

NONCOMMUTATIVE RINGS

Michael Artin

class notes, Math 251, Berkeley, fall 1999

I began writing notes some time after the semester began, so the beginning of the course (diamond lemma, Peirce decomposition, density and Wedderburn theory) is not here. Also, the first chapter is sketchy and unreadable. The remaining chapters, though not in good shape, are a fair record of the course except for the last two lectures, which were on graded algebras of GK dimension two and on open problems.

I. Morita equivalence	
1. Hom	3
2. Bimodules	3
3. Projective modules	4
4. Tensor products	5
5. Functors	5
6. Direct limits	6
7. Adjoint functors	7
8. Morita equivalence	10
II. Localization and Goldie's theorem	
1. Terminology	13
2. Ore sets	13
3. Construction of the ring of fractions	15
4. Modules of fractions	18
5. Essential submodules and Goldie rank	18
6. Goldie's theorem	20
III. Central simple algebras and the Brauer group	
1. Tensor product algebras	23
2. Central simple algebras	25
3. Skolem-Noether theorem	27
4. Commutative subfields of central simple algebras	28
5. Faithful flatness	29
6. Amitsur complex	29
7. Interlude: analogy with bundles	32
8. Characteristic polynomial for central simple algebras	33
9. Separable splitting fields	35
10. Structure constants	36
11. Smooth maps and idempotents	37
12. Azumaya algebras	40
13. Dimension of a variety	42
14. Background on algebraic curves	42
15. Tsen's theorem	44

IV. Maximal orders	
1. Lattices and orders	46
2. Trace pairing on Azumaya algebras	47
3. Separable algebras	49
4. Maximal orders in separable algebras	51
5. Hensel's lemma	52
6. Maximal orders over complete discrete valuation rings	54
7. Recovering data from the completion	57
8. Addendum to Chapter III	59
V. Irreducible representations	
1. Definition	62
2. Examples	62
3. Standard identities	64
4. Central polynomials	66
5. Polynomial identity rings	68
6. Kaplansky's theorem	69
7. Amitsur's theorem	70
8. Posner's theorem	72
9. Intrinsic characterization of Azumaya algebras	72
10. Irreducible representations of the free ring	74
11. The case of two 2×2 matrices X, Y	76
12. Some tensor notation	77
13. The main theorem of invariant theory	78
14. Procesi's theorem	79
15. The spectrum of the ring of generic matrices	82
VI. Growth of algebras	
1. Growth functions	85
2. Warfield's theorem	87
3. Path algebras	88
4. Bergman's gap theorem	88
5. Theorem of Stephenson and Zhang	92
6. Projective covers	93
7. Hilbert functions of graded algebras of finite global dimension	95
8. Modules with linear growth	98
9. Theorem of Small and Warfield	99

I. MORITA EQUIVALENCE

This discusses when the module categories over two different rings A and B are equivalent.

I.1. Hom

When possible, we write homomorphisms $f : P_B \longrightarrow N_B$ of right B -modules as acting on the left: $p \mapsto fp$. Writing homomorphisms on the left has the advantage of keeping the scalars out of the way. B -linearity becomes the associative law:

$$f(pb) = (fp)b, \quad \text{and} \quad f(p_1 + p_2) = fp_1 + fp_2.$$

For the same reason, we try to write homomorphisms $g : {}_A M \longrightarrow {}_A M$ of left modules as acting on the right: $m \mapsto mg$.

The set $\text{Hom}_B(P, N)$ of module homomorphisms is an abelian group under addition of functions: $[f + g]p = fp + gp$.

Composition of operators makes $\text{Hom}_B(P, N)$ into a covariant functor of N and a contravariant functor of P : If $u : N \longrightarrow N$ is a homomorphism, then the map $\text{Hom}_B(P, N) \longrightarrow \text{Hom}_B(P, N)$ is defined by sending $f \mapsto u \circ f$, and similarly, if $v : P \longrightarrow P$, then $\text{Hom}_B(P, N) \longleftarrow \text{Hom}_B(P, N)$ sends f to $f \circ v$.

Homomorphisms from a module P_B to itself are called endomorphisms, and we denote the set $\text{Hom}_B(P, P)$ by $\text{End } P_B$. It is a ring, multiplication being composition of functions.

I.2. Bimodules

Let A, B be rings. An A, B -bimodule is an abelian group with a structure of right B -module and also a structure of left A -module, such that the operations of A and B commute:

$$(I.2.1) \quad (ap)b = a(pb)$$

for all $a \in A, p \in P, b \in B$. There are other ways to interpret such a structure: If we consider the right B -module structure P_B , then left multiplication λ_a by an element $a \in A$, the map defined by $\lambda_a p = ap$, is an endomorphism of P_B . So we obtain a map $A \longrightarrow \text{End } P_B$ by sending $a \mapsto \lambda_a$, and this map is a ring homomorphism. Conversely, any ring homomorphism $A \longrightarrow \text{End } P_B$ gives a bimodule structure on P :

Corollary I.2.2. *An A, B -bimodule ${}_A P_B$ is given by a right module P_B together with an arbitrary homomorphism $A \longrightarrow \text{End } P_B$. In particular a right B -module P has a canonical E, B -bimodule structure, where $E = \text{End } P_B$. \square*

Similarly, ${}_A P_B$ is also determined by a left module ${}_A P$ and a homomorphism $B \longrightarrow \text{End}_A P$.

Lemma I.2.3. (i) If ${}_A P_B$ is a bimodule, then $\text{Hom}_B(P, N)$ is a right A -module by the rule $[fa]p = f(ap)$.

(ii) If ${}_R N_B$ is a bimodule then $\text{Hom}_B(P, N)$ is a left R -module, with the operation defined by $[rf]p = r(fp)$.

(iii) $\text{Hom}_B(P, -)$ is a left exact functor on $\text{Mod } B$. More precisely, sequence

$$0 \longrightarrow N_1 \longrightarrow N_2 \longrightarrow N_3$$

is exact if and only if

$$0 \longrightarrow \text{Hom}_B(P, N_1) \longrightarrow \text{Hom}_B(P, N_2) \longrightarrow \text{Hom}_B(P, N_3)$$

is exact for all $P \in \text{Mod } B$.

(iv) $\text{Hom}_B(-, N)$ is a right exact contravariant functor on $\text{Mod } B$. A sequence

$$P_1 \longrightarrow P_2 \longrightarrow P_3 \longrightarrow 0$$

is exact if and only if

$$0 \longrightarrow \text{Hom}_B(P_1, N) \longrightarrow \text{Hom}_B(P_2, N) \longrightarrow \text{Hom}_B(P_3, N)$$

is exact for all $N \in \text{Mod } B$. \square

Remark I.2.4: Let A be a commutative ring. By *central* A -module we mean the A, A -bimodule obtained from a right A -module by decreeing that the left and right actions are the same: $am = ma$. In commutative algebra, it is customary to move scalars from left to right informally, i.e., to work with this bimodule, calling it a module. This is not a good idea when A isn't commutative, because the associative law for scalar multiplication screws things up: If we declare that $am = ma$ for all $a \in A$ and $m \in M$, then for $a, a' \in A$, $(aa')m = m(aa')$. But the associative law requires that $(aa')m = a(a'm) = a(ma') = (ma')a = m(a'a)$. Thus right multiplication by aa' and by $a'a$ are forced to be equal, which means that the actions of A on M factor through a commutative quotient ring of A .

I.3. Projective modules

A right B -module P is called *projective* if it satisfies any one of the conditions of the proposition below. For example, a free module $\bigoplus B$ is projective.

Proposition I.3.1. *The following conditions on a right B -module P are equivalent:*

(i) Let $N \xrightarrow{v} N$ be a surjective map of right B -modules, and let $P \xrightarrow{\phi} N$ be any map. There exists a lifting of ϕ to N , a map $P \xrightarrow{\phi'} N$ such that $\phi = v \circ \phi'$.

(ii) Every surjective map $Q \xrightarrow{f} P$ of right B -modules splits, i.e., there is a map $Q \xleftarrow{s} P$ such that $f \circ s = \text{id}_P$.

(iii) P is a summand of a free module F , i.e., $F \approx P \oplus P$ for some module P .

(iv) The functor $\text{Hom}_B(P, -)$ is exact: If $0 \longrightarrow N_1 \longrightarrow N_2 \longrightarrow N_3 \longrightarrow 0$ is an exact sequence, then

$$0 \longrightarrow \text{Hom}_B(P, N_1) \longrightarrow \text{Hom}_B(P, N_2) \longrightarrow \text{Hom}_B(P, N_3) \longrightarrow 0$$

is also exact. \square

I.4. Tensor products

If M_A and ${}_A P$ are right and left A -modules, we can form the tensor product $M \otimes_A P$. It is defined as the abelian group generated by elements $m \otimes p$, with the relations

$$(m_1 + m_2) \otimes p = m_1 \otimes p + m_2 \otimes p, \quad m \otimes (p_1 + p_2) = m \otimes p_1 + m \otimes p_2, \quad \text{and} \quad ma \otimes p = m \otimes ap.$$

There is no structure of A -module on this tensor product: the actions of A have been used up in the middle.

Lemma I.4.1. (i) *The tensor product is functorial in M and in P .*

(ii) *If ${}_A P_B$ is a bimodule, then $M \otimes_A P$ becomes a right B -module.*

(iii) *The functor $- \otimes_A P$ is right exact: If*

$$M_1 \longrightarrow M_2 \longrightarrow M_3 \longrightarrow 0$$

is an exact sequence of right A -modules, then

$$M_1 \otimes_A P \longrightarrow M_2 \otimes_A P \longrightarrow M_3 \otimes_A P \longrightarrow 0$$

is also exact. Similarly, if

$$P_1 \longrightarrow P_2 \longrightarrow P_3 \longrightarrow 0$$

is exact, then so is

$$M \otimes_A P_1 \longrightarrow M \otimes_A P_2 \longrightarrow M \otimes_A P_3 \longrightarrow 0$$

\square

I.5. Functors

We are interested in additive functors $\text{Mod } A \xrightarrow{F} \text{Mod } B$, but most considerations carry over to abelian categories. So let's write the functor neutrally as

$$\mathcal{C} \xrightarrow{F} \mathcal{D}.$$

Recall that a functor associates to every object $X \in \mathcal{C}$ an image $FX \in \mathcal{D}$, and for every pair X_1, X_2 of objects in \mathcal{C} , a map $\text{Hom}_{\mathcal{C}}(X_1, X_2) \longrightarrow \text{Hom}_{\mathcal{D}}(FX_1, FX_2)$, these maps on Hom being compatible with composition of morphisms in \mathcal{C} .

The adjective ‘‘additive’’ means that the maps on morphisms are homomorphisms of abelian groups, which is a mild restriction. A functor is additive if and only if $F(X_1) \oplus F(X_2) \approx F(X_1 \oplus X_2)$ (see [McL]). But never mind: Let's just assume it.

On the other hand, an additive functor needn't be compatible with infinite direct sums. Given an indexed family $M_i \in \text{Mod } A$ for $i \in I$, there is a canonical map $\bigoplus F(M_i) \longrightarrow F(\bigoplus M_i)$, but when

I is infinite it may not be an isomorphism. If this map is an isomorphism for all indexed families M_i , we say that F is *compatible with direct sums*.

Exercise: Give an example of a functor which is not compatible with direct sums.

A functor $\mathcal{C} \xrightarrow{F} \mathcal{D}$ is *fully faithful* if the map $\text{Hom}_{\mathcal{C}}(M_1, M_2) \longrightarrow \text{Hom}_{\mathcal{D}}(FM_1, FM_2)$ is bijective for all $M_1, M_2 \in \mathcal{C}$, and F is *essentially surjective* if every $N \in \mathcal{D}$ is isomorphic to FM for some $M \in \mathcal{C}$.

A *quasi-inverse* of a functor F is a functor $\mathcal{C} \xleftarrow{G} \mathcal{D}$ such that $FG \approx id_{\mathcal{D}}$ and $GF \approx id_{\mathcal{C}}$, where the symbols \approx stand for isomorphisms of functors. The usual proof of uniqueness of inverses shows that a quasi-inverse is unique up to isomorphism of functors. A functor which has a quasi-inverse is called an *equivalence of categories*.

Example I.5.1: Let A be a commutative ring, and let L be an invertible A -module, or an invertible central bimodule. This means that there is an inverse module L^{-1} , such that $L \otimes_A L^{-1} \approx L^{-1} \otimes_A L \approx A$. Then $F = - \otimes_A L$ is an equivalence of categories $\text{Mod } A \longrightarrow \text{Mod } A$, an *autoequivalence* of $\text{Mod } A$, with quasi-inverse $G = - \otimes_A L^{-1}$. Note that GF is not the identity functor, because $M \otimes_A L \otimes_A L^{-1}$ is not the same module as M . However, it is canonically isomorphic to M , and this canonical isomorphism yields the isomorphism of functors $GF \approx id$.

Proposition I.5.2. (i) A functor F between small categories is an equivalence if and only if it is fully faithful and essentially surjective.

(ii) An equivalence of categories is an exact functor, and it commutes with direct sums.

We will sketch the proof of the most interesting part, which is that a fully faithful, essentially surjective functor F is an equivalence. To prove this, we have to define a quasi-inverse G . Let $N \in \mathcal{D}$. We choose an arbitrary isomorphism $FM \xrightarrow{\phi} N$, where M is an object of \mathcal{C} , and we set $GN = M$. This defines G on objects. To define G on maps, let $N_1 \xrightarrow{f} N_2$ be given. Then $\phi_2^{-1} \circ f \circ \phi_1$ is a map $FM_1 \longrightarrow FM_2$. Since F is fully faithful, there is a unique map $M_1 \xrightarrow{g} M_2$ such that $F(g) = \phi_2^{-1} f \phi_1$. We set $G(f) = g$. \square

I.6. Direct limits

By abstract *diagram* I we mean a category whose objects form a (small) set. For example, a directed graph, a collection of vertices and of arrows connecting vertices, gives rise to a diagram if one adds identity maps and compositions of arrows. By diagram *in a category* \mathcal{C} we mean a functor $X : I \longrightarrow \mathcal{C}$ from an abstract diagram to \mathcal{C} . Given such a diagram in \mathcal{C} and an object $i \in I$, we may write X_i for the image of i . So a diagram in \mathcal{C} may be viewed as a collection of objects X_i and of maps between them, identities and compositions of maps being included.

A map $X \xrightarrow{\Phi} N$ from a diagram in \mathcal{C} to an object N of \mathcal{C} is by definition a collection of maps $X_i \xrightarrow{\phi_i} N$ which are compatible with all maps in the diagram. More precisely, if $i \xrightarrow{u} j$ is a map in I , then X_i maps to X_j by u (strictly speaking, by X_u), and we require that $\phi_j = \phi_i \circ u$. This is to be true for all arrows u in I .

The *direct limit* $\overline{X} = \varinjlim_I X$ of a diagram X in \mathcal{C} is the universal object for maps from X to objects of \mathcal{C} . So it is defined by the rule

$$\mathrm{Hom}_{\mathcal{C}}(\overline{X}, N) \approx \{\text{maps } X \longrightarrow N\}.$$

It is not hard to construct this direct limit for an arbitrary diagram. We start with the direct sum $U = \bigoplus_I X_i$. Then a map $U \longrightarrow N$ corresponds to an arbitrary collection of maps from the X_i to N :

$$\mathrm{Hom}_{\mathcal{C}}(U, N) \approx \prod_{i \in I} \mathrm{Hom}_{\mathcal{C}}(X_i, N).$$

The direct limit is a quotient of U , which is obtained by introducing the compatibility conditions: Let $i \xrightarrow{u} j$ in I . Then $X_i \xrightarrow{u} X_j$. Let K denote the kernel of the map $(u, -1) : X_i \oplus X_j \longrightarrow X_j$. If \mathcal{C} is a module category, then K consists of pairs $(x, -ux)$ with $x \in X_i$. To make the two maps ϕ_i, ϕ_j compatible with u requires killing K . So \overline{X} is obtained from U by killing all of these subobjects K .

Corollary I.6.1. *Let $F : \mathcal{C} \longrightarrow \mathcal{D}$ be a right exact functor which is compatible with direct sums. Then F is compatible with arbitrary direct limits. In other words, if $X : I \longrightarrow \mathcal{C}$ is a diagram in \mathcal{C} and if FX denotes the composed diagram $F \circ X$ in \mathcal{D} , then $\varinjlim FX \approx F(\varinjlim X)$.*

Proof. This follows from the fact that $\varinjlim X$ is constructed as a cokernel of a direct sum, and that F is compatible with cokernels and direct sums. \square

Exercise: Describe $\varinjlim X$ when I has two objects.

I.7. Adjoint functors

Most functors $F : \mathcal{C} \longrightarrow \mathcal{D}$ are not equivalences. For one thing, if we are given an object $Y \in \mathcal{D}$, there is no reason to suppose that an object $X \in \mathcal{C}$ such that $FX \approx Y$ exists. But we may still ask for a best approximation to such an object. So among objects $X \in \mathcal{C}$ such that FX maps to Y , we may look for a universal one.

Let I_Y denote the category whose objects are pairs (X, ϕ) , where $X \in \mathcal{C}$ and $\phi : FX \longrightarrow Y$ is a map in \mathcal{D} . A morphism $(X_1, \phi_1) \longrightarrow (X_2, \phi_2)$ in I_Y is defined to be a morphism $X_1 \xrightarrow{u} X_2$ in \mathcal{C} which is compatible with ϕ_i , i.e., such that $\phi_2 = \phi_1 \circ Fu$.

By the defining property of direct limits, $\varinjlim_{(X, \phi) \in I_Y} FX$ maps to Y , provided that the limit exists. Suppose that F is compatible with direct limits. Then we obtain a map $F\overline{X} \approx \varinjlim_{I_Y} FX \longrightarrow Y$, call it $\overline{\phi}$. This map is the universal one we are looking for.

We can define a functor $G : \mathcal{C} \longleftarrow \mathcal{D}$ by setting $GY = \overline{X}$. Morphisms can be defined in a natural way, using the fact that if a map $Y_1 \xrightarrow{v} Y_2$ in \mathcal{D} is given, then any map $FX \xrightarrow{\phi} Y_1$ yields a map $FX \xrightarrow{v\phi} Y_2$ by composition.

Notice the following property for the functor G constructed in this way: Given a map $FX \longrightarrow Y$ in \mathcal{D} , there is a canonical map $X \longrightarrow \overline{X} = GY$, and conversely. In other words, we have a bijection

$$\mathrm{Hom}_{\mathcal{C}}(FX, Y) \longrightarrow \mathrm{Hom}_{\mathcal{D}}(X, GY).$$

Two functors $\mathcal{C} \xrightarrow{F} \mathcal{D}$ and $\mathcal{C} \xleftarrow{G} \mathcal{D}$ are said to be *adjoint* if there is a natural bijection

$$\mathrm{Hom}_{\mathcal{C}}(FX, Y) \xrightarrow{\theta} \mathrm{Hom}_{\mathcal{D}}(X, GY),$$

for $X \in \mathcal{C}$ and $Y \in \mathcal{D}$. (The word “natural” means that θ is compatible with maps in both variables.) Then F is called a *left adjoint* of G and G is a *right adjoint* of F . The above construction leads to the adjoint functor theorem of Kan and Freyd (see [McL], p. 125), which is quoted below. For the statement, we need to know one more definition: A *generator* for a category \mathcal{C} is an object U such that every object is a quotient of a direct sum of copies of U . For example, the module A_A is a generator for the category $\mathrm{Mod} A$.

Theorem I.7.1. *Let \mathcal{C}, \mathcal{D} be abelian categories with exact direct limits. Suppose that \mathcal{C} has a set of generators. Then a functor $F : \mathcal{C} \rightarrow \mathcal{D}$ has a right adjoint if and only if it is right exact and compatible with direct sums. \square*

Exercise: State the theorem for existence of left adjoints, and describe their construction by reversing arrows appropriately in the above discussion.

Proposition I.7.2. *The functors $- \otimes_A P$ and $\mathrm{Hom}_B(P, -)$ are adjoint. In other words, for $M \in \mathrm{Mod} A$ and $N \in \mathrm{Mod} B$, there are natural bijections*

$$\mathrm{Hom}_B(M \otimes_A P, N) \xrightarrow{\theta} \mathrm{Hom}_A(M, \mathrm{Hom}_B(P, N)).$$

Example I.7.3: (*extension and restriction of scalars*) Suppose we are given a ring homomorphism $\phi : A \rightarrow B$. A right B -module N_B can be made into a right A -module $N|_A$ by *restriction of scalars*, which means that A acts via the homomorphism ϕ : $na = n(\phi a)$. Similarly, we can restrict scalars in a left module, so we can make ${}_B B_B$ into an A, B bimodule ${}_A B_B$ by restricting the left action. Then the right A -module $\mathrm{Hom}_B({}_A B_B, N_B)$ is just $N|_A$. So, setting $P = {}_A B_B$, the adjointness formula reads

$$\mathrm{Hom}_B(M \otimes_A B, N) \approx \mathrm{Hom}_A(M, N_A).$$

Here $- \otimes_A B$ is the *extension of scalars* from A to B : Extension and restriction are adjoint functors.

Proof of Proposition I.7.4. We already know that $- \otimes_A P$ is right exact and compatible with direct sums, and that $\mathrm{Mod} A$ has a generator. So a right adjoint exists. Unfortunately this fact is not of much help in identifying the adjoint.

We have to find a functorial bijection between the two sets $\mathrm{Hom}_B(M \otimes_A P, N)$ and $\mathrm{Hom}_A(M, \mathrm{Hom}_B(P, N))$. There is a general method to verify such a formula in three steps: Case 1: $M = A_A$. Case 2: M is a free module, a direct sum of copies of A . Case 3: The general case.

Suppose that $M = A$. Since $A \otimes_A P \approx P$ and $\mathrm{Hom}_A(A, X) \approx X$, the two sets are canonically equivalent, as required.

Next, suppose that $M = \bigoplus_I A$ is free. Since tensor product commutes with direct sums, $(\bigoplus_I A) \otimes P \approx \bigoplus_I (A \otimes P) \approx \bigoplus_I P$. Also, by the defining property of direct sums, $\mathrm{Hom}_A(\bigoplus_I X, Y) \approx$

$\Pi_I \text{Hom}_A(X, Y)$. Combining these two facts and setting $H = \text{Hom}_B(P, N)$, we obtain $\text{Hom}_B(M \otimes P, N) \approx \Pi H$, and $\text{Hom}_A(M, H) = \Pi \text{Hom}_A(A, H) \approx \Pi H$, as required.

Finally, for arbitrary M , choose a presentation of M by free modules, say

$$\mathcal{M} := M_1 \longrightarrow M_2 \longrightarrow M_3 \longrightarrow 0,$$

where M_1, M_2 are free and $M_3 = M$. Then $\mathcal{M} \otimes_A P$ is exact, hence $\text{Hom}_B(\mathcal{M} \otimes_A P, N)$ is left exact. Also, setting $H = \text{Hom}_B(P, N)$ as before, $\text{Hom}_A(\mathcal{M}, H)$ is left exact. So we have a commutative diagram

$$\begin{array}{ccccccc} 0 & \longrightarrow & \text{Hom}_B(M_3 \otimes_A P, N) & \longrightarrow & \text{Hom}_B(M_2 \otimes_A P, N) & \longrightarrow & \text{Hom}_B(M_1 \otimes_A P, N) \\ & & \downarrow & & \downarrow & & \downarrow \\ 0 & \longrightarrow & \text{Hom}_A(M_3, H) & \longrightarrow & \text{Hom}_A(M_2, H) & \longrightarrow & \text{Hom}_A(M_1, H) \end{array}$$

in which the second and third vertical arrows are isomorphisms. It follows that the left hand vertical arrow is also an isomorphism.

This argument isn't complete, because it isn't clear that the bijections described for free modules are functorial (i.e. compatible with maps). Without that, we don't know that the right hand square is commutative. So to complete this argument, we will have to prove functoriality. Functoriality is "easy to see". (When you see "easy to see" in mathematics, it usually means either "I'm too lazy to think this through, but I think it is true", or else "This is too grungy to write down". It rarely means easy to see in the literal sense. If something were truly easy to see, then the author would probably have told us how to see it, or else wouldn't have thought it necessary to emphasize the point.)

An alternative method of proof is to write the bijection down in a functorial way at the start. This approach is less elementary, and it replaces one "easy to see" with another one. Here is the description of the functorial map θ . Given a B -map $\phi : M \otimes_A P \longrightarrow N$, we have to exhibit its image, an A -map $\psi : M \longrightarrow H$. So given ϕ and an element $m \in M$, we need to define $\psi_m \in H$. We set $\psi_m(p) = \phi(m \otimes p)$. This formula also defines the inverse of θ . The notation has become cumbersome, and there are many points to check, but all the verification is straightforward. This method is perhaps better. The first approach as a good heuristic argument. \square

Lemma I.7.5. (i) Two right adjoints G_1, G_2 of a functor F are isomorphic functors.
(ii) If F is an equivalence and G is a quasi-inverse of F , then G is a left and also a right adjoint of F .

Proof. (i) This follows from the fact that an object Z in a category \mathcal{C} is determined by the functor $\text{Hom}_{\mathcal{C}}(-, Z)$, up to canonical isomorphism (the *Yoneda Lemma* [McL]). The adjoint property determines the functor $\text{Hom}_{\mathcal{C}}(-, G_i Y)$. \square

Proposition I.7.6. The following are equivalent for a functor $F : \text{Mod } A \rightarrow \text{Mod } B$.

- (i) There is an A, B -bimodule ${}_A P_B$ and an isomorphism of functors $- \otimes_A P \approx F$.
- (ii) F has a right adjoint.
- (iii) F is right exact and compatible with direct sums.

Proof. (ii) \Leftrightarrow (iii) follows from the adjoint functor theorem. (i) \Rightarrow (ii) is the exactness property of $-\otimes_A P$.

(ii) \Rightarrow (i): Let G be a right adjoint of F . Then for $N \in \text{Mod } B$,

$$GN \approx \text{Hom}_A(A, GN) \approx \text{Hom}_B(FA, N) = \text{Hom}_B(P, N).$$

So $G \approx \text{Hom}_B(P, -)$. Since the left adjoint of $\text{Hom}_B(P, -)$ is $-\otimes_A P$, the functors F and $-\otimes_A P$ are isomorphic. \square

I.8. Morita equivalence

Proposition I.8.1. *A functor $F : \text{Mod } A \rightarrow \text{Mod } B$ is an equivalence of categories if and only if there are bimodules ${}_A P_B$ and ${}_B Q_A$ such that*

(i) $F \approx -\otimes_A P$, and

(ii) $P \otimes_B Q$ is isomorphic as bimodule to ${}_A A_A$ and also $Q \otimes_A P$ is isomorphic to ${}_B B_B$.

If these conditions hold, then $-\otimes_B Q$ is a quasi-inverse of $-\otimes_A P$.

It follows that A, B, P, Q describes a Morita context:

$$\begin{pmatrix} A & P \\ Q & B \end{pmatrix}.$$

Proof. Suppose that F is an equivalence. Then F is exact and compatible with direct sums, so it has the form $-\otimes_A P$. Similarly, the quasi-inverse G of F has the form $-\otimes_B Q$. The isomorphism $GF \approx \text{id}_{\text{Mod } A}$ gives us an isomorphism $Q \otimes_B P \approx A \otimes_A P \otimes_B Q \approx A_A$ which is functorial, hence compatible with left multiplication, i.e. a bimodule isomorphism. Similarly, $FG \approx \text{id}_{\text{Mod } B}$ yields ${}_B B_B \approx Q \otimes_A P$.

Conversely, suppose that conditions (i) and (ii) are satisfied. Then using the isomorphisms $M \approx M \otimes_A P \otimes_B Q$ and $N \approx N \otimes_B Q \otimes_A P$ one shows that F is fully faithful and essentially surjective, hence an equivalence. Then the isomorphisms (ii) show that $-\otimes_B Q$ is a quasi-inverse of $-\otimes_A P$. \square

Proposition I.8.2. *(the main example) The category of right modules over a ring B is equivalent to the category of right modules over the algebra $M_n(B)$ of $n \times n$ matrices with entries in B .*

Proof. Let $F = B^n$ be the right module of n -dimensional row vectors with entries in the ring B , let $F^* = B^n = \text{Hom}_B(F, B)$ be the left module of n -dimensional column vectors, and let $A = M_n(B) = \text{End } F_B$ be the matrix algebra. So ${}_B F_A$ and ${}_A F_B^*$ are bimodules. Matrix multiplication defines an obvious bijection $F^* \otimes_B F \xrightarrow{\eta} A$, and also a surjective map $F \otimes_A F^* \xrightarrow{\theta} B$. These maps define a Morita context

$$\begin{pmatrix} B & F \\ F^* & A \end{pmatrix}.$$

Since θ is surjective, it is bijective (see Proposition I.8.4 below). Therefore the previous proposition applies. \square

Proposition I.8.3. *Let P, Q be as in proposition I.8.1. Then*

- (i) *For $N \in \text{Mod } B$, there are natural isomorphisms $N \otimes_B Q \approx \text{Hom}_B(P, N)$.*
- (ii) *$Q \approx \text{Hom}_B(P, B)$ is the dual module of P_B ,*
- (iii) *$A \approx \text{Hom}_B(P, P) = \text{End } P_B$,*
- (iv) *P_B is a projective B -module,*
- (v) *P_B is a generator for $\text{Mod } B$, and*
- (vi) *P_B is a finitely generated B -module.*

Proof. (i) Both functors in question are right adjoints of $- \otimes_A P$.

(ii) Set $N = B$ in (i), to obtain $Q \approx B \otimes_B Q \approx \text{Hom}_B(P, B)$.

(iii) Set $N = P$ in (i), to obtain $A \approx P \otimes_B Q \approx \text{Hom}_B(P, P)$.

(iv,v) These assertions follow F is an equivalence of categories, A_A is a projective generator of $\text{Mod } A$, and $P = FA$.

(vi) This is the only tricky part of the proposition. We'll use the non-obvious part of the description of Morita context A, B, P, Q , which is that the two ways of contracting $P \otimes_B Q \otimes_A P \longrightarrow P$ are equal. Let's label the isomorphisms I.8.1(ii), say $P \otimes_B Q \xrightarrow{\theta} A$ and $Q \otimes_A P \xrightarrow{\eta} B$. On tensors $p \otimes q \otimes p$, the equality of the two contractions reads $ap = pb$, where $a = \theta(p \otimes q)$ and $b = \eta(q \otimes p)$. (Note: for these contractions to be equal, θ and η have to be chosen compatibly.)

First, since $P \otimes_B Q \approx A$, we can find finite sets $p_i \in P$ and $q_i \in Q$ such that $\theta(\sum p_i \otimes q_i) = 1_A$. We claim that the set $\{p_i\}$ generates P_B . To see this, let $p \in P$. Then $p = 1p = \sum \theta(p_i \otimes q_i)p = \sum p_i \eta(q_i \otimes p) = \sum p_i b_i$, as required. \square

The reasoning used above also shows the following

Proposition I.8.4. *Let A, B, P, Q be a Morita context. If the canonical map $\theta : P \otimes_B Q \longrightarrow A$ is surjective, then it is bijective.*

Proof. As above, we use surjectivity to write $\theta(\sum p_i \otimes q_i) = 1_A$. Suppose that x is in the kernel of θ , and say $x = \sum u_j \otimes v_j$, with $u_j \in P$ and $v_j \in Q$. So $\theta(\sum u_j \otimes v_j) = 0$. We look at the element $z = \sum p_i \otimes q_i \otimes u_j \otimes v_j$ in the four fold tensor product $P \otimes_B Q \otimes_A P \otimes_B Q$. We have $1 \otimes 1 \otimes \theta = 1 \otimes \eta \otimes 1 = \theta \otimes 1 \otimes 1$. Evaluating the left map on z yields 0, while the right map yields x . Thus $x = 0$. \square

Proposition I.8.1 identifies the bimodules P which define equivalences of categories as invertible bimodules. The following theorem gives another characterization of such bimodules.

Theorem I.8.5. *Let P_B be a finitely generated projective generator of $\text{Mod } B$, and let $Q = \text{Hom}_B(P, B)$ and $A = \text{End } P_B$. Then P is an A, B -bimodule, and the functor $F := - \otimes_A P$ is an equivalence of categories.*

Proof. Note that the previous proposition is essentially the converse to this theorem. The first step is

Lemma I.8.6. *Let P_B be an arbitrary right B -module, let $Q = \text{Hom}_B(P, B)$ and $A = \text{End}_P B$. There is a canonical Morita context*

$$\begin{pmatrix} A & P \\ Q & B \end{pmatrix}.$$

Proof. The fact that ${}_A P_B$ and ${}_B Q_A$ are bimodules follows from our discussion of the functor Hom . We need to define the maps $P \otimes_B Q \xrightarrow{\theta} A$ and $Q \otimes_A P \xrightarrow{\eta} B$. We can evaluate an element $q \in Q = \text{Hom}_B(P, B)$ on $p \in P$. Let us denote the result by $\langle q, p \rangle$. Then our map $Q \otimes_A P \xrightarrow{\eta} B$ is defined on a tensor $q \otimes p$ by $\eta(q \otimes p) = \langle q, p \rangle$. Next the map $P \otimes_B Q \xrightarrow{\theta} A$ is defined as follows: Given a tensor $p \otimes q$, we have to exhibit an element of A , i.e., an endomorphism of P_B . The rule is $\theta(p \otimes q)p = p\langle q, p \rangle$. These maps obviously satisfy the requirement that the two contractions $P \otimes_B Q \otimes_A P \rightarrow P$ are equal. For the two contractions of a tensor $q \otimes p \otimes q$, we note that the results are elements of Q , which are determined by their actions on elements $p \in P$. To evaluate on p means to contract the four-fold tensor $q \otimes p \otimes q \otimes p$ to an element of B . Starting with $\eta \otimes 1 \otimes 1$ leads to $\langle q, p \rangle \langle q, p \rangle$, and the second contraction is

$$q \otimes (p \otimes q) \otimes p \mapsto q \otimes \theta(p \otimes q) \otimes p = q \otimes p\langle q, p \rangle \mapsto \langle q, p \rangle \langle q, p \rangle.$$

So the two are equal, as required. \square

Next, we use the fact that P_B is a generator.

Lemma I.8.7. *With the above notation, the map $Q \otimes_A P \xrightarrow{\eta} B$ is bijective.*

Proof. Since P_B generates the category $\text{Mod } B$, there is a surjective map $\bigoplus_I P \rightarrow B_B$. In order for such a map to be surjective, it suffices that 1 be in its image, which will be true for a finite sum. So we may assume that I is finite. The map $\bigoplus_I P \rightarrow B$ is determined by a collection of maps $\{q_i : P \rightarrow B\}_{i \in I}$, a collection of elements of Q . So we can write $1_B = \sum \langle q_i, p_i \rangle$ for some $p_i \in P$. Therefore η is surjective, and by Proposition I.8.4, it is bijective. \square

Next, we use the fact that P_B is a finitely generated projective module. There is a finitely generated free module, say $F = B^n$, such that $F \approx P \oplus P$. What can we deduce from proposition 8/2? With the notation of that proposition, $F^* \approx Q \oplus Q$, where $Q = \text{Hom}_B(P, B)$, and, denoting $\text{Hom}_B(P, P)$ by (P, P) for short, we have a Peirce decomposition

$$M_n(B) \approx \begin{pmatrix} (P, P) & (P, P) \\ (P, P) & (P, P) \end{pmatrix}.$$

The projection $P \rightarrow F \rightarrow P$ is the relevant idempotent element in $M_n(B)$. So $A = (P, P) = eM_n(B)e$, $P = eF$, and $Q = F^*e$. Since $F \otimes_B F^* \approx M_n(B)$, we also have $P \otimes_B Q = eF \otimes_B F^*e \approx A$. This shows that ${}_A P_B$ and ${}_B Q_A$ are inverse bimodules, and completes the proof \square

Exercise: Show that if A, B are commutative and Morita equivalent, then they are isomorphic.

Exercise: Let A, B be rings. Prove that if their categories of right modules are equivalent, then so are their categories of left modules.

Reference:

[McL] S. MacLane, *Categories for the working mathematician*, Springer 1971.

II. LOCALIZATION and GOLDIE'S THEOREM

This discusses the possibility of using fractions in noncommutative algebra. For example, we would like to embed a domain A in a skew field of fractions.

II.1. Terminology:

An element $s \in A$ is *regular* if it is not a left or right zero divisor, i.e., if $as = 0$ implies $a = 0$ and also $sa = 0$ implies $a = 0$. If s is regular, then we can cancel: $sa = sb$ implies that $a = b$ and $as = bs$ implies $a = b$.

The *right annihilator* of a subset X of a module M_A is the right ideal $rann(X) = \{a \in A \mid Xa = 0\}$. Similarly, the *left annihilator* of a subset $X \subset_A M$ is the left ideal $lann(X) = \{a \in A \mid aX = 0\}$.

Thus $s \in A$ is regular element if and only if $rann(s) = 0$ and $lann(s) = 0$. (I can never remember whether $rann(s) = 0$ means that s is left regular or that it is right regular.)

Lemma II.1.1. *If a regular element s has a right inverse, then it is invertible.*

Proof. Suppose that $st = 1$ and that its right annihilator $rann(s)$ is zero. Then we can cancel s on the left in the equation $sts = s$ to obtain $ts = 1$. \square

II.2. Ore Sets:

We will work with *right fractions* as^{-1} , where $a, s \in A$. The difficulty with fractions is that the formal product of two right fractions has no obvious interpretation as a fraction: To interpret $(bs^{-1})(at^{-1})$ as a fraction, we need to have a method to “move s^{-1} past a ”. In other words, we need to be able to rewrite a left fraction $s^{-1}a$ as a right fraction, say as $a_1s_1^{-1}$.

Suppose for the moment that A is a domain. Working formally with the relation $s^{-1}a = a_1s_1^{-1}$, we multiply on left by s and on right by s_1 , obtaining the peculiar equation

$$(II.2.1) \quad sa_1 = as_1.$$

In order to have a hope of success, for given a and s , there must exist elements a_1, s_1 , with $s_1 \neq 0$, such that (II.2.1) is true. This is called the *Ore condition*.

Now in many cases one can't solve (II.2.1) for s_1, a_1 . For example, let k be a field and let $A = k\langle x, y \rangle$ be the free ring of noncommutative polynomials in x, y . Setting $a = x, s = y$, the equation (II.2.1) becomes $yf = xg$, which has no nonzero solution for f, g in A . The correct method of embedding A into a skew field does not use fractions. (See P.M. Cohn, *Free rings and their relations*) The next proposition exhibits the close relation between the Ore condition and free rings.

Proposition II.2.2. (*Jategaonkar*) *Let A be a domain which is an algebra over a field k . Let $a, s \in A$, with $s \neq 0$. If the right Ore condition (II.2.1) fails for this pair of elements, then they generate a free subring of A , so A contains a free ring.*

Proof. Suppose that a, s do not generate a free ring, so that $f(a, s) = 0$, where $f(x, y)$ is a nonzero, noncommutative polynomial with coefficients in k . We choose f of minimal degree, and write

$f(x, y) = c + xu + yv$, where $c \in k$ is the constant term of f , and where $u = u(x, y)$ and $v = v(x, y)$ are noncommutative polynomials. Multiplying on the right by y , we obtain $fy = cy + xuy + yvy$, hence $yq = xp$ in A , with $p = uy$ and $q = -(c + vy)$. Setting $a_1 = q(a, s)$ and $s_1 = p(a, s)$, we obtain the required equation $sa_1 = as_1$ in A unless $s_1 = 0$ in A . Now if $s_1 = u(a, s)s = 0$, in A , then because $s \neq 0$ and A is a domain, $u(a, s) = 0$. Since f has minimal degree, $u(x, y) = 0$, $f = c + yv$, and $0 = c + sv(a, s)$. If $c \neq 0$, then s is invertible (1.1), and (II.2.1) can be solved with $s_1 = 1$. If $c = 0$, then $v(a, s) = 0$, hence $v(x, y) = 0$. But then $f = 0$, which is a contradiction. \square

In spite of the negative evidence, Oystein Ore in the 1930s investigated the problem of deciding when fractions could be used. And thanks to the work of Goldie in the 1950s, fractions have become a cornerstone of the theory of noncommutative noetherian rings.

Let A be a ring. We'll call a subset $S \subset A$ a *right Ore set* if it has the following properties:

(II.2.3)

- (a) S is closed under multiplication, and $1 \in S$.
- (b) The elements of S are regular.
- (c) *right Ore condition*: For all $a \in A$ and $s \in S$, there exist $a_1 \in A$ and $s_1 \in S$ such that $sa_1 = as_1$.

The Ore condition (c) can be restated by saying

$$sA \cap aS \neq \emptyset.$$

As in commutative algebra, the requirement that S consists of regular elements can be relaxed, but never mind.

Exercise: Let $A = k\langle x, y \rangle / (yx - xy - 1)$ be the Weyl algebra. Show that the set S of powers of x is a right Ore set.

Exercise: Let R be a ring and let $\phi : R \rightarrow R$ be an injective ring homomorphism. Let $A = R[x, \phi]$ be the Ore extension, the polynomial ring in which scalars commute with x by the action of ϕ , i.e., $xa = a^\phi x$. Show that the set S of powers of x is a left Ore set, and that it is a right Ore set if and only if ϕ is bijective.

