, v) : V \to k$ given by $u \mapsto B(u, v)$ is nonzero.

Definition 5.8. Two quadratic forms $q, q' : V \to k$ are equivalent if there is an invertible linear transformation $T \in \text{GL}(V)$ of $V$ such that $q'(v) = q(Tv)$ for all $v \in V$.

Example 5.9. The forms $x^2 + y^2$ and $5x^2 + 5y^2$ are equivalent over $\mathbb{Q}$, since
\[ (2x + y)^2 + (x - 2y)^2 = 5x^2 + 5y^2. \]

Proposition 5.10. Every quadratic form $q(x_1, \ldots, x_n)$ over $k$ is equivalent to a diagonal form
\[ a_1 x_1^2 + \cdots + a_n x_n^2 \]
for some $a_i \in k$.

Proof. We proceed by induction on $n$, i.e. the dimension of the vector space $V$ on which $q$ is defined, the case $n \leq 1$ being obvious. If $q \equiv 0$, then we are done (just let every $a_i = 0$). Otherwise, let $v \in V$ be such that $q(v) \neq 0$. Then $B(\cdot, v) : V \to k$ is surjective, so its kernel $v^\perp = \{ u \in V : B(u, v) = 0 \}$ has dimension $\dim v^\perp = n - 1$. Moreover, $v \notin v^\perp$ by our hypothesis on $v$. Hence, we have $V = kv \oplus v^\perp$.

For any $u = u_1 + u_2 \in V$ with $u_1 = x_1 v$ and $u_2 \in v^\perp$, we have
\[ q(u) = B(u_1 + u_2, u_1 + u_2) = B(u_1, u_1) + B(u_2, u_1) + 2B(u_1, u_2) = q(u_1) + q(u_2) \]
since $B(u_1, u_2) = 0$ by definition of $v^\perp$. By inductive hypothesis, the quadratic form $q|_{v^\perp}$ is equivalent to a diagonal form, and $q(u_1) = a_1 x_1^2$ with $a_1 = q(v)$. This shows that $q$ is equivalent to a diagonal form, completing the induction. \( \square \)

Remark. If $q$ is equivalent to $a_1 x_1^2 + \cdots + a_n x_n^2$, then the rank of $q$ is $\# \{ a_i : a_i \neq 0 \}$.

Definition 5.11. A quadratic form $q$ on $V/k$ represents $c \in k$ if $q(v) = c$ for some nonzero $v \in V$.

Proposition 5.12. If a nondegenerate quadratic over $k$ represents 0, then it represents any $c \in k$.

Proof. Let $e \in V$ be a nonzero vector such that $q(e) = 0$. Since $q$ is nondegenerate, there exists some nonzero $f \in V$ such that $B(e, f) \neq 0$; note that $e$ and $f$ must be linearly independent. Now,
\[ q(ax + by) = axy + by^2 = (ax + by)y \]
for some $a, b \in k$, with $a = 2B(e, f) \neq 0$. For any $c \in k$, we can solve $(ax + by)y = c$ by setting $y = 1$ and $x = (c - b)/a$. \( \square \)
5.2. **Hasse-Minkowski theorem.** We now begin the proof of the Hasse-Minkowski theorem.

**Theorem 5.13** (Hasse-Minkowski). A quadratic form over \(\mathbb{Q}\) represents 0 if and only if it represents 0 over \(\mathbb{Q}_p\) for every \(p \leq \infty\).

**Remark.** Actually, this result was proved by Minkowski alone; Hasse later generalized the above result to the setting where \(\mathbb{Q}\) is replaced by an arbitrary number field.

Note that the “only if” direction is clear, so we will focus on the “if” direction. So let us a fix a quadratic form \(q(x_1, \ldots, x_n) \in \mathbb{Q}[x_1, \ldots, x_n]\) which represents 0 over \(\mathbb{Q}_p\) for every \(p \leq \infty\). We shall consider the cases of different numbers \(n\) of variables separately.

**Proof in the case** \(n = 2\). Without loss of generality, we may assume that \(q(x, y) = x^2 - ay^2\) for some \(a \in \mathbb{Q}\). Note that \(q(x, y)\) represents 0 if and only if \(a \) is a square. Now, since \(q\) represents 0 over \(\mathbb{R}\), we have \(a > 0\). Since \(q\) represents 0 over \(\mathbb{Q}_p\), we have that \(v_p(a)\) is even. Combining these, we see that

\[
q = \prod_p p^{v_p(a)} = \left(\prod_p p^{v_p(a)/2}\right)^2
\]

is a square and hence \(q\) represents 0, as desired. \(\Box\)

To prove the case \(n = 3\) of the Hasse-Minkowski theorem, we record the following lemma.

**Lemma 5.14.** Let \(a, b \in k\) where \(k\) is a field of characteristic zero. Let \(N : k(\sqrt{a}) \to k\) be the norm map defined by:

- If \(a\) is not a square in \(k\), then \(N(x + y\sqrt{a}) = x^2 - ay^2\).
- If \(a\) is a square in \(k\), then \(N(x) = x\).

Then \(q(x, y, z) = x^2 - ay^2 - bz^2\) represents 0 over \(k\) if and only if \(b = N(\alpha)\) for some \(\alpha \in k(\sqrt{a})\).

**Proof.** Consider first the case where \(a\) is not a square in \(k\). If \(b = N(x + y\sqrt{a})\), then we have \(x^2 - ay^2 - b \cdot 1^2 = 0\). Conversely, if \(x^2 - ay^2 - bz^2\) represents zero, then a nontrivial solution \((x, y, z)\) must satisfy \(b = 0\), whence \(b = N\left(\frac{x}{z} + \frac{y}{z}\sqrt{a}\right)\).

Consider next the case where \(a = c^2\) is a square with \(c \in k\). We claim both statements involved always hold true. Indeed, first \(x^2 - ay^2 = (x + cy)(x - cy)\) which is equivalent to \(xy\). The latter represents everything in \(k\), so \(x^2 - ay^2 - bz^2 = 0\) has a solution with \(b = 1\). Next, \(N(b) = b\). \(\Box\)

6. Lecture 6

6.1. **Hasse-Minkowski theorem (continued).** We continue the notation and proof from the previous lecture.

**Proof of HM for** \(n = 3\). Without loss of generality, we may assume \(q(x, y, z) = x^2 - ay^2 - bz^2\) where \(a, b \in \mathbb{Z}\) are nonzero squarefree integers. We shall assume that \(q\) represents 0 over \(\mathbb{Q}_p\) for all \(p \leq \infty\), and show that it must represent 0 over \(\mathbb{Q}\). We shall proceed by induction on \(m := |a| + |b|\). First, suppose that \(m \leq 2\). Then \(q\) must be one of the forms

\[
x^2 + y^2 + z^2, \quad x^2 + y^2 - z^2, \quad x^2 - y^2 + z^2, \quad x^2 - y^2 - z^2.
\]

The first quadratic form does not represent 0 over \(\mathbb{R}\), so is excluded. The remaining quadratic forms all clearly represent 0 over \(\mathbb{Q}\) so we are done.

So suppose now that \(m > 2\). Without loss of generality, we may assume \(|b| \geq |a|\) and in particular \(|b| \geq 2\). If \(p\) is a prime number dividing \(b\), then by our assumption there is a primitive triple.
\((x, y, z) \in \mathbb{Z}_p^3\) such that \(x^2 - ay^2 - bz^2 = 0\). We claim that \(a\) is a square modulo \(p\). Indeed, otherwise,

\[
\begin{align*}
x^2 - ay^2 &\equiv 0 \mod p \implies x, y \equiv 0 \mod p \\
&\implies b \equiv x^2 - ay^2 \equiv 0 \mod p^2 \\
&\implies p | z \text{ (since } b \text{ is squarefree)},
\end{align*}
\]

contradicting the fact that \((x, y, z) \in \mathbb{Z}_p^3\) is primitive. Thus, \(a\) is a square modulo \(p\) for any prime divisor \(p\) of \(b\), and by the Chinese remainder theorem it follows that \(a\) is a square modulo \(b\). In other words, there exists \(t \in \mathbb{Z}\) such that

\[
t^2 - a = bb'
\]

for some \(b' \in \mathbb{Z}\), which we may assume is nonzero (since otherwise \(a\) is a square so \(x^2 - ay^2 - bz^2\) represents 0 over \(\mathbb{Q}\) by taking \((x, y, z) = (\sqrt{a}, 1, 0)\)). Moreover, by adding to \(t\) a suitable multiple of \(b\) we may assume that \(|t| \leq |b|/2\). Note that

\[
|b'| = \left| \frac{t^2 - a}{b} \right| \leq \frac{|t|^2}{|b|} + \frac{|a|}{|b|} \leq \frac{|b|}{4} + 1 < |b|
\]

where the last inequality follows from our assumption that \(|b| \geq 2\). Now, let \(p \leq \infty\). Since \(bb' = t^2 - a\) is a norm in \(\mathbb{Q}(\sqrt{a})\), this implies \(bb'\) is a norm in \(\mathbb{Q}_p(\sqrt{a})\). Since \(b\) is a norm in \(\mathbb{Q}(\sqrt{a})\) by our hypothesis on the quadratic form \(q\) and Lemma 5.14, it follows that \(b' = bb'/b\) is a norm in \(\mathbb{Q}_p(\sqrt{a})\). In other words, the quadratic form

\[
x^2 - ay^2 - b'z^2
\]

represents 0 in \(\mathbb{Q}_p\) for every \(p \leq \infty\). Since \(|a| + |b'| < |a| + |b|\), by our induction hypothesis it follows that \(x^2 - ay^2 - b'z^2\) represents 0 over \(\mathbb{Q}\), or in other words \(b'\) is a norm in \(\mathbb{Q}(\sqrt{a})\) by Lemma 14. It follows that \(b = bb'/b\) is a norm in \(\mathbb{Q}(\sqrt{a})\), and by Lemma 5.14 we conclude that \(x^2 - ay^2 - bz^2\) represents 0 over \(\mathbb{Q}\), as desired. This completes the induction and the proof. \(\square\)

6.2. Application to sums of three squares. We begin by recording the following corollary of the Hasse-Minkowski theorem.

**Corollary 6.1.** Let \(a \in \mathbb{Q}\). If \(q\) is a quadratic form on \(\mathbb{Q}\), then \(q\) represents \(a\) over \(\mathbb{Q}\) if and only if \(q\) represents \(a\) over \(\mathbb{Q}_p\) for all \(p \leq \infty\).

*Proof.* This was assigned as homework in Problem Set 2. \(\square\)

**Example 6.2.** A nonzero \(a \in \mathbb{Q}\) is represented by \(x^2 + y^2 + z^2\) over \(\mathbb{Q}\) if and only if \(a > 0\) and \(a\) is not of the form \(4^mu\) with \(m \in \mathbb{Z}\) and \(u \in 7 + 8\mathbb{Z}_2\).

*Proof.* Let \(a \in \mathbb{Q}^\infty\) be given. By the Hasse-Minkowski theorem (Corollary 6.1), \(x^2 + y^2 + z^2\) represents \(a\) over \(\mathbb{Q}\) if and only if it represents \(a\) over \(\mathbb{Q}_p\) for all \(p \leq \infty\).

By Proposition 4.8 (proved using the Chevalley-Warning theorem and Hensel’s lemma), we know that \(x^2 + y^2 + z^2\) represents 0 over \(\mathbb{Q}_p\) for every odd prime \(p\), and hence \(x^2 + y^2 + z^2\) represents \(a\) over all such \(\mathbb{Q}_p\) by Proposition 5.12. Next, it is clear that \(x^2 + y^2 + z^2\) represents \(a\) over \(\mathbb{R}\) if and only if \(a > 0\). It remains to verify that \(x^2 + y^2 + z^2\) represents \(a\) over \(\mathbb{Q}_2\) if and only if \(a\) is not of the form \(4^mu\) with \(m \in \mathbb{Z}\) and \(u \in 7 + 8\mathbb{Z}_2\).

Now, from homework in Problem Set 2 we know that an element \(u \in \mathbb{Z}_2^2\) is a square if and only if it belongs to \(1 + 8\mathbb{Z}_2\). This shows that the image of \(\mathbb{Z}_2^2 \times \mathbb{Z}_2 \times \mathbb{Z}_2\) under \(x^2 + y^2 + z^2\) is a union of cosets of \(8\mathbb{Z}_2\). Since the squares in \(\mathbb{Z}/8\) are precisely the congruence classes of \(\{0, 1, 2, 4\}\) and \(\{1 + a + b : a, b \in \{0, 1, 2, 4 \mod 8\}\}\), we see \(a \in \mathbb{Q}_2\) is of the form \(a = x^2 + y^2 + z^2\) with primitive \((x, y, z) \in \mathbb{Z}_2^3\) if and only if

\[
a \in \{1, 2, 3, 5, 6\} + 8\mathbb{Z}_2.
\]
Now, a nonzero \((x, y, z) \in \mathbb{Q}^3\) is obtained from a primitive triple \((x, y, z) \in \mathbb{Z}^3\) by multiplication with a power of 2, which multiplies the output of \(x^2 + y^2 + z^2\) by a power of 4. Thus, if \(a\) is represented by \(x^2 + y^2 + z^2\) over \(\mathbb{Q}_2\) then \(a\) cannot be of the form \(4^m u\) with \(u \in 7 + 8\mathbb{Z}_2\). Conversely, if \(a\) is not of the form \(4^m u\) with \(u \in 7 + 8\mathbb{Z}_2\), then we can write \(a = 4^m a'\) with \(n \in \mathbb{Z}\) and \(v_2(a') = 0\) or 1 (whence \(a' \equiv 1, 2, 3, 5, \) or 6 mod 8); our work so far shows that \(a'\) (and hence \(a\)) is represented by \(x^2 + y^2 + z^2\) over \(\mathbb{Q}_2\), as desired. This proves the result. \(\square\)

**Theorem 6.3 (Gauss).** A positive integer \(a\) is a sum of three integer squares if and only if \(a\) is not of the form \(4^m(8n + 7)\) for any \(m, n \in \mathbb{Z}_{\geq 0}\).

**Proof.** By the previous example, it suffices to prove the following:

**Claim 6.4.** If \(a \in \mathbb{Z}\), then \(x^2 + y^2 + z^2 = a\) has a solution in \(\mathbb{Q}\) if and only if it has a solution in \(\mathbb{Z}\).

Note the formal similarity between the above claim and the fact that \(x^2 + y^2 = a\) with \(a \in \mathbb{Z}\) is solvable in \(\mathbb{Z}\) if and only if it is solvable in \(\mathbb{Q}\) (see Problem Set 1). For the latter, the arithmetic of \(\mathbb{Z}[i]\) was useful; to draw analogy with this, we first give a sketch of the claim using the arithmetic of quaternions. Consider the quaternion algebra \(B = \mathbb{Q} \oplus \mathbb{Q} i \oplus \mathbb{Q} j \oplus \mathbb{Q} k\) with multiplication rules \(i^2 = j^2 = k^2 = ijk = -1\). Suppose \((x, y, z) \in \mathbb{Q}^3\) with \(x^2 + y^2 + z^2 = a\). If we set \(\alpha = xi + yj + zk\), then note that \(\alpha^2 + a = 0\), showing that \(\alpha\) is integral over \(\mathbb{Z}\). By arithmetic of quaternions, there exists some \(b \in B^x\) such that

\[
\beta b^{-1} = \alpha_0
\]

belongs to the maximal order (analogous to the ring of integers \(\mathbb{Z}[i]\) of \(\mathbb{Q}(i)\))

\[
\mathbb{Z} + \mathbb{Z}i + \mathbb{Z}j + \mathbb{Z}^{-1} + i + j + k = \frac{2}{2}.
\]

Since \(\text{tr}(\alpha_0) = \text{tr}(\alpha) = 0\), in fact \(\alpha_0 \in \mathbb{Z}i + \mathbb{Z}j + \mathbb{Z}k\) and writing

\[
\alpha_0 = x_0 i + y_0 j + z_0 k
\]

with \((x_0, y_0, z_0) \in \mathbb{Z}^3\) we have \(x_0^2 + y_0^2 + z_0^2 = a\), as desired.

In the remainder of this proof, we give a self-contained approach to the claim using elementary geometric arguments. (We remark that the same argument can be applied to the problem of solving \(x^2 + y^2 = a\) in \(\mathbb{Z}\) vs \(\mathbb{Q}\).) Suppose that

\[
(x, y, z) = \left(\frac{m_1}{d}, \frac{m_2}{d}, \frac{m_3}{d}\right) \in \mathbb{Q}^3
\]

is a rational solution of \(x^2 + y^2 + z^2 = a\) with \(m_i, d \in \mathbb{Z}\) and \(d > 0\). Let us write \(m_i = dq_i + r_i\) with \(q_i, r_i \in \mathbb{Z}\) and \(|r_i| \leq d/2\). If \(r_i = 0\) for all \(i\), then \((x, y, z)\) is integral and we are done, so let us assume \(r_i \neq 0\) for some \(i\). We shall denote \(\tilde{v} = (x, y, z)\) and

\[
\tilde{v} = \tilde{q} + \frac{1}{d} \tilde{r}
\]

where \(\tilde{q} = (q_1, q_2, q_3)\) and \(\tilde{r} = (r_1, r_2, r_3) \neq 0\). Now,

\[
\tilde{v} \cdot \tilde{v} = \left(\tilde{q} + \frac{1}{d} \tilde{r}\right) \cdot \left(\tilde{q} + \frac{1}{d} \tilde{r}\right) = \tilde{q} \cdot \tilde{q} + 2 \frac{\tilde{q} \cdot \tilde{r}}{d} + 1 \frac{\tilde{r} \cdot \tilde{r}}{d^2}.
\]

Multiplying both sides by \(d/(\tilde{r} \cdot \tilde{r})\), we get

\[
\frac{d(\tilde{v} \cdot \tilde{v})}{\tilde{r} \cdot \tilde{r}} = \frac{d(\tilde{q} \cdot \tilde{q})}{\tilde{r} \cdot \tilde{r}} + 2 \frac{\tilde{q} \cdot \tilde{r}}{\tilde{r} \cdot \tilde{r}} + \frac{1}{d}.
\]
Now, consider the vector $\vec{w} = \tau_r(\vec{v})$ obtained by reflection of $\vec{v}$ across the plane through the origin normal to $\vec{r}$. Since reflection is an isometry, we have $\vec{w} \cdot \vec{w} = \vec{v} \cdot \vec{v} = a$. Now, note that

$$\vec{w} = \tau_r(\vec{v}) = \tau_r\left( \vec{q} + \frac{1}{d} \vec{r} \right) = \vec{q} - 2 \frac{\vec{q} \cdot \vec{r}}{\vec{r} \cdot \vec{r}} \frac{1}{d} \vec{r},$$

$$= \vec{q} - \frac{d(\vec{v} \cdot \vec{v} - \vec{q} \cdot \vec{q})}{\vec{r} \cdot \vec{r}} \vec{r} = \vec{q} - \frac{\vec{v} \cdot \vec{v} - \vec{q} \cdot \vec{q}}{\vec{r} \cdot \vec{r}} \vec{r}.$$

Now, note that the denominator $\vec{r} \cdot \vec{r} / d$ is given by

$$\frac{\vec{r} \cdot \vec{r}}{d} = d(\vec{v} \cdot \vec{v} - \vec{q} \cdot \vec{q}) - 2\vec{q} \cdot \vec{r}$$

which is an integer. Moreover,

$$\frac{\vec{r} \cdot \vec{r}}{d} \leq \frac{r_1^2 + r_2^2 + r_3^2}{d} \leq 3 \frac{d^2}{4d} = \frac{3}{4} d,$$

so $\vec{w}$ has a smaller denominator than $\vec{v}$. By induction on the magnitude of the denominator $d$, we thus conclude that $x^2 + y^2 + z^2 = a$ has a solution in $\mathbb{Z}$.

**Corollary 6.5** (Lagrange). Every $a \in \mathbb{Z}_{\geq 0}$ is a sum of four squares.

**Proof.** If $a$ is a sum of three squares, then we are done by letting 0 be the fourth square. Otherwise, we have $a = 4^m(8n + 7)$ for some $m, n \in \mathbb{Z}_{\geq 0}$. Let us write $8n + 6$ as a sum of three squares. Then $8n + 7$ is a sum of four squares, and so is $a$. □

**Corollary 6.6** (Gauss). Every $a \in \mathbb{Z}_{\geq 0}$ is a sum of three triangular numbers (i.e. those of the form $m(m + 1)/2$ for $m \in \mathbb{Z}_{\geq 0}$).

**Proof.** Note that, if $x = 2m + 1$, then

$$\frac{x^2 - 1}{8} = \frac{m(m + 1)}{2}.$$

Now, let $x_1, x_2, x_3 \in \mathbb{Z}$ be such that $x_1^2 + x_2^2 + x_3^2 = 8a + 3$. Congruence modulo 4 shows that all of the $x_i$ must be odd, say $x_i = 2m_i + 1$. Then

$$\frac{m_1(m_1 + 1)}{2} = \frac{m_2(m_2 + 1)}{2} + \frac{m_3(m_3 + 1)}{2} = a. \quad \square$$

7. Lecture 7

**7.1. Field extensions.**

**Definition 7.1.** Let $L/k$ be an extension of fields.

1. The degree $[L : k]$ of $L/k$ is the dimension of $L$ as a vector space over $k$. (Note that if $M/L/k$ are field extensions then $[M : k] = [M : L][L : k]$.) We say that the extension $L/k$ is **finite** if $[L : k] < \infty$.

2. An element $\alpha \in L$ is **algebraic** over $k$ if there is a nonzero $f \in k[x]$ such that $f(\alpha) = 0$; otherwise, $\alpha$ is **transcendental** over $k$. We say that $L/k$ is **algebraic** if every $\alpha \in L$ is algebraic over $k$.

3. A subset $S \subseteq L$ is **algebraically independent** over $k$ if, for any choice of $s_1, \ldots, s_m \in k$ and nonzero $f \in k[x_1, \ldots, x_m]$, we have

$$f(s_1, \ldots, s_m) \neq 0.$$

A **transcendence basis** for $L/k$ is a set maximally algebraically independent over $k$.

**Note.** A set $S \subset L$ algebraically independent over $k$ is a transcendence basis for $L/k$ if and only if $L/k(S)$ is algebraic.
Proposition 7.2. Any two transcendence bases for a field extension $L/k$ have the same cardinality.

Proof. We shall prove this in the case where $L/k$ has a finite transcendence basis. Let us suppose $S = \{s_1, \ldots, s_m\}$ is a transcendence basis for $L/k$ with minimal cardinality. It suffices to show that, for every field extension $L/k$, there exists a transcendence basis $T = \{t_1, \ldots, t_n\} \subseteq L$ of minimal cardinality. Without loss of generality, we may assume $\mathbb{A}^n_k(L)$ is algebraically independent over $k(S_1)$ where $S_1 = \{t_1, s_2, \ldots, s_m\}$. Since $L/k(S_1)$ is algebraic, by minimality of $m$ we deduce that $S_1$ is a transcendence basis for $L/k$. Replacing $S$ and $T$ by $S_1$ and $T_1 = \{t_2, \ldots, t_n\}$, we repeat the above argument (using algebraic independence of any subset of $T$) to conclude that we must have $n \leq m$, as desired. □

Definition 7.3. The transcendence degree of an extension of fields $L/k$ is the cardinality of any (and hence every) transcendence basis for $L/k$.

7.2. Affine varieties and Hilbert Nullstellensatz. Let $k$ be a field.

Definition 7.4. The affine space $\mathbb{A}^n_k$ over $k$ is (given by) the assignment

$$\mathbb{A}^n_k(L) = L^n$$

for every field extension $L/k$.

