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:

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

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 integral 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.

\[ Date: \text{October 6, 2019}. \]
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. $\nabla f(P) = (\frac{\partial f}{\partial x}(P), \frac{\partial f}{\partial y}(P)) \neq 0$ 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.** Let us determine $V_f(\mathbb{Q})$ where

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

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 give us

$$x^2 + t^2(x + 1)^2 = 1 \implies (x + 1)((1 + t^2)x - (1 - t^2)) = 0.$$

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 $(a : \cdots : a_n)$.

**Example 2.** As a set,

$$\mathbb{P}^1(L) = \{[1 : a] : a \in L\} \cup \{[0 : 1]\} = L \sqcup \{\infty\}.$$

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

$$\mathbb{P}^n(L) = \{[1 : a_1 : \cdots : a_n] : a_i \in L\} \cup \{[0 : a_1 : \cdots : a_n] : a_i \in L\} = \mathbb{A}^n(L) \sqcup \mathbb{P}^{n-1}(L).$$

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

$$V_F(L) = \{(a_0 : a_1 : a_2) \in \mathbb{P}^2(L) : F(a_0, a_1, a_2) = 0\}$$

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 $V_F$ as a projective curve defined over $\mathbb{Q}$. 

3
Example 3. Let \( f(x, y) = x^2 + y^2 - 1 \) as in Example 1. The zero set \( V_F(C) \) in \( \mathbb{P}^2(C) \) of the homogenization
\[
F(X_0, X_1, X_2) = X_0^2 f(X_1/X_0, X_2/X_0) = X_1^2 + X_2^2 - X_0^2
\]
is given by the union of two sets
\[
\{(1 : a_1 : a_2) \in \mathbb{P}^2(C) : F(1, a_1, a_2) = f(a_1, a_2) = 0\} = V_f(C)
\]
and
\[
\{(0 : a_1 : a_2) \in \mathbb{P}^2(C) : F(0, a_1, a_2) = a_1^2 + a_2^2 = 0\} = \{(0 : 1 : i), (0 : 1 : -i)\}.
\]
Thus, we have
\[
V_F(C) = V_f(C) \sqcup \{(0 : 1 : i), (0 : 1 : -i)\}
\]
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
\[
\mathbb{P}^1(C) \to V_F(C) \subset \mathbb{P}^2(C)
\]
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 \( V_F \). 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 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 5.** 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 6.** 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 $n_2, n_3, n_4, \cdots \in \mathbb{Z}$ (indexed by primes, with all but finitely many $n_p$ zero) such that
   \[ n = u \prod_{p \text{ prime}} p^{n_p}. \]

**Definition 7.** For each prime $p$, the $p$-adic valuation on $\mathbb{Q}$ is the function
\[ v_p : \mathbb{Q}^{\times} \to \mathbb{Z} \]
given by $v_p(n) = n_p$. We extend $v_p$ to a function $v_P : \mathbb{Q} \to \mathbb{Z} \cup \{+\infty\}$ by setting $v_p(0) = +\infty$.

**Lemma 8.** 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 $x = p^n \frac{u}{s}$ and $y = p^m \frac{u}{v}$ with $u, v, r, s \in \mathbb{Z}$ not divisible by $p$. Without loss of generality, $v_p(x) = n \leq m = v_p(y)$. Then
\[ x + y = p^n \left( \frac{r}{s} + \frac{m-n}{v} \right) = p^n \frac{rv + pm - nus}{sv}. \]

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

**Corollary-Definition 9.** The function $| \cdot |_p : \mathbb{Q} \to \mathbb{R}_{\geq 0}$ 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 10.** 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 11.** $\mathbb{Q}$ is not complete with respect to the usual absolute value $| \cdot |_{\infty}$, but $\mathbb{R}$ is.

**Definition 12.** Two sequences $(a_i)$ and $(b_i)$ in $K$ are equivalent if $\lim_{i \to \infty} |a_i - b_i| = 0$.

**Definition 13.** The completion $\bar{K}$ 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 \( (a_i) \sim 0 \), 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 \( K \hookrightarrow \hat{K} \) 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 14.** Let \( L \) be a field complete with respect to an absolute value \( |\cdot| \), and let \( K \subseteq L \) be an arbitrary subfield (so that \( K \) inherits an absolute value from \( L \)).

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

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

**Definition 15.** 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 16.** Two absolute values \( |\cdot| \) and \( |\cdot|' \) 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 \( (K, |\cdot|) \simeq (K, |\cdot|') \).