Theorem II.2.4. (Ore) *Let $S \subset A$ be a right Ore set. There is a ring AS^{-1} of right fractions and an injective ring homomorphism $A \rightarrow AS^{-1}$ such that*

- (a) *the image of every element $s \in S$ is invertible in AS^{-1} , and*
- (b) *every element of AS^{-1} can be written as a product as^{-1} .*

Moreover, any homomorphism $A \xrightarrow{f} R$ such that the images of elements of S are invertible in R factors uniquely through AS^{-1} .

Corollary II.2.5. *Suppose that S is both a left and a right Ore set. Then the rings of left fractions and of right fractions are isomorphic \square*

The next theorem is one of the most important tools for the study of noetherian rings. The proof, which is in sections 5 and 6, has been analyzed many times, but it still seems remarkable. It is a triumph of abstract ring theory.

Theorem II.2.6. (Goldie) *Let A be a right noetherian, semiprime ring. The set S of regular elements in A is a right Ore set, and AS^{-1} is a semisimple ring, a sum of matrix algebras over a skew fields.*

II.3. Construction of the Ring of Fractions

Let S be a right Ore set in a ring A . To construct the ring of fractions $B = AS^{-1}$, we use the Ore condition to change left fractions to right fractions. However, we must verify that this procedure is consistent, and the number of points which have to be checked is unpleasantly large. Moreover, some of the points can be confusing. So we proceed indirectly as follows: We first construct B as a left A -module, which is easy. Then we show that B is a ring by identifying it as the ring of endomorphisms $E = \text{End}_A B$ of the module we have constructed. This step requires a bit of work, but the approach has benefits: For one thing, we know from the start that E is a ring.

Lemma II.3.1. (i) *Suppose that s, t are elements of a right Ore set S , and that for some $x \in A$, $sx = t$. Then x is regular.*

(ii) (existence of common denominators) *Let $s, s \in S$. There is a common multiple $t \in S$, i.e., $t = sx$ and $t = s x$ for some regular elements $x, x \in A$.*

(iii) (variation of the elements a_1, s_1 in the Ore condition) *With the notation of the Ore condition, suppose that $sa_1 = as_1$ and also $sa_2 = as_2$, with $s_1, s_2 \in S$. There are regular elements $x_1, x_2 \in A$ such that $s_1x_1 = s_2x_2$ and $a_1x_1 = a_2x_2$.*

Proof. (i) Because $\text{rann}(t) = 0$, the equation $sx = t$ implies $\text{rann}(x) = 0$. To show that $\text{lann}(x) = 0$ requires the Ore condition. We set $s = t$ and $a = s$ in the (II.2.3c), obtaining $ta_1 = ss_1$, with $s_1 \in S$. Then $sxa_1 = ta_1 = ss_1$, and we may cancel s to obtain $xa_1 = s_1$. Since $\text{lann}(s_1) = 0$, $\text{lann}(x) = 0$ too.

(ii) We set $s = s$ and $a = s$ in the Ore condition, obtaining $sa_1 = s s_1$. Since S is closed under multiplication, $t = s s_1 \in S$. So we may take $x = a_1$ and $x = s_1$. The fact that x is regular follows from (i).

(iii) We choose a common multiple $s_1x_1 = s_2x_2$ in S with x_i regular. Then $sa_1x_1 = sa_2x_2$. Since s is a regular element, $a_1x_1 = a_2x_2$. \square

Of course, the elements of the ring we are looking for are equivalence classes of fractions, not the fractions themselves. Two fractions as^{-1} and $a s^{-1}$ are defined to be equivalent if there are elements $x, x \in A$ such that $sx \in S$, $sx = s x$, and $ax = a x$. One has to verify that this is indeed an equivalence relation. I like to replace these verifications by describing the equivalence classes as elements of a direct limit. If you don't want to do this, just skip ahead.

To interpret fractions as a limit, we'll make the set S into a category, by defining a map $s \rightarrow t$ between elements $s, t \in S$ to be an element $x \in A$ such that $sx = t$.

Lemma II.3.2. *There is at most one arrow $s \rightarrow t$ for $s, t \in S$, and any two elements $s, s \in S$ admit maps to some $t \in S$. Hence the category S is filtering.*

Proof. First, given $s, t \in S$, there is at most one arrow $s \rightarrow t$. Indeed, if $sx = t$ and $sy = t$, then $sx = sy$, and because s is regular, $x = y$. Next, given $s, s \in S$, we must find an element t and maps $s \rightarrow t$ and $s \rightarrow t$. We take a common multiple $t = sx = s x$. \square

Let As^{-1} denote the left A -module generated freely by an element denoted s^{-1} . The elements of As^{-1} are our formal fractions as^{-1} . So ${}_A A$ is canonically isomorphic to ${}_A As^{-1}$, by $a \mapsto as^{-1}$. We define a functor from S to the category of left A -modules by sending $s \mapsto As^{-1}$. If $s \xrightarrow{x} t$, i.e., $t = sx$, the map $As^{-1} \rightarrow At^{-1}$ corresponds to right multiplication by x on A .

$$\begin{array}{ccc} s & As^{-1} & \xleftarrow{s^{-1}} A \\ x \downarrow & & \downarrow x \\ t & At^{-1} & \xleftarrow{t^{-1}} A \end{array}$$

This reflects the heuristic computation

$$as^{-1} = (ax)(sx)^{-1} = (ax)t^{-1}.$$

We define $B = \varinjlim As^{-1}$. Elements of B are represented by formal fractions as^{-1} , and two fractions as^{-1} , $a's^{-1}$ represent the same element b if there is a common multiple $t = sx = s'x'$ in S such that $ax = a'x'$, i.e., if they are equivalent fractions in the sense defined above. As is customary in arithmetic, we may refer to a fraction informally as an element of B . An element $a \in A$ can stand for the fraction $a1^{-1}$ and s^{-1} for the fraction $1s^{-1}$.

Lemma II.3.3. (i) B is a left A -module, the product of $\alpha \in A$ with a fraction as^{-1} being $(\alpha a)s^{-1}$. There are canonical injective A -linear maps $As^{-1} \rightarrow B$, in particular, there is an injective map ${}_A A \rightarrow {}_A B$

(ii) Left multiplication by $s \in S$ on B is injective. \square

Since B is a left A -module, there is a canonical bijection of left modules ${}_A B \rightarrow \text{Hom}({}_A A, {}_A B)$. It sends an element $\beta \in B$ to right multiplication by β , acting on A , and its inverse evaluates a homomorphism $A \rightarrow B$ on the identity element 1_A . The main point of the construction of the ring of fractions is the next proposition.

Proposition II.3.4. For every $\beta \in B$, right multiplication by β on A extends uniquely to an endomorphism $\tilde{\beta}$ of ${}_A B$. This extension provides a bijection $B \rightarrow \text{End}_A B$.

Proof. Case 1: $\beta \in A$. Let us rename β as α . To define $\tilde{\alpha} : B \rightarrow B$, we represent an element $b \in B$ by a fraction as^{-1} . Then heuristically, we want to define $b\tilde{\alpha}$ as the product $as^{-1}\alpha$. To get an element of B , we must “move s^{-1} past α ”. We apply the Ore condition (II.2.3c), writing $s\alpha_1 = \alpha s_1$, and we set

$$(II.3.5) \quad b\tilde{\alpha} = (a\alpha_1)s_1^{-1}.$$

We must show that this is well-defined by verifying independence of the two choices made:

- a) the choice of α_1, s_1 in the Ore condition, and
- b) the choice of the fraction as^{-1} representing b .

a) We may assume that a second choice $s\alpha_2 = \alpha s_2$ has the form $\alpha_2 = \alpha_1 x$, $s_2 = s_1 x$ (see II.3.1). Our procedure yields $(a\alpha_2)(s_2)^{-1} = (a\alpha_1 x)(s_1 x)^{-1}$, which is a fraction equivalent with (II.3.5).

b) Because of the existence of common multiples, it suffices to treat the case that a second fraction representing b has the form $a s^{-1}$ where $a = ax$ and $s = sx$. Our procedure tells us to write $s \alpha_2 = \alpha s_2$ and to use the fraction $(a \alpha_2) s_2^{-1}$. Taking common multiples again, we may assume that $s_2 = s_1$, where s_1 is as in (II.3.5). This is allowable because a) has been verified. Then $\alpha s_1 = s \alpha_1$ and also $\alpha s_1 = s \alpha_2 = s x \alpha_2$. Since s is regular, $s \alpha_1 = s x \alpha_2$ implies that $\alpha_1 = x \alpha_2$. So $(a \alpha_2) s_2^{-1} = (a x \alpha_2) s_1^{-1} = (a \alpha_1) s_1^{-1}$, as required.

To show that the map $\tilde{\beta}$ we have constructed is A -linear, we may represent a pair of elements $b, b' \in B$ by fractions using a common denominator. Then linearity is clear from the formula (II.3.5).

Case 2: $\beta = \sigma^{-1}$. This case does not require the Ore condition: With notation as above, we simply identify the formal product $a s^{-1} \sigma^{-1}$ in the obvious way, as the fraction $a(\sigma s)^{-1}$. Independence of the choice of fraction representing b and A -linearity of $\tilde{\beta}$ are easy to check.

Case 3: $\beta \in B$ is arbitrary. We represent β by a fraction $\alpha \sigma^{-1}$. Then right multiplication by β on A is the composition of the two maps $A \xrightarrow{\alpha} A \xrightarrow{\sigma^{-1}} B$. This composition extends to B in two steps:

$$(II.3.6) \quad \begin{array}{ccccc} A & \xrightarrow{\alpha} & A & \xrightarrow{\sigma^{-1}} & B \\ \downarrow & & \downarrow & & \downarrow \\ A & \xrightarrow{\alpha} & B & \xrightarrow{\tilde{\sigma}^{-1}} & B \\ \downarrow & & \downarrow & & \downarrow \\ B & \xrightarrow{\tilde{\alpha}} & B & \xrightarrow{\tilde{\sigma}^{-1}} & B \end{array}$$

To complete the discussion, we must check that $\tilde{\beta}$ is the only A -linear extension of $A \xrightarrow{\beta} B$. This follows from the next lemma.

Lemma II.3.7. *An endomorphism ϕ of ${}_A B$ is determined uniquely by the element $\beta := 1\phi$.*

Proof. Let $b = a s^{-1}$ be an element of B , and let $x = s^{-1}\phi$. Since ϕ is A -linear, $b\phi = ax$. We have $sx = s(s^{-1}\phi) = (ss^{-1})\phi = \beta$. Since s is a regular element, sx determines x for every $x \in B$. Thus x is determined uniquely by β , and so is $b\phi$. \square

Proof of Theorem II.2.4. Since $E = \text{End}_A B$ is a ring, the bijection $B \approx E$ sending $\beta \mapsto \tilde{\beta}$ defines a ring structure on B . By definition, multiplication in B corresponds to composition of operators in E . We verify the assertions of the theorem:

(i) The bijection is an isomorphism of the (previously defined) left module structures ${}_A B \longrightarrow {}_A E$.

Let $\beta_3 = \alpha_1 \beta_1 + \alpha_2 \beta_2$ in B . To show that $\tilde{\beta}_3 = \alpha_1 \tilde{\beta}_1 + \alpha_2 \tilde{\beta}_2$, it suffices to show that these two operators take the same values on the element 1, which is true.

(ii) The canonical map $A \longrightarrow B$ (II.3.3) is a ring homomorphism.

This follows from (i).

(iii) The elements of S are invertible in the ring B , and every element of B is a product as^{-1} , where $a \in A$ and $s \in S$.

Clear, I think.

(iv) Multiplication in B is the unique associative law which agrees with left multiplication by A on ${}_A B$.

Let $*$ denote such a law. Agreement with the left multiplication by A means that $a * b = ab$ for $a \in A$ and $b \in B$. Then $s(s^{-1} * (a s^{-1})) = s * s^{-1} * a * s^{-1} = a * s^{-1} = a s^{-1}$, and also $s(s^{-1} a s^{-1}) = a s^{-1}$. Cancelling s , $s^{-1} * (a s^{-1}) = s^{-1} a s^{-1}$. Then to show that $b * b = bb$, we represent by fractions: $b * b = (a s^{-1}) * (a s^{-1})$, and apply what has been shown.

(v) A ring homomorphism $\phi : A \rightarrow R$ such that the images of elements of S are invertible extends to B .

We'll omit the verification of this statement. \square

II.4. Modules of Fractions

Let S be a right Ore set in A , and let M be a right A -module. We may form a module of fractions $M S^{-1}$ in two ways: First, as a direct limit $M_1 = \varinjlim M s^{-1}$ analogous to the construction of $A S^{-1}$. But note that since M has no left A -module structure, neither does $M s^{-1}$. So M_1 is, a priori, only an abelian group. Also, right multiplication by $s \in S$ need not be injective on M , so the canonical maps $M s^{-1} \rightarrow M_1$ may not be injective. However, M_1 has the advantage of being an exact functor of M , because \varinjlim is exact.

The second way to construct a module of fractions is as the tensor product $M_2 = M \otimes_A A S^{-1}$. This yields a right $A S^{-1}$ -module, but a priori, tensor product is only right exact.

Proposition II.4.1. (i) *There is a canonical bijection of abelian groups $M_1 \rightarrow M_2$, sending $m s^{-1} \mapsto m \otimes s^{-1}$.*

(ii) *$A S^{-1}$ is flat over A , i.e., $- \otimes_A A S^{-1}$ is an exact functor.*

Proof. (i) Independence of the fraction representing an element of M_1 , and linearity of the map are easy and similar to verification done in the previous section. It is clear that every element of $M \otimes_A A S^{-1}$ can be represented by a tensor $m \otimes s^{-1}$, so the map is surjective. To show that it is bijective, we note that both constructions are right exact and compatible with direct sums, and that the map is bijective when $M = A_A$. Hence it is bijective in general. (ii) follows from (i). \square

II.5. Essential Submodules and Goldie Rank

A submodule E of a right A -module M is *essential* if for every nonzero submodule X of M , $X \cap E \neq 0$.

Exercise: In a prime ring A , every nonzero two-sided ideal I is an essential right ideal.

Proposition II.5.1. (i) Let $M_1 \subset M_2 \subset M_3$ be modules. If M_1 is essential in M_2 and M_2 is essential in M_3 , then M_1 is essential in M_3 .

(ii) Let $f : M \rightarrow M$ be a homomorphism of right A -modules. If W is an essential submodule of M , then $f(W) := f^{-1}(W)$ is essential in M .

(iii) If E_i is an essential submodule of a right module M_i for $i = 1, \dots, n$, then $E_1 \oplus \dots \oplus E_n$ is an essential submodule of $M_1 \oplus \dots \oplus M_n$. \square

Proof. Assertion (i) is clear.

(ii) Let X be a nonzero submodule of M . Then $f(X)$ is a submodule of M . If $f(X) \neq 0$, then $f(X) \cap W \neq 0$, and $f(X) \cap W = f(X \cap W)$. If $f(X) = 0$, then $X \subset W$. In either case, $X \cap W \neq 0$.

(iii) By (ii), if E is essential in M , then $E \oplus N$ is essential in $M \oplus N$ for every N . Hence we may replace one E_i at a time by M_i in the direct sum $E_1 \oplus \dots \oplus E_n$, and apply (i). \square

Proposition II.5.2. Let M_1 be any submodule of a module M . There exists a complementary submodule M_2 of M , one such that $M_1 \cap M_2 = 0$ and $M_1 \oplus M_2$ is essential.

Proof. (i) By Zorn, we may take for M_2 a maximal submodule such that $M_1 \cap M_2 = 0$. Then $M_1 \oplus M_2$ is essential. \square

Corollary II.5.3. A module M has no proper essential submodule if and only if it is semisimple.

Proof. This follows from (II.5.2) and from the characterization of semisimple modules by the splitting of submodules. \square

A module M is *uniform* if it is not zero and if every nonzero submodule of M is essential. For example, if A is a commutative domain with field of fractions K , then a torsion-free module M is uniform if and only if $M \otimes_A K$ has dimension 1 over K . Also, $M = k[t]/(t^n)$ is an essential $k[t]$ -module. This definition seems a bit arbitrary from the point of view of commutative algebra, but it works well.

Exercise: Verify the above assertions, and describe all uniform modules over a commutative ring A .

Exercise: A right module U is uniform if and only if for every pair of nonzero elements $u_1, u_2 \in U$, there exist $a_1, a_2 \in A$ such that $u_1 a_1$ and $u_2 a_2$ are equal and nonzero.

Lemma II.5.4. A nonzero module M which is not uniform contains a direct sum of nonzero submodules.

Proof. This follows from (II.5.2). \square

A module M_A has *finite Goldie rank* if it does not contain an infinite direct sum of nonzero submodules. A noetherian module has finite Goldie rank.

Theorem II.5.5. (Goldie) Let M be a module of finite Goldie rank.

(i) If M is not zero, then it contains a uniform submodule.

(ii) M contains an essential submodule which is a direct sum of uniform submodules.

(iii) Let E be an essential submodule of M which is a direct sum of uniform submodules U_1, \dots, U_r , and suppose given another submodule which is a direct sum of nonzero submodules N_1, \dots, N_s . Then $r \geq s$. Hence the number r , called the Goldie rank $\text{goldie}(M)$, depends only on M .

(iv) Let M' be a submodule of M . Then the $\text{goldie}(M') \leq \text{goldie}(M)$, with equality if and only if M' is an essential submodule of M .

(v) $\text{goldie}(M) = 0$ if and only if $M = 0$.

Goldie rank is also called *uniform dimension*.

Exercise: Determine the Goldie rank of A_A when A is a matrix algebra over a skew field.

Exercise: Let A be a commutative domain with field of fractions K , and let M be a torsion-free A -module. Show that $\text{goldie}(M) = \dim_K(A \otimes_A M)$.

Proof of the theorem. (i),(ii) If M is not uniform, it contains an essential direct sum of nonzero submodules (II.5.4). If one of these submodules is not uniform, it contains an essential direct sum, etc... Since M does not contain an infinite direct sum, the process of replacing a non-uniform module by an essential direct sum stops.

(iii) Let W be a submodule of M , and $W_i = W \cap U_i$. If $W_i \neq 0$ for $i = 1, \dots, n$, then W_i is essential in U_i , hence (II.5.1) $W_1 \oplus \dots \oplus W_r$ is essential in E , and W is essential in M .

We may assume that $s > 0$. Then $W = N_2 \oplus \dots \oplus N_s$ is not essential in M , hence $W \cap U_i = 0$ for some i , say for $i = 1$. This allows us to replace N_1 by U_1 without changing s . Continuing, we may replace every N_j by an appropriate U_i , which implies that $s \leq r$.

(iv) This follows from (iii).

(v) This follows from (i). \square

II.6. Goldie's Theorem

Throughout this section, we assume that A is a right noetherian, semiprime ring. We recall the statement of (II.2.6).

Theorem II.6.1. (*Goldie*) *Let A be a semiprime, right noetherian ring. The set S of regular elements in A is a right Ore set, and $Q = AS^{-1}$ is a semisimple ring.*

The main part of the proof is to show that the set S of regular elements of A is an Ore set. Conditions (a) and (b) of the definition (II.2.3) hold, so we need only verify the Ore condition (c). This is done by the sequence of lemmas below.

Lemma II.6.2. *The left annihilator of an essential right ideal is zero.*

Proof. Suppose the contrary. Turning the statement around, there is a nonzero left ideal L whose right annihilator R is essential. We choose $L \neq 0$ such that the right annihilator R is essential and maximal among right annihilators of nonzero left ideals, and we obtain a contradiction by showing that $L = 0$.

Since A is semiprime, it suffices to show that $L^2 = 0$. If $L^2 \neq 0$, there are elements $x, y \in L$ such that $xy \neq 0$. Then yA is a nonzero right ideal and since R is essential, $yA \cap R \neq 0$: There is an element $a \in A$ such that $xya = 0$ but $ya \neq 0$. This implies that $\text{rann}(xy) > \text{rann}(y)$. But

$rann(y) = rann(Ay) \supset R$. Because R was chosen to be maximal, $rann(xy) = A$, and so $xy = 0$, which is a contradiction. \square

Exercise: Show that if N is an essential right ideal of a right noetherian ring A , then the left annihilator of N is nilpotent.

Lemma II.6.3. *Let $s \in A$. If $rann(s) = 0$, then s is a regular element and sA is an essential right ideal.*

Proof. If $rann(s) = 0$, then sA is isomorphic to A as right module, and so the Goldie ranks of A and sA are equal. By (II.5.8 iv), this implies that sA is an essential right ideal. Therefore (II.6.2) $lann(sA) = lann(s) = 0$, and s is regular. \square

Lemma II.6.4. *Every right ideal N of A contains an element x such that $rann(x) \cap N = 0$.*

Proof. Case 1: N is uniform. Since A is semiprime, $N^2 \neq 0$, and we may choose $x, y \in N$ such that $xy \neq 0$. We claim that then $W := rann(x) \cap N = 0$. Else, if $W \neq 0$, then because N is uniform, W is essential in N . Consider the homomorphism $\lambda_y : A_A \rightarrow N_A$ sending $\alpha \mapsto y\alpha$. By (II.5.1ii), $W = \lambda_y^{-1}(W)$ is an essential right ideal. But $yW \subset W$, hence $xyW = 0$. Since $xy \neq 0$, the left annihilator of W is not zero. This contradicts (II.6.2).

Case 2: The general case. We look at submodules $V \subset N$ which contain elements v with $rann(v) \cap V = 0$, and we choose V maximal among such submodules. We'll show that $V = N$ by showing that $rann(v) \cap N = 0$. Else, if $rann(v) \cap N \neq 0$, we choose a uniform submodule $U \subset rann(v) \cap N$. By what has been shown, there exists an element $u \in U$ such that $rann(u) \cap U = 0$. We set $x = u + v$, and we claim that $rann(x) \cap (U + V) = 0$. Since $U + V > V$, this will provide the contradiction that we are after.

We note that $U \cap V \subset rann(v) \cap V = 0$. So the sum $U + V$ is a direct sum. Suppose that $x \in rann(x) \cap (U \oplus V)$. So $xx = 0$ and $x = u + v$. Because the sum $U \oplus V$ is direct, $ux = 0$ and $vx = 0$, or $uu + uv = 0$ and $vu + vv = 0$. We know that $vu = 0$ because $U \subset rann(v)$. Hence $vv = 0$, and because $rann(v) \cap V = 0$, this implies that $v = 0$. Then $uu = 0$, which for an analogous reason implies that $u = 0$, i.e., $x = 0$, as required. \square

Lemma II.6.5. *Every essential right ideal E contains a regular element.*

Proof. Let $x \in E$ be as in Lemma (II.6.4): $rann(x) \cap E = 0$. Then because E is essential, $rann(x) = 0$ (II.6.2), which implies that x is regular (II.6.3). \square

We now verify that the right Ore condition $sA \cap aS \neq \emptyset$ holds for the set S of regular elements in a right noetherian, semiprime ring A . If s is regular, then sA is an essential right ideal (II.6.3). Consider the map $\lambda_a : A_A \rightarrow A_A$ which sends $\alpha \mapsto a\alpha$. By (II.5.1 ii), $\lambda_a^{-1}(sA)$ is also an essential right ideal, so it contains a regular element, say s_1 . Then $as_1 \in sA$, as required.

To complete the proof of Goldie's Theorem, we must show that the ring of fractions $Q = AS^{-1}$ is semisimple.

Lemma II.6.6. *Let S be a right Ore set in a ring A , and let $Q = AS^{-1}$.*

(i) *A is an essential submodule of $Q_A = AS_A^{-1}$.*

(ii) *If N is an essential right ideal of Q , then $N \cap A$ is an essential right ideal of A .*

Proof. (i) Since the inclusion $A \rightarrow Q$ is a ring homomorphism, A_A is a submodule of Q_A . Let X_A be a nonzero submodule of Q_A , and let $q = as^{-1} \in X$ be nonzero. Then $a = qs \in X \cap A$, and $a \neq 0$.

(ii) Let X be a nonzero right ideal of A . Then XS^{-1} is a nonzero right ideal of Q , hence $XS^{-1} \cap N \neq 0$. Clearing denominators shows that $X \cap N \neq 0$, and $X \cap N = X \cap (N \cap A)$. \square

To show that Q is semisimple, we show that Q has no proper essential right ideal (II.5.3). Let N_Q be an essential right ideal of Q . By (II.5.1 ii), $N \cap A$ is an essential right ideal of A , so it contains a regular element s (II.6.5). Then s is a unit in Q , hence N contains a unit, so $N = Q$. \square

Analysis of the above proof shows that one can weaken the noetherian hypothesis on A slightly. A ring A is called a *right Goldie ring* if it has finite Goldie rank, and also has the ascending chain condition on right annihilators, meaning that any ascending chain $R_1 \subset R_2 \subset \cdots$ of right ideals, each of which is the right annihilator of a subset X_i , is eventually constant.

Theorem II.6.7. (*Goldie*). *A semiprime right Goldie ring A has a semisimple right ring of fractions.*

III. CENTRAL SIMPLE ALGEBRAS and the BRAUER GROUP

III.1. Tensor product algebras

If A and B are algebras over a commutative ring R , then the tensor product $A \otimes B$ (we are writing \otimes for \otimes_R here) is made into a ring by the rule $(a \otimes b)(a' \otimes b') = (aa') \otimes (bb')$. There are canonical ring homomorphisms $A \rightarrow A \otimes B$ sending $a \mapsto a \otimes 1$, and $B \rightarrow A \otimes B$ sending $b \mapsto 1 \otimes b$. These maps are often injective - for example they are injective if R is a field and A, B are not zero. If so, we may speak informally of A and B as subrings of $A \otimes B$, using the notation a for the element $a \otimes 1$ and b for $1 \otimes b$. By definition of multiplication in $A \otimes B$, the images of A and B commute: $(a \otimes 1)(1 \otimes b) = a \otimes b = (1 \otimes b)(a \otimes 1)$.

The first part of the next lemma shows that tensor product makes the set of isomorphism classes of R -algebras into a commutative semigroup, and that the class of R is the identity element.

Lemma III.1.1. *Let R be a commutative ring, and let A, B, C be R algebras.*

(i) $A \otimes B \approx B \otimes A$, $A \otimes (B \otimes C) \approx (A \otimes B) \otimes C$, and $R \otimes A \approx A \approx A \otimes R$.

(ii) Let $M_n(A)$ denote the matrix algebra over A . Then $M_n(A) \otimes B \approx M_n(A \otimes B)$. \square

There are many games with tensor product algebras, and we are going to need quite a few of them in the next pages. The first is to turn an (A, B) -bimodule ${}_A M_B$ into a right module over a tensor product algebra. When A and B are given as algebras over a commutative ring R , an (A, B) -bimodule gets a structure of (R, R) -bimodule by restriction of scalars. In this situation, it is customary to assume tacitly that the left and right actions of the commutative ring R are the same: $rm = mr$. Then, since the operations of A and B on a bimodule commute, we could view the bimodule as a module over the tensor product except for one problem: A operates on the left and B on the right. If λ_a denotes left multiplication of $a \in A$ on M and if $a, a' \in A$, then $\lambda_{a'} \lambda_a$ means first a , then a' . To move this operation to the right side requires reversing the order of multiplication.

The *opposite ring* A° of a ring A is defined to be the ring whose underlying set is in bijective correspondence with A , but whose multiplication is reversed: If $x \in A$, we'll write x° for the corresponding element of A° . The multiplication rule in A° is $x^\circ y^\circ = (yx)^\circ$.

Lemma III.1.2. (i) *Let A be a ring. A right A -module M is also a left A° -module by the rule $a^\circ m := ma$. Similarly, a left A -module is a right A° -module.*

(ii) *Let A, B be R -algebras. A bimodule ${}_A M_B$ is the same thing as a right module over the tensor product algebra $A^\circ \otimes B$. The two scalar actions are related by the formula $m(a^\circ \otimes b) = amb$. \square*

A special case is that of an (A, A) -bimodule ${}_A M_A$. Such a bimodule is a right module over the ring $E = A^\circ \otimes A$. For instance, A is a right module over E . The ring E is called the *enveloping algebra* of the algebra A .

Let S be a subalgebra of an R -algebra A . The *centralizer* of S in A is defined to be the set S' of elements of A which commute with all elements of S :

$$S' = \{a \in A \mid as = sa \text{ for all } s \in S\}.$$

The *center* $Z(A)$ of A is the set of elements which commute with every element of A . We could also say, pedantically, that $Z(A)$ is the centralizer of A in A .

Proposition III.1.3. *Let A, B be algebras over a field K .*

(i) *The center $Z(A \otimes B)$ is the tensor product of the centers $Z(A) \otimes Z(B)$.*

(ii) *Let $S \subset A$ be a subalgebra, with centralizer S in A . The centralizer of $S \approx S \otimes R$ in $A \otimes B$ is $S \otimes B$.*

Proof. (i) Let $x \in A \otimes B$. We can write

$$(III.1.4) \quad x = \sum_1^n a_i \otimes b_i,$$

where b_i are K -independent elements of B (i.e., are linearly independent over K). In this expression, the elements a_i are uniquely determined in terms of x and the K -independent set $\{b_i\}$.

An element $x \in A \otimes B$ is in the center if and only if it commutes with all elements of the forms $\alpha \otimes 1$ and $1 \otimes \beta$. Suppose x is expressed as in (III.1.4). Then $[\alpha \otimes 1, x] = \alpha x - x\alpha = \sum[\alpha, a_i] \otimes b_i$. Because the expression (III.1.4) is uniquely determined by the set $\{b_i\}$, $[\alpha, x] = 0$ if and only if $[\alpha, a_i] = 0$ for all i . If x is contained in the center, then this is true for all $\alpha \in A$, hence $a_i \in Z(A)$ for all i . This shows that $Z(A \otimes B)$ is in the center of $Z(A) \otimes B$. Similarly, $Z(Z(A) \otimes B) \subset Z(A) \otimes Z(B)$. Combining these two inclusions, $Z(A \otimes B) \subset Z(A) \otimes Z(B)$. The opposite inclusion is clear.

The proof of (ii) is analogous. \square

Corollary III.1.5. (i) *If R is the center of an algebra A and if S is a commutative R -algebra, then $Z(A \otimes S) = S$.*

(ii) *If $Z(A) = R$ and $Z(B) = R$, then $Z(A \otimes B) = R$.* \square

Proposition III.1.6. *Let A be a simple algebra whose center is a field K , and let B be any K -algebra. Every ideal of $A \otimes B$ has the form $A \otimes J$, where J is an ideal of B . In fact, taking into account the canonical injection $B \rightarrow A \otimes B$, we can identify J as $I \cap B$.*

Lemma III.1.7. *Let I be an ideal of $A \otimes B$, where A, B are as in Proposition III.1.6. Let b_1, \dots, b_n be K -independent in B , and let $x = \sum a_i \otimes b_i \in I$ with $a_1 \neq 0$. There is an element $x = \sum a_i \otimes b_i$ in I such that $a_1 = 1$.*

Proof. Because A is a simple ring, 1 is in the ideal generated by a_1 , i.e., $1 = \sum u_\nu a_1 v_\nu$, with $u_\nu, v_\nu \in A$. Then $x = \sum u_\nu x v_\nu = 1 \otimes b_1 + \sum_2^n a_i \otimes b_i$ is in I , where $a_i = \sum u_\nu a_i v_\nu$. \square

Proof of the proposition. Let I be an ideal of $A \otimes B$, and let I_0 denote the ideal of $A \otimes B$ generated by $I \cap B$. Let $x \in I$. To show that $x \in I_0$, we write $x = \sum_1^n a_i \otimes b_i$, where $\{b_i\}$ are K -independent, and we use induction on n . We may assume that a_1 is not zero. Let x be as in the lemma. If $a_i \in K$ for all i , then $x \in I_0$. In that case it suffices to show that $x - a_1 x$ is in I_0 , which is true by induction because the terms $a_1 \otimes b_1$ cancel. Else if, say $a_2 \notin K$, there is an element $w \in A$ such that $[w, a_2] = wa_2 - a_2 w \neq 0$. Then $y = [w, x] = \sum_2^n c_i \otimes b_i$ is in I , with $c_i = [w, a_i]$. The element y has fewer than n terms, and $c_2 \neq 0$. Lemma III.1.7 shows that there is an element $y = \sum_2^n c_i \otimes b_i$ in I with $c_2 = 1$. By induction, $y \in I_0$. Then $x - a_2 y$ is in I_0 by induction too. \square

Corollary III.1.8. *If A, B are simple algebras over a field K and if $Z(A) = K$, then $A \otimes B$ is simple.* \square

III.2. Central simple algebras

Let's agree that when we say informally that a ring is a matrix algebra over a ring R , we mean that it is isomorphic to a matrix algebra over R .

Let K be a field. A K -algebra is called *central simple* if it is a finite algebra over K , is a simple ring, and if $Z(A) = K$. By Wedderburn's theorem, a central simple K algebra A is a matrix algebra over a division ring D , where D is finite over K . It is easily seen that $Z(D) = K$ as well, so D will also be a central simple K -algebra.

Lemma III.2.1. *A central simple K -algebra A is semisimple, and all simple right A -modules are isomorphic. If $A \approx M_n(D)$, then D^n is a simple A -module. \square*

Lemma III.2.2. *Let D be a division ring which is a finite algebra over a field K . Any element $x \in D$ generates a commutative subfield $L = K[x]$.*

Proof. $K[x]$ is commutative because it is a quotient of a polynomial ring. It is a field because it is a domain of finite dimension over K . \square

Proposition III.2.3. *Let A be a K -algebra.*

(i) *Let L be a field extension of K . Then $A_L = A \otimes L$ is a central simple L -algebra if and only if A is a central simple K -algebra.*

(ii) *Let \overline{K} be the algebraic closure of K . If A is a central simple K -algebra, then $A_{\overline{K}} = A \otimes \overline{K}$ is a matrix algebra over \overline{K} .*

Proof. (i) Corollaries III.1.5 and III.1.8.

(ii) Since $A_{\overline{K}}$ is a central simple \overline{K} -algebra, it is a matrix algebra over a division ring D finite over \overline{K} . Lemma III.2.2 shows that the only such division ring is \overline{K} itself. \square

Example III.2.4. Cyclic algebras. Let n be an integer not divisible by the characteristic of K , and suppose that K contains a primitive n -th root of unity ζ . Let $a, b \in K$. The K -algebra A generated by two elements x, y , with the relations

$$x^n = a, \quad y^n = b, \quad yx = \zeta xy$$

is central simple, and of rank n^2 over K . To show this, Proposition III.2.3(i) allows us to replace K by a field extension. So we may assume that b has an n -th root β in K . In that case, A is isomorphic to the matrix algebra $M_n(K)$. The isomorphism is illustrated here for $n = 3$:

$$(III.2.5) \quad x = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ a & 0 & 0 \end{pmatrix}, \quad y = \beta \begin{pmatrix} 1 & 0 & 0 \\ 0 & \zeta & 0 \\ 0 & 0 & \zeta^2 \end{pmatrix}.$$

If the n -th roots of unity are not in K , then deciding whether or not this algebra is a matrix algebra can be tricky.

Exercise: Let a, b be variables, and let K be the field of rational functions $k(a, b)$. Prove that the cyclic algebra (III.2.4) is a division ring.

We will use the notation

$$(III.2.6) \quad [A : K] = \dim_K A.$$

Proposition III.2.7. (i) *The K -dimension of a central simple K -algebra is a square. In particular, the dimension of a division ring with center K and finite over K is a square.*

(ii) *Suppose that $\dim_K A = n^2$. The dimension of a nonzero right A -module is at least n , and A is a matrix algebra over K if and only if there is a module of K -dimension n .*

Proof. (i) $\dim_K A = \dim_{\overline{K}} A_{\overline{K}}$.

(ii) If $A \approx M_r(D)$ and $\dim_K D = d^2$, then $n^2 = d^2 r^2$. The simple right A -modules are isomorphic to D^r , which has dimension $d^2 r$. This number is larger than n unless $d = 1$. \square

Proposition III.2.8. *Let A, B be central simple K -algebras.*

(i) *$A \otimes B$ is central simple.*

(ii) *$E = A^o \otimes A$ is a matrix algebra over K .*

Proof. (i) Proposition III.1.3 and Corollary III.1.8.

(ii) E is a central simple algebra, and it has K -dimension n^4 . Moreover, A is a right E -module, and $\dim_K A = n^2$. Proposition III.2.7 shows that E is a matrix algebra over K . \square

Lemma III.2.9. *Let A, B be central simple algebras over K . The following are equivalent:*

(a) *A and B are matrix algebras over the same division ring D .*

(b) *There are integers r, s such that the matrix algebras $M_r(A)$ and $M_s(B)$ are isomorphic.*

Proof. A matrix algebra $A = M_n(D)$ over a division ring D determines D up to isomorphism because one can identify D as the ring $\text{End } V_A$, where V is a simple module. The assertion follows from this and from the fact that $M_r(M_s(A)) \approx M_{rs}(A)$. \square

We are now in position to define the *Brauer group* $\text{Br } K$ of a field K . The elements of $\text{Br } K$, called *Brauer classes*, can be thought of in two equivalent ways:

(a) K -isomorphism classes $[D]$ of division rings with center K and finite over K , or

(b) Equivalence classes $[A]$ of central simple algebras over K , where two algebras A, A' are called equivalent if they are matrix algebras over the same division ring D , or equivalently, if there are integers r, s such that the matrix algebras $M_r(A)$ and $M_s(A')$ are K -isomorphic.

The first of these is perhaps more appealing, but the second is a little easier to work with. With the definition (b), the product $[A][B]$ of two elements of the Brauer group is defined to be the class $[A \otimes B]$ of the tensor product. The fact that this is independent of the choice of the representatives A, B in the Brauer classes follows easily from Lemma III.1.1.

The law of composition is associative and commutative and has a unit element $[K]$ (III.1.1), and (III.2.8) shows that $[A^o] = [A]^{-1}$, hence that $\text{Br } K$ is an abelian group.

Proposition III.2.10. (i) *The Brauer group of an algebraically closed field is the trivial group.*

(ii) *The Brauer group of the field \mathbb{R} of real numbers is cyclic of order 2. A division algebra of finite dimension over its center \mathbb{R} is isomorphic either to \mathbb{R} or to the algebra of quaternions \mathbb{H} .*

Proof. (ii) There are many proofs of this fact. The following one is due to Palais. Let D be a division ring finite over \mathbb{R} and different from \mathbb{R} . Let $x \in D$ be an element not in \mathbb{R} . The field $L := \mathbb{R}[x]$ is commutative because x commutes with \mathbb{R} and with itself. The only possibility is that L is isomorphic to the field of complex numbers. So L contains a square root of -1 , call it i .

Let ϕ denote the operator of conjugation by i on D , i.e., $\phi(y) = iyi^{-1}$. Because i commutes with elements of L , ϕ is an L -linear operator on the right L -module D_L . Then ϕ^2 is conjugation by $i^2 = -1$. Because -1 is in the center \mathbb{R} , ϕ^2 is the identity, and ϕ is a diagonalizable operator with eigenvalues ± 1 . Let D_+ and D_- denote the two eigenspaces of ϕ , so that $D \approx D_+ \oplus D_-$. Note that $D_+ = L$. The reason is that since L is algebraically closed, there is no larger commutative subfield of D . If $y \in D_+$, then y commutes with i , and hence it commutes with every element of L , so $L[y]$ is commutative. Since L is maximal, $L[y] = L$, i.e., $y \in L$.

Next, left multiplication by i defines a bijection $D_+ \longrightarrow D_-$, so the dimensions of D_+ and D_- are equal. It follows that $[D : \mathbb{R}] = 4$.

Let $z \in D_-$ and let $s = z^2$. Then $s \in D_+ = L$. The commutative field $L = \mathbb{R}[z]$ is also isomorphic to \mathbb{C} , and for dimension reasons, $L \cap L = \mathbb{R}$. Therefore $s \in \mathbb{R}$ but its square root is not real. So $s < 0$. We normalize z so that $s = -1$, and we rename $z = j$. Then $ij = -ji$, and $D = \mathbb{R} + \mathbb{R}i + \mathbb{R}j + \mathbb{R}ij$ is the algebra of quaternions \square

The next theorem is one of the main results in the theory of the Brauer group. Unfortunately it would take too long to prove here.

Theorem III.2.11. (*Merkuriev-Suslin*) *Suppose that K is a field of characteristic zero which contains all roots of unity. Then the Brauer group $\text{Br } K$ is generated by the classes of cyclic algebras (III.2.4).*

III.3. Skolem-Noether theorem

Theorem III.3.1. *Let A be a central simple K algebra, and let $\theta : B \longrightarrow B$ an isomorphism between subalgebras of A which are simple rings. Then θ is induced by an inner automorphism of A , i.e., there is an invertible element $u \in A$ such that θ is conjugation by u . In particular, any two isomorphic simple subalgebras of A are conjugate.*

Proof. Let's write b^θ for $\theta(b)$. We can make A into a right B -module in two ways: First, we can simply restrict scalars to B . We'll denote this module by A_B . The second way is to let B act through θ . Let's write this operation as $*$, and denote the module thus obtained by A_B^* . So the action of $b \in B$ on A^* is given by

$$(III.3.2) \quad x * b = xb^\theta.$$

Both structures are compatible with the operation of left multiplication by A , so we have two (A, B) -bimodules, ${}_A A_B$ and ${}_A A_B^*$. We can view them as right modules over the ring $T := A^o \otimes B$. Proposition III.1.3 and Corollary III.1.8 show that T is a central simple L -algebra, where L is the center of B .