For any $f \in k[x_1, \ldots, x_n]$ and $P \in \mathbb{A}^n_k(L)$, the value $f(P) \in L$ is well-defined. We say that $k[x_1, \ldots, x]$ is the coordinate ring of $\mathbb{A}^n_k$.

Definition 7.5. For any subset $S \subseteq k[x_1, \ldots, x_n]$, the associated (closed) affine variety $V_S \subseteq \mathbb{A}^n_k$ is (given by) the assignment

$$V_S(L) = \{P \in \mathbb{A}^n_k(L) : f(P) = 0 \text{ for all } f \in S\}$$

for every field extension $L/k$.

Note. We note the following.

1. If $(S) \subseteq k[x_1, \ldots, x_n]$ is the ideal generated by $S$, the $V_S = V((S))$.

Proof. Clearly, $V(S)(L) \subseteq V_S(L)$ for all field extensions $L/k$. Conversely, suppose $P \in V_S(L)$. If $f \in (S)$ is of the form $f = \sum_i g_i h_i$ with $g_i \in (S)$ and $h_i \in S$, then

$$f(P) = \sum_i g_i(P) h_i(P) = 0$$

which implies $P \in V(S)(L)$. □

2. We have $V(0) = \mathbb{A}^n_k$ and $V(1) = \emptyset$.

3. Given an affine variety $V \subseteq \mathbb{A}^n_k$ over $k$ and a morphism of fields $L_1 \to L_2$ over $k$, we have a natural map

$$V(L_1) \to V(L_2).$$

In fact $V$ defines a functor $V : \text{Field extensions}/k \to \text{Sets}$.

Definition 7.6. If $I \subseteq R$ is an ideal in a ring $R$, then

$$\sqrt{I} = \{f \in R : f^m \in I \text{ for some } m \geq 0\}$$

is an ideal of $R$, called the radical of $I$. We say that an ideal $I \subseteq R$ is radical if $I = \sqrt{I}$.

Note. We have the following.

1. We have $V_I = V_{\sqrt{I}}$ for every ideal $I \subseteq k[x_1, \ldots, x_n]$. 
(2) Let us choose an algebraic closure $\bar{k}$ of $k$. If $V \subseteq \mathbb{A}_{k}^{n}$ is an affine variety over $k$, then

$$I(V) := \{ f \in k[x_{1}, \ldots, x_{n}] : f(P) = 0 \text{ for all } P \in V(\bar{k}) \}$$

is a radical ideal in $k[x_{1}, \ldots, x_{n}]$. (Note $\bar{k}$.) We remark that $I(V)$ is independent of the choice of the algebraic closure $\bar{k}$.

**Theorem 7.7** (Hilbert Nullstellensatz). There is an order-reversing bijection

$$\{ \text{Radical ideals } I \subseteq k[x_{1}, \ldots, x_{n}] \} \leftrightarrow \{ \text{Affine varieties } V \subseteq \mathbb{A}_{k}^{n} \text{ over } k \}$$

given by $I \mapsto V_{I}$ with inverse $V \mapsto I(V)$.

**Proof.** Clearly, the map $I \mapsto V_{I}$ is surjective. To show that this map is injective, it suffices to show that if $I \subseteq k[x_{1}, \ldots, x_{n}]$ is an ideal then $I(V_{I}) = \sqrt{I}$. So let $I \subseteq k[x_{1}, \ldots, x_{n}]$ be an ideal. By the Hilbert basis theorem, the ideal $I$ is finitely generated, say $I = (f_{1}, \ldots, f_{m})$.

Suppose $f \in I(V_{I})$, i.e. $f(P) = 0$ for every $P \in \mathbb{A}_{k}^{n}(\bar{k})$ with $f_{1}(P) = \cdots = f_{m}(P) = 0$. (Here, we have fixed an algebraic closure $\bar{k}$ of $k$.) Then $f_{1}, \ldots, f_{m}, 1 - x_{0}f \in k[x_{0}, \ldots, x_{m}]$ (polynomials in $n + 1$ variables) have no common zero in $\bar{k}$. We now cite the following commutative algebra fact:

**Fact 7.8.** If $L/k$ is a field extension which is finitely generated as a $k$-algebra, then $L/k$ is algebraic.

Using this, we claim that $(f_{1}, \ldots, f_{m}, 1 - x_{0}f)$ is not contained in any maximal ideal of $k[x_{0}, \ldots, x_{n}]$ so it must be $(1)$. Indeed, suppose $\mathfrak{m}$ is a maximal ideal of $k[x_{0}, \ldots, x_{n}]$ containing $(f_{1}, \ldots, f_{m}, 1 - x_{0}f)$. Then since the quotient field

$$L = k[x_{0}, \ldots, x_{n}] / \mathfrak{m}$$

is finitely generated as a $k$-algebra it follows that $L/k$ is algebraic, and without loss of generality we may assume $L \subseteq \bar{k}$. But if $P(a_{0}, a_{1}, \ldots, a_{n}) \in L^{n+1}$ with each $a_{i}$ the image of $x_{i}$ in $L$ under the projection $k[x_{0}, \ldots, x_{n}] \to k[x_{0}, \ldots, x_{n}] / \mathfrak{m} = L$, then $f_{1}(a_{1}, \ldots, a_{n}) = \cdots = f_{m}(a_{1}, \ldots, a_{n}) = 1 - a_{0}f(a_{1}, \ldots, a_{n}) = 0$ since $(f_{1}, \ldots, f_{m}, 1 - x_{0}f) \subseteq \mathfrak{m}$, contradicting the fact that $f_{1}, \ldots, f_{m}, 1 - x_{0}f$ have no common zeros in $\bar{k}$. Thus there exist $g, g_{1}, \ldots, g_{m} \in k[x_{0}, x_{1}, \ldots, x_{n}]$ such that

$$1 = g(1 - x_{0}f) + \sum_{i=1}^{m} g_{i} f_{i}.$$ 

Consider the morphism $k[x_{0}, \ldots, x_{n}] \to k(x_{1}, \ldots, x_{n})$ given by $x_{0} = 1/f$ and $x_{1} \mapsto x_{1}, \ldots, x_{n} \mapsto x_{n}$. We have in $k(x_{1}, \ldots, x_{n})$

$$1 = \sum_{i=1}^{m} g_{i}(1/f, x_{1}, \ldots, x_{n}) f_{i}.$$ 

Finding common denominators for the $g_{i}$, we have

$$1 = \frac{\sum_{i} g_{i} f_{i}}{f^{r}}$$

for some $r \geq 0$ and $g_{i} \in k[x_{1}, \ldots, x_{n}]$, so $f^{r} = \sum_{i} g_{i} f_{i} \in I$. This shows that $f \in \sqrt{I}$, so $I(V_{I}) \subseteq \sqrt{I}$. The other direction is obvious, so we have $I(V_{I}) = \sqrt{I}$ and we are done. \qed

8. Lecture 8

Fix a field $k$ and algebraic closure $\bar{k}$. Recall from last time:

**Theorem 8.1** (Hilbert Nullstellensatz). There is an order-reversing bijection

$$\{ \text{Radical ideals } I \subseteq k[x_{1}, \ldots, x_{n}] \} \leftrightarrow \{ \text{Affine varieties } V \subseteq \mathbb{A}_{k}^{n} \text{ over } k \}$$

given by $I \mapsto V_{I}$ with inverse $V \mapsto I(V) = \{ f \in k[x_{1}, \ldots, x_{n}] : f(P) = 0 \text{ for all } P \in V(\bar{k}) \}$.

**Corollary 8.2.** If $V, W \subseteq \mathbb{A}^{n}$ are affine varieties over $k$ with $V(\bar{k}) = W(\bar{k}) \subseteq \mathbb{A}^{n}(\bar{k})$, then $V = W$. 

Proof. If \( V(\bar{k}) = W(\bar{k}) \), then \( I(V) = I(W) \) so \( V = V_I(V) = V_I(W) = W \). \( \square \)

**Note.** The above corollary can fail if \( \bar{k} \) is replaced by a smaller field. For example, in \( \mathbb{A}^2 \)
\[
V_{(x^2 + y^2 + 1)}(\mathbb{R}) = V_{(1)}(\mathbb{R}) = \emptyset
\]
but \( V_{(x^2 + y^2 + 1)}(\mathbb{C}) \neq \emptyset \) while \( V_{(1)}(\mathbb{C}) = \emptyset \), so \( V_{(x^2 + y^2 + 1)} \neq V_{(1)} \).

The corollary tells us that one can recover \( V \) from its set \( V(\bar{k}) \) of \( \bar{k} \)-points. Thus, as long as one is only dealing with \( \bar{L} \)-points of \( V \) where \( \bar{L} \) is a subfield of \( k \), one may define \( V \) directly as its set \( V(\bar{k}) \) of \( \bar{k} \)-points. This is the classical approach. In the remainder of this lecture, we will tacitly do this and often drop \( \bar{k} \) from notation. We will lean how to deal with \( V = V(\bar{k}) \) as a geometric object; this means we define the notion of (i) topology on \( V \) and (ii) functions on \( V \).

### 8.1. Zariski topology.

**Definition 8.3.** A *topology* on a set \( X \) is a collection \( \tau \subseteq \mathcal{P}(X) \) of subsets of \( X \), defined to be the open subsets, such that

1. \( X \) and \( \emptyset \) are open (i.e. \( \in \tau \)),
2. If \( U \) and \( V \) are open subsets of \( X \), then so is \( U \cap V \).
3. If \( \{U_i\}_{i \in I} \) is an arbitrary family of open sets, then so is \( \bigcup_{i \in I} U_i \).

A topological space is a pair \((X, \tau)\) given by a set \( X \) and a topology \( \tau \) on \( X \). A subset \( F \subseteq X \) of a topological space is closed if \( F = X \setminus U \) for some open subset \( U \subseteq X \).

**Remark.** One can equally define a topology on \( X \) in terms of closed sets instead of open sets, with complementary axioms (i.e. \( X \) and \( \emptyset \) are closed; a union of two closed sets is closed; an arbitrary intersection of closed sets is closed).

**Example 8.4.** We have the following examples.

1. If \( X \) is any set, the topology \( \tau = \mathcal{P}(X) \) is called the *discrete topology* on \( X \), and the topology \( \tau = \{\emptyset, X\} \) is called the *trivial topology* on \( X \).
2. If \( X = \mathbb{R}^n \), let us say that \( U \subseteq \mathbb{R}^n \) is open if for every \( p \in U \) there is an open ball \( B(p, r) \) of some positive radius \( r > 0 \) centered at \( p \) such that \( B(p, r) \subseteq U \). This defines a topology on \( \mathbb{R}^n \), called the *Euclidean topology*.

**Proposition 8.5.** We have the following.

1. \( \mathbb{A}^n = V_{(0)} \) and \( \emptyset = V_{(1)} \).
2. If \( I, J \subseteq k[x_1, \ldots, x_n] \), then \( V_I \cup V_J = V_{IJ} \) where \( IJ = \{\sum_i f_i g_i : f_i \in I, g_i \in J\} \).
3. If \( \{I_\alpha : \alpha \in A\} \) is an arbitrary collection of ideals in \( k[x_1, \ldots, x_n] \), then
\[
\bigcap_{\alpha \in A} V_{I_\alpha} = V_{(\bigcup_{\alpha \in A} I_\alpha)}
\]

**Proof.** (1) is clear. For (2), note first that clearly \( V_I \cup V_J \subseteq V_{IJ} \) and hence \( V_I \cup V_J \subseteq V_{IJ} \). Conversely, suppose \( P \in V_{IJ} \). Suppose \( P \notin V_I \), so we must show \( P \notin V_J \). Since \( P \notin V_I \), there exists \( f \in I \) such that \( f(P) \neq 0 \). Now, for any \( g \in J \) we must have \( f(P)g(P) = 0 \) since \( fg \in IJ \) and \( P \in V_{IJ} \). This shows that \( g(P) = 0 \) for every \( g \in J \), i.e. \( P \notin V_J \), as desired. Finally, if \( \{I_i\}_{i \in A} \) is an arbitrary collection of ideals in \( k[x_1, \ldots, x_n] \), then
\[
\bigcap_{i \in A} V_{I_i} = \{P \in \mathbb{A}^n : f(P) = 0 \text{ for all } f \in I_i \text{ for all } i \in A\} = V_{(\bigcup_{i \in A} I_i)}.
\]

**Definition 8.6.** We define the following.

1. The *Zariski topology* on \( \mathbb{A}^n \) over \( k \) is the topology where the Zariski closed sets are precisely the closed subvarieties \( V \subseteq \mathbb{A}^n \) over \( k \).
(2) The Zariski topology on an affine variety $V \subseteq \mathbb{A}^n$ over $k$ is the subspace topology induced by the Zariski topology on $\mathbb{A}^n$. Equivalently, the Zariski close subsets of $V$ are precisely the closed subvarieties $W \subseteq V \subseteq \mathbb{A}^n$.

**Note.** If $L \subseteq \bar{k}$ is a field extension of $k$ and $V \subseteq \mathbb{A}^n$ is an affine variety over $k$, then the Zariski topology $\tau_L$ on $V$ over $L$ is finer than the Zariski topology $\tau_k$ on $V$ over $k$, i.e. $\tau_k \subseteq \tau_L$. For example, a point $P \in V(\bar{k})$ is closed (i.e. Zariski closure of $\{P\}$ is itself) under the Zariski topology over $k$ if and only if $P \in V(k)$.

**Proposition 8.7.** The affine space $\mathbb{A}^n$ over $k$ is quasicompact with respect to the Zariski topology over $k$, i.e. every open cover has a finite subcover.

**Proof.** Reformulating in terms of closed sets, we need to show that if $\{V_\alpha : \alpha \in A\}$ is a collection of affine varieties $V_\alpha \subseteq \mathbb{A}^n$ over $k$ with $\bigcap_\alpha V_\alpha = \emptyset$ then there is a finite subset $\{\alpha_1, \ldots, \alpha_r\} \subseteq A$ with $\bigcap_{i=1}^r V_{\alpha_i} = \emptyset$. Now, letting $I_\alpha = I(V_\alpha)$ we see that $\bigcap_\alpha V_\alpha = \emptyset$ implies $(\bigcup_{\alpha \in A} I_\alpha) = (1)$ by Hilbert Nullstellensatz. In particular, there exists a finite set $\{f_1, \ldots, f_r\}$ of elements in $\bigcup_{\alpha \in A} I_\alpha$ such that $(f_1, \ldots, f_r) = (1)$. Choosing $\alpha_i$ so that $f_i \in I_{\alpha_i}$, we have $\bigcap_{i=1}^r V_{\alpha_i} = \emptyset$. \qed

8.2. Coordinate ring.

**Definition 8.8.** The coordinate ring of an affine variety $V \subseteq \mathbb{A}^n$ over $k$ is the $k$-algebra

$$k[V] = k[x_1, \ldots, x_n]/I(V).$$

An element of $f \in k[V]$ is called a regular function on $V$ over $k$.

**Note.** We have the following.

1. The coordinate ring $k[V]$ of an affine variety $V \subseteq \mathbb{A}^n_k$ is reduced, i.e. satisfies $\sqrt{(0)} = (0)$.
2. Hilbert’s Nullstellensatz shows that $k[V]$ along with its presentation as a quotient of $k[x_1, \ldots, x_n]$ determines the affine variety $V \subseteq \mathbb{A}^n_k$. In fact, we have the following.

**Proposition 8.9.** If $V \subseteq \mathbb{A}^n_k$ is an affine variety over $k$, we have a natural bijection

$$\text{Hom}_{k-alg}(k[V], L) \overset{\sim}{\to} V(L)$$

for every field extension $L/k$, given by $\varphi \mapsto (\varphi(x_1), \ldots, \varphi(x_n)) \in \mathbb{A}^n_k(L)$.

**Proof.** Given $P = (a_1, \ldots, a_n) \in V(L)$, there is a unique ring homomorphism $\varphi : k[x_1, \ldots, x_n] \to L$ given by $\varphi(x_i) = a_i$, so it suffices to show that $\varphi$ factors through $k[V]$, i.e. maps $I(V)$ to zero. Since $V = V_1(V)$ by Hilbert Nullstellensatz, this means that $\varphi(f) = f(\varphi(x_1), \ldots, \varphi(x_n)) = f(a_1, \ldots, a_n) = 0$ for every $f \in I(V)$, and therefore $\varphi$ factors through $k[V]$ as desired. \qed

8.3. Irreducibility and dimension.

**Definition 8.10.** A nonempty topological space $X$ is irreducible if, whenever $X = F_1 \cup F_2$ where $F_1$ and $F_2$ are closed subsets of $X$, we must have $F_1 = X$ or $F_2 = X$.

**Proposition 8.11.** Let $V \subseteq \mathbb{A}^n$ be an affine variety over $k$. The following are equivalent:

1. The set $V(\bar{k})$ is irreducible with respect to the Zariski topology over $k$.
2. The ring $k[V]$ is an integral domain.
3. The ideal $I(V) \subseteq k[x_1, \ldots, x_n]$ is prime.

We say that $V/k$ is irreducible if one of the above conditions holds.

**Proof.** (2) $\iff$ (3) is elementary. We will show (1) $\iff$ (3). Suppose (3) holds, and let $W_1$ and $W_2$ be closed affine subvarieties of $V$ with $V = W_1 \cup W_2$. This means that $I(V) = \sqrt{I(W_1)I(W_2)}$ and in particular $I(W_1)I(W_2) \subseteq I(V)$. If $I(W_1) \subseteq I(V)$, then $W_1 \supseteq V$ and hence $W_1 = V$. So suppose there exists $f \in I(W_1)$ with $f \not\in I(V)$. Then for any $g \in I(W_2)$ we have $fg \in I(W_1)I(W_2) \subseteq I(V)$.
and therefore \( g \in I(V) \) since \( I(V) \) is prime. This shows \( I(W_2) \subseteq I(V) \), we deduce that \( V = W_2 \). This being true for any closed subvarieties \( W_i \) with \( W_1 \cup W_2 = V \), we conclude that \( V \) is irreducible.

Conversely, suppose (1) holds, and let \( f, g \in \mathbb{k}[x_1, \ldots, x_n] \) with \( fg \in I(V) \). Let \( I_1 = I(V) + (f) \) and \( I_2 = I(V) + (g) \). Then \( I_1 \cap I_2 \subseteq I(V) \), and we see that \( V_{I_1} \) and \( V_{I_2} \) are subvarieties of \( V \) satisfying \( V_{I_1} \cup V_{I_2} = V \). By irreducibility of \( V(\bar{k}) \) over \( \mathbb{k} \), we must have \( V_{I_1} = V \) or \( V_{I_2} = V \). Without loss of generality, say \( V_{I_1} = V \). Then \( \sqrt{I(V) + (f)} = \sqrt{I_1} = I(V) \). In particular, \( f \in I(V) \). This being true for any \( f, g \) with \( fg \in I(V) \), we deduce that \( I(V) \) is prime. \( \square \)

**Definition 8.12.** In general, an affine variety \( V \subseteq \mathbb{A}^n \) over \( \mathbb{k} \) is a finite union of irreducible affine varieties over \( \mathbb{k} \). We can write

\[
V = V_1 \cup \cdots \cup V_r
\]

where each \( V_i \) is an irreducible affine variety over \( \mathbb{k} \) with the property that each \( V_i \) is maximal with respect to inclusion among irreducible closed subvarieties of \( V \); then each \( V_i \) is called an irreducible component of \( V \).

**Example 8.13.** \( V_{(xy)} \subseteq \mathbb{A}^2 \) is not irreducible, since it is a union of proper closed subsets \( V(x) \) and \( V(y) \). Each of these latter is an irreducible component of \( V_{(xy)} \).

**Definition 8.14.** The dimension \( \dim X \) of a topological space \( X \) is the supremum of all \( n \) such that there is a chain

\[
Z_0 \subsetneq Z_1 \subsetneq \cdots \subsetneq Z_n
\]

of distinct irreducible closed subsets of \( X \). The dimension of an affine variety \( V/k \), denoted \( \dim V/k \), is the dimension of \( V(\bar{k}) \) equipped with the Zariski topology over \( k \).

**Definition 8.15.** Let \( V \subseteq \mathbb{A}^n \) be an irreducible variety over \( \mathbb{k} \). The function field of \( V/k \) is

\[
k(V) = \text{Frac} k[V].
\]

**Fact 8.16.** Suppose that \( V \subseteq \mathbb{A}^n \) is an irreducible affine variety over \( \mathbb{k} \). We have the following:

1. We have \( \dim V/k = \text{trdeg}_k k(V) \).
2. If \( f \in k[V] \) is a nonzero element which is not a unit, then every irreducible component of \( V_f = \{ P \in V : f(P) = 0 \} \subseteq V \) has dimension equal to \( \dim(V) - 1 \).

9. Lecture 9

9.1. Projective spaces and projective varieties. Let \( k \) be a field.

**Definition 9.1.** The projective \( n \)-space \( \mathbb{P}_k^n \) over \( k \) is (given by) the assignment

\[
\mathbb{P}_k^n(L) = \frac{L^{n+1} \setminus \{(0, \ldots, 0)\}}{\sim}
\]

with equivalence relation \( \sim \) given by \( (a_0, \ldots, a_n) \sim (b_0, \ldots, b_n) \) iff \( (a_0, \ldots, a_n) = (\lambda b_0, \ldots, \lambda b_n) \) for some \( \lambda \in L \). The class of nonzero \( (a_0, \ldots, a_n) \in L^{n+1} \) in \( \mathbb{P}_k^n(L) \) is denoted \( (a_0 : \cdots : a_n) \).

If \( f \in k[X_0, \ldots, X_n] \) is homogeneous, the condition \( f(P) = 0 \) makes sense independently of choice of representative for \( P \).

**Definition 9.2.** Given any subset \( S \subseteq k[X_0, \ldots, X_n] \) consisting of homogeneous elements, the associated (closed) projective variety \( \mathbb{V}_S \subseteq \mathbb{P}_k^n \) over \( k \) is (given by) the assignment

\[
\mathbb{V}_S(L) = \{ P \in \mathbb{P}_k^n(L) : f(P) = 0 \text{ for all } f \in S \}
\]

for every field extension \( L/k \).

**Definition 9.3.** An ideal \( I \subseteq k[X_0, \ldots, X_n] \) is homogeneous if it has a generating set consisting of homogeneous polynomials.
Note. We note the following.

1. If $S \subseteq k[X_0, \ldots, X_n]$ is a set consisting of homogeneous elements, then $V_S = V(S)$.
2. The ideal $m = (X_0, \ldots, X_n) \subseteq k[X_0, \ldots, X_n]$ is called the irrelevant ideal. We have $V_m = \emptyset$.
3. If $I \subseteq k[X_0, \ldots, X_n]$ is homogeneous, then its radical $\sqrt{I}$ is homogeneous. (This will be an exercise in Problem Set 3.) Moreover, $V_I = V(\sqrt{I})$.

4. Fix an algebraic closure $\overline{k}$ of $k$. Given a projective variety $V \subseteq \mathbb{P}^n$ over $k$, let $I(V)$ be the ideal generated by the homogeneous polynomials $f \in k[X_0, \ldots, X_n]$ satisfying $f(P) = 0$ for all $P \in V(\overline{k})$. Then $I(V)$ is a homogeneous radical ideal of $k[X_0, \ldots, X_n]$.