**Theorem 17** (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 \( |b| = b^\alpha = |b|_\infty^\alpha \). 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
\[\|n\| \leq \|a_0\| + \|a_1\|\|b\| + \cdots + \|a_s\|\|b\|^s \]
\[\leq 1 + b^\alpha + \cdots + b^{\alpha s}\] (since \(\|a_i\| \leq 1\) by minimality of \(b\))
\[= (1 + b^{-\alpha} + \cdots + b^{-\alpha s})b^{\alpha s}\]
\[\leq c\alpha^n\]
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
\[\|n\| \geq \|b^{s+1}\| - \|b^{s+1} - n\| \geq b^{(s+1)\alpha} - (b^{s+1} - n)^\alpha\]
(by what we proved earlier)
\[\geq b^{((2+1)\alpha)} - b^{(s+1)\alpha} = b^{(s+1)\alpha} \left(1 - \left(1 - \frac{1}{b}\right)^\alpha\right) \geq c\alpha^n\]
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
\[\|n\| = \prod_q q^{v_q(n)} = \|p\|^{v_p(n)} = p^{-\alpha v_p(n)} = |n|_p^{\alpha} \quad \Box\]

3. Lecture 3

**3.1. Discrete valuations.**

**Definition 18.** 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 19. 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 $C(t)$ be the field of rational functions in one variable $t$ over $\mathbb{C}$. For every $a \in C$, the function
  
  $$v_a(f) = \text{order of vanishing of } f \text{ at } a \quad \text{for } f \in C(t)^\times$$

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

Definition-Lemma 20. Let $v : K \to \mathbb{Z}$ be a nontrivial discrete valuation.

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

Proof. (a) Since $v(0) = +\infty$ and $v(1) = 0$, we have $0, 1 \in O_v$. If $x, y \in O_v$ 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 $xy \in O_v$ and $x + y \in O_v$. This shows that $O_v$ is a subring of $K$.

(b) Arguing as in (a), we see that $m_v$ is an ideal of $O_v$. If $x \in O_v \setminus m_v$, 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 $1/x \in O_v$ and $x \in O_v^\times$. It follows that $O_v^\times = O_v \setminus m_v$, the inclusion $O_v^\times \subseteq O_v \setminus m_v$ being obvious. This shows that $m_v$ is the unique maximal ideal of $O_v$, as desired. (Recall that $R^\times = R \setminus \bigcup_m \{ \text{maximal ideal } m \}$ 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 $x \in m_v$,

$$v \left( \frac{x}{\pi} \right) = v(x) - v(\pi) \geq 0 \implies x/\pi \in O_v \implies x \in \pi O_v.$$ 

This shows that $m_v = \pi O_v$. 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 $v(\pi) \mid v(x)$ for every $x \in m_v$. So if $I \subseteq O_v$ is any nonzero ideal, then $\min \{ v(x) : x \in I \} = k \cdot v(\pi)$ for some $k \in \mathbb{Z}_{\geq 0}$, and $I = \pi^k O_v$ by arguing as above.

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

(d) Let $\pi \in O_v$ be a uniformizer. For any $x \in K$, we have $v(\pi^N x) = Nv(\pi) + v(x) \geq 0$ for $N \geq 0$ and hence $\pi^N x \in O_v$ and $x \in O_v[1/\pi]$.

3.2. $p$-adic integers.

Definition 21. 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 22. 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 $\mathbb{Z}_p^\times = \mathbb{Z}_p \setminus p\mathbb{Z}_p$.
3. Every $x \in \mathbb{Z}_p$ is of the form $p^n u$ with $n = v_p(x) \in \mathbb{Z}_{\geq 0}$ and $u \in \mathbb{Z}_p^\times$.

Proposition 23. 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 $a \in \mathbb{Z}_p$. 
2. (Cont'd)
(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 $p\mathbb{Z}$, 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

$$a_i = \frac{b_i}{c_i}$$

with $b_i, c_i \in \mathbb{Z}$ such that $\gcd(b_i, c_i) = 1$ and $p \nmid c_i$.

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

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

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

**Proposition 24** (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 23, there exists $b'_0 \in \mathbb{Z}$ such that $|a - b'_0|_p < 1$. Then $a = b'_0 + pa'_1$ for some $a'_1 \in \mathbb{Z}_p$. 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 $\mathbb{Q}_p = \mathbb{Z}_p[1/p]$.

\hfill \Box

**Lemma 25.** For each $n \geq 1$, the inclusion $\mathbb{Z} \to \mathbb{Z}_p$ induces a natural isomorphism

$$\mathbb{Z}/p^n\mathbb{Z} \isom \mathbb{Z}_p/p^n\mathbb{Z}_p$$

compatible with the projections

$$\begin{align*}
\mathbb{Z}/p^{n+1}\mathbb{Z} &\longrightarrow \mathbb{Z}/p^n\mathbb{Z} \\
\mathbb{Z}_p/p^{n+1}\mathbb{Z}_p &\longrightarrow \mathbb{Z}_p/p^n\mathbb{Z}_p
\end{align*}$$

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 $\mathbb{Z} \cap p^n\mathbb{Z}_p = p^n\mathbb{Z}$. \hfill \Box
3.3. \(\mathbb{Z}_p\) as an inverse limit.