Since T is central simple over L , there is one isomorphism class V of simple right T -modules. By semisimplicity, A_T is isomorphic to $\bigoplus_r V$ for some integer r , and A_T^* is isomorphic to $\bigoplus_s V$ for some s . Since $\dim_K A = \dim_K A^*$, $r = s$. So A_T and A_T^* are isomorphic modules. Let $\phi : A \longrightarrow A^*$ denote an isomorphism between them. The T -linearity of ϕ means linearity with respect to left multiplication by A and right multiplication by B . So for $x \in A$, $\phi(ax) = a\phi(x)$,

while $\phi(xb) = \phi(x) * b = \phi(x)b^\theta$. Let $u = \phi(1)$. Then for $b \in B \subset A$, $\phi(1b) = \phi(1) * b = ub^\theta$. But because B is a subring of A , we also have $\phi(b1) = b\phi(1) = bu$. Thus

$$ub^\theta = bu,$$

or $b^\theta = u^{-1}bu$. Since this is true for all $b \in B$, θ is conjugation by u , as required. \square

Exercise: Fill the gap in the above proof by showing that u is an invertible element of A .

Theorem III.3.3. (Skolem-Noether) *Every automorphism of a central simple K -algebra A is inner.*

Proof. This follows by setting $B = B = A$ in Theorem III.3.1. \square

III.4. Commutative subfields of central simple algebras

Let A be a central simple K -algebra.

Lemma III.4.1. (i) *Let A be any K -algebra, and let $x \in A$. The subalgebra $K[x]$ of A generated by x is commutative.*

(ii) *Let L be a commutative subalgebra of a K -algebra A , and let L denote its centralizer in A . Then $L \subset L$. If $x \in L$, then $L[x]$ is a commutative subring of A . Hence L is a maximal commutative subalgebra if and only if $L = L$. \square*

Theorem III.4.2. *Let A be a central simple algebra of rank n^2 over K , and let L be a commutative subalgebra of A which is a field. Then $[L : K] \leq n$. Moreover, L is a maximal commutative subalgebra of A if and only if $[L : K] = n$.*

Proof. Say that $[L : K] = r$. The operation of (left) multiplication on itself embeds L as a subalgebra of the matrix algebra $B = \text{End } L_K \approx M_r(K)$. Let $\tilde{L} \subset B$ denote its image, and let $T = A \otimes B$. Then L embeds in two ways as a subalgebra of T , namely using the embeddings $L \subset A$ and $L \approx \tilde{L} \subset B$. Let's denote the images of these two embeddings by L_1, L_2 respectively.

The algebra T is central simple. Because L is a field, it is a simple ring. Theorem III.3.1 shows that L_1 and L_2 are conjugate.

We have four centralizers: the centralizer L of L in A , the centralizer \tilde{L} of \tilde{L} in B , and the two centralizers L_i of L_i in T .

Lemma III.4.3. *With the above notation, $\tilde{L} = \tilde{L}$.*

Proof. First, because \tilde{L} is commutative, $\tilde{L} \subset \tilde{L}$. Second, \tilde{L} is a (\tilde{L}, \tilde{L}) -bimodule, hence $\tilde{L} \subset \text{End}_{\tilde{L}} \tilde{L} = \tilde{L}$. To check that \tilde{L} is a bimodule as claimed, let $b \in \tilde{L}$ and $a \in \tilde{L}$. Then for $x \in \tilde{L}$, $(ax)b = (xa)b = x(ab) = x(ba) = (xb)a = a(xb)$. \square

Proposition III.1.3(ii) tells us that $L_1 = L \otimes B$ and $L_2 = A \otimes \tilde{L} = A \otimes \tilde{L}$. So $[L_1 : K] = [L : K]r^2$ and $[L_2 : K] = n^2r$. Because L_i are conjugate, so are their centralizers. Therefore $[L : K]r^2 = n^2r$. Because $L \subset L$, this implies that $r \leq n$, and that if $r < n$, then $L < L$. The theorem follows from III.4.1(iii). \square

Theorem III.4.4. (Wedderburn) *A finite division ring is a field, i.e., it is commutative.*

Proof. Let D be a finite division ring with center K . Say that $[D : K] = n^2$ and that $|K| = q$. Every commutative subalgebra of D is a field, and according to Theorem III.4.2, every maximal commutative subalgebra has order q^n . Let L be one of these maximal subfields. Since all fields of order q^n are isomorphic, Theorem III.3.1 applies. It shows that the maximal commutative subfields are all conjugate to L . Because every element x of D must be contained in at least one maximal commutative subfield, the conjugates of L cover D completely. An elementary fact about finite groups allows us to conclude that $L = D$, hence that D is commutative.

Lemma III.4.5. *If H is a proper subgroup of a finite group G , the conjugates of H do not cover G .*

Proof. The number of conjugate subgroups is the index $[G : N]$, where N is the normalizer of H (the stabilizer of H for the operation of conjugation). Moreover, $H \subset N$. Therefore the number of conjugate subgroups is at most equal to the index $[G : H]$. This index is also equal to the number of left cosets of H in G . The cosets cover G without overlap, while the conjugates of H do overlap, because they contain 1. So the conjugates can not cover the group. \square

To apply this lemma to our situation, we let G and H be the groups obtained by deleting 0 from D and L respectively. \square

For example, this theorem shows that if K is a finite field which contains a primitive n -th root of unity ζ , where p does not divide n , then the cyclic algebra III.2.4 is a matrix algebra over K . This is not obvious at all.

III.5. Faithful flatness

Let R be a ring. A left R -module N is called *flat* if the functor $-\otimes_R N$ is exact, i.e., if for any exact sequence $M_1 \longrightarrow M_2 \longrightarrow M_3$ of right R -modules, the sequence $M_1 \otimes N \longrightarrow M_2 \otimes N \longrightarrow M_3 \otimes N$ is also exact. The analogous definitions are made for left modules and left flat ring extensions S .

The functor $-\otimes_R N$ is always right exact. It is exact if and only if $M_1 \otimes_R N \longrightarrow M_2 \otimes_R N$ is injective whenever $M_1 \longrightarrow M_2$ is injective.

A left module N is called *faithfully flat* if it is flat and if $M \otimes N = 0$ implies that $M = 0$.

Lemma III.5.1. *A flat left R -module N is faithfully flat if and only if the following condition is satisfied: A sequence $M_1 \longrightarrow M_2 \longrightarrow M_3$ of right R -modules is exact if and only if the sequence $M_1 \otimes N \longrightarrow M_2 \otimes N \longrightarrow M_3 \otimes N$ is exact. \square*

If R is commutative and M is a right R -module, then it can also be viewed as a left module, and M_R is flat if and only if ${}_R M$ is flat. This applies in particular to the case of a homomorphism of commutative rings $R \longrightarrow S$: S_R is flat if and only if ${}_R S$ is flat.

We omit the proof of the next proposition.

Proposition III.5.2. *Let $R \longrightarrow S$ be a homomorphism of commutative rings such that S is flat over R . Then S is faithfully flat if and only if the map $\text{Spec } S \longrightarrow \text{Spec } R$ is surjective \square .*

Exercises: (1) If R is a field, then any nonzero module M_R is right faithfully flat.

- (2) Let Σ be a right Ore set in R and $S = R\Sigma^{-1}$. Then ${}_R S$ is flat.
- (3) Let R be a commutative ring and let $s_1, \dots, s_n \in R$. Let S be the direct sum of the localizations $R_i = R[s_i^{-1}]$. then S is faithfully flat over R if and only if the ideal $(s_1, \dots, s_n)R$ is the unit ideal.
- (4) Suppose that ${}_R N$ is faithfully flat. If $\phi : M \rightarrow M$ is a homomorphism of R -modules such that $M \otimes N \rightarrow M \otimes N$ is an isomorphism, then ϕ is an isomorphism.

III.6. The Amitsur complex

The *Amitsur complex* $\mathcal{A}(S/R)$ is a cosimplicial complex associated to an arbitrary ring homomorphism $\theta : R \rightarrow S$. A cosimplicial complex \mathcal{A} consists of a set \mathcal{A}_n for $n = 0, 1, 2, \dots$, the set of “cosimplices” of dimension n , together with certain maps between them called face and degeneracy maps. (The word “cosimplicial” means that the arrows go in the opposite direction from those in a simplicial complex.)

We now define the complex \mathcal{A} : In dimension n , \mathcal{A}_n is the (S, S) -bimodule $S \otimes \cdots \otimes S = S^{\otimes n+1}$. The face maps are maps $d^i : \mathcal{A}_n \rightarrow \mathcal{A}_{n+1}$ for $i = 0, \dots, n+1$. They are defined by inserting a 1 in the i th position of a tensor:

$$(III.6.1) \quad x_0 \otimes \cdots \otimes x_n \mapsto x_0 \otimes \cdots \otimes x_i \otimes 1 \otimes x_{i+1} \otimes \cdots \otimes x_n.$$

The degeneracies s^i , defined for $i = 0, \dots, n-1$, are maps $\mathcal{A}_n \rightarrow \mathcal{A}_{n-1}$ which multiply the i th and $(i+1)$ th entries in a tensor:

$$(III.6.2) \quad x_0 \otimes \cdots \otimes x_n \mapsto x_0 \otimes \cdots \otimes x_i x_{i+1} \otimes \cdots \otimes x_n.$$

It is customary to suppress the index n in these operators. So there is a map d^i for each $n \geq i-1$. These maps satisfy certain *standard identities* which tell us when compositions of two face or degeneracy operators are equal, and when such a composition is the identity. No one can remember for more than a day what these identities are, and exactly how they come out depends on whether the operations are written on the left or on the right. However, they are “trivial”. In order to put composition of functions in the natural order, we will write operators on the right. So, for example, we write $(x_0 \otimes x_1)d^2s^0 = (x_0 \otimes x_1 \otimes 1)s^0 = x_0x_1 \otimes 1$. Some of the standard identities are:

$$d^2s^0 = s^1d^1, \quad d^1s^0 = id, \quad d^1d^2 = d^1d^1, \quad s^2s^0 = s^0s^1.$$

Exercise: Write down the standard identities which generalize the above formulas to other indices.

The face and degeneracy maps are S -bimodule homomorphisms, except for d^0 and $d^n : \mathcal{A}_n \rightarrow \mathcal{A}_{n+1}$. On the left, d^0 is only R -linear, and similarly, d^n is only R -linear on the right.

If R and S are commutative rings, then \mathcal{A}_n are rings and the face and degeneracy maps are ring homomorphisms.

The map $\theta : R \rightarrow S$ is called an *augmentation* of the Amitsur complex. The composed maps are equal:

$$(III.6.3) \quad \theta d^0 = \theta d^1.$$

We will use the shorthand notation $S \otimes S = SS$ and $S \otimes S \otimes S = SSS$. Similarly, if M is a left R -module, we may write $S \otimes M = SM$, etc. With this notation, the face operators in the augmented Amitsur complex look like this:

$$(III.6.4) \quad R \rightarrow S \rightrightarrows SS \rightrightarrows SSS \cdots .$$

This cosimplicial complex yields a *complex* of R -modules

$$(III.6.5) \quad 0 \longrightarrow R \xrightarrow{\theta} S \xrightarrow{\delta^0} SS \xrightarrow{\delta^1} SSS \cdots ,$$

where $\delta^n = d^0 - d^1 + \cdots \pm d^{n+1}$. We'll denote this complex by \mathcal{A} too.

Exercise: Prove that this sequence is a complex using the standard identities.

Theorem III.6.6. (Grothendieck) *Let $\theta : R \rightarrow S$ be ring homomorphism such that S_R is faithfully flat. Then the Amitsur complex \mathcal{A} is a resolution of R , i.e., III.6.5 is exact. Moreover, this complex remains exact when tensored on the right by an arbitrary left R -module M :*

$$0 \longrightarrow M \longrightarrow SM \longrightarrow SSM \longrightarrow \cdots .$$

The analogous theorem is true when ${}_R S$ is faithfully flat: If M is a right R -module, then

$$0 \longrightarrow M \longrightarrow MS \longrightarrow MSS \longrightarrow \cdots$$

is exact.

Proof of the theorem. Grothendieck's trick is to note that III.6.5 is a sequence of left R -modules. To prove it exact, it suffices to show that the sequence obtained by tensoring on the left with S_R is exact. When we tensor III.6.5 with S_R , we obtain the Amitsur complex again, except that the face d^0 and degeneracy s^0 are missing, $i \otimes \theta = d^1$, and all indices in the remaining face maps are increased by 1. Let's denote the maps in the complex $S \otimes \mathcal{A}$ by $\bar{\delta}$:

$$\bar{\delta}^n := i \otimes \delta^{n-1} = d^1 - d^2 + \cdots \pm d^{n+1} : S \otimes S^{\otimes n} \longrightarrow S \otimes S^{\otimes n+1}.$$

This new complex is homotopically trivial, the homotopy being given by the missing degeneracy: $h = s^0$, applied in each degree n . In other words, $h\bar{\delta} + \bar{\delta}h = \text{identity}$. The existence of such a homotopy shows that the sequence is exact. Namely, if $x \in \ker \bar{\delta}$, i.e., $x\bar{\delta} = 0$, then

$$x = x(h\bar{\delta} + \bar{\delta}h) = xh\bar{\delta},$$

so x is in the image of $\bar{\delta}$.

The fact that $h = s^0$ is a homotopy is checked directly:

$$\begin{aligned} (y \otimes x_0 \otimes \cdots \otimes x_n) s^0 \bar{\delta} &= (yx_0 \otimes x_1 \otimes \cdots \otimes x_n) (d^1 - d^2 + \cdots) \\ &= yx_0 \otimes 1 \otimes x_1 \otimes \cdots - yx_0 \otimes x_1 \otimes 1 \otimes x_2 \otimes \cdots + \cdots , \end{aligned}$$

while

$$\begin{aligned} (y \otimes x_0 \otimes \cdots \otimes x_n) \bar{\delta} s^0 &= (y \otimes 1 \otimes x_0 \otimes x_1 \otimes \cdots - y \otimes x_0 \otimes 1 \otimes x_1 \otimes \cdots + \cdots) s^0 \\ &= y \otimes x_0 \otimes x_1 \otimes \cdots - yx_0 \otimes 1 \otimes x_1 \otimes \cdots + \cdots . \end{aligned}$$

This shows that III.6.5 is exact, and if we tensor on the right by an arbitrary left R -module M , the same proof shows that the resulting sequence is exact.

The analogous proof works for a left faithfully flat homomorphism, except for a slight notation complication coming from the fact that the "last" face and degeneracy operators have varying indices. \square

Corollary III.6.7. (*descent for elements of a module*) Assume that ${}_R S$ is faithfully flat. Let M be a right R -module, let $M = M \otimes S$, and let $M = M \otimes S \otimes S$. Then d^0 and d^1 are maps $M \rightrightarrows M$. An element $x \in M$ has the form $x = 1 \otimes y$ for some unique $y \in M$ if and only if $xd^0 = xd^1$.

This corollary restates the exactness of the augmented Amitsur complex at the first two terms. It is usually stated this way:

$$(III.6.8) \quad \text{An element } x \text{ of } N \text{ lies in } M \text{ if and only if } xd^0 = xd^1$$

To show that this is not an automatic requirement, we'll write the condition out in terms of tensors. Say that $x = \sum_{\nu} a_{\nu} \otimes m_{\nu} \in SM$. Then $xd^0 = xd^1$ reads

$$\sum_{\nu} 1 \otimes a_{\nu} \otimes m_{\nu} = \sum_{\nu} a_{\nu} \otimes 1 \otimes m_{\nu}.$$

Because of this descent principle, Grothendieck was led to define exactness for a diagram

$$(III.6.9) \quad X \longrightarrow Y \rightrightarrows Z$$

of arbitrary sets. The *kernel* of a pair of maps of sets $Y \rightrightarrows Z$ is the set of elements $y \in Y$ whose images under the two maps are equal. The diagram (III.6.9) is *exact* if the arrow $X \longrightarrow Y$ maps X bijectively onto the kernel of the pair of maps $Y \rightrightarrows Z$. Thus the first three terms of the Amitsur complex III.6.4 form an exact sequence, provided that ${}_R S$ or S_R is faithfully flat.

There are many applications of the descent principle to situations in which an element x is uniquely determined by certain properties. If x is uniquely determined, and if we can apply this fact to its images xd^0 and xd^1 in $M \otimes S \otimes S$, it will (hopefully) show that these images are equal. When S is commutative, $S \otimes S$ is also a commutative ring, and then $M \otimes S \otimes S$ is the module obtained from M by extension of scalars $R \longrightarrow S \otimes S$. We have a good chance of interpreting the properties in that case. The principle is harder to apply when S is not commutative, because then $S \otimes S$ is only a bimodule. I don't know a simple interpretation of $M \otimes S \otimes S$ in the general case.

Exercise: Let S be a commutative, faithfully flat R -algebra, and let A be any R -algebra, not necessarily associative or with unit element. Prove that A is associative, or commutative, or has an identity element if and only if the same is true for $A \otimes S$. Prove that if A has an identity element, then an element a invertible in A if and only if $a \otimes 1$ is invertible in $A \otimes S$.

III.7. Interlude: Analogy with bundles

Let A be a central simple algebra over the field K , and let \overline{K} be the algebraic closure of K . We want to explain why the fact that $\overline{A} = A \otimes \overline{K}$ is a matrix algebra is analogous to the concept of a *bundle* in topology.

Let V denote the vector space \mathbb{C}^n . Roughly speaking, a complex *vector bundle* of dimension n over a space X is a map $\pi : E \longrightarrow X$ of topological spaces, all of whose fibres $\pi^{-1}(x)$ are vector spaces isomorphic to V . There are some topological conditions which we will suppress. The *trivial bundle* over X is $E = X \times V$, and we call any bundle isomorphic to the trivial bundle trivial too. Every vector bundle is *locally trivial*, which means that there is an open covering $\{U_i\}$ of X such that

the restriction of E to each U_i is the trivial bundle $U_i \times V$. If we let Y denote the disjoint union of the U_i , then we have a surjective map $Y \rightarrow X$ such that the pullback of E to Y is trivial.

In algebra, the arrows are reversed by imagining a commutative ring as a ring of functions on the space $\text{Spec } R$. The analogue of vector bundle over a commutative ring R is a projective R -module E of constant rank n , and the analogue of the trivial bundle is the free bundle of rank n . It is a fact that every such projective module is locally free: There are elements s_1, \dots, s_n in R which generate the unit ideal, such that the localization E_{s_i} of E is a free module over $R_{s_i} = R[s_i^{-1}]$. Passing to spectra, $U_i = \text{Spec } R_{s_i}$ is an open set in $X = \text{Spec } R$, and because the s_i generate the unit ideal, the open sets cover X . The analogy with the topological notion is perfect.

However, the concept of bundle in topology also allows the fibres $\pi^{-1}(x)$ to have a structure other than that of a vector space. One may consider bundles of matrix algebras: all fibres are algebras over \mathbb{C} which are isomorphic to $A = M_n(\mathbb{C})$. With the correct topological hypotheses, any such bundle will be locally isomorphic to $X \times M_n(\mathbb{C})$, i.e., will be locally trivial.

Here the algebraic analogue is more complicated, because for any central simple algebra A over a field K , $A \otimes_K \overline{K}$ is a matrix algebra. This shows that one should broaden the concept of local triviality in algebra, and include not only localizations (the adjunction of inverses), but also field extensions such as $K \rightarrow \overline{K}$. The Amitsur complex provides a formalism which allows one to recover information from such a generalized notion of localization.

III.8. Characteristic polynomial for central simple algebras

The characteristic polynomial is one of the main invariants of a matrix. We now use descent to show that the characteristic polynomial is defined for elements of an arbitrary central simple algebra A , hence, in particular, that the *trace* and *determinant* of any element $a \in A$ are defined, and that they are elements of K . Usually $\text{trace}(a)$ is referred to as the *reduced trace*, and $\det(a)$ the *reduced norm* of a . I prefer the words trace and determinant, though they are somewhat ambiguous.

In order to focus attention, let's concentrate on the determinant. The same reasoning will work for all of the coefficients of the characteristic polynomial.

Let A be a central simple K -algebra. So $A \otimes \overline{K}$ is a matrix algebra which, as we know, means that it has a set of matrix units $\{e_{ij}\}$. To find these matrix units, it is not necessary to go all the way to the algebraic closure. They will already be in $A \otimes L$ for some subfield L of \overline{K} which is a finite extension of K . So we consider a finite extension L such that $A \otimes L$ is a matrix algebra $M_n(L)$. Of course L is a faithfully flat extension of K . Let's drop the tensor symbol, writing A_L for $A \otimes L$, LL for $L \otimes L$, etc.

Lemma III.8.1. (i) Let R be a commutative ring, and let $A \xrightarrow{\phi} M_n(R)$ be an isomorphism of A with a matrix algebra. Then $\det(a\phi)$ is independent of ϕ . Hence there is a uniquely defined map $\det : A \rightarrow R$.

(ii) Let $R \xrightarrow{f} S$ be a homomorphism of commutative rings. Then $M_n(R) \otimes_R S \approx M_n(S)$.

(iii) Let $R \xrightarrow{f} S$ be a homomorphism of commutative rings, and let us denote the induced homomorphism $M_n(R) \rightarrow M_n(S)$ by f too. Then $\det(f(a)) = f(\det(a))$ for all matrices $a \in M_n(R)$. Hence \det is invariant under extension of scalars. \square

(Sorry: This lemma is out of order. Logically, It should come after Theorem 11.2.)

Now let A be a central simple K -algebra. We choose a finite field extension L such that A_L is isomorphic to a matrix algebra over L , and we define $\det(a)$ for an element $a \in A$ as

$$(III.8.2) \quad \det(a) = \det(a \otimes 1),$$

where $a \otimes 1$ is the image of a in $A_L = A \otimes L$. Lemma III.8.1(iii) shows that this is independent of the choice of L .

Proposition III.8.3. *The function \det on A takes its values in K . Hence there is a uniquely defined map $\det : A \rightarrow K$.*

Proof. We have an exact diagram

$$K \longrightarrow L \rightrightarrows LL,$$

and $x = \det(a) \in L$. To show that x is in K , it suffices to show that its two images xd^0 and xd^1 in LL are equal. We note that A_{LL} is obtained from A_L by extension of scalars using either of the two maps $L \xrightarrow{d^i} LL$, namely $A_{LL} \approx A \otimes_K L \otimes_L LL$. Since A_L is a matrix algebra over K , A_{LL} is a matrix algebra over LL . Therefore $\det : A_{LL} \rightarrow LL$ is defined uniquely. Then since $(a \otimes 1)d^i = a \otimes 1 \otimes 1$ is true for $i = 0, 1$, III.8.1(iii) shows that

$$xd^i = \det((a \otimes 1)d^i) = \det(a \otimes 1 \otimes 1),$$

i.e., $xd^0 = xd^1$ as required. \square

We now make a slight improvement in the above result, by showing that the determinant on a central simple algebra can be obtained by evaluating a polynomial with coefficients in K .

Let's choose a basis for the central simple algebra A as K -vector space, say $\{u_1, \dots, u_{n^2}\}$, where $n^2 = [A : K]$. Also, let e_{ij} denote the matrix units in $A_{\overline{K}} = M_n(\overline{K})$. Of course $\{u_\alpha\}$ is also a basis for $A_{\overline{K}}$, so the two bases are related by a linear change of variable, say

$$(III.8.4) \quad u_\alpha = \sum p_{\alpha ij} e_{ij},$$

with $p_{\alpha ij} \in \overline{K}$.

Working in \overline{K} for the moment, we write an undetermined element a of $A_{\overline{K}}$ as a matrix with variable entries y_{ij} , i.e., as a linear combination $a = \sum y_{ij} e_{ij}$. Then of course $\det(a)$ is the standard homogeneous polynomial of degree n

$$(III.8.5) \quad \sum (-1)^\sigma y_{\sigma(1)1} \cdots y_{\sigma(n)n}.$$

If we also write a as linear combination $a = \sum x_\alpha u_\alpha$, then the coefficients x_α are obtained from the y_{ij} by a linear change of variable using the formula III.8.4. So there is also a homogeneous polynomial Δ of degree n with coefficients in \overline{K} , such that

$$(III.8.6) \quad \det(a) = \Delta(x_1, \dots, x_{n^2}).$$

Proposition III.8.7. *The polynomial $\Delta(x)$ has coefficients in K . Thus $\det(a)$ can be obtained by evaluating a homogeneous polynomial with coefficients in K , and which, over \overline{K} , is obtained from the standard polynomial for $\det(a)$ by a linear change of variable.*

Proof. It is clear that this is true if A is a matrix algebra over K , because then the whole computation can be made over K . In view of this, Wedderburn's Theorem III.4.4 takes care of the case that K is a finite field, and we may assume that K is infinite. In that case, the result follows from Lemma III.8.9 below.

Let x_1, \dots, x_r be variables. Every polynomial $f(x_1, \dots, x_r) \in K[x_1, \dots, x_r]$ defines a function $K^r \rightarrow K$ by "evaluation". If we denote the ring of all functions $K^r \rightarrow K$ by \mathcal{F} , then evaluation yields a homomorphism $K[x_1, \dots, x_r] \rightarrow \mathcal{F}$.

Lemma III.8.8. *If K is an infinite field, the homomorphism $K[x_1, \dots, x_r] \rightarrow \mathcal{F}$ is injective.*

Lemma III.8.9. *Let K be an infinite field, and let $f(x_1, \dots, x_r)$ be a polynomial with coefficients in a field extension L of K . If $f(a_1, \dots, a_r) \in K$ for all a_1, \dots, a_r in K , then the coefficients of f lie in K , i.e., $f \in K[x_1, \dots, x_r]$.*

Exercise: Prove Lemmas III.8.8 and III.8.9.

III.9. Separable splitting fields

A *splitting field* for a central simple algebra A over K is a field extension L such that A_L is a matrix algebra.

Theorem III.9.1. *Every central simple algebra A has a splitting field L which is a finite separable extension of K .*

Instead of basing a proof of this fact on special properties of central simple algebras, we will deduce it using the fact that the determinant is obtained by evaluating a polynomial. Wedderburn's theorem allows us to assume that K is infinite.

Let's review the concept of separability. A polynomial $f(y) \in K[y]$ is *separable* if it has distinct roots in the algebraic closure \overline{K} .

Proposition III.9.2. *Let L be a finite ring extension of K .*

(i) *The following two conditions are equivalent, and L is called a separable extension of K they hold:*

(a) *Every element of L is the root of a separable polynomial with coefficients in K , or*

(b) *$L \otimes_K \overline{K}$ is a direct sum of finitely many copies of \overline{K} .*

(ii) *Let $f(y) \in K[y]$ be a polynomial in one variable. Assume that f and its derivative f' are relatively prime. Then $L = K[y]/(f)$ is a separable extension of K .*

(iii) *A separable ring extension of K is a direct sum of finitely many separable field extensions. \square*

Proof of Theorem III.9.1. We use induction on $[A : K]$. We know that A is isomorphic to a matrix algebra over a division ring D , which is also central simple over K , so to split A it suffices to split D . This lowers the degree $[A : K]$ unless A is itself a division ring. So we are done by induction unless A is a division ring.

If A is a division ring, then the determinant of every nonzero element $a \in A$ is nonzero, because a is invertible in A and in the matrix algebra $A_{\overline{K}}$. We will look for a finite separable extension L of K such that A_L contains an element α with determinant zero. Then A_L is not a division ring, and the induction hypothesis will complete the proof, because a composition of finite separable extensions is separable.

Let $\Delta(x)$ be the polynomial (III.8.6) which computes the determinant. A nonzero element of A_L with zero determinant corresponds to a solution of the polynomial equation $\Delta(x) = 0$ for x_1, \dots, x_{n^2} in L , with x_α not all zero. We compute the partial derivative of the standard polynomial (III.8.5) with respect to the variable y_{11} , to check that it is not identically zero. Therefore the partial derivatives of Δ are not all identically zero. The next Proposition completes the proof. \square

Proposition III.9.3. *Let K be a field, and let $f(x)$ and $g(x)$ be polynomials in $K[x_1, \dots, x_n]$. Assume that f and g have no common factor, and that the partial derivatives $\partial f/\partial x_i$ are not all identically zero. There exists a finite separable field extension L of K and elements a_1, \dots, a_n in L such that $f(a) = 0$ but $g(a) \neq 0$.*

Proof. If K is finite, then every finite field extension is separable. The theorem is easy to prove in that case, so we assume that K is infinite.

We may assume that f is irreducible, and that it does not divide g or $\partial f/\partial x_n$. Let's rename x_n to y , and write f for $\partial f/\partial x_n$. Let \mathcal{K} be the fraction field $K(x_1, \dots, x_{n-1})$. In the polynomial ring $\mathcal{K}[y]$, the greatest common divisor of f and g is 1, and also the gcd of f and $\partial f/\partial y$ is 1. Writing these gcd's as linear combinations and clearing denominators yields expressions of the form $p_1 f + q_1 g = u_1$ and $p_2 f + q_2 \partial f/\partial y = u_2$, where p_i, q_i are polynomials in x_1, \dots, x_{n-1}, y and u_i are polynomials in x_1, \dots, x_{n-1} .

We may choose $a_1, \dots, a_{n-1} \in K$ so that $u_i(a) \neq 0$. Then $f(a, y)$ is relatively prime to $g(a, y)$ and to $\partial f/\partial y(a, y)$. It follows that $L = K[y]/(f(a, y))$ is a separable ring extension of K , and that the residue of $g(a, y)$ is invertible in L . \square

III.10. Structure constants:

Let R be a commutative ring, and let A be an R -algebra which is a free R -module, say of rank n . Let u_1, \dots, u_n be an R -basis for A . We don't need to make any assumptions about A here.

The *multiplication table* for A , with respect to this basis, is obtained by writing the products $u_i u_j$ in terms of the basis, say

$$(III.10.1) \quad u_i u_j = \sum c_{ijk} u_k,$$

where the *structure constants* c_{ijk} are elements of R . Every element of A is a linear combination $a = \sum r_i u_i$, with $r_i \in R$, and multiplication in A can be carried out using the table.

If S is a commutative R -algebra and $A_S = A \otimes S$, then the same set u_1, \dots, u_n is an S -basis for A_S , provided that we identify u_i with $u_i \otimes 1$. The multiplication table for A_S is also given by (III.10.1), if we interpret the structure constants as elements of S by taking their images. So every element of A_S is a linear combination with coefficients in S , and multiplication is carried out the

same way. The only change in passing from A to A_S is that the coefficients in a linear combination come from a different ring.

Exercise: What if S is a noncommutative R -algebra?

III.11. Smooth maps and idempotents

Let's work for the moment over the field of complex numbers. Consider a commutative \mathbb{C} -algebra U which is defined by a single polynomial relation: $U = \mathbb{C}[x_1, \dots, x_n]/(f)$. Let \mathcal{X} denote the locus of zeros of $f(x)$ in the \mathbb{C}^n . The locus \mathcal{X} , or the algebra U , is called *smooth* if at every point p of \mathcal{X} , at least one partial derivative $f_{x_i} = \partial f / \partial x_i$ is not zero. Saying the same thing, \mathcal{X} is smooth if the polynomials $f, f_{x_1}, \dots, f_{x_n}$ have no common zeros in \mathbb{C}^n . The Nullstellensatz tells us that this is equivalent to saying that these polynomials generate the unit ideal.

This definition is borrowed from analysis. When \mathcal{X} is smooth, the implicit function theorem can be applied to conclude that \mathcal{X} is a manifold of complex dimension $n - 1$, i.e., is locally homeomorphic to \mathbb{C}^{n-1} . Since the implicit function theorem is not available in polynomial algebra, we state the definition of smoothness in terms of partial derivatives.

By analogy, we define the concept of smoothness for any algebra of the form $U = R[x_1, \dots, x_n]/(f)$, where R is a noetherian commutative ring and f is a polynomial: U is a *smooth R -algebra* if $f, f_{x_1}, \dots, f_{x_n}$ generates the unit ideal.

Example III.11.1: Idempotents in R . The relevant polynomial is $f(x) = x^2 - x$, $f' = df/dx = 2x - 1$. It is easy to see that f and f' generate the unit ideal over the ring of integers, hence over any ring R . In fact, $f'^2 - 4f = 1$.

Going back to the case $R = \mathbb{C}$, suppose that we are given two defining relations for a commutative \mathbb{C} -algebra: $U = \mathbb{C}[x_1, \dots, x_n]/(f_1, f_2)$, and let \mathcal{X} be the locus of solutions of the system of equations $f_1(x) = f_2(x) = 0$ in \mathbb{C}^n . The implicit function theorem again provides a condition under which \mathcal{X} is a manifold of complex dimension $n - 2$. It is that at any point of \mathcal{X} , some 2×2 minor M of the jacobian matrix $J = \partial f_i / \partial x_j$ must be invertible. So the condition becomes that f together with all determinants $\det M$ of the 2×2 -minors M of the jacobian, should generate the unit ideal.

The generalization to more equations seems pretty clear at first glance. However, closer inspection reveals that there are serious problems: In almost every case

- (a) the equations defining the locus are redundant, and to make matters worse,
- (b) we don't know the dimension of \mathcal{X} , so we don't know what size minors of J to inspect.

Example III.11.2: Idempotents in a finite algebra. We'll set this example up with an arbitrary commutative ring of scalars R for future reference, and then look at the case that $R = \mathbb{C}$. Let A be an R -algebra which is a free module with basis u_1, \dots, u_n . We write a variable element of A in the form $a = \sum x_i u_i$, where x_i are indeterminates. The equation for an idempotent, $a^2 = a$, becomes $\sum_{ij} x_i x_j u_i u_j = \sum x_k u_k$, and using the multiplication table (III.10.1), this expands to

$$(III.11.3) \quad \sum_{ijk} c_{ijk} x_i x_j u_k = \sum_k x_k u_k.$$

Since $\{u_k\}$ is a basis, this equation holds if and only if the n equations in n unknowns

$$(III.11.4) \quad \sum_{ij} c_{ijk} x_i x_j - x_k = 0$$

hold, for $k = 1, \dots, n$. We will see below that the locus \mathcal{X} of zeros of this system of polynomial equations is always smooth. But its dimension depends on the structure of the algebra A , and \mathcal{X} may have different dimensions at different points.

Let's go back to the case $R = \mathbb{C}$, and look at the locus when A is the matrix algebra $M_r(\mathbb{C})$. So $n = r^2$. We have several sorts of idempotents, e_{11} and $e_{11} + e_{22}$ for instance. The operation of the general linear group by conjugation moves them around. The centralizer of e_{11} is the set of invertible matrices of the form (illustrated for $r = 3$)

$$p = \begin{pmatrix} * & 0 & 0 \\ 0 & * & * \\ 0 & * & * \end{pmatrix}.$$

The centralizer has codimension $2r - 2$ in GL_r , so this is also the dimension of the orbit of e_{11} . This orbit is one component of the locus \mathcal{X} of zeros of III.11.3. The idempotent $e_{11} + e_{22}$ lies on a different component which has dimension $4r - 8$. For the matrix algebra, \mathcal{X} is a smooth locus consisting of several disconnected pieces of varying dimensions. The thought of using partial derivatives to prove smoothness is discouraging.

Fortunately, Grothendieck gave an alternative definition of smoothness which is much easier to handle in such cases. To explain it, we need the concept of points with values in a commutative algebra.

Let R be a noetherian commutative ring, and let U be a finitely generated commutative R -algebra. So there is some presentation of U , of the form $U = R[x_1, \dots, x_n]/(f_1, \dots, f_r)$. Let $X = \text{Spec } U$. This scheme is analogous to the locus of zeros of a polynomial in \mathbb{C}^n , and it will play mainly an intuitive role.

Let S be a commutative R -algebra. A *point* p of X (or of U) *with values in* S is a homomorphism of algebras $U \xrightarrow{p} S$.

The *affine space* \mathbb{A}_R^n over R is the spectrum of the polynomial ring $U = R[x_1, \dots, x_n]$. A point of \mathbb{A}_R^n with values in S is given by an arbitrary collection of elements a_1, \dots, a_n in S , i.e., by a "point" (a_1, \dots, a_n) whose coordinates are in S . The map $U \xrightarrow{p} S$ sends $x_i \mapsto a_i$. So the points of \mathbb{A}^n with values in S form a free S -module of rank n .

If U is presented as above, $U = R[x_1, \dots, x_n]/(f_1, \dots, f_r)$, then a point of X with values in S is given by a solution of the system of equations $f(x) = 0$ in S , a collection of elements $a_1, \dots, a_n \in S$ such that $f_1(a) = \dots = f_r(a) = 0$. So a point of X is a "point" (a_1, \dots, a_n) in affine space at which the polynomials $f_\nu(x)$ take the value zero.

Note that if $S \rightarrow S'$ is a homomorphism of R -algebras and if p is a point of X with values in S , then we obtain a point with values in S' by composition. In terms of coordinates, if our point is (a_1, \dots, a_n) with $a_i \in S$, then the coordinates of the image point (a_1, \dots, a_n) are the images of the a_i in S' .

The data necessary for Grothendieck's definition is

$$(III.11.5) \quad N \subset S \longrightarrow S,$$

where S is a commutative R -algebra, $N \subset S$ is an ideal with $N^2 = 0$, and $S = S/N$. We say that U has the *lifting property* for this data if every point p of U with values in S is induced by a point p with values in S , in other words, every diagram

$$\begin{array}{ccc} & & S \\ & & \downarrow \\ U & \xrightarrow{p'} & S \end{array}$$

of maps of R -algebras can be completed by a map $U \xrightarrow{p} S$.

In terms of coordinates, the lifting property says this: Given a solution (a_1, \dots, a_n) of the equations $f_1 = \dots = f_r = 0$ with $a_i \in S$, there exists a solution (a_1, \dots, a_n) , with $a_i \in S$, such that $a_i \equiv a_i$ (modulo N).

Definition III.11.6: A finitely generated commutative R -algebra U is *smooth* if it has the lifting property for all data III.11.5.

It is a fact that, in order to verify smoothness, it is enough to prove the lifting property in the case that S is a finitely generated, artinian R -algebra.

Example III.11.7: We use Newton's method to verify that a polynomial $f(x) \in R[x]$ defines a smooth algebra according to Grothendieck's definition if F and f generate the unit ideal. A point of $U = R[x]/(f)$ with values in S is a solution $x = a$ of the equation $f(x) = 0$ with $a \in S$. With notation as in (III.11.5), let $x = a$ be a solution in S . Looking for a solution in S which lifts a , we begin by choosing an arbitrary lifting, say to $a \in S$. This arbitrary lifting is unlikely to be a solution, but since $a \equiv a$ (modulo N) and $f(a) = 0$, we can conclude that $f(a) \equiv 0$ (modulo N). We are allowed to adjust our original choice a by adding an element of N , so we try to solve the equation $f(a + h) = 0$, with $h \in N$. Because $h^2 = 0$, Taylor's formula reduces to $f(a + h) = f(a) + hf'(a)$. So we look for h so that $hf'(a) = -f(a)$. Now since f, f' generate the unit ideal, $pf' + qf = 1$ for some $p, q \in R[x]$. So $pf'^2 + qff' = f$. Substituting $x = a$, the first term lands in N^2 , hence is zero. So $q(a)f'(a)f'(a) = f(a)$, and we may set $h = -q(a)f'(a)$.

We'll omit the proof of the next proposition.

Proposition III.11.11. *A smooth algebra or scheme is flat over R .*

Exercise: Let $U = \mathbb{C}[x, y]/(xy)$, and define $S_n = \mathbb{C}[t]/(t^{n+1})$. Starting with the point $U \xrightarrow{p_0} S_0$ whose coordinates are $(0, 0)$, determine its liftings to S_1 and to S_2 , and verify that U does not satisfy Grothendieck's definition of smoothness.

Proposition III.11.8. *Let A be an R -algebra which is a free module, and let U be the algebra defined by the idempotent relations (III.11.4). Then U is a smooth R -algebra.*

Proof. As was illustrated above in the case of one equation, one can try to use the explicit equations and Newton's method to lift solutions, which amounts to using a jacobian criterion for smoothness. That method is hard to apply in our case. Also, the equations (III.11.4) are not intrinsic to the problem of finding idempotents, because they depend on the choice of basis.

We will present Grothendieck's general method for verifying the lifting property for an algebra U . The method requires us to have a usable interpretation of the meaning of a point of U with

values in a commutative algebra S . In our case, a point of U with values in an R -algebra S has an intrinsic interpretation: It corresponds to an idempotent element of the algebra A_S . That is how the equations are defined. A solution (a_1, \dots, a_n) of (III.11.4), with values in S , yields an idempotent element e in $A_{S'}$, and lifting this to a solution (a_1, \dots, a_n) with values in S is equivalent with the problem of lifting the idempotent e from the algebra $A_{S'}$ to an idempotent of A_S . The algebra $A_{S'}$ is obtained by killing the square zero ideal NA_S in A_S , so we know that the lifting is possible. Therefore U is smooth! \square

A similar analysis can be made for the matrix unit equations $e_{ij}e_{kl} = \delta_{ij}e_{il}$ and $e_{11} + \dots + e_{rr} = 1$. If A is a free R -algebra of rank r^2 with a basis u_α , we can write down a system of “matrix unit equations” $f_\nu(x) = 0$ in r^3 variables $x_{ij\alpha}$, such that a solution of the system in S yields matrix units, hence yields an isomorphism of A_S with the matrix algebra $M_r(S)$. The exact nature of the equations is not important, but they are

$$(III.11.9) \quad \sum_{\alpha\beta\gamma} c_{\alpha\beta\gamma} x_{ij\alpha} x_{kl\beta} = \delta_{jk} x_{il\gamma}, \quad \text{and} \quad \sum_{\alpha} x_{ij\alpha} = 1.$$

Now it is known that if $S = S/N$ where $N^2 = 0$, then a system e_{ij} of matrix units in A_S can be lifted to a system e_{ij} in A_S . This shows

Proposition III.11.10. *The system of matrix unit equations defines a smooth algebra $U = R[x_{ij\alpha}]/(f_\nu)$. \square*

To close this section, we note that any commutative R -algebra U has a tautological point with values in U , namely the one given by the identity map $U \rightarrow U$. It is worthwhile spending a moment to reinterpret this point. First, if $U = R[x_1, \dots, x_n]/(f_1, \dots, f_r)$, then a point $U \xrightarrow{p} S$ with values in S corresponds to a solution (a_1, \dots, a_n) of the system of equations $f(x) = 0$ in S . The solution corresponding to the identity map $U \rightarrow U$ takes for a_i the residue of x_i in U . Also, if we can interpret the meaning of a point with values in S , then we can apply that interpretation to the point $U = U$. Thus if U is the algebra defined by the matrix unit equations (III.11.9), then a point of U with values in S corresponds to a system of matrix units in the algebra A_S . So the identity map $U = U$ corresponds to a set of matrix units in A_U , and A_U is isomorphic to a matrix algebra over U .

Now most algebras aren't isomorphic to $r \times r$ matrix algebras, and most can't be made into matrix algebras by changing the ground ring. For instance, if A happens to be commutative, then $A \otimes U$ will be commutative for all commutative U , whereas if $r > 1$, the matrix algebra $M_r(U)$ is not commutative - unless U is the zero ring. The matrix algebra over the zero ring is the only commutative one. Thus, for a random algebra A , it is quite likely that our ring U will be the zero ring, i.e., the polynomials (III.11.9) will generate the unit ideal.

Exercise: The identity map $U \rightarrow U$, the corresponding solution of the system of equations in U , and their interpretation in case of a problem such as matrix units, all have universal properties. Discuss.

III.12. Azumaya algebras

Azumaya algebras are analogous to bundles of matrix algebras in topology. As we have seen, we should allow more general forms of “localization” than the adjunction of inverses.

Proposition III.12.1. *Let R be a noetherian commutative ring, and let A be an R -algebra which is a finite R -module. The following are equivalent. If one of them holds, then A is called an Azumaya algebra over R .*

(i) *There is a faithfully flat, commutative R -algebra R' such that $A \otimes_R R'$ is isomorphic to a matrix algebra $M_n(R')$.*

(ii) *There is a smooth commutative R -algebra R' such that the map $\text{Spec } R' \rightarrow \text{Spec } R$ is surjective, and that $A \otimes_R R'$ is isomorphic to a matrix algebra $M_n(R')$.*

(iii) *A is a projective R -module of constant rank, and for every map $R \rightarrow \overline{K}$ where \overline{K} is an algebraically closed field, $A \otimes_R \overline{K}$ is isomorphic to a matrix algebra over \overline{K} .*

(iv) *A is a projective R -module of constant rank, and the canonical map $A^\circ \otimes_R A \rightarrow \text{End}_R A$ is bijective.*

Proof. (i) \Rightarrow (iii): Let R' be as in (i), and let $R' \rightarrow \overline{K}$ be a homomorphism. Let $S = R' \otimes_R \overline{K}$. Then S is not the zero algebra because R' is faithfully flat over R . Hence S is faithfully flat over the field \overline{K} . Also, $A_S = A \otimes_R S \approx A_{R'} \otimes_{R'} S$. Since $A_{R'}$ is a matrix algebra, so is A_S . We replace R by \overline{K} and R' by S . The fact that A_S is a matrix algebra means that the matrix unit equations (III.11.9) have a solution in some nonzero ring. Hence the algebra U (III.11.10) is not the zero ring, and by the Nullstellensatz, U has a point with values in the algebraically closed field \overline{K} . Thus $A_{\overline{K}}$ is a matrix algebra.

(iii) \Rightarrow (iv): Let $E_1 = A^\circ \otimes A$ and $E = \text{End}_R A$. The right action of E_1 on A commutes with the R -action, which we view as a left action. This defines the canonical map $f : E_1 \rightarrow E$ referred to in the statement of the theorem. Note that E_1 and E are locally free R -modules of the same ranks. A map $f : E_1 \rightarrow E$ between locally free modules is bijective if and only if the induced map $E_1 \otimes \overline{K} \rightarrow E \otimes \overline{K}$ is bijective for every homomorphism $R \rightarrow \overline{K}$ to a field. So to prove (iii), we may replace R by \overline{K} . Then (ii) tells us that A is a matrix algebra, for which we verify directly that f is bijective.

(iv) \Rightarrow (iii): This reduces to showing that if R is an algebraically closed field and (iii) holds, then A is a matrix algebra. Since R is a field, $\text{End}_R A$ is a matrix algebra, hence a central simple R -algebra. Applying (1.5) and (1.8), we see that the center of A is R and that A is a simple ring. So A is central simple. Since R is algebraically closed, A is a matrix algebra.

(iii) \Rightarrow (ii): We look at the matrix unit equations (III.11.9) and the algebra U they define. This algebra is smooth. It suffices to show that the map $\text{Spec } U \rightarrow \text{Spec } R$ is surjective (III.5.2). Then we can take $R' = U$, and use the solution corresponding to the identity map $U \rightarrow U$. Surjectivity of the map of spectra is equivalent with saying that every map $R \rightarrow \overline{K}$, where \overline{K} is an algebraically closed field, has a point of U with values in \overline{K} lying over it. Translating this, for every map $R \rightarrow \overline{K}$, there is a system of matrix units in the algebra $A_{\overline{K}}$. This is true because $A_{\overline{K}}$ is a matrix algebra.

(ii) \Rightarrow (i): This follows from (III.11.11) and (III.5.2). \square

We now take a second look at the proof of Theorem III.3.1, and adapt it to extend the Skolem-Noether theorem.

Theorem III.12.2. (*Skolem-Noether*) *Let R be a commutative, noetherian local ring, and let A be an Azumaya algebra over R . Then every R -automorphism of A is inner.*

Proof. For the moment, R can be arbitrary. Let θ be an R -automorphism of A . We make A into an (A, A) -bimodule in two ways: First in the usual way, and second by letting A act on the right through θ . We denote this second module by A^* . Thus for $x \in A^*$, right multiplication is defined by $x * a = xa^\theta$. Let $E = A^o \otimes A \approx \text{End}_R A$. Let us suppose for the moment that A and A^* are isomorphic right E -modules, and let $\phi : A \rightarrow A^*$ be an isomorphism. The proof of (III.3.1) shows that θ is inner. Conversely, if θ is inner, say $a^\theta = u^{-1}au$ where u is invertible in A , then right multiplication by u is an E -isomorphism $A \rightarrow A^*$.

We need to conclude that, if R is a local ring, then A and A^* are isomorphic. Since A is a projective R -module, we have a Morita equivalence of categories

$$(III.12.3) \quad \text{Mod } R \xrightarrow{-\otimes_R A} \text{Mod } E.$$

It has an inverse $-\otimes_E Q$, where $Q = \text{Hom}_E(A, E)$. Then $L = A^* \otimes_E Q$ is a projective R -module. If R is a local ring, then L is free, and one sees that its rank must be 1, hence it is isomorphic to $R \approx A \otimes_E Q$. Because $-\otimes_E Q$ is an equivalence, A and A^* are isomorphic. \square

Exercise: Prove the following proposition:

Proposition III.12.4. *An Azumaya algebra A over a commutative ring R has a well defined determinant $\det : A \rightarrow R$. This determinant is compatible with extension of scalars $R \rightarrow R'$, and it agrees with the usual determinant when A is a matrix algebra.*

III.13. Dimension of a variety

Let k be an algebraically closed field. Every nonempty variety X over k has a *dimension*, with these properties:

Theorem III.13.1. (i) *If $\dim X = 0$, then X consists of a finite set of points. If $\dim X > 0$, then X contains infinitely many points.*

(ii) *The affine space $\mathbb{A}_k^n = \text{Spec } k[x_1, \dots, x_n]$ has dimension n .*

(iii) *If X is irreducible and if Y is a proper closed subvariety of X , then $\dim Y < \dim X$.*

(iv) *If $f : X \rightarrow Y$ is a morphism of varieties and if y is a point of Y , then the fibre $f^{-1}(y)$, a closed subvariety of X , has dimension at least $\dim X - \dim Y$ unless it is empty.*

III.14. Background on algebraic curves

Let k be an algebraically closed field. A *function field in one variable* K over k is a field extension of transcendence degree 1, which is finitely generated as field extension. We'll call such a field a *function field* for short.

Every function field K has a unique smooth projective model X , an algebraic curve. Calling it projective means that it can be embedded into projective space in some way. We won't use an

explicit embedding. For example, if $K = k(t)$ is a pure transcendental extension, then X is the projective line \mathbb{P}^1 over k .

The points of X are in bijective correspondence with discrete valuations of K . We'll denote the valuation associated to a point p of X by v_p . So for $a \in K$, its valuation at p , denoted by $v_p(a)$, is an integer which is zero for all but a finite number of points p . If $v_p(a)$ is positive, it is called the *order of zero* of a at p . If $v_p(a)$ is negative, then $-v_p(a)$ is called the *order of pole* of a at p .

A *divisor* on X is an integer linear combination of points: $D = \sum r_i p_i$, the sum being finite. The *degree* of a divisor $D = \sum r_i p_i$ is the integer $\sum r_i$.

The divisor of a function a is defined to be

$$(III.14.1) \quad \sum_p v_p(a) p,$$

“the zeros minus the poles of a ”. It is usually denoted by the overused symbol (a) , and is often split into its positive and negative parts: $(a) = (a)_0 - (a)_\infty$, where $(a)_0$ is the divisor of zeros, and $(a)_\infty$ is the divisor of poles. The degree of the divisor of a function is zero:

$$(III.14.2) \quad \sum v_p(a) = 0.$$

A function has the same number of zeros as poles.

The Riemann-Roch theorem computes the space of elements of K whose poles are bounded by a fixed divisor D . This set becomes a vector space over k if the zero element (which does not have a divisor) is added. A critically important fact is that the dimension of this vector space is finite for any divisor D .

The Riemann-roch theorem makes reference to a certain invariant of the curve X , its *genus* g . The genus of \mathbb{P}^1 is zero.

Theorem III.14.3. (*Riemann-Roch*) *Let D be a divisor of degree d on X , and let V be the vector space of elements of K whose poles are bounded by D :*

$$V = \{f \in K \mid (f)_\infty < D\}.$$

Then if $d > 2g - 2$,

$$\dim V = d + 1 - g.$$

In any case, $\dim V \geq d + 1 - g$.

Exercise: Prove the Riemann-Roch theorem for \mathbb{P}^1 .

III.15 Tsen's theorem

Theorem III.15.1. (Tsen) *Let K be a function field in one variable over an algebraically closed field k . Every division ring D which is a K -algebra, finite as K -module, is commutative. Hence the Brauer group of K is the trivial group.*

Proof. We may assume that K is the center of D , so that D is a central simple algebra over K . We try to adapt the method of proof of Theorem III.9.2. Since D is a division ring, the determinant of any nonzero element $a \in D$ is nonzero. If $[D : K] = n^2$, then the determinant is computed by evaluating a polynomial Δ with coefficients in K , which is homogeneous of degree n in n^2 variables x_i . The statement that $\det(a) \neq 0$ for all $a \neq 0$ means that this polynomial does not take the value zero for any substitution $x_i = a_i$, with $a_i \in K$ and not all zero. The next theorem shows that the only possibility is that $n = 1$ and therefore $D = K$. \square

Theorem III.15.2. (Tsen-Lang) *Let K be a function field in one variable over an algebraically closed field k , and let $F(x_1, \dots, x_r)$ be a homogeneous polynomial in $K[x_1, \dots, x_r]$ of degree $d < r$. There exist elements a_1, \dots, a_r in K , not all zero, such that $F(a_1, \dots, a_r) = 0$.*

Proof. We will prove this theorem using Riemann-Roch. To do so, we look for a zero (a_1, \dots, a_r) such that the elements a_i have poles only at a fixed point p of X , say of degree $\leq n$. We'll allow n to be as large as necessary, in particular, larger than $2g - 2$. By Riemann-Roch, the space V of all such functions has dimension $n + 1 - g$. The space V^r of r -tuples (a_1, \dots, a_r) of such functions obviously has dimension

$$(III.15.3) \quad \dim V^r = r(n + 1 - g) = rn + \text{const}(n).$$

Now if $(a_1, \dots, a_r) \in V^r$, what are the poles of $F(a)$? First of all, F has coefficients in the function field K , and these coefficients will have poles. The coefficients of F are fixed elements of K . Let Z be a divisor larger than the poles of all of these coefficients, and let z be the degree of Z . Second, each monomial of degree d in $\{a_i\}$ contributes poles only at p , and of order $\leq dn$. Thus the poles of $F(a)$ are bounded above by $Z + dn p$, which is a divisor of degree $z + dn$. So $F(a)$ is contained in the vector space W of all functions with at most these poles. Riemann-Roch tells us that

$$(III.15.4) \quad \dim W = dn + z + 1 - g = dn + \text{const}(n).$$

Since $r > d$, we find that $\dim V^r > \dim W$ if n is sufficiently large.

Now we make a slight contortion. We regard the vector spaces V^r and W as varieties. To be precise, a k -vector space W of dimension N is isomorphic to the space k^N via a choice of basis, and k^N has another incarnation, as the space of points of affine space \mathbb{A}_k^N with values in k (see Section 10). So W is isomorphic to the space of points of an affine space too. At the risk of confusion, we'll denote the affine spaces associated to our two vector spaces by V^r and W too.

Lemma III.15.5. *The polynomial F defines a morphism of varieties $V^r \rightarrow W$ which sends (a_1, \dots, a_r) to $F(a)$.*

Assume the lemma. The dimension formula (III.13.iv) shows that the nonempty fibres of F have positive dimension. The fibre over 0 in W is not empty because $F(0, \dots, 0) = 0$. Thus it has positive dimension, and contains infinitely many nonzero points (a_1, \dots, a_r) . \square

Proof of the lemma. What has to be verified is that, whereas F has coefficients in the function field K , its restriction to V^r is described by some polynomial functions with coefficients in the ground field k .

We choose a k -basis for V , say v_1, \dots, v_m , and write the elements a_1, \dots, a_r of V in terms of this basis: $a_i = \sum y_{ij} v_j$, where y_{ij} are undetermined coefficients which are to be evaluated in k . The substitution $x_i = a_i$ into F yields a polynomial $\Phi(y)$ in the variables y_{ij} , say, in multi-index notation,

$$(III.15.6) \quad \Phi(y) = \sum c_\alpha y^\alpha.$$

The coefficients c_α are obtained by algebraic operations from the coefficients of F and the v_j . So they are elements of the function field K . Let W denote the finite dimensional k -subspace of K spanned by the coefficients and by W . Since the monomials y^α are polynomials, formula III.15.6 exhibits the restriction of F to V^r as a polynomial map $V^r \rightarrow W$. Since its image is contained in W , we also obtain a polynomial map to W by projection. \square

IV. MAXIMAL ORDERS

Terminology: R : a commutative, noetherian, often a Dedekind domain, and K the field of fractions of A .

A_K : a K -algebra which is a finite K -module, and which is often central simple.

For many reasons, it is useful to “extend” A_K over $\text{Spec } R$. This means finding an algebra A over R such that $A \otimes K \approx A_K$. For instance, if K is the field of rational numbers, we may want to reduce A_K modulo p , and in order for this to make sense, the defining equations for A_K must have integer coefficients.

Extending to $\text{Spec } R$ amounts to this: choosing generators and defining relations for A_K , and manipulating them so that the defining relations have coefficients in R . The most primitive approach would be simply to take a set of defining equations, and clear denominators from the coefficients. Such a direct approach is unworkable. Complications arise because the procedure is far from unique. One needs to choose the generators and relations very carefully.

Suppose that A_K is central simple. The nicest situation is when we can choose the relations so that A is an Azumaya algebra. But this may not be possible, even when the defining relations are chosen optimally. In that case we say that the algebra A_K “ramifies”, and we’d like a structure theorem for the extended algebra. There is structure theorem when R is a Dedekind domain.

As a simple example, the equations defining the algebra of quaternions, which are $\mathbf{i}^2 = -1, \mathbf{j}^2 = -1, \mathbf{j}\mathbf{i} = -\mathbf{i}\mathbf{j}$, have integer coefficients. The \mathbb{Z} -algebra $A = \mathbb{Z}\langle \mathbf{i}, \mathbf{j} \rangle / (\mathbf{i}^2 = \mathbf{j}^2 = -1, \mathbf{j}\mathbf{i} = -\mathbf{i}\mathbf{j})$, whose elements are the integer linear combinations $a + b\mathbf{i} + c\mathbf{j} + d\mathbf{i}\mathbf{j}$, is a fairly nice \mathbb{Z} -algebra. Its reduction modulo p is a matrix algebra for every integer prime $p \neq 2$. But when we reduce modulo 2, the algebra becomes commutative. So A is not an Azumaya algebra. We will discuss this example in more detail later.

Exercise: Prove that $A \otimes \mathbb{F}_p$ is a matrix algebra for all $p \neq 2$.

IV.1. Lattices and orders

An R -lattice L in a finitely generated right K -module V is a finitely generated R -submodule which generates V as K -module. We may refer to an R -lattice simply as a *lattice*, for short.

A reminder: If S denotes the multiplicative set of nonzero elements of the domain R , then $L \otimes K = LS^{-1}$. So L generates V if and only if $L \otimes K \approx V$.

Proposition IV.1.1. (i) Suppose that V is generated as K -module by elements v_1, \dots, v_n . Then the set $L = v_1R + \dots + v_nR$ is a lattice. Every lattice has this form.

(ii) If $R \neq K$ and $V \neq 0$, then V contains infinitely many lattices.

(iii) If L is an R -lattice in V and $u \in K$ is not zero, then Lu is an R -lattice.

(iv) Let L be an R -lattice in V and let M be any finitely generated R -submodule of V . There is a nonzero element $s \in R$ such that $Ms \subset L$.

(v) Let M be a finitely generated, torsion-free R -module. The canonical map $M \rightarrow V := M \otimes K$ is injective, and M is an R -lattice in V . \square

An R -order A in the finite K -algebra A_K is an R -subalgebra which is a lattice in A_K , i.e., which is a finitely generated R -module such that $A \otimes K \approx A_K$. We may also refer to an R -order simply as an *order*.

Constructing R -orders is not as easy as constructing lattices, because the R -algebra generated by a finite set of elements need not be a finite module. There are far fewer orders than lattices. For example, if $A_K = K$, then the only R -order is R itself. The \mathbb{Z} -algebra generated by the fraction $\frac{1}{2}$ is not an order, though the \mathbb{Z} -module $\frac{1}{2}\mathbb{Z}$ is a lattice.

There is a tricky way to construct an order, and thereby to show that orders always exist. I find this method extremely elegant. The first step is to choose an arbitrary finite, faithful, right A_K -module V , and a lattice L in V . For example, we may choose $V = A_K$. The second step is to take for A the subset of A_K of elements which carry the lattice L to itself:

$$(IV.1.2) \quad A = \{a \in V \mid La \subset L\}.$$

Proposition IV.1.3. *The set (IV.1.2) is an R -order. Hence every finite K -algebra A_K contains an order.*

Proof. It is easily seen that A is a ring, that $R \subset A$, and that A is an R -algebra. To see that A is a finite R -module, we note that right multiplication by $a \in A$, the map $\rho_a : L \rightarrow L$, is R -linear because $ar = ra$ for $a \in A$ and $r \in R$. It is not the zero map unless $a = 0$, because $\rho_a \otimes K$ is right multiplication by a on V , and V is faithful. So sending $a \mapsto \rho_a$ defines an injective map $A \rightarrow \text{End}_R L$. Since L is a finite R -module, so is $\text{End}_R L$. Since R is noetherian, A is also a finite R -module.

Next, to show that $A \otimes K \approx A_K$, let $a \in A_K$. If v_1, \dots, v_r is a set of generators for L as R -module, then $v_1 a, \dots, v_r a$ generate La . So La is a finitely generated R -module, and there is a nonzero element $s \in R$ such that $Las \subset L$. Thus $as \in A$, and $a \in A \otimes K$. \square

If A is an R -order, then we may also define the concept of A -lattice in a finite right A_K -module V . An A -lattice is an A -submodule of V which is also an R -lattice. Note that since the operation of R is central, we may consider an A -lattice as an (R, A) -bimodule. The analogues of the assertions of IV.1.1 for R -lattices carry over.

Lemma IV.1.4. *Let A be an R -order in a finite K -algebra A_K .*

(i) *Let V be a finite A_K -module, generated by elements v_1, \dots, v_n . The set $L = v_1 A + \dots + v_n A$ is an A -lattice. Moreover, every A -lattice has this form.*

(ii) *If $R \neq K$ then any nonzero A_K -module V contains infinitely many lattices.*

(iii) *If L is an A -lattice in V and $u \in K$ is not zero, then Lu is an A -lattice.*

(iv) *Let L be an A -lattice in V and let M be any finitely generated A -submodule of V . There is a nonzero element $s \in R$ such that $Ms \subset L$.*

(v) *Let A be an R -algebra which is a finite, torsion-free R -module. then A is an R -order in $A \otimes K$.*
 \square

Exercise: (1) Prove that if M, N are finite R -modules, then $\text{Hom}_R(M, N)$ is a finite R -module.

(2) Let L be a lattice in V and let $V = L \otimes K$. Show that $\text{End}_R L$ is an order in $\text{End}_K V$.

IV.2. The trace pairing on an Azumaya algebra

Let M, N be finite, locally free R -modules. The concept of an R -bilinear form $\langle -, - \rangle : M \times N \rightarrow R$ is defined exactly as when R is a field. If M and N are free modules with bases $\{u_i\}$ and $\{v_j\}$ respectively, then the form is determined by the matrix $\langle u_i, v_j \rangle$, which can be an arbitrary matrix with entries in R .

Let $M^* = \text{Hom}_R(M, R)$ be the dual module. A bilinear form defines a map $N \rightarrow M^*$, by the rule $y \mapsto \langle -, y \rangle$. This is an R -linear map.

Suppose that $M = N$ and that the form is symmetric. If the map $M \rightarrow M^*$ defined by the form is bijective, then the form is called *nondegenerate*. It is important to notice that this is stronger than saying that the nullspace of the form is zero. The nullspace of the form is zero whenever the map $M \rightarrow M^*$ is injective, and if R is a field, this implies bijectivity. Not otherwise.

In case M is free, with basis $\{u_i\}$, the form is nondegenerate if and only if its *discriminant* $\delta = \det(\langle u_i, u_j \rangle)$ is an invertible element of R . A change of basis, say $v = Pu$, where P is an invertible R -matrix, changes the discriminant by the factor $(\det P)^2$. Since $\det P$ is invertible, this is the square of a unit. So the discriminant is determined only locally, and locally only up to the square of a unit factor.

Proposition IV.2.1. *Let A be an Azumaya algebra over an arbitrary commutative noetherian ring S . The map $A \times A \rightarrow S$, defined by*

$$\langle a, b \rangle = \text{trace}(ab)$$

is a nondegenerate symmetric bilinear form on the S -module A .

Proof. In the case that A is a matrix algebra, this lemma is proved by taking a look, and it follows in the general case by descent. \square

Proposition IV.2.2. *Let R be a commutative, noetherian, normal domain, and let A be an R -order in a central simple K -algebra A_K . For any $a \in A$, the characteristic polynomial $p(t)$ of a is in R .*

Proof. Let $R[a]$ denote the commutative R -subalgebra of A generated by a . This is a finite R -module because it is an R -submodule of A , A is a finite module, and R is noetherian. Therefore a is integral over R , and is the root of a monic polynomial, say $g(t)$, with coefficients in R .

Going over to K , we know that the kernel I_K of the map $K[t] \rightarrow K[a]$ is a principle ideal, generated by a monic polynomial $f(t)$. Going further, to the algebraic closure \overline{K} , we have $K[a] \otimes \overline{K} \subset A \otimes \overline{K}$ because \overline{K} is flat over K , so $K[a] \otimes \overline{K}$ is the \overline{K} -subalgebra of $A \otimes \overline{K} = M_n(\overline{K})$ generated by a . It follows that f also generates the kernel of the map $\overline{K}[t] \rightarrow \overline{K}[a]$. This shows that f is the minimal polynomial of the matrix a . The minimal polynomial and the characteristic polynomial $p(t)$ have the same roots. So $p(t)$ divides a power of $f(t)$ in $\overline{K}[t]$ and in $K[t]$.

Next, suppose that R is a discrete valuation ring, so that, in particular, it is a unique factorization domain. Let f_0 and p_0 denote the primitive polynomials, determined up to unit factor, associated with f and p . Then f_0 generates the kernel I_R of the map $R[t] \rightarrow R[a] \subset K[a]$. (If f divides h in $K[t]$ and if $h \in R[t]$, then f_0 divides h in $R[t]$.) We saw above that the kernel I_R contains a monic polynomial. Therefore f_0 must also be monic (up to unit factor), i.e., $f_0 = f$. Since p

divides a power of f , p_0 divides a power of f_0 , which shows that p_0 is monic and equal to p , hence that $p \in R[t]$.

Finally, to treat the case of a general normal domain R , we apply the next lemma, a standard result from commutative algebra. \square

Recall that the height one prime ideals P in a domain R are those which are minimal among nonzero prime ideals.

Lemma IV.2.3. *Let R be a commutative, noetherian, normal domain.*

(i) *The local ring R_P at a height one prime ideal P is a discrete valuation ring.*

(ii) *$R = \bigcap_P R_P$, where P runs over the height one primes of R , and R_P is the localization of R at P .*

The second part of this lemma becomes an important tool when one tries to extend results from Dedekind domains to domains of higher dimension.

IV.3. Separable algebras

Let K be a field. A finite K -algebra A_K is *separable* if it is a direct sum $A_1 \oplus \cdots \oplus A_r$, where each A_i is a central simple algebra over a separable field extension K_i of K . Equivalently, A_K is separable if and only if $A_K \otimes \overline{K}$ is a direct sum of matrix algebras over \overline{K} .

Proposition IV.3.1. *Let K be a field, and let A_K be a finite K -algebra. Suppose that $\tau : A_K \rightarrow K$ is a linear function such that the bilinear form*

$$\langle a, b \rangle = \tau(ab),$$

is a symmetric and nondegenerate form on A_K . Then A_K is a separable K -algebra.

Proof. The linear form extends to \overline{K} , so we may assume that K is algebraically closed. The next lemma shows that every ideal yields a decomposition of A_K into a direct sum of subrings, hence that A_K decomposes as a direct sum of simple rings. By Wedderburn's theorem, each of these simple rings is a matrix algebra. \square

Exercise: Prove the converse: If A_K is a separable algebra, then such a linear form τ exists.

Lemma IV.3.2. *Let I be an ideal of A_K , and let I^\perp be its orthogonal complement. Then I^\perp is an ideal, and $I I^\perp = I^\perp I = 0$. So $A_K = I \oplus I^\perp$ is a decomposition of A_K as a direct sum of rings.*

Proof. By definition, $I^\perp = \{y \in A_K \mid \langle y, I \rangle = 0\}$. Then for $a \in A_K$, $\langle ya, I \rangle = \tau(yaI) = \langle y, aI \rangle \subset \langle y, I \rangle = 0$. This shows that I^\perp is a right ideal. Since $\langle I, y \rangle = 0$ as well, I^\perp is also a left ideal. It is clear that $I I^\perp = I^\perp I = 0$, and the fact that I and I^\perp are subrings follows. \square

The decomposition of 1 in $I \oplus I^\perp$ yields a pair of idempotents which are in the center of A_K .

Proposition IV.3.3. *Let A be an R -order in a central simple K -algebra A_K . Assume that A is a locally free R -module. Then A is an Azumaya algebra if and only if $\langle a, b \rangle = \text{trace}(ab)$ defines a nondegenerate pairing $A \times A \rightarrow R$.*

Proof. The “only if” part is contained in Proposition IV.2.1. Suppose that the trace pairing is nondegenerate. To prove that A is an Azumaya algebra, it suffices to show that $A_{\overline{L}}$ is a matrix

algebra over \bar{L} for every homomorphism $R \longrightarrow \bar{L}$ to an algebraically closed field (CSA, III.12.1). Tensoring with \bar{L} , the trace pairing is the composition of multiplication in $A_{\bar{L}}$, composed with the linear form $\tau = \text{trace} \otimes \bar{L}$. But we can't conclude that the linear form is trace until we show that $A_{\bar{L}}$ is a matrix algebra, which is what we are trying to prove. Anyway, we can apply the previous proposition to this pairing, to conclude that $A_{\bar{L}}$ is a separable algebra. To show that it is a matrix algebra, it suffices to show that $A_{\bar{L}}$ doesn't contain any central idempotent other than 0, 1.

There are probably more elegant ways to prove this, but I'm too lazy to think of one. At this stage, we can get through just by beating it to death. Suppose that $\bar{e} \in A_{\bar{L}}$ is a central idempotent different from 0 and 1. Let U be the smooth R -algebra defined by the idempotent equations (CSA, III.11.4). The idempotent \bar{e} corresponds to a point of U with values in \bar{L} . Let P be the kernel of the map $U \longrightarrow \bar{L}$, let S be the local ring U_P , and let L be the residue field of S . So we have a sequence of maps $U \longrightarrow S \longrightarrow L \longrightarrow \bar{L}$, and corresponding idempotents $e_U \mapsto e_S \mapsto e_L \mapsto e_{\bar{L}} = \bar{e}$.

Lemma IV.3.4. *Let S be a local ring with residue field L and let \bar{L} be the algebraic closure of L . Let e be an idempotent element of a finite S -algebra A_S . If the image of e in $A_{\bar{L}}$ is central, then e is central.*

Proof. Say that $e = e_1$ and that $e_1 + e_2 = 1$. Consider the Peirce decomposition $A = \bigoplus A_{ij}$, where $A_{ij} = e_i A e_j$. Because $e_1 e_2 = e_2 e_1 = 0$ and e_1 is the identity element in A_{11} , e_1 is central in A if and only if $A_{12} = A_{21} = 0$. The Peirce decomposition of A induces the Peirce decompositions of A_L and $A_{\bar{L}}$ by tensor product, and by the Nakayama lemma, $A_{12} = 0$ if and only if $A_{12} \otimes L = 0$, which is true if and only if $A_{12} \otimes \bar{L} = 0$. \square

Scholium IV.3.5. Suppose that A is an R -order in a central simple K -algebra A_K . There are two obstructions to A being an Azumaya algebra. First, A must be a locally free R -module, and second, the trace pairing must be nondegenerate. Both of these are open conditions on $\text{Spec } R$.

If A is locally free, then whether or not a given form is nondegenerate is determined by its discriminant δ , computed with respect to a local basis $\{u_i\}$ of A . The points where A is an Azumaya algebra are those at which δ does not take the value 0. This is an open set.

To see that the condition of being locally free is open, let M be a finite module over the commutative, noetherian, domain R , and let $n = \dim_K(M \otimes K)$. If $p \in \text{Spec } R$ is a point corresponding to a prime ideal P of R , let $k(p)$ be the residue field $k(p) = R_P/P_P$, and let

$$n(p) = \dim_{k(p)}(M \otimes k(p)).$$

Lemma 2.7. *Let M be a finite module over a commutative noetherian domain R . With the above notation,*

- (i) *The function $n(p)$ is upper semicontinuous on $\text{Spec } R$. In particular, $n(p) \geq n$ for all points p .*
- (ii) *M is locally free (or projective) at a point p if and only if $n(p) = n$.*

Thus the locus of points at which the order A is locally free is an open set in $\text{Spec } R$.

Proof. (i) We choose a free resolution of M , say

$$R^r \xrightarrow{P} R^s \longrightarrow M \longrightarrow 0,$$

where P is an $r \times s$ matrix with entries in R (operating on the right). Then $n(p) \geq d$ if and only if the rank of the matrix $P \otimes k(p)$ is at most $s - d$, i.e., if and only if the determinants of all $(s - d + 1)$ -rowed minors of P are zero. The vanishing of these determinants is a closed condition.

(ii) We may replace R by the local ring at p . The Nakayama lemma shows that we may choose the presentation so that $r = n(p)$. If some entry of P is not zero, then the rank of $P \otimes K$ is less than r , so $n(p) > n$. If $P = 0$, then M is free. \square

IV.4. Maximal orders in central simple algebras

An R -order A in a finite K -algebra A_K is called a *maximal order* if it is not properly contained in another order. Maximal orders in central simple algebras are analogues to integrally closed domains, and it is natural to concentrate the study of orders on them. A change that occurs when A_K is not commutative is that the maximal order is often not unique.

Theorem IV.4.1. *Let A_K be a central simple algebra. Every R -order in A_K is contained in a maximal order.*

Example IV.4.2: Some orders in a matrix algebra. Let's take the case that $R = k[t]$ and $K = k(t)$, where k is an algebraically closed field. Let $A_K = M_2(K)$. Then $A = M_2(R)$ is a maximal order (see (IV.4.5) below). It contains many nonmaximal orders which are isomorphic to A except at the point $t = 0$, for instance the order

$$(IV.4.3) \quad \begin{pmatrix} R & R \\ tR & R \end{pmatrix}.$$

There are also other maximal orders. If L is an arbitrary A -lattice in A_K , then $A = \text{End } L_A$ will be maximal (see (IV.4.5)). This yields examples such as

$$(IV.4.4) \quad L = \begin{pmatrix} tR & tR \\ R & R \end{pmatrix}, \quad A = \begin{pmatrix} R & tR \\ t^{-1}R & R \end{pmatrix}.$$

Proof of Theorem IV.4.1. The proof is similar to the proof that the integral closure of R in a finite separable field extension is a finite R -module. It uses the trace pairing defined in the previous section. Proposition 2.2 shows that the form $\langle a, a \rangle = \text{trace}(aa)$ takes its values in R if $a, a \in A$, and so it defines a pairing $A \times A \rightarrow R$. This in turn defines an R -linear map $A \rightarrow A^* = \text{Hom}_R(A, R)$, which is injective because the pairing is nondegenerate on A_K . Since A is a finite R -module, so is A^* .

Now suppose that $A \subset B$ are two orders. Then because B is an order, the trace pairing carries $B \times B \rightarrow R$. We can restrict this pairing to obtain a pairing $A \times B \rightarrow R$. This pairing defines an injective map $B \rightarrow \text{Hom}_R(A, R) = A^*$. Then we have inclusions $A \subset B \subset A^*$, which hems B in, because A^* is a finite R -module, and hence noetherian. \square

Corollary IV.4.5. *An Azumaya algebra over R is a maximal order.*

Proof. If A is an Azumaya algebra, the trace pairing is nondegenerate, and therefore $A = A^*$. Going back to the proof of the theorem, if $A \subset B$ are orders and A is Azumaya, then $A \subset B \subset A^*$, which shows that $A = B$. \square

Exercise: Show by example that an Azumaya algebra A over R needn't be Morita equivalent to R .

Example IV.4.6. Integer quaternions. Here $R = \mathbb{Z}$, and A is the order of quaternions with integer coefficients. In order to differentiate the complex number i from the element of A usually denoted by the same letter, we'll use bold face \mathbf{i} and \mathbf{j} for the quaternions. So A is a free \mathbb{Z} -module with basis $1, \mathbf{i}, \mathbf{j}, \mathbf{ij}$.

Let $R = \mathbb{Z}[i]$ be the ring of Gauss integers. There is a matrix representation of $A_{R'}$, given by

$$(IV.4.7) \quad \mathbf{i} \mapsto \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad \mathbf{j} \mapsto \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix},$$

and the formulas

$$e_{11} = \frac{1}{2}(1 - \mathbf{ii}), \quad e_{22} = \frac{1}{2}(1 + \mathbf{ii}), \quad e_{12} = -e_{11}\mathbf{j}, \quad e_{21} = e_{22}\mathbf{j}$$

show that $A_{R'}$ becomes isomorphic to the 2×2 matrix algebra when 2 is inverted, i.e., except at the prime 2. This shows that $A[\frac{1}{2}]$ is Azumaya algebra over $\mathbb{Z}[\frac{1}{2}]$, hence a maximal order. However A is not a maximal \mathbb{Z} -order. To obtain a maximal order, we must adjoin the element $\alpha = \frac{1}{2}(1 + \mathbf{i} + \mathbf{j} + \mathbf{ij})$. This is analogous to the fact that $\frac{1}{2}(1 + \sqrt{-3})$ is an algebraic integer.

Exercise: Prove that the \mathbb{Z} -lattice B having basis $1, \mathbf{i}, \mathbf{j}, \alpha$ is an order.

It is not hard to see that B is a maximal order, though it is not an Azumaya algebra. According to Proposition 2.2, the determinant of an element $\beta \in B$ must be an integer. Writing $\beta = a + b\mathbf{i} + c\mathbf{j} + d\mathbf{ij}$, we compute the determinant by using the matrix representation (IV.4.7):

$$(IV.4.8) \quad \beta = \begin{pmatrix} a + i & -c - di \\ c - di & a - i \end{pmatrix},$$

and

$$(IV.4.9) \quad \det(\beta) = a^2 + b^2 + c^2 + d^2.$$

Since $A[\frac{1}{2}]$ is maximal, $A[\frac{1}{2}] = B[\frac{1}{2}]$, and so if $\beta \in B$, then the denominators of a, b, c, d are powers of 2. Then in order for $\det(\beta)$ to be an integer, the denominators, if not 1, must all be 2. Therefore β is in the lattice spanned by $1, \mathbf{i}, \mathbf{j}, \alpha$.

Exercise: Compute the discriminant of the trace pairing on B , and determine the structure of $B \otimes \mathbb{F}_2$.

Exercise: Let $R = k[t]$, $K = k(t)$, and let A_K be the K -algebra $K[x]/(x^2)$. Prove that there is no maximal R -order in A_K .

Proposition IV.4.7. *Let A be a maximal R -order in the matrix algebra $A_K = M_n(K)$.*

(i) *Let L be an A -lattice in $V = K^n$. Then $A \approx \text{End}_R L$.*

(ii) *Suppose that R is a Dedekind domain. Then A has the form $\text{End}_R L$, where L is a projective R -module.*

Proof. (i) Every lattice is an ${}_R L_A$ bimodule, hence corresponds to a map $A \rightarrow \text{End}_R L$. Then $\text{End}_R L$ is also an order in $M_n(K) = \text{End}_K V$. Since A is maximal, $\text{End}_R L = A$.

(ii) This is true because every finite, torsion free module over a Dedekind domain is projective. \square

IV.5. Hensel's Lemma

Let R be a complete local ring with residue field k and maximal ideal \mathfrak{m} . For $f(t) \in R[t]$, we denote by \bar{f} the polynomial in $k[t]$ obtained by reducing the coefficients of f .

Proposition IV.5.1. *(Hensel's Lemma) Let $f(t) \in R[t]$, and suppose that in $k[t]$, $\bar{f} = \bar{g}\bar{h}$, where \bar{g} and \bar{h} are relatively prime, and \bar{g} is a monic polynomial. Then f factors accordingly: There are unique polynomials $g, h \in R[t]$ such that $f = gh$, g is monic, $\bar{g} = \bar{g}$, and $\bar{h} = \bar{h}$. Moreover, g, h generate the unit ideal in $R[t]$.*

Proof. We proceed by lifting the solution $\bar{f} = \bar{g}\bar{h}$ from R/\mathfrak{m} to each R/\mathfrak{m}^d step by step. This suffices because $R = \varprojlim R/\mathfrak{m}^d$. The general lemma which allows the lifting is the following:

Lemma IV.5.2. *Let S be a commutative ring, N an ideal of S such that $N^2 = 0$, and $f \in S[t]$. Let \bar{f} denote its image in $\bar{S}[t]$, where $\bar{S} = S/N$. Suppose given a factorization $\bar{f} = \bar{g}\bar{h}$ in $\bar{S}[t]$, such that \bar{g} is monic, and that \bar{g}, \bar{h} generate the unit ideal in $\bar{S}[t]$, i.e., $\bar{p}\bar{g} + \bar{q}\bar{h} = 1$ for some $\bar{p}, \bar{q} \in \bar{S}[t]$. There are unique polynomials g, h in $S[t]$ such that $f = gh$, g is monic, $\bar{g} = \bar{g}$, and $\bar{h} = \bar{h}$. Moreover, g, h generate the unit ideal in $S[t]$.*

Proof. This is done by Newton's method. We begin by choosing arbitrary liftings of g, h to $S[t]$. Then $f \equiv gh \pmod{N}$, so we may write $f + w = gh$, where w is a polynomial in $N[t]$, i.e., all of its coefficients are in N . We are free to adjust g, h by polynomials in $N[t]$, say to $g_1 = g + u$, and $h_1 = h + v$. Then because $N^2 = 0$, $g_1 h_1 = gh + gv + uh = f + w + gv + hu$. To factor f , we must solve the equation $gv + hu + w = 0$. Now since $N^2 = 0$, $N[t]$ is an $\bar{S}[t]$ module, i.e., $S[t]$ acts on $N[t]$ through $\bar{S}[t]$. So we may as well write $\tilde{g}v + \tilde{h}u + w = 0$, and this equation can be solved because \tilde{g}, \tilde{h} generate the unit ideal.

This procedure may not yield a monic polynomial g . To show that we can lift the solution while keeping g monic, we start with any lifting. The leading coefficient of g , say $b \in S$, has the property that $b \equiv 1 \pmod{N}$. Since $N^2 = 0$, b is a unit. So we may change g to $b^{-1}g$ and h to bh , and we are done.

The fact that g, h generate the unit ideal follows from the next lemma. Since we do not need the uniqueness here, we leave the proof of uniqueness as an exercise. \square

Lemma IV.5.3. *Let R be a local ring with residue field k , and let $g, h \in R[t]$ be polynomials, with g monic. If \bar{g}, \bar{h} generate the unit ideal in $k[t]$, then g, h generate the unit ideal in $R[t]$.*

Proof. because g is monic, $R[t]/(g)$ is a finite R -module, and so is $S = R[t]/(g, h)$. By the Nakayama lemma, $S \otimes k = k[t]/(\bar{g}, \bar{h}) = 0$ implies that $S = 0$. \square

We remark that there is a system of equations which expresses the fact that f is a product gh , and that g, h generate the unit ideal, say $pg+qh = 1$, where g, h, p, q are polynomials with undetermined coefficients, g being monic. The degrees of g and h have to add up to $\deg f$. The degrees of p and q are fixed and arbitrary. Lemma IV.5.2 shows that this is a smooth system of equations.

Exercise: Write down the system of equations.

Corollary IV.5.4. *Let f be a primitive and irreducible polynomial in $R[t]$ such that \bar{f} is not a constant. Then the leading coefficient of f is a unit.*

Proof. We write $\bar{f} = \tilde{g}\tilde{h}$, where $\tilde{h} \in k$ is the leading coefficient of \bar{f} and \tilde{g} is monic. By Hensel's Lemma, f factors accordingly: $f = gh$, where g is monic. Since \tilde{g} is not constant, neither is g . Then since f is irreducible, h is a constant, and since $\tilde{h} \neq 0$, h is a unit. \square

IV.6. Maximal orders over complete discrete valuation rings

In this section, R denotes a complete discrete valuation ring with field of fractions K , and D is a division ring with center K and finite over K .

Theorem IV.6.1. *Let R be a complete discrete valuation ring, and let D be a central division ring over its field of fractions K .*

(i) *There is a unique maximal order A in D , and it is the set*

$$A = \{a \in D \mid \det(a) \in R\}.$$

(ii) *Let A be the maximal order described in (i). Every maximal order in $M_r(D)$ is isomorphic to $M_r(A)$.*

Proof of Theorem IV.6.1(i). We already know that the determinant of any element of an order is in R (2.2). So it is clear that A , as defined above, contains every order, and the only thing to be proved is that A , as defined above, is an order.

Lemma IV.6.2. *If $a \in D$ and if $\det(a) \in R$, then the characteristic polynomial of a is in $R[t]$, and a is integral over R .*

Proof. Because D is a division ring, $R[a]$ is a commutative domain. Therefore the minimal monic polynomial $f(t)$ for a over K is irreducible. The minimal polynomial and the characteristic polynomial $p(t)$ have the same roots, so since f is irreducible, p is a power of f . By hypothesis, the constant coefficient $\det(a)$ of $p(t)$ is in R . Therefore the constant coefficient of $f(t)$ is in R too. Let f_0 denote the associated primitive polynomial in $R[t]$, which is obtained from f by multiplying by a power of a generator x for the maximal ideal \mathfrak{m} of R to clear denominators.

We claim that the polynomial \bar{f}_0 obtained by reducing f_0 modulo \mathfrak{m} is not constant. Since f_0 is primitive, at least one of its coefficients is a unit of R , and the degree of \bar{f}_0 is the largest degree

having a unit coefficient. If f does not have coefficients in R , then multiplication by x is necessary, and because the constant term of f is already in R , the constant term of f_0 will be divisible by x . So it is not a unit, and \bar{f}_0 has positive degree in this case. Else, if f itself has coefficients in R , then $f_0 = f$, and \bar{f} has degree n because f is monic.

Corollary IV.5.4 shows that f_0 is monic, so $f_0 = f$, and p , being a power of f , has coefficients in R . \square

Lemma IV.6.3. *The set A is an R -algebra.*

Proof. It is clear that A is closed under multiplication. To see that $R \subset A$, it is enough to note that, if $[D : K] = n^2$, then

$$(IV.6.4) \quad \det(c) = c^n$$

for any element $c \in K$, in particular for $c \in R$.

It is less clear that A is closed under addition, but there is an amazing trick to show this. Let $a, b \in A$, and say that the order of zero of $\det(a)$ is not greater than that of $\det(b)$. Then $u = a^{-1}b \in A$. So by the previous lemma, u is integral over R , and every element of the commutative ring $R[u]$ is integral, hence is in A . In particular, $1 + u$ is in A . Then $a(1 + u) = a + b$ is in A too, as was to be shown. \square

Lemma IV.6.5. *The set A is an R -order in D .*

Proof. It is clear from the definition of A that $A \otimes K = D$. So taking into account the previous lemma, what remains to be proved is that A is a finite R -module. This uses the trace pairing, and is similar to the proof of IV.4.1. Let $a, b \in A$. Then $ab \in A$, and by Lemma IV.6.2, $\langle a, b \rangle = \text{trace}(ab) \in R$. We may choose an R -lattice L in D which is a submodule of A . The trace pairing defines a map $L \times A \rightarrow R$, hence an injective map $A \subset L^*$. Since L^* is a finite R -module, so is A . \square

The next proposition lists some of the nice properties of the maximal order IV.6.1. The proofs are all easy, and we omit them.

Proposition IV.6.6. *Let A be the order IV.6.1, and let $\pi \in A$ be an element such that $\det(\pi)$ is not a unit of R , but $v(\det(\pi))$ is minimal among such elements of A .*

(i) *Every nonzero element of D has the form $\pi^k u$, where k is an integer and u is an invertible element of A .*

(ii) *π is a normal element of A , i.e., $\pi A = A\pi$. The two-sided ideal $J = \pi A$ of A is the Jacobson radical of A , and is a maximal right ideal and a maximal left ideal. $\Delta := A/J$ is a division ring.*

(iii) *The nonzero ideals of A are the powers of J . Every right or left ideal is a two-sided ideal.* \square

Proposition IV.6.7. *Every finite, torsion-free A -module M_A is free.*

Proof. Since Δ is a division ring, $\bar{M} := M \otimes_A \Delta$ is a free Δ -module, and we may lift a basis $\{\bar{m}_i\}$ of \bar{M} to a set $\{m_i\}$ of elements of M . This set generates M by the Nakayama Lemma, and it is a basis because M is torsion-free: If $\sum m_i a_i = 0$ is a nontrivial relation, then because M is torsion-free, we may cancel a power of π so that some a_ν is not in J . Because the set $\{\bar{m}_i\}$ is a basis, this implies that $a_\nu = 0$, which is a contradiction. \square

Proposition IV.6.8. *Let v_K denote the \mathfrak{m} -adic valuation of K . This valuation defines a valuation v_D on D , by the rule*

$$v_D(x) = v_K(\det(x)).$$

So for $x \in A$, $v_D(x) = k$ if $x \in J^k$ and $x \notin J^{k+1}$.

Proof. The defining properties of a valuation are

$$v_D(xy) = v_D(x) + v_D(y), \text{ and}$$

$$v_D(x + y) \geq \min\{v_D(x), v_D(y)\}.$$

The first of these is clear. For the second one, we use the trick of the proof of IV.6.3. Say that $v_D(x) \leq v_D(y)$. Then $v_D(x^{-1}y) \geq 0$, so $x^{-1}y \in A$. Then $1 + x^{-1}y \in A$ too, so $v_D(1 + x^{-1}y) \geq 0$. Multiplying by x , $v_D(x + y) \geq v_D(x)$. \square

Note that for $r \in R$,

$$(IV.6.9) \quad v_D(R) = n v_K(R)$$

People often normalize the valuation v_D differently. One can divide by n so that v_D agrees with v_R on R , or else one can set $v_D(\pi) = 1$, so that all integer values are taken on. The first has the disadvantage that the values are not necessarily integers, and the second complicates the definition of the valuation. But these are minor points.

Let R be the integral closure of R in a separable extension K of k , and let \mathfrak{m} be its maximal ideal. Then $k = R/\mathfrak{m}$ is an extension field of k . The extension K/k is called *unramified* if $\mathfrak{m} = \mathfrak{m}R$ and k/k is separable. In that case $k = R \otimes k$. We may say that k is obtained from K by reduction modulo \mathfrak{m} .

Proposition IV.6.10. *Let R be a complete discrete valuation ring with perfect residue field k , and let D be a central simple K -algebra which is a division ring.*

(i) Every finite field extension \bar{L} of k which is obtained as above from an unramified extension L of K . If $\bar{L} \subset \Delta$, then L can be realized as a subfield of D .

(ii) There is a maximal commutative subfield L of D which is an unramified extension of K .

Proof. (i) Let $\bar{L} \subset \Delta$ be a separable commutative field extension of k . We may choose a primitive element, and write $\bar{L} = k[\bar{a}] \approx k[t]/(\bar{f})$, where \bar{f} is an irreducible monic polynomial in $k[t]$ and where $f, df/dt$ are relatively prime. We lift \bar{f} arbitrarily to a monic polynomial $f \in R[t]$. Taking into account Lemma IV.5.3, it suffices to show that the root \bar{a} of \bar{f} can be lifted to a root a of f in A . Because $A = \varprojlim A/J^n$, it suffices to lift the root step by step to A/J^n for every n . This is done by Newton's method, using the fact that $\bar{f}, d\bar{f}/dt$ are relatively prime.

For the proof of (ii), we use induction on $[D : K]$, together with the next lemma.

Lemma IV.6.11. *Suppose that $D \neq K$. To prove (ii), It suffices to find any unramified field extension L of K contained in D , and which is not equal to K .*

Proof. We choose L to be maximal. Let L be the centralizer of L in D . So L is a division ring, and L is contained in the center of L . By (A.1), the centralizer of L is L , and since $L \subset L$, L

is the center of L . By induction, there is an unramified extension L_1 of L which is a maximal commutative subfield of L . By maximality of L , $L = L_1$, and this is possible only if $L = L$ (CSA, 4.2). \square

It remains to find the unramified extension as in the Lemma.

Case 1: $\Delta > k$. We choose an arbitrary element $\bar{a} \in \Delta$ which is not in k . Because k is perfect, \bar{a} is separable over k . We may write $k = k[\bar{a}] \approx k[t]/(\bar{f})$, where $\bar{f}, d\bar{f}/dt$ generate the unit ideal in $k[t]$. Let $f \in R[t]$ be a monic polynomial whose residue modulo \mathfrak{m}_R is \bar{f} . Lemma IV.5.3 shows that $f, df/dt$ generate the unit ideal in $R[t]$, hence that $R = R[t]/(f)$ is a finite etale R -algebra. Because \bar{f} is irreducible, so is f . So $K = K[t]/(f)$ is the required unramified extension of K .

Case 2: $\Delta = k$. In this case we show that $A = R$, hence that $D = K$. Let $\pi \in A$ be an element of minimal positive valuation, and $J = A\pi$. Let $a = a_0 \in A$ be arbitrary. Because $\Delta = k$, we can find an element $r_0 \in R$ such that $a_0 \equiv r_0$ (modulo $A\pi$), or $a = r_0 + a_1\pi$. The same reasoning, applied to a_1 , yields $a = r_0 + r_1\pi + a_2\pi^2$, and continuing, we obtain $a = r_0 + r_1\pi + r_2\pi^2 + \dots$, with $r_i \in R$. On the other hand, π is integral over R , so it satisfies a monic polynomial relation with coefficients in R . If the degree of this relation is d , then the process can be terminated at degree d , to show that a can be written as a polynomial in π with coefficients in R . This shows that $A \subset R[\pi]$, so A and A_K are commutative. Since K is the center of A_K , $A_K = K$. \square

Proof of Theorem IV.6.1(ii). We note that D^r is a $(D, M_r(D))$ -bimodule. If B is an order in $M_r(D)$, then $A^o \otimes B$ is an order in $D^o \otimes M_r(D)$. Choose an $A^o \otimes B$ -lattice L in D^r . By xxx, L is a free left A -module. So this lattice defines an operation of B on A_A^r , hence it identifies B as a subalgebra of $M_r(A) = \text{End } A_A^r$. Since there exists a maximal order, $M_r(A)$ must be maximal. \square

Corollary IV.6.12. (i) Let L be a maximal subfield of D and \bar{L} a maximal subfield of k . Then $[L : K] = [\bar{L} : k]$.

(ii) Let $\pi \in A$ be an element of minimal positive valuation. Say that $v_D(\pi) = f$ and $[D : K] = n^2$. Also, let k be the center of Δ . Then f divides n , say $ef = n$. Moreover $[k : k] = e$ and $[\Delta : k] = f^2$.

Proof. (ii) Let p be a generator for the maximal ideal \mathfrak{m} , and let $e = n/f$. Then (IV.6.9) $v_D(p) = n$, hence $p = \pi^e u$, where u is invertible in A . This shows that $e = n/f$ is an integer and that $pA = J^e$. Because A is a free R -module, $\dim_k(A/pA) = \dim_K D = n^2$.

We have a chain of ideals $A \supset J \supset J^2 \supset \dots \supset J^e = pA$, and J^i/J^{i+1} is a free right $A/J = \Delta$ -module of rank 1, generated by π^i . So $n^2 = \dim_k(A/pA) = e[\Delta : k]$, and $[\Delta : k] = ef^2$.

Let $e = [k : k]$, and $f^2 = [\Delta : k]$, so that $[\Delta : k] = ef^2$. Then a maximal subfield of Δ has degree f over k , hence degree ef over k , and (i) shows that $ef = n = ef$. This implies that $e = e$ and $f = f$. \square

More can be said. We state the next theorem without proof.

Theorem IV.6.13. With the above notation,

(i) The field extension k of k is a Galois extension with cyclic Galois group.

(ii) Let \widehat{R} be the integral closure of R in an unramified extension K of K which splits D . Then $A := A \otimes R$ is isomorphic to a standard hereditary order in the matrix algebra $M_n(K)$, as described below.

The description of the order: Let $M = M_r(R)$, and let p denote a generator for the maximal ideal of R and of \widehat{R} . Then A is made up of a $e \times e$ array of $f \times f$ blocks. The blocks below the diagonal are pM , and the rest are M . The Jacobson radical $J = J \otimes R$ has all blocks on the diagonal equal to pM as well, as illustrated below for the case $e = 3$.

$$(IV.6.14) \quad A = \begin{pmatrix} M & M & M \\ pM & M & M \\ pM & pM & M \end{pmatrix}, \quad J = \begin{pmatrix} pM & M & M \\ pM & pM & M \\ pM & pM & pM \end{pmatrix}.$$

IV.7. Recovering data from the completion

We want to use the results of the previous section to describe maximal orders over a Dedekind domain R which is not complete. If we are given a central simple K -algebra A_K , then we can as first approximation, choose any R -order, maximal or not, and ask where it is Azumaya. As we have remarked, being Azumaya is an open condition, so in the case of a Dedekind domain, it will be the complement of a finite set of points. The problem is to describe the order at these “bad” points. We can treat each of these points separately, and thereby reduce to the case that a maximal order has been described except at a single point p . (The uses of the symbols p and t have changed.) There is a general principle which tells us that the order, or any other structure on a finite R -module, is determined by its restriction to the complement of the point on $\text{Spec } R$, and by its completion at p . We describe this principle here.

Let R be a Dedekind domain which is not a field, and let \mathfrak{m} be a maximal ideal of R . Let's call p the corresponding point of $\text{Spec } R$. For simplicity, we'll assume that \mathfrak{m} is a principal ideal of R , generated by an element t . Let $\widehat{R} = R[t^{-1}]$, and let \widehat{R} denote the completion of R at \mathfrak{m} . So \widehat{R} is a complete discrete valuation ring, and the valuations of R and of \widehat{R} at p agree on the field of fractions K of R . We pose the problem of recovering information about R from \widehat{R} and R . We know that \widehat{R} and R are flat over R , and by (CSA, 5.2) that $\widehat{R} \oplus R$ is a faithfully flat R -algebra.

We consider the category \mathcal{C} of pairs (M, \widehat{M}) , where M is a finite R -module and \widehat{M} is a finite \widehat{R} -module, such that $\widehat{M} \otimes_{\widehat{R}} K = \widehat{M}$ and $M \otimes_{R'} K$ are isomorphic. We suppress the isomorphism in our notation in order to avoid clutter, but it is to be understood that the isomorphism is assigned by the pair. A morphism $(M, \widehat{M}) \rightarrow (N, \widehat{N})$ in \mathcal{C} is a pair of maps $M \rightarrow N$ and $\widehat{M} \rightarrow \widehat{N}$ such that the induced maps $\widehat{M} \rightarrow \widehat{N}$ are equal.

We have a functor $\Phi : (\text{mod } R) \rightarrow \mathcal{C}$, which carries a finite R -module M to the pair $(M \otimes R, M \otimes \widehat{R})$. The isomorphism is the canonical isomorphism

$$M \otimes R \otimes_{R'} K \approx M \otimes K \approx M \otimes R \otimes_{\widehat{R}} K.$$

If M is a finite R -module, we will write $M = M \otimes R$, $\widehat{M} = M \otimes \widehat{R}$, and $\widehat{M} = M \otimes K$. However, we also use the notation (M, \widehat{M}) for an element of \mathcal{C} which, a priori, does not come from a finite R -module M .

Theorem IV.7.1. (i) Let M be a finite R -module. The sequence

$$(IV.7.2) \quad 0 \longrightarrow M \longrightarrow M \oplus \widehat{M} \xrightarrow{(+,-)} \widehat{M}$$

is exact, where $(+, -)$ denotes the difference of the canonical maps. In other words, elements of M correspond bijectively to pairs of elements (m, m) whose images in \widehat{M} are equal.

(ii) The functor Φ is an equivalence of categories.

Proof. The analogues of this theorem are true in higher dimensions, but since we are stating the result for Dedekind domains, we may as well use their special properties.

(i) The sequence IV.7.2 is compatible with direct sums. So since every finite R -module is a direct sum of a finite length module and a projective module, we may treat the cases of a finite length module and of a projective module separately.

The map $M \longrightarrow M \oplus \widehat{M}$ is injective because $R \oplus R$ is faithfully flat (CSA, xxx). Also, the composition of the two maps in the sequence is zero.

If M has finite length and is supported at p , then $M = 0$, $\widehat{M} = 0$, and $M \approx \widehat{M}$. If M has finite length and its support is away from p , then $M \approx M$, and $\widehat{M} = \widehat{M} = 0$. The sequence is exact in both cases.

Suppose that M is projective. Then M is a summand of a free module, and it suffices to verify the assertion for a free module. This in turn reduces to the case that $M = R$. The assertion for R can be restated in this way: If the image in K of an element $x \in R$ lies in R , then $x \in R$. To say that $x \in R$ is the same as saying that its valuation $v_p(x)$ is not negative. The valuation on the completion \widehat{R} is the same as the valuation on R at p . So this is true if $x \in R$.

Lemma IV.7.3. Let R be a commutative noetherian ring, and let M, N be finite R -modules.

(i) $\text{Hom}_R(M, N)$ is a finite R -module. More generally, for all q , $\text{Ext}_R^q(M, N)$ is a finite R -module.

(ii) Let S be a flat R -algebra, and let $M_S = M \otimes S$. The canonical maps $\text{Ext}_R^q(M, N) \otimes S \longrightarrow \text{Ext}_S^q(M_S, N_S)$ are bijective. In particular, $\text{Hom}_R(M, N) \otimes S \approx \text{Hom}_S(M_S, N_S)$ is bijective. \square

Lemma IV.7.6. (i) The functor Φ is fully faithful.

(ii) The category of R -modules of finite length supported at p is equivalent to the category of finite length R -modules, and also to the full subcategory of \mathcal{C} of objects of the form $(0, F)$.

(iii) Let F be an R -module of finite length supported at p . Then for any finite R -module N , $\text{Ext}_R^1(F, N) \approx \text{Ext}_{\widehat{R}}^1(F, N) \approx \text{Ext}_{\mathcal{C}}^1((0, F), (N, N))$.

Proof. (i) Let $H = \text{Hom}_R(M, N)$. By Lemma xxx, $\text{Hom}_{R'}(M, N) \approx H$, $\text{Hom}_{\widehat{R}}(\widehat{M}, N) \approx H$, and $\text{Hom}_{\widehat{K}}(\widehat{M}, N) \approx H$. By part (i) of the theorem, the sequence

$$0 \longrightarrow H \longrightarrow H \oplus H \longrightarrow H$$

is exact. This identifies H with $\text{Hom}_{\mathcal{C}}((M, \widehat{M}), (N, N))$.

(ii) This follows because a finite length R -module supported at p is a module over $R/t^n R$ for some n , and $R/t^n R \approx R/t^n R$.

(iii) We choose a resolution of F as R -module, say

$$(IV.7.7) \quad 0 \longrightarrow P \longrightarrow R^k \longrightarrow F \longrightarrow 0.$$

Then because R is a Dedekind domain, P is projective. So the exact sequence

$$(IV.7.8) \quad 0 \longrightarrow \text{Hom}_R(R^k, N) \longrightarrow \text{Hom}_R(P, N) \longrightarrow \text{Ext}_R^1(F, N) \longrightarrow 0$$

computes Ext^1 . Both sequences remain exact when tensored with R , and P is a projective R -module. The first two terms of the second sequence, tensored with R , are identified by Lemma xxx. This shows that $\text{Ext}_R^1(F, N) \otimes R \approx \text{Ext}_{\widehat{R}}^1(F, N)$. The isomorphism $\text{Ext}_{\widehat{R}}^1(F, N) \approx \text{Ext}_{\mathcal{C}}^1((0, F), (N, N))$ is verified by inspection. \square

Going back to the proof of the second part of the theorem, let (M, \widehat{M}) be an object of \mathcal{C} . There is a finitely generated R -submodule N of M such that $N \approx M$, and then $N \approx \widehat{M}$ too. If we replace N by $t^k N$ for suitably large k , N will be t -torsion free, and the map $N \longrightarrow \widehat{M}$ will be induced by an injection $N \subset \widehat{M}$. Then setting $F = M/N$, \widehat{M} becomes an R -module extension of N by F . By Lemma xxx(iii), this extension is induced by an R -module extension of M by F , and one sees that $\Phi(M) \approx (M, \widehat{M})$. \square

IV.8. Addendum to the handout on Central simple algebras

The following is one of Brauer's theorems, and it should have been included in (CSA, Section 4). At the time, I thought we wouldn't need this assertion, so I suppressed it. You can find a proof in any reference which treats central simple algebras.

Theorem A.1. *Let A be a central simple K -algebra, and let B be a simple subalgebra. Let B be the centralizer of B in A , and B the centralizer of B . Then (i) B is simple. (ii) $B = B$, and (iii) $[A : K] = [B : K][B : K]$.*

This discussion should have been in (CSA, Section 12). The Brauer group of a ring R is defined as the group of equivalence classes $[A]$ of Azumaya algebras A over R , the operation being tensor product. The only point that needs discussion is the definition of the equivalence relation. The clue is given by (CSA III.12.1(iv)), which tells us that $A^o \otimes_R A$ is isomorphic to $E := \text{End}_R A$. If we expect $[A^o]$ to invert $[A]$, as it does in the case of a field, then we must make $[E]$ equal to the trivial class $[R]$. It is therefore natural to define the equivalence class of R to be the set of algebras isomorphic to $\text{End}_R P$, where P is any finite, locally free (or projective) R -module. Then we must say that A and B are equivalent if there are finite projective modules P, Q such that $A \otimes \text{End}_R P$ and $B \otimes \text{End}_R Q$ are isomorphic. This works, and it defines a group.

Exercise: Show that the tensor product of Azumaya algebras is Azumaya, and verify that $Br(R)$ is an abelian group.

V. IRREDUCIBLE REPRESENTATIONS

V.1. Definition

Let A be an algebra over a field k , and let K be a field extension of k . We'll be talking exclusively about finite-dimensional representations of A , so when we say representation, the hypothesis that it is finite-dimensional is there.

An n -dimensional *matrix representation* over K is a k -homomorphism $\rho : A \rightarrow M_n(K)$ to the matrix algebra over K . Two matrix representations ρ, ρ' are *equivalent* if one is obtained from the other by an inner automorphism of the matrix algebra, i.e., $\rho' = u\rho u^{-1}$ for some invertible matrix $u \in M_n(K)$. The representation is *irreducible* if the image $\rho(A)$ generates the matrix algebra as K -algebra, or if the map $A \otimes K \rightarrow M_n(K)$ is surjective. We often study the case that $k = K$. In that case a representation $\rho : A \rightarrow M_n(k)$ is irreducible if and only if it is a surjective homomorphism.

Lemma V.1.1. *Let $\rho : A \rightarrow M_n(K)$ be a matrix representation, and let L be a field extension of K . Then ρ is irreducible if and only if the representation $A \rightarrow M_n(L)$ obtained by composition with the inclusion map $M_n(K) \rightarrow M_n(L)$ is irreducible. \square*

Suppose that A is a finitely generated algebra over k , say $A = k\langle x_1, \dots, x_r \rangle / I$, where I is an ideal of the free ring $k\langle x \rangle$. Then the representation ρ is determined when we assign the images $\rho(x_i)$, which have to satisfy the relations imposed by the ideal I .

A second definition of an n -dimensional *representation* is as a module V over the tensor product $A \otimes_k K$, which is n -dimensional as K -module. Since K and A commute in the tensor product, the module structure defines a homomorphism $A \rightarrow \text{End}_K V$, and by the choice of a basis, $\text{End}_K V$ becomes a matrix algebra. Conversely, a homomorphism $A \rightarrow M_n(K)$ defines an $A \otimes K$ -module structure on $V = K^n$. Equivalence of representations carries over to isomorphism of modules. When $k = K$, V is just an A -module which has dimension n over k .

Proposition V.1.2. *Let \overline{K} denote the algebraic closure of K . Via the above correspondence, irreducible matrix representations correspond to finite dimensional $A \otimes K$ -modules V such that $V \otimes \overline{K}$ is a simple $A \otimes \overline{K}$ -module.*

Proof. We'll show that if ρ is not irreducible, then $V \otimes_K \overline{K}$ is not simple. Suppose that the map $A \otimes K \rightarrow M_n(K)$ induced by a choice of basis for V_K is not surjective. Then $A \otimes \overline{K} \rightarrow M_n(\overline{K})$ is not surjective either. We replace K by \overline{K} , thereby reducing to the case that K is algebraically closed. Let $B \subset M_n(K)$ denote the image of $A \otimes K$, and let J denote the Jacobson radical of B . Since V is simple and $V/VJ \neq 0$ by Nakayama, it follows that $VJ = 0$ and that V is a B/J -module. By Wedderburn, B/J is a sum of matrix algebras, and if $B \neq M_n(K)$, the ranks of these matrix algebras are smaller than n . Then the simple B/J -modules have dimensions $< n$ too, and it follows that V is not simple. \square

There is a third way to look at irreducible representations, namely in terms of the kernel of the homomorphism $A \rightarrow M_n(K)$. Suppose that $k = K$, and let ρ be an irreducible representation. The kernel P of the surjective map $A \rightarrow M_n(k)$ is a maximal ideal because $M_n(k)$ is a simple ring, and so A/P is isomorphic to the matrix algebra. By the Skolem-Noether theorem, the isomorphism is determined up to conjugation. So P determines the equivalence class of the representation. This

shows that equivalence classes of irreducible representations over k correspond to certain maximal ideals P in A : those such that A/P is a matrix algebra.

The kernel characterizes the equivalence class of a representation when K is any field, but to use it we must understand when matrix representations over different fields should be called equivalent. It is better to leave this point for later, when some theory has been developed (see Theorem 8.7).

V.2. Examples

We'll work out the irreducible representations of three rings in this section. We assume that the ground field k is algebraically closed and of characteristic zero, and we'll study representations by k -vector spaces.

Example V.2.1 *The enveloping algebra of \mathfrak{sl}_2 .* The Lie algebra \mathfrak{sl}_2 of traceless matrices is spanned by

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

and we denote these matrices by x, y, h respectively. They satisfy the relations

$$(V.2.2) \quad [y, x] = h, \quad [x, h] = 2x, \quad [y, h] = -2y,$$

and the enveloping algebra $A = U(\mathfrak{sl}_2)$, which is the associative algebra generated by x, y, h , is defined by the same relations:

$$(V.2.3) \quad xy - yx = h, \quad hx - xh = 2x, \quad hy - yh = -2y.$$

Exercise: Use the Diamond Lemma to show that the ordered monomials $h^i x^j y^k$ form a basis of U .

Let V be an irreducible representation of A , and let $v \in V$ be an eigenvector for the operation of h on V , say $vh = \lambda v$. Then

$$v_x h = v[x, h] + v_h x = 2vx + \lambda vx = (\lambda + 2)vx.$$

So vx , if not zero, is an eigenvector for h with eigenvalue $\lambda + 2$. Similarly, $vyh = (\lambda - 2)vy$.

Choose an eigenvector $v = v_0$ for h with an eigenvalue $\lambda = r$, and so that $r + 2$ is not an eigenvalue, hence so that $vx = 0$. For $n \geq 0$, set $v_n = vy^n$. Then

$$v_n h = (r - 2n)v_n,$$

$$(V.2.4) \quad v_n y = v_{n+1},$$

$$v_n x = \alpha_n v_{n-1},$$

where $\alpha_n = nr - n(n - 1)$. The last formula is proved by induction, using the initial condition $\alpha_0 = 0$:

$$v_{n+1} x = v_n y x = v_n [y, x] + v_n x y = v_n h + \alpha_n v_{n-1} y = (r - 2n + \alpha_n) v_n.$$

So $\alpha_{n+1} = \alpha_n + r - 2n$, from which the formula follows.

These formulas show that $v_{r+1} x = 0$, but that v_n are independent if $n \leq r$. This allows us to mimic our representation by using the same formulas (V.2.4) for the vector space with basis $\{v_0, \dots, v_r\}$, but setting $v_{r+1} = 0$. Our representation maps to this one, so since V is irreducible, the map must be an isomorphism. Working this out yields the following:

Proposition V.2.5. *There is, up to isomorphism, a unique irreducible representation of $U(\mathfrak{sl}_2)$ of dimension $r + 1$ for every $r \geq 0$. \square*

Example V.2.6. *Varying the q -plane.* Let A denote the ring generated by elements x, y, t, t^{-1} , with relations

$$(V.2.7) \quad yx = txy \quad , \quad tx = xt \quad , \quad ty = yt \quad .$$

The 1-dimensional representations of A are the homomorphisms $A \rightarrow k$. Since k is commutative, we can factor such a homomorphism through the ring \tilde{A} obtained by forcing commutativity. Since t is already central, the only relation we need to add is $yx = xy$. Combining with the first relation (V.2.7), this yields the defining equation for \tilde{A} , namely $(t - 1)xy = 0$. So the one dimensional representations of A are the k -valued points of the commutative ring \tilde{A} , which are the points on the three planes $t = 1$, $x = 0$, and $y = 0$ in (x, y, t) -space.

We now describe the higher dimensional irreducible representations. Since t is in the center of A and since the center of $M_n(k)$ is the set of scalar matrices, the matrix representing t must be a scalar. So we can begin by assigning a value for q in k . Setting $t = q$ yields the quantum polynomial ring $k\langle x, y \rangle / (yx - qxy)$. We may as well work with the quantum polynomial ring $A_q = k\langle x, y \rangle / (yx - qxy)$, where q denotes any nonzero element of k . Elements of A_q can be represented uniquely by ordered polynomials $\sum a_{ij} x^i y^j$.

Lemma V.2.8. *Suppose that $q \in k$ is not a root of unity. Then every nonzero ideal of A_q contains a monomial $x^i y^j$.*

Proof. To show this, we start with any nonzero element $f(x, y) = \sum a_{ij} x^i y^j$ in an ideal I . We note that $(x^i y^j)x = q^j x(x^i y^j)$. So $q^r x(x^i y^j) - (x^i y^j)x = 0$ if and only if $j = r$. Replacing f by $q^r x f - f x$ kills the term $x^i y^j$ in f if $j = r$, and replaces it by the nonzero term $(q^r - q^j)x^{i+1} y^j$ if $j \neq r$. In this way we can reduce number of powers of y which appear in f to 1, without killing f . Similarly, operations of the form $q^r f y - y f$ reduce the number of powers of x to 1, and we are left with a polynomial of the form $a x^i y^j$ with $a \neq 0$. \square

Lemma V.2.9. *If q is not a root of unity, then the only irreducible representations of A_q are one-dimensional.*

Proof. Let $\rho : A_q \rightarrow M_n(k)$ be a surjective homomorphism. Because $M_n(k)$ is a prime ring, the kernel P of ρ is a prime ideal. It is not zero because A_q is infinite-dimensional. By the previous lemma, P contains $x^i y^j$ for some i, j . Then because the monomials span A_q , $x A_q x^{i-1} y^j \subset P$ too, and by induction, x or y is in P . Therefore A_q/P is commutative, and $n = 1$. \square

The quantum polynomial ring A_q has a \mathbb{Z}^2 -grading, and a corresponding two-parameter family of automorphisms, defined by $x \rightarrow \lambda x$ and $y \rightarrow \mu y$, $\lambda, \mu \neq 0$.

Lemma V.2.10. *Suppose that q is a primitive n th root of unity.*

(i) *The map sending x to the $n \times n$ diagonal matrix whose diagonal entries are the powers of ζ in order, and y to the $n \times n$ cyclic permutation matrix, defines an irreducible representation of A_q . (See (CSA, 2.5), and set $a = 1$.)*

(ii) *Every irreducible representation of A_q has dimension 1 or n , and the representations of dimension n are obtained from the representation (i) by the automorphisms $x \mapsto \lambda x$ and $y \mapsto \mu y$.*

Exercise V.2.11: Prove Lemma V.2.10.

Example V.2.12. *The Weyl algebra.* This algebra is $A = k\langle x, y \rangle / (yx = xy + 1)$. The equation $yx = xy + 1$ has no solution in $n \times n$ -matrices, because $\text{trace}(xy) = \text{trace}(yx)$ and $\text{trace}(1) = n$. Therefore this algebra has no finite-dimensional representations at all.

The Weyl algebra exhibits the limitations inherent in the program of studying algebras by means of their representation theory. What is worse, there are no general methods for handling representations of all dimensions simultaneously. The methods available analyze the representations whose dimensions are bounded by a fixed integer n . This is a serious problem. However, for representations of dimensions $\leq n$ there is a big theory, and we will look at two aspects of it. The first one is abstract, and it leads to the concept of *polynomial identity rings*. The second one is the direct approach of describing a matrix representation ρ of a finitely presented ring $A = k\langle x_1, \dots, x_r \rangle / I$ as the point in n^2r -dimensional space whose coordinates are the entries of the matrices $\rho(x_i)$. This approach leads back to classical invariant theory.

V.3. The standard identities

The commutator $[x, y] = xy - yx$ has analogues for more variables, called *generalized commutators*

$$(V.3.1) \quad S_n(x_1, \dots, x_n) = \sum (-1)^\sigma x_{\sigma_1} x_{\sigma_2} \cdots x_{\sigma_n},$$

where σ runs over the group of all permutations. Thus $S_2(x, y) = [x, y]$, and

$$S_3(x, y, z) = xyz + yzx + zxy - yxz - xzy - zyx.$$

The generalized commutators are multilinear and alternating polynomials in the variables, i.e., they are linear in each variable and zero if two variables are equal. A general multilinear polynomial in x_1, \dots, x_n has the form

$$(V.3.2) \quad p(x_1, \dots, x_n) = \sum c_\sigma x_{\sigma_1} x_{\sigma_2} \cdots x_{\sigma_n},$$

where the coefficients c_σ are elements of the ground field k .

Our first fundamental result is the following theorem of Amitsur and Levitski:

Theorem V.3.3. (*Amitsur-Levitski*) *Let R be a commutative ring, and let r be an integer.*

- (i) *If $r \geq 2n$, then $S_r(a_1, \dots, a_r) = 0$ for every set a_1, \dots, a_r of $n \times n$ matrices with entries in R .*
- (ii) *Let $p(x_1, \dots, x_r)$ be a nonzero multilinear polynomial. If $r < 2n$, there exist $n \times n$ matrices a_1, \dots, a_r such that $p(a_1, \dots, a_r) \neq 0$. In particular, $S_r(x_1, \dots, x_r)$ is not identically zero.*

The identity $S_{2n} \equiv 0$ is called the *standard identity* of $n \times n$ matrices. For instance $S_2 \equiv 0$ is the commutative law, which holds for 1×1 matrices but not for $n \times n$ matrices if $n > 1$.

Note that since $M_n(R)$ is a free R -module of rank n^2 , it satisfies the identity $S_{n^2+1} = 0$, simply because the generalized commutator is multilinear and skew symmetric. However, the precise bound provided by the Amitsur-Levitski theorem is useful.

Here is the reason that the theorem is important: Suppose that we wish to study the representations $A \longrightarrow M_n(k)$ of a k -algebra A . Let $I \subset A$ denote the ideal generated by all substitutions of elements

of A into S_{2n} , and let $\tilde{A} = A/I$. Part (i) of the theorem tells us that the identity $S_{2n} = 0$ is true in $M_d(k)$ if $d \leq n$. So every d -dimensional representation of A factors through \tilde{A} . Part (ii) tells us that S_{2n} is not true if $d > n$. So no irreducible d -dimensional representation $A \rightarrow M_d(k)$ factors through \tilde{A} . Killing I has the effect of keeping the representations of dimensions $\leq n$, while cutting out all irreducible representations of higher dimension. The ring \tilde{A} is an example of a *polynomial identity* ring, and we need to develop machinery to handle these rings.

Proof of Theorem V.3.3. (ii) We leave the reduction to the case $r = 2n - 1$ as an exercise. Suppose that $r = 2n - 1$. We rearrange the variables so that the monomial $x_1 \cdots x_{2n-1}$ has a nonzero coefficient c in p . Consider the string of $2n - 1$ matrices $e_{11}, e_{12}, e_{22}, e_{23}, \dots, e_{n-1n}, e_{nn}$. The product of the string in the order indicated is e_{1n} , while the product in any other order is zero. Substituting the above string into p yields $c e_{1n}$, which is not zero.

The proof that $n \times n$ matrices do satisfy the identity $S_{2n} = 0$ is much trickier. Since S_{2n} is linear in each of its variables, it would suffice to verify the identity when each x_ν is some matrix unit e_{ij} , but this direct approach is awkward. The following proof by Rosset is the best one available.

Extending a smaller matrix by zeros shows it is enough to prove the assertion in the case $r = 2n$. Next, since the entries of S_{2n} are polynomials in the matrix entries of the x_ν , we are trying to prove some identities among polynomials. Checking such identities can be done over the complex numbers. So it suffices to prove $S_{2n} \equiv 0$ when the x_ν are complex $n \times n$ matrices.

Let $A = M_n(\mathbb{C})$, let U be a complex vector space of dimension $2n$, say with basis $\{u_1, \dots, u_{2n}\}$, and let E denote the exterior algebra on U . Let $T = A \otimes E$ denote the tensor product algebra. Let x_ν be indeterminate elements of A , and set

$$(V.3.4) \quad \alpha = \sum x_\nu \otimes u_\nu.$$

Also, let $\beta = \alpha^2$. Collecting terms, we find

$$(V.3.5) \quad \alpha^r = \sum S_r(x_{i_1}, \dots, x_{i_r}) \otimes u_{i_1} \wedge \cdots \wedge u_{i_r},$$

the sum being over all multi-indices $i_1 < \cdots < i_r$. In particular,

$$(V.3.6) \quad \beta^n = \alpha^{2n} = S_{2n}(x_1, \dots, x_{2n}) \otimes u_1 \wedge \cdots \wedge u_{2n}.$$

It suffices to show $\beta^n = 0$. Now E is a graded algebra, and the part E_0 of even degree is commutative. The tensor product inherits the grading, and our element β is even. So we can work in the algebra $A \otimes E_0$, which is isomorphic to the matrix algebra $M_n(E_0)$ over the commutative ring E_0 .

Lemma V.3.7. *Let R be a commutative \mathbb{C} -algebra, and let $\beta \in M_n(R)$ be a matrix such that $\text{trace}(\beta^i) = 0$ for all $i = 1, \dots, n$. Then $\beta^n = 0$.*

Proof. In characteristic zero we can recover the coefficients of the characteristic polynomial of β from these traces. So the characteristic polynomial is t^n , and the Cayley-Hamilton theorem says that $\beta^n = 0$. \square

Now to prove the Amitsur-Levitski theorem, it suffices to verify that $\text{trace}(\beta^i) = 0$ for $1 \leq i \leq n$. We look at formula (V.3.5). Since trace is linear, it suffices to show that $S_{2i}(y_1, \dots, y_{2i})$ has trace zero for all $i \leq 2n$ and all matrices y_i .