**Theorem 9.4 (Projective nullstellensatz).** We have an inclusion-reversing bijection

$$
\{\text{Radical homogeneous ideals } I \subseteq k[X_0, \ldots, X_n] \text{ not containing } (X_0, \ldots, X_n) \}
\leftrightarrow \{\text{Projective varieties } V \subseteq \mathbb{P}^n \text{ over } k\}
$$

given by $I \mapsto V_I$ with inverse $V \mapsto I(V)$.

As in the case of affine varieties, we can deduce from the Nullstellensatz that a projective variety $V \subseteq \mathbb{P}^n$ over $k$ is determined by its set $V(\overline{k})$ of $\overline{k}$-points, and we may identify $V$ with the latter. We define the Zariski topology on $\mathbb{P}^n$ and projective varieties in the same way as we did for affine varieties (Zariski closed subsets are given by closed subvarieties).

**Definition 9.5.** The homogeneous coordinate ring of a projective variety $V \subseteq \mathbb{P}^n$ over $k$ is the graded $k$-algebra

$$S_V = k[X_0, \ldots, X_n]/I(V).$$

**Proposition 9.6.** Let $V \subseteq \mathbb{P}^n$ be a projective variety over $k$. The following are equivalent:

1. $V(\overline{k})$ is irreducible under the Zariski topology over $k$.
2. The ideal $I(V)$ is a prime ideal.
3. The homogeneous coordinate ring $S_V$ is an integral domain.

**Proof.** The proof is the same as for affine varieties. \qed

9.2. **Affine patches and projective closures.** Given $\mathbb{P}^n$, the standard hyperplane $H_i \subset \mathbb{P}^n$ is given by $H_i = \mathbb{V}(X_i)$, i.e.

$$H_i(L) = \{(a_0 : \cdots : a_n) \in \mathbb{P}^n(L) : a_i = 0\} \text{ for every field extension } L/k.$$

The open complement $U_i = \mathbb{P}^n \setminus H_i$, consisting of points of the form

$$(x_0 : \cdots : x_{i-1} : 1 : x_{i+1} : \cdots : x_n),$$

is a copy of $\mathbb{A}^n$. Let us write $k[U_i] = k[x_0, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n]$ for the coordinate ring of $U_i \simeq \mathbb{A}^n$.

**Definition 9.7.** Fix $i \in \{0, \ldots, n\}$. Given $f \in k[x_0, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n]$ of total degree $d$, its homogenization is

$$F(X_0, \ldots, X_n) = X_i^d f \left( \frac{X_0}{X_i}, \ldots, \frac{X_{i-1}}{X_i}, \frac{X_{i+1}}{X_i}, \ldots, \frac{X_n}{X_i} \right) \in k[X_0, \ldots, X_n].$$

Given a homogeneous polynomial $F \in k[X_0, \ldots, X_n]$, its dehomogenization (with respect to $X_i$) is

$$f(x_0, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) = F(x_0, \ldots, x_{i-1}, 1, x_{i+1}, \ldots, x_n) \in k[U_i].$$

**Definition 9.8.** Fix $i \in \{0, \ldots, n\}$. We make the following definitions.

1. Let $V \subseteq \mathbb{P}^n$ be a projective variety over $k$. The $i$-th affine patch of $V$ is the affine variety given as the intersection $V \cap U_i$.
2. Let $V \subseteq \mathbb{A}^n = U_i \subset \mathbb{P}^n$ be an affine variety. The projective closure of $V$ in $\mathbb{P}^n$ is the smallest projective variety $\overline{V} \subseteq \mathbb{P}^n$ over $k$ containing $V$. 

24
Proposition 9.9. Fix \( i \in \{0, \ldots, n\} \). We have the following.

1. Let \( I = I(V) \leq k[X_0, \ldots, X_n] \) be the homogeneous ideal of a projective variety \( V \subseteq \mathbb{P}^n \) over \( k \). Then the ideal \( I(V \cap U_i) \) of the \( i \)-th patch \( V \cap U_i \) is the ideal \( I_i \) generated by the dehomogenizations of the elements of \( I \) with respect to \( X_i \).

2. Let \( I = I(V) \leq k[U_i] \) be the ideal of an affine variety \( V \subseteq \mathbb{A}^n = U_i \) over \( k \). Then the ideal \( I(\overline{V}) \) of the projective closure \( \overline{V} \) in \( \mathbb{P}^n \) is the ideal \( I^h \) generated by the homogenizations of the elements of \( I \).

Proof. (1) Suppose \( P = (a_0, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n) \in \mathbb{P}^n(k) \). Viewing \( P \) as a point of \( \mathbb{P}^n \), we have \( P = (a_0 : \cdots : a_{i-1} : 1 : a_{i+1} : \cdots : a_n) \) in homogeneous coordinates. Clearly, by definition of dehomogenization we have \( P \in V \) in \( \mathbb{P}^n \) if and only if \( f(P) = 0 \) on \( \mathbb{A}^n = U_i \) for the dehomogenization \( f \) of any homogeneous polynomial in \( I \). The latter is equivalent to saying that \( P \in V(I_i) \) on \( U_i \). It thus only remains to show that the ideal \( I_i \) is radical.

So suppose \( f \in k[U_i] \) satisfies \( f^m \in I_i \) for some \( m \geq 0 \). Then \( f^m = \sum_{i=1}^r h_ig_i \) for some \( h_i \in k[U_i] \) and \( g_i \) the dehomogenization of some homogeneous \( G_i \in I \). Let \( N = \max\left\{ m \deg f, \deg h_1 + \deg g_1, \ldots, \deg h_r + \deg g_r \right\} \). Then in \( k[X_0, \ldots, X_n] \) we have

\[
X_i^{N-m \deg f} f^m = \sum_i X_i^{N-\deg h_i - \deg g_i} H_i G_i
\]

where \( F \) and \( H_i \) are homogenizations of \( f \) and \( h_i \), etc. This shows \( X_i^{N-m \deg f} f^m \in I \) and multiplying by \( X_i^{(N-m \deg f)(m-1)} \) gives us

\[
(X_i^{N-m \deg f} F)^m \in I \implies X_i^{N-m \deg f} F \in I
\]

since \( I = I(V) \) is radical. The dehomogenization of \( X_i^{N-m \deg f} F \) is \( f \), so \( f \in I_i \) as desired.

(2) Note that the dehomogenization of every \( F \in I^h \) with respect to \( X_i \) must vanish on \( V \). By minimality of \( \overline{V} \), it then follows that \( I^h \subseteq I(V) \). Conversely, if \( F \in I(\overline{V}) \) then its dehomogenization \( f \) belongs to \( I \) by part (1) and the fact that we must have \( \overline{V} \cap U_i = V \) by minimality of \( \overline{V} \). Then the homogenization \( F_0 \in I^h \) of \( f \) satisfies \( F = X_i^{\deg F - \deg F_0} F_0 \), so \( F \in I^h \).

Example 9.10. The projective closure of the affine plane curve (with coordinates \( x = X/Z, y = Y/Z \))

\[
y^2 = x(x-1)(x-3)
\]

in \( \mathbb{P}^2 \) (with homogeneous coordinates \( X, Y, Z \)) is given by

\[
Y^2Z = X(X-Z)(X-3Z).
\]

Definition 9.11. Let \( V \subset \mathbb{P}^n \) be a projective variety over \( k \).

1. The dimension of \( V/k \) is the maximum of the dimensions of its affine patches.

2. The function field \( k(V) \) of \( V/k \) is the function field of any one of its affine patches. (It turns out that they are all isomorphic.)

10. Lecture 10

10.1. Remark on irreducibility. Let \( V \) be a variety over a field \( k \) with algebraic closure \( \bar{k} \). Recall that \( V/k \) is irreducible if and only if \( V(\bar{k}) \) is irreducible with respect to the Zariski topology over \( k \). We note that the notion of irreducibility depends on the base field \( k \).
Example 10.1. Let $V = V(x^2 - 2y^2) \subseteq \mathbb{A}^2$ defined over $\mathbb{Q}$. Since $x^2 - 2y^2$ is irreducible in $\mathbb{Q}[x, y]$, it follows that $V/\mathbb{Q}$ is irreducible. On the other hand, since $x^2 - 2y^2 = (x - \sqrt{2}y)(x + \sqrt{2}y)$ in $\mathbb{Q}(\sqrt{2})[x, y]$ we see that

$$V = V(x - \sqrt{2}y) \cup V(x + \sqrt{2}y)$$

is a union of two irreducible components.

Definition 10.2. An algebraic variety over a field $k$ is geometrically irreducible if it is irreducible when viewed as an algebraic variety over $\overline{k}$.

10.2. Hypersurfaces. Let $k$ be a field.

Definition 10.3. A hypersurface in $\mathbb{A}^n$ (resp. $\mathbb{P}^n$) over a field $k$ is the zero set of a nonzero nonunit $f \in k[x_1, \ldots, x_n]$ (resp. nonzero homogeneous $F \in k[X_0, \ldots, X_n]$).

By Fact 8.16, a hypersurface in $\mathbb{A}^n$ or $\mathbb{P}^n$ has dimension $n-1$. Conversely:

Proposition 10.4. Every irreducible subvariety $V \subseteq \mathbb{A}^n$ or $\mathbb{P}^n$ of dimension $n-1$ is a hypersurface.

Proof. We will prove this for $V \subseteq \mathbb{A}^n$ of dimension $n-1$. We will use the fact that $k[x_1, \ldots, x_n]$ is a unique factorization domain. Let $p \subseteq k[x_1, \ldots, x_n]$ be a prime ideal such that $V = V_p$. Since $p \neq (0)$ (otherwise $V = \mathbb{A}^n$ has dimension $n$), we have $f \in p$ for some nonzero nonunit $f$. Let

$$f = g_1 \cdots g_r$$

with each $g_i \in k[x_1, \ldots, x_n]$ irreducible. Then we have $g_1 \cdots g_r = f \in p$ showing that $g_i \in p$ for some $i$, since $p$ is prime. Replacing $f$ by this $g_i$, we may assume without loss of generality that $f$ is irreducible. Now, $V_f$ is an irreducible proper closed subvariety of $\mathbb{A}^n$. We claim that $V = V_f$. Indeed, otherwise, since dim $V = n-1$, there is a chain

$$Z_0 \subset Z_1 \subset \cdots \subset Z_{n-1} = V \not\subseteq V_f \not\subseteq \mathbb{A}^n$$

of irreducible closed subvarieties of $\mathbb{A}^n$, showing that dim $\mathbb{A}^n \geq n+1$, a contradiction. It follows that $V = V_f$, as desired.

10.3. Tangent spaces and smoothness. We begin by providing motivation for our definitions. Given an open domain $U \subseteq \mathbb{R}^n$ and a smooth function $f : U \to \mathbb{R}$, the gradient or differential of $f$ at $P \in U$ is

$$\nabla f(P) = \left( \frac{\partial f}{\partial x_1}(P), \ldots, \frac{\partial f}{\partial x_n}(P) \right).$$

Recall that $\nabla f(P)$ provides a first-order approximation of $f$ around $P$ in that, by Taylor’s theorem,

$$f(Q) = f(P) + \nabla f(P) \cdot (Q - P) + \text{(higher order terms)}.$$

Given a tangent vector $\tilde{v}$, the directional derivative of $f$ at $P$ along $\tilde{v}$ is given by

$$\nabla_{\tilde{v}} f(P) = (\nabla f(P)) \cdot \tilde{v} \in \mathbb{R}.$$

In particular, intuitively, $\tilde{v}$ provides a functional on the space of functions over a “first-order” neighborhood of $P$. We make the following definition.

Definition 10.5. Let $V \subseteq \mathbb{A}^n$ be an affine variety over an algebraically closed field $k$. Let $P \in V(\overline{k})$ be a point, with corresponding maximal ideal $m_P \subseteq k[V]$. Let

- The differential of $f \in k[V]$ at $P$ is the element $df_P := f - f(P) \in m_P/m_P^2$.
- The cotangent space of $V$ at $P$ is the $k$-vector space $T_P^*V := m_P/m_P^2$.
- The tangent space of $V$ at $P$ is the dual $k$-vector space $T_PV := (m_P/m_P^2)^*$.  

Lemma 10.6. Let $P \in V(k)$ where $V \subseteq \mathbb{A}^n$ is an affine variety over $k = \overline{k}$.

1. $d_P(c) = 0$ for any $c \in k$.  

26
(2) \( d_P(fg) = d_P(f) \cdot g(P) + f(P) \cdot d_P(g) \) for every \( f, g \in k[V] \). (Leibniz rule)

Proof. (1) is clear, so it remains to prove (2). To prove (2), note that

\[
d_P(fg) = f g - f(P)g(P) = (f - f(P))(g - g(P)) + f(P)(g - g(P)) + (f - f_P)g(P)
\]

\[
= d_P(f) \cdot g(P) + f(P) \cdot d_P(g).
\]

\[\square\]

**Fact 10.7.** Let \( V \subseteq \mathbb{A}^n \) be an affine variety over \( k = \overline{k} \). If \( P \in V(k) \) is a point, then

\[
\dim_k m_P/m_P^2 \geq \max \{ \dim V_i \}
\]

where \( V_i \) runs over all irreducible components of \( V \) containing \( V \).

**Definition 10.8.** Let \( k \) be a field and \( \overline{k} \) its algebraic closure. Let \( V/\overline{k} \) be an affine variety.

- A point \( P \in V(\overline{k}) \) is a *singular point* if \( \dim T_P V > \max \{ \dim V_i \} \) where \( V_i \) runs over all irreducible components of \( V \) containing \( P \), and a *regular point* otherwise.
- \( V/\overline{k} \) is said to be *smooth* if every \( P \in V(\overline{k}) \) is a regular point.

More generally, we will say that an affine variety \( V/k \) is *smooth* if it is smooth when viewed as a variety over \( k \). A projective variety \( V/k \) is *smooth* if each of its affine patches is smooth.

**Example 10.9.** Let \( V = V(xy) \subset \mathbb{A}^2 \) defined over \( \overline{Q} \).

- At \( P = (0, 0) \in V(\overline{Q}) \), we have \( m_P = (x, y) \mod (xy) \) and \( m_P^2 = (x^2, xy, y^2) \mod (xy) \). A brief computation shows that we have

\[
m_P/m_P^2 = \overline{Q}x \oplus \overline{Q}y
\]

showing that \( \dim T_P V = 2 > 1 = \dim V \), and hence \( P \) is a singular point of \( V \).

- At \( Q = (1, 0) \in V(\overline{Q}) \), we have \( m_Q = (x - 1, y) \mod (xy) \) and

\[
m_Q^2 = ((x - 1)^2, (x - 1)y, y^2) \mod (xy) = ((x - 1)^2, y) \mod (xy).
\]

We have \( m_Q/m_Q^2 = (x - 1)\overline{Q} \), so \( \dim T_Q V = 1 = \dim V \) and \( Q \) is a regular point of \( V \).

**Proposition 10.10** (Jacobian criterion). Let \( V = V(f_1, \ldots, f_m) \subset \mathbb{A}^n \) be an affine variety over \( K \). A point \( P \in V(\overline{k}) \) is a regular point of \( V \) if and only if

\[
\left( \frac{\partial f_i}{\partial x_i}(P) \right) \in M_{m \times n}(\overline{k})
\]

has rank \( n - \max \{ \dim V_i \} \) where \( V_i \) runs over all irreducible components of \( V \) containing \( P \).

**Example 10.11.** Let \( C \subset \mathbb{A}^2 \) be given by \( f = y^2 - x^3 - x^2 \). Then

\[
\left( \frac{\partial f}{\partial x} , \frac{\partial f}{\partial y} \right) = (-3 - 2x, 2y)
\]

so \( f(P) = \frac{\partial f}{\partial x}(P) = \frac{\partial f}{\partial y}(P) = 0 \) if and only if \( P = (0, 0) \). It follows that \( C \) is singular with unique singular point \((0, 0)\).

We end with the following result (in view of Question 3.(e) in Problem Set 2).

**Definition 10.12.** A *curve* is a 1-dimensional algebraic variety.

**Theorem 10.13** (Faltings). Let \( C \subset \mathbb{P}^2 \) be a smooth projective geometrically irreducible curve defined over a number field \( K \). If the genus of the curve (given by \( g := (d - 1)(d - 2)/2 \) where \( d \) is the degree of a polynomial \( F \in K[X_0, X_1, X_2] \) with \( C = V_F \)) is at least \( 2 \), then \( C(K) \) is finite.
11. Lecture 11

11.1. Morphisms of affine varieties. Let $k$ be a field.

**Definition 11.1.** Let $X \subseteq \mathbb{A}^n$, $Y \subseteq \mathbb{A}^m$ be affine varieties over $k$. A morphism $f : X \to Y$ is given by $f = (f_1, \ldots, f_m) \in k[X]^m$ such that

$$(f_1(P), \ldots, f_m(P)) \in Y(L)$$

for all field extensions $L/k$ and for all $P \in X(L)$. Given morphisms $f : X \to Y$ and $g : Y \to Z \subseteq \mathbb{A}^l$ of affine varieties over $k$, their composition $g \circ f : X \to Z$ is given by

$$g \circ f = (\tilde{g}_1(f_1, \ldots, f_m), \ldots, \tilde{g}_l(f_1, \ldots, f_m)) \in k[X]^l$$

where $\tilde{g}_i \in k[y_1, \ldots, y_m]$ is any representative $g_i \in k[Y] = k[y_1, \ldots, y_m]/I(Y)$. (It can be shown that $f \circ g$ is independent of the choices of the $\tilde{g}_i$s and we may write $g_i(f_1, \ldots, f_m) = \tilde{g}_i(f_1, \ldots, f_m) \in k[X].$)

**Proposition 11.2.** Given affine varieties $X \subseteq \mathbb{A}^n$ and $Y \subseteq \mathbb{A}^m$ over $k$, there is a bijection

$$k\text{-algebra homomorphisms}$$

$$\{ \varphi : k[Y] = k[y_1, \ldots, y_m]/I(Y) \to k[X] \} \simeq \{ \text{morphisms } f : X \to Y \}$$

given by $\varphi \mapsto f = (\varphi(y_1), \ldots, \varphi(y_m))$.

**Proof.** We have the chain of equalities and bijections

$$\{ f : X \to Y \} = \{(f_1, \ldots, f_m) \in k[X]^m : (f_1(P), \ldots, f_m(P)) \in Y \forall P \in X \}$$

$$= \{(f_1, \ldots, f_m) \in k[X]^m : g(f_1(P), \ldots, f_m(P)) = 0 \forall g \in I(Y) \forall P \in X \}$$

$$= \{(f_1, \ldots, f_m) \in k[X]^m : g(f_1, \ldots, f_m) = 0 \forall g \in I(Y) \}$$

$$\leftrightarrow \{ \varphi : k[y_1, \ldots, y_m] \to k[X] : \varphi(g) = g(\varphi(y_1), \ldots, \varphi(y_m)) = 0 \forall g \in I(Y) \}$$

$$\leftrightarrow \{ \varphi : k[Y] \to k[X] \}. \quad \Box$$

**Corollary 11.3.** Let $X \subseteq \mathbb{A}^n$ and $Y \subseteq \mathbb{A}^m$ be as above.

1. $k[X]$ can be identified with the set of morphisms $X \to \mathbb{A}^1$ over $k$.
2. Under the identification in part (1), given a morphism $f : X \to Y$ the associated $k$-algebra homomorphism $f^* = \varphi : k[Y] \to k[X]$ is just

$$k[Y]\{ \text{morphisms } Y \to \mathbb{A}^1 \} \to \{ \text{morphisms } X \to \mathbb{A}^1 \} = k[X]$$

given by $(g : Y \to \mathbb{A}^1) \mapsto (f^*(g) = g \circ f : X \to Y \to \mathbb{A}^1)$.

**Corollary 11.4.** The category of affine varieties over $k$ is equivalent to the opposite category of finitely generated $k$-algebras.

11.2. Rational maps. Let $k$ be a field.

**Definition 11.5.** Let $X/k$ be an irreducible variety. A rational map $f : X \to \mathbb{P}^n$ is an equivalence class $f = (f_0 : \cdots : f_n)$ of $(n+1)$-tuples $(f_i) \in k(X)^{n+1}$, not all identically 0. The equivalence relation is

$$(f_0 : \cdots : f_n) = (\lambda f_0 : \cdots : \lambda f_n)$$

for any $\lambda \in k(V)^\times$.

**Definition 11.6.** We make the following definitions.

1. A rational map $f : X \to \mathbb{P}^n$ is defined at $P \in X(L)$ (for any field extension $L/k$) if there exists a representative $(f_i)$ of $f = (f_0 : \cdots : f_n)$ such that $f(P) := (f_0(P) : \cdots : f_n(P)) \in \mathbb{P}^n(L)$ is well-defined (i.e. all $f_i(P)$ are defined with not all zero).
2. Let $Y \subseteq \mathbb{P}^n$ be a variety over $k$. A rational map $f : X \to Y \subseteq \mathbb{P}^n$ is a rational map $f : X \to \mathbb{P}^n$ such that $f(P) \in Y$ for every $P \in X$ at which $f$ is defined.
Note. A rational map $X \to \mathbb{P}^m$ (resp. $X \to Y$) is just an element of $\mathbb{P}^m(k(X))$ (resp. $Y(k(X))$).

**Proposition 11.7.** Let $f : X \to Y \subseteq \mathbb{P}^m$ be a rational map.

1. There is a (Zariski dense) open subset $U \subseteq X$ on which $f$ is defined.
2. Suppose $X \subseteq \mathbb{A}^m$ and $Y \subseteq \mathbb{A}^m = U_i \subseteq \mathbb{P}^m$ are affine varieties and $f$ is defined everywhere on $X$, then $f : X \to Y$ is a morphism of affine varieties.

The above proposition allows us to extend the notion of morphisms to those between varieties that are not necessarily affine.

**Definition 11.8.** A rational map $f : X \to Y$ which is defined at every $P \in X(L)$ for every field extension $L/k$ is called a morphism, and denoted $f : X \to Y$.

**Example 11.9.** The map $\mathbb{P}^1 \to \mathbb{P}^2$ given by

$$(X : Y) \mapsto (X^2 : XY : Y^2)$$

is a morphism. Strictly speaking, the map should be written $(t^2 : t : 1) = (1 : t^{-1} : t^2)$ where $t = X/Y \in k(\mathbb{P}^1)$. The image in $\mathbb{P}^2$ is the projective curve in $\mathbb{P}^2$ given (in homogeneous coordinates) $X_0, X_1, X_2$ on $\mathbb{P}^2$ by $X_1^2 = X_0X_2$.

**Example 11.10.** Let $C$ be the affine curve $x^2 + y^2 = 1$ in $\mathbb{A}^2$ over $\mathbb{Q}$ and $\overline{C}$ its projective closure $X^2 + Y^2 = Z^2$ in $\mathbb{P}^2$. In previous lectures, we constructed a rational map $f : \mathbb{P}^1 \to \overline{C} \subseteq \mathbb{P}^2$ by

$$(t : 1) \mapsto \left(\frac{1 - t^2}{1 + t^2} : \frac{2t}{1 + t^2} : 1\right).$$

This can be written as

$$(X : Y) \mapsto (X^2 - Y^2 : 2XY : X^2 + Y^2).$$

The map $f$ is defined everywhere (i.e. is a morphism) since $X^2 - Y^2 = X^2 + Y^2 = 2XY = 0$ if and only if $X = Y = 0$. On the other hand, we have the rational map $g : \overline{C} \to \mathbb{P}^1$ given by

$$(x : y : 1) \mapsto \left(\frac{y}{x + 1} : 1\right)$$

which can be written as $(X : Y : Z) \mapsto (Y : X + Z) = (X - Z : Y)$. (The equality on the right hand side holds on $\overline{C}$ from the equation $X^2 + Y^2 = Z^2$.) Since $(Y : X + Z)$ is defined everywhere on $\overline{C}$ away from $(1 : 0 : -1)$ and $(X - Z : Y)$ is defined everywhere on $\overline{C}$ away from $(1 : 0 : 1)$, it follows that the above rational map is defined everywhere (i.e. is a morphism). It can be verified that $g \circ f = \text{id}_{\mathbb{P}^1}$ and $f \circ g = \text{id}_{\overline{C}}$.