**Definition 26.** An inverse system of sets is a sequence \((A_n)\) of sets and maps \((f_n)\) of the form
\[
\cdots \to A_{n+1} \xrightarrow{f_n} A_n \to \cdots \to A_2 \xrightarrow{f_1} A_1 \xrightarrow{f_0} A_0.
\]
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 \(\epsilon_n : A \to A_n\) 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 27.** The sequence of projections \(\cdots \to \mathbb{Z}/p^{n+1}\mathbb{Z} \to \mathbb{Z}/p^n\mathbb{Z} \to \cdots \to \mathbb{Z}/p^2\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z}\) forms an inverse system of rings.

**Proposition 28.** The projections \(\mathbb{Z}_p \to \mathbb{Z}_p/p^n\mathbb{Z}_p\) induce an isomorphism
\[
\mathbb{Z}_p \simeq \lim_{\leftarrow} \mathbb{Z}_p/p^n\mathbb{Z}_p \simeq \lim_{\leftarrow} \mathbb{Z}/p^n\mathbb{Z}.
\]

**Proof.** Under the natural map \(\mathbb{Z}_p \to \lim_{\leftarrow} \mathbb{Z}_p/p^n\mathbb{Z}_p\), each element \(([b_0], [b_0+b_1p], [b_0+b_1p+b_2p^2], \ldots) \in \lim_{\leftarrow} \mathbb{Z}_p/p^n\mathbb{Z}_p\) is in the image of \(\sum_{i=0}^{\infty} b_ip^i \in \mathbb{Z}_p\). 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 \(f \in \mathbb{Z}_p[x_1, \cdots, x_N]\), note that reduction modulo \(p^n\) gives us an inverse system
\[
\cdots \to V_f(\mathbb{Z}/p^{n+1}\mathbb{Z}) \xrightarrow{\text{mod } p^n} V_f(\mathbb{Z}/p^n\mathbb{Z}) \to \cdots \to V_f(\mathbb{Z}/p\mathbb{Z}).
\]

**Lemma 29.** For any \(f \in \mathbb{Z}_p[x_1, \cdots, x_N]\), we have an isomorphism
\[
V_f(\mathbb{Z}_p) \simeq \lim_{\leftarrow} V_f(\mathbb{Z}/p^n\mathbb{Z})
\]
given by \(P \mapsto (P \mod p^n)\).

**Proof.** Note that
\[
\lim_{\leftarrow} V_f(\mathbb{Z}/p^n\mathbb{Z}) \subseteq \lim_{\leftarrow} ((\mathbb{Z}/p^n\mathbb{Z})^N) \simeq (\lim_{\leftarrow} \mathbb{Z}/p^n\mathbb{Z})^N \simeq \mathbb{Z}_p^N.
\]
So for any \((P_n) \in \lim V_f(\mathbb{Z}/p^n\mathbb{Z})\) 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 \(\tilde{f}(P) \equiv f(P_n) \equiv 0 \mod p^n\) 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 30.** 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 \( T_{k,n} = \text{Im}(A_{n+k} \to \cdots \to A_n) \subseteq A_n \). 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 \( (E_n) \) 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
\[
(e_n) \in \liminf E_n \subseteq \lim A_n
\]
so in particular \( \lim \inf A_n \) is nonempty. \( \square \)

**Corollary 31.** For any \( f \in \mathbb{Z}_p[x_1, \cdots, x_n] \), 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 32.** 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 \( P = (a_1, \cdots, a_N) \) are multiples of \( p \) in \( \mathbb{Z}_p \).

**Lemma 33.** Let \( F \in \mathbb{Z}[X_0, \cdots, X_N] \) be homogeneous. The following are equivalent:

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

(b) \( V_F(\mathbb{Z}/p^n\mathbb{Z}) \) contains \( n+1 \)-tuples of points with a primitive point.

(c) \( V_F(\mathbb{Z}/p^n\mathbb{Z}) \) contains \( n+1 \)-tuples of points with a primitive point for all \( n \geq 1 \).

Proof. (a) \( \iff \) (b). If \( (a_0, \cdots, a_N) \in V_F(\mathbb{Z}_p) \) is primitive then \( (a_0, \cdots, a_N) \not\equiv 0 \) 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 \( (a_0, \cdots, a_N) \in \mathbb{Z}_p^{N+1} \) and \( m = \min\{v_p(a_0), \cdots, v_p(a_N)\} \), then we have \( (p^{-m}a_0, \cdots, p^{-m}a_N) \in \mathbb{Z}_p^N \) 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 30 and the observation that
\[
V_F(\mathbb{Z}_p)_{\text{prim}} = \liminf V_F(\mathbb{Z}/p^n\mathbb{Z})_{\text{prim}}.
\]

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

**Lemma 34** (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 \( (a_n) \in \liminf \mathbb{Z}/p^n\mathbb{Z} \) 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
\[
a_{n+1} = \bar{a}_{n+1} + p^n z
\]
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
\[
f(a_{n+1}) = f(\bar{a}_{n+1} + p^n z) = f(\bar{a}_{n+1}) + f'(\bar{a}_{n+1}) p^n z + \cdots
\]
where the remaining terms \((\cdots)\) in the expansion are divisible by \(p^{2n}\) and hence zero modulo \(p^{n+1}\). We have \(f(\bar{a}_{n+1}) = p^n\) for some \(c \in \mathbb{Z}/p\mathbb{Z}\) since \(f(\bar{a}_{n+1}) \equiv f(a_n) \equiv 0 \mod p^n\). Thus,
\[
    f(a_{n+1}) = p^n c f'(\bar{a}_{n+1}) + p^n z = p^n (c + f'(\bar{a}_{n+1}) z).
\]
Since \(f'(\bar{a}_{n+1}) \equiv f'(a) \not\equiv 0 \mod p\), it follows that \(c + f'(\bar{a}_{n+1}) z\) 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 \(\bar{a} \in \mathbb{Z}/p\mathbb{Z}\) uniquely corresponding to \((a_n) \in \lim_{\leftarrow} \mathbb{Z}/p^n\mathbb{Z}\) satisfies \(f(\bar{a}) \equiv f(a_n) \equiv 0 \mod p^n\) for every \(n \geq 1\), showing that \(f(\bar{a}) = 0\). Finally, by construction we have \(\bar{a} \equiv a \mod p\).

**Corollary 35.** Let \(f \in \mathbb{Z}[x_1, \ldots, x_n]\). If \(P \in V_f(\mathbb{Z}/p\mathbb{Z})\) and \(\nabla f(P) = (\frac{\partial f}{\partial x_1}(P), \ldots, \frac{\partial f}{\partial x_n}(P)) \neq 0\) in \(\mathbb{Z}/p\mathbb{Z}\), then there exists \(\bar{P} \in V_f(\mathbb{Z}_p)\) such that \(\bar{P} \equiv P \mod p\). In particular, \(V_f(\mathbb{Z}_p) \neq \emptyset\).

**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 \(\bar{a}_j \in \mathbb{Z}_p\) such that \(\bar{a}_j \equiv a_j \mod p\) for \(j = 2, \ldots, n\). Apply Hensel’s lemma to \(f(x, \bar{a}_2, \ldots, \bar{a}_n) \in \mathbb{Z}_p[x]\). □

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

**Proposition 36.** 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 \(V_F(\mathbb{Q}_p) \neq 0\).

**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 \(\bar{P} \in V_F(\mathbb{Z}_p)\) such that \(\bar{P} \equiv P \mod p\). Finally, note that \(\bar{P}\) is primitive since \(P\) is primitive. Thus \(V_F(\mathbb{Q}_p)\) is nonempty by Lemma 33. □

**4.3. 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 \(\mathbb{Z}_p^\times \supset U_1 \supset U_2 \supset \cdots\).

**Lemma 37.** We have the following.

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

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

**Corollary 38.** For each \(n \geq 1\), we have \(\#U_1/U_n = p^{-n-1}\).

**Proposition 39.** 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 \(\mathbb{Z}_p^\times = U_1 \times \mu_{p-1}\).

**Proof.** We have \(\mu_{p-1} = \ker(\mathbb{Z}_p^\times \xrightarrow{x^{p-1}} \mathbb{Z}_p^\times) \leq \mathbb{Z}_p^\times\). 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 \(\bar{a} \in \mu_{p-1}\) such that \(\bar{a} \equiv a \mod p\), so the composition \(\mu_{p-1} \hookrightarrow \mathbb{Z}_p^\times \to \mathbb{F}_p^\times\) is an isomorphism. Since \(\mu_{p-1} \cdot U_q = \mathbb{Z}_p^\times\) and \(\mu_{p-1} \cap U_1 = \{1\}\), we deduce that \(\mathbb{Z}_p^\times = U_1 \times \mu_{p-1}\). □

**Remark.** We have \(\mu_{p-1} \simeq \mathbb{F}_p^\times \simeq \mathbb{Z}/(p-1)\mathbb{Z}\).
Proposition 40. Suppose that \( p \) is an odd prime. Then we have an isomorphism \( \mathbb{Z}_p \cong U_1 \) 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 \( a = 1 + kp^n \) with \( k \in \mathbb{Z}_p^\times \), then we have

\[
\alpha^p = (1 + kp^n)^p = 1 + \left( \begin{array}{c} p \\ 1 \end{array} \right) kp^n + \left( \begin{array}{c} p \\ 2 \end{array} \right) k^2 p^{2n} + \cdots + k^p p^{pn} \equiv 1 + kp^{n+1} \mod p^{n+2}.
\]

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 \( u \) 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

\[
\mathbb{Z}/p^{n-1}\mathbb{Z} \cong U_1/U_n
\]

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

\[
\mathbb{Z}_p \cong \varprojlim \mathbb{Z}/p^n\mathbb{Z} \cong \varprojlim U_1/U_n \cong U_1. \quad \square
\]

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

Corollary 41. If \( p \) is an odd prime, then \( \mathbb{Z}_p^\times \cong \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}_p \).

5. Lecture 5

5.1. Quadratic forms. Let \( k \) be a field of characteristic different from 2.

Definition 42. 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 43. A quadratic form on a finite-dimensional vector space \( V/k \) is a function \( q : V \to k \) such that there is an isomorphism \( k^n \cong V \) of vector spaces \( (n = \dim V) \) so that the composition

\[
k^n \cong V \to k
\]

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 \( q(x_1 e_1 + \cdots + x_n e_n) \) is a homogeneous polynomial of degree 2 in \( x_1, \ldots, x_n \).)

Definition 44. 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 45. For each finite-dimensional vector space \( V/k \), we have a bijection

\[
\{ \text{quadratic forms on } V \} \leftrightarrow \{ \text{symmetric bilinear forms on } V \}
\]

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

\[
q(v) := B(v, v) \mapsto (B : V \times V \to k).
\]
Example 46. If \( q(x, y) = x^2 + y^2 \) is a quadratic form on \( k^2 \), the corresponding bilinear form is

\[
B((x_1, y_1), (x_2, y_2)) = \frac{1}{2} \left( (x_1 + x_2)^2 + (y_1 + y_2)^2 - (x_1^2 + y_1^2) - (x_2^2 + y_2^2) \right) = x_1 x_2 + y_1 y_2.
\]

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 47. We have

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

**Definition 48.** 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 \( q(x_1, \cdots, x_n) \) is **nondegenerate** if one of the following equivalent conditions holds:
\[
\begin{align*}
& \text{Rank of } q \text{ is } n. \\
& \text{The associated matrix } A \text{ is invertible.} \\
& \text{For every } v \in V, \text{ the linear map } B(\cdot, v) : V \to k \text{ given by } u \mapsto B(u, v) \text{ is nonzero.}
\end{align*}
\]

**Definition 49.** 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 50.** 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 51.** Every quadratic form \( q(x_1, \cdots, x_n) \) 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 \( u_1 + u_2 \in V \) with \( u_1 = x_1 v \) and \( u_2 \in v_{\perp} \), we have

\[
q(u) = B(u_1 + u_2, u_1 + u_2) = B(u_1, u_1) + B(u_2, u_2) + 2B(u_1, u_2) = q(u_1) + q(u_2)
\]
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 \( q(u_1) = a_1 x^2 \) 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 52.** A quadratic form \( q \) on \( V/k \) represents \( c \in k \) if \( q(v) = c \) for some nonzero \( v \in V \).

**Proposition 53.** 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 54** (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. \( \square \)

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

**Lemma 55.** 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 \( b \neq 0 \), whence \( b = N(\frac{a}{x} + \frac{a}{z}\sqrt{a}) \).

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, \( N(b) = b \). \( \square \)

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,

\[
x^2 - ay^2 \equiv 0 \mod p \implies x, y \equiv 0 \mod p
\]

\[
\implies bz^2 = x^2 - ay^2 \equiv 0 \mod p^2
\]

\[
\implies p \mid z \text{ (since } b \text{ is squarefree),}
\]

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 55, it follows that \(b' = bb'/b\) is a norm in \(\mathbb{Q}_p(\sqrt{a})\). In other words, the quadratic form

\[
x^2 - ay^2 - bz^2
\]

represents 0 in \(\mathbb{Q}_p\) for every \(p \leq \infty\). Since \(|a| + |b'| < |a| + |b|\), by our induction hypothesis it follows that \(x^2 - ay^2 - bz^2\) represents 0 over \(\mathbb{Q}\), or in other words \(b'\) is a norm in \(\mathbb{Q}(\sqrt{a})\) by Lemma 55. It follows that \(b = bb'/b\) is a norm in \(\mathbb{Q}(\sqrt{a})\), and by Lemma 55 we conclude that \(x^2 - ay^2 - bz^2\) represents 0 over \(\mathbb{Q}\), as desired. This completes the induction and the proof. □

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

Corollary 56. 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. □

Example 57. 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 \(4^m u\) with \(m \in \mathbb{Z}\) and \(u \in 7 + 8\mathbb{Z}_2\).

Proof. Let \(a \in \mathbb{Q}^\times\) be given. By the Hasse-Minkowski theorem (Corollary 56), \(x^2 + y^2 + z^2\) represents a over \(\mathbb{Q}\) if and only if it represents a over \(\mathbb{Q}_p\) for all \(p \leq \infty\).

By Proposition 36 (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 53. 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 \(4^m u\) 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 \(u \in \mathbb{Z}_2^\times\) is a square if and only if it belongs to \(1 + 8\mathbb{Z}_2\). This shows that the image of \(\mathbb{Z}_2^\times \times \mathbb{Z}_2 \times \mathbb{Z}_2\) 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 \(\{1 + a + b : a, b \in \{0, 1, 2, 4 \mod 8\}\}\) \(= \{1, 2, 3, 5, 6 \mod 8\}\) in \(\mathbb{Z}/8\mathbb{Z}\), 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 \((x, y, z) \in \mathbb{Q}^3_2\) is obtained from a primitive triple \((x, y, z) \in \mathbb{Z}^3_2\) 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 \(a = 4^n a'\) with \(n \in \mathbb{Z}\) and \(v_2(a') = 0\) or 1 (whence \(a' \equiv 1, 2, 3, 5, \text{ or } 6 \mod 8\)); our work so far shows that \(a'\) (and hence \(a\)) is represented by \(x^2 + y^2 + z^2\) \(\mathbb{Q}_2\), as desired. This proves the result. \(\square\)

**Theorem 58 (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 59.** 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 \(b \in B^\times\) such that 
\[
ba^{-1} = \alpha_0
\]
belongs to the maximal order (analogous to the ring of integers \(\mathbb{Z}[i]\) of \(\mathbb{Q}(i)\))
\[
\mathbb{Z} + \mathbb{Z} i + \mathbb{Z} j + \mathbb{Z} \frac{1 + i + j + k}{2}.
\]

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
\[
\alpha_0 = x_0i + y_0j + z_0k
\]
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 \(\vec{v} = (x, y, z)\) and
\[
\vec{v} = \vec{q} + \frac{1}{d} \vec{r}
\]
where \(\vec{q} = (q_1, q_2, q_3)\) and \(\vec{r} = (r_1, r_2, r_3) \neq 0\). Now,
\[
\vec{v} \cdot \vec{v} = \left(\vec{q} + \frac{1}{d} \vec{r}\right) \cdot \left(\vec{q} + \frac{1}{d} \vec{r}\right) = \vec{q} \cdot \vec{q} + \frac{2}{d^2} \vec{q} \cdot \vec{r} + \frac{1}{d^2} \vec{r} \cdot \vec{r}.
\]
Multiplying both sides by \(d/(\vec{r} \cdot \vec{r})\), we get
\[
\frac{d(\vec{v} \cdot \vec{v})}{\vec{r} \cdot \vec{r}} = \frac{d(\vec{q} \cdot \vec{q})}{\vec{r} \cdot \vec{r}} + \frac{2 \vec{q} \cdot \vec{r}}{\vec{r} \cdot \vec{r}} + \frac{1}{d}.
\]
Now, consider the vector \( \vec{w} = \tau_{\vec{r}}(\vec{v}) \) 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

\[
\vec{w} = \tau_{\vec{r}}(\vec{v}) = \tau_{\vec{r}}\left(\vec{q} + \frac{1}{d}\vec{r}\right) = \vec{q} - \frac{2\vec{q} \cdot \vec{r}}{\vec{r} \cdot \vec{r}} \cdot \frac{1}{d}\vec{r} = \vec{q} - \frac{d(\vec{v} \cdot \vec{v} - \vec{q} \cdot \vec{q})}{\vec{r} \cdot \vec{r}} \vec{r} = \vec{q} - \frac{\vec{v} \cdot \vec{v} - \vec{q} \cdot \vec{q}}{\vec{r} \cdot \vec{r} / d} \vec{r}.
\]

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,

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

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 \( \mathbb{Z} \). \( \square \)

**Corollary 60** (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 \( a \). \( \square \)

**Corollary 61** (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 62.** 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 \( f(x) = 0 \); 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 63. 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, if $T = \{t_1, \ldots, t_n\} \subseteq L$ is algebraically independent over $k$, then $n \leq m$. Note that, by maximality of $S$, we see $\{t_1, s_1, \ldots, s_m\}$ is algebraically dependent over $k$. Since $t_1$ is transcendental over $k$, without loss of generality we may thus assume $s_1$ is algebraic over $k(S_1)$ where $S_1 = \{s_2, \ldots, s_m\}$. 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 64. 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 65. The affine space $\mathbb{A}_k^n$ over $k$ is (given by) the assignment

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

for every field extension $L/k$.

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

Definition 66. For any subset $S \subseteq k[x_1, \ldots, x_n]$, the associated (closed) affine variety $V_S \subseteq \mathbb{A}_k^n$ is (given by) the assignment

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

for every field extension $L/k$.

Note. We note the following.

1. If $(S) \leq k[x_1, \ldots, x_n]$ 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 $g_i \in k[x_1, \ldots, x_n]$ 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 $V_{(0)} = \mathbb{A}_k^n$ and $V_{(1)} = \emptyset$.

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

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

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

Definition 67. If $I \leq R$ 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 $I \leq k[x_1, \ldots, x_n]$. 

19
(2) Let us choose an algebraic closure $\overline{k}$ of $k$. If $V \subseteq \mathbb{A}^n_k$ is an affine variety over $k$, then
\[ I(V) := \{ f \in k[x_1, \ldots, x_n] : f(P) = 0 \text{ for all } P \in V(\overline{k}) \} \]
is a radical ideal in $k[x_1, \ldots, x_n]$. (Note $\overline{k}$.) We remark that $I(V)$ is independent of the choice of the algebraic closure $\overline{k}$.

**Theorem 68** (Hilbert Nullstellensatz). There is an order-reversing bijection
\[ \{ \text{Radical ideals } I \leq k[x_1, \ldots, x_n] \} \leftrightarrow \{ \text{Affine varieties } V \subseteq \mathbb{A}^n_k \text{ over } k \} \]
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 $I \leq k[x_1, \ldots, x_n]$ is an ideal then $I(V_I) = \sqrt{I}$. So let $I \leq k[x_1, \ldots, x_n]$ 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 $P \in \mathbb{A}^n_k(\overline{k})$ with $f_1(P) = \cdots = f_m(P) = 0$. (Here, we have fixed an algebraic closure $\overline{k}$ of $k$). Then $f_1, \ldots, f_m, 1-x_0f \in k[x_0, \ldots, x_m]$ (polynomials in $n+1$ variables) have no common zero in $\overline{k}$. We now cite the following commutative algebra fact:

**Fact 69.** 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 $(f_1, \ldots, f_m, 1-x_0f)$ is not contained in any maximal ideal of $k[x_0, \ldots, x_n]$ so it must be $(1)$. Indeed, suppose $m$ is a maximal ideal of $k[x_0, \ldots, x_n]$ containing $(f_1, \ldots, f_m, 1-x_0f)$. Then since the quotient field
\[ L = k[x_0, \ldots, x_n]/m \]
is finitely generated as a $k$-algebra it follows that $L/k$ is algebraic, and without loss of generality we may assume $L \subseteq \overline{k}$. 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 $k[x_0, \ldots, x_n] \to k[x_0, \ldots, x_n]/m = L$, then $f_1(a_1, \ldots, a_n) = \cdots = f_m(a_1, \ldots, a_n) = 1-a_0f(a_1, \ldots, a_n)$ is not zero, contradicting the fact that $f_1, \ldots, f_m, 1-x_0f$ have no common zeros in $\overline{k}$. Thus there exist $g, g_1, \ldots, g_m \in k[x_0, x_1, \ldots, x_n]$ such that
\[ 1 = g(1 - x_0f) + \sum_{i=1}^m g_i f_i. \]

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 $g_i$s, we have
\[ 1 = \frac{\sum_i \tilde{g}_i f_i}{f^r} \]
for some $r \geq 0$ and $\tilde{g}_i \in k[x_1, \ldots, x_n]$, so $f^r = \sum_i \tilde{g}_i f_i \in I$. 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. \(\square\)
Fix a field $k$ and algebraic closure $ar{k}$. Recall from last time:

**Theorem 70** (Hilbert Nullstellensatz). There is an order-reversing bijection

\[
\{\text{Radical ideals } I \subseteq k[x_1, \ldots, x_n]\} \leftrightarrow \{\text{Affine varieties } V \subseteq \mathbb{A}^n_k \text{ over } k\}
\]

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 71.** If $V, W \subseteq \mathbb{A}^n$ are affine varieties over $k$ with $V(\bar{k}) = W(\bar{k}) \subseteq \mathbb{A}^n_k(\bar{k})$, 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 $L$-points of $V$ where $L$ 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 learn 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 72.** 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 73.** 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 74.** 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

\[\bigcap_{\alpha \in A} V_{I_\alpha} = V_{\bigcup_{\alpha \in A} I_\alpha}\]
Definition 75. 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 $P \in V(k)$ 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 76. 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 $\bigcap_{i=1}^r V_{\alpha_i} = \emptyset$. $\Box$

8.2. Coordinate ring.

Definition 77. 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 $V \subseteq \mathbb{A}^n$ 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 $V \subseteq \mathbb{A}^n$. In fact, we have the following.