Lemma V.3.8. For any $n \times n$ matrices y_1, \dots, y_{2i} , $S_{2i}(y_1, \dots, y_{2i})$ has trace zero.

Proof by example. Trace is preserved by cyclic permutations of the variables, and cyclic permutations of an even number of indices are odd. So we can group as follows:

$$S_4(x, y, z, w) = (xyzw - wxyz) + (-yxzw + wyxz) + \cdots .$$

□

V.4. Central polynomials

In about 1970, Formanek and Razmyslov independently found polynomials which, when evaluated on $n \times n$ matrices, always yield scalar matrices, but which aren't identically zero. Such polynomials are called *central polynomials*. Their discovery simplified the theory of PI rings greatly. We'll present Razmyslov's central polynomial here.

Let $N = n^2$, let $v = \{v_i\}$, $i = 1, \dots, N$ be a basis for the matrix algebra $M_n(k)$, and let $v^* = \{v_i^*\}$ denote the dual basis, with respect to the trace form $\langle x, y \rangle = \text{trace}(xy)$. Let I denote the identity matrix in $M_n(k)$.

Proposition V.4.1. For any matrix z , $\sum_i v_i z v_i^* = \text{trace}(z)I$.

Proof. We first check the formula for the case that v is the standard basis $e = \{e_{\mu\nu}\}$, taken in an arbitrary order. The dual basis is $e_{\mu\nu}^* = e_{\nu\mu}$. Then, writing $z = \sum z_{\mu\nu} e_{\mu\nu}$,

$$\sum e_{\mu\nu} z e_{\mu\nu}^* = \sum e_{\mu\nu} z e_{\nu\mu} = \sum e_{\mu\nu} z_{\nu\nu} e_{\nu\nu} e_{\nu\mu} = \sum z_{\nu\nu} e_{\mu\mu} = \text{trace}(z)I.$$

Next, for an arbitrary basis $v = \{v_i\}$, we write $v = Pe$ and $v^* = e^*Q$, for some $n^2 \times n^2$ matrices P, Q . Then $I_{n^2} = \langle v, v^* \rangle = P \langle e, e^* \rangle Q = PQ$. So $PQ = I$. Now

$$\sum v_i z v_i^* = \sum p_{ij} e_j z e_k^* q_{ki} = \sum q_{ki} p_{ij} e_j z e_k^* = \sum \delta_{kj} e_j z e_k^* = \sum e_j z e_j^* = \text{trace}(z)I,$$

as required. □

So to find a central polynomial, it is enough to express the dual basis x_i^* of a basis x_i as a polynomial $f_i(x, y)$ in x_i and some auxiliary variables y_i . Then $\sum x_i z f_i$ will be a central polynomial.

We need a multilinear and alternating polynomial in x_1, \dots, x_N which isn't identically zero when evaluated on $n \times n$ matrices. The *Capelli polynomial* is the **not true** universal such polynomial. It is obtained by interspersing $s + 1$ auxiliary variables y_0, \dots, y_N between the terms in the products $x_{\sigma_1} x_{\sigma_2} \cdots x_{\sigma_n}$. The auxiliary variables are needed so that the polynomial does not vanish identically on $n \times n$ matrices.

$$(V.4.2) \quad d = d(x, y) = \sum (-1)^\sigma y_0 x_{\sigma_1} y_1 x_{\sigma_2} y_2 \cdots x_{\sigma_s} y_N.$$

Let M denote the space of noncommutative polynomials in some variables x, y, \dots which are linear in the variable x . We define an operation τ_x on M as follows: Any monomial appearing in f has the form uxv , where u, v are monomials in the remaining variables. We set

$$(V.4.3) \quad (uxv)^{\tau_x} = vu.$$

This definition extends linearly to an operation τ_x on the space M .

Theorem V.4.4. *Let x_1, \dots, x_N be a basis of $M_n(k)$ and let y_0, \dots, y_N be arbitrary matrices in $M_n(k)$. Let $a = \text{trace}(d)$. The dual basis is*

$$x_i^* = a^{-1}d^{\tau_i},$$

where d is the Capelli polynomial and $\tau_i = \tau_{x_i}$.

Proof. Since $\text{trace}(uxv) = \text{trace}(xvu)$ for all matrices u, x, v and since trace is linear, it follows that for all $f \in M$,

$$(V.4.5) \quad \text{trace}(xf^{\tau_x}) = \text{trace}(f)$$

whenever matrices are substituted for the variables. Substituting another variable for x yields a trivial variant of this identity:

$$(V.4.6) \quad \text{trace}(wf^{\tau_x}) = \text{trace}(f|_{\{x=w\}}),$$

where $f|_{\{x=w\}}$ denotes the polynomial obtained by substituting w for x into f . Again, this formula is valid whenever matrices are substituted for the variables.

We apply these formulas to the Capelli polynomial $d = d(x, y)$. (V.4.5) shows that

$$(V.4.7) \quad \langle x_i, d^{\tau_i} \rangle = \text{trace}(x_i d^{\tau_i}) = \text{trace}(d) = a.$$

If $i \neq j$, then $d|_{\{x_j=x_i\}} = 0$ because d is alternating. Therefore V.4.6 shows that

$$\langle x_i, d^{\tau_j} \rangle = \text{trace}(x_i d^{\tau_j}) = 0$$

for all $i \neq j$. \square

Theorem V.4.8. *Let $x_1, \dots, x_N; y_0, \dots, y_N; z$ be variables. The polynomial*

$$c_n(x, y, z) = \sum x_i z d^{\tau_i}$$

is a homogeneous multilinear central polynomial for $n \times n$ matrices. More precisely, evaluating this polynomial on $n \times n$ matrices, $c_n(x, y, z) = \text{trace}(d)\text{trace}(z)I$. Moreover, $\text{trace}(d)$ is not identically zero on $M_n(k)$, hence $c_n(x, y, z)$ is not identically zero either.

Proof. The formula $c_n(x, y, z) = \text{trace}(d)\text{trace}(z)I$ follows from Proposition V.4.1 and Theorem V.4.4. It remains to show that $\text{trace}(d)$ is not identically zero. For this, we take for x the matrix unit basis $e_{i_\nu j_\nu}$, $1 \leq \nu \leq s$, in an arbitrary order. We set $y_0 = e_{1i_1}$, $y_N = e_{j_N 1}$, and for $0 < \nu < s$, $y_\nu = e_{j_\nu i_{\nu+1}}$. Then

$$y_0 x_1 y_1 \cdots y_{n-1} x_n y_n = e_{1i_1} e_{i_1 j_1} e_{j_1 i_2} e_{i_2 j_2} \cdots e_{i_N j_N} e_{i_N 1} = e_{11},$$

while $y_0 x_\sigma y_1 \cdots y_{n-1} x_\sigma y_n = 0$ if $\sigma \neq 1$. Then $\text{trace}(d) = \text{trace}(e_{11}) = 1$. \square

Proposition V.4.9. *The central polynomial $c_n(x, y, z)$ is identically zero on $r \times r$ -matrices with $r < n$.*

Proof. We can view an $r \times r$ matrix as the upper left block of an $n \times n$ matrix with zeros in the remaining positions. This is compatible with addition and multiplication. Then c_n evaluates to a scalar matrix with the bottom right entry zero, hence to zero. \square

Example V.4.10. A central polynomial for 2×2 -matrices. There is a simpler central polynomial for 2×2 matrices. The Cayley-Hamilton theorem for 2×2 matrices z shows that $z^2 = \text{trace}(z)z - \det(z)$. Since $\text{trace}[x, y] = 0$, for any 2×2 matrices x, y , it follows that $[x, y]^2 = -\det[x, y]$, hence that $[x, y]^2$ is a central polynomial of degree 4. The Razmyslov polynomial has degree 10.

V.5. Polynomial identity rings

A *polynomial identity* for an algebra A over a field k is a nonzero, noncommutative polynomial $f(x_1, \dots, x_r)$ with coefficients in k which vanishes identically on A , i.e., such that $f(a_1, \dots, a_r) = 0$ for all $a_1, \dots, a_r \in A$.

Proposition V.5.1. *Let k be an infinite field, and let A be a k -algebra.*

(i) *If A satisfies an identity $f(x_1, \dots, x_r) = 0$, then the homogeneous components of f also vanish identically on A .*

(ii) *If A satisfies a homogeneous polynomial identity of positive degree, then it also satisfies a multilinear identity of the same degree.*

Proof. This is proved by a method called *polarization*. Let t_1, \dots, t_r be indeterminate scalars. Then $f(t_1x_1, \dots, t_rx_r)$ expands to

$$(V.5.2) \quad \sum_{(i)} t_1^{i_1} \cdots t_r^{i_r} f_{(i)}(x),$$

where $f_{(i)}$ is the multihomogeneous part of f of degree i_ν in the variables x_ν . If f vanishes identically on A , so does $f(t_1x_1, \dots, t_rx_r)$, and the homogeneous parts can be recovered by making sufficiently many substitutions for t_ν in k . So they also vanish identically. This proves (i).

(ii) Suppose that f is multihomogeneous, say of degree d in x_1 . We substitute $x_1 = t_1y_1 + \cdots + t_dy_d$ into f , obtaining a polynomial $f(t_1y_1 + \cdots + t_dy_d, x_2, \dots, x_r)$ in the variables $y_1, \dots, y_d, x_2, \dots, x_r$, which vanishes identically, and we extract the coefficient of $t_1 \cdots t_d$. This is a multilinear polynomial in y_1, \dots, y_d . We continue with the other variables. \square

Corollary V.5.3. *If k is infinite and if A is not the zero ring, then $M_n(A)$ satisfies no polynomial identity of degree $< 2n$.*

Proof. This follows from the second part of the the Amitsur-Levitski theorem. \square

Proposition V.5.4. *Let A be a k -algebra, with or without unit element, which satisfies a multilinear identity $f(x_1, \dots, x_r) = 0$. Then for any commutative k -algebra R , $A \otimes R$ satisfies the same identity, as does any quotient of $A \otimes R$. \square*

We will call a k -algebra A a *polynomial identity algebra*, or a PI algebra for short, if $A \otimes \bar{k}$, where \bar{k} is the algebraic closure of k , satisfies a polynomial identity.

Thus commutative algebras and finite-dimensional algebras are a PI algebras:

V.6. Kaplansky's theorem

The next result is a Corollary of the Jacobson density theorem.

Proposition V.6.1. *Let A be a ring, let V be a simple right A -module, and let $D = \text{End } V_A$.*

(i) *If V has dimension $\geq n$ over the division ring $D = \text{End } V_A$, then the matrix algebra $M_n(D)$ is a subquotient of A , i.e., there is a subring B of A and a surjective map $B \longrightarrow M_n(D)$.*

(ii) *If A satisfies a homogeneous polynomial identity of degree $2r$, then $\dim_D V \leq r$, and if $\dim_D V = n$, then $A \approx M_n(D)$.*

Proof. (i) We choose a D -submodule U of dimension n , and we let B be the subring of A of elements which carry U to itself. Then restriction to U defines a homomorphism $B \longrightarrow \text{End}_D U \approx M_n(D)$, and the density theorem tells us that this homomorphism is surjective.

(ii) This follows from (i), Corollary V.5.3, and the density theorem \square

We now come to Kaplansky's theorem, the main theorem in the theory of PI algebras. Remarkably, the exact identity which holds is of no importance.

Theorem V.6.2. *(Kaplansky) Let A be a PI algebra over a field k , which satisfies a multilinear identity of degree $2r$ and which has a faithful simple right module V (so A is a primitive ring). Then A is a central simple algebra over its center, a field K , and $[A : K] \leq r^2$.*

We first treat the case that A is a division ring.

Lemma V.6.3. *A division ring D which satisfies a multilinear polynomial identity is finite over its center K .*

Proof. We may choose a multilinear identity $p(x_1, \dots, x_r)$ of degree $2r$ for D . Let L be a maximal commutative subfield of D , let $B = L^o \otimes D$ (which is equal to $L \otimes D$ because L is commutative), and let $V = D$. The L, D -bimodule structure on V makes V into a right B -module. Then

- (a) $[B : L] = [D : K]$.
- (b) B satisfies the same identity (V.5.4).
- (c) B is a simple ring (CSA,1.8), and its center is L (CSA,1.3).
- (d) $V := D$ is a simple B -module, because it is a simple D -module.
- (e) $\text{End } V_B = L$.

For the last assertion, we note that $\text{End } V_D = D$, hence $\text{End } V_D$ is the centralizer of L in D , which is L itself because L is maximal (CSA,4.1).

Then B is dense in $\text{End}_L V$, and by Proposition V.6.1, $\dim_L V \leq r$, and $[B : L] = [D : K] \leq r^2$. \square

Proof of Theorem V.6.2. Let $D = \text{End } V_A$, and say that $\dim_D V \geq n$. Because V is a faithful module, A embeds into $\text{End}_D V$, and by Proposition V.6.1, $A \approx M_n(D)$. Since $D \subset M_n(D)$, D also satisfies a polynomial identity, and is a finite module over its center K . This shows that A is a central simple algebra over K . If \overline{K} denotes the algebraic closure of K , then $A \otimes \overline{K}$ is a matrix

algebra, say $M_n(\overline{K})$ over \overline{K} (CSA,2.3), and it also satisfies the same identity (V.5.4). So, $n \leq r$ and $[A : K] \leq r^2$. \square

V.7. A theorem of Amitsur

The symbol $A[t]$ denotes the ring of polynomials with coefficients in A , and in which t is a central variable, i.e., the elements of A commute with t .

Theorem V.7.1. (*Amitsur*) *Let A be a k -algebra which contains no nonzero nil ideal, for example a prime algebra. Then the Jacobson radical of $A[t]$ is zero.*

Example V.7.2. Let k be algebraically closed, and let A be the (commutative) power series ring $k[[x]]$. The Jacobson radical of $A[t]$ is the intersection of its maximal ideals, and the theorem asserts that this intersection is zero. This seems a bit mysterious at first, because the obvious maximal ideals of $A[t]$ are the ideals $(x, t - c)$ for $c \in k$. Their intersection is $xA[t]$. However, there are other maximal ideals. For instance, the polynomial $xt - 1$ generates a maximal ideal. The ring $A[t]/(xt - 1)$ can be viewed as the ring obtained from $k[[x]]$ by adjoining the inverse of x , and inverting x in $k[[x]]$ gives us the field of fractions $k((x))$.

Exercise: Describe all maximal ideals of the ring $A[t]$ in the above example.

Proof of the theorem. We argue by contradiction. Let J be the Jacobson radical of $A[t]$, and let $\alpha = \alpha(t)$ be a nonzero element of J . Then $-\alpha t \in J$ as well, and therefore $1 - \alpha t$ is invertible in $A[t]$, say $(1 - \alpha t)^{-1} = g(t)$. If A were commutative, the next lemma would show that the coefficients of α are nilpotent.

Lemma V.7.3. *Let R be a commutative ring. A polynomial $f(t) = a_0 + a_1t + \cdots + a_nt^n$ is invertible in the polynomial ring $R[t]$ if and only if a_i is nilpotent for every $i > 0$, and a_0 is a unit of R . Hence, if $f(t) = 1 - \alpha(t)t \in R[t]$ is invertible, then $\alpha(t)$ is nilpotent. \square*

It is harder to interpret the existence of an inverse in the noncommutative case. But there is a trick: Let $S = \{i_1, \dots, i_r\}$ be an ordered set of indices for which there exists a nonzero element $\alpha = a_{i_1}t^{i_1} + \cdots + a_{i_r}t^{i_r}$ in J . We choose S so that r is minimal, and we compute the commutator $[a_{i_\nu}, \alpha]$. The commutator is in J , and its coefficient of t^{i_ν} is $[a_{i_\nu}, a_{i_\nu}] = 0$. Since r is minimal, we can conclude that $[a_{i_\nu}, \alpha] = 0$, which means that $[a_{i_\nu}, a_{i_\mu}] = 0$ for all μ, ν . So the ring $R = k\langle a_{i_\nu} \rangle$ generated by the coefficients of α is commutative, and $1 - \alpha t \in R[t]$.

Now we regard $A[t]$ as subring of the ring of formal power series $A[[t]]$. In the ring $A[[t]]$, we have $R[t] = A[t] \cap R[[t]]$. The polynomial $1 - \alpha t$ is invertible in $A[t]$ and in $R[[t]]$. Therefore it is invertible in $R[t]$, and this shows that the coefficients a_{i_ν} of α are nilpotent.

Finally, for this particular set S , the vector space N of coefficients a_{i_1} of lowest degree among elements of J which are linear combinations of t^{i_1}, \dots, t^{i_r} is an ideal of A , and as we have seen, it is a nil ideal. \square

Corollary V.7.4. *Let A be a PI algebra which contains no nil ideal. Then A has a faithful family of irreducible representations of bounded degree. In other words, there is a family of matrix representations $\rho_i : A \longrightarrow M_{n_i}(K_i)$, with n_i bounded, such that the product map $A \longrightarrow \prod M_{n_i}(K_i)$ is injective.*

Proof. We may replace A by $A[t]$ (V.5.4), hence assume that the Jacobson radical of A is zero. A ring whose Jacobson radical is zero has a faithful family of simple modules: For every maximal right ideal M of A , A/M is a simple module in which M is the annihilator of the residue of 1. Since J is the intersection of these maximal right ideals, the operation of A on the product of simple modules is faithful. To be specific, we may take one simple module V_i from each isomorphism class $[V_i]$, and set $D_i = \text{End}(V_i)_A$. Then the map

$$(V.7.5) \quad A \longrightarrow \text{End}_{D_i} V_i$$

is injective.

Suppose, as in our case, that A is also a PI ring. Then by Kaplansky's theorem, $\text{End}_{D_i} V_i$ is a central simple algebra of bounded rank, which we can split in a finite extension K_i of the center K_i . So in this case A embeds as subring into a direct product of matrix algebras $M_{r_i}(K_i)$, with r_i bounded, as claimed. \square

V.8. Posner's theorem

Theorem V.8.1. (*Posner*) *A prime PI algebra has a right ring of fractions which is a central simple algebra. More precisely, let Z denote the center of A , and let K be the field of fractions of Z . Then $A \otimes_Z K$ is a central simple K -algebra.*

The original proof of this theorem showed that, though A needn't be noetherian, Goldie's conditions were satisfied. Since central polynomials became available, the proof has been simplified.

Theorem V.8.2. (*Rowan*) *Let A be a prime PI algebra with center Z , and let I be a nonzero ideal of A . Then $I \cap Z \neq 0$.*

Proof. Since A is prime, it has a faithful family of irreducible representations, and I has a nonzero image in at least one of the corresponding matrix algebras, say in $B = M_n(K)$. Thus $A \otimes K \longrightarrow B$ is surjective, and the image of I in B is not zero. We choose such a B with n maximal. The right ideal IB is the image of $I \otimes K$ in B , so it is a two-sided ideal, and $IB = B$ because B is a simple ring.

Let $c(x, y)$ be a multilinear central polynomial for $n \times n$ matrices. then c has a nonzero evaluation in B , so it is not an identity in B . The fact that there is a surjective map $I \otimes K \longrightarrow B$ shows that c is not an identity in $I \otimes K$. By Proposition V.5.4, it is not an identity in I either. So there are elements $u, v \in I$ with $c(u, v) \neq 0$. If $A \longrightarrow M_r(L)$ is an irreducible representation of A with $r \leq n$, then the image of $c(u, v)$ is in the center because c is a central polynomial (in fact, it is zero if $r < n$). And if $r > n$, then the image of $c(u, v)$ is zero because $c(u, v) \in I$ and n was chosen maximal. Hence $c(u, v)$ maps to the center in every irreducible representation, which shows that it is in the center of A . \square

Proof of Posner's theorem. With the notation of the theorem, $A \otimes_Z K$ is a localization of A , and it is a PI ring whose center is K (CSA,1.3). So to prove the theorem, we may replace A by $A \otimes_Z K$. This reduces the theorem to the next lemma.

Lemma V.8.3. *A prime PI ring A whose center is a field K is a central simple algebra over K .*

Proof. Theorem V.8.2 shows that A is a simple ring. So it has a faithful irreducible module, and (V.6.2) shows that it is central simple. \square

Corollary V.8.4. *Let A be a prime PI algebra which satisfies the standard identity $S_{2n} = 0$. Then A satisfies all identities which hold for $n \times n$ matrices over \bar{k} .*

Proof. By Posner's theorem, A embeds into a central simple algebra which, by (V.6.4) has rank at most n^2 over its center. \square

Lemma V.8.5. (i) *Let R be a commutative ring, and let I be an ideal of an R -algebra A , and let B be a faithfully flat R -algebra. If $I \otimes B$ is a prime ideal of $A \otimes B$, then I is a prime ideal of A .*
(ii) *The kernel P of an irreducible representation ρ is a prime ideal.*

Proof. (i) First, $I \otimes B$ is an ideal of $A \otimes B$ because B is flat. Suppose that $I \otimes B$ is a prime ideal. Let $a, a \in A$. If $axa = 0$ for all $x \in A$, then $(a \otimes 1)(x \otimes b)(a \otimes 1) = 0$ in $A \otimes B$ for all b , hence $a \otimes 1 = 0$ or $a \otimes 1 = 0$, i.e., $a = 0$ or $a = 0$. Therefore I is prime.

(ii) The kernel of the surjective map $A \otimes K \longrightarrow M_n(K)$ is prime because $M_n(K)$ is a prime ring. \square

Definition V.8.6. Two matrix representations $\rho : A \longrightarrow M_n(K)$ and $\rho' : M_n(K)$ are called *equivalent* if there is a common field extension L of K and K' such that the induced representations $\rho \otimes_{K'} L : A \longrightarrow M_n(L)$ and $\rho' \otimes_{K'} L : A \longrightarrow M_n(L)$ are conjugate.

Theorem V.8.7. *Two irreducible matrix representations ρ, ρ' are equivalent if and only if their kernels are equal.*

Proof. It is clear from the definition that equivalent representations have the same kernel. To prove the converse, let P be the kernel of a pair of irreducible representations ρ, ρ' , which is a prime ideal. We may replace A by the prime ring A/P , hence suppose that the kernel is zero. Then because the center of $M_n(K)$ is a field, every nonzero element of the center Z of A becomes invertible in $M_n(K)$. So ρ factors through the ring of fractions $A \otimes_Z K_0$, where K_0 is the field of fractions of Z . By Posner's theorem, $A \otimes_Z K_0$ is a central simple algebra. this reduces us to the case that A is a central simple algebra over a field K_0 .

When A is central simple, the representation ρ defines an injection of centers $Q \longrightarrow K$, and ρ is a homomorphism of Q -algebras. So $A \otimes_Q K \approx M_n(K)$. The Skolem-Noether theorem tells us that two isomorphisms to the matrix algebra differ by an inner automorphism, hence they define equivalent representations. So ρ and ρ' become equivalent in any common field extension L . \square

V.9. An intrinsic characterization of Azumaya algebras

Matrix algebras satisfy identities which are not consequences of the standard identity, but though much work has been done to describe them, they are still not completely understood. One example of such an identity for 2×2 matrices is obtained from the central polynomial $[x, y]^2$ (4.10):

$$(V.9.1) \quad [[x, y]^2, z] = 0.$$

is an identity.

Exercise: Show that V.9.1 is not a consequence of the standard identity.

On the other hand, these identities have not been of very much importance, because the study of PI rings has emphasized prime rings. As we know from Posner's theorem, a prime ring A which satisfies the standard identity $S_{2n} = 0$ actually satisfies all identities of $n \times n$ matrices over \bar{k} . We'll say that a prime ring A has *PI degree* n if $S_{2n} \equiv 0$ holds in A .

Theorem V.9.2. *Let A be a k -algebra which satisfies all identities of $n \times n$ matrices over \bar{k} , and which has no irreducible representation of dimension less than n . Then A is an Azumaya algebra of rank n^2 over its center.*

Corollary V.9.3. *Let A be a prime k -algebra of PI degree n , and let γ be a nonzero evaluation of the Razmyslov polynomial in A . Then $A = A[\gamma^{-1}]$ is an Azumaya algebra.*

Proof of the Corollary. Let $\rho : A \rightarrow M_r(K)$ be a representation of dimension $r < n$, and let $P = \ker(\rho)$. By Proposition 4.9, $\gamma \in P$. Therefore P generates the unit ideal in A . This shows that ρ does not extend to a representation of A . On the other hand, every representation of A restricts to a representation of A of the same dimension. So A has no representation of dimension $r < n$ at all, and A is Azumaya. \square

Lemma V.9.4. *Let R be a commutative ring, and V an R -module. Suppose given elements $v_1, \dots, v_n \in V$ and $\eta_1, \dots, \eta_n \in V^* = \text{Hom}_R(V, R)$ such that $x = \sum \eta_i(x)v_i$ for all $x \in V$. Then V is a projective R -module.*

Proof. The formula shows that v_i generate V , so there is a surjection $\pi : R^n \rightarrow V$ which sends the standard basis vector e_i to v_i . The map $\theta : V \rightarrow R^n$ defined by $\theta(x) = \sum \eta_i(x)e_i$ splits this projection: $\pi\theta(x) = \sum \eta_i(x)v_i = x$. So V is projective. \square

Proof of Theorem V.9.2. (Schelter) We'll prove the theorem here in the case of a prime ring A . Posner's theorem shows that a prime ring A satisfies the identities of $n \times n$ matrices if and only if it satisfies any multilinear identity of degree $2n$.

Let $c(x_1, \dots, x_N, y)$ be the Razmyslov polynomial, which we know is multilinear and alternating in the variables x_i , and also involves some other variables, which we call y . Because A embeds into an $n \times n$ matrix algebra (by Posner's theorem), every evaluation of c in A yields an element of its center. Notice that s is the largest number of variables in which a polynomial can be multilinear and alternating without vanishing identically on $n \times n$ matrices. Schelter considers the following expression:

$$(V.9.5) \quad \phi(x_0, \dots, x_N; y) = \sum (-1)^\nu c(x_0, \dots, x_\nu, \dots, x_N; y)x_\nu.$$

This is a multilinear alternating polynomial in the $N + 1$ variables x_0, \dots, x_N , and in some auxiliary variables y . Evaluating on $n \times n$ matrices and fixing y , it becomes a multilinear map $\bigoplus^{N+1} M_n(K) \rightarrow M_n(K)$, or a linear map $\Lambda^{N+1} M_n(K) \rightarrow M_n(K)$, which is therefore identically zero.

Suppose first that there is an evaluation of c in A which is an invertible central element, say $\gamma = c(\alpha_1, \dots, \alpha_N; \beta)$, where α, β are $n \times n$ matrices. We substitute $x_0 = x$, $x_i = \alpha_i$ for $i = 1, \dots, N$, and $y = \beta$ into ϕ , obtaining the expression

$$(V.9.6) \quad \gamma x = \sum (-1)^{\nu+1} c(x, \alpha_1, \dots, \alpha_\nu, \dots, \alpha_N; \beta)\alpha_\nu.$$

Let

$$(V.9.7) \quad \eta_\nu(x) = (-1)^{\nu+1}c(x, \alpha_1, \dots, \alpha_\nu, \dots, \alpha_N; \beta).$$

Then $\eta_\nu(x)$ is a central element for all $x \in A$. Since it is also a linear function of x , we can view it as an element of $A^* = \text{Hom}_R(A, R)$. Then

$$(V.9.8) \quad \gamma x = \sum \eta_\nu(x)\alpha_\nu.$$

Lemma V.9.4 shows that A is a projective R -module, generated by α_ν . So the rank of A as R -module is at most n^2 at any point. Since the rank of a projective module is constant on every connected component of $\text{Spec } R$, and since $R \subset A$, the rank is positive.

Let $R \rightarrow \overline{K}$ be any homomorphism to an algebraically closed field. Then $A \otimes_R \overline{K}$ is not zero, and it has dimension at most n^2 . Because there is no irreducible representation of dimension $< n$, it follows from Wedderburn's theorem that $A \otimes_R \overline{K} \approx M_n(\overline{K})$. Therefore A is an Azumaya algebra. This completes the proof in the case that there is an evaluation of the central polynomial which is a unit.

In the general case, we use the next lemma:

Lemma V.9.9. *With the hypotheses of the theorem, there is a finite set of evaluations γ_i of the central polynomial $c(x; y)$ in A which generate the unit ideal in A , i.e., such that for some $a_i \in A$,*

$$\sum \gamma_i a_i = 1.$$

Proof. It suffices to show that the evaluations of $c(x; y)$ do not all lie in any maximal ideal M of A . If M is a maximal ideal, then by Kaplansky's theorem, $\tilde{A} = A/M$ is central simple over its center K , and by hypothesis, $\tilde{A} \otimes \overline{K} \approx M_n(K)$. Therefore there is a nonzero evaluation of $c(x; y)$ in \tilde{A} , which we can lift to A . \square

Going back to the formula (V.9.8), we may write

$$(V.9.10) \quad \gamma_i x = \sum_\nu \eta_{i,\nu}(x)\alpha_{i,\nu}$$

for each i , and since $x = \sum \gamma_i x a_i$,

$$(V.9.11) \quad x = \sum_{i,\nu} \eta_{i,\nu}(x)\alpha_{i,\nu} a_i.$$

This shows again that A is a finitely generated projective R -module. The elements c_i do not vanish identically on $\text{Spec } R$ because they don't vanish identically at any maximal ideal of A , and every maximal ideal of R has a maximal ideal of A lying over it. So $\{c_i\}$ generates the unit ideal in R too. By what has been proved, $A[c_i^{-1}]$ is an Azumaya algebra for every i . It follows that A is Azumaya. \square

V.10. Irreducible representations of the free ring

A matrix representation of the free ring $k\langle x_1, \dots, x_m \rangle$ is given by assigning arbitrary matrices as images of the variables. In itself, this is not particularly interesting. However, when one asks for equivalence classes of irreducible representations, one is led to an interesting problem in invariant theory. This is the topic of the remaining sections.

In principle, this discussion will also apply to finitely generated k -algebras which are presented as quotients of the free ring. Their representations form certain loci in the spaces of representations of the free ring.

In order to simplify our discussion, we assume that our ground field k is algebraically closed, and we study representations into $M_n(k)$. By points of a variety X we will mean closed points, i.e., k -valued points here, and if R is a commutative ring, we'll write $\text{Spec } R$ for the space of maximal ideals. This restriction to an algebraically closed field is minor.

A more important assumption will be that the characteristic of the field k is zero. This is needed in order to apply classical invariant theory, and is essential for much of what follows.

The two assumptions will be in force throughout the rest of these notes.

Let $m \geq 2$ be an integer, and let X_1, \dots, X_m be $n \times n$ matrices with entries x_{ij}^ν which are independent central variables. The subring of the matrix algebra $M_n(k[x_{ij}^\nu])$ generated by the matrices X_ν is called, rather ambiguously, the *algebra of generic matrices*. We'll denote it by $k\langle X_1, \dots, X_m \rangle$. Of course there is a canonical homomorphism

$$(V.10.1) \quad k\langle x_1, \dots, x_m \rangle \xrightarrow{\pi} k\langle X_1, \dots, X_m \rangle$$

from the free ring on variables x_1, \dots, x_m to this ring.

If u_1, \dots, u_m are $n \times n$ matrices with entries in a commutative k -algebra R , we can substitute u_j for X_j , and thereby obtain a homomorphism $k\langle X_1, \dots, X_m \rangle \longrightarrow M_n(R)$.

Proposition V.10.2. (i) *A polynomial $f(x_1, \dots, x_m)$ is in the kernel of the map π if and only if it vanishes identically on $M_n(R)$ for every commutative k -algebra R , and this is true if and only if f vanishes identically on $M_n(k)$.*

(ii) *The (irreducible) matrix representations of the free ring $k\langle x_1, \dots, x_m \rangle$ of dimension $\leq n$ correspond bijectively to the (irreducible) matrix representations of the ring of generic matrices, \square*

Lemma V.10.3. *Let u_1, \dots, u_N be a basis for the matrix algebra $M_n(\overline{K})$, and let z_1, \dots, z_N be indeterminates. The entries of the matrix $Z = \sum z_j u_j$ are algebraically independent. \square*

Theorem V.10.4. (Amitsur) *The algebra $k\langle X_1, \dots, X_m \rangle$ of generic matrices is a domain.*

Proof. We choose a field extension K of k and a division ring D with center K , such that $[D : K] = n^2$, and we let u_1, \dots, u_N be a K -basis for D . Let z_{ij} be central indeterminates. The polynomial ring $D[z_{ij}] = D \otimes k[z_{ij}]$ is a domain. Over the algebraic closure \overline{K} , $D \otimes \overline{K} \approx M_n(\overline{K})$, so we can view the elements u_j as $n \times n$ matrices, and they form a basis of the matrix algebra. The lemma tells us that the matrices $Z_i = \sum z_{ij} u_j$ have algebraically independent entries. So the map $k\langle X_1, \dots, X_m \rangle \longrightarrow M_n(\overline{K}[z_{ij}])$ which sends $X_i \mapsto Z_i$ is injective, and its image is in the domain $D[z_{ij}]$. So $k\langle X \rangle$ is a domain too.

The only thing to be verified is that such a division ring D exists. The simplest construction of D is as a cyclic algebra. Let $L = k(w_1, \dots, w_n)$ be a pure transcendental extension of k , and consider the automorphism σ of L which permutes w_1, \dots, w_n cyclically. Let B denote the Ore extension $L[y; \sigma]$. This is a Noetherian domain, which has a division ring of fractions D . The center of D is $K = L^\sigma(v)$ where $v = y^n$, and $[D : K] = n^2$. \square

Exercise: Use Amitsur's theorem to prove (V.9.2) for algebras A which are not necessarily prime.

Because an n -dimensional matrix representation ρ of the free ring $k\langle x_1, \dots, x_m \rangle$ is determined by assigning arbitrary images to the variables, it corresponds to a point in the affine space whose coordinates are the n^2m matrix entries $x_{ij}^{(\nu)}$ of the variable matrices X_ν . This affine space is the space $\text{Spec } R$ of maximal ideals of the polynomial ring $R = k[x_{ij}^{(\nu)}]$. So in what follows, we will identify points of the affine space $\text{Spec } R$ with n -dimensional matrix representations.

The general linear group $G = GL_n(k)$ operates on the matrices (X_1, \dots, X_m) by simultaneous conjugation. An invertible matrix $p \in G$ operates as

$$(V.10.5) \quad p(X_1, \dots, X_m)p^{-1} = (pX_1p^{-1}, \dots, pX_mp^{-1}).$$

This induces an action on the affine space $\text{Spec } R$, and on the commutative ring R . Viewing points of the affine space as representations, the orbits for the action of G are the equivalence classes of representations. We would like to parametrize the equivalence classes by constructing a variety whose points are the G -orbits, and classical invariant theory provides a reasonably satisfactory solution to this problem.

To be specific about the action of G on the polynomial ring R , the operation of $p \in G$ sends the variable $x_{ij}^{(\nu)}$ to the i, j entry of the matrix $pX_\nu p^{-1}$. This rule induces the operation on $R = k[x_{ij}^{(\nu)}]$.

In order minimize indices while avoiding ambiguity, we use the following notation: If X, Y are indeterminate matrices, we write a commutative polynomial f in the variable matrix entries x_{ij}, y_{ij} as $f([X], [Y])$. Then the action of p on R sends a polynomial $f([X_1], \dots, [X_m])$ to

$$(V.10.6) \quad f^p([X_1], \dots, [X_m]) = f([pX_1p^{-1}], \dots, [pX_mp^{-1}]).$$

The polynomials $f \in R$ which are invariant under this action are constant on the G -orbits. So we can try to use invariant polynomials to parametrize the orbits.

Some of the simplest invariant polynomials are $\text{trace}(X_1)$, $\text{trace}(X_1X_2)$, $\det(X_1)$ and $\det(X_1 + X_2)$.

The next theorem is an application of the Hilbert-Mumford theory of invariants for reductive group actions.

Theorem V.10.7. (*Hilbert-Mumford*)

(i) *The ring of invariants R^G is a finitely generated, integrally closed k -algebra.*

(ii) *The map $X = \text{Spec } R \rightarrow \text{Spec } R^G = Y$ is surjective. The fibre X_y over a point $y \in Y$ is a union of G -orbits, and it contains exactly one closed orbit. Thus the closed orbit is in the closure of every other orbit in the fibre.*

V.11. The case of two 2×2 matrices X, Y

Let X, Y denote two variable 2×2 matrices, and let $R = k[x_{ij}, y_{ij}]$. Let

$$(V.11.1) \quad t_1 = \text{trace}(X), t_2 = \text{trace}(Y), t_3 = \text{trace}(X^2), t_4 = \text{trace}(Y^2), t_5 = \text{trace}(XY).$$

We can recover traces from determinants by the formula

$$\text{trace}(Z) = \det(Z + 1) - \det(Z) - 1.$$

And since the characteristic is zero, we can recover determinants from traces:

$$\det(Z) = \frac{1}{2}(\text{trace } Z)^2 - \text{trace}(Z^2).$$

Proposition V.11.2. (i) $[X, Y]^2 = t_1^2 t_4 + t_2^2 t_3 - 2t_3 t_4 - 2t_1 t_2 t_5 + 2t_5^2$.

(ii) Two 2×2 matrices u, v with entries in k generate the matrix algebra $M_2(k)$ as k -algebra if and only if $1, u, v, uv$ form a basis of $M_2(k)$. This is true if and only if $[u, v]^2 \neq 0$.

Proposition V.11.3. The five traces t_1, \dots, t_5 are algebraically independent, and they generate the ring R^G of all invariant functions in the polynomial ring $R = k[x_{ij}, y_{ij}]$.

It is a good exercise to prove these assertions without appealing to Theorem V.10.7. One can start by showing that two different semisimple representations can be distinguished by these trace functions. Then, once it is proved that t_1, \dots, t_5 are algebraically independent, the proof that they generate R^G is simplified because the ring that they generate is integrally closed.

It is very unusual for a ring of invariants to be a polynomial ring, and this does not happen when $n > 2$.

V.12. Some tensor notation.

This section contains some notation and some formulas that we need. The verifications of the formulas are straightforward. If you don't carry these verifications out, you risk becoming lost in the notation.

Let V be an n -dimensional vector space over k , and let $V^* = \text{Hom}_k(V, k)$, and $E = \text{End } V$. Thus $E \approx M_n(k)$. We let V^* and E act on the left on V . If $v \in V$ and $\alpha \in V^*$, we write

$$(V.12.1) \quad \alpha(v) = \{\alpha, v\}.$$

There is a canonical isomorphism $V \otimes V^* \approx E$, where a tensor $u \otimes \alpha$ acts on a vector $v \in V$ by the formula

$$(V.12.2) \quad [u \otimes \alpha]v = u\{\alpha, v\},$$

$\{\alpha, v\}$ being a scalar factor. Similarly, E acts on the right on V^* , by

$$(V.12.3) \quad \beta[u \otimes \alpha] = \{\beta, u\}\alpha.$$

Multiplication in E is given by the formula

$$(V.12.4) \quad (u \otimes \alpha)(v \otimes \beta) = u \otimes \beta \{\alpha, v\},$$

where as before, $\{\alpha, v\}$ is a scalar factor.

The trace map $E \rightarrow k$ is given by

$$(V.12.5) \quad \text{trace}(u \otimes \alpha) = \{\alpha, u\}.$$

Therefore the trace pairing $\text{trace}(pq)$ on E is

$$(V.12.6) \quad \langle u \otimes \alpha, v \otimes \beta \rangle = \{\beta, u\}\{\alpha, v\}.$$

The trace pairing defines an isomorphism $E \approx E^*$. This isomorphism is the same as the one defined by the symmetry of the tensor product:

$$(V.12.7) \quad E^* = V^* \otimes V \approx V \otimes V^* = E.$$

Denoting by $V^{\otimes m}$ the m -th tensor power of V , we have

$$(V^{\otimes m})^* \approx (V^*)^{\otimes m}$$

and

$$(V.12.8) \quad E^{\otimes m} \approx V^{\otimes m} \otimes V^{*\otimes m} \approx \text{End } V^{\otimes m}.$$

Thus $E^{\otimes m}$ is an $n^m \times n^m$ matrix algebra. The trace pairing on $E^{\otimes m}$ is given on tensors by the formula

$$(V.12.9) \quad \langle p_1 \otimes \cdots \otimes p_m, q_1 \otimes \cdots \otimes q_m \rangle = \langle p_1, q_1 \rangle \cdots \langle p_m, q_m \rangle.$$

V.13. The main theorem of invariant theory

The group $G = GL_n(k)$ acts on V , on V^* , and on E by conjugation. An element $p \in G$ acts by

$$(V.13.1) \quad v \mapsto pv, \quad \alpha \mapsto \alpha p^{-1}, \quad \text{and} \quad v \otimes \alpha \mapsto pv \otimes \alpha p^{-1}.$$

The induced action of p on $E^{\otimes m}$ is conjugation by the matrix $p^{\otimes m} = p \otimes \cdots \otimes p$.

We consider two subalgebras of the algebra $E^{\otimes m}$: the group algebra \mathcal{B} of the symmetric group S_n , which operates on $V^{\otimes m}$ by permuting its factors, and the subalgebra $\mathcal{A} = \text{Symm}^m(E)$ consisting of symmetric tensors. The symmetric tensors are those elements which are invariant under permutation of the factors of $E^{\otimes m}$. They are linear combinations of *orbit sums* of the form

$$(V.13.2) \quad \sum_{\sigma} x_{\sigma 1} \otimes \cdots \otimes x_{\sigma m},$$

with $x_1, \dots, x_m \in E$.

Proposition V.13.3. $\mathcal{A} = \text{Sym}^m(E)$ is spanned by elements of the form $p^{\otimes m}$, with $p \in GL_n$.

Proof. First, the elements $p^{\otimes m}$ with $p \in GL_n$ are dense in the space of all elements $x^{\otimes m}$ with $x \in E$, so they span the same subspace. To show that the orbit sums (V.13.2) can be written in terms of these elements, we polarize. Let $x_t = x_1 t_1 + \cdots + x_m t_m$, where t_i are indeterminate scalars. We expand $x_t^{\otimes m}$ formally as a polynomial in t_1, \dots, t_m . The coefficient of $t_1 \cdots t_m$ is (V.13.2), and it can be recovered using a finite number of substitutions for t_i in k . \square

Theorem V.13.4. (Main theorem of invariant theory) The algebras \mathcal{A} and \mathcal{B} are semisimple, \mathcal{B} is the centralizer of \mathcal{A} , and \mathcal{A} is the centralizer of \mathcal{B} .

Proof. The group algebra \mathcal{B} is semisimple by Maschke's theorem. So its centralizer \mathcal{B} is semisimple, and the centralizer of \mathcal{B} is \mathcal{B} . [(MO,A.1) is still not stated in sufficient generality. :(] So the only thing that needs to be verified is that the centralizer of \mathcal{B} is \mathcal{A} .

Of course an element commutes with \mathcal{B} if and only if it commutes with every permutation. Perhaps it should be obvious from the definitions that such an element is a symmetric tensor, but I don't see why it is obvious. So, let σ be a permutation. By definition, $\sigma(v_1 \otimes \cdots \otimes v_m) = v_{\sigma 1} \otimes \cdots \otimes v_{\sigma m}$. This allows us to write σ explicitly as element of E , in terms of the matrix units:

$$(V.13.5) \quad \sigma = \sum_{(i)} e_{i_{\sigma 1} i_1} \otimes \cdots \otimes e_{i_{\sigma m} i_m},$$

the sum being over all multi-indices (i_1, \dots, i_m) . To verify this formula, we compute the product with tensors $e_{j_1} \otimes \cdots \otimes e_{j_m} \in V^{\otimes m}$. A similar computation shows that σ^{-1} operates on the right on $V^{*\otimes m}$ by the permutation σ .

If $x = (e_{i_1} \otimes \epsilon_{j_1}) \otimes \cdots \otimes (e_{i_m} \otimes \epsilon_{j_m}) = e_{i_1 j_1} \otimes \cdots \otimes e_{i_m j_m}$, then

$$(V.13.6) \quad \sigma x \sigma^{-1} = e_{i_{\sigma 1} j_{\sigma 1}} \otimes \cdots \otimes e_{i_{\sigma m} j_{\sigma m}}.$$

A linear combination

$$(V.13.7) \quad \sum c_{(i),(j)} e_{i_1 j_1} \otimes \cdots \otimes e_{i_m j_m}$$

commutes with σ if and only if

$$(V.13.8) \quad c_{\sigma(i),\sigma(j)} = c_{(i),(j)}$$

for all multi-indices $(i), (j)$. \square

V.14. Procesi's theorem

To state Procesi's results, we need to define the semisimple representation associated to a representation $\rho : A \rightarrow M_n(k)$. As always, we assume that k is algebraically closed, but the algebra A can be arbitrary here. Recall that a right A -module V is semisimple if it is a direct sum of simple submodules. Any module V which is finite-dimensional as k -module has finite length, and so there is a filtration by A -submodules $0 = V_0 \subset V_1 \subset \cdots \subset V_r = V$, such that the successive quotients $\bar{V}_i = V_i/V_{i-1}$ are simple. The Jordan-Hölder theorem says that, though the filtration is not canonical, the associated graded module $\bar{V} = \bigoplus \bar{V}_i$ is determined up to isomorphism. And of

course $\dim \bar{V} = \dim V$. We call \bar{V} the *associated semisimple representation* of the representation V .

If we choose a basis v_i of V compatibly with the filtration, then the corresponding matrix representation ρ has a block triangular form. There are r diagonal blocks $\bar{\rho}_i$, where ρ_i is the matrix representation associated to \bar{V}_i with the induced basis. It is an irreducible matrix representation of dimension $\dim \bar{V}_i$. The representation ρ is zero below the blocks. The entries above the blocks needn't be zero, and the associated semisimple representation is obtained by replacing these entries by zero, i.e., it is the block diagonal representation

$$(V.14.1) \quad \rho^{ss} = \bar{\rho}_1 \oplus \cdots \oplus \bar{\rho}_r.$$

Thus, if a matrix representation $\rho : A \rightarrow M_n(k)$ is given, the associated semisimple representation ρ^{ss} is determined only up to equivalence.

Theorem V.14.2. (*Procesi*)

- (i) The ring of invariants R^G is generated by traces of monomials in the matrices X_1, \dots, X_m .
- (ii) The closed orbits are the equivalence classes of semi-simple representations. A matrix representation ρ has the same image in $\text{Spec } R^G$ as its associated semisimple representation ρ^{ss} . Thus the orbit of ρ has the closed orbit of ρ^{ss} in its closure.

Proof. There are two steps in the proof of (i). We first reduce the question to multilinear invariants, and then apply the main theorem.

Lemma V.14.3. *It suffices to prove the theorem for G -invariant polynomials which are multilinear.*

Proof. First of all, the number of variables appearing is unimportant. If we could write an invariant function in R^G using auxiliary variables, then we could replace those variables by zero. (We're doing commutative algebra here.)

Next, conjugation $X \mapsto pXp^{-1}$ is a linear operation, so the action of G on R is multi-homogeneous: If $(d) = (d_1, \dots, d_m)$ is a multi-index and if $R_{(d)}$ denotes the set of polynomials whose degree in $[X_\nu] = \{x_{ij}^{(\nu)}\}$ is d_ν , then G carries $R_{(d)}$ to itself. Therefore $R^G = \bigoplus R_{(d)}^G$.

We reduce homogeneous invariant polynomials to multilinear ones by polarization. Let $f([X], [Y], \dots)$ be a multihomogeneous, G -invariant polynomial function of some matrix entries $[X], [Y], \dots$, which is homogeneous of degree d in $[X]$. Let X_1, \dots, X_d be indeterminate matrices, and let t_1, \dots, t_d be scalar indeterminates. The polynomial $f([t_1X_1 + \cdots + t_dX_d], [Y], \dots)$ is G -invariant too. Moreover, if we collect terms of like multi-degree in t , and write

$$f([t_1X_1 + \cdots + t_dX_d], [Y], \dots) = \sum_{(i)} f_{(i)}([X_1], \dots, [X_d], [Y], \dots)t^{(i)},$$

then each $f_{(i)}$ is G -invariant. Setting $(1) = (1, \dots, 1)$, the invariant polynomial $f_{(1)}([X_1], \dots, [X_d], [Y], \dots)$ is G -invariant and multilinear in X_1, \dots, X_d , and we can recover p by the formula

$$(V.14.4) \quad d! f([X], [Y], \dots) = f_{(1)}([X], [X], \dots, [X], [Y], \dots).$$

Doing this for all the variables shows that we can replace p by a multilinear polynomial. \square

Now a multilinear polynomial can be interpreted either as a multilinear map $\phi : \bigoplus E \longrightarrow k$, or as a linear map $E^{\otimes m} \longrightarrow k$. And since $E^{*\otimes m} \approx E^{\otimes m}$, we can also regard ϕ as an element of $E^{\otimes m}$. As we have seen, the operation by $p \in GL_n$ on $E^{\otimes m}$ is conjugation by $p^{\otimes m}$. So (V.13.4) the invariant elements are the elements in the centralizer of $\text{Sym}^m(E)$, which is the group algebra \mathcal{B} . The group algebra is generated by permutations. To prove the theorem, it suffices to show that the linear map $\phi_\sigma : \bigoplus E \longrightarrow k$ defined by a permutation σ can be expressed in terms of traces. Moreover, it suffices to do this for a cyclic permutation, say $\sigma = (1\ 2 \dots r)^{-1}$.

Lemma V.14.5. *Let $\sigma = (1\ 2 \dots r)^{-1}$. The multilinear map $\bigoplus^m E \longrightarrow k$ defined by σ is*

$$\phi_\sigma([X_1], \dots, [X_m]) = \text{trace}(X_1 \cdots X_r).$$

Proof. Since both sides are multilinear, it suffices to verify this formula when X_ν are matrix units.

We write σ as element of $E^{\otimes m}$ using the formula (V.13.5). The associated linear map $\phi_\sigma : E^{\otimes m} \longrightarrow k$ is given by the trace pairing, which is

$$(V.14.6) \quad \langle X_1 \otimes \cdots \otimes X_m, Y_1 \otimes \cdots \otimes Y_m \rangle = \langle X_1, Y_1 \rangle \cdots \langle X_m, Y_m \rangle,$$

and where $\langle X, Y \rangle = \text{trace}(XY)$. Set $X_1 \otimes \cdots \otimes X_m = e_{\xi_1 \eta_1} \otimes \cdots \otimes e_{\xi_m \eta_m}$. With σ written as in (V.13.5),

$$(V.14.7) \quad \langle \sigma, X_1 \otimes \cdots \otimes X_m \rangle = \sum_{(i)} \left(\text{trace}(e_{\sigma i_\nu i_\nu} e_{\xi_\nu \eta_\nu}) \right).$$

For fixed (i) , the term on the right side is zero unless $(i) = (\xi)$. If $(i) = (\xi)$, it is

$$\prod_\nu \text{trace}(e_{\sigma(\xi_\nu)\eta_\nu})$$

which is zero if $\sigma(\xi) \neq (\eta)$ and 1 if $\sigma(\xi) = (\eta)$. So

$$(V.14.8) \quad \langle \sigma, X_1 \otimes \cdots \otimes X_m \rangle = \delta_{\sigma(\xi)\eta}.$$

When σ is the cyclic permutation $(1\ 2 \cdots r)^{-1}$, this yields 1 if $\xi_{i-1} = \eta_i$ for $i = 1, \dots, r$ and is zero otherwise. Evaluating $\text{trace}(e_{\xi_1 \eta_1} \cdots e_{\xi_r \eta_r}) = 1$ leads to the same result. \square

The second part of Procesi's theorem follows from the main theorem and from next two lemmas.

Lemma V.14.9. *Let ρ be an n -dimensional representation of the ring A of generic matrices. The associated semisimple representation ρ^{ss} is in the closure of the orbit of ρ .*

Proof by example. Say that ρ is a reducible 2-dimensional representation. Then with a suitable choice of basis, the representation will have the triangular form

$$(V.14.10) \quad \rho(u) = \begin{pmatrix} a(u) & b(u) \\ 0 & d(u) \end{pmatrix},$$

for $u \in A$. Let

$$p = \begin{pmatrix} t & 0 \\ 0 & 1 \end{pmatrix}.$$

then

$$(V.14.11) \quad p\rho p^{-1} = \begin{pmatrix} a(u) & tb(u) \\ 0 & d(u) \end{pmatrix}.$$

For every nonzero element $t \in k$, the representation $p\rho p^{-1}$ is in the orbit of ρ . Therefore the representation obtained by setting $t = 0$ in V.14.11 is in the closure of this orbit. \square

Lemma V.14.12. *Two semisimple n -dimensional representations ρ and ρ are equivalent if $\text{trace}(\rho(u)) = \text{trace}(\rho(u))$ for every $u \in A$.*

Proof. Let $\bar{\rho}_1, \dots, \bar{\rho}_r$ denote the equivalence classes of distinct simple representations which occur as factors of ρ and ρ , and say that $\dim \bar{\rho}_\nu = d_\nu$. These simple representations correspond to distinct maximal ideals $\mathfrak{m}_1, \dots, \mathfrak{m}_r$ of A , such that the map $A \rightarrow A/\mathfrak{m}_\nu \approx M_{d_\nu}(k)$ is the representation $\bar{\rho}_\nu$. Unless ρ and ρ are equivalent, at least one of the representations $\bar{\rho}_\nu$, say $\bar{\rho}_1$, occurs with a different multiplicity in them.

By the Chinese remainder theorem, the map $A \rightarrow \prod A/\mathfrak{m}_\nu$ is surjective. So there is an element $a \in A$ such that $\bar{\rho}_1(a) = 1$ and $\bar{\rho}_\nu(a) = 0$ for $\nu \neq 1$. The trace of this element distinguishes the two representations. \square

V.15. The spectrum of the ring of generic matrices

As in the previous sections, we assume that k is algebraically closed. If A is a finitely generated PI algebra over k , $\text{Spec } A$ will denote the space of its maximal ideals. We know that if \mathfrak{m} is a maximal ideal, then A/\mathfrak{m} is isomorphic to a matrix algebra $M_r(k)$ of bounded rank. We define $\text{Spec}_r A$ to be the subset of $\text{Spec } A$ of maximal ideals such that A/\mathfrak{m} is isomorphic to $M_r(A)$. Then

$$(V.15.1) \quad \text{Spec } A = \text{Spec}_1 A \cup \text{Spec}_2 A \cup \dots \cup \text{Spec}_n A,$$

for suitable n . If A is a prime ring, then n is the PI degree of A , the largest integer such that the standard identity $S_{2n} = 0$ holds.

Because they are defined by the identities $S_{2r} \equiv 0$, the subsets $\text{Spec}_{\leq r} A = \text{Spec}_1 A \cup \dots \cup \text{Spec}_r A$ are closed in $\text{Spec } A$ in the Zariski topology. Consequently $U = \text{Spec}_n A$, being the complement of a closed set, is open in $\text{Spec } A$.

We now describe the intrinsic structure of the set U . Let's suppose for simplicity that A is a prime ring of PI degree n , and let γ be a nonzero evaluation of the Razmyslov polynomial. Then (V.9.3) $A[\gamma^{-1}]$ is an Azumaya algebra of rank n^2 over its center.

Next, $U = \text{Spec } A$ can be identified with the open subset of $U = \text{Spec}_n A$ consisting of the maximal ideals which do not contain γ . Moreover, for every maximal ideal $\mathfrak{m} \in U$, there is an evaluation γ of the Razmyslov polynomial which is not in \mathfrak{m} (V.9.9). This shows that U can be covered by open subsets U which are spectra of Azumaya algebras.

If γ_1 and γ_2 are two evaluations of the Razmyslov polynomial, with corresponding Azumaya algebras A_1 and A_2 , then the intersection of the open sets U_1 and U_2 is the spectrum of the common localization $A_{12} = A_1[\gamma_2^{-1}] = A_2[\gamma_1^{-1}]$.

Now a central localization of an Azumaya algebra is just given by a localization of the center. The common localizations A_{12} described above defines data for gluing the spectra $V_i = \text{Spec } Z_i$ of the centers A_i . In this way one constructs a scheme V which is covered by affine opens of the form $\text{Spec } Z$. In general, this scheme will not be affine itself. However, it comes equipped with a coherent sheaf of Azumaya algebras \mathcal{A} over its structure sheaf \mathcal{O}_V , the sheaf whose algebra of sections on the open set V is A . We may think of V as the "center" of U .

Recall that the two sided ideals of an Azumaya algebra A correspond bijectively to two sided ideals of its center Z , because two-sided ideals are modules over $A^o \otimes A \approx \text{End}_Z A$, which is a ring Morita equivalent to Z . Because of this, U and V are homeomorphic, *as topological spaces*. They are two incarnations of the space of irreducible n -dimensional representations. However, *as schemes*, one is commutative and the other is not.

The subsets $\text{Spec}_r A$ can be described similarly.

Though the above remarks don't depend on the characteristic of k , we need to assume that the characteristic is zero to continue this discussion. Let $A = k\langle X_1, \dots, X_n \rangle$ be the ring of generic $n \times n$ matrices, and let $R = k[x'_{ij}]$ be the polynomial ring on the matrix entries, as before. Then the space $\text{Spec}_n A$ has yet a third incarnation, namely as the subset Y_n of the spectrum $Y = \text{Spec } R^G$ of the invariant ring which corresponds to the irreducible representations.

Proposition V.15.2. *Let $Y = \text{Spec } R^G$ be the spectrum of the invariant ring, and let $Y_n \subset Y$ be the subset of equivalence classes of irreducible representations. Then Y_n is an open subscheme of Y , and it is isomorphic to the scheme V described above.*

Proof. An evaluation γ of the Razmyslov polynomial is central, hence it is a G -invariant element of A : For any $\alpha, \beta \in A$, $c(\alpha, \beta) = pc(\alpha, \beta)p^{-1} = c(p\alpha p^{-1}, p\beta p^{-1})$. It is also an element of R , hence it is in R^G . The set of common zeros of these evaluations in Y is a closed subset, and Y_n is the complement of that set. So Y_n is open in Y .

Let γ be an evaluation of the Razmyslov polynomial, and let $A = A[\gamma^{-1}]$ be the corresponding Azumaya algebra. The trace of any element $\alpha \in A$ is in its center Z (B.12.4). Since R^G is generated by traces, we obtain a canonical homomorphism $R^G \rightarrow Z$. These canonical maps glue to define a morphism $V \rightarrow Y$. Since both V and Y_n correspond bijectively to irreducible n -dimensional representations, the map carries V bijectively to the open subscheme Y_n .

The center Z is a domain because A is a prime ring. Hence V is reduced and irreducible. Moreover, R^G is integrally closed (V.10.7) hence Y_n is a normal scheme. Now the fact that the map $V \rightarrow Y_n$ is an isomorphism follows from Zariski's main theorem. \square

What remains to be done here is to understand the relation between the whole spectrum $\text{Spec } A$ of the ring of generic matrices, and the spectrum of the invariant ring $\text{Spec } R^G$. Let $U_r = \text{Spec}_r A$ for $r \leq n$, and let V_r be the center of U_r , constructed as for the case $r = n$ above. Then as topological spaces, $U_r \approx V_r$.

We also have $\text{Spec } A = U_1 \cup \dots \cup U_n$, which, set-theoretically, is in bijective correspondence with the disjoint union union

$$(V.15.3) \quad V = V_1 \cup \dots \cup V_n.$$

(I don't like to write $U \approx V$ because U is a noncommutative scheme of Azumaya algebras whose center is V .)

On the other hand, the scheme Y is the space of semisimple n -dimensional representations of A . This space is closely related to V , but it is not the same space. When $n = 2$, a semisimple 2-dimensional representation, if reducible, is the sum of two (possibly equivalent) 1-dimensional

representations. It corresponds to an unordered pair of points of V_1 , i.e., a point of the symmetric square $S^2(V_1)$. So in this case

$$(V.15.4) \quad Y = Y_{1,1} \cup Y_2 \approx S^2(V_1) \cup V_2.$$

For $n = 3$,

$$(V.15.5) \quad Y = Y_{1,1,1} \cup Y_{1,2} \cup Y_3 \approx S^3(V_1) \cup (V_1 \times V_2) \cup V_3.$$

Exercise: Work out the case $n = 4$.

VI. GROWTH OF ALGEBRAS

VI.1. Growth functions

In these notes, k is a field. By algebra, we will mean k -algebra unless otherwise specified.

Let A be a finitely generated algebra over k . If V is a subspace of A , we denote by V^n the subspace spanned by all products of elements of V of length n . By *generating subspace* V , we will mean a finite-dimensional subspace of A which generates A as algebra, and which contains 1. Then V^n is spanned by the products of length $\leq n$ of elements of V , and because V generates, $A = \bigcup V^n$.

The *growth function* associated to a generating subspace is

$$(VI.1.1) \quad f(n) = f_V(n) := \dim_k V^n.$$

An equivalent definition is to take a finite subset X of A which generates A and which contains 1. Let X^n denote the set of products of length n of elements of X . As before, this includes all shorter products. Let V^n be the vector space spanned by X^n . Then we set $f_X(n) = \dim V^n$, which is the number of linearly independent products of length $\leq n$ of elements of X .

An algebra A is said to have *polynomial growth* if there are positive real numbers c, r such that

$$(VI.1.2) \quad f(n) \leq cn^r$$

for all n . We will see below (1.5) that this is independent of the choice of the generating subspace V . And A has *exponential growth* if there is a real number $s > 1$ so that $f(n) \geq s^n$ for infinitely many n .

The *Gelfand-Kirillov dimension* of an algebra with polynomial growth is the infimum of the real numbers r such that (1.2) holds for some c :

$$(VI.1.3) \quad \text{gk}(A) = \inf \{ r \mid f(n) \leq cn^r \}.$$

If A does not have polynomial growth, then $\text{gk}(A) = \infty$. Again, (1.5) shows that the GK dimension is independent of the choice of the subspace V .

Exercise: Show that $\text{gk}(A) = \inf \{ r \mid f(n) \leq p(n) \}$ for some polynomial p of degree r .

In contexts in which the generating set is given naturally, for example if the algebra is graded, we may also want to keep track of the constant coefficient c . In that case it would be more natural to ask for the smallest c such that (1.2) holds it for large n . Such a modification would not change r .

Exercise: Show that $\text{gk}(A) = \inf \{ r \mid f(n) \leq n^r \text{ for large } n \}$.

Examples VI.1.4. 1. The commutative polynomial ring $A = k[x_1, \dots, x_d]$ has Gelfand-Kirillov dimension d . If we let V be the space spanned by $\{1, x_1, \dots, x_d\}$, then V^n is the space of polynomials of degree $\leq n$. The dimension of this space is $\binom{n+d}{d}$, which is a polynomial of degree d in n .

2. If $d > 1$, the free ring $A = k\langle x_1, \dots, x_d \rangle$ has exponential growth, hence infinite GK dimension, because the number of noncommutative monomials of degree n is d^n .

Exercise: Let A be a finitely generated, commutative domain of Krull dimension d . Prove that $\text{gk}(A) = d$.

The next proposition shows, as promised, that the Gelfand-Kirillov dimension is independent of the choice of the generating space V .

Proposition VI.1.5. *Let A be a finitely generated k -algebra, let V, W be generating subspaces, and set $f = f_V, g = f_W$. If $f(n) \leq cn^r$ for all n , then there is a c such that $g(n) \leq c n^r$ for all n .*

Proof. Assume that $f(n) \leq cn^r$. Since $A = \bigcup V^n$ and W is finite-dimensional, $W \subset V^s$ for some s . Then $W^n \subset V^{sn}$, hence $g(n) \leq f(sn) \leq cs^n n^r = c n^r$. \square

For comparison purposes, it would be useful to have a lower bound on the growth of f_V . The only general one that I know is the following:

Proposition VI.1.6. *Suppose that A is finitely generated but not finite dimensional over k . Let $f = f_V$ be a growth function for A . Then f is a strictly increasing function, hence $f(n) \geq n$.*

Proof. We have $f(n+1) \geq f(n)$ because $V^{n+1} \supset V^n$. If $f(n+1) = f(n)$ for some n , then $VV^n = V^n$, and this implies that $V^{n+m} = V^n$ for all $m \geq 0$. Then $A = \bigcup V^{n+m} = V^n$, so A is finite dimensional. \square

Corollary VI.1.7. *Let A be a finitely generated k -algebra. Then $gk(A) = 0$ if and only if A has finite dimension over k . If A is infinite dimensional, then $gk(A) \geq 1$. \square*

The next three propositions derive elementary properties of the GK dimension.

Proposition VI.1.8. *Let $A[t]$ denote the algebra of polynomials in t with coefficients in A , where t is a central variable. Then $gk(A[t]) = gk(A) + 1$.*

Proof. Let V be a subspace of A which generates A and contains 1. The subspace W spanned by V and t generates $A[t]$, and $W^n = V^n \oplus V^{n-1}t \oplus \dots \oplus Vt^{n-1} \oplus kt^n$. Suppose that $f(n) \leq cn^r$. Then $f_W(n) = f(0) + \dots + f(n) \leq c(1^r + \dots + n^r) \leq c n^{r+1}$. Thus $gk(A[t]) \leq gk(A) + 1$.

On the other hand, $W^{2n} \supset V^n t^n \oplus V^{n-1} t^{n-1} \oplus \dots \oplus V^n t^0$, hence $\dim W^{2n} \geq n \dim V^n$. Then if $\dim V^n \geq cn^r$ for infinitely many n , $\dim W^{2n} \geq cn^{r+1}$ for infinitely many n too. This shows that $gk(A[t]) \geq gk(A) + 1$. \square

Proposition VI.1.9. *If γ is a regular central element of a finitely generated algebra A , then $gk(A) = gk(A[\gamma^{-1}])$. The same is true if γ is a normalizing element.*

Proof. Let V be a generating subspace of A , and let $W = \text{Span}\{V, \gamma^{-1}\}$, which is a generating subspace for $A[\gamma^{-1}]$. Then $V^n \subset W^n$ and $\gamma^n W^n \subset V^{2n}$. So $f_V(n) \leq f_W(n) \leq f_V(2n)$. The assertion follows. \square

Proposition VI.1.10. *Let A be a finitely generated graded algebra. Let $r = gk(A)$, and let*

$$s = \inf\{s \mid \dim A_n \leq cn^s\}.$$

(i) $r \leq s + 1$.

(ii) *If there is an integer $d > 0$, such that $\dim A_n \leq \dim A_{n+d}$ for all sufficiently large n , then $r = s + 1$.*

(iii) *The condition of (ii) is satisfied in each of the following cases:*

(a) *$\dim A_n$ is monotonic,*

(b) *$\dim A_{kd}$ is a polynomial function of k , or*

(c) *A contains a homogeneous left or right regular element of degree d .*

Proof. Let $a_n = \dim_k A_n$. We may take as generating subspace for A the space $V = A_0 + \cdots + A_k$ for some sufficiently large k . Then $V^n = A_0 \oplus \cdots \oplus A_{nk}$, and $f_V(n) = a_0 + a_1 + \cdots + a_{nk}$. If $a_n \leq cn^s$, then $\dim V^n \leq \sum_{i=0}^n kci^s \leq cn^{s+1}$. This proves (i).

Suppose that the hypothesis of (ii) is satisfied. Replacing d by a multiple and increasing k , we arrange things so that $d = k \geq 2$. Then

$$V^{2n} = A_0 \oplus \cdots \oplus A_{2nd} \supset A_{n+d} \oplus A_{n+2d} \oplus \cdots \oplus A_{n+nd},$$

hence $\dim V^{2n} \geq na_n$. If $s < s'$, then $a_n > cn^{s'}$ for infinitely many n , and then $f_V(2n) \geq cn^{s'} = c(2n)^{s'}$. Hence $\text{gk}(A) > s + 1$. The proof of (iii) is easy. \square

VI.2. Warfield's theorem

Theorem VI.2.1. (Warfield) *For any real number $r \geq 2$, there is a finitely generated graded algebra R with $\text{gk}(R) = r$.*

Proof. Because of Proposition VI.1.8, it suffices to construct a ring R with $\text{gk}(R) = r$ when $2 \leq r \leq 3$.

Let $F = k\langle x, y \rangle$ be the free ring on two generators, and let I denote the two-sided ideal of F generated by the element y . Let $A = F/I^3$ and $B = F/I^2$. Then A_n has a monomial basis consisting of the monomials of degree n which have ≤ 2 in y . There are $\binom{n}{k}$ monomials of degree k in y . Thus

$$(VI.2.2) \quad \dim A_n = \binom{n}{2} + \binom{n}{1} + \binom{n}{0} = \frac{1}{2}(n^2 + n + 2),$$

and similarly,

$$(VI.2.3) \quad \dim B_n = n + 1.$$

So $\text{gk}(A) = 3$ and $\text{gk}(B) = 2$. The algebra we construct will have the form $R = F/J$, where J is a monomial ideal, and $I^3 \subset J \subset I^2$. To define J , we must decide which elements $x^i y x^j y x^k$ from I^2 to put into J_n .

Let q be a real number with $0 \leq q \leq 1$. Set $u_n = [n^q]$ and $v_n = n - u_n$, where $[c]$ denotes the integer part of c . Warfield's theorem follows from the next lemma:

Lemma VI.2.4. (i) *The sequences u_n and v_n are monotonic.*

(ii) *Let J_n denote the span of all monomials of degree n which have degree ≥ 3 in y , together with the monomials of the form $x^i y x^j y x^k$ with $j < v_n$. Then $J = \bigoplus J_n$ is an ideal of the free ring F . If $A = F/J$, then $\dim A_n = \binom{u_n}{2} + n + 1$, this dimension is monotonic, and $\text{gk}(A) = 2 + q$.*

Proof. (ii) That J is an ideal follows from the fact that v_n is monotonic. The monomials of degree n not in J_n are those of degree ≤ 1 in y , together with the monomials $x^i y x^j y x^k$ with $j \geq v_n$. The number of these monomials is $\binom{u_n}{2} + n + 1$. This verifies the formula for $\dim A_n$, and monotonicity follows from the fact that u_n is monotonic. The value $\text{gk}(A) = q + 2$ follows from Proposition VI.1.10. \square

Note that the definition of J requires only that u_n and v_n be monotonic. So within these bounds, any function is possible. This illustrates the fact that, without some hypotheses on the algebra, growth functions can be pretty arbitrary.

VI.3. Path algebras

Let Γ be a finite oriented graph, and let its edges, the *arrows*, be labeled x_1, \dots, x_m . The *path algebra* A is the graded algebra with basis the oriented paths in Γ . The product uv of two paths is defined to be the composed path $u \cdot v$ if the paths can be composed, i.e., if v starts where u ends. If the paths can not be composed, then $uv = 0$.

The algebra is graded by the *length* of a path, the number of its arrows: A_n is the vector space spanned by paths of length n .

It is customary to adjoin paths of length zero which represent the vertices of the graph. They are idempotent elements in the path algebra A , and their sum is 1.

For example, the paths in the graph $\circ \xrightarrow{x_1} \circ \xrightarrow{x_2} \dots \xrightarrow{x_m} \circ$ are the words $x_i x_{i+1} \dots x_j$ for $1 \leq i \leq j \leq m$. This path algebra is the algebra of upper triangular $m+1 \times m+1$ matrices, with $x_i = e_{i-1, i}$.

The path algebra is finite dimensional if the graph Γ has no oriented loops, or *cycles*. But a cycle gives us a path, say u , whose powers u^n are all distinct and nonzero. The next proposition shows how the GK dimension of the path algebra can be read off from the geometry of the graph.

Proposition VI.3.1. *Let A be the path algebra of a finite oriented graph Γ .*

- (i) *If Γ contains two cycles which have a vertex in common, then A has exponential growth.*
- (ii) *Suppose that Γ contains r cycles, none of which have a vertex in common. Then $\text{gk}(A) \leq r$. Moreover, $\text{gk}(A) = r$ if and only if there is a path that traverses all of the cycles.*

Sketch of the proof. (i) Say that a vertex is in common to two distinct cycles, and let u, v be the paths which traverse these loops, starting and ending at a common vertex. Then the words in u, v represent distinct paths. In fact, they represent distinct elements in the fundamental group of the graph. So A contains the free ring $k\langle u, v \rangle$.

(ii) Suppose for instance that there are exactly two distinct cycles, and say that u, v are paths which traverse them. There may or may not be some paths y connecting u to v , i.e., such that $uyv \neq 0$. But if such a path exists, then there can be no path in the opposite direction, because if $vzu \neq 0$, then yz would be a cycle yz having a vertex in common with u . If y exists, then the paths which can be built using u, y, v are $u^i y v^j$. Since we can not return to u from v , every path has the form $w u^i y v^j w$, where each of the subpaths w, w, y is a member of a finite set. This leads to very regular quadratic growth. \square

VI.4. Bergman's Gap Theorem

Theorem VI.4.1. *(Bergman) There is no algebra A whose GK dimension is in the open interval $1 < r < 2$. If A is a finitely generated algebra of GK dimension 1, then every growth function $f = f_V$ for A is bounded by a linear function, i.e., $f(n) \leq cn$ for some c .*

Let $\{x_1, \dots, x_m\}$ be indeterminates. We'll refer to a monomial w in $\{x_\nu\}$ also as a *word*, and we will use notation of the form $w = a_1 \cdots a_r$, where each a_i denotes one of the variables x_ν . A *subword* of a word w is a word v such that $w = yvz$ for some words y, z . The *length* of a word w , the number of its letters, will be denoted by $|w|$.

An *infinite word* is a sequence of letters indexed by \mathbb{Z} . A *period* of an infinite word $h = \{a_n\}$ is an integer p such that $a_n = a_{n+p}$ for all $n \in \mathbb{Z}$. The periods of h form a subgroup of \mathbb{Z}^+ , and if h has a non-zero period, then the positive generator for this subgroup is called the *minimal period* of h .

A (finite) word w will also be called *periodic*, of period p , provided that $|w| \geq p$ and that $a_i = a_{i+p}$ for all relevant indices. Thus *abcabca* has period 3. The periodic infinite word $h = \{a_n\}$ with period p and having $w = a_1 \cdots a_s$ as a subword is uniquely determined by w . It is the sequence defined by $a_{pk+r} = a_r$ for $r = 1, \dots, p$.

If an infinite word has period p , then verifying that it also has another period q requires checking periodicity on a sufficiently large finite subword. The next lemma determines the length of that subword.

Lemma VI.4.2. *Let $h = \{a_n\}$ be an infinite word which has a period p , and let q be an integer $< p$. Suppose that a subword w of length $\geq p+q-1$ is periodic of period q . Then q is also a period of h .*

Proof. We may assume $w = a_1 \cdots a_{p+q-1}$. Because w has period q , any subword of length q is a complete cycle, and it differs from the initial subword $u = a_1 \cdots a_q$ by a cyclic permutation. The string $v = a_p \cdots a_{p+q-1}$ is such a subword.

In the infinite word h , $u = a_1 \cdots a_q = a_{p+1} \cdots a_{p+q}$ because h has period p . So u and v have a common subword $x = a_1 \cdots a_{q-1} = a_{p+1} \cdots a_{p+q}$. Then since both u and v are complete q -cycles, the remaining letters of u and v , which are $a_q = a_{p+q}$ and a_p , must also be equal. The equality $a_p = a_{p+q}$ shows that the periodicity of w extends one letter to the right, to the word $w = a_1 \cdots a_{p+q}$. By induction, it extends indefinitely. Inverting the order shows that the periodicity also extends to the left. \square

The (graded) *lexicographic order* on monomials: If v, w are words, then $v < w$ if either $|v| < |w|$, or if $|v| = |w|$ and v appears earlier in the dictionary.

Lemma VI.4.3. *Let y, u, u, z be words. Then $yz < yuz$ if and only if $u < u$. \square*

Let A be a finitely generated algebra which is presented as a quotient of the free algebra $F = k\langle x_1, \dots, x_n \rangle$, say $A = F/I$. The *standard monomial basis* for A is the lowest basis in lexicographic order. It is obtained recursively: We assign degree 1 to the variables x_i . Let $V = \text{Span}(1, x_1, \dots, x_m)$. We take the smallest basis W_{n-1} for V^{n-1} , and add monomials of degree n as needed, in lexicographic order, to obtain a basis W_n of V^n . Then $W = \bigcup(W_n)$ is the standard monomial basis for A . Thus the standard monomial basis for the commutative polynomial ring $k[x, y]$ is $\{x^i y^j\}$.

Lemma VI.4.4. *A set W of monomials is a standard monomial basis for a quotient algebra $A = F/I$ if and only if every subword of a word in W is also in W . If so, then the set N of words not in W spans a monomial ideal J , and W is the standard monomial basis for the graded algebra F/J . \square*

So to measure the growth of an arbitrary algebra $A = F/I$, we may replace it by the graded algebra defined by a suitable monomial ideal J . Unfortunately, the property that I is finitely generated, if it holds, may not carry over to the monomial ideal J .

We now examine more closely the case of a monomial ideal J which is generated by words of some fixed length d (and hence is a finitely generated ideal).

Lemma VI.4.5. *Let S be an arbitrary set of words of length d , let J be the ideal generated by the words of length d not in S , and let $A = F/J$. The standard monomial basis for A consists of all words w such that every subword of w of length d is in S . \square*

Let S be a set of words of length d . We form an oriented graph Γ whose vertices are the elements of S and whose edges, the *arrows* are as follows: If $u, v \in S$, there is an arrow $u \rightarrow v$ if $ub = av$ for some letters a, b . This is equivalent with saying that there is a common subword x of u and v of length $d - 1$, such that $u = ax$ and $v = xb$. So $u = a_1 \cdots a_d$, $v = a_2 \cdots a_{d+1}$, and $x = a_2 \cdots a_d$, where $a = a_1$, $b = a_{d+1}$. Since a and b are uniquely determined by u and v , there is at most one arrow $u \rightarrow v$.

Given an oriented path $u_0 \rightarrow u_1 \rightarrow \cdots \rightarrow u_r$ in Γ of length r , we form a word as follows. We write $u_0 = a_1 \cdots a_d$, $u_1 = a_2 \cdots a_{d+1}$, ..., $u_r = a_{r+1} \cdots a_{d+r}$. Then the word $w = a_0 \cdots a_{d+r-1}$ of length $d + p$ has the property that every subword of length d is in S .

Lemma VI.4.6. *Let W be the standard monomial basis for the above algebra A . The oriented paths of length n in Γ correspond bijectively to elements of W_{n+d} . \square*

Let $u = a_1 \cdots a_d$ be a word with minimal period $p \leq d$. Then we obtain an oriented *cycle* of length p in Γ :

$$(VI.4.7) \quad u = u_0 \rightarrow u_1 \rightarrow \cdots \rightarrow u_p = u$$

as follows: We take the infinite periodic word $h = \{a_n\}$ of which u is a subword, and we define

$$(VI.4.8) \quad u_0 = a_1 \cdots a_d, \quad u_1 = a_2 \cdots a_{p+1}, \quad \dots, \quad u_p = a_{p+1} \cdots a_{p+d} = u_0.$$

We call u_i the *translates* of u . For instance, let $u = abcabca$. The period is 3, and the cycle is $u_0 \rightarrow u_1 \rightarrow u_2 \rightarrow u_0$, where $u_0 = u$, $u_1 = bcabcab$, and $u_2 = cabcab$.

Lemma VI.4.9. *The oriented cycles of length p in Γ are determined by words u of period $p \leq d$ whose translates u_i are in W for all i .*

Proof. Let $u_0 \rightarrow u_1 \rightarrow \cdots \rightarrow u_p = u_0$ be a cycle, where $u_0 = a_1 \cdots a_d$, and where $u_i \neq u_0$ for $1 < i < p$. Then $u_1 = a_2 \cdots a_d b_1$, $u_2 = a_3 \cdots a_d b_1 b_2$, and $u_p = a_{p+1} \cdots a_d b_1 \cdots b_p = u_0$, for suitable elements b_i taken from among the elements a_1, \dots, a_p . Since $u_0 = u_p$, we have $a_1 \cdots a_{d-p} = a_{p+1} \cdots a_d$. Thus u is periodic of period p as claimed. \square

The next proposition, which is analogous to Proposition VI.3.1, shows that the structure of the graph Γ reflects the GK dimension of the algebra A .

Proposition VI.4.10. *Let Γ denote the graph of an algebra A defined by monomial relations of degree d .*

- (i) *If Γ contains two cycles which have a vertex in common, then A has exponential growth.*
- (ii) *Suppose that Γ contains r cycles, none of which have a vertex in common. Then $gk(A) \leq r$. Moreover, $gk(A) = r$ if and only if there is a path that traverses all of the cycles. \square*

Theorem VI.4.11. *If $|S| \leq d$, an oriented path in Γ can traverse at most one cycle. It can traverse that cycle a finite number of times in succession, but it can not leave the cycle and then return to it.*

Proof. The method is to show that any path which traverses two cycles has length at least $d + 1$, unless the two cycles are equal and are traversed in immediate succession.

We may assume that our path starts at a cyclic word u of period p , traverses the corresponding cycle C_u to return to u , then proceeds to another cyclic word v of period q along a path P , and finally traverses the cycle C_v . There are p vertices on the cycle C_p and q vertices on C_q , and the endpoints of the path P are among them. So we must show that the length s of P is at least $d - p - q + 2$. We may assume that $q \leq p$. The other case is treated by passing to the opposite graph.

If $w_1 \rightarrow w_2$ is an arrow in Γ , then w_2 is obtained from w_1 by removing the first letter, and putting some other letter at the end. To obtain v from u via the path P , we repeat this process s times. So if $u = a_1 \cdots a_d$, then $v = a_{s+1} \cdots a_d b_1 \cdots b_s$ for some b_i . Thus, in the notation of VI.4.8, the word v is obtained from u_s by replacing the last s letters. By hypothesis, u and u_s have period p , while v has period q .

If $s < d - p + q + 2$, then $d - s \geq p + q - 1$. We may embed u into an infinite periodic word h and apply Lemma VI.4.2. The lemma shows that u has period q . Since $q \leq p$ and p is the minimal period of u , we conclude that $q = p$. The first segment of our path P leaves the cycle determined by u , hence it replaces $u = a_1 \cdots a_d$ by $a_2 \cdots a_d b$, where in the notation of VI.4.8, $b \neq a_{d+1}$. This breaks the symmetry of u , and because $a_{d+1} = a_{d+1-p}$, there is no way to obtain a periodic word without replacing the letter a_{d+1-p} . This requires a path of length at least $d - p + 1$. \square

Proof of Bergman's theorem. Let A be a finitely generated algebra with $1 \leq gk(A) < 2$, let $\{x_1, \dots, x_m\}$ be a generating subset. Lemma VI.4.4 shows that the standard monomial basis W for A is also the monomial basis for an algebra defined by a monomial ideal J . So we may assume that A is the algebra $A = F/J$, in particular that A is graded and generated by elements of degree 1. Then $V = \text{Span}(1, x_1, \dots, x_m)$ is a generating subspace.

If $a_n = \dim A_n \geq n$ for all n , then $\dim V^n = a_0 + \cdots + a_n \geq \binom{n+1}{2}$, and $gk(A) \geq 2$. Since this is not the case, $\dim A_d \leq d$ for some d . This is the only property of the algebra A that we use in what follows.

Let S denote the set of words of degree d which are in the standard monomial basis W , and let T be the words of length d not in S . Thus T is a subset of the monomial ideal J . Let J be the ideal generated by T , and let $A = F/J$. Then because $T \subset J$, A is a quotient of A , and $\dim A_n \geq \dim A_n$, while $A_d = A_d$. So we are reduced to considering the algebra discussed in Theorem VI.4.11, and the fact that a path can traverse only one loop shows that the growth is linear VI.4.10. \square

VI.5. Theorem of Stephenson and Zhang

Theorem VI.5.1. (*Stephenson – Zhang*) *A connected graded, right noetherian k -algebra A has subexponential growth.*

A sequence $f = \{f(n)\}$ of real numbers has exponential growth if, for some $s > 1$, there are infinitely many n so that $f(n) > s^n$. A function which does not have exponential growth is said to have *subexponential growth*.

To determine exponential growth, it is convenient to use the measure $\overline{\lim} f(n)^{1/n}$. The growth is exponential if and only if

$$(VI.5.2) \quad \overline{\lim} f(n)^{1/n} > 1.$$

Lemma VI.5.3. *Let A be a finitely generated graded algebra. Let $V = A_0 \oplus \cdots \oplus A_k$ be a generating subspace, with growth function $f(n)$, let $a(n) = \dim_k A_n$ and $s(n) = a(0) + \cdots + a(n)$. The assertions of exponential growth of the sequences $a(n)$, $s(n)$, and $f(n)$ are equivalent.*

Proof. Let $\bar{a}(n)$ denote the maximum value of $a(i)$ for $i \leq n$. Then $a(n) \leq \bar{a}(n) \leq s(n) \leq (n+1)\bar{a}(n)$. The sequences \bar{a} and s are monotonic, and unless the sequence $a(n)$ is bounded, $a(n) = \bar{a}(n)$ infinitely often. Then

$$\overline{\lim}(a(n))^{1/n} \leq \overline{\lim}(s(n))^{1/n} \leq \overline{\lim}(n+1)^{1/n}(\bar{a}(n))^{1/n} = \lim(\bar{a}(n))^{1/n} = \overline{\lim}(a(n))^{1/n}.$$

Finally, $f(n) = s(nk)$. Since $s(n)$ is monotonic, $(\overline{\lim}(f(n))^{1/n})^k = \overline{\lim}(s(n))^{1/n}$. \square

Proof of the theorem. Suppose that A has exponential growth. By Lemma VI.5.3, there is a real number $\alpha > 1$ so that $a(n) > \alpha^n$ infinitely often. To show that A is not right noetherian, we use the next lemma to construct an infinite increasing chain of right ideals.

Lemma VI.5.4. *There is a sequence of integers $0 < r_1 < r_2 < \cdots$ such that for every k ,*

$$(VI.5.5) \quad a(r_k) > \sum_{i=1}^{k-1} a(r_k - r_i),$$

In particular, $a(r_1) > 0$.

Suppose that this sequence has been found. Then we construct a chain of right ideals as follows: We choose $y_1 \in A_{r_1}$, with $y_1 \neq 0$, and we set $I_1 = y_1 A$. This is possible because $a(r_1) > 0$. Then because $a(r_2) > a(r_2 - r_1)$, it follows that $y_1 A_{r_2 - r_1} \neq A_{r_2}$, and $A_{r_2} \not\subset I_1$. We choose $y_2 \in A_{r_2}$ and not in I_1 , and we set $I_2 = I_1 + y_2 A$. Then $I_1 < I_2$. Next, because $a(r_3) > a(r_3 - r_1) + a(r_2 - r_2)$, it follows that $y_1 A_{r_3 - r_1} + y_2 A_{r_3 - r_2} \neq A_{r_3}$, and $A_{r_3} \not\subset I_2$. We choose $y_3 \in A_{r_3}$ not in I_2 , and we set $I_3 = I_2 + y_3 A$, so that $I_2 < I_3$, and so on. Continuing in this way, we obtain the required strictly increasing sequence of right ideals. \square

Proof of the lemma. We will construct the sequence r_i with the additional property that $\alpha^{r_i} > 2^i$. This simply requires choosing each r_i large enough. We can find r_1 . So suppose that $r_1 < r_2 < \cdots < r_{k-1}$ have been found.

Sublemma VI.5.6. *There are infinitely many integers n so that, for every $i < k$, $a(n) > a(n - r_i)\alpha^{r_i}$.*

Proof. Suppose the contrary: For every $n > n_0$, there is an i with $1 \leq i \leq k - 1$, so that $a(n) \leq a(n - r_i)\alpha^{r_i}$. If $n - r_i > n_0$, we repeat: There is a j so that $a(n - r_i) \leq a(n - r_i - r_j)\alpha^{r_j}$, hence $a(n) \leq a(n - r_i - r_j)\alpha^{r_i+r_j}$ and so on. By induction, $a(n) \leq a(n)\alpha^{n-n'}$, where $n \leq n_0$. Therefore $a(n) < c\alpha^n$ for all n . This contradicts exponential growth. \square

The sublemma allows us to choose r_k so that $a(r_k) > a(r_k - r_i)\alpha^{r_i}$ for $i = 1, \dots, k - 1$, and also so that $\alpha^{r_k} > 2^k$. Then $a(r_k - r_i) < a(r_k)\alpha^{-r_i} < a(r_k)2^{-i}$, and $\sum a(r_k - r_i) < a(r_k)(\sum 2^{-i}) < a(r_k)$. \square

VI.6. Projective covers

Let $A = k \oplus A_1 \oplus \dots$ be a noetherian, connected graded algebra. The term connected just means that $A_0 = k$. In the next two sections we work primarily with graded right A -modules. By *finite* module we mean a finitely generated module. A *map* $\phi : M \rightarrow N$ of graded modules is a homomorphism which sends $M_n \rightarrow N_n$ for every n . The modules we consider will all be *left bounded*, which means that $M_n = 0$ if $n \ll 0$.

The *shift* $M(r)$ of a module M is defined to be the graded module whose term of degree n is $M(r)_n = M_{n+r}$. In other words $M(r)$ it is equal to M except that the degrees have been shifted. The reason for introducing these shifts is to keep track of degrees in module homomorphisms. For example, if $x \in A_d$ is a homogeneous element of degree d , then right multiplication by x defines a map of graded modules $A(r) \xrightarrow{\rho_x} A(r+d)$. Since all linear maps $A_A \rightarrow A_A$ are given by left multiplication by A , this identifies the set of maps:

Corollary VI.6.1. $\text{Hom}_{gr}(A(r), A(s)) = A_{s-r}$. \square

If M is a graded right module and L is a graded left module, the tensor product $M \otimes_A N$ is a graded vector space, the degree d part of which is generated by the images of $\{M_n \otimes_k L_{d-n}\}$.

The symbol k will also denote the left or right A -module $A/A_{>0}$. It is a graded module, concentrated in degree zero, i.e., $k_0 = k$ and $k_n = 0$ for $n \neq 0$. For any module M , $MA_{>0}$ is a submodule, and

$$(VI.6.2) \quad M \otimes k = M \otimes (A/A_{>0}) \approx M/MA_{>0}.$$

This is a graded vector space, and it is finite dimensional if M is finitely generated.

Proposition VI.6.3. (*Nakayama Lemma*) (i) *Let M be a left bounded module. If $M \otimes k = 0$, then $M = 0$.*

(ii) *A map $\phi : M \rightarrow N$ of left bounded graded modules is surjective if and only if the map $M \otimes_A k \rightarrow N \otimes_A k$ is surjective.*

Proof. (i) Assume that M is not the zero module, and let d be the smallest degree such that $M_d \neq 0$. Then $(MA_{>0})_d = 0$, so

$$(M \otimes k)_d \approx M_d/(MA_{>0})_d = M_d.$$

The second assertion follows from the right exactness of tensor product. \square

Definition VI.6.4: A map $P \rightarrow M$ of finite graded modules is a *projective cover* of M if P is projective and if the induced map $P \otimes k \rightarrow M \otimes k$ is bijective.

Proposition VI.6.5. (i) Let $\phi : M \longrightarrow N$ be a surjective map of finite graded modules. If N is projective, then ϕ is bijective.

(ii) Every finite graded projective A -module is isomorphic to a finite direct sum of shifts of A_A : $P \approx \bigoplus A(r_i)$.

(iii) If $P \longrightarrow P \longrightarrow M \longrightarrow 0$ is an exact sequence of finite graded modules with P, P projective, then P is a projective cover of M if and only if the map $P \otimes k \longrightarrow P \otimes k$ is the zero map. \square

Proposition VI.6.6. Let

$$(VI.6.7) \quad \mathcal{P} \longrightarrow M := \{ \cdots \longrightarrow P_1 \longrightarrow P_0 \longrightarrow M \longrightarrow 0 \}$$

be a projective resolution of a finite module M , and define M_i by $M_0 = M$ and $M_i = \ker(P_{i-1} \longrightarrow M_{i-1})$ for $i > 0$. The following conditions are equivalent. If they hold, the resolution is said to be a minimal resolution.

(a) P_i is a projective cover of M_i for all i ,

(b) if $P_0 \longrightarrow M$ is a projective cover of M and for all $i > 0$, the induced maps $P_i \otimes k \longrightarrow P_{i-1} \otimes k$ are zero. \square

Corollary VI.6.8. Let $\mathcal{P} \longrightarrow M$ be a minimal projective resolution of a module M . Then $P_i \otimes k \approx \text{Tor}_i^A(M, k)$.

Proof. The Tor are computed as the homology of the complex $\mathcal{P} \otimes k$. Since the maps in this complex are zero, $H_i(\mathcal{P} \otimes k) = P_i \otimes k$. \square

Corollary VI.6.9. Let $\mathcal{P} \longrightarrow k \longrightarrow 0$ be a minimal projective resolution of the right module k , and say that

$$(VI.6.10) \quad P_i \approx \bigoplus_j A(-r_{ij}).$$

The minimal projective resolution of k as left module has the same shape, i.e., the number of summands and the shifts r_{ij} which appear are the same.

Proof. $P_i \otimes k \approx \text{Tor}_i^A(k, k)$, and $\text{Tor}_i^A(k, k)$ can be computed using either a projective resolution of the left module k or a projective resolution of the right module k . \square

Scholium VI.6.11. Let $P = \bigoplus A(p_i)$ and $P = \bigoplus A(q_j)$ be finite projective modules. Corollary VI.6.1 shows that

$$\text{Hom}_{gr}(P, P) = \bigoplus_{i,j} A(q_j - p_i).$$

The term $A(q_j - p_i)$ is zero unless $p_i \leq q_j$, because $A_n = 0$ if $n < 0$. If $\phi : P \longrightarrow P$ is a map, then $\phi \otimes k = 0$ is zero if and only if no entry ϕ_{ij} is a nonzero constant. This means $\phi_{ij} \in A_{q_j - p_i}$ is zero unless $p_i < q_j$.

Suppose that ϕ appears in a minimal projective resolution of some module. Then for every p_i , the summand $A(p_i)$ of P must have a nonzero image in P . Together with the condition that $\phi \otimes k = 0$, this implies that p_i must be strictly less than at least one index q_j . So with the notation (VI.6.10), the indices $-r_{ij}$ are decreasing with i . However, because various shifts can appear, the overlapping indices confuse the situation.

VI.7. Hilbert functions of graded algebras of finite global dimension

Let A be a noetherian connected graded algebra. The *Hilbert function* of A is the sequence $a_n = \dim_k A_n$. As we have seen (1.10), the Hilbert function is closely related to the growth of the algebra. We also consider the power series

$$(VI.7.1) \quad h(t) = \sum_{n=0}^{\infty} a_n t^n,$$

which is called the *Hilbert series* of A .

Lemma VI.7.2. *The radius of convergence r of the Hilbert series $h(t)$ is < 1 if and only if the Hilbert function has exponential growth.*

Proof. The root test tells us that $r = \overline{\lim}(a_n)^{1/n}$. \square

Suppose that A has *finite global dimension* d . This means that every finite graded module has a graded projective resolution of length $\leq d$. Then one can obtain a recursive formula for the Hilbert function in terms of a resolution of the A -module k . (It is a fact that if k has a finite projective resolution, then A has finite global dimension, i.e, every finite A -module has a finite projective resolution, but never mind.)

Say that the minimal projective resolution is

$$(VI.7.3) \quad 0 \longrightarrow P_d \xrightarrow{f_d} \dots \xrightarrow{f_2} P_1 \xrightarrow{f_1} P_0 \longrightarrow k \longrightarrow 0,$$

where each P_i is a finitely generated graded projective, hence is a sum of shifts of A . We note that $P_0 = A$ in this case, and we write $P_i = \bigoplus A(-r_{ij})$ as in (VI.6.10).

Lemma VI.7.4. *If $0 \longrightarrow V_d \longrightarrow V_{d-1} \longrightarrow \dots \longrightarrow V_0 \longrightarrow 0$ is an exact sequence of finite-dimensional vector spaces, then $\sum (-1)^i \dim V_i = 0$. \square*

Applying this lemma to the terms of degree n in the resolution (VI.7.3), we obtain the formula, valid for all $n > 0$,

$$(VI.7.5) \quad a_n - \sum_{i=1}^d (-1)^{i+1} (\sum_j a_{n-r_{ij}}) = 0,$$

in which all r_{ij} are positive (VI.6.11). This recursive formula, together with the initial conditions $a_n = 0$ for $n < 0$ and $a_0 = 1$, determines the Hilbert function.

Examples VI.7.6.

1. The q -polynomial ring $A = k_q[x, y]$, is defined by the relation $yx = qxy$. Writing operators on the right, the resolution of k is

$$(VI.7.7) \quad 0 \longrightarrow A(-2) \xrightarrow{(y, -qx)} A(-1)^2 \xrightarrow{\begin{pmatrix} x \\ y \end{pmatrix}} A \longrightarrow k \longrightarrow 0.$$

The recursive formula is $a_n = 2a_{n-1} - a_{n-2}$, and the Hilbert function is that of the commutative polynomial ring (as was clear from the start).

2. Let $A = k\langle x, y \rangle / I$, where I is the ideal generated by the two elements $[x, [x, y]] = x^2y - 2xyx + yx^2$ and $[[x, y], y] = xy^2 - 2yxy + y^2x$. The global dimension is three, and the resolution has the form

$$(VI.7.8) \quad 0 \longrightarrow A(-4) \xrightarrow{f^{(2)}} A(-3)^2 \xrightarrow{f^{(1)}} A(-1)^2 \xrightarrow{f^{(0)}} A \longrightarrow k \longrightarrow 0,$$

where ,

$$f^{(0)} = \begin{pmatrix} x \\ y \end{pmatrix}, \quad f^{(1)} = \begin{pmatrix} yx - 2xy & x^2 \\ y^2 & xy - 2yx \end{pmatrix}, \quad f^{(2)} = (y, x)$$

The recursive formula for the Hilbert function is $a_n - 2a_{n-1} - 2a_{n-3} + a_{n-4} = 0$.

3. Let $A = k\langle x, y \rangle / I$, where I is the ideal generated by the element $y^2x + x^2y - x^3$. The global dimension is 2, and the resolution has the form

$$(VI.7.9) \quad 0 \longrightarrow A(-3) \xrightarrow{(y^2-x^2, x^2)} A(-1)^2 \xrightarrow{\begin{pmatrix} x \\ y \end{pmatrix}} A \longrightarrow k \longrightarrow 0.$$

The recursive formula is $a_n - 2a_{n-1} + a_{n-3} = 0$.

Exercise: Using the Diamond lemma, prove that the resolutions VI.7.7-9 are exact.

We can also describe the Hilbert series $h(t)$ conveniently in terms of the recursive formula. Because signs alternate, we can gather the terms in (VI.7.5) together, to obtain a formula of the general shape

$$(VI.7.10) \quad a_n - \sum a_{n-r_i} + \sum a_{n-s_j} = 0,$$

Let

$$(VI.7.11) \quad q(t) = 1 - \sum t^{r_i} + \sum t^{s_j}.$$

The next proposition is proved by computing the product $q(t)h(t)$.

Proposition VI.7.12. $h(t) = 1/q(t)$. Hence the Hilbert function is a rational function. \square

Exercise: Prove Hilbert's theorem, that the Hilbert series of any finitely generated commutative graded ring is a rational function. Do it by writing A as a quotient of a polynomial ring P , and resolving A as a P -module.

Having expressed $h(t)$ as a rational function, we can determine the growth of the algebra. We write $q(t) = \prod (1 - \lambda_i t)$, where λ_i are the reciprocal roots of $q(t)$ – the reciprocals of the roots.

Theorem VI.7.13. Let A be a finitely generated, connected graded algebra of finite global dimension, and let $h(t) = 1/q(t)$ be its Hilbert series.

(i) a_n has exponential growth if and only if $q(t)$ has a reciprocal root λ with $|\lambda| > 1$.

(ii) If every reciprocal root of $q(t)$ has absolute value ≤ 1 , then the reciprocal roots are roots of unity, and q is a product of cyclotomic polynomials.

(iii) If the reciprocal roots of q are roots of unity, then A has polynomial growth, and its GK dimension is the multiplicity of the reciprocal root 1, the order of pole of $h(t)$ at $t = 1$. Moreover, the order of pole of h at $t = 1$ is its maximal order of pole.

Proof. (i) The radius of convergence r of the rational function $h(t)$ is the minimum absolute value of its poles. So $r < 1$ if and only if $q(t)$ has a root λ of absolute value < 1 .

(ii) The reciprocal roots are the nonzero roots of the polynomial $t^n q(t^{-1}) = t^n - \sum t^{n-r_i} + \sum t^{n-s_j}$. This is a monic polynomial with integer coefficients. So first of all, the product of the reciprocal roots is an integer with absolute value ≥ 1 . If $|\lambda_i| \leq 1$ for all i , then $|\prod \lambda_i| \leq 1$ and so $|\lambda_i| = 1$ for all i . Lemma (VI.7.23) below completes the proof.

(iii) Let k denote an integer such that $\lambda_i^k = 1$ for all i , and let ζ be a primitive k th root of 1. Also, let p denote the largest multiplicity among the roots of $q(t)$. We write $h(t)$ in terms of partial fractions, say

$$(VI.7.14) \quad h(t) = \frac{1}{q(t)} = \sum_{i,j} \frac{c_{ij}}{(1 - \zeta^i t)^j}$$

with $i = 0, \dots, k-1$ and $j = 1, \dots, p$, where c_{ij} are complex numbers. The binomial expansion for a negative power is

$$(VI.7.15) \quad \frac{1}{(1-t)^j} = \sum \binom{n+j-1}{j-1} t^n.$$

This yields the formula

$$(VI.7.16) \quad a_n = \sum c_{ij} \binom{n+j-1}{j-1} \zeta^{in},$$

where $j = 1, \dots, p$. Thus the value of a_n cycles through k polynomial functions. For $\nu = 0, \dots, k-1$,

$$(VI.7.17) \quad a_n = \gamma_\nu(n) := \sum_{i,j} c_{ij} \zeta^{i\nu} \binom{n+j-1}{j-1}, \quad \text{if } n \equiv \nu \pmod{k}.$$

Because a_n takes real values at the integers $n \equiv \nu$, $\gamma_\nu(n)$ is a real polynomial. Its degree is at most $p-1$, so $gk(A) \leq p$.

The coefficient of n^{p-1} in $\gamma_\nu(n)$ is

$$(VI.7.18) \quad \frac{\gamma_{\nu p}}{(p-1)!} = \sum_i \frac{c_{ip} \zeta^{i\nu}}{(p-1)!}.$$

It is non-negative because a_n takes non-negative values.

Since $h(t)$ has a pole of order p , at least one of the coefficients c_{ip} is nonzero. The coefficient vector $(\gamma_{0p}, \gamma_{1p}, \dots, \gamma_{k-1p})$ is obtained from the vector $(c_{0p}, \dots, c_{k-1p})$ by multiplying by the nonsingular matrix $[\zeta^{i\nu}]$. Therefore at least one coefficient γ_{ip} is positive, and the sum $\gamma = \gamma_{0p} + \dots + \gamma_{k-1p}$ is positive too. Since

$$(VI.7.19) \quad \gamma = \sum_{i,\nu} c_{ip} \zeta^{i\nu} = k c_{i0},$$

It follows that $c_{i0} > 0$, which implies that h has a pole of order p at $t = 1$. Then (1.10) shows that $gk(A) = p$. \square

Examples VI.7.20. For the algebra with resolution (VI.7.8), $q(t) = 1 - 2t + 2t^3 - t^4 = (1-t)^3(1+t)$. The algebra has GK dimension 3. For the algebra with resolution (VI.7.9), $q(t) = 1 - 2t + t^3 = (1-t)(t^2 - t - 1)$, which has the root $\frac{1}{2}(1 + \sqrt{5})$: exponential growth, hence by the theorem of Stephenson and Zhang, not noetherian.

Corollary VI.7.21. (Stephenson – Zhang) *A connected graded, right noetherian algebra A of finite global dimension has polynomial growth.*

Conjecture VI.7.22. (Anick) *If A is connected graded and right noetherian, then*

(i) *$gk(A)$ is equal to the global dimension of A , the length of the minimal projective resolution VI.7.3.*

(ii) *The Hilbert series has the form $h(t) = 1/q(t)$, where $q(t)$ is a product of polynomials of the form $1 - t^k$.*

The series $h = 1/q$, where $q(t) = (1 - t^{k_1}) \cdots (1 - t^{k_a})$ is the Hilbert series of the commutative polynomial ring on d variables y_i , where the degree of y_i is k_i . So these series do arise.

Lemma VI.7.23. *Let $f(x)$ be a monic polynomial with integer coefficients. If $|\alpha| = 1$ for every complex root α of f , then f is a product of cyclotomic polynomials, and so its roots are roots of unity.*

Proof. We may assume that f is irreducible, of degree n . Let its complex roots be $\alpha_1, \dots, \alpha_n$, and let K be the field obtained from \mathbb{Q} by adjoining an abstract root \bar{x} of f . Then K has n embeddings ϕ_i into \mathbb{C} , defined by $\phi_i(\bar{x}) = \alpha_i$. The roots of f are algebraic integers in K . What we need to know from algebraic number theory is that the algebraic integers in K form a ring \mathcal{O} , and that the map (ϕ_1, \dots, ϕ_n) embeds this ring as a lattice (not a full lattice) into \mathbb{C}^n . So \mathcal{O} is a discrete subset of \mathbb{C}^n . The set of points $(z_1, \dots, z_n) \in \mathbb{C}^n$ such that $|z_i| = 1$ for all i is compact and closed under multiplication. Therefore it contains only finitely many algebraic integers, and they form a finite group. \square

VI.8. Modules with linear growth

Let A be a finitely generated algebra. We fix a finite dimensional subspace V which generates A and which contains 1. Let M be a finite A -module, and let U be a finite dimensional subspace which generates M as A -module. Then $M = \bigcup UV^n$. The growth function of M associated to the two subspaces U and V is

$$(VI.8.1) \quad g(n) = \dim UV^n.$$

Lemma VI.8.2. *Let U_1, U_2 be two generating subspaces of a finite module M , and let $g_i(n) = \dim_k U_i V^n$. There is an integer r such that*

$$g_1(n - r) \leq g_2(n) \leq g_1(n + r).$$

Proof. Since $U_1 V^n$ exhaust M , $U_2 \subset U_1 V^r$ for some integer r . Then for all n , $V^n U_2 \subset V^{n+r} U_1$, which implies that $g_2(n) \leq g_1(n + r)$. \square

Lemma VI.8.3. *Let M be a finite A -module which has infinite dimension over k , and let $g(n) = \dim_k UV^n$ as above. Then $g(n)$ is a strictly increasing function, and in particular, $g(n) \geq n$. \square*

Lemma VI.8.4. *Let $N \subset M$ be finitely generated modules, with N infinite dimensional. Let V be a generating subspace for A , and U a generating subspace of M . Suppose that the corresponding growth function for M has a linear bound, say $g(n) = \dim UV^n \leq rn + s$. Let \bar{U} be the image of U in $\bar{M} = M/N$. Then $\dim \bar{U}V^n \leq (r-1)n + s$.*

Proof. Since N is finitely generated, $W = UV^k \cap N$ is a generating subspace for suitable k . Then for $n > k$, there is a surjective map $(UV^n/WV^{n-k}) \rightarrow \bar{U}V^n$. By (VI.8.3), $\dim WV^m \geq m$. So $\dim \bar{U}V^n \leq \dim UV^n - \dim WV^{n-k} \leq rn + s - (n-k) = (r-1)n + (s+k)$. \square

Corollary VI.8.5. *Let $0 \subset M_0 \subset \dots \subset M_n = M$ be a filtration of submodules. Suppose that M is finitely generated, and has a growth function $g(n)$ bounded by a linear function $rn + s$. There are at most r indices i such that M_i/M_{i-1} is finitely generated and infinite dimensional.*

Proof. Let i_0 be the first index such that M_{i_0}/M_{i_0-1} is finitely generated and infinite dimensional. We may replace M by M/M_{i_0-1} and reindex, to reduce to the case that $i_0 = 1$. By the previous lemma, M/M_1 has a growth function bounded by $(r-1)n + s$, so induction on r completes the proof. \square

VI.9. Theorem of Small and Warfield

Theorem VI.9.1. *(Small – Warfield) Let A be a finitely generated prime k -algebra of GK dimension 1. Then A is a finite module over its center Z , and Z is a finitely generated commutative domain of dimension 1.*

Corollary VI.9.2. *Suppose that k is an algebraically closed field, and let A be a finitely generated k -algebra which is a domain, and which has GK dimension 1. Then A is commutative.*

Proof. The corollary follows from the theorem because the field of fractions K of the center Z is a function field in one variable, and $D = A \otimes_Z K$ is a division ring finite over K . Tsen's theorem implies that D , and hence A , is commutative. \square

We have three preliminary propositions before beginning the proof.

Proposition VI.9.3. *Let Q be a k -algebra with center k . If $Q \otimes K$ is semisimple for all field extensions $K \supset k$, then $[Q : k] < \infty$, and hence Q is a central simple algebra over k .*

Proof. We may assume k algebraically closed. Since Q is semisimple with center k , it is a matrix algebra over a division ring D , and we must show that $D = k$.

Let K be the rational function field $k(t)$. Together with the monomials t^ν , the elements $1/(t-a)^j$ with $a \in k$ and $j = 1, 2, \dots$ form a k -basis of K . Let x be a nonzero element of D . It is easily checked that $x+t$ is a regular element of $Q \otimes K$. So because $Q \otimes K$ is semisimple, $x+t$ is invertible. We write its inverse w explicitly using the basis for K , say

$$(VI.9.4) \quad w = \sum \frac{c_{ij}}{(t-a_i)^j} + \sum b_\nu t^\nu,$$

with $a_i \in k$ and $c_{ij}, b_\nu \in Q$. It is clear that w is not a polynomial in t , so some c_{ij} is nonzero. Computing $(x+t)w$ gives us the relations

$$(x+a_i)c_{ij} + c_{i,j+1} = 0,$$

for $j \geq 1$. Since $c_{i,j+1} = 0$ for sufficiently large j , we obtain some relation of the form $(x+a)c = 0$ with $a \in k$ and $c \in Q$ nonzero. Since $x \in D$, this implies that $x+a = 0$. \square

Proposition VI.9.5. *Let S be a right Ore set of regular elements in a ring A , and let $Q = AS^{-1}$.*

(i) *If A is a prime ring, so is Q .*

(ii) *If Q is a simple ring, then A is a prime ring.*

Proof. (i) If I is an ideal of Q , then $(I \cap A)Q = I$, so $I \cap A \neq 0$, and $I \cap A$ is an ideal of A . Then if I, J are nonzero ideals of Q , $(I \cap A)(J \cap A) \neq 0$ and $IJ \neq 0$.

(ii) Suppose that Q is simple, and let I be a nonzero ideal of A . Then $QIQ = Q$, so $1 = \sum u_i s_i^{-1} v_i t_i^{-1}$ with $s_i, t_i \in S$ and $v_i \in I$. We may replace the t_i by a common multiple t . (The elements $v_i \in I$ will change.) Then $t = \sum u_i s_i^{-1} v_i \in QI$, and since t is invertible, $QI = Q$. If J is another nonzero ideal, $QIJ = QJ \neq 0$, so $IJ \neq 0$. \square

Proposition VI.9.6. *Let A be a prime, infinite dimensional k -algebra.*

(i) *Every nonzero right ideal M of A is infinite dimensional.*

(ii) *Let $M < M'$ be right ideals such that M is a right annihilator. There is an ideal N with $M \subset N \subset M'$ such that N/M is generated by one element and is infinite dimensional.*

Proof. (i) The right annihilator J of a finite dimensional right ideal M has finite codimension in A . If M is not zero, the left annihilator I of J is not zero because it contains M . From $IJ = 0$, and $I \neq 0$ we conclude $J = 0$, which shows that A is finite dimensional.

(ii) Say that M is the right annihilator of the set $X \subset A$. Then there is an element $x \in X$ such that $xM \neq 0$. We set $N = xM + M$, so that N/M is generated by the residue of x , and we note that N/M is isomorphic to the right ideal xM . So (i) applies. \square

Lemma VI.9.7. *A finitely generated prime algebra A of GK-dimension ≤ 1 is a Goldie ring.*

Proof. The fact that A satisfies acc on right annihilators follows from Lemmas VI.8.5 and VI.9.6. The fact that A has finite Goldie rank follows similarly: Every uniform right ideal U of A contains a uniform right ideal which is finitely generated, and then Corollary VI.8.6 applies. \square

Lemma VI.9.8. *Let A be a finitely generated prime k -algebra of GK dimension 1, and with ring of fractions $AS^{-1} = Q$. A right ideal M of A is essential in A if and only if it has finite codimension.*

Proof. Since A is Goldie, a right ideal M is essential if and only if it contains a regular element s . Then $sA \subset M \subset A$, and because sA has the same growth as A , it has finite codimension. \square

Lemma VI.9.9. *Let A be a finitely generated prime k -algebra of GK dimension ≤ 1 , let K be a field extension of k , and let $A' = A \otimes K$.*

(i) *A' is a finitely generated K -algebra, and its GK dimension as K -algebra is ≤ 1 .*

(ii) *If k is the center of A , then A' is a prime ring.*

Proof. The first assertion is clear. For (ii), let Q be the ring of fractions of A . Then $Q \otimes K$ is a simple ring with center K (CSA, 1.5), so (VI.9.5ii) applies. \square

Lemma VI.9.10. *Let A be a finitely generated prime k -algebra of GK dimension ≤ 1 , with ring of fractions Q . Let K be the center of Q , and let $A' = AK$ denote the subring of Q generated by A and K . Then A' is finitely generated prime K -algebra of GK dimension ≤ 1 , its center is K , and its ring of fractions is Q .*

Proof. Since A' is a quotient of $A \otimes K$, it is a finitely generated K -algebra of GK dimension ≤ 1 . Its center is K because $K \subset A' \subset Q$ and K is the center of Q . Finally, the set of regular elements of A' is also an Ore set in A' , with ring of fractions Q , so A' is a prime ring by (VI.9.5ii). \square

Lemma VI.9.11. *Let A be a finitely generated prime k -algebra of GK dimension ≤ 1 , and with ring of fractions Q . Then Q is a finite module over its center.*

Proof. Lemma VI.9.10 allows us to assume that k is the center of A and of Q . We assume that $[A : k] = \infty$, and we derive a contradiction. By Proposition VI.9.3, it suffices to show that $Q = Q \otimes K$ is semisimple for every field extension K of k , and $A = A \otimes K$ is a finitely generated prime algebra of GK dimension ≤ 1 (VI.9.9). Also, Q is a simple ring (CSA, 1.5).

The set S of regular elements of A is an Ore set in A , and $Q = A S^{-1}$. Since A is a Goldie ring (VI.9.7), so is Q . To show Q semisimple, it suffices to show that every regular element of Q is invertible.

Let $t \in Q$ be a regular element, so that $tQ = N$ is an essential right ideal of Q , and let $M = N \cap A$. Then $MS^{-1} = N$, and M is an essential right ideal of A . Indeed, if P is a nonzero right ideal of A , then $(P \cap M)S^{-1} = PS^{-1} \cap MS^{-1}$, which is not zero because $MS^{-1} = N$ is essential.

By VI.9.8, M has finite codimension in A . Since A is infinite dimensional, the right annihilator of $L = A/M$ is not zero. Then since Q is a simple ring, $L \otimes_{A'} Q = 0$. This shows that $M \otimes_{A'} Q = MS^{-1} = Q$, and that t is invertible. \square

Lemma VI.9.12. *A finitely generated prime algebra A of GK dimension 1 satisfies a polynomial identity. \square*

Lemma VI.9.13. *Let A be a k -algebra, and let $\gamma_1, \dots, \gamma_r \in A$ be regular central elements which generate the unit ideal in A . If $A[\gamma_i^{-1}]$ is finitely generated for each $i = 1, \dots, r$, then A is finitely generated.*

Proof. If γ_i generate the unit ideal, then for each $n \geq 1$, $\gamma_1^n, \dots, \gamma_r^n$ also generate A . This is seen by raising the equation $a_1\gamma_1 + \dots + a_r\gamma_r = 1$ to a sufficiently large power. Thus for every n , there is an equation $a_1\gamma_1^n + \dots + a_r\gamma_r^n = 1$ with $a_i \in A$.

We choose elements $x_1, \dots, x_m \in A$ such that $A[\gamma_i^{-1}]$ is generated by $\{x_1, \dots, x_m; \gamma_i^{-1}\}$ and such that the subalgebra B of A generated by x_1, \dots, x_m contains γ_i , for each i . If $\alpha \in A$, then $\alpha \in A[\gamma_i^{-1}]$, hence $\alpha\gamma_i^n \in B$ for some n , which we may choose independently of i . Then the equation $a_1\gamma_1^n + \dots + a_r\gamma_r^n = 1$ shows that $\alpha \in B$. \square

Lemma VI.9.14. *Let A be an Azumaya algebra over a commutative ring Z . If A is finitely generated, so is Z .*

Proof. If A is a finitely generated k -algebra, then the ring $A^o \otimes_Z A = \text{End}_Z A$ is also a finitely generated k -algebra, and its center is Z . This reduces us to the case that $A = \text{End}_Z V$ for some locally free Z -module V . The previous lemma reduces us to the case that V is a free module, i.e., that A is a matrix algebra over Z . When A is a matrix algebra, $Z = e_{11}Ae_{11}$, and if $\{x_i\}$ generate A , $\{e_{11}x_i e_{11}\}$ generate Z . \square

Lemma VI.9.15. *Let K be a function field in one variable over k , let $R \subset R'$ be subrings of K with field of fractions K . If R' is a finitely generated k -algebra, so is R .*

Proof. Let R be a Dedekind domain with field of fractions K . An element $x \in K$ lies in R if and only if it has no pole at any prime ideal $p \in S = \text{Spec } R$. Thus R is determined by its spectrum. If $R \subset R'$ are two finitely generated Dedekind domains, the map $\text{Spec } R' = S' \rightarrow S = \text{Spec } R$ identifies S' as

the complement of a finite set of primes in S . So there are finitely many possibilities for a spectrum in between, hence finitely many intermediate Dedekind domains. (Actually every intermediate ring is a Dedekind domain, but never mind.)

The assertion to be proved is equivalent with the following one: Given a finitely generated algebra R with field of fractions K , the set of finitely generated subrings R of R with field of fractions K has the ascending chain condition. We may assume that R is a Dedekind domain. Given a chain of subrings $B_1 \subset B_2 \subset \cdots \subset R$, the normalizations R_i of B_i form a chain of subrings which are Dedekind domains. So by what has been shown, this chain stabilizes, and we may assume that the normalizations R_i are all equal, say to R . Then because R is a finite module over B_1 and $B_i \subset R$ for all i , the chain stabilizes again. \square

Lemma VI.9.16. *Let A be a finitely generated prime PI algebra of GK dimension 1. The center Z of A is a finitely generated k -algebra of Krull dimension 1.*

Proof. The fact that A is prime implies that its center is a domain. Let γ be a nonzero evaluation of a central polynomial in A . Then $A = A[\gamma^{-1}]$ is an Azumaya algebra over its center Z (REP, 9.3), and $Z = Z[\gamma^{-1}]$. By Lemma VI.9.15, Z is finitely generated. By Lemma VI.1.9, $gk(A) = gk(A) = 1$. Therefore $gk(Z) \leq 1$, and since A is not a finite k -module, neither is Z , so $gk(Z) = 1$. This implies that Z is a finitely generated domain of Krull dimension 1. Lemma VI.9.15 shows that Z is finitely generated. \square

Lemma VI.9.17. *Let Z be a noetherian domain whose integral closure R is a Dedekind domain and a finite Z -module, and let K be the field of fractions of Z . Let A_K be a central simple K -algebra and let A be a subring which generates A_K over K and which contains Z . Then A is a finite Z -module if and only if $\text{trace}(a) \in R$ for every $a \in A$.*

Proof. Let $A = AR$ be the ring generated by A and R . If A is a finite Z -module, then A is an R -order, and $\text{trace}(a) \in R$ for every $a \in A$ (MO, 2.2). We must show the converse, so we suppose that $\text{trace}(a) \in R$ for every $a \in A$, and we show that A is a finite Z -module. The hypothesis tells us that the trace pairing $a, b \mapsto \text{trace}(ab)$ takes its values in R . We restrict this pairing to an arbitrary Z -lattice $V \subset A$, obtaining a map $V \otimes A \rightarrow R$, hence a map $A \rightarrow \text{Hom}_Z(V, R)$, which is injective because A_K is a simple ring. Since $\text{Hom}_R(V, R)$ is a finite module, so is A . \square

Lemma VI.9.18. *With the notation of (VI.9.17), suppose that Z is the center of A . Then A is a finite Z -module.*

Proof. This is hack work. We first reduce to the case that $Z = R$, i.e., that Z is a Dedekind domain. Let $A = AR$ as above. For this reduction, it suffices to show that the center R of A is equal to R . Then if the lemma is proved when the center is R , A will be a finite R -module. Hence A and A will be finite Z -modules.

Suppose that $R \neq R$. Then there is a prime $p \in S = \text{Spec } R$ which is not in $S = \text{Spec } R$, and there are elements $x \in R$ with arbitrarily large pole at p . We may write $x = \sum a_i r_i$, with $a_i \in A$ and $r_i \in R$. Let c be a nonzero element of the conductor of R in Z . Then $xc \in A$, hence $xc \in Z$. But if the pole of x at p is large enough, then $xc \notin R$, which is a contradiction.

It remains to treat the case that the center of A is a Dedekind domain R . We choose a finite field extension K of K which splits the central simple algebra A_K , and we let R be the integral closure of R in K . The extension $R \rightarrow R$ is faithfully flat because R is a Dedekind domain. So R is

the center of $A = A \otimes_R R$, A is a subring of $A_{K'}$, and A is a finite R -module if and only if A is a finite R -module. This reduces us to the case that A_K is a matrix algebra. Moreover, a further finite field extension is permissible.

It suffices to show that $\text{trace}(a) \in R$ for all $a \in A$. Suppose not. We choose a with $\text{trace}(a) \notin R$. So $\text{trace}(a)$ has a pole at some point $p \in \text{Spec } R$. We replace R by the local ring at p , a discrete valuation ring. If the ground field k is infinite, then inspecting the characteristic polynomial shows that there is a $c \in k$ such that $\det(c + a) \notin R$. In any case, we can achieve this at the cost of making a finite field extension of k and K . So we obtain an element a such that $\det(a) \notin R$.

Let t be a generator for the maximal ideal of R . Since A generates A_K , there is an integer r such that $t^r e_{ij} \in A$ for all i, j . Since $\det(a) \notin R$, at least one of its matrix entries is not in R . Replacing a by a power, we may assume that some matrix entry a_{ij} has a pole of order $> 2r$. Then $t^{-1}e_{ii} \in A$ for every i , which shows that $t^{-1} \in R$, contrary to hypothesis. \square

Theorem VI.9.1 follows from Lemmas VI.9.12, VI.9.16 and VI.9.18.

MIT OpenCourseWare
<https://ocw.mit.edu>

18.706 Noncommutative Algebra
Spring 2023

For information about citing these materials or our Terms of Use, visit: <https://ocw.mit.edu/terms>.