**Definition 11.11.** We have the following.

1. A map $f : X \to Y$ of varieties over a field $k$ is an isomorphism if there is a morphism $g : Y \to X$ such that $g \circ f = \text{id}_X$ and $f \circ g = \text{id}_Y$. We say $X$ and $Y$ are isomorphic if an isomorphism between them exists.
2. A rational map $f : X \to Y$ of irreducible varieties over a field $k$ is a birational map if there is a rational map $g : Y \to X$ such that $g \circ f = \text{id}_X$ and $f \circ g = \text{id}_Y$ where defined. We say $X$ and $Y$ are birationally equivalent if a birational map between them exists.

Note. Two irreducible varieties $X, Y$ are birationally equivalent if and only if $k(X) \cong k(Y)$. Birational geometry is the field of algebraic geometry which aims to classify algebraic varieties up to birational equivalence.
12. Lecture 12

12.1. Interlude: first view of elliptic curves. We review our progress and set the compass for the remainder of the course. Let $C \subset \mathbb{P}^2$ be a geometrically irreducible smooth projective algebraic curve over $\mathbb{Q}$. By Proposition 10.4, we have $C = \mathbb{V}_F$ for some homogeneous $F \in \mathbb{Q}[X,Y,Z]$ of some degree $\deg F = d \geq 1$. We consider the arithmetic of $C(\mathbb{Q})$ and how it depends on $d$.

Case $d = 1$. $C \subset \mathbb{P}^2$ is a projective line over $\mathbb{Q}$, and $C(\mathbb{Q})$ is trivially infinite (by linear algebra).

Case $d = 2$. We established the following results.

1. Local-to-global principle holds. Namely, by Hasse–Minkowski theorem,

\[ C(\mathbb{Q}) \neq \emptyset \iff C(\mathbb{Q}_p) \neq \emptyset \text{ for all } p \leq \infty. \]

Moreover, the right hand side is a finite check, by arithmetic in finite fields, the Chevalley–Warning theorem, Hensel’s lemma, etc.

2. Suppose $C(\mathbb{Q}) \neq \emptyset$, then there is a birational map $\mathbb{P}^1 \to C$ over $\mathbb{Q}$, which is in fact an isomorphism (this will follow from later lectures). In particular, we have $\mathbb{P}^1(\mathbb{Q}) \cong C(\mathbb{Q})$.

We now move into new territory ($d \geq 3$).

Case $d = 3$. It turns out that the following holds.

1. The local-to-global principle (also called the Hasse principle) fails! For example, if

\[ F(X, Y, Z) = 3X^3 + 4X^3 + 5X^3 \]

then $C(\mathbb{Q}_p) \neq \emptyset$ for all $p \leq \infty$ but $C(\mathbb{Q})$ is empty; this example is due to Selmer.

2. Suppose $C(\mathbb{Q}) \neq \emptyset$, say $P \in C(\mathbb{Q})$. Then it turns out that (by coordinate transformations) there is an isomorphism over $\mathbb{Q}$ from $C$ to a projective cubic curve of the form $E = E_{A,B}$ defined by the equation

\[ E : \quad Y^2Z = X^3 + AXZ^2 + BZ^3 \]

for some $A, B \in \mathbb{Q}$, sending the rational point $P$ to the rational point $O = (0 : 1 : 0) \in E(\mathbb{Q})$. In affine coordinates $x = X/Z$ and $y = Y/Z$, the equation for $E$ is

\[ y^2 = x^3 + Ax + B. \]

A curve of the form $E_{A,B}$ is an elliptic curve. How do we study $E(\mathbb{Q})$? Mimicking the study of conics by drawing lines through them, we may consider lines through various points of $E$. This leads to the following:

Observation. If $P_1, P_2 \in E(\mathbb{Q})$ and the line through $\overline{P_1 P_2}$ intersects $E$ at a third point $P_3$, then we have $P_3 \in E(\mathbb{Q})$. [This makes sense even when one of the points involved is $O$. For instance, if $P_1 = O$, then $\overline{P_1 P_2}$ is the vertical line through $P_2$.]

We define a composition law on $E(\mathbb{C})$ by defining the sum of two points $P_1, P_2 \in E(\mathbb{C})$ to be the reflection across the $x$-axis of the point $P_3 \in E(\mathbb{C})$ obtained as the third point of intersection between $\overline{P_1 P_2}$ and $E$. We will later prove the following:

Proposition 12.1. The above composition law equips $E(\mathbb{C})$ with the structure of an abelian group, with identity element $O$.

Poincaré observed around 1900 that $E(\mathbb{Q})$ forms a subgroup of $E(\mathbb{C})$, and raised questions about the structure of $E(\mathbb{Q})$, such as whether $E(\mathbb{Q})$ is finitely generated. This question was later resolved by Mordell; part of the second half of the course will be devoted to the proof of his result below.

Theorem 12.2 (Mordell, 1922). The group $E(\mathbb{Q})$ of rational points of an elliptic curve $E/\mathbb{Q}$ is a finitely generated abelian group. In particular, $E(\mathbb{Q})$ is of the form

\[ E(\mathbb{Q}) \cong \mathbb{Z}^r \oplus T \]

where $T$ is a finite abelian group and $r = \text{rk}(E/\mathbb{Q})$ is called the rank of $E/\mathbb{Q}$.
Remark. We do not yet have a general algorithm to determine $\text{rk}(E/Q)$ from $E/Q$. A celebrated conjecture of Birch and Swinnerton-Dyer relates $\text{rk}(E/Q)$ to the behavior of the counts of the mod $p$ points $E(\mathbb{F}_p)$ of $E$ over primes $p$.

Case $d \geq 4$. In this case, we have $g := (d - 1)(d - 2)/2 \geq 2$, so $C(Q)$ is finite by Faltings’ theorem.

12.2. Valuations and curves. Let $k$ be a field, and let $C/k$ be an irreducible curve.

**Definition 12.3.** Let $P \in C(k)$. We make the following definitions.

1. $\mathcal{O}_P = \{ f \in k(C) : f \text{ is defined at } P \}$ is the local ring of $C$ at $P$.
2. $\mathfrak{m}_P = \{ f \in \mathcal{O}_P : f(P) = 0 \}$ is the unique maximal ideal of $\mathcal{O}_P$.

**Remark.** If $C_0 \subset C$ is an affine patch of $C$ containing $P$ (so in particular $k(C) = \text{Frac} \, k[C_0]$), then $f$ is defined at $P$ if and only if there exist $g, h \in k[C_0]$ with $h(P) \neq 0$ such that $f = g/h$ in $k(C)$. Note in particular that $k[C_0] \subsetneq \mathcal{O}_P$.

**Example 12.4.** Let $C = \mathbb{A}_k^1$ be the affine line over $k$ and $P = 0$ the origin. Then:

$$k(C) = \left\{ \frac{p(t)}{q(t)} : p(t), q(t) \in k[t], q(t) \neq 0 \right\} =: k(t),$$

$$\mathcal{O}_P = \left\{ \frac{p(t)}{q(t)} : q(0) \neq 0 \right\},$$

$$\mathfrak{m}_P = \left\{ \frac{p(t)}{q(t)} : p(0) = 0, q(0) \neq 0 \right\}.$$ 

Note that $\mathcal{O}_P$ is the valuation ring for the discrete valuation $v_P : k(t) \to \mathbb{Z} \cup \{\infty\}$ associated to the irreducible polynomial $t$ (see Question 4.(b) of Problem Set 1). More generally, we have the following.

**Theorem 12.5.** Let $C/k$ be an irreducible curve and $P \in C(k)$ such that $C$ is regular at $P$ (viewed as a $k$-point). Then there exists a valuation $v_P : k(C) \to \mathbb{Z} \cup \{\infty\}$ such that

$$\mathcal{O}_P = \{ f \in k(C) : v_P(f) \geq 0 \} \quad \text{and} \quad \mathfrak{m}_P = \{ f \in k(C) : v_P(f) = 0 \}.$$ 

**Proof sketch.** We prove the theorem in the case where $k = \mathbb{C}$. Let $C_0 \subset C$ be an affine patch of $C$ containing $P$, so $k(C) = \text{Frac} \, k[C_0]$ and $k[C_0] \subset \mathcal{O}_P$. Let $\mathfrak{m} \subset k[C_0]$ be the maximal ideal corresponding to $P$; in other words, $\mathfrak{m}$ is the kernel of the evaluation map $k[C_0] \to k$ given by $f \mapsto f(P)$. The latter extends to a homomorphism $\mathcal{O}_P \to k$, whose kernel is $\mathfrak{m}_P$ by definition. One can verify that $\mathfrak{m}_P = \mathfrak{m} \mathcal{O}_P$ (i.e. the image of $\mathfrak{m}$ under $k[C_0] \to \mathcal{O}_P$ generates $\mathfrak{m}_P$), and moreover $\mathfrak{m}/\mathfrak{m}^2 = \mathfrak{m}_P/\mathfrak{m}_P^2$. Now, we have

$$\dim_k \mathfrak{m}_P/\mathfrak{m}_P^2 = \dim_k \mathfrak{m}/\mathfrak{m}^2 = 1$$

where the second inequality follows from our hypothesis that $C$ is regular at $P$. Let $t \in \mathcal{O}_P$ be such that $\mathfrak{m}_P/\mathfrak{m}_P^2 = k \cdot t$ (i.e. the class of $t$ in $\mathfrak{m}_P/\mathfrak{m}_P^2$ spans the entire $k$-vector space). In particular, any element of $\mathfrak{m}_P$ is of the form $ct + g$ for some unique $c \in k$ and $g \in \mathfrak{m}_P^2$. By considering products of such elements we conclude that, if $f \in \mathfrak{m}_P^n$, then $f \equiv ct^n \mod \mathfrak{m}_P^{n+1}$ for some (unique) $c \in k$, so $\mathfrak{m}_P^n/\mathfrak{m}_P^{n+1} = k \cdot t^n$. Thus, we may write each element $f \in \mathcal{O}_P$ formally as a power series

$$f = \sum_{n \geq 0} c_n t^n$$

for some $c_n \in k$ uniquely determined by $f$, and this furnishes for us an embedding of $k$-algebras $\mathcal{O}_P \to k[[t]]$ into the formal power series ring over $k$ in the variable $t$. This induces a field embedding $k(C) \hookrightarrow k((t))$. The valuation $v_P$ on $k(C)$ is the restriction of the usual $t$-adic valuation on $k[[t]]$ to $k(C)$ under this embedding. More explicitly, for $f \in \mathcal{O}_P$ we have $v_P(f) = \max\{ n : f \in \mathfrak{m}_P^n \}$. 

\[\square\]
Remark. We make the following observations.

(1) In the above proof, further works show that in fact \( k((t)) \) is isomorphic to the completion of \( k(C) \) with respect to \( v_P \), and \( k[[t]] \) is the closure of \( \mathcal{O}_P \). Geometrically, this tells us that the “formal neighborhoods” of smooth points on curves all look alike (and look like the formal neighborhood of \( \mathbb{A}^1 \) around the origin).

(2) The above proof shows that, if \( t \in \mathfrak{m}_P \) is such that \( k \cdot t = \mathfrak{m}_P / \mathfrak{m}_P^2 \), then \( t \) is a uniformizer of the valuation ring \( \mathcal{O}_P \).

Suppose that \( C \) is an affine plane curve defined by \( f(x, y) = 0 \) in \( \mathbb{A}^2 \), and suppose \( P = (a, b) \in C(\bar{k}) \) is a smooth point. By the Jacobian criterion (Proposition 10.10), this means that \( \frac{\partial f}{\partial x}(a, b) \neq 0 \) or \( \frac{\partial f}{\partial y}(a, b) \neq 0 \). Now, since \( \mathfrak{m}_P = (x - a, y - b) \), we have \( \mathfrak{m}_P / \mathfrak{m}_P^2 = \text{Span}_k \{(x - a), (y - b)\} \). On the other hand, \( f = 0 \) on \( C \) implies in particular

\[
0 = df = \frac{\partial f}{\partial x}(a, b)(x - a) + \frac{\partial f}{\partial y}(a, b)(y - a)
\]

in \( \mathfrak{m}_P / \mathfrak{m}_P^2 \). This implies the following, which may be viewed as an algebraic version (or approximation) of the implicit function theorem in calculus.

- If \( \frac{\partial f}{\partial y}(a, b) \neq 0 \) (i.e. tangent line is not vertical), then \( y - b = -\frac{\partial f/\partial x}{\partial f/\partial y}(x - a) \) in \( \mathfrak{m}_P / \mathfrak{m}_P^2 \), which shows that the class of \( x - a \) spans \( \mathfrak{m}_P / \mathfrak{m}_P^2 \), so \( x - a \) is a uniformizer.
- If \( \frac{\partial f}{\partial x}(a, b) \neq 0 \), then \( y - b \) is a uniformizer.

**Example 12.6.** Suppose \( C \subset \mathbb{A}^2 \) is given by \( y^2 = x^3 - x \) over \( \mathbb{Q} \) and \( P = (0, 0) \). Let \( f(x, y) = y^2 - x^3 + x \) so that \( C = V(f) \). We have \( \partial f/\partial x = -3x^2 + 1 \) which is nonzero at \( P \), so \( y \) is a uniformizer at \( P \); in other words, \( v_P(y) = 1 \). On the other hand, we have

\[
x = y^2 \left( \frac{1}{x^2 - 1} \right)
\]

and \( 1/(x^2 - 1) \in \mathcal{O}_P^*, \) so \( v_P(x) = 2v_P(y) + v_P(1/(x^2 - 1)) = 2 \).

12.3. Closed points.

**Definition 12.7.** Let \( V \) be an algebraic variety over \( k \). A **closed point** in \( V \) is a 0-dimensional irreducible subvariety of \( V \) over \( k \).

**Note.** If \( V \subset \mathbb{A}^n \) is an affine variety over \( k \), then Hilbert Nullstellensatz gives us a bijection

\[
\{\text{closed points of } V/k\} \leftrightarrow \{\text{maximal ideals of } k[V] = k[x_1, \ldots, x_n]/I(V)\}.
\]

In particular, if \( k = \bar{k} \) then the closed points of \( V/k \) are in bijection with \( V(k) \). If \( k \) is a field of characteristic zero (or more generally a perfect field), then it turns out that the closed points of \( V/k \) are in bijection with \( \text{Gal}((\bar{k}/k)) \)-orbits in \( V(\bar{k}) \).

**Example 12.8.** The closed points of the affine line \( \mathbb{A}^1_k \) over \( \mathbb{R} \) come in two types: the real points \( \mathbb{A}^1(\mathbb{R}) \subset \mathbb{R} \) of \( \mathbb{A}^1/\mathbb{R} \) corresponding to the maximal ideals of the form \((x - a) \subset \mathbb{R}[x]\), and the 0-dimensional irreducible subvarieties over \( \mathbb{R} \) corresponding to maximal ideals of the form \((x^2 + ax + b) \subset \mathbb{R}[x]\) where \( b^2 - 4ac < 0 \). The closed points of \( \mathbb{A}^1 \) over \( \mathbb{R} \) are in bijection with orbits of the complex points \( \mathbb{A}^1(\mathbb{C}) \subset \mathbb{C} \) under complex conjugation (the only nontrivial element of \( \text{Gal}(\mathbb{C}/\mathbb{R}) \)).

**Definition 12.9.** If \( P \) is a closed point of a curve \( C/k \), then one can define \( \mathcal{O}_P \) and \( \mathfrak{m}_P \) as before. The **residue field** \( k(P) = \mathcal{O}_P/\mathfrak{m}_P \) (also the function field of \( P \) as an irreducible variety) is a finite field extension of \( k \), and \( \deg P = [k(P) : k] \) is called the **degree** of \( P \) over \( k \). If \( C \) is smooth at a closed point \( P \) (i.e. \( C \) is smooth at any \( Q \in C(\bar{k}) \) belonging to \( P(\bar{k}) \subset C(\bar{k}) \)), then there is a discrete valuation \( v_P : k(C) \to \mathbb{Z} \cup \{\infty\} \) associated to \( P \) as before, and the same theory can be developed.
13.1. Rational maps between curves.

**Theorem 13.1.** If \( \phi : C \to X \) is a rational map from a smooth irreducible curve to a projective variety, then \( \phi \) is defined everywhere, i.e. \( \phi \) is a morphism.

**Proof.** It suffices to check that \( \phi \) is defined every closed point \( P \) of \( C \). Suppose \( X \subseteq \mathbb{P}^n \), and write \( \phi = (f_0 : \cdots : f_n) \) with \( f_i \in k(C) \). Since \( P \) is a smooth point of \( C \), there is an associated discrete valuation \( v_P \). Let \( f \) be the \( f_i \) such that \( v_P(f) = \min \{v_P(f_0), \ldots, v_P(f_n)\} \). Then \( \phi = (f_0/f, \ldots, f_n/f) \) where \( v_P(f_j/f) \geq 0 \) for all \( j \) (in particular \( f_j/f \) is defined at \( P \)) and \( f_i/f = 1 \) (in particular nonzero at \( P \)) so \( \phi(P) \in \mathbb{P}^n \) is well-defined. This shows that \( \phi \) extends to a well-defined morphism \( \phi : C \to \mathbb{P}^n \). The locus \( \phi^{-1}(X) \subseteq C \) is a Zariski dense closed subset of \( C \), and hence is equal to \( C \), i.e. the morphism \( \phi : C \to \mathbb{P}^n \) factors through \( X \), and \( \phi : C \to X \) is a morphism. \( \square \)

**Example 13.2.** Smoothness of \( C \) is necessary. For example, let \( C \) be the affine plane curve over \( \mathbb{Q} \) given by \( y^2 = x^2(x + 1) \) and consider the rational map \( C \to \mathbb{A}^1 \) given by \( (x,y) \mapsto y/x \). This is a birational map with birational inverse \( \mathbb{A}^1 \to C \) given by \( (1 - t^2, t - t^3) \). But \( \phi \) cannot be defined at \((0,0)\). Indeed, over \( \mathbb{R} \), \( \phi \) does not even extend to a continuous function at \((0,0)\).

**Remark.** If \( C/k \) is an irreducible curve, then there is a smooth projective curve \( C'/k \) and a birational map \( C' \to C \) over \( k \). This is an instance of resolution of singularities. Higher-dimensional analogue of this is due to Hironaka when the base field \( k \) has characteristic zero, and is still open for base fields of positive characteristic.

**Definition 13.3.** A nonconstant rational map \( \phi : C \to C' \) of irreducible curves over \( k \) induces a field homomorphism \( \phi^* : k(C') \to k(C) \) given by \( f \mapsto f \circ \phi \). This presents \( k(C) \) as a finite field extension of \( k(C') \), and \( \deg \phi := [k(C)/k(C')] \) is defined to be the degree of \( \phi \).

**Remark.** A nonconstant rational map \( \phi : C \to C' \) of irreducible curves over \( k \) induces a partially defined map \( C(k) \to C'(k) \) (defined everywhere if \( \phi \) is a morphism) which is surjective outside of finitely many points in \( C'(k) \). For all but finitely many \( p \in C'(k) \), the number of \( k \)-points in the fiber \( \phi^{-1}(p) \) is equal to the separable degree of \( k(C)/k(C') \). If \( k(C)/k(C') \) is a separable extension (which is always the case if \( k \) has characteristic zero), then this number is equal to \( \deg \phi = [k(C) : k(C')] \).

**Definition 13.4.** We define the following.

1. A nice over \( k \) if it is a smooth, projective, geometrically irreducible curve over \( k \).
2. A 1-dimensional function field over \( k \) is a field extension \( K/k \) such that \( K \) is a finite extension of \( k(t) \) and \( K \) contains no finite extension of \( k \).

**Theorem 13.5.** There is an equivalence of categories

\[
\left\{ \begin{array}{c}
\text{nice curves over } k \\
\text{with nonconstant rational maps}
\end{array} \right\} \cong \left\{ \begin{array}{c}
\text{1-dimensional function fields over } k \\
\text{with } k\text{-linear field homomorphisms}
\end{array} \right\}^{\text{op}}
\]

given by \( C \mapsto k(C) \).

**Example 13.6.** Let \( C \subseteq \mathbb{P}^2 \) be the projective closure of the affine curve \( C_0 \subseteq \mathbb{A}^2 \) given by \( y^2 = x^3 - x \) over \( \mathbb{Q} \). One can show that \( C/\mathbb{Q} \) is a nice curve. The morphism \( C_0 \to \mathbb{A}^1 \) given by \( (x,y) \mapsto x \) extends to a rational map \( C \to \mathbb{P}^1 \), with induced homomorphism

\[
\mathbb{Q}(\mathbb{P}^1) = \mathbb{Q}(x) \to \text{Frac} \mathbb{Q}[x,y]/(y^2 - x^3 + x) = \mathbb{Q}(x,y) = \mathbb{Q}(\sqrt[3]{x^3 - x}).
\]

This is a degree 2 extension. On the other hand, the fiber of a point \( x \in \mathbb{A}^1(\overline{\mathbb{Q}}) \subseteq \mathbb{P}^1(\overline{\mathbb{Q}}) \) under \( \phi \) is given by \( \{(x,\sqrt[3]{x^3-x}), (x,-\sqrt[3]{x^3-x})\} \), which has size 2 for all but finitely many \( x \).
13.2. Divisors. Let $k$ be a field and $C/k$ an irreducible curve.

**Definition 13.7.** The divisor group $\text{Div}(C) = \text{Div}(C/k)$ is given by

$$\text{Div}(C) = \bigoplus_{P \text{ closed points of } C \text{ over } k} \mathbb{Z} \cdot P.$$ 

A divisor on $C/k$ is an element of $\text{Div}(C)$, i.e. a formal sum $\sum n_P P$ with $n_P \in \mathbb{Z}$ and all but finitely many $n_P$ equal to zero.

We have a partial order on $\text{Div}(C)$ given by $\sum n_P P \geq m_P Q$ if and only if $n_P \geq m_P$ for all $P$.

**Definition 13.8.** A divisor $D = \sum n_P P \in \text{Div}(C)$ is effective if $D \geq 0$ (i.e. $n_P \geq 0$ for all $P$).

**Definition 13.9.** Given $D = \sum n_P P \in \text{Div}(C)$, we define its support by $\text{Supp}(D) = \{ P : n_P \neq 0 \}$.

Given $D \in \text{Div}(C)$, there exist effective divisors $D_1, D_2 \geq 0$ such that $D = D_1 - D_2$. Moreover, $D_1$ and $D_2$ are uniquely determined if we further require that $\text{Supp}(D_1) \cap \text{Supp}(D_2) = \emptyset$. Recall that, if $P$ is a closed point of $C$, then $\deg P = [k(P) : k]$ where $k(P) = \mathcal{O}_P/\mathfrak{m}_P$. We extend this notion to divisors on $C/k$.

**Definition 13.10.** Given $D = \sum n_P P \in \text{Div}(C)$, we define its degree to be $\deg D = \sum n_P \deg(P)$.

The kernel of the homomorphism $\deg : \text{Div}(C) \to \mathbb{Z}$ is denoted $\text{Div}^0(C)$.

**Remark.** If $k = \bar{k}$, then $k(P) = k$ for every closed point $P$ of $C/k$, and in fact the set of closed points of $C$ are in bijection with $C(k)$. In this case, we have $\deg(\sum n_P P) = \sum n_P$.

**Example 13.11.** Let $C \subset \mathbb{P}^2$ be the projective curve given by $X^2 + Y^2 = Z^2$ over $\mathbb{Q}$. Let $P = (1 : 0 : 1)$ and $Q = (3 : 4 : 5)$. Then $2P - 3Q \in \text{Div}(C)$, and $\deg(2P - 3Q) = 2\deg(P) - 3\deg(Q) = 2 - 3 = -1$.

We note that the definition of $\text{Div}(C)$ is sensitive to the base field $k$. We make the following definition.

**Definition 13.12.** If $V$ is a variety over $k$, and $L/k$ is a field extension, the base extension $V_L$ is the variety over $L$ obtained by viewing the polynomials defining $X$ as those with coefficients in $L$.

**Assume now that $k$ is a perfect field** (i.e. $k$ has characteristic zero or, in case $k$ has characteristic $p > 0$, then the Frobenius endomorphism $x \mapsto x^p$ is surjective). Let $C/k$ be a geometrically irreducible curve. If $P$ is a closed point of $C/k$, then its base extension $P_L$ decomposes into a finite union of $\bar{k}$-points $P_1, \ldots, P_n$ where $n = \deg P$. We have a homomorphism

$$\text{Div}(C) \to \text{Div}(C_{\bar{k}})$$

defined by sending $P \mapsto P_1 + \cdots + P_n$ for each closed point $P$ and $P_i$ as above, and extending this linearly to $\text{Div}(C)$.

**Proposition 13.13.** Let $C/k$ be as above with $k$ perfect. The homomorphism $\text{Div}(C) \to \text{Div}(C_{\bar{k}})$ is injective and its image is the group $\text{Div}(C_{\bar{k}})^{\text{Gal}(k/\bar{k})}$ of divisors of $C_{\bar{k}}$ that are invariant under the action of the Galois group $\text{Gal}(k/\bar{k})$.

**Example 13.14.** Let $C$ be the affine plane curve defined by $x^2 + y^1 = 1$ over $\mathbb{Q}$. Let $P = (1/2, \sqrt{3}/2)$ and $Q = (1/2, -\sqrt{3}/2)$ be points on $C(k)$. The divisor $P + Q$ is invariant under $\text{Gal}(\mathbb{Q}/\mathbb{Q})$-action, so it arises from an element of $\text{Div}(C)$. In fact, it is the image of the closed point $R$ of $C/\mathbb{Q}$ given by the equation $x = 1/2$ on $C$. Note that $k(R) = \mathbb{Q}[x, y]/(x^2 + y^2 - 1, x - 1/2) \cong \mathbb{Q}(\sqrt{3})$. 

34
13.3. **Principal divisors.** Let $k$ be a field. Let $C$ be a nice (i.e. smooth, projective, geometrically irreducible) curve over $k$. Given $f \in k(C)^\times$, let us define
\[
\text{div}(f) = (f) := \sum_P v_P(f)P \in \text{Div}(C).
\]
One can show that $v_P(f) = 0$ for all but finitely many $P$, so the above sum is finite.

**Definition 13.15.** A divisor $D \in \text{Div}(C)$ is **principal** if $D = (f)$ for some $f \in k(C)^\times$. We denote by $\text{PDiv}(C)$ the group of principal divisors. In other words, $\text{PDiv}(C)$ is the image of the group homomorphism
\[
k(C)^\times \rightarrow \text{Div}(C)
\]
given by $f \mapsto (f)$.

**Example 13.16.** Let $C = \mathbb{P}^1_k$, so $k(C) = k(t)$. Then for any closed point $P$ of $\mathbb{P}^1$ corresponding to an irreducible polynomial $p(t) \in k[t]$, the valuation $v_P(f)$ is the exponent of $p(t)$ in the complete factorization of $f \in k(t)^\times$. Thus, $(f)$ keeps track of the factorization of $f$.

**Fact 13.17.** We have the following.

1. For any nice curve $C/k$ and $f \in k(C)^\times$, we have
   \[
   \deg(\text{div}(f)) = 0.
   \]
   In other words, we have $\text{PDiv}(C) \leq \text{Div}^0(C)$. Moreover, writing $\text{div}(f) = D_1 - D_2$ with $D_1, D_2 \geq 0$ such that $\text{Supp}(D_1) \cap \text{Supp}(D_2) = \emptyset$, the following numbers are equal:
   - degree of the rational map $f : 1 : C \rightarrow \mathbb{P}^1$.
   - $\deg D_1$ (the number of “zeros” of $f$ counted with multiplicity).
   - $\deg D_2$ (the number of “poles” of $f$ counted with multiplicity).

2. We have $\text{PDiv}(\mathbb{P}^1) = \text{Div}(\mathbb{P}^1)$, i.e. every divisor of degree zero on $\mathbb{P}^1$ is principal.

14. **Lecture 14**

Let $C$ be a nice curve over a field $k$.

**Definition 14.1.** Two divisors $D_1$ and $D_2$ are **(linearly) equivalent** (and written $D_1 \sim D_2$) if $f \in k(C)^\times$ such that $D_1 - D_2 = \text{div}(f)$. A divisor class $[D]$ is an equivalence class of divisors, i.e. an element of the group
\[
\text{Pic}(C) = \frac{\text{Div}(C)}{\text{PDiv}(C)}
\]
called the Picard group of $C$.

**Example 14.2.** If $C = \mathbb{P}^1_k$, then $\text{PDiv}(C) = \text{Div}^0(C)$ and we have $\deg : \text{Pic}(C) \simeq \mathbb{Z}$.

In general, we have an exact sequence
\[
0 \rightarrow k^\times \rightarrow k(C)^\times \xrightarrow{\deg} \text{Div}(C) \rightarrow \text{Pic}(C) \rightarrow 0.
\]
Since $\text{PDiv}(C) \leq \text{Div}^0(C)$, the morphism $\deg : \text{Pic}(C) \rightarrow \mathbb{Z}$ given by $[D] \mapsto \deg(D)$ is well-defined. We denote $\text{Pic}^0(C) = \ker\{\deg : \text{Pic}(C) \rightarrow \mathbb{Z}\}$.

**Example 14.3.** Let $E \subset \mathbb{P}^2$ be the projective closure of the affine curve $E_0 \subset \mathbb{A}^2$ defined over $\mathbb{Q}$ by
\[
y^2 = x(x-1)(x-7).
\]
One can show that $E/\mathbb{Q}$ is a nice curve. We shall show that $\text{Pic}^0(E)$ contains a point of exact order $2$. Note that the homogeneous equation with respect to $X, Y, Z$ (where $x = X/Z$ and $y = Y/Z$) for $E$ is
\[
Y^2Z = X(X - Z)(X - 7Z).
\]
Intersecting with the “hyperplane at infinity” \( Z = 0 \), we get \( X = 0 \), so \( O = (0 : 1 : 0) \) is the unique point of \( E \) not contained in \( E_0 \). Consider \( \text{div}(x) \) where \( x = X/Z \in \mathbb{Q}(E_0)^* = \mathbb{Q}(E)^* \). On \( E_0 \), \( x \) is a regular function \( (\text{i.e. } v_P(x) \geq 0 \text{ for all closed points } P \text{ of } E_0) \) and vanishes only at \( P = (0,0) \). Since \( \deg(\text{div}(x)) = 0 \), we must have \( \text{div}(x) = nP - nO \) for some \( n \in \mathbb{Z}_{\geq 1} \). To find \( n \), note that \( \frac{\partial}{\partial x} (y^2 - x(x-1)(x-7)) \neq 0 \) at \( P \) so that \( y \) is a uniformizer of \( E \) at \( P \), and
\[
\frac{1}{(x-1)(x-7)} y^2 \implies v_P(x) = 2.
\]
This shows that \((x) = 2P - 2O \). Let \( D = P - O \). Then \( 2D = (x) \) so \([D] \in \text{Pic}^0(C) \) satisfies \( 2[d] = 0 \). To conclude, we must show that \( D \) is not principal. So suppose otherwise that \( D = (f) \) for some \( f \in \mathbb{Q}(E)^* \). Since \( \text{div}(f) = P - O \) this shows that \( f : E \to \mathbb{P}^1 \) has degree 1, i.e. \([\mathbb{Q}(E) : \mathbb{Q}(t)] = 1 \). This shows that \( E \) and \( \mathbb{P}^1 \) are birationally equivalent over \( \mathbb{Q} \), and hence isomorphic (since both curves are nice). But \( E(\mathbb{R}) \) has two connected components while \( \mathbb{P}^1(\mathbb{R}) \) only has one; this is a contradiction. This shows that we have \([D] \neq 0 \) in \( \text{Pic}^0(E) \).

14.1. **Genus.** Suppose \( C \) is a nice curve over \( \mathbb{C} \). Then \( C(\mathbb{C}) \) has the structure of a closed one-dimensional complex manifold \( (\text{i.e. compact Riemann surface}) \). In particular, the underlying surface looks like the surface of a donut with \( g \) “donut” holes \((g = 0 \text{ case being the sphere})\).

**Definition 14.4.** The **genus** of \( C/\mathbb{C} \) is the number \( g \) associated to \( C(\mathbb{C}) \) as above.

**Remark.** In fact, there is a way to define the genus of any nice curve \( C/k \) over any field \( k \), which recovers the definition given above.

**Fact 14.5.** If \( C \subset \mathbb{P}^2 \) is a nice curve in the projective plane defined by an irreducible polynomial of degree \( d \), then \( C \) has genus equal to
\[
g = \frac{(d - 1)(d - 2)}{2}.
\]

15. **Lecture 15**

15.1. **Genus and Newton polygons.** Recall that, given a nice curve \( C/\mathbb{C} \), the genus of \( C \) is the number of “donut holes” on the Riemann surface \( C(\mathbb{C}) \). We will give a method to compute the genus of a nice curve from the defining equation of its birational planar model.

**Definition 15.1.** We make the following definitions.

- A **lattice point** in \( \mathbb{R}^2 \) is an element of \( \mathbb{Z}^2 \subset \mathbb{R}^2 \).
- A **convex lattice polygon** \( P \) in \( \mathbb{R}^2 \) is the convex hull of a finite subset of \( \mathbb{Z}^2 \). The length \( \text{length}(s) \) of a side \( s \) of \( P \) is \( n - 1 \) where \( n \) is the number of lattice points on \( s \), including its end points.

Let \( k \) be a field. Suppose \( C/k \) is a nice curve birational to the affine plane curve defined by the equation \( f(x, y) = 0 \) where
\[
f(x, y) = \sum_{i,j} a_{ij} x^i y^j \in k[x, y].
\]

**Definition 15.2.** The **Newton polygon** of \( f \) is the convex hull of \( \{(i, j) \in \mathbb{Z}^2 : a_{ij} \neq 0\} \).

Let \( P \) be the Newton polygon of \( f \). Choose an orientation of each side \( s \) of \( P \), and for each \( s \) let \( f_s(t, u) \in k[t, u] \) be the homogeneous polynomial of degree equal to \( \text{length}(s) \) where the \( \text{length}(s) + 1 \) coefficients of \( f_s \) are the coefficients of \( f \) corresponding to the lattice points of \( s \).

**Theorem 15.3.** Let \( C/k \), \( f = \sum a_{ij} x^i y^j \) and \( P \) be as above. Suppose that

1. the affine curve defined by \( f(x, y) = 0 \) is smooth, and
(2) for each side $s$ of $P$, the polynomial $f_s$ is squarefree.

Then the genus of $C$ is equal to the number of lattice points in the interior of $P$.

**Example 15.4.** Let $C \subset \mathbb{P}^2$ be the projective closure of affine plane curve $C_0 \subset \mathbb{A}^2$ defined by $y^2 = x^3 - 2$ over $\mathbb{Q}$. One can show that $C/Q$ is a nice curve (exercise). The Newton polygon of $f(x, y) = y^2 - x^3 + 2$ is the convex hull of $(0, 0)$, $(3, 0)$, and $(0, 2)$. The length of the side $s_1 = (0, 2)(0, 0)$ joining $(0, 2)$ and $(0, 0)$ is 2. The length of the side $s_2 = (0, 0)(3, 0)$ is 3, and the length of the side $s_3 = (0, 2)(3, 0)$ is 1. We have

\[
\begin{align*}
    f_{s_1}(t, u) &= t^2 + 2u^2, \\
    f_{s_2}(t, u) &= 2t^3 - u^3, \\
    f_{s_3}(t, u) &= t - u.
\end{align*}
\]

The Jacobian criterion shows that the plane curve defined by $f(x, y) = 0$ is smooth. By the theorem above, it follows that $C/k$ has genus 1.

**Example 15.5.** Let $C/Q$ be a nice curve birational to the affine plane curve $C_0$ defined by

\[y^2 = p(x)\]

where $p(x) = a_dx^d + \cdots + a_1x + a_0 \in \mathbb{Q}[x]$ is a squarefree polynomial of degree $d \geq 1$. For simplicity, let us assume that $a_0 \neq 0$. The Jacobian criterion shows that $C_0$ is smooth. The Newton polygon of $f(x, y) = y^2 - p(x)$ is the convex hull of $(0, 0)$, $(d, 0)$, and $(0, 2)$. Let $s_1 = (0, 0)(0, 2)$, $s_2 = (0, 0)(d, 0)$, and $s_3 = (0, 2)(d, 0)$. Then we have

\[
\begin{align*}
    f_{s_1}(t, u) &= t^2 - a_0u^2, \\
    f_{s_2}(t, u) &= \text{homogenization of } p, \\
    f_{s_3}(t, u) &= \begin{cases} 
        t^2 - a_du^2 & \text{if } d \text{ is even,} \\
        t - a_d & \text{if } d \text{ is odd.}
    \end{cases}
\end{align*}
\]

By the theorem above, we have

\[\text{genus}(C) = \left\lfloor \frac{d - 1}{2} \right\rfloor.
\]

**Example 15.6.** Let $C \subset \mathbb{P}^2$ be a nice curve obtained as the projective closure of an affine curve $f(x, y) = 0$ of degree $d$. If $f(x, y)$ is “sufficiently generic,” the Newton polygon of $f$ is the convex hull of $(0, 0)$, $(d, 0)$, and $(0, d)$ and the genus of $C$ is computed by the above theorem to be

\[\text{genus}(C) = \frac{(d - 1)(d - 2)}{2}.
\]

This verifies the degree-genus formula mentioned in the previous lecture in the generic case.

15.2. **Riemann-Roch theorem.** Let $C/k$ be a nice curve.

**Definition 15.7.** Given $D \in \text{Div}(C)$, let $L(D) = \{f \in k(C)^\times : (f) + D \geq 0\} \cup \{0\}$.

**Proposition 15.8.** $L(D)$ is a $k$-linear subspace of $k(C)$.

**Proof.** Let us write $D = \sum n_PP$. Then $f \in L(D)$ if and only if $v_P(f) \geq -n_P$ for all $P$. For each $P$, it is easy to see that the set $V_P = \{f \in k(C) : v_P(f) \geq -n_P\}$ forms a $k$-linear subspace of $k(C)$. Since $L(D) = \bigcap_P V_P$, the conclusion follows. \(\Box\)

**Example 15.9.** If $D = 0$, then $L(0) = \{0\} \cup \{f \in k(C)^\times : (f) \geq 0\}$. But since $\deg((f)) = 0$, we have $(f) \geq 0 \iff (f) = 0 \iff f \in k^\times$. This shows that $L(0) = k$. 

37
Example 15.10. If \( D = 2P \) for a closed point \( P \), then \( L(D) \) is the space of \( f \in k(C) \) such that \( f \) has at most a double pole at \( P \) and is defined at every other closed point of \( C \). If \( D = 3P - 2Q \) for closed points \( P \) and \( Q \), then \( L(D) \) is the space of \( f \in k(C) \) such that \( f \) has at most triple pole at \( P \), no other poles, and at least a double zero at \( Q \).

Example 15.11. Let \( C = \mathbb{P}^1 \to A^1 \), and let \( \infty \) be the point \( \mathbb{P}^1 \setminus A^1 \), so \( v_\infty(p) = -\deg(p) \) for every polynomial \( p(t) \in k[t] = k[A^1] \). Let \( D = d\infty \) with \( d \geq 0 \). By definition, if
\[
f(t) = \frac{p(t)}{q(t)} \in k(t)
\]
with \( p, q \in k[t] \) coprime, then we have
\[
f \in L(d\infty) \iff v_\infty(p) - v_\infty(q) \geq -d \text{ and } v_P(p) \geq v_P(q) \text{ for all } P \neq \infty
\]
\[
\iff q \in k^x, f(t) = p(t)/q \text{ and } v_\infty(p) = -\deg p \geq -d, \text{ i.e. } \deg p \leq d.
\]
It follows that \( L(d\infty) = \{ f \in k[t] : \deg f \leq d \} \) and \( \dim_k L(d\infty) = d + 1 \).

More generally, it turns out that \( L(D) \) has finite dimension over \( k \) for any \( D \in \text{Div}(C) \).

Definition 15.12. For each \( D \in \text{Div}(C) \), define \( \ell(D) := \dim_k L(D) \in \mathbb{Z}_{\geq 0} \).

Proposition 15.13. We have the following observations.

1. If \( \deg D < 0 \), then \( L(D) = \{0\} \) and \( \ell(D) = 0 \).
2. If \( D \sim D' \), then \( \ell(D) = \ell(D') \).

Proof. If \( f \in L(D) \) is nonzero, then we must have \( \deg(D) \deg((f) + D) \geq 0 \) since \( (f) + D \geq 0 \). This proves the contrapositive of part (1). For (2), let \( D = D' + (g) \) for some \( g \in k(C)^x \). Then, for \( f \in k(C)^x \), we have
\[
(f) + D \geq 0 \iff (f) + (g) + D' \geq 0 \iff (fg) + D' \geq 0.
\]
So multiplication by \( g \) gives a \( k \)-linear map \( L(D) \to L(D') \) whose inverse map is given by multiplication by \( g^{-1} \), showing that \( L(D) \cong L(D') \). \( \square \)

Theorem 15.14 (Riemann-Roch). Let \( C/k \) be a nice curve of genus \( g \). There is a divisor class, consisting of divisors \( K \) called canonical divisors, such that
\[
\ell(D) - \ell(K - D) = \deg D + 1 - g
\]
for every \( D \in \text{Div}(C) \).

Corollary 15.15. Let \( K \) be any fixed canonical divisor on a nice curve \( C/k \).

1. \( \ell(K) = g \).
2. \( \deg(K) = 2g - 2 \).
3. If \( \deg D > 2g - 2 \), then \( \ell(D) = \deg D + 1 - g \).

Proof. (1) Setting \( D = 0 \), Riemann-Roch gives us \( 1 - \ell(K) = 0 + 1 - g \), whence \( \ell(K) = g \). (2) Setting \( D = K \), Riemann-Roch gives us \( \ell(K) - 1 = \deg K + 1 - g \) whence \( \deg K = g - 2 + \ell(K) = 2g - 2 \) by part (1). (3) If \( \deg D > 2g - 2 \), then \( \deg(K - D) < 0 \) so \( \ell(K - D) = 0 \). Riemann-Roch the gives us \( \ell(D) = \deg D + 1 - g \). \( \square \)

Example 15.16. Let \( C = \mathbb{P}^1 \). We showed that for \( d \geq 0 \) we have
\[
L(d\infty) = \{ p(t) \in k[t] : \deg p \leq d \} \implies \ell(d\infty) = d + 1.
\]
On the other hand, Riemann-Roch shows that $\ell(d\infty) = d+1-g$ for $d \gg 0$, verifying that $g = 0$ for the Riemann sphere $\mathbb{P}^1(\mathbb{C})$. Let $D \in \text{Div}(C)$ be a divisor of degree $d \in \mathbb{Z}$. Since $\text{PDiv}(\mathbb{P}^1) = \text{Div}^0(\mathbb{C})$, we have $[D] = d[\infty]$ in $\text{Pic}(C)$, and

$$\ell(D) = \begin{cases} 0 & \text{if } d < 0, \\ d + 1 & \text{if } d \geq 0. \end{cases}$$

**Proposition 15.17.** If $C$ is a nice curve of genus 0 over $k$ and $C(k) \neq \emptyset$, then $C \simeq \mathbb{P}^1$ over $k$.

**Proof.** Let $P \in C(k)$. Then $\deg P = 1 > -2$ showing that $\ell(P) = 2$ by Riemann-Roch. Now, $L(0) = k \subset L(P)$. Thus, there exists some $f \in L(P) - L(0)$, showing that $f$ is nonconstant with a simple pole at $P$ and no other poles. This shows that the degree of $(f : 1) : C \to \mathbb{P}^1$ is 1, i.e. is a birational map. We conclude that $C \simeq \mathbb{P}^1$ since both $C$ and $\mathbb{P}^1$ are nice. \hfill $\Box$

### 16. Lecture 16

16.1. **Elliptic curves.** Let $k$ be a perfect field of characteristic $\neq 2, 3$.

**Definition 16.1.** A (short) **Weierstrass equation** is a polynomial equation of the form

$$y^2 = x^3 + Ax + b$$

for some $A, B \in k$. (The long Weierstrass equation is of the form

$$y^2 + a_1 xy + a_3 y = x^3 + a_2 x^2 + a_4 x + a_6$$

for some $a_i \in k$. Since $\text{char } k \neq 2, 3$ by our assumption, we can complete squares in $y$ and cubes in $x$ to make $a_1, a_2, a_3 = 0$.)

**Proposition 16.2.** Let $E \subset \mathbb{P}^2$ be the projective closure of the affine plane curve $E_0$ defined by a Weierstrass equation $y^2 = x^3 + Ax + B$. The following are equivalent:

1. $E_0$ is smooth.
2. $E$ is smooth.
3. The polynomial $x^3 + Ax + B$ is separable (i.e. squarefree).
4. The discriminant $-16(A^3 + 27B^2)$ is nonzero.

If these conditions holds, then $E/k$ is a nice curve of genus 1 with a single point $(0 : 1 : 0)$ at infinity. Otherwise, $E$ has a unique singular point, and $E$ is birational to $\mathbb{P}^1$.

The proof of the above proposition is left as an exercise.

**Definition 16.3.** An **elliptic curve** over $k$ is a nice genus 1 curve over $k$ with a distinguished $k$-rational point $O \in E(k)$ (labelled the origin).

**Theorem 16.4.** We have the following.

1. Given a Weierstrass equation $y^2 = x^3 + Ax + B$ with $x^3 + Ax + B$ squarefree, the projective closure of the associated affine curve defines an elliptic curve over $k$ with distinguished $k$-rational point $(0 : 1 : 0)$.
2. Every elliptic curve over $k$ is isomorphic to a curve arising in this way.

**Proof.** (1) follows from the previous proposition. (2) Let $E$ be an elliptic curve, and let $P \in E(k)$ be the origin of $E$. By the corollary to Riemann-Roch, we have $\ell(nP) = n$ for all $n \geq 1$. So we have $L(0) = L(P) = k = \langle 1 \rangle$, and

- $L(2P) = \langle 1, x \rangle$ for some $x \in k(E)$, $v_P(x) = -2$,
- $L(3P) = \langle 1, x, y \rangle$ for some $y \in k(E)$, $v_P(y) = -3$. 

39
Since \( v_P(x^2) = -4 \) and \( v_P(xy) = -5 \), we have \( L(4P) = \langle 1, x, y, x^2 \rangle \) and \( L(5P) = \langle 1, x, y, x^2, xy \rangle \). Now, \( \ell(6P) \) so the functions \( 1, x, y, x^2, xy, x^3, y^2 \) spanning \( L(6P) \) must be linearly dependent, and moreover this linear relation must involve both \( x^3 \) and \( y^2 \) since they both have valuation \(-6\) at \( P \). Up to scaling \( x \) and \( y \), the linear relation gives the long Weierstrass equation. By a change of variables, we may assume that \( x, y \in k(E) \) satisfy the short Weierstrass equation.

We have a rational map \( (x : y : 1) : E \to \mathbb{P}^2 \) with image lying in the curve \( C \subset \mathbb{P}^2 \) given by the Weierstrass equation (with function field \( k(C) = k(x, y) \)). We have \( [k(E) : k(x)] = 2 \) and \( [k(E) : k(y)] = 3 \) since \( x \in L(2P) - L(P) \) and \( y \in L(3P) - L(2P) \). Since 2 and 3 are coprime, we must have \( [k(E) : k(C)] = 1 \), showing that \( E \simeq C \). Note that this isomorphism maps \( P \) to \((0 : 1 : 0)\).

### 16.2. Group law on elliptic curves.

**Theorem 16.5.** Let \( E \) be an elliptic curve with origin \( O \). Then the map of sets

\[
E(k) \to \text{Pic}^0(E)
\]

given by \( P \mapsto [P - O] \) is a bijection.

**Proof.** To show injectivity, suppose \( P, Q \in E(k) \) are such that \([P - O] = [Q - O]\). Then \( P - Q = (f) \) for some \( f \in k(C)^\times \). If \( P \neq Q \) then \( \deg((f : 1) : E \to \mathbb{P}^1) = 1 \), so \( E \simeq \mathbb{P}^1 \), a contradiction.

To show surjectivity, let \([D] \in \text{Pic}^0(E)\), with \( D \in \text{Div}^0(E) \). Then \( \ell(D + O) = 1 \) by Riemann-Roch. Thus, there exists \( f \) such that \((f) + D + O \geq 0 \). But \( \deg((f) + D + O) = 1 \), so \((f) + D + O = P \) for some \( P \in E(k) \), and \([D] = [P - O] \).

The above result endows \( E(k) \) with the structure of an abelian group. Let \( E \subset \mathbb{P}^2 \) be the elliptic curve given by a Weierstrass equation

\[
Y^2Z = X^3 + AXZ^2 + BZ^3.
\]

Given a line \( L \) in \( \mathbb{P}^2 \) given by a linear equation \( \ell(X, Y, Z) = 0 \), the intersection \( L \cap E \) consists of three \( k \)-points (counted with multiplicity), and defines an effective divisor of degree 3.

**Example 16.6.** Let \( L = \mathbb{V}_Z \) be the line at infinity. Then \( E \cap L = 3 \cdot (0 : 1 : 0) \).

If \( L_1 \) and \( L_2 \) are lines in \( \mathbb{P}^2 \) given by linear forms \( \ell_1 \) and \( \ell_2 \), then the ratio \( \ell_1/\ell_2 \) gives a rational function on \( E \) such that

\[
(f) = (L_1 \cap E) - (L_2 \cap E) \in \text{Div}^0(E).
\]

In particular, if \( L_2 \) is the line at infinity and \( L_1 \cap E = P + Q + R \) with \( P, Q, R \in E(k) \), then

\[
0 = [(f)] = [P - O] + [Q - O] + [R - O].
\]

**Proposition 16.7.** Let \( E \subset \mathbb{P}^2 \) be an elliptic curve in Weierstrass form, with origin \( O = (0 : 1 : 0) \), then

1. \( O \) is the identity for the group law on \( E(k) \), and
2. \( P, Q, R \in E(k) \) are collinear, i.e. \( L \cap E = P + Q + R \) for some linear \( L \subset \mathbb{P}^2 \), then \( P + Q + R = 0 \) in the group \( E(k) \).

**Proof.** This is obvious from the group law on \( E(k) \) defined by the bijection \( E(k) \simeq \text{Pic}^0(E) \) sending \( P \mapsto [P - O] \).

In fact, it is possible to define \( E \times E \) and addition morphism \( E \times E \to E \) and inverse morphism \( E \to E \) inducing the group law on \( E(k) \).
17. Lecture 17

17.1. Torsion on elliptic curves. Let $k$ be a perfect field.

**Definition 17.1.** Let $E/k$ be an elliptic curve, and $n \geq 1$. We say that $P \in E(L)$ (for some field extension $L/k$) is an $n$-torsion point if $nP = O$ in $E(L)$. The $n$-torsion subgroup $E[n]$ of $E(k)$ is the kernel of the map $[n] : E(k) \to E(\bar{k})$ given by $P \mapsto nP$.

**Fact 17.2.** If $C \to C'$ is a nonconstant morphism of nice curves, then the induced map $C(\bar{k}) \to C'(\bar{k})$ is surjective. It can be showed that $[n]$ is induced by a morphism $[n] : E \to E$, and hence is surjective on $E(\bar{k})$.

**Example 17.3.** Let $k$ be a perfect field of characteristic $\neq 2$, and let $E$ be the projective closure of $y^2 = f(x)$ where $f(x) \in k[x]$ is a squarefree polynomial of degree 3. Then $E[2] = \{O, (\alpha, 0) : \alpha \in \bar{k}, f(\alpha) = 0\}$. In particular, $E[2] \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ as groups. If moreover $f(x) = (x - e_1)(x - e_2)(x - e_3)$ for distinct $e_i \in k[x]$, then $E[2] \subseteq E(k)$.

17.2. Mordell’s theorem: outline.

**Theorem 17.4** (Mordell, 1922). Let $E/\mathbb{Q}$ be an elliptic curve. Then $E(\mathbb{Q})$ is a finitely generated abelian group.

By the structure theory of finitely generated abelian groups, Mordell’s theorem implies that $E(\mathbb{Q}) \cong \mathbb{Z}_r \times T$ for some integer $r \in \mathbb{Z}_{\geq 0}$ (rank of $E(\mathbb{Q})$) and some finite abelian group $T$ (torsion subgroup of $E(\mathbb{Q})$). The Mordell-Weil theorem is a generalization of Mordell’s theorem, due to Weil (1928), replacing $\mathbb{Q}$ by any number field $K$ and $E$ by any abelian variety $A/K$.

In this course, we will prove Mordell’s theorem under the additional hypothesis that $E[2] \subseteq E(\mathbb{Q})$, i.e. $E$ has rational 2-torsion. The proof follows two steps:

1. **Theorem** (Weak Mordell-Weil). If $E/\mathbb{Q}$ is an elliptic curve, then $E(\mathbb{Q})/2E(\mathbb{Q})$ is finite.
2. **Theorem of height functions on $E(\mathbb{Q})$**.

17.3. Weak Mordell-Weil. Let $E/\mathbb{Q}$ be an elliptic curve with $E[2] \subseteq E(\mathbb{Q})$, given by Weierstrass equation

$$y^2 = x^3 + Ax + B = (x - e_1)(x - e_2)(x - e_3)$$

with $A, B, e_i \in \mathbb{Q}$. Substituting $x = x'/d^2$ and $y = y'/d^8$ and multiplying the equation by $d^6$, we get an isomorphic curve with a new Weierstrass equation. By choosing $d$ so that the denominator of each $e_i$ divides $d$, we may assume henceforth that $e_i \in \mathbb{Z}$ throughout.

Our strategy to prove finiteness of $E(\mathbb{Q})/2E(\mathbb{Q})$ is the construct an injection

$$E(\mathbb{Q})/2E(\mathbb{Q}) \to \text{Hom}_{\text{cts}}(\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}), E[2])$$

and show that the image is finite by “local” computations.

Let $P \in E(\mathbb{Q})$ be given, and let $Q \in E(\overline{\mathbb{Q}})$ be such that $[2]Q = P$ in $E(\overline{\mathbb{Q}})$. The fiber $[2]^{-1}(P)$ decomposes over $\overline{\mathbb{Q}}$ into four $\overline{\mathbb{Q}}$-points

$$Q + E[2] = \{Q, Q + R_1, Q + R_2, Q + R_3\}$$

where $E[2] = \{O, R_1 = (e_1, 0), R_2 = (e_2, 0), R_1 + R_2 = (e_3, 0)\}$. Now, given any $\sigma \in \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, we have $[2]^\sigma(Q) = \sigma([2]Q) = \sigma P = P$, so $\sigma Q \in Q + E[2]$, showing that $\sigma Q - Q \in E[2]$.

**Lemma 17.5.** The map $\psi_P : \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}) \to E[2]$ given by $\sigma \mapsto \sigma Q - Q$ is a continuous group homomorphism independent of $Q$.  

41
Proof. Given $\sigma, \tau \in \text{Gal}(\mathbb{Q}/\mathbb{Q})$, we have

$$\psi_P(\sigma \tau) = \sigma \tau Q - Q = \sigma \tau Q - \sigma Q + \sigma Q - Q = \sigma (\tau Q - Q) + (\sigma Q - Q).$$

Since $\tau Q - Q \in E[2] \subset E(\mathbb{Q})$, we have $\sigma (\tau Q - Q) = \tau Q - Q$ so it follows that

$$\psi_P(\sigma \tau) = (\tau Q - Q) + (\sigma Q - Q) = \psi_P(\sigma) + \psi_P(\tau)$$

showing that $\psi_P$ is a group homomorphism (continuity of $\psi_P$ follows from the profinite topology on $\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$). Since we have $\sigma R_i = R_i \in E[2] \subset E(\mathbb{Q})$ for any $i$, we have

$$\sigma(Q + R_i) = (Q + R_i) = (\sigma Q - Q) + (\sigma R_i - R_i) = \sigma Q - Q$$

and showing that $\psi_P$ is independent of the choice of $Q \in [2]^{-1}P$. \qed

**Proposition 17.6.** The assignment $P \mapsto \psi_P$ gives us an injective group homomorphism

$$\frac{E(\mathbb{Q})}{2E(\mathbb{Q})} \rightarrow \text{Hom}_{cts}(\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}), E[2]).$$

Proof. If $P_1, P_2 \in E(\mathbb{Q})$ and $Q_1, Q_2 \in E(\overline{\mathbb{Q}})$ are such that $P_1 = 2Q_1$, then $P_1 + P_2 = 2(Q_1 + Q_2)$ so

$$(\psi_{P_1} + \psi_{P_2})(\sigma) = \psi_{P_1}(\sigma) + \psi_{P_2}(\sigma) + (\sigma Q_1 - Q_1) + (\sigma Q_2 - Q_2) = \sigma(Q_1 + Q_2) - (Q_1 + Q_2) = \psi_{P_1 + P_2}(\sigma).$$

This shows that $E(\mathbb{Q}) \rightarrow \text{Hom}_{cts}(\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}), E[2])$ given by $P \mapsto \psi_P$ is a group homomorphism. To prove the proposition, it suffices to show that $\psi_P = 0$ if $P \in 2E(\mathbb{Q})$. So suppose that $P = 2Q$ for some $Q \in E(\mathbb{Q})$. Then $\psi_P(\sigma) = \sigma Q - Q = Q - O$ for every $\sigma \in \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$, and hence $\psi_P = 0$ as desired. \qed

By Galois theory and basic algebra, there is a bijection

$$\text{Hom}_{cts}(\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q}), E[2]) \approx \begin{cases} \text{Finite Galois extensions } K/\mathbb{Q} \\
\text{with } \text{Gal}(K/\mathbb{Q}) \rightarrow E[2] \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \end{cases} \approx \frac{\mathbb{Q}^\times}{\mathbb{Q}^\times 2} \times \frac{\mathbb{Q}^\times}{\mathbb{Q}^\times 2},$$

where the bijection from the first term to the second is given by $\psi \mapsto \overline{\text{ker}(\psi)}$ and the bijection from the third term to the second is given by $(a, b) \mapsto \mathbb{Q}(\sqrt{a}, \sqrt{b})$.

**Fact 17.7.** The composition $E(\mathbb{Q})/2E(\mathbb{Q}) \rightarrow \cdots \rightarrow (\mathbb{Q}^\times/\mathbb{Q}^\times 2)^2$ is given for $P \in E(\mathbb{Q}) \setminus E[2]$ by

$$P = (x_0, y_0) \mapsto (\psi_P(\sigma) = \sigma Q - Q) \mapsto \mathbb{Q}(Q) = \mathbb{Q}(\sqrt{x_0 - e_1}, \sqrt{x_0 - e_2}) \mapsto (x_0 - e_1, x_0 - e_2).$$

Here, $Q \in E(\overline{\mathbb{Q}})$ is such that $2Q = P$. Additionally, under $E(\mathbb{Q})/2E(\mathbb{Q}) \rightarrow (\mathbb{Q}^\times/\mathbb{Q}^\times 2)^2$ we have

$$O \mapsto (1, 1)$$

$$R_1 = (e_1, 0) \mapsto ((e_1 - e_2)(e_1 - e_3), e_1 - e_2)$$

$$R_2 = (e_2, 0) \mapsto (e_2 - e_1(e_2 - e_3))$$

$$R_3 = (e_3, 0) \mapsto (e_3 - e_1(e_3 - e_2)).$$

Given $P \in E(\mathbb{Q})$, an element $\sigma \in \text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ belongs to $\text{ker}(\psi_P)$ if and only if $\sigma Q = Q$, and one can show that $\overline{\text{ker}(\psi_P)} = \mathbb{Q}(Q)$. The identification $\mathbb{Q}(Q) = \mathbb{Q}(\sqrt{x_0 - e_1}, \sqrt{x_0 - e_2})$ follows by explicit computations involving the rational map $[2]$, which is omitted here for brevity.

**Proposition 17.8.** Let $S$ be the set of primes $p$ such that $p \mid (e_i - e_j)$ for some distinct $i, j$. Let $Q(S, 2) \subset \mathbb{Q}^\times/\mathbb{Q}^\times 2$ be the finite subgroup generated by $-1$ and the primes in $S$. Then the image of the homomorphism

$$\phi : E(\mathbb{Q})/2E(\mathbb{Q}) \rightarrow (\mathbb{Q}^\times/\mathbb{Q}^\times 2)^2$$

constructed above is contained in $Q(S, 2) \times Q(S, 2)$. 

42
Proof sketch. Suppose \( P = (x_0, y_0) \in E(\mathbb{Q}) \setminus E[2] \), so \( \phi(x_0, y_0) = (x_0 - e_1, x_0 - e_2) \). Saying that \( x_0 - e_1 \in \mathbb{Q}(S, 2) \) is equivalent to saying that \( v_p(x_0 - e_1) \) is even for every \( p \notin S \). Fix \( p \notin S \).

Suppose first that \( v_p(x_0) < 0 \). Then \( v_p(x_0 - e_1) = v_p(x_0) \) for \( i = 1, 2, 3 \) since \( e_i \in \mathbb{Z} \). Now

\[
2v_p(y_0) = v_p(y_0^2) = v_p((x_0 - e_1)(x_0 - e_2)(x_0 - e_3)) = 3v_p(x_0)
\]

so \( 2 \mid v_p(x_0) \) as desired. Now, suppose \( v_p(x_0) \geq 0 \). Then \( p \) divides at most one of \( x_0 - e_1, x_0 - e_2, \) and \( x_0 - e_3, \) since otherwise \( p \) must divide \( e_i - e_j \) for some \( i \neq j \), contradicting the hypothesis that \( p \notin S \). On the other hand, \( v_p((x_0 - e_1)(x_0 - e_2)(x_0 - e_3)) \) is even by the same argument as above, so \( v_p(x - e_i) \) is even for each \( i \).

\[ \square \]

18. Lecture 18

18.1. Review of Galois theory. We review Galois theory to supplement the previous lecture.

Definition 18.1. Given an extension of fields \( L/k \), let \( \text{Aut}(L/k) \) be the group of field automorphisms of \( L \) fixing \( k \) elementwise.

Definition 18.2. An algebraic field extension \( L/k \) is Galois if

\[ L^{\text{Aut}(L/k)} := \{ x \in L : \sigma(x) = x \text{ for all } \sigma \in \text{Aut}(L/k) \} = k, \]

and we denote \( \text{Gal}(L/k) = \text{Aut}(L/k) \).

Remark. One can show that the above notion is equivalent to saying that \( L/k \) is

1. normal, i.e. every irreducible \( f \in k[x] \) with a zero in \( L \) factors completely into linear factors in \( L[x] \); and
2. separable, i.e. for every \( x \in L \) its minimal polynomial \( f \) over \( k \) satisfies \( \gcd(f, f') = 1 \) in \( k[x] \).

Fact 18.3. Let \( L/k \) be a Galois extension. Let \( I \) denote the collection of finite Galois extensions \( F/k \) with \( k \subseteq F \subseteq L \).

1. If \( F, F' \in I \), then \( FF' \in I \), where \( FF' \) denotes the subfield of \( L \) generated by \( F \) and \( F' \).
2. If \( k \subseteq E \subseteq L \) with \( E/k \) finite, then \( E \in F \) for some \( F \in I \).
3. \( \bigcup_{F \in I} F = L \).
4. If \( F, F' \in I \) with \( F \subseteq F' \), then we have a surjection

\[ \text{Gal}(F'/k) \twoheadrightarrow \text{Gal}(F/k) \]

given by restriction \( \sigma \mapsto \sigma|_F \).
5. If \( F \in I \), then \( |\text{Gal}(F/k)| = [F : k] \).

Proposition 18.4. The collection \( \{ \text{Gal}(F/k) : F \in I \} \) forms an inverse system of groups, and

\[ \text{Gal}(L/k) \cong \varprojlim \text{Gal}(F/k). \]

This endows \( \text{Gal}(L/k) \) with the profinite topology (also called the Krull topology in this context).

Theorem 18.5 (Fundamental theorem of Galois theory). Let \( L/k \) be a Galois extension. There is an inclusion-reversing bijection

\[ \{ \text{Field subextensions } k \subset E \subset L \} \leftrightarrow \{ \text{Closed subgroups } H \leq \text{Gal}(L/k) \} \]

given by \( E \mapsto \text{Gal}(L/E) \) with inverse map \( H \mapsto L^H \). Moreover, if \( E \leftrightarrow H \), then \( E/k \) is Galois if and only if \( H \) is normal, and \( E/k \) is finite if and only if \( H \) is open (or, equivalently, closed of finite index).
18.2. Infinite descent and height. Given $E/\mathbb{Q}$ elliptic curve with $E[2] \subseteq E(\mathbb{Q})$, we showed that $E(\mathbb{Q})/2E(\mathbb{Q})$ is finite. This by itself does not imply that $E(\mathbb{Q})$ is finitely generated. Indeed, there exist abelian groups $G$ with $G/2G$ finite but $G$ not finitely generated (e.g. $G = \mathbb{Q}$).

**Proposition 18.6.** Suppose $G$ is an abelian group such that $G/2G$ is finite. Suppose $h : G \to \mathbb{R}$ is a function satisfying the following conditions:

(i) For each $P_0 \in G$, we have

$$h(P + P_0) \leq 2h(P) + O_{P_0}(1) \quad \forall P \in G.$$ 

(ii) We have $h(2P) = 4h(P) + O(1)$ for all $P \in G$.

(iii) For each $B \in \mathbb{R}$, the set $\{ P \in G : h(P) \leq B \}$ is finite.

Then $G$ is finitely generated.

**Proof.** Let $R \subseteq G$ be a set of coset representatives of $G/2G$. If we consider only additions by the finitely many elements $P_0 \in R$, the $O(1)$ in (i) will be uniform. Now, given $Q_0 \in G$, write $Q_0 = 2Q_1 + r_1$ with $Q_1 \in G$ and $r_1 \in R$. Then

$$4h(Q_1) + O(1) = h(2Q_1) \leq 2h(Q_0) + O(1)$$

by (i) and (ii). This shows that

$$h(Q_1) \leq \frac{1}{2} h(Q_0) + O(1) \leq \frac{2}{3} h(Q_0)$$

if $h(Q_0) \gg 0$. Let us choose $B$ such that the above inequality holds whenever $h(Q_0) > B$. Without loss of generality, we may assume $h(r) \leq B$ for every $r \in R$. Now, let $S = \{ P \in G : h(P) \leq B \}$. We claim that $G = \langle S \rangle$. Fix $Q_0 \in G$, and write $Q_0 = 2Q_1 + r_1$. To show $Q_0 \in \langle S \rangle$, it suffices to show that $Q_1 \in \langle S \rangle$. We keep repeating the process by writing $Q_1 = 2Q_2 + r_2$, etc. We have

$$h(Q_1) \leq \frac{2}{3} h(Q_0)$$

unless $h(Q_0) \leq B$. If the former holds, we have $h(Q_2) \leq \frac{2}{3} h(Q_1)$ unless $h(Q_1) \leq B$, and so on. We thus have a sequence $Q_0, Q_1, Q_2, \ldots$, and we must ultimately have $h(Q_n) \in B$ for some $n \gg 0$. Then $Q_n \in S$, and we conclude that $Q_0 \in \langle S \rangle$ (This is an instance of Fermat’s infinite descent). □

The above result motivates us to introduce the notion of a height function $h : E(\mathbb{Q}) \to \mathbb{R}$, and show that it satisfies the conditions (i)-(iii) above.

**Definition 18.7.** Let $t = a/b \in \mathbb{Q}$ with $a, b \in \mathbb{Z}$ coprime. The exponential height of $t$ is

$$H(t) = \max(|a|, |b|).$$

We extend this notion to $t \in \mathbb{P}^1(\mathbb{Q}) = \mathbb{Q} \cup \{ \infty \}$ by setting $H(\infty) = 1$.

**Definition 18.8.** The (logarithmic) height of $t \in \mathbb{Q} \cup \{ \infty \}$ is $h(t) = \log H(t)$.

**Example 18.9.** We have $h(100) < h(1001/1000)$.

**Proposition 18.10** (Northcott property). For any $B > 0$, the set $\{ t \in \mathbb{Q} : H(t) \leq B \}$ is finite.

**Proof.** The set $t = a/b \in \mathbb{Q}$ with $a, b \in \mathbb{Z}$ satisfying $|a|, |b| \leq B$ is finite. □
19. Lecture 19

19.1. Heights and rational functions. Last time, we introduced the notion of height of a rational number \( t = a/b \in \mathbb{Q} \) in lowest terms:

\[ h(t) = \log H(t) \quad \text{where} \quad H(t) = \max\{|a|, |b|\} \]

and observed the Northcott property, i.e. for all \( B > 0 \) the set \( \{ t \in \mathbb{Q} : H(t) \leq B \} \) is finite. We now study the behavior of height under rational maps.

**Definition 19.1.** The degree of a rational function \( f(x) = p(x)/q(x) \in \mathbb{Q}(x) \) in lowest terms is \( \deg(f) = \max\{\deg p, \deg q\} \) (This is equal to the degree of \( f \) as a rational map \((f : 1) : \mathbb{P}^1 \to \mathbb{P}^1\).)

**Theorem 19.2.** If \( f(x) \) is a rational function of degree \( d \), then \( h(f(t)) = dh(t) + O_f(1) \) for all \( t \in \mathbb{Q} \).

**Remark.** Note that we have \( h(f(t)) = dh(t) \) if \( f(x) = x^d \). The theorem states that this fact continues to hold, up to a bounded error term that depends only on \( f \), for any rational function of degree \( d \).

**Proof.** Write \( f(x) = p(x)/q(x) \) where \( p, q \in \mathbb{Z}[x] \) have greatest common divisor 1, and let us write \( t = a/b \) in lowest terms. Let \( P(X, Y) = Y^d p(X/Y) \) and \( Q(X, Y) = Y^d q(X/Y) \) be homogeneous polynomials of degree \( d \) so that \( f(t) = f(a/b) = P(a, b)/Q(a, b) \).

We first show that \( h(f(t)) \leq dh(t) + O_f(1) \). Note that

\[ H(f(t)) \leq \max\{|P(a, b)|, |Q(a, b)|\} \leq C \max\{|a|, |b|\}^d \leq CH(t)^d \]

for some \( C \) depending only on \( f \). Taking logarithms of both sides, we have the desired inequality. It remains to show that \( h(f(t)) \geq dh(t) + O_f(1) \). For this, first note that since \( P(X, Y) \) and \( Q(X, Y) \) have no common zero in \( \overline{\mathbb{Q}} \) except at \((0, 0)\) we have, by nullstellensatz,

\[ \sqrt{(P, Q)} = (X, Y). \]

This shows that the exists an integer \( n \geq d \) such that \( X^n, Y^n \in (P, Q) \) in \( \mathbb{Q}[X, Y] \). Clearing denominators, we see that

\[ f_1 P + f_2 Q = cX^n, \quad g_1 P + g_2 Q = dY^n \]

for some \( f_i, g_i \in \mathbb{Z}[X, Y] \) and \( c, d \in \mathbb{Z}_{\geq 1} \) (note that these choices depend only on the original function \( f \)). Without loss of generality, we may assume that \( f_i \) and \( g_i \) are homogeneous of degree equal to \( n - d \). Evaluating at \( (X, Y) = (a, b) \) we get

\[ \gcd(P(a, b), Q(a, b)) | \gcd(ca^n, db^n) | cd \]

since \((a, b) = 1\). Informally, this means that \( P(a, b) \) and \( Q(a, b) \) are “almost coprime,” and thus we have

\[ H(f(t)) = H(P(a, b)/Q(a, b)) \sim \max\{|P(a, b)|, |Q(a, b)|\} \]

where \( \sim \) means that the ratio of both sides is bounded from above and away from 0 by a constants that depend only on \( f \). Now, we have

\[ c|a|^n \leq C_1 \max\{|a|, |b|\}^{n-d} \max\{|P(a, b)|, |Q(a, b)|\} \quad \text{and} \quad d|b|^n \leq C_2 \max\{|a|, |b|\}^{n-d} \max\{|P(a, b)|, |Q(a, b)|\} \]

for some constants \( C_1 \) and \( C_2 \) that depend only on \( f_i \) and \( g_i \) and hence on \( f \). Since \( c, d \geq 1 \) this shows that

\[ \max\{|a|, |b|\}^n \leq \max\{|a|, |b|\}^{n-d} \max\{|P(a, b)|, |Q(a, b)|\} \implies H(t) \leq CH(f(t)) \]

for some constant \( C \) depending only on \( f \). Taking logarithms, we get the desired inequality. \( \square \)
19.2. Heights on elliptic curves. Let \( E/Q \) be an elliptic curve with Weierstrass equation
\[
y^2 = (x - e_1)(x - e_2)(x - e_3) = x^3 + Ax + B, \quad A, B, e_i \in \mathbb{Z}.
\]

**Definition 19.3.** For \( P \in E(Q) \), define
\[
h_x(P) := h(x(P)).
\]
By convention, \( h_x(O) = 0 \). The *naive height* of \( P \) is defined to be \( h_{\text{naive}}(P) = \frac{1}{2} h_x(P) \).

We will show that \( h_x : E(Q) \to \mathbb{R} \) satisfies the three requisite conditions for descent in Proposition 18.6, thereby proving Mordell’s theorem. We note that property (iii) in Proposition 18.6 holds for \( h_x \) by the Northcott property of heights on rational functions (and the fact that the morphism \( E \to \mathbb{P}^1 \) given by the \( x \)-coordinate has finite fibers). We note prove property (ii).

**Proposition 19.4.** For all \( P \in E(Q) \), we have
\[
h_x(2P) = 4h_x(P) + O_E(1).
\]

*Proof.* One can show by explicit computation that if \( P = (x, y) \) then \( x(2P) = r(x) \) for a rational function \( r(x) \) of degree 4; indeed,
\[
x(2P) = \frac{x^4 - 2Ax^2 - 8Bx + A^2}{4(x^3 + Ax + B)}.
\]

Alternatively, we have a commutative diagram
\[
\begin{array}{ccc}
E & \xrightarrow{[2]} & E \\
\downarrow x & & \downarrow x \\
\mathbb{P}^1 & \xrightarrow{r} & \mathbb{P}^1
\end{array}
\]
inducing a diagram of inclusions of respective function fields. Since the field extensions associated to the vertical arrows are of degree 2 and \([2]\) has degree 4, it follows that \( r \) has degree 4 as well. Thus, by the previous subsection, we have \( h_x(2P) = h(r(x)) = 4h(x) + O_E(1) \) as desired. \( \Box \)

It remains to establish property (i) in Proposition 18.6 for \( h_x \). We will need two lemmas.

**Lemma 19.5.** Every rational point of \( E \) other than \( O \) is of the form \((a/d^2, b/d^3)\) for some \( a, b, d \in \mathbb{Z} \) with \( \gcd(a, d) = \gcd(b, d) = 1 \).

*Proof.* We have \( y^2 = x^3 + Ax + B \), so that denominator(\( y \))^2 = denominator(\( x \))^3. This shows that there exists \( d \geq 1 \) such that denominator(\( y \)) = \( d^3 \) and denominator(\( x \)) = \( d^2 \). \( \Box \)

**Lemma 19.6.** If \( P, Q \in E(Q) \setminus \{O\} \) satisfy \( x(P) + x(Q) \), then
\[
x(P) + x(Q) + x(P + Q) = \left( \frac{y(Q) - y(P)}{x(Q) - x(P)} \right)^2.
\]

*Proof.* Let \( y = mx + b \) be the line through \( P \) and \( Q \), so that
\[
m = \frac{y(Q) - y(P)}{x(Q) - x(P)}.
\]
If \( R \) is the third point of intersection of this line with \( E \), then \( x(P), x(Q), \) and \( x(R) = x(P + Q) \) are solutions to the equation
\[
(mx + b)^2 = x^3 + Ax + B
\]
so that
\[
x^3 - m^2x^2 + (A - 2mb)x + (B - b^2) = (x - x(P))(x - x(Q))(x - x(R)).
\]
Looking at the coefficient of \( x^2 \), we get \( x(P) + x(Q) + x(P + Q) = m^2 \). \( \Box \)
Proposition 19.7. Fix $P_0 \in E(\mathbb{Q})$. For every $P \in E(\mathbb{Q})$, we have
\[ h_x(P + P_0) \leq 2h_x(P) + O_{E,P_0}(1). \]

Proof. This is shown by explicit computation. Without loss of generality, we may assume $P_0 \neq O$ and $P \neq O, \pm P_0$. Write
\[ P = (x_0, y_0) = \left( \frac{a_0}{d_0^2}, \frac{b_0}{d_0^3} \right), \quad P = (x, y) = \left( \frac{a}{d^2}, \frac{b}{d^3} \right). \]
Note that we have
\[ x(P + P_0) = \left( \frac{y - y_0}{x - x_0} \right)^2 - x - x_0 \]
by previous lemma. We thus have
\[
x(P + P_0) = \frac{y^2 - 2yy_0 + y_0^2 - (x + x_0)(x - x_0)^2}{(x - x_0)^2} = \frac{x^3 + Ax + B + x_0^3 + Ax_0 + B - 2yy_0 - (x + x_0)(x - x_0)^2}{(x - x_0)^2} = \frac{(xx_0 + A)(x + x_0) + 2B - 2yy_0}{(x - x_0)^2} = \frac{(aa_0 + 4d^2a_0^2)(a_0d_0^2 + ad_0^2) + 2Bd_0^4d_0^4 - 2bd_0^5d_0}{(ad_0^2 - ad_0^2)^2}
\]
showing that
\[ H(x(P + P_0)) = O_{E_1,P_0}(1) \cdot \max\{|a|^2, |ad^2|, |d|^4, |bd|\}. \]
We have $|a|, |d^2| \leq H(x)$ on one hand, and on the other hand $y^2 = x^3 + Ax + B$ gives
\[ b^2 = a^3 + Aad^4 + Bd^6 \implies |b|^2 \leq O_{E,P_0}(1)H(x)^3. \]
We conclude that $H(x(P + P_0)) \leq O_{E,P_0}(1)H(x)^2. \]

Remark. In fact, one can show that there is a choice of height function $h : E(\mathbb{Q}) \to \mathbb{R}_{\geq 0}$ (called the Néron-Tate height) better than $h_x$, that satisfies the parallelogram identity
\[ h(P + Q) + h(P - Q) = 2h(P) + 2h(Q) \]
for every $P, Q \in E(\mathbb{Q})$. Note that the above identity implies both properties (i) and (ii) of Proposition 18.6, and equips $E(\mathbb{Q}) \otimes \mathbb{R}$ with the structure of a Euclidean vector space.

20. Lecture 20

Main reference for the next few lectures is “Lecture notes on Diophantine analysis” by Zannier. We will study integer solutions to Diophantine equations of the form
\[ f(x, y) = k \]
where $f(x, y) \in \mathbb{Z}[x, y]$ is homogeneous and $k \in \mathbb{Z}$. The above equation defines an algebraic curve $C$, and we may denote the set of solutions $(x, y) \in \mathbb{Z}^2$ of the above equation by $C(\mathbb{Z})$.

Note. If $V_1 \subseteq \mathbb{A}^{n_1}$ and $V_2 \subseteq \mathbb{A}^{n_2}$ are varieties over $\mathbb{Q}$ with an isomorphism $f : V_1 \to V_2$ over $\mathbb{Q}$, then we have $f : V_1(\mathbb{Q}) \cong V_2(\mathbb{Q})$ but it is not necessarily true that $f$ establishes a bijection (or even a map) between $V_1(\mathbb{Z})$ and $V_2(\mathbb{Z})$. (For example, consider affine plane curves $V_1 = V_{xy-1}$ and $V_{xy-2}$ and the map $V_1 \to V_2$ given by $(x, y) \mapsto (x, 2y)$. This can be remedied by requiring that both $f$ and its inverse $f^{-1}$ are defined as polynomial functions with integral coefficients. A rigorous treatment of this will require the introduction of notion of schemes over $\mathbb{Z}$ (generalizing varieties over $\mathbb{Q}$).
Remark. We may assume without loss of generality that \( f(x, y) \in \mathbb{Z}[x, y] \) is irreducible. (If coefficient of \( x^{\deg(f)} \) is nonzero, this is equivalent to saying that \( f(x, 1) \in \mathbb{Z}[x] \) is irreducible.) Indeed, otherwise we have \( f(x, y) = g(x, y)h(x, y) \) for some homogeneous \( g, h \in \mathbb{Z}[x, y] \), so solving \( f(x, y) = k \) amounts to solving the system

\[
\begin{align*}
g(x, y) &= k_1 \\
h(x, y) &= k_2
\end{align*}
\]

for finitely many \( k_1, k_2 \in \mathbb{Z} \), \( k_1k_2 = k \).

20.1. Linear equations. Consider the equation

\[ ax + by = k \]

where \( a, b, k \in \mathbb{Z} \), to be solved for \((x, y) \in \mathbb{Z}^2 \).

**Proposition 20.1.** The above equation is solvable over \( \mathbb{Z} \) if and only if \( \gcd(a, b) \mid k \). If this holds, then it has infinitely many integer solutions parametrized by \( \mathbb{Z}^1(\mathbb{Z}) = \mathbb{Z} \).

**Proof.** Clearly if \( ax + by = k \) has an integral solution then \( \gcd(a, b) \mid k \). Conversely, suppose that \( \gcd(a, b) \mid k \). To prove our claim it suffices to prove the case \( k = \gcd(a, b) \), since if \( k = \gcd(a, b) \cdot k_1 \) for some \( k_1 \in \mathbb{Z} \) and \( ax + by_1 = \gcd(a, b) \) then \( a(k_1x_1) + b(k_2x_2) = k \). So we want to show that

\[ ax + by = \gcd(a, b) \]

has an integral solution, which we can do using *Euclid’s method (or algorithm)* as follows. If \( a = 0 \) or \( b = 0 \) then the claim is obvious, so assume otherwise. We proceed by induction on \( |a| + |b| \). Without loss of generality, we have \( b \geq |a| > 0 \). We may write \( b = aq + r \) for some integers \( q, r \) with \( 0 \leq r < a \). Note that \( \gcd(a, r) = \gcd(a, b) \) and \( r \) is an integral linear combination of \( a \) and \( b \), so it suffices to show that

\[ ax + ry = \gcd(a, r) \]

has an integral solution. But \( |a| + r < |a| + |b| \), so this is true by inductive hypothesis, and we are done. Now, if \( ax + by = k \) has an integral solution \((x_0, y_0) \in \mathbb{Z}^2 \), then \((x_0 - bm, y_0 + am) \) is also a solution for any \( m \in \mathbb{Z} \). \( \square \)

*Remark.* Note the similarity with the proof of Hasse-Minkowski theorem from earlier in the course.

20.2. Quadratic equations. Consider the equation

\[ ax^2 + bxy + cy^2 = k \]

where \( a, b, c, k \in \mathbb{Z} \) and \( \Delta = b^2 - 4ac \) is not a square. Integer solutions to equations of this type can be determined effectively using Dirichlet’s method and quadratic irrationals. We will illustrate in a special case as follows.

**Proposition 20.2.** Let \( d \in \mathbb{Z} \) be a nonsquare integer. Then the integer solutions to

\[ x^2 - dy^2 = k \]

for any fixed \( k \in \mathbb{Z} \) can be effectively determined, and are “finitely generated.”

**Proof.** If \( d < 0 \), then the region \( x^2 + |d|y^2 \leq B \) in \( \mathbb{R}^2 \) is compact for any \( B \geq 0 \), so \( x^2 - dy^2 = k \) will have only finitely many integer solutions which can be found effectively. So suppose that \( d > 0 \). We claim it is enough to find a nontrivial integral solution to \( x^2 - dy^2 = 1 \) with \( y \neq 0 \). Indeed, suppose that

\[ p^2 - dq^2 = 1 \]

The answer is 48.
for some \((p, q) \in \mathbb{Z}^2\) with \(q \neq 0\). Write \(\eta = p + q\sqrt{d}\), so that \(\eta\) has norm \(N(\eta) = p^2 - dq^2 = 1\). Suppose that \((x, y) \in \mathbb{Z}\) is a solution to \(x^2 - dy^2 = k\), and write \(\alpha = x + y\sqrt{d}\). We may assume without loss of generality that \(\eta > 1\) and \(\alpha > 0\). Let \(c = \sqrt{|k|}\). Since \(\eta > 1\), there exists \(m \in \mathbb{Z}\) such that
\[
c^{-1} \leq \alpha \eta^m < c.
\]
Let \(\beta = \alpha\eta^k = u + v\sqrt{d} \in \mathbb{Z}[\sqrt{d}]\). Then \(\beta\bar{\beta} = N(\beta) = N(\alpha\eta^m) = k\). Note that
\[
|\beta| = \alpha\eta^k < c, \quad \text{and} \quad |\bar{\beta}| = \frac{|k|}{\beta} \leq \frac{|k|}{\alpha} < c.
\]
This shows that
\[
u = \frac{|\beta - \beta'|}{2\sqrt{d}} \leq \frac{|\beta| + |\beta'|}{2\sqrt{d}} < \frac{c}{\sqrt{d}},
\]
showing that \((u, v)\) lies in a bounded subset of \(\mathbb{R}^2\) and hence has only finitely many possible values. Thus, admitting the existence of \(\eta = p + q\sqrt{d}\) with \(p, q \in \mathbb{Z}\) and \(N(\eta) = 1\) (which will be proved below), to check if \(x^2 - dy^2 = k\) has a solution we only need to check finitely many points \((u, v) \in \mathbb{Z}^2\) satisfying the above bounds. Once the possible points \((u, v)\) are determined, all other solutions will be given in the form \((x, y)\) with \(x + y\sqrt{d} = (u + v\sqrt{d})\eta^m\) for some \(m \in \mathbb{Z}\).

**Theorem 20.3** (Pell equation). Let \(d\) be a positive nonsquare integer. Then the equation
\[
x^2 - dy^2 = 1
\]
admits a solution in integers \((x, y) \in \mathbb{Z}^2\) with \(y \neq 0\).

**Proof.** We will be using the following lemma:

**Lemma 20.4** (Dirichlet). Let \(\alpha \in \mathbb{R}\) and let \(Q > 0\) be a positive integer. Then there exist \(p, q \in \mathbb{Z}\) such that \((p, q) = 1\), \(0 < q \leq Q\), and
\[
\left| \alpha - \frac{p}{q} \right| \leq \frac{1}{q(Q + 1)}.
\]

**Corollary 20.5.** There are infinitely many \((x, y) \in \mathbb{Z}\) with \(y > 0\) such that
\[
\left| \frac{x}{y} - \sqrt{d} \right| \leq \frac{1}{y^2}.
\]

Now, suppose \((x, y) \in \mathbb{Z}^2\) satisfies the conclusion of the corollary. Then we have \(|x - \sqrt{d}y| \leq \frac{1}{y} \leq 1\), so \(x \leq 1 + y\sqrt{d}\). This shows that
\[
|x^2 - dy^2| \leq \frac{|x + y\sqrt{d}|}{y} \leq \frac{2|y|\sqrt{d} + 1}{|y|} \leq 2\sqrt{d} + 1.
\]
Since \(|x^2 - dy^2|\) is an integer, the above corollary shows that there exists an integer \(M\) in the interval \((-2 - 2\sqrt{d}, 2\sqrt{d})\) such that \(x^2 - dy^2 = M\) admits infinitely many solutions \((x, y) \in \mathbb{Z}^2\) with \(y > 0\). Moreover, we can partition the set
\[
\{(x, y) \in \mathbb{Z}^2 : x^2 - dy^2 = M\}
\]
into \(\mathbb{Z}/M\mathbb{Z} \times \mathbb{Z}/M\mathbb{Z}\) sets
\[
\{(x, y) \in \mathbb{Z}^2 : x^2 - dy^2 = M, (x, y) \equiv (c_1, c_2) \mod M\}.
\]
and one of these sets must contain infinitely many elements. Suppose that \((x_1, y_1)\) and \((x_2, y_2)\) are in this set. Write \((x_1 + y_1\sqrt{d})(x_2 - y_2\sqrt{d}) = A + B\sqrt{d}\), so that \(A = x_1x_2 - dy_1y_2 \in \mathbb{Z}\) and \(B = x_2y_1 - x_1y_2 \in \mathbb{Z}\). We may assume WLOG that \(B \neq 0\). Then

\[
A \equiv x_1^2 - dy_1^2 \equiv 0 \mod M \quad \text{and} \quad B \equiv x_1y_1 - x_1y_1 \equiv 0 \mod M
\]

showing that \(A = MA'\) and \(B = MB'\) for some \(A', B' \in \mathbb{Z}\). Then \(A^2 - dB^2 = (x_1^2 - dy_1^2)(x_2^2 - dy_2^2) = M^2\) and \(A'^2 - dB'^2 \equiv (A^2 - dB^2)/M^2 = 1\). Moreover, \(B' \neq 0\). This proves the theorem.

It remains for us to prove Dirichlet’s lemma.

**Proof of Dirichlet’s lemma.** Let \(\alpha\) and \(Q\) be given as in the lemma. Consider the \(Q + 1\) numbers

\[
0, \{\alpha\}, \ldots, \{Q\alpha\}
\]

where \(\{x\} = x - \lfloor x \rfloor\) denotes the fractional part of \(x\). Divide the half-open interval \([0, 1)\) into \(Q + 1\) intervals \(I_j = [j/(Q + 1), (j + 1)/(Q + 1)]\). Each of these intervals contains exactly one of the elements in the above sequence, then there exists \(0 < q \leq Q\) such that

\[
\frac{Q}{Q + 1} \leq \{q\alpha\} < 1.
\]

This means that there exists \(p \in \mathbb{Z}\) such that \(|q\alpha - p| \leq \frac{1}{Q + 1}\), giving the desired result. What remains is the case where there is an interval \(I_j\) with \(\{ro\}, \{so\} \in \mathbb{Z}\) for some \(0 \leq s < r \leq Q\). Then

\[
|\{ro\} - \{so\}| < \frac{1}{Q + 1}
\]

and arguing as above, we get the desired result. \(\square\)

21. LECTURE 21

**Theorem 21.1** (Thue, 1909). Let \(f(x, y) \in \mathbb{Z}[x, y]\) be a homogeneous polynomial of degree \(d \geq 3\), irreducible over \(\mathbb{Q}\). Then for any \(k \in \mathbb{Z}\) the equation

\[
f(x, y) = k
\]

has at most finitely many integral solutions \((x, y) \in \mathbb{Z}^2\).

The idea of Thue’s proof, using Diophantine approximation, is the following. Let \(C\) denote the curve given by the above equation, and let \(\overline{C} \subset \mathbb{P}^2\) its projective compactification. If \(C(\mathbb{Z})\) is infinite, then there is a sequence \(\{P_n\}\) in \(C(\mathbb{Z})\) that converges to a point \(P \in \overline{C} \setminus C\), say \(\lim_{n \to \infty} P_n = P\). Now, each point \(P \in \overline{C} \setminus C\) correspond to the slope (inverse) of an asymptotic direction of \(C\) (given by a root \(\alpha\) of \(f(x, 1)\)). The points \(P_n = (x_n, y_n)\) then give rise to rational approximation \(\frac{x_n}{y_n}\) of \(\alpha\), which by virtue of \(P_n\) lying on \(C\) will be “very good.” On the other hand, Thue’s result in Diophantine approximation (discussed below) will show that such good rational approximations cannot exist; so \(C(\mathbb{Z})\) must be finite.

21.1. **Diophantine approximation.** We recall that, by Dirichlet’s lemma, if \(\alpha \in \mathbb{R}\) is irrational then there are infinitely many \((p, q) \in \mathbb{Z}^2\) coprime with \(q > 0\) such that

\[
\left|\frac{p}{q} - \alpha\right| \leq \frac{1}{q^2}.
\]

Given \(\alpha\), one may ask if one can replace the exponent 2 by a larger number \(\delta\) and still get an infinitude of pairs \((p, q) \in \mathbb{Z}^2\) as above such that \(|\alpha - p/q| \leq 1/q^\delta\). This is not always possible, as the following example shows.
Example 21.2. Suppose $\alpha = \sqrt{d}$ for some nonsquare integer $d > 0$. Then for any $p + q\sqrt{d}$ with $p, q \in \mathbb{Z}$ and $q > 0$ we have

$$(p + q\sqrt{d})(p - q\sqrt{d}) = p^2 - dq^2 \neq 0 \implies |p - q\sqrt{d}| \geq \frac{1}{|p + q\sqrt{d}|}.$$ 

Suppose $|p - q\sqrt{d}| \leq 1$. Then $|p + q\sqrt{d}| \leq 1 + 2q\sqrt{d} \leq 3q\sqrt{d}$, so

$$\frac{|p - \sqrt{d}|}{q} = \frac{|p - q\sqrt{d}|1}{q} \geq \frac{1}{q}|p + q\sqrt{d}| \geq \frac{1}{3\sqrt{d}q^2}.$$ 

This shows that the exponent 2 cannot be improved for $\alpha = \sqrt{d}$. A similar argument shows that the same is true for any irrational quadratic number $\alpha \in \mathbb{R}$.

Theorem 21.3 (Liouville, 1844). Let $\alpha \in \overline{\mathbb{Q}}$ be an algebraic number of degree $d$. There exists an effectively computable constant $c = c(\alpha) > 0$ such that, for any $p, q \in \mathbb{Z}$ with $q > 0$ and $p/q \neq \alpha$,

$$\left| \alpha - \frac{p}{q} \right| \geq \frac{c}{q^d}.$$ 

Proof. Let $f(x) = a_0 x^d + \cdots + a_d \in \mathbb{Z}[x]$ be the minimal polynomial of $\alpha$ over $\mathbb{Q}$, with $a_0 > 0$. We may assume $\alpha \in \mathbb{R}$ and $|\alpha - p/q| \leq 1$, since otherwise the claim is trivial. By the mean value theorem,

$$f(\alpha) - f(p/q) = \left( \alpha - \frac{p}{q} \right) P'(\eta)$$

for some $\eta \in (\alpha - 1, \alpha + 1)$. So $|f'(\eta)| < 1/c$ for some $c = c(\alpha) > 0$. Since $f(\alpha) = 0$ we have then

$$|f(p/q)| = |\alpha - p/q| \cdot |f'(\eta)| \implies |\alpha - p/q| \geq c \cdot |f(p/q)|.$$ 

Since $f(p/q) \neq 0$ and $q^d f(p/q) \in \mathbb{Z}$, we have $|f(p/q)| \geq 1/q^d$, finishing the proof. 

Note. Liouville’s theorem can be used to construct examples of transcendental numbers.

Example 21.4. The number

$$\alpha = \sum_{n=1}^{\infty} 10^{-n!}$$

is transcendental. Indeed, let $p_k = 10^{k+1} \sum_{n=k}^{\infty} 10^{-n!}$ and $q_k = 10^{k!}$ for $k = 1, 2, \ldots$. Then $p_k, q_k$ are relatively prime, and

$$\left| \frac{p_k}{q_k} - \alpha \right| = \sum_{n=k+1}^{\infty} 10^{-n!} < 10^{-(k+1)!} \sum_{n=0}^{\infty} 10^{-n} = \frac{10}{9} q_k^{-(k+1)} < \frac{1}{q_k^{k+1}}.$$ 

By Liouville’s theorem, $\alpha$ must be transcendental.

Definition 21.5. Given $\alpha \in \mathbb{R}$, the exponent of approximation $e(\alpha)$ is

$$e(\alpha) = \inf \left\{ \delta \in \mathbb{R} : \left| \alpha - \frac{p}{q} \right| < \frac{1}{q^\delta} \text{ has only finitely many solutions in rational } p/q \right\}.$$ 

Note. We have $e(\alpha) = 1$ for every $\alpha \in \mathbb{Q}$. For, if $\alpha = r/s$ then

$$\left| \frac{r}{s} - \frac{p}{q} \right| = \left| \frac{rq - ps}{sq} \right| \geq \frac{1}{sq}$$

unless $r/s = p/q$. By our previous discussion, $e(\alpha) = 2$ for $\alpha \in \mathbb{R}$ a quadratic irrational. Liouville’s theorem shows that $e(\alpha) \leq d$ if $\alpha \in \mathbb{R}$ is algebraic of degree $d$. The following theorem of Thue gave a first improvement of Liouville’s bound.
Theorem 21.6 (Thue, 1909). If \( \alpha \in \mathbb{R} \) is algebraic of degree \( d \geq 3 \), then \( e(\alpha) \leq 1 + \frac{d}{2} \). In other words, for each \( \epsilon > 0 \) there exists \( \gamma(\alpha, \epsilon) > 0 \) such that, for all \( p, q \in \mathbb{Z} \) with \( q > 0 \),
\[
\left| \alpha - \frac{p}{q} \right| > \frac{\gamma}{q^{1 + \frac{d}{2} + \epsilon}}.
\]

Remark. Thue’s proof is ineffective in the it does not give a procedure to determine \( \gamma(\alpha, \epsilon) \). Thue’s theorem was later generalized (with ineffectivity persisting) as follows. Let \( \alpha \in \mathbb{R} \) be algebraic of degree \( d \geq 2 \).

- Siegel: \( e(\alpha) \leq 2\sqrt{d} \).
- Gelfond, Dyson: \( e(\alpha) \leq \sqrt{2d} \).
- Roth (1955): \( e(\alpha) = 2 \).

We now deduce Thue’s theorem on Diophantine equations from his approximation theorem.

Proof of Theorem 21.1. Let \( f \in \mathbb{Z}[x, y] \) be homogeneous of degree \( d \geq 3 \), irreducible over \( \mathbb{Q} \). If \( f(x, y) = k \) for some \((x, y) \in \mathbb{Z}^2\), then we claim there is a root \( \alpha \) of \( f(x, 1) \) with
\[
\left| \alpha - \frac{x}{y} \right| \leq \frac{B}{|y|^d}
\]
where \( B \) depends only on \( f \) and \( k \). By Thue’s theorem on Diophantine approximation, this will show that there must be only finitely integral solutions \((x, y) \in \mathbb{Z}^2\).

To see the above inequality, let us write \( f(x, y) = a_0 \prod_{i=1}^d (x - \alpha_i y) \) where \( \alpha_i \) are distinct algebraic numbers and \( a_0 \in \mathbb{Z} \). Let us write \( \eta = \min_{i \neq j} |\alpha_i - \alpha_j| > 0 \). Let \( \alpha \) be one of the \( \alpha_i \) such that \( |x - \alpha y| \) is minimum. Then from \( f(x, y) = k \) we get
\[
|x - \alpha y|^d \leq \frac{|k|}{|a_0|} \leq |k|.
\]
Let us set \( B = |k|/(2\eta)^{d-1} \). We have two cases to consider:

1. Suppose first that \( |y| \leq 2|x - \alpha y|/\eta \). Then \( |x - \alpha y| \cdot |y|^{d-1} \leq |x - \alpha y|^d \cdot (2/\eta)^{d-1} \leq |k|(2/\eta)^{d-1} \), whence
\[
\left| \alpha - \frac{x}{y} \right| = \frac{|x - \alpha y|}{|y|} \leq \frac{1}{|y|^d} |k|(2/\eta)^{d-1} = \frac{|B|}{|y|^d}
\]
as desired.

2. Suppose next that \( |y| > 2|x - \alpha y|/\eta \). If \( \alpha_j \neq \alpha \) then
\[
|x - \alpha_j y| = |(x - \alpha y) - (\alpha - \alpha_j) y| \\
\geq |\alpha - \alpha_j| |y| - |x - \alpha y| \\
\geq \eta |y| - |x - \alpha y| \geq \eta |y|/2.
\]

Whence
\[
|k| = |a_0| \prod_{i \neq j} |x - \alpha_j y| \geq \left( \frac{\eta}{2} |y| \right)^{d-1} |x - \alpha y|
\]
from which the desired inequality follows. \( \square \)

22. Lecture 22

22.1. Runge’s method. We describe one (effective) method of getting finiteness of integral points for some affine algebraic curves which does not involve Diophantine approximation. Let \( C \subset \mathbb{A}^2 \) be an affine plane curve over \( \mathbb{Q} \), given by \( f(x, y) = 0 \) where \( f \in \mathbb{Z}[x, y] \). Let \( \overline{C} \subset \mathbb{P}^2 \) be the projective closure, and let \( D = \overline{C} \setminus C \) be the divisor at infinity. For simplicity of exposition, we shall assume that \( \overline{C} \) is nice.
Theorem 22.1 (Runge, 1887). Suppose that $D = D_1 + D_2$ for some divisors $D_1, D_2 \in \text{Div}(\overline{C}/\mathbb{Q})$ over $\mathbb{Q}$ such that $D_1, D_2 > 0$ and $\text{Supp}(D_1) \cap \text{Supp}(D_2) = \emptyset$. Then $C(\mathbb{Z})$ is finite.

Proof. Suppose that $\{P_n\}$ is a sequence in $C(\mathbb{Z})$ converging to some $P \in \overline{C} \setminus C$. Note that $P \in C(\overline{\mathbb{Q}})$. Without loss of generality, we have $P \in \text{Supp}(D_1)$ if we view $D_1$ as a divisor over $\overline{C}_{\overline{\mathbb{Q}}}$ under the inclusion $\text{Div}(\overline{C}) \hookrightarrow \text{Div}(\overline{C}_{\overline{\mathbb{Q}}})$. Thus, $P \notin \text{Supp}(D_2)$.

By Riemann-Roch, we have $\ell(N \cdot D_2) > 0$ for some $N \gg 0$, so there exists $f \in \mathbb{Q}(C)$ such that $(f) + N \cdot D_2 \geq 0$. Since $f$ has no pole on $C$, we must have $f \in \mathbb{Q}[C]$. Up to multiplying $f$ by a suitable integer, we may assume that $f$ is in the image of $\mathbb{Z}[x,y] \to \mathbb{Q}[C]$, so in particular $f$ takes integral values on $P_n \in C(\mathbb{Z})$. Now, note that $f(P)$ is finite since $f$ has no pole along $D_1$. Since $\lim_{n \to \infty} f(P_n) = f(P)$ it follows that the sequence $f(P_n)$ must be bounded. There then exists $M \in \mathbb{Z}$ such that if $f(P_n) = M$ for infinitely many of the elements $P_n$, showing that $f$ must be constant, a contradiction.

Example 22.2. Consider the affine plane curve $C$ given by the equation

$$x^4 - y^4 - x^3 = 2.$$ 

We see that the divisor $D = \overline{C} \setminus C$ at infinity is given by

$$D = \{(X : Y : 0) \in \mathbb{P}^2(\overline{\mathbb{Q}}) : X^4 + Y^4 = 0\} = (1 : 1 : 0) + (1 : -1 : 0) + ((1 : i : 0) + (1 : -i : 0)) \in \text{Div}(\overline{C}).$$

We see that if $C$ has infinitely many points then there must be a sequence of them converging to $(1 : 1 : 0)$ or $(1 : -1 : 0)$. But note that

$$x - y = \frac{2 + x^3}{(x + y)(x^2 + y^2)}$$

is bounded near $(1 : 1 : 0)$, and

$$x + y = \frac{2 + x^3}{(x - y)(x^2 + y^2)}$$

is bounded near $(1 : -1 : 0)$. This shows that $C(\mathbb{Z})$ must be finite.

22.2. Mordell’s equation. The equation

$$y^2 = x^3 + k$$

for $k \in \mathbb{Z}$ is commonly referred to as Mordell’s equation. Mordell used Thue’s theorem and the theory of binary cubic forms to prove the following result.

Theorem 22.3. Given any nonzero integer $k \in \mathbb{Z}$, the equation $y^2 = x^3 + k$ has only finitely many solutions $(x, y) \in \mathbb{Z}^2$.

We will present the proof of Mordell’s theorem in the next lecture. For now, we illustrate the case $k = -2$ of the above theorem (considered by Fermat) using basic algebraic number theory.

Example 22.4. The only integral solutions to $y^2 = x^3 - 2$ are $(x, y) = (3, \pm 5)$.

Proof. Suppose $(x, y) \in \mathbb{Z}^2$ is a solution. If $x$ is even then $y^2 \equiv -2 \mod 8$, which is impossible; so but $x$ and $y$ must be odd. Below, we will use the fact that $\mathbb{Z}[\sqrt{-2}]$ is a UFD as well as that the units of $\mathbb{Z}[\sqrt{-2}]$ are $\{\pm 1\}$. We have

$$x^3 = y^2 + 2 = (y + \sqrt{-2})(y - \sqrt{-2}).$$

We claim that the two factors on the right hand side are relatively prime. Indeed, if $\delta \mid y + \sqrt{-2}$ and $\delta \mid y - \sqrt{-2}$ in $\mathbb{Z}[\sqrt{-2}]$ then $\gamma \mid 2\sqrt{-2}$ This shows $N(\gamma) \mid N(2\sqrt{-2}) = 8$. But $N(\delta) \mid N(y + \sqrt{-2})y^2 + 2$
which is odd, so we must have \( N(\delta) = 1 \) and \( \delta \) is a unit. By unique factorization in \( \mathbb{Z}[\sqrt{-2}] \), we conclude that each of \( y + \sqrt{-2} \) and \( y - \sqrt{-2} \) must be a cube up to a unit, whence must be a cube (since all units of \( \mathbb{Z}[\sqrt{-2}] \) are cubes). So we can write
\[
y + \sqrt{-2} = (m + \sqrt{-n})^3
\]
for some \( m, n \in \mathbb{Z} \). Expanding both sides, we find two equations
\[
y = m^3 - 6mn^2 = m(m^2 - 6n^2) \quad \text{and} \quad 1 = 3m^2n - 2n^3 = n(3m^2 - 2n^2).
\]
The second equation shows that \( n = \pm 1 \). If \( n = 1 \), then \( 1 = 3m^2 - 2 \) by the first equation so that \( m = \pm 1 \). We then have \( y = \pm 5 \) and \( x = 3 \). The case \( n = -1 \) cannot occur, since otherwise \( 1 = -3(m^2 - 2) \) so \( 3m^2 - 1 \), which is impossible. \( \Box \)

**Remark.** In general, Baker showed that an integral solution \((x, y) \in \mathbb{Z}^2\) to the Mordell equation \( y^2 = x^3 + k \) must satisfy
\[
\max\{|x|, |y|\} < \exp(10^{10}|k|^{10^4})
\]
making the solution of Mordell’s equation effective in principle (if not in practice).

23. Lecture 23

Less is known about integer points on affine varieties of dimension greater than one. Varieties with arithmetic group actions form a class of varieties where basic structure results have been established.

23.1. Binary quadratic forms.

**Definition 23.1.** An (integral) binary quadratic form is a homogeneous polynomial of degree 2 in two variables:
\[
Q(x, y) = ax^2 + bxy + cy^2 \in \mathbb{Z}[x, y].
\]
We say a binary quadratic form \( Q \) represents an integer \( k \) if \( Q(x, y) = k \) has a solution \((x, y) \in \mathbb{Z}^2\).

Study of binary quadratic forms and their representations of integers formed the basis of algebraic number theory (consider e.g. Fermat’s work on \( x^2 + y^2 = k \)).

**Definition 23.2.** Let
\[
\text{SL}_2(\mathbb{Z}) = \left\{ A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} : a, b, c, d \in \mathbb{Z}, \det(A) = ad - bc = 1 \right\}
\]
be the group (under matrix multiplication) of integral \( 2 \times 2 \) matrices with trivial determinant. We say that two binary quadratic forms \( Q_1, Q_2 \) are equivalent if there exists \( A \in \text{SL}_2(\mathbb{Z}) \) such that
\[
Q_1(x, y) = Q_2(x', y') \quad \text{where} \quad \begin{bmatrix} x \\ y \end{bmatrix} = A \begin{bmatrix} x' \\ y' \end{bmatrix}.
\]

**Note.** Equivalent forms represent the same integers.

**Definition 23.3.** The discriminant of \( Q(x, y) = ax^2 + bxy + cy^2 \in \mathbb{Z}[x, y] \) is \( \text{disc}(Q) = b^2 - 4ac \).

**Lemma 23.4.** If two binary quadratic forms are equivalent, they have the same discriminant.

**Proof.** A binary quadratic form \( Q(x, y) = ax^2 + bxy + cy^2 \) is given by
\[
Q(x, y) = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}
\]
where \( \text{disc}(Q) = -4 \det(M) \). Action of an element in \( \text{SL}_2(\mathbb{Z}) \) replaces \( M \) by \( A'MA \) for some \( A \in \text{SL}_2(\mathbb{Z}) \), which has the same discriminant. \( \Box \)
Note. The set of binary quadratic forms of fixed discriminant $k$ can be viewed as the $\mathbb{Z}$-points $V_k(\mathbb{Z})$ of the affine variety

$$V_k := V_{b^2 - 4ac - k} \subseteq \mathbb{A}^3_{a,b,c}$$

of dimension 2 given by $b^2 - 4ac = k$. The group $\text{SL}_2(\mathbb{Z})$ acts on $\mathbb{A}^3_{a,b,c}$ by linear transformations preserving each $V_k$, and the equivalence classes of binary quadratic forms with discriminant $k$ are parametrized by the orbits $\text{SL}_2(\mathbb{Z})\setminus V_k(\mathbb{Z})$.

**Theorem 23.5.** For each $k \neq 0$, the number $h(k) = |\text{SL}_2(\mathbb{Z})\setminus V_k(\mathbb{Z})|$ of equivalence classes of binary quadratic forms of fixed discriminant $k$ is finite.

**Remark.** In fact, for $k < 0$, Gauss established that for any $Q \in V_k(\mathbb{Z})$ there exists a unique equivalent form $Q'(x, y) = ax^2 + bxy + cy^2$ with $|b| \leq a \leq c$ where $b \geq 0$ if $a = c$ or $a = |b|$. Moreover, for any squarefree $k \in \mathbb{Z}$ he showed that there is a subset $H(k) \subseteq \text{SL}_2(\mathbb{Z})\setminus V_k(\mathbb{Z})$ consisting of equivalence classes of *primitive* forms (i.e. those with relatively prime coefficients) forms a finite abelian group - which turns out essentially to be the ideal class group of (an order in) $\mathbb{Q}(\sqrt{k})$.

### 23.2. Binary cubic forms.

**Definition 23.6.** An (integral) binary form of degree $d$ is a homogeneous polynomial of degree 2 in two variables:

$$f(x, y) = a_d x^d + a_{d-1} x^{d-1} y + \cdots + a_0 y^d \in \mathbb{Z}[x, y].$$

**Definition 23.7.** The discriminant of a binary form $f(x, y)$ of degree $d$ is

$$D_f = a_d^{2d-2} \prod_{i<j} (r_i - r_j)^2$$

where $r_i$ are the roots of $f(x, 1)$.

**Example 23.8.** If $f(x, y) = ax^3 + 3bx^2y + 3cxy^2 + dy^3$ with $a, b, c, d \in \mathbb{Z}$, then one can compute

$$D_f = 27(-a^2d^2 + 6abcd + 3b^2c^2 - 4ac^3 - 4db^3).$$

**Definition 23.9.** To binary forms $f$ and $g$ are *equivalent* if there exists $A \in \text{SL}_2(\mathbb{Z})$ such that

$$f(x, y) = g(x', y'), \quad \begin{bmatrix} x' \\ y' \end{bmatrix} = A \begin{bmatrix} x \\ y \end{bmatrix}.$$

**Lemma 23.10.** If $f(x, y) = ax^3 + 3bx^2y + 3cxy^2 + dy^3$ represents 1, then $f$ is equivalent to a form of the shape $g(x, y) = x^3 + 3bx^2y + 3c'xy^2 + d'y^3$.

**Proof.** If $f(x_0, y_0) = 1$, then $x_0$ and $y_0$ must be coprime, so there exists $A \in \text{SL}_2(\mathbb{Z})$ such that

$$A \begin{bmatrix} x_0 \\ y_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$ 

Then $g(x, y) = f(A^{-1}(x, y))$ satisfies $g(1, 0) = 1$, so the coefficients of $g$ have the desired form. \qed

**Definition 23.11.** Given $f(x, y) = ax^3 + 3bx^2y + 3cxy^2 + dy^3$, define

$$H^f(x, y) = -\frac{1}{36} \det \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix} \in \mathbb{Z}[x, y]$$

$$G^f(x, y) = \frac{1}{3} \det \begin{bmatrix} f_x & f_y \\ H_x & H_y \end{bmatrix} \in \mathbb{Z}[x, y]$$

binary forms of degree 2 and 3, respectively.
Remark. The action of $\text{SL}_2(\mathbb{C})$ on binary forms of degree $d$ induces a linear action on the space of coefficients $k_{a_0}^{d+1} \ldots, a_d$. A polynomial in $\mathbb{Z}[x, y, a_d, \ldots, a_0]$ is a covariant if it is invariant under the induced action of $\text{SL}_2(\mathbb{C})$. A covariant of degree 0 in $x, y$ is an invariant.

**Example 23.12.** The discriminant is an invariant of a binary form. The polynomials $H$ and $G$ are covariants of a binary cubic form.

We record the following facts about binary cubic forms, the first of which is a restatement of the fact that $H$ and $G$ are covariants of a binary cubic form.

**Theorem 23.13.** If $g(x, y) = f(A(x, y))$ for some $A \in \text{SL}_2(\mathbb{Z})$, then $H^g(x, y) = H^f(A(x, y))$ and $G^g(x, y) = G^f(A(x, y))$.

**Theorem 23.14.** Let $f$ be a binary cubic form, and let $H = H^f$, $G = G^f$, $D_1 = \frac{1}{27}D_f$. Then

\[ G^2 + D_1f^2 = 4H^3. \]

**Theorem 23.15.** For fixed nonzero integer $d$, there are only finitely many equivalence classes of integral binary cubic forms of discriminant $d$.

23.3. **Mordell’s equation.** We shall use the above facts about binary cubic forms and Thue’s theorem to prove that Mordell’s equation admits only finitely many integer solutions.

**Theorem 23.16.** Let $k \in \mathbb{Z}$ be nonzero. Then

\[ y^2 = x^3 + k \]

has at most finitely many solutions $(x, y) \in \mathbb{Z}^2$.

**Proof.** Let $f(x, y)$ be a cubic form (with middle coefficients divisible by 3) of discriminant $D_f = 27 \cdot 4 \cdot k$ and assume $f(x, y) = 1$ has a solution $(x_0, y_0) \in \mathbb{Z}^2$. Let $x = H(x_0, y_0)$ and $y = G(x_0, y_0)/2$. Then $(x, y)$ is a solution of Mordell’s equation.

Suppose $(p, q)$ is a solution to Mordell’s equation. Let

\[ f(x, y) = x^3 - 3pxy^2 + 2qy^3. \]

Then $f(1, 0) = 1$ and

\[ H^f(x, y) = -\frac{1}{36} \begin{bmatrix} 6x & 6py \\ 6py & 12qy - 6px \end{bmatrix} = px^2 - 2qxy + p^2y^2, \]

\[ G^f(x, y) = \frac{1}{3} \begin{bmatrix} 3x^2 - 3py^2 & -6pxy + 6qy^2 \\ 2px - 2qy & -2qx + 2p^2y \end{bmatrix} = 2(-qx^3 + 3p^2x^2y - 3pqxy^2 + (-p^2y)^3). \]

So we have $D_f = 27 \cdot 4 \cdot k$ and $p = H^f(1, 0)$ and $q = -G^f(1, 0)/2$. Conversely, if $f(x, y)$ is a binary cubic form (with middle coefficients divisible by 3) with discriminant $D_f = 27 \cdot 4 \cdot k$ such that $f(x, y) = 1$ admits a solution $(x_0, y_0) = 1$, then letting

\[ x = H(x_0, y_0), \quad y = -G(x_0, y_0)/2 \]

we find that $(x, y)$ is a solution of Mordell’s equation by Theorem 23.14. (Moreover, by Lemma 23.10 and completing the cube, up to equivalence we can take $f$ to be of the form $(\ast)$ above and $(x_0, y_0) = (1, 0).$) This gives us a bijection

\[ \left\{ (x, y) \in \mathbb{Z}^2 \text{ solving } y^2 = x^3 + k \right\} \leftrightarrow \left\{ (x_0, y_0) \in \mathbb{Z}^2 \text{ and } f \text{ with } D_f = 108k \text{ such that } f(x_0, y_0) = 1 \right\}/\text{SL}_2(\mathbb{Z}). \]

Now, by Theorem 23.15 the number of equivalence classes of binary cubic forms of discriminant 108$k$ is finite. For each form $f$ representing an equivalence class, the number of integer solutions for $f(x, y) = 1$ is finite by Thue’s theorem. Combining these, we get the desired result. \[\square\]