Proposition 78. If $V \subseteq \mathbb{A}^n_k$ is an affine variety over $k$, we have a natural bijection

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

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 $V = V(I(V))$ 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. $\Box$
8.3. Irreducibility and dimension.

**Definition 79.** 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 80.** 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 $I(V) = \{f \in k[x_1, \ldots, x_n] \mid f \text{ is not a unit} \}$ 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 $f \notin I(V)$. 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 $f, g \in k[x_1, \ldots, x_n]$ with $fg \in I(V)$. Let $I_1 = I(V) + (f)$ and $I_2 = I(V) + (g)$. Then $I_1I_2 \subseteq I(V)$, 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 $k$, we must have $V_{I_1} = V$ or $V_{I_2} = V$. Without loss of generality, say $V_{I_1} = V$. Then $\sqrt{I(V) + (f)} = \sqrt{I_{I_1}} = I(V)$. 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. \hfill $\square$

**Definition 81.** In general, an affine variety $V \subseteq \mathbb{A}^n$ over $k$ is a finite union of irreducible affine varieties over $k$. We can write

$$V = V_1 \cup \cdots \cup V_r$$

where each $V_i$ is an irreducible affine variety over $k$ 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 82.** $V_{(xy)} \subseteq \mathbb{A}^2$ 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 83.** 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 84.** Let $V \subseteq \mathbb{A}^n$ be an irreducible variety over $k$. The function field of $V/k$ is

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

**Fact 85.** Suppose that $V \subseteq \mathbb{A}^n$ is an irreducible affine variety over $k$. 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 86.** 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 $\mathbb{P}(L)$ is denoted $(a_0 : \ldots : a_n)$.

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 87.** 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 88.** An ideal $I \leq k[X_0, \ldots, X_n]$ 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 $\mathbb{V}_S = \mathbb{V}(S)$.
2. The ideal $m = (X_0, \ldots, X_n) \leq k[X_0, \ldots, X_n]$ is called the irrelevant ideal. We have $\mathbb{V}_m = \emptyset$.
3. If $I \leq k[X_0, \ldots, X_n]$ is homogeneous, then its radical $\sqrt{I}$ is homogeneous. (This will be an exercise in Problem Set 3.) Moreover, $\mathbb{V}_I = \mathbb{V}_{\sqrt{I}}$.
4. Fix an algebraic closure $\bar{k}$ 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 $P \in V(k)$. Then $I(V)$ is a homogeneous radical ideal of $k[X_0, \ldots, X_n]$.

**Theorem 89** (Projective Nullstellensatz). *We have an inclusion-reversing bijection

$$\{\text{Radical homogeneous ideals } I \leq k[X_1, \ldots, X_n] \text{ not containing } (X_0, \ldots, X_n)\}$$

$$\leftrightarrow \{\text{Projective varieties } V \subseteq \mathbb{P}^n \text{ over } k\}$$

given by $I \mapsto \mathbb{V}_I$ 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 $V(k)$ of $k$-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 90.** 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 91.** Let $V \subseteq \mathbb{P}^n$ be a projective variety over $k$. The following are equivalent:

1. $V(\bar{k})$ 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. □
9.2. Affine patches and projective closures. Given $\mathbb{P}^n$, the standard hyperplane $H_i \subset \mathbb{P}^n$ is given by $H_i = \mathcal{V}(x_i)$, i.e.

$$H_i(L) = \{(a_0 : \cdots : a_n) \in \mathbb{P}^n(L) : a_i = 0\}$$

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 92.** 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 93.** 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 $V \subseteq \mathbb{A}^n = U_i \subseteq \mathbb{P}^n$ 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 $V$.

**Proposition 94.** 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 $P = (a_0, \ldots, a_{i-1}, a_{i+1}, \ldots, a_n) \in \mathbb{A}^n(\overline{k}) = U_i(\overline{k})$. 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

$$N = \max\{m \deg f, \deg h_1 + \deg g_1, \ldots, \deg h_r + \deg g_r\}.$$

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

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

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 $V \cap U_i = V$ by minimality of $V$. Then the homogenization $F_0 \in I^h$ of $f$ satisfies $F = X^{\deg F - \deg F_0} F_0$, so $F \in I^h$. □

**Example 95.** 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

$$Y^2 Z = X(X - Z)(X - 3Z).$$

**Definition 96.** 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.)