Associative rings and their adjoint groups.

Ya. Sysak

One of purposes of these lectures is to state some results and methods (included some recent of them) concerning investigations of relationships between the structure of an associative ring, in particular its Lie-structure, and the structure of the adjoint group and semigroup of this ring.

1 Preliminaries. Irreducible modules and the density theorem of Jacobson.

Let R be a ring, not necessarily with a unity, and M an abelian group. Recall that M is a *right R-module* if a mapping $M \times R \to M$ whose image is usually denoted by $\{mr \mid m \in M, r \in R\}$ is determined and for all $m, n \in M$ and all $r, s \in R$ the following conditions hold:

1) $(m+n)r = mr + nr$,

2) $m(r+s) = mr + ms$, and

3) $m(rs) = (mr)s$.

If R has a unity 1 and $m \cdot 1 = m$ for all m, then M is called *unital*. Obviously every abelian group can be viewed as a unital $\mathbb{Z}$-module.

As usually, if V is a non-empty subset of M, then VR denotes the subgroup of M generated by the set $\{vr \mid v \in V, r \in R\}$. A subgroup N of M is a *submodule*, if $NR \subseteq N$. The factor group M/N with the mapping $(m+N)r = mr + N$ for all $m \in M$ and $r \in R$ is a right R-module called the *factor module* of M modulo N. A module M is a *direct sum* of modules M_i with i from a set $\mathcal{I}$ if M, regarded as a group, is the direct sum of the subgroups M_i.

Example. 1. If $mr = 0$ for all $m \in M$ and $r \in R$, then M is an R-module called *trivial*. Every subgroup of M is a submodule in M.

2. If $M = R^+$ and and the mapping $M \times R \to M$ is the multiplication in R, then R can be considered as a right R-module whose submodules coincide with the right ideals of R.

A module M is *simple* if 0 and M are the only submodules of M. A simple module M is *irreducible* if $mR = M$ for some (and so for any) non-zero $m \in M$. In particular, every simple module is either irreducible or trivial. A module M is *semisimple* if it is a direct sum of irreducible submodules.

Lemma 1.1. *Let M be a semisimple R-module. Then every submodule N of M is complemented in M, i.e. there exists a submodule V of M such that $M = N \oplus V$ (which means $M = N + V$ and $N \cap V = 0$). Moreover, if $m \in M$, then $m \in mR$.*

Proof. By Zorn's lemma, there exists a submodule V which is a direct sum of irreducible submodules of M and is maximal with respect to the condition $N \cap V = 0$. If $N + V \neq M$, then there exists an irreducible submodule W of V such that $(N + V) \cap W = 0$ and so $N \cap (V + W) = 0$ which contradicts to the choice of V. Thus $M = N \oplus V$.

Let $0 \neq m \in M$ and $N = mR + m\mathbb{Z}$. Then N is a submodule of M such that $NR \subseteq mR$. By proved above, $N = mR \oplus W$ for a submodule W of N. Since $WR \subseteq mR \cap W = 0$, this implies $W = 0$ and so $m \in mR$. □

For two right R-modules M_1 and M_2, a group homomorphism $\alpha : M_1 \to M_2$ is called an *R-homomorphism* or a *module homomorphism* from M_1 into M_2 if $(mr)^\alpha = m^\alpha r$ for all $m \in M_1$ and $r \in R$. Clearly the kernel $\operatorname{Ker}\alpha$ and the image $\operatorname{Im}\alpha$ of the R-homomorphism α are submodules of M_1 and M_2, respectively, and the factor module $M_1/\operatorname{Ker}\alpha$ is isomorphic to $\operatorname{Im}\alpha$.

If M is a right R-module, then the set $\operatorname{End}_R M$ of all R-homomorphisms from M into M (called R-endomorpisms of M) forms a ring if the addition and the multiplication of two elements α and β of $\operatorname{End}_R M$ are defined by the rules

$$m^{\alpha+\beta} = m^\alpha + m^\beta \quad \text{and} \quad m^{\alpha\beta} = (m^\alpha)^\beta$$

for every $m \in M$. The identity mapping on M is a unity of this ring and M can be viewed as a unital right $\operatorname{End}_R M$-module if we put $m\alpha = m^\alpha$ for all $m \in M$ and $\alpha \in \operatorname{End}_R M$.

For a non-empty subset V of M, the set $\operatorname{Ann}_R(V) = \{r \in R \mid vr = 0 \text{ for all } v \in V\}$ is a right ideal of R which is called the *annihilator* of V in R. Moreover, the annihilator $\operatorname{Ann}_R(M)$ of M in R is an ideal of R. The module M is *faithful* if $\operatorname{Ann}_R(M) = 0$.

Lemma 1.2. *Let M be an irreducible right R-module and $P = \operatorname{Ann}_R(m)$ for a non-zero element $m \in M$. Then the following statements hold:*

1) *P is a maximal right ideal of R such that the modules M and R/P are isomorphic,*

2) *there exists an element $e \in R \setminus P$ such that $r - er \in P$ for every $r \in R$, and*

3) *the ideal I of R which is maximal with respect to the condition $I \subseteq P$ coincides with the annihilator $\operatorname{Ann}_R M$.*

Conversely, if for a maximal right ideal P of R there exists an element $e \in R \setminus P$ such that $r - er \in P$ for every $r \in R$, then the factor module R/P is irreducible.

Proof. Since the mapping $r \mapsto mr$ with $r \in R$ is a module homomorphism from R onto M, its kernel P is a right ideal of R and the image mR is a submodule of M. As M is irreducible and $mR \neq 0$, we have $mR = M$. Therefore M is isomorphic to the factor module R/P and so P is a maximal right ideal of R.

As $m = me$ for some $e \in R$, we have $mr = mer$ for every $r \in R$ which implies $m(r - er) = 0$ and so $r - er \in P$. Clearly $e \neq P$ because $m \neq 0$.

Finally, from $\mathrm{Ann}_R(M) \subseteq P$ it follows $\mathrm{Ann}_R(M) \subseteq I$. On the other hand, $MI = (mR)I \subseteq mI = 0$, so that $\mathrm{Ann}_R M = I$

Conversely, if P is a maximal right ideal of R and $r - er \in P$ for all $r \in R$ and some $e \in R \setminus P$, then $(e + P)R = R$ and so the R-module R/P is irreducible. □

Corollary 1.3. *If a ring R, regarded as a right R-module, is irreducible, then R is a division ring.*

Proof. By Lemma 1.2, R has a left unity e. Therefore $re = re^2$ and hence $(r - re)e = 0$ for every $r \in R$. Since $\mathrm{Ann}_R(e) = 0$, it follows that $r = re$, so that e is a unity of R. Finally, every non-zero element $r \in R$ is left invertible in R because $rR = R$. Thus, r must be invertible in R and so R is a division ring. □

Let R be a ring, M a right R-module and $K = \mathrm{End}_R M$. Then M is a unital right K-module, so that we can consider the ring $L = \mathrm{End}_K M$. For each $r \in R$, we define the mapping $\hat{r} : M \to M$ by the rule $m^{\hat{r}} = mr$ for every $m \in M$. Then $\hat{r}$ is a K-endomorphism of M because

$$(m\alpha)^{\hat{r}} = (m^{\alpha})r = (mr)^{\alpha} = m^{\hat{r}}\alpha$$

for all $m \in M$ and $r \in R$. The mapping $r \mapsto \hat{r}$ determines a ring homomorphism $\hat{}\colon R \to L$ whose kernel coincides with $\mathrm{Ann}_R M$ and whose image $\hat{R}$ is a subring of L. In particular, $\hat{R}$ is isomorphic to R provided that M is a faithful R-module.

Let $E = \mathrm{End}\, M$ be the ring of all group endomorphisms of M. Clearly all rings K, L and $\hat{R}$ are subrings of E. Moreover, $K = C_E(\hat{R})$ and $L = C_E(K)$, so that $\hat{R} \subseteq L = C_E(C_E(\hat{R}))$. It turns out that in many important cases $\hat{R}$ coincides with L. A precise description of this situation is given by the *density theorem of Jacobson.*

Theorem 1.4. *Let M be a semisimple R-module and $K = \mathrm{End}_R M$. Then for each $f \in \mathrm{End}_K M$ and any elements $m_1, \ldots, m_k$ of M there exists $r \in R$ such that $m_i^f = m_i r$ for all $1 \leq i \leq k$.*

Proof. Let $k = 1$. We want to find an element $r \in R$ such that $m_1^f = m_1 r$. It follows from Lemma 1.1 that $M = m_1 R \oplus N$ for some submodule N of M and $m \in mR$ for every $m \in M$. Let $\pi : M \to m_1 R$ be the projection from M onto $m_1 R$, i.e. $m - m^{\pi}$ belongs to N for every $m \in M$. Obviously π is an R-endomorphism of M, so that $\pi \in K$. Therefore $m_1^f = (m_1^{\pi})^f = (m_1^f)^{\pi} = m_1 r$ for some $r \in R$.

Consider now the general case. Let $M^k = M \oplus \ldots \oplus M$ with k summands. Then M^k is a semisimple R-module. The mapping $f^k : M^k \to M^k$ with $(x_1, \ldots, x_k)^{f^k} = (x_1^f, \ldots, x_k^f)$ for any $x_1, \ldots, x_k \in M$ is a K-endomorphism of M^k. Hence, by proved above, there exists $r \in R$ such that

$$(m_1, \ldots, m_k)^{f^k} = (m_1, \ldots, m_k)r.$$

Thus $m_i^f = m_i r$ for every $1 \leq i \leq k$, as required. □

Corollary 1.5. *Let M be a faithful irreducible R-module and $K = \mathrm{End}_R M$. Then K is a division ring, so that M is a vector space over K. If $\dim_K M < \infty$, then R is isomorphic to $\mathrm{End}_K M$.*

Proof. Let $m_1, \ldots, m_k$ be a basis of M over K and $f \in \mathrm{End}_K M$. Then there exists $r \in R$ such that $m_i^f = m_i r$ for every $1 \leq i \leq k$ by the density theorem 1.4. This implies $m^f = mr$ for every $m \in M$ and so $f = \hat{r}$. Thus $\hat{R} = \mathrm{End}_K M$. □

Corollary 1.6. *Let M be a finite-dimensional vector space over an algebraic closed field F and let R be a subring of $\mathrm{End}_F M (\cong \mathrm{Mat}_k(F))$. If M is an irreducible R-module, then R is isomorphic to $\mathrm{End}_F M$.*

Proof. Since M is irreducible, then the ring $K = \mathrm{End}_R M$ is a division ring by Schur's lemma. Clearly every element of K is algebraic over F because $K = C_{\mathrm{End}_F M}(R)$ and every matrix over F is a root of its characteristic polynomial. Thus, for each $\alpha \in K$, there exists a minimal polynomial $p(x)$ over F such that $p(\alpha) = 0$. Since F is algebraic closed and the polynomial $p(x)$ is irreducible over F, we have $\deg p(x) = 1$ and so $\alpha \in F$. Therefore $K = F$. Let $m_1, \ldots, m_k$ be a basis of M over F and $f \in \mathrm{End}_K M = \mathrm{End}_F M$. Then there exists $r \in R$ such that $m_i^f = m_i r$ for all $1 \leq i \leq k$ and hence $m^f = mr$ for every $m \in M$. Thus $f = r$ and so $R = \mathrm{End}_F M$. □

A ring R is called (right)*primitive* if there exists a faithful irreducible (right) R-module. It is known that right primitivity does not imply left primitivity. An example of such a ring was given by G. Bergman (1964).

Example. 1. If R is a ring and M is an irreducible right R-module, then the factor ring $R/\mathrm{Ann}_R M$ is primitive because M can be viewed as a faithful irreducible $R/\mathrm{Ann}_R M$-module.

2. Every division ring is primitive.

3. The ring $\mathbb{Z}$ and any ring with zero multiplication are not primitive.

Proposition 1.7. *If a commutative ring R is primitive, then R is a field.*

Proof. If M is a faithful irreducible R-module, then $\mathrm{Ann}_R(m) = 0$ for each non-zero $m \in M$ by Lemma 1.2.3) and so R is a field by Corollary 1.3. □

Proposition 1.8. *Let R be a primitive ring. Then either R is isomorphic to the ring $\mathrm{Mat}_n(D)$ of all $(n \times n)$-matrices over a division ring D for some integer $n \geq 1$ or for each $n \geq 1$ there exists a subring S_n of R such that $\mathrm{Mat}_n(D)$ is a homomorphic image of S_n.*

Proof. Let M be a faithful irreducible R-module and $D = \mathrm{End}_R M$. Then M is a vector space over D. If M is finite-dimensional over D and $\dim_D M = n$ for some $n \geq 1$, then R is isomorphic to $\mathrm{End}_D M$ by Corollary 1.5 and so to $\mathrm{Mat}_n(D)$.

Let M be infinite-dimensional and V its subspace of dimension $n \geq 1$. Take a basis $v_1, \ldots, v_n$ of V and denote by W a vector subspace of M such that $M = V \oplus W$. Every non-zero element $f \in \mathrm{End}_D V$ can be considered as an element of $\mathrm{End}_D M$ if we put $W^f = 0$. Thus, by the density theorem 1.4, there exists $r \in R$ such that $v_i^f = v_i r$ for all $1 \leq i \leq n$. Clearly $Vr \subseteq V$. Put $S_n = \{s \in R \mid Vs \subseteq V\}$. Then S is a subring of R and the factor ring $S_n / \mathrm{Ann}_{S_n} V$ is isomorphic to $\mathrm{End}_D V$ and so to $\mathrm{Mat}_n(D)$. □

The following lemma gives an internal description of the division ring D from Proposition 1.8 by means of the structure of R itself. If S is a subring of R, then the set

$$\mathrm{Id}_R S = \{r \in R \mid rS \subseteq S \quad \text{and} \quad Sr \subseteq S\}$$

is a subring of R in which S is an ideal. The subring $\mathrm{Id}_R S$ is called the *idealizer* of S in R.

Lemma 1.9. *Let M be a faithful irreducible right R-module and P the annihilator of a non-zero element of M in R. If $S = \mathrm{Id}_R P$, then the factor ring S/P is a division ring whose multiplicative group is isomorphic to the multiplicative group of $\mathrm{End}_R M$.*

Proof. By Lemma 1.2, P is a maximal right ideal of R such that there exists an element $e \in R \setminus P$ with $r - er \in P$ for every $r \in R$. In particular, $e - e^2 \in P$ and $eP \subseteq P$, so that $e \in S$. Show first that $s - se \in P$ for all $s \in S$. Since $(s - se)e = s(e - e^2) \in P$, it suffices to prove that the inclusion $se \in P$ implies $s \in P$ if $s \in S$.

Indeed, in this case $ser \in P$ and $sr - ser = s(r - er) \in P$ for every $r \in R$, so that $sR \subseteq P$. As $eR \nsubseteq P$ and the set $\{r \in R \mid rR \subseteq P\}$ is a right ideal of R containing P, we have $P = \{r \in R \mid rR \subseteq P\}$ and so $s \in P$, as desired.

Thus, e is a unity of S modulo P. Next, for each $s \in S \setminus P$, we have $sR + P = R$ and so $e = sr + x$ for some $r \in R$ and $x \in P$. Therefore $e - sr \in P$ and, in particular, $sr \in S$. This implies $r \in S$ because otherwise $R = rP + P$ and so $sR = srP + sP \subseteq P$, contrary to the above. Clearly $r \notin P$, so that there exists $t \in S$ with $e - rt \in P$. Since the elements $(e - sr)t$, $s(e - rt)$, $t - et$ and $s - se$ are all contained in P, we have $s - t \in P$ and so $e - rs \in P$.

Hence the factor ring $D = S/P$ is a division ring. To prove that the multiplicative groups of D and $\mathrm{End}_R M$ are isomorphic it suffices to assume that $M = R/P$ because the right R-modules M and R/P are isomorphic by Lemma 1.2. For each $d = s + P$ of D and every $m = r + P$ of M, we put $dm = sr + P$ and show that M is a left vector space over D and the mapping $\hat{d} : m \mapsto dm$ is an R-endomorphism of M. .

Indeed, if $d = s_1 + P$ and $m = r_1 + P$ for some $s_1 \in S$ and $r_1 \in R$, then both elements $s - s_1$ and $r - r_1$ belong to P, so that $sr - s_1 r_1 =$

$(s-s_1)r+s_1(r-r_1)\in P$ and therefore $s_1r_1+P=sr+P$. Furthermore, for all $d,d_1\in D$ and $m,m_1\in M$, the relations $d(m+m_1)=dm+dm_1$, $(d+d_1)m=dm+d_1m$ and $(dd_1)m=d(d_1m)$ are immediately verified. Finally, for every $t\in R$, we have $(mt)^{\hat{d}}=d(mt)=srt+P=(dm)t=(m^{\hat{d}})t$, so that $\hat{d}\in \mathrm{End}_R M$.

It is easily seen that the mapping $d^{-1}\mapsto \hat{d}$ determines a group isomorphism $\widehat{\ }$ from $D\setminus\{0\}$ into the multiplicative group of $\mathrm{End}_R M$. Now, if α is a non-zero element of $\mathrm{End}_R M$ and $v=e+P$, then $v^\alpha=s+P=vs$ for some $s\in R\setminus P$. As $P=\mathrm{Ann}_R(v)$ and $v(sP)=(v^\alpha)P=(vP)^\alpha=0$, it follows that $sP\subseteq P$ and so $s\in S\setminus P$. Take the element $d=s+P\in D\setminus\{0\}$. Since $s+P=se+P$ and $m=r+P=vr$, we have $m^\alpha=(vr)^\alpha)=(v^\alpha)r=sr+P=dm$ and so $\alpha=\hat{d}$. Thus the group homomorphism $\widehat{\ }$ is surjective and so the multiplicative groups of D and $\mathrm{End}_R M$ are isomorphic. □

2 The Jacobson radical and its properties.

For a ring R, the intersection of the annihilators of all simple right R-modules is an ideal $J(R)$ of R called the *Jacobson radical* of R. As we shall see further, $J(R)$ coincides also with the intersection of the annihilators of all simple left R-modules. But before we give a characterization of $J(R)$ by internal properties of the ring R.

Consider on R the "circle multiplication" $r\circ s=r+s+rs$ for all $r,s\in R$ under which the set of all elements of R forms a semigroup R^{ad} with neutral element $0\in R$. The group of all invertible elements of this semigroup is called the *adjoint group* of R and denoted by R°. The following lemma describes a relationship between $J(R)$ and R°.

Lemma 2.1. *Let R be a ring and J a right ideal of R. Then $J\subseteq J(R)$ if and only if $J\subseteq R^\circ$.*

Proof. Let $J\subseteq J(R)$ and assume that $J\not\subseteq R^\circ$. Then there exists an element $a\in J$ such that the set $P=\{r+ar\mid r\in R\}$ is a proper right ideal of R.

Indeed, otherwise for every $a\in J$ there exist elements $b,c\in R$ such that $b+ab=-a$ and $c+bc=-b$. This means that $a\circ b=b\circ c=0$ and so $a=a\circ b\circ c=c$. Therefore $a\circ b=b\circ a=0$ and hence $a\in R^\circ$, contrary to the assumption.

Clearly $R=aR+P$ and thus a does not belong to P, so that by Zorn's lemma there exists a maximal right ideal S of R containing P. Since the element $-a$ is a left unity modulo S, the right R-module $M=R/S$ is irreducible by Lemma 1.2 and $Ma=M\neq 0$, contrary to the inclusion $a\in J\subseteq J(R)$. Thus $J\subseteq R^\circ$.

Conversely, let $J\subseteq R^\circ$ and $J\not\subseteq J(R)$. Then there exists an irreducible right R-module M such that $MJ\neