ALGEBRAIC GEOMETRY
JAN 04

Contents

0.1. **The course 797W, Algebraic geometry.** This course is an introduction to the *vocabulary* and *methods* of algebraic geometry, geared towards the use of algebraic geometry in various areas of mathematics: number theory, representation theory, combinatorics, mathematical physics.

The basic vocabulary will evolve from systems of polynomial equations to algebraic varieties and schemes. We will also get introduced to the topology of algebraic varieties: cohomology and algebraic cycles.

0.1.1. *Examples.* Our first example will be the *algebraic curves.* This is the best understood part of algebraic geometry since it deals with one-dimensional objects. The highlights: Riemann Roch theorem and the relation to number theory ("geometric class field theory").

The second example are the *flag varieties*, i.e., the homogeneous compact algebraic varieties. We will consider maps of curves into flag varieties, a topic of current interest related to mathematical physics ("quantum cohomology") and representation theory.

0.1.2. *Text.* A source with classical and "elementary" flavor is:

Shafarevich Igor R., Basic algebraic geometry (Springer-Verlag).

1. Varieties in projective space, and 2. Schemes and complex manifolds. There is a soft-cover as well as the hard-cover edition.

A (more) modern treatment is:

Hartshorne Robin, Algebraic geometry
(Graduate Texts in Mathematics, No. 52. Springer-Verlag).

I also hope to type course notes.

0.1.3. *Prerequisites.* Some familiarity with algebra on the level of 611 will be helpful.

Part 0. Idea of space (*A picture book of algebraic geometry*)

This is an introductory part. After this we start from the beginning and develop ideas slowly, clearly, precisely.

The moral here will be that for a given problem you may want to find the notion of a geometric space that will be useful (set, manifold, algebraic variety, scheme, ...).

One of the characteristics of algebraic geometry is that it has gone particularly far in developing more and more abstract notions of space that are on the other hand *useful*, i.e., the new way of thinking solves old problems. I want to mention various notions of space that are used in today's mathematics, and why these classes of spaces were introduced, what do they do for us. However, this "big picture" is likely to be much bigger then this course.

0.2. **Formation of spaces useful for a given problem.** The idea is that in a given problem you may want to understand some object X from a *geometric point of view* , and for this you may want to encode some of its properties into saying that X is a **geometric space of the kind $\mathfrak{X}$**. Here, $\mathfrak{X}$ could be something like

- Set,
- Topological space,
- Manifold,
- Algebraic variety,
- Scheme,
- Stack or n-stack,
- Differential Graded Scheme
- Non-commutative space a la Connes, etc.

Rather then going through encyclopedia of definitions, let us try to see the *principles* which historically pushed the introduction of certain classes of spaces. Some ideas we will emphasize:

- *Observation Principle: Space is what you observe.* It will lead to thinking of a space X in terms of the algebra of functions on X, hence to use of ALGEBRA in GEOMETRY.
- *Understanding solutions of systems of polynomial equations.* This is the origin of algebraic geometry.
- *Stability of the set of solutions under perturbations of the system.* This wish leads to the use of algebraically closed fields, projective spaces, infinitesimals, homological algebra, etc.
- *Formation of moduli.* We want to make the set of isomorphism classes of objects of a certain kind into a geometric space. This is roughly the same question as being able to make quotients X/G of spaces by groups that act on them.

- *Include the number theory.* This wish forces us to give a geometric meaning to all commutative rings.

0.3. **Space is what you observe.** First, we are likely to imagine a space consisting of points and so it is a *set of points.*] However we are usually interested in situations that have more organization then just a set. The simplest form of additional organization may be *topology*, i.e., a vague prescription of what is close to what. In the next step we often use the *Observation Principle*

we think of X as a space of kind $\mathfrak{X}$ if on X we observe objects of class $\mathfrak{X}$.

This is a part of terminology in physics: we study a system through *observables*, i.e., things that can be observed, i.e., measured. An observable on our object X will be some kind of a function on X so that it can be measured at each point.

For instance on the real line $\mathbb{R}$ we have studied

- All $\mathbb{R}$-valued functions,
- Continuous functions $C(\mathbb{R})$,
- Smooth (infinitely differentiable) functions $C^\infty(\mathbb{R})$,
- Polynomials $\mathcal{O}(\mathbb{R}) = \mathbb{R}[x]$,
- Analytic functions $\mathcal{O}_{an}(\mathbb{R})$

etc, and then $\mathfrak{X}$ depended on what functions we were interested in – we would think of $\mathbb{R}$ respectively as a set, topological space, manifold, algebraic variety, analytic manifold. On a plane $\mathbb{R}^2$ we would also have holomorphic functions $\mathcal{O}_{an}(\mathbb{C})$ and Holomorphic polynomials $\mathcal{O}(\mathbb{C}) = \mathbb{C}[z]$, so we could think of it as 2d real manifold or a 1d complex manifold, etc.

0.4. **Algebraic Geometry: combine A and G.** So we will view X as a space of kind $\mathfrak{X}$ if on X we can observe *observables* (functions) of kind $\mathfrak{X}$. If our observables are functions $\mathcal{O}(X)$ on X with values in a ring $\Bbbk$ (something like $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$), then $\mathcal{O}(X)$ is a ring (one adds and multiplies functions pointwise). This puts us (in a very general sense) in *Algebraic Geometry*, since we can combine the geometric understanding of X with the algebraic analysis of the ring $\mathcal{O}(X)$.

0.5. **Local and global.** One of the fundamental geometric ideas is the *Relation of local and global objects*. For instance the analysis on $\mathbb{R}^n$ or $\mathbb{C}^n$ is the local precursor of the global subject of analysis on manifolds. In fact, even the notion of a manifold (a "global object"), is obtained by gluing together some open pieces of $\mathbb{R}^n$ or $\mathbb{C}^n$ (the "local pieces" of our LEGGO game).

0.5.1. In algebraic geometry one often introduces a class $\mathcal{C}$ of spaces in two stages.

(1) The *affine $\mathcal{C}$-spaces* X are the ones that are completely controlled by the algebra of ("global") functions $\mathcal{O}(X)$ on X.

From this point of view defining the class of *affine $\mathcal{C}$-spaces* is the same as defining a certain class $\mathcal{A}$ of commutative rings: the rings which appear as rings of functions on affine $\mathcal{C}$-varieties.

(2) Now the class $\mathcal{C}$ is defined as the class of spaces obtained by gluing together the affine $\mathcal{C}$-spaces.

So, general $\mathcal{C}$-spaces are "global" objects obtained by gluing together several affine $\mathcal{C}$-spaces, so we consider from this point of view the affine $\mathcal{C}$-spaces as the "local version" of the notion of the $\mathcal{C}$-space.

Two examples of this philosophy are the notions of Algebraic Varieties and Schemes.

0.5.2. *Affine, Projective, Quasiprojective and Algebraic varieties.* The class of spaces here is the class $\mathcal{A}lg\mathcal{V}ar_{\Bbbk}$ of Algebraic Varieties over an algebraically closed field $\Bbbk$. The summary bellow will only make sense later.

- One starts with the local version, the class $\mathcal{A}ff\mathcal{V}ar_{\Bbbk}$ of Affine Varieties (short for: Affine Algebraic Varieties). It consists of subsets of affine spaces $\mathbb{A}^n(\Bbbk) = \Bbbk^n$ given by systems of polynomial equations.[1]
- Now $\mathcal{A}lg\mathcal{V}ar_{\Bbbk}$ consists of spaces that have a finite open cover by affine varieties: $X = U_1 \cup \cdots \cup U_n$.
- Projective Varieties form a subclass $\mathcal{P}roj\mathcal{V}ar_{\Bbbk}$ of $\mathcal{A}lg\mathcal{V}ar$ that one can describe directly as subsets of projective spaces $\mathbb{P}^n(\Bbbk)$ given by systems of *homogeneous* polynomial equations.
- The class $q\mathcal{P}roj\mathcal{V}ar_{\Bbbk}$ of Quasiprojective Varieties consists of all open subvarieties of Projective Varieties. One has

$$\mathcal{P}roj\mathcal{V}ar_{\Bbbk} \ \cup\ \mathcal{A}ff\mathcal{V}ar_{\Bbbk} \subseteq q\mathcal{P}roj\mathcal{V}ar_{\Bbbk} \subseteq \mathcal{A}lg\mathcal{V}ar_{\Bbbk}.$$

  This is the most useful generality.

0.5.3. *Schemes.* First, *affine schemes* are the class of geometric spaces that corresponds to *all commutative rings.* Then, *schemes* are spaces that have an open cover by affine schemes.

However, the main point here will be the first step: finding a geometric way to think of all commutative rings. We will start with the in examples like

(1) the dual numbers $\Bbbk[X]/X^2$,
(2) formal power series $\Bbbk[[X]]$,
(3) integers $\mathbb{Z}$.

[1] As we will see, the corresponding class of commutative rings are the finitely generated $\Bbbk$-algebras without nilpotents.

1. **Algebraic Varieties**

Algebraic geometry historically started with polynomial functions on affine spaces $\mathbb{A}^n$.

1.0.4. *Affine spaces* $\mathbb{A}^n$. We start with a commutative ring $\Bbbk$ (something like $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$) and define the n-dimensional affine space $\mathbb{A}^n = \mathbb{A}^n(\Bbbk)$ as the set $\Bbbk^n$ of n-tuples of numbers from $\Bbbk$, with the ring of functions $\mathcal{O}(\mathbb{A}^n) = \Bbbk[X_1, ..., X_n]$ given by the polynomial functions.

1.0.5. *Affine algebraic varieties.* An *affine algebraic variety* X over $\Bbbk$[2] is a subset X of some $\mathbb{A}^n(\Bbbk)$ that can be described by several polynomial equations

$$X = \{a = (a_1, ..., a_n) \in \Bbbk^n;\ 0 = F_j(a),\ 1 \leq j \leq m\}$$

for some polynomials $F_j \in \mathcal{O}(\mathbb{A}^n)$. The definition offers at least three points of view on affine algebraic varieties

(1) Sets: X is a subset of $\mathbb{A}^n(\Bbbk)$,
(2) Algebra: On X one naturally has a $\Bbbk$-algebra $\mathcal{O}(X)$ of "polynomial functions on X", which one define as all restrictions of polynomials to X:

$$\mathcal{O}(X) \stackrel{\text{def}}{=} \{f|X;\ f \in \mathcal{O}(\mathbb{A}^n)\}.$$

(3) System of polynomial equations: X is described by equations $F_j = 0,\ 1 \leq j \leq m$.

1.0.6. *Varieties and schemes.* In the world of *algebraic varieties*, the first point is basic. We use algebra but when it gives different picture from sets, we adjust it to fit the sets.

In 1950s, Grothendieck discovered that varieties lie in the next world, the larger world of *schemes.* Here one trusts algebra completely and when differences arise, we massage the set theory.

We will spend most time on varieties and just rudiments of schemes, because schemes become useful when one finds difficulties in working with varieties.

1.1. **Relations between algebraic varieties are reflected in algebras of functions.**

For an affine algebraic variety $X \subseteq \mathbb{A}^n$, inclusion $X \subseteq \mathbb{A}^n$ is reflected in the restriction morphism of algebras $\mathcal{O}(\mathbb{A}^n) \xrightarrow{\rho} \mathcal{O}(X),\ \rho(f) = f|X$. Its kernel is the ideal $I_X \subseteq \mathcal{O}(\mathbb{A}^n)$ that consists of all polynomials that vanish on X. For instance I_X contains the defining equations F_j.

Since the restriction map is surjective by the definition of $\mathcal{O}(X)$, we find that the functions on X are described by

$$\mathcal{O}(X) = \mathcal{O}(\mathbb{A}^n)/I_X.$$

[2]Actually this is the standard terminology only if $\Bbbk$ is an algebraically closed field.

Example: Circles in $\mathbb{A}^2$. For instance, consider the "circle"

$$X = S_{a,b}(r) \stackrel{\text{def}}{=} \{(x,y) \in \mathbb{k}^2;\ (x-a)^2 + (y-b)^2 = r^2\} \subseteq \mathbb{A}^2,$$

In this case, the ideal $I_X \subseteq \mathcal{O}(\mathbb{A}^2)$ is generated by the above defining function $F = (X - a)^2 + (Y-b)^2 - r^2$, i.e., $I_X = \mathbb{k}[X,Y]/F{\cdot}\mathbb{k}[X,Y]$. So, the $\mathbb{k}$-algebra of functions on X has two generators X, Y related by one relation $(X-a)^2 + (Y-b)^2 - r^2 = 0$. So, $\mathcal{O}(X)$ has a basis

$$X^j Y^j,\ 0 \le i,\ 0 \le j \le 1.$$

1.1.1. *Maps of varieties and maps of algebras.* To any map of varieties $f : X \to Y$ there corresponds a morphism of algebras of functions in the opposite direction

$$\mathcal{O}(Y) \xrightarrow{f^*} \mathcal{O}(X),$$

given by the *pull-back of functions*, i.e.,

$$f^*(\phi) = \phi \circ f.$$

Actually, this gives an identification

$$Map(X,Y) \ni f \mapsto f^* \in \mathrm{Hom}_{\mathbb{k}-alg}[\mathcal{O}(Y), \mathcal{O}(X)].$$

Example: maps into affine spaces. A map $f : Y \to \mathbb{A}^n$ consists of n component functions $f = (f_1, ..., f_n),\ f_i \in \mathcal{O}(Y)$. The corresponding map $\mathbb{k}[x_1, ..., x_n] = \mathcal{O}(\mathbb{A}^n) \xrightarrow{f^*} \mathcal{O}(X)$, sends generator x_i to $f^* x_i = x_i \circ f = f_i$. So,

The dictionary between maps of varieties $f : Y \to \mathbb{A}^n$ and morphism of algebras $\mathbb{k}[x_1, ..., x_n] = \mathcal{O}(\mathbb{A}^n) \xrightarrow{F} \mathcal{O}(X)$, is:

- *F gives $f = (F(x_1), ..., F(x_n))$, and*
- *f gives F such that $F(x_i)$ is the i^{th} component function f_i of f.*

1.1.2. *Constructions in geometry and algebra.* Set theoretic operations have algebraic incarnations. For $X, Y \subseteq \mathbb{A}^n$, the equations of the *intersection* $X \cap Y$ are obtained by taking the union of equations of X and of Y, for the algebras it will turn out to involve the operation of tensoring

$$\mathcal{O}(X \cap_{\mathbb{A}^n} Y) = \mathcal{O}(X) \otimes_{\mathcal{O}(\mathbb{A}^n)} \mathcal{O}(Y).$$

The equations of the *union* $X \cup Y$ are obtained by multiplying the equations of X and of Y.

2. Stability of solutions (intersections)

By the stability of the intersections of algebraic varieties $X, Y \subseteq \mathbb{A}^n$ we mean that a small motion (perturbation) should usually cause no change in the nature of the intersection.

2.0.3. *Intersections of circles.* Our motivation in this section comes from intersecting lines and circles in an affine plane $\mathbb{A}^2(\Bbbk)$.

We start with the most familiar $\Bbbk = \mathbb{R}$. If two circles $X = \{f = 0\}$ and $Y = \{g = 0\}$ in $\mathbb{A}^2(\mathbb{R}) = \mathbb{R}^2$ meet, they are likely to meet in two points. If we move them a little, they still meet in two points. However, if we move them more, we get two more behaviors:

$$X \cap Y = \begin{cases} \text{two points,} \\ \text{one point,} \\ \text{no points;} \end{cases}$$

so our stability seems to break.

2.1. Passage to algebraically closed fields.

Why is it that from the situations of having nonempty intersection we get to empty intersection; i.e. from having solutions to the system $f = g = 0$, to no solutions?

The first observation is that such things happen in a simpler case, for $c \in \mathbb{R}$

$$Z_c = \{x \in \Bbbk = \mathbb{R};\ x^2 = c\} \subseteq \mathbb{A}^1(\mathbb{R}) = \mathbb{R}.$$

While there are two points in Z_c (i.e., two solutions) for $c > 0$, there are none for $c < 0$. The problem is familiar: $\mathbb{R}$ is not algebraically closed, i.e. there are polynomial equations over $\mathbb{R}$ that have no solutions over $\mathbb{R}$.

This historically led to the introduction of complex numbers, and it turns out that passing from $\mathbb{R}$ to $\mathbb{C}$ increases the stability solves our problem: for generic circles X and Y in $\mathbb{A}^2(\mathbb{C})$, the intersection of X and Y consists of two points.

2.1.1. *Generic point.* Here, *generic* means "not in a very special position"[3]

2.1.2. *Fewer exceptions over* $\mathbb{C}$. Over $\mathbb{C}$ there are still exceptions:[4]

- (i) $X \neq Y$ but the centers are the same,
- (ii) X and Y are tangent at one point,

[3] It is a standard idea in geometry but it is not easy to give it a precise general meaning. Grothendieck did it elegantly.

[4] So many exceptions! So, did we improve the situation by passing from $\mathbb{R}$ to $\mathbb{C}$? Yes, because in some sense over $\mathbb{R}$ there are as many bad positions as good, but over $\mathbb{C}$ the bad ones are a thin subset of all positions! Do you see this? The same behavior happens in a simpler situation:

The set $U(\mathbb{R})$ of $c \in \mathbb{R}$ such that $X^2 = c$ has two solutions over $\mathbb{R}$ is $[0, +\infty)$, and it is of the same size as its complement (the bad c's). However, the set $U(\mathbb{C})$ of $c \in \mathbb{C}$ such that $X^2 = c$ has two solutions over $\mathbb{C}$ has a small complement $\{0\}$.

- (iii) $X = Y$

Later we will come back and resolve even these exceptions.

2.1.3. *The moral.* It is easier to work over an algebraically closed field. Even if you are interested in what happens over $\mathbb{R}$, you may get the basic orientation by first understanding the solution over $\mathbb{C}$, and then you check what part of the solution appears over $\mathbb{R}$.

2.2. **Passage from affine varieties to projective varieties.** Two lines in $\mathbb{A}^2(\mathbb{R})$ are likely to meet in one point, however they may be parallel. In practice, this makes reasoning more complicated since in a situation with a bunch of lines we need to discuss various cases when some of them are parallel.

2.2.1. *Making parallel lines meet.* One can try to solve this by following the railroad track intuition: two parallel lines in a plane should meet, though only at ∞. So we try passing to a larger space then $\mathbb{A}^2(\mathbb{R})$ by adding something at ∞ of $\mathbb{A}^2$. (We hope that our problem is: "$\mathbb{A}^2$ has a hole at ∞".)

What should we add? If we add just one point, $\mathbb{R}^2 \cup \{\infty\}$, then all lines should go through it and the size of $L_1 \cap L_2$ could be 2. Not good, the infinite points of lines that meet in $\mathbb{A}^2$ should be different. So we add one line for each class of parallel lines. Since each such class contains precisely one line through the origin, we can say this in a simpler way: we add one point per each line through the origin.

2.2.2. *Projective spaces* $\mathbb{P}(V)$. For a vector space V over a field $\mathbb{k}$ we denote by $\mathbb{P}(V)$ the set of lines through the origin, i.e., the 1-dimensional vector subspaces. With this notation, we are passing from $\mathbb{A}^2$ to $\mathbb{A}^2 \sqcup \mathbb{P}(\mathbb{k}^2)$.

This actually works for any affine space $\mathbb{A}^n$: we can add $\mathbb{P}^(\mathbb{k}^n)$ and think of this as adding one point per each class of parallel lines in $\mathbb{A}^n$. This turns out to work beautifully – the new object is natural (i.e., it does not have to be explained starting from $\mathbb{A}^n$).

2.2.3. *Lemma.* $V \sqcup \mathbb{P}(V) \ \cong \mathbb{P}(V \oplus \mathbb{k})$.

2.2.4. *Projective coordinates.* We first introduce the "*projective coordinates*" on $\mathbb{P}(V)$. A basis e_i of V gives coordinates x_i on V and we denote the line through a vector $x = (x_1, ..., x_n)$ by $k{\cdot}x = [x_1 : \cdots : x_n]$. Then:

- A line $[x_1 : \cdots : x_n]$ is given when not all x_i are 0.
- Multiplying all projective coordinates by the same scalar $c \in \mathbb{k}^* \stackrel{\text{def}}{=} \mathbb{k} \backslash \{0\}$, does not change the line: $[cx_1 : \cdots : cx_n] = \ [x_1 : \cdots : x_n]$.

2.2.5. *Proof.* Coordinates $x_1, ..., x_n$ on V give coordinates $x_0, x_1, ..., x_n$ on $\Bbbk \oplus V$. Now, $\mathbb{P}(k \oplus V)$ breaks into a subset $x_0 = 0$ which is really $\mathbb{P}(V)$, and a subset $x_0 \neq 0$, which is isomorphic to V (all lines here have unique presentation of the form $[1; x_1 : \cdots : x_n]$ with $x_i \in \Bbbk$ arbitrary). **QED**

2.2.6. $\mathbb{P}^n$. By the "n-dimensional projective space" (over $\Bbbk$), we mean

$$\mathbb{P}^n \stackrel{\text{def}}{=} \mathbb{P}(\Bbbk^{n+1}).$$

By the lemma,the completion of $\mathbb{A}^n$ obtained by making the parallel lines meet at ∞ is just $\mathbb{A}^n \sqcup \mathbb{P}^{n-1} = \mathbb{P}^n$.

2.2.7. *Corollary.* $\mathbb{P}^n \cong \mathbb{A}^n \sqcup \mathbb{A}^{n-1} \sqcup \cdots \sqcup \mathbb{A}^1 \sqcup \mathbb{A}^0$.

Here, the embedding of $\mathbb{A}^i$ into $\mathbb{P}^n$, given by a repeated use of the lemma sends a point $(b_1, ..., b_i) \in \mathbb{A}^i$ to a point $[0 : \cdots : 1 : b_1 : \cdots : b_i]$ in $\mathbb{P}^n$.

2.2.8. *Projective algebraic varieties.* A *projective algebraic variety* Y over $\Bbbk$ is a subset Y of some $\mathbb{P}^n(\Bbbk)$ that can be described by several *homogeneous* polynomial equations

$$Y = \{(x_0 : \cdots : x_n) \in \mathbb{P}^n(\Bbbk);\ 0 = G_j(x_0, ..., x_n),\ 1 \leq j \leq m\}$$

for some *homogeneous* polynomials $G_j \in \Bbbk[X_0, ..., X_n]$.

2.2.9. *Scarcity of functions.* Observe that in $\mathbb{P}^1 \cong \mathbb{A}^1 \sqcup \mathbb{A}^0$ viewed as lines in $\mathbb{A}^2$ (through $(0,0)$, the first part is given by lines $L_k = \{(x, Y);\ y = kx$ with a slope $k \in \Bbbk = \mathbb{A}^1$ and the second part is a point, the vertical line $x = 0$ of slope ∞.

Functions on $\mathbb{P}^1$ are functions on $\mathbb{A}^1$ that extend over ∞, i.e., polynomials $P(x)$ that have a finite value $\lim_{x \to \infty} P(x)$ at ∞, but these are just constants. Actually, in general $\mathcal{O}(\mathbb{P}^n) = \Bbbk$[5] Later we will remove this problem by noticing that there are many *local* functions though the only *global* ones are constants.

Because of this scarcity of functions on projective spaces we did not use functions to define projective subvarieties of $\mathbb{P}^n$ (as in the case of $\mathbb{A}^n$). If $G \in \Bbbk[X_0, ..., X_n]$ is homogeneous of degree d then $G(cx_0, ..., cx_n) = c^d \cdot G(x_0, ..., x_n)$, so the value of G on the line $[x_0 : \cdots : x_n]$ does not make sense, so it is not a function on $\mathbb{P}^n$.[6] However "$\underline{G = 0 \text{ on } [x_0 : \cdots : x_n]}$" still does make sense, and this is what we used above.

[5]The same holds for all connected projective varieties.

[6]However, such G's will be seen to be *sections of a line bundle on* $\mathbb{P}^n$, and in some sense they will turn out to be generalizations of functions.

2.2.10. *Completion of affine varieties to projective varieties.* Passing from $\mathbb{A}^n$ to $\mathbb{P}^n$ we need to pass somehow from all polynomials to homogeneous polynomials.

The degree of a polynomial $F = \sum_I c_I x^I \in \Bbbk[X_1, ..., X_n]$ is the maximal degree $|I| = I_1 + \cdots + I_n$ of the monomials that appear (i.e., $c_I \neq 0$). If $F \in \Bbbk[X_1, ..., X_n]$ has a degree d we can use it to produce a *homogeneous* polynomial of the same degree but with one more variable

$$\widetilde{F} = \sum_I c_I{\cdot}X^I{\cdot}X_0^{d-|I|} \in \Bbbk[X_0, X_1, ..., X_n].$$

Now if an affine variety $X{\subseteq}\mathbb{A}^n$ is given by equations $F_j = 0$ then the equations $\widetilde{F_j} = 0$ give a projective subvariety of $|P^n$ that we will call $\overline{X}$. Notice that $\bar{X} \cap \mathbb{A}^n$ consists of lines $[x_0; \cdots; x_n]$ with $0 = \widetilde{F_j}(x_0, ..., x_n)$ and $x_0 \neq 0$. After rescaling x_0 we see that these are the lines $[1; y_1; \cdots; y_n]$ with with $0 = \widetilde{F_j}(1; y_1, ..., y_n)$, i.e., $0 = F_j(y_1, , ..., y_n)$ (since $\widetilde{F}|_{x_0=1} = F$!). So,

$$\bar{X} \cap \mathbb{A}^n = X.$$

When thinking of $\mathbb{P}^n = \mathbb{A}^n \sqcup \mathbb{P}^{n-1}$ as a completion of $\mathbb{A}^n$, I will call $\mathbb{P}^{n-1}$ the boundary $\partial\mathbb{A}^n$ of $\mathbb{A}^n$ in $\mathbb{P}^n$. Similarly, $\partial X \stackrel{\text{def}}{=} \overline{X}\backslash X = \overline{X} \cap \partial\mathbb{A}^n = \{x_0 = 0 \text{ in } \overline{X}\}$ will be called the boundary ∂X of X in $\overline{X}$.

2.2.11. *Examples.* (a) The boundary of line is a point.

A line L is given by $aX + bY = c$ with $a \neq 0$ or $b \neq 0$. (Here $X = X_1$, $Y = X_2$.) So, $F = aX + bY - c$ and $\widetilde{F} = aX + bY - cX_0$. ∂X consists of lines $[x_0 : x : y]$ such that $X_0 = 0 = \widetilde{F}$, i.e., lines $[0; x : y]$ such that $\widetilde{F}|_{X_0=0}$ vanishes. Here, $\widetilde{F}|_{X_0=0} = aX + bY$, notice that in genera that $\widetilde{F}|_{X_0=0}$ is the top degree homogeneous part F_{top} of F. So, ∂X consists of lines $[x : y]$ that satisfy $ax + by = 0$. This gives precisely one line which one can describe as $[b, -a]$.

(b) The boundary of a circle in $\mathbb{A}^2(\mathbb{C})$ consists of two points.

A circle C is given by $F = (X-a)^2+(Y-b)^2-\gamma$, hence $\widetilde{F} = (X-aX_0)^2+(Y-bX_0)^2-\gamma X_0^2$, and ∂C is given by lines killed by $\widetilde{F}|_{X_0=0} = F_{top} = X^2+Y^2$. These are two lines $[1, \pm i]$.

Notice that

(1) over $\mathbb{R}$, $\partial C = \emptyset$, as expected since a circle does not stretch to ∞.
(2) Over $\mathbb{C}$, all circles pass through the same two points at ∞ !

2.3. **Include infinitesimals.** Bravely, we turn from intersecting two lines to intersecting a circle and a line. Over $\mathbb{R}$ we get two points (secant line), one point (tangent line), or no points. Over $\mathbb{C}$ there are only two case: two points or one point (the tangent case).

2.3.1. *Idea of a double point.* One way to describe this is that when a line degenerates from the generic position with respect to the circle to the special ("degenerate") case of a tangent line, the two points in the intersection degenerate to one point. Traditionally, geometers would go around this instability in the number of solutions by saying that the one point intersection in the tangent case should be counted twice as it is a limit of a pair of points, so it is a *double point.*

Nice but hazy! If this makes you unhappy you can try to study the situation in algebra.

2.3.2. *Functions on the intersection of a line and a circle; the algebraic calculation.* To do an algebraic calculation we choose coordinates conveniently. So the circle is the standard circle $C = \{X^2+Y^2=1\}$, and the line is the horizontal line L_c on height c, i.e., $L_c = \{Y=c\}$. The intersection $C\cap L_c$ is obtained by imposing both equations, so $Y=c$ and $X^2=1-c^2$. Therefore, the algebra of functions on the intersection is obtained as the quotient of $\Bbbk[X,Y]=\mathcal{O}(\mathbb{A}^2)$ obtained by imposing both equations:

$$\mathcal{O}(C\cap L_c) = \Bbbk[X,Y]|_{Y=c \text{ and } X^2=1-c^2} = \Bbbk[X]|_{X^2=1-c^2} = \Bbbk[X]/(X^2-(1-c^2))\cdot\Bbbk[X,Y].$$

So, the algebra is two dimensional: $\mathcal{O}(C\cap L_c) = \{a+bX;\ a,b\in\Bbbk\}$ and $X^2=1-c^2$.

This sounds roughly right: the intersection usually consists of two points and the functions are therefore two dimensional (can choose value at each point). However, for $c=1$ the line is tangent and the intersection is one point $(0,1)$, while the algebra we got:

$$\mathcal{O}(C\cap L_1) = \{a+bX;\ a,b\in\Bbbk\} \quad \text{with} \quad X^2=0,$$

is two dimensional.

2.3.3. *A mistake! (If we are really calculating functions on the variety $C\cap L_1$).* C and L_c are affine varieties given by one equation each. The intersection of these two subsets of $\mathbb{A}^2$ is the affine subvariety given by two equation. The definition of functions on the affine variety $C\cap L_c$ is:

$$\mathcal{O}(C\cap L_c) \stackrel{def}{=} \text{algebra of all restrictions of polynomials to the } \underline{\text{set}}\ C\cap L_c.$$

So, I have actually made a mistake in the algebraic calculation. When $c=\pm1$, then $X^2=1-c^2$ has one solution $X=0$ and the algebra of functions on $C\cap L_1$ is the algebra of restrictions $\Bbbk[X]|_{X^2=1-c^2} = \Bbbk[X]|_{X=0}$ of polynomials to the point $X=0$, so it is one dimensional: $\Bbbk[X]/X\cdot\mathbb{K}[X] \cong \Bbbk$. The mistake was that I was just imposing algebraic conditions rather then checking what happens on the level of sets as I should have if I am working with algebraic varieties (by definition, they are subsets of $\mathbb{A}^n$).

2.3.4. *Why should I believe that the algebraic calculation was correct in some world?, i.e., that the Double Point really exists?* The calculation with a mistake was better in the sense that the result was more stable since the dimension of functions on the intersection of C and L_c was independent of c ! This offers a way out:

- Algebra suggest that there is a world in which the intersection of C and L_c inside $\mathbb{A}^2$, is literally more then a point. This intersection we will denote $C\cap_{\mathbb{A}^2} L_c$ and it will be a space[7], characterized by its ring of functions:

$$\mathcal{O}(C\cap_{\mathbb{A}^2} L_c) \overset{def}{=} \textit{take the quotient of } \mathcal{O}(\mathbb{A}^2) \textit{ by imposing the equations of both } C \textit{ and } L_c.$$

  This space we will call a *double point*, and the algebra of functions on a double point is isomorphic to $\Bbbk[X]/X^2{\cdot}\Bbbk[X]$.

Now everything fits:

- The functions on a double point are expressions $a+bX,\ a,b\in\Bbbk$, with $X^2=0$.
- We certainly expect to have constant functions (even if there is only one point). More precisely, one should think that the double point contains an ordinary point p, because of the quotient map

$$\mathcal{O}(\textit{Double Point}) = \Bbbk[X]/X^2 \xrightarrow{X\to 0} \Bbbk = \mathcal{O}(pt) = \mathcal{O}(p),$$

  which can be viewed as restriction of functions to a point p.
- What is X and what is the meaning of $X^2=0$? (There are no such elements in a field such as $\mathbb{R}$ or $\mathbb{C}$, except for 0!) The (intuitive) explanation is that X measures the distance from p in the Double Point. Now $X^2=0$ shows that a Double Point is just slightly more then a point – we move so little from p that the function X only has infinitesimally small values, they are so small that X^2 is not only "negligible" but actually 0.

So, we can make sense of the double point. We lifted it from the vague idea that some point should be counted twice to a precise mathematical object (an algebra). But there is a small price: we need to find the geometric way of thinking about rings more general then the rings of polynomials ($=\mathcal{O}(\mathbb{A}^n)$), and their quotients $\mathcal{O}(X)$ obtained by restricting polynomials to algebraic subvarieties X of affine spaces $\mathbb{A}^n$. Because

2.3.5. *A double point is not an algebraic variety.* Notice that our double point is *not* an algebraic variety because

The algebra of functions $\mathcal{O}(Y)$ on an algebraic subvariety $Y\subseteq\mathbb{A}^n$ over a field $\Bbbk$ has no nilpotents.

This is so because $f\in\mathcal{O}(Y)$ is the restriction of some polynomial function F from $\mathbb{A}^n$ to $\Bbbk$, to the subset $Y\subseteq\mathbb{A}^n$. So, f is a function from Y to $\Bbbk$. Now $f^e=0$ implies that for all $y\in Y$ one has in $\Bbbk$: $0=f^n(y)$, i.e., $0=f(y)^e$. But since $\Bbbk$ is a field this implies that $f(y)=0,\ y\in Y$, i.e., $f=0$.

[7]This use of the word *space* means that we do not yet know what we want.

2.3.6. *Schemes.* So far we are making the case for the existence of a larger world that includes varieties, but also more things, for instance the double point. This will be true in

Grothendieck's world of $\mathcal{SCHEMES}$.

2.3.7. *Calculus.* At the break of the calculus dawn, the great minds calculated with infinitesimally small quantities. Later we fixed our view on real numbers and decided infinitesimals are nonsense, but we can fix the problem by translating the original formulations into the ε, δ-language. However, algebra allows quantities h which are "infinitesimally small" in the sense that $h^n = 0$ for some n – they just do not live in $\mathbb{R}$ but in some larger ring such as $\mathbb{R}[h]/h^n$.

Trying to do derivatives in this way, we first rewrite $f'(a) \stackrel{\text{def}}{=} \lim_{h\to 0} \frac{f(a+h)-f(a)}{h}$, as

$$f'(a)\d{h} = f(a+h) - f(a) + \mathcal{O}(h^2),$$

where $\mathcal{O}(h^2)$ denotes a quantity which goes to 0 faster then h^2. (This reformulation is anyway necessary in higher dimension.) Then one would say

- $f'(a)$ is the coefficient of h in $f(a+h) - f(a) \ \in \ \mathbb{R}[h]/h^2$.

Admittedly, one needs f to be in some sense algebraic so that $f(a+h)$ makes sense, i.e., that f extends naturally from $\mathbb{R}$ to $\mathbb{R}[h]/h^2$. (For instance rational functions and power series have this property.)

2.3.8. *Infinitesimal neighborhoods.* If we believe in infinitesimally small objects, we can ask what is the part of the line $\mathbb{A}^1$ which is infinitesimal close to 0? If there really is such space, the *infinitesimal neighborhood* X *of* 0 *in* $\mathbb{A}^1$, then by taking the clue from the Observation Principle, the question

"what is the infinitesimal neighborhood X of 0?",

can be restated as

"what is the algebra of function $\mathcal{O}(X)$ on the infinitesimal neighborhood X of 0?".

If we are in the algebraic setting, we want something like polynomials, but larger since polynomials make sense on all of $\mathbb{A}^1$, there should be things that make sense only close to 0. This reminds us that the *convergent power series* are series $\sum_0^\infty a_n x^n$ which converge on some small interval $(-\delta, \delta)$ around 0. We may want something even larger, both in order to distinguish the infinitesimal neighborhood X from actual neighborhoods $(-\delta, \delta)$, and to have a purely algebraic notion (i.e., no use of topology in $\mathbb{R}$). This leads to using *all formal power series*

$$\mathcal{O}(X) \stackrel{\text{def}}{=} \mathbb{R}[[T]],$$

the idea being that they will all converge when we are *infinitesimally* close to 0.

That's it. We've made sense of infinitesimals. The price is as mentioned before, develop a geometric way of thinking about rings like $\Bbbk[h]/h^n$ or $\Bbbk[[h]]$. Actually, it turns out that

Grothendieck's notion of schemes gives a geometric way of thinking about <u>all commutative rings</u>.

This goes very far since we (or most of us) prefer geometric way of thinking then the algebraic way.[8]

2.3.9. *Non-commutative geometry.* Next, we would like to have a geometric approach to all rings, including the scary non-commutative rings. This is in progress. There is a number of partial approaches. The most promising one by Connes.

3. **Spot a theorem**

We have looked at examples $\overline{X}\cap\overline{Y}$ of intersections in $\mathbb{P}^2$ of projective completions of affine curves $X = \{f(X,Y) = 0\}\subseteq\mathbb{A}^2$ and $Y = \{g(X,Y) = 0\}\subseteq\mathbb{A}^2$, given by two polynomials f and g.

3.0.10. *Question.* How many points are there in the intersection $X \cap Y$?

We saw that the number $|X \cap Y|$ behaves better if $\Bbbk = \mathbb{C}$ then when $\Bbbk = \mathbb{R}$ (fewer exceptions from the expected behavior). So let us work over $\mathbb{C}$.

3.0.11. *Examples with lines and circles.* Two lines meet at one point (unless $X = Y$). Line and circle meet at two points (unless tangent). Two circles usually meet at four points: two in $\mathbb{A}^2$ and the two common points at ∞.

3.0.12. *Degree of a planar curve.* The number of solutions of a polynomial in one variable is given by its degree. Seeing that this is likely to be important, we will say that a curve $X\subseteq\mathbb{A}^2$ has degree d if it is given by a polynomial f (in two variables), of degree d.

This pushes us to notice that instead of "circles" we should talk of (non-degenerate) *quadrics*, i.e., curves given by a polynomial (in two variables) of degree 2:

3.0.13. *There are only three quadrics over* $\mathbb{C}$. A quadric will mean a curve given by a quadratic equation $F = aX^2+bXY+cY^2+dX+eY+g$ (i.e., degree 2). By a linear change of coordinates (i.e. just a change of point of view on $\mathbb{A}^2$ and $\mathbb{P}^2$), we can rewrite F using completion to a square, first as $\alpha X'^2+\gamma Y'^2+\delta X'+\varepsilon Y'+\phi$ and then as $AX''^2+BY''^2+C$. The degenerate case is when one of A, B, C is zero, this reduces to two lines $X'''Y''' = 0$ or a double line $(X''')^2 = \rho$. So, non-degenerate quadrics are really circles (over $\mathbb{C}$!).

[8] Because we can draw or visualize pictures.

The projective extension $\overline{C}$ of a quadric curve C is given by a homogeneous quadratic polynomial $G = \sum_{0\leq i\leq j\leq 2} c_{ij}x_ix_j$. If we want to count[9] the projective quadrics (it is not difficult to see that this is really the same problem as affine quadrics), the question has more symmetry. The answer is that (over $\mathbb{C}$), after a linear change of coordinates any quadratic form G diagonalizes to one of the form $\sum_{0\leq i\leq k} Y_i^2$ with $0 \leq k \leq 2$. So we again get three quadrics.

Over $\mathbb{R}$, after a linear change of coordinates any quadratic form G diagonalizes to one of the form $(\sum_{1\leq i\leq k_+} Y_i^2) - (\sum_{1\leq j\leq k_-} Z_j^2)$, with $1 \leq k_+ + k_- \leq 3$. So there are 4 non-degenerate quadrics and 9 all together.

What are these three quadrics? For completeness we check geometrically that

Lemma. The three quadric curves are quite different

(1) $Q_2 = \{[Y_0 : Y_1 : Y_2];\ Y_0^2 + Y_1^2 + Y_2^2 = 0\} \cong \mathbb{P}^1$.
(2) $Q_1 = \{[Y_0 : Y_1 : Y_2];\ Y_0^2 + Y_1^2 = 0\}$ consist of two projective lines $L_\pm \xrightarrow{\cong} \mathbb{P}^1$ that intersect in a point.
(3) $Q_0 \stackrel{\text{def}}{=} \{[Y_0 : Y_1 : Y_2];\ Y_0^2 = 0\} \cong \mathbb{P}^1$ should be counted as a *double line.*

Proof. The affine part Q^o of $Q = Q_2 = \{[Y_0 : Y_1 : Y_2];\ Y_0^2 + Y_1^2 + Y_2^2 = 0\}$ is given by $Y_0 = 1$, i.e., $Y_1^2 + Y_2^2 = -1$ and in term of $Y_\pm = Y_1 \pm 1iY_2$ this is $-1 = Y_+ \cdot Y_-$ which is $\mathbb{k}^*$. The boundary is given by $\mathbb{P}(\{Y_0 = 0\})$, i.e., the lines in $Y_+ \cdot Y_- 0$ and these are two points. All together we see that$Q^o \cong \mathbb{A}^1 - \{0\}$ and $\partial Q^o = \{0, \infty\}$ glues to $Q \cong \mathbb{P}^1$.

For $Q = Q_1 = \{[Y_0 : Y_1 : Y_2];\ Y_0^2 + Y_1^2 = 0\}$ the affine part $Q^o = \{Y_0 = 1\} = \{Y_1^2 = -1\}$ consists of two affine lines $L^o_\pm = \{Y_1 \in \mathbb{A}^o \text{ and } Y_2 = \pm i\}$. The boundary $\partial Q^o = \mathbb{P}(\{Y_0 = 0\}) = \mathbb{P}\{[Y_1; Y_2];\ Y_1^2 = 0\}$, is one point common to both lines $L^o_\pm$. So, Q consists of two projective lines $L_\pm \xrightarrow{\cong} \mathbb{P}^1$ that intersect in a point.

Finally, $Q = Q_0 \stackrel{\text{def}}{=} \{[Y_0 : Y_1 : Y_2];\ Y_0^2 = 0\}$ is just one projective line $\mathbb{P}^1$, however it should be counted as a *double line.*

3.0.14. *Bezout's theorem.* So in the examples above the expected number is the product of degrees:

Conjecture. $|X \cap Y| = deg(f) \cdot \deg(g)$.

This turns out to be always true – once one accounts for exceptions by counting intersection points with multiplicities !

Theorem. $\sum_{p\in X\cap Y} mult_p(X, Y) = deg(X) \cdot \deg(Y)$.

[9]The count of quadrics here really means the count of *different quadrics*, i.e., the isomorphism classes of quadrics.

4. **Include the number theory**

4.1. **The spectrum of $\mathbb{Z}$.** Number theory starts with studying the ring $\mathbb{Z}$ of integers. What one studies are polynomial equations, i.e., algebro geometric questions: solve $x^2 + y^2 = z^2$ in $\mathbb{Z}$, or $x^3 + y^3 = z^3$ in $\mathbb{Z}$, etc. To think of this really as algebraic geometry, we need a space. So, we would like the ring of integers $\mathbb{Z}$ to be the ring of functions on some space S: $\mathcal{O}(S) = \mathbb{Z}$. We will call S the spectrum of $\mathbb{Z}$: $S = \mathrm{Spec}(\mathbb{Z})$. Then, doing algebra in $\mathbb{Z}$ will be the same as geometry on S. The geometric way of thinking involves some basic questions

4.1.1. *What are the points of Spec*($\mathbb{Z}$)*?* A point of an affine space $a \in \mathbb{A}^n(\Bbbk)$ over a field $\Bbbk$ is a vantage point from which we can observe the observables i.e., functions $f \in \mathcal{O}(\mathbb{A}^n)$. Of course, by "observing functions at a", I mean evaluating functions at a. What is the evaluation at a in terms of the algebra $\mathcal{O}(\mathbb{A}^n)$? It is a map from functions to the ground field $\Bbbk$ $ev_a : \mathcal{O}(\mathbb{A}^n) \to \Bbbk$, $f \mapsto f(a)$. We can also think of it as a restriction $\mathcal{O}(\mathbb{A}^n) \to \mathcal{O}(\{a\}) \cong \Bbbk$.

So, algebraically, a point of X is a

(1) homomorphism of rings $\mathcal{O}(X) \to A$,
(2) it is surjective, i.e., it is a quotient map: $A \cong \mathcal{O}(X)/I$ for some ideal I (actually $I = I_a$),
(3) the target is a field, i.e. A is a field.

The last requirement intuitively means that the algebra A is "small". For instance for any algebraic subvariety $X \subseteq \mathbb{A}^n$, the restriction map $\mathcal{O}(\mathbb{A}^n) \to \mathcal{O}(X)$, satisfies (1) and (2) but it satisfies (3) only if X is a single point.

So, a point x of S should be homomorphism of algebras from $\mathcal{O}(S) = \mathbb{Z}$ to some quotient field A. All quotients are of the form $A = \mathbb{Z}/n\mathbb{Z}$, $n \in \mathbb{Z}_+$; and this is a field iff n is a prime. So

The points of Spec($\mathbb{Z}$) *are primes!*

4.1.2. *Question.* So the quotient fields of $\mathbb{Z}$ should be thought of as points of $\mathrm{Spec}(\mathbb{Z})$. What is the geometric meaning of the fraction field $\mathbb{Q}$ of $\mathbb{Z}$?

4.1.3. *Psychological maturity.* The first great triumph of this way of thinking in number theory was (as much as I can remember) I guess the proof of the Mordel conjecture by Faltings. I find it remarkable that it took something like 30 years.

4.2. **Affine schemes.** Thanks to Grothendieck, one can think of any commutative ring A as the ring of functions on a space $\mathrm{Spec}(A)$. Such spaces that correspond to commutative rings, are called *affine schemes.* Of course, if $A = \Bbbk[X_1, ..., X_n]$ then $\mathrm{Spec}(A)$ should be really $\mathbb{A}^n(\Bbbk)$, in the sense that it is an object that contains the same information.

Unfortunately, we have to accept that we may talk about it a bit differently when we think of it as a scheme rather then as a variety.

4.2.1. *Points of schemes and varieties.* One uses differently the word *point* when one talks about schemes then when one talks of varieties. Since varieties are schemes this introduces some *confusion*, and we will make peace with this in a moment!

4.2.2. *The geometric space Spec*(A). Now we will describe the structure of Spec(A) in stages. To avoid abstraction shock, we will now only explain the first two out of three levels of structure on Spec(A) :

(1) Spec(A) is a set:

The points of Spec(A) *are prime ideals* P *of* A.

(2) Spec(A) is a topological space:

The open sets in Spec(A) *are subsets Spec*$(A)_P$, $P \in$ *Spec*(A), *where Spec*$(A)_P$ *is the complement of the subset* $V_P \stackrel{def}{=} \{Q \ \ Q \supseteq P\} \subseteq$ *Spec*(A).

(3) Spec(A) is a ringed space (i.e., a topological space supplied with a sheaf of rings):

The ring of function on Spec(A) *is* A. *Moreover, there is a ring of functions* $\mathcal{O}(U)$ *for any open* $U \subseteq$ *Spec*(A), *and together they form a sheaf of rings* $\mathcal{O}$ *on Spec*(A). *For instance,* $\mathcal{O}(\text{Spec}(A)_P) = A_P$ *is the localization of* A *at* P, *i.e. one inverts all elements in* $A \backslash P$.

Of course, all of this will only make sense much later.

4.3. **Set Spec**(A)**.**

4.3.1. *Maximal ideals as cpoints of Spec*(A). We have decided above that the algebraic way to think of a point of a variety X as a surjective map $\phi : \mathcal{O}(X) \twoheadrightarrow l$ where l is a field. Then $I = \text{Ker}(\phi)$ is an ideal in $\mathcal{O}(X)$ and $l \cong \mathcal{O}(X)/I$. So, all information is contained in an ideal I of $\mathcal{O}(X)$ such that $\mathcal{O}(X)/I$ is a field. However,

Lemma. A/I is a field iff the ideal I in A is maximal.

4.3.2. *Corollary.* The points of an affine variety X are the maximal ideals in $\mathcal{O}(X)$.

So we would like to say that the points of the spectrum Spec(A) (of any ring A), are the maximal ideals of A.

Actually, according to the standard terminology these are not all points of Spec(A) but only the points of a special kind. We will say that

The cpoints[10] *of Spec*(A) *are maximal ideals of* A.

[10]Closed points.

4.3.3. *The points of Spec*(A). We will use another, more general, notion of points of Spec(A) – we will allow more ideals

The points *of Spec*(A) *are all* <u>*prime*</u> *ideals of* A.

4.3.4. *Prime ideals.* An ideal $P \subseteq A$ is said to be prime if $a, b \in A$ and $ab \in P$ implies that $a \in P$ or $b \in P$.

Lemma. (a) P is prime iff A/P has no zero divisors.

(b) Maximal ideals are prime.

(c) Zero ideal is prime iff A has no zero divisors ("A is integral").

4.3.5. *Lemma.* (a) For the ring $\mathbb{Z}$:

(1) all ideals are *principal* i.e., of the form $(n) \stackrel{\text{def}}{=} n\mathbb{Z}$ for some $n \in \mathbb{Z}$.
(2) maximal ideals = all $p\mathbb{Z}$ with p a prime.
(3) prime ideals = maximal ideals $\sqcup \{0\}$.

(b) For the ring $\mathcal{O}(\mathbb{A}^1) = \mathbb{k}[X]$

(1) All ideals in are principal, i.e. of the form $I = (P) = P \cdot \mathbb{k}[X]$ for some polynomial P. Actually, if we ask that P is monic this will make it unique.
(2) Ideal (P) is prime iff $P = 0$ or P is irreducible.
(3) If $\mathbb{k}$ is closed the only irreducible polynomials are of the form $X - a,\ a \in \mathbb{A}^1$. The corresponding ideal $(X - a)$ is exactly the ideal I_a of all functions that vanish at a.
(4) So, if $\mathbb{k}$ is closed, prime ideals are ideals of the form I_Y for one of the following subvarieties Y of $\mathbb{A}^1$:
  - Y is a point (maximal ideals correspond to points!),
  - $Y = \mathbb{A}^1$.

(c) Consider the ring $\mathcal{O}(X)$ for the affine variety $X = \{xy = 0\ in\ \mathbb{A}^2\}$.

(1) It has a basis $..., y^2, y, 1, x, x^2, ...$, and $xy = 0$.
(2) Maximal ideals correspond to points.
(3) 0 is not prime but there are two more prime ideals $(y) = I_{x-axis}$ and $(y) = I_{y-axis}$.

(c) Prime ideals in $\mathcal{O}(\mathbb{A}^2)$ correspond to

- (0) points,
- (2) $\mathbb{A}^2$ itself and
- (1) irreducible curves in X ("those that consist of one piece").

4.3.6. *Prime ideals and irreducible components.* We say that an affine variety X is *irreducible* if $\mathcal{O}(X)$ has no zero divisors.

Having zero divisors $f, g \neq 0 = fg$ in $\mathcal{O}(X)$ means that there are two affine subvarieties $Y = \{f = 0\}$ and $Z = \{g = 0\}$, such that $Y \cup Z = X$ and neither $Y \subseteq Z$, nor $Z \subseteq Y$. A basic example is $X = \{xy = 0 \ in \ \mathbb{A}^2\}$ which is the union of the x and y axes in the plane. Therefore, *irreducible* means the opposite, i.e., that we can not decompose X into two smaller affine subvarieties.

4.3.7. *Functoriality of the spectrum requires prime ideals.* To a map of varieties $f : X \to Y$ there corresponds the morphism of algebras of functions in the opposite direction

$$\mathcal{O}(Y) \xrightarrow{f^*} \mathcal{O}(X),$$

given by the *pull-back of functions*, i.e., $f^*(\phi) = \phi \circ f$. Actually, this gives an identification

$$Map(X, Y) \ni f \mapsto f^* \in \mathrm{Hom}_{\Bbbk-alg}[\mathcal{O}(Y), \mathcal{O}(X)].$$

Therefore, we expect for general rings A and B to have a bijection

$$Map[\mathrm{Spec}(A), \mathrm{Spec}(B)] \leftrightarrow \mathrm{Hom}_{\mathcal{R}ings}[B, A].$$

So to a map of rings $F : B \to A$, we expect to associate a map of sets $f = \mathrm{Spec}(F)$: $\mathrm{Spec}(A), \mathrm{Spec}(B)$.

We have agreed that maximal ideals in A should be points of $\mathrm{Spec}(A)$. So we expect to attach to any maximal ideal $I \subseteq A$ some maximal ideal $f(I) \subseteq B$. The map $F : B \to A$ gives just one way of associating to an ideal I in A an ideal J in B – this is the pull-back $J = F^{-1}I = \{a \in A;\ F(a) \in I\}$. So we want

$$\textit{If } I \subseteq A \textit{ is maximal then } F^{-1}I \subseteq B \textit{ is maximal.}$$

However, it is easy to find counterexamples.

How bad is this? If we insist that maximal ideals in A are points of $\mathrm{Spec}(A)$, we are forced to allow more points in $\mathrm{Spec}(B)$ then we expect, not only the maximal ideals in B but also any ideal which is the pull-back of a maximal ideal.

How much more is this? Actually, the pull-back of a maximal ideal is always a prime ideal, and more is true:

Lemma. The pull-back of a prime ideal is always prime!

Proof. Let $F : B \to A$ and let P be a prime ideal in A. If $ab \in F^{-1}P$, i.e., $P \ni F(ab) = F(a) \cdot F(b)$, since P is prime we know that either $P \ni F(a)$ or $P \ni F(b)$, i.e., $F^{-1}P \ni a$ or $F^{-1}P \ni b$).

Conclusion. Adding all prime ideals to maximal idels solves the functoriality (naturality) problem!

4.3.8. *The scheme-theoretic points of a variety.* For an affine variety $X/\Bbbk$, what are the points of the associated scheme $\mathrm{Spec}(\mathcal{O}(X))$? The answer is simple – instead of only looking at the points of X, which are the smallest subvarieties of X, we are force to look at all subvarieties at once. Actually, *all* is slightly more then is needed – we are not interested in the ones that consist of several pieces, if X is a union of two smaller subvarieties Y and Z we omit X from the list.

Lemma. The points of $\mathrm{Spec}(\mathcal{O}(X))$, i.e., the prime ideals in $\mathcal{O}(X)$ are the same as *irreducible* subvarieties Y of X.

Proof. The prime ideals in $\mathcal{O}(X)$ are precisely the ideals I_Y corresponding to *irreducible* subvarieties $Y\subseteq X$.

4.4. **Topological space Spec(A).**

4.4.1. *Zariski topology on affine varieties.* We define the Zariski topology on an affine variety X so that the closed subsets are precisely the affine subvarieties of Y.

So, $X\subseteq\mathbb{A}^n$ is given by finitely many polynomial equations $X=\{0=F_1=\cdots=F_c\}$. We say that a subset $Y\subseteq X$ is Zariski closed if it is given by a few more additional polynomial equations. This is natural in the sense that if we *think* of polynomials as continuous functions the the subsets $\{G=0\}$ should be closed! All-together, if say $\Bbbk=\mathbb{C}$ then the *Zariski closed* subsets of $\mathbb{A}^n=\mathbb{C}^n$ are the closed subsets which can be described using polynomials.

Lemma. A family $\mathcal{C}$ of subsets of a set X is the set of close subsets in some topology $\mathcal{T}$ on X iff

(1) $\mathcal{C}\ni\emptyset, X$, and
(2) $\mathcal{C}$ is closed under finite unions and arbitrary intersections.

Proof. If $\mathcal{T}=\{X-F;\ F\in\mathcal{C}\}$, the conditions on $\mathcal{C}$ translates into conditions on $\mathcal{T}$ which are precisely the definition of a topology.

Proposition. Zariski topology on an affine variety is well defined.

Proof. This is almost a tautology but there is one thing to check.[11] At the moment we will postpone the proof and prove a more general version for schemes (that one really is a tautology!).

[11] What?

4.4.2. *Zariski topology on Spec(A).* Any ideal $I\subseteq A$ defines a subset

$$V_I \stackrel{\text{def}}{=} \{P\in \text{Spec}(A);\ P\supseteq I\}.$$

We define the Zariski topology on Spec(A) so that the closed subsets are precisely the subsets V_I given by ideals I.

Example. Let us see what this means when $A=\mathcal{O}(X)$ for an affine variety $X/\Bbbk$. The interesting ideals $I\subseteq A$ are the ideals $I_Y = \ \{f\in\mathcal{O}(X);\ f|Y=0\}$, corresponding to the affine subvarieties $Y\subseteq X$. We will see that V_{I_Y} is an incarnation of Y itself. In particular, the Zariski closed sets in the scheme Spec$[\mathcal{O}(X)]$ correspond to the Zariski closed sets in X, so $I\mapsto V_I$ is the correct generalization of the Zariski topology on varieties to schemes.

For this we gauge the subset V_{I_Y} of the set of prime ideals of A by looking at the prime ideals that make sense geometrically, and these are the maximal ideals I_a in $\mathcal{O}(X)$ corresponding to the points a of X. I_a lies in V_{I_Y} if $I_a\subseteq I_Y$, i.e., if each function that vanishes on Y also vanishes at a. This happen precisely if $a\in Y$. So,

- To any subvariety Y of X we attach an ideal I_Y and therefore also a closed subset V_{I_Y} of Spec(A).
- One has
$$V_{I_Y}\ \cap\ X\ =\ Y,$$
i.e., the intersection of V_{I_Y} with the points of X (viewed as maximal ideals of $\mathcal{O}(X)$), consists precisely of points of Y.

4.4.3. *Proposition.* Zariski topology is well defined.

Proof. We have to check that $\mathcal{C}=\{V_I,\ I \text{ an ideal in } A\}$ satisfies the conditions from the lemma 4.4.1.... It all follows from the next lemma.

4.4.4. *Lemma.* (a) For ideals $I,J\subseteq A$,

$$V_{I\cap J}=V_I\cup V_J.$$

(b) For ideals $I_p\subseteq A$,

$$\cap_p\ V_{I_p}\ =\ V_{\sum_p I_p}.$$

4.4.5. *The closed points of a scheme.* In particular we will see that the points of an affine variety X can be described as the closed scheme-theoretic points of X.

Lemma. (a) The closure of a point $P\in\text{Spec}(A)$ is $\overline{\{P\}}=V_P$.

(b) The closed points in Spec(A) are precisely the maximal ideals.[12]

(c) For an affine variety $X/\Bbbk$, the closed points in the associated scheme Spec$(\mathcal{O}(X))$ are the same as the points of X.

[12]i.e., what we called the cpoints of Spec(A).

Proof. (a) $\overline{\{P\}} = V_P$ is the smallest closed set V_I that contain P, i.e., the smallest V_I such that $I \subseteq P$. Since $I \subseteq P \Rightarrow V_P \subseteq V_I$, the smallest one is V_P.

Lemma. The prime ideals in $\mathcal{O}(X)$ are precisely the ideals I_Y corresponding to *irreducible* subvarieties $Y \subseteq X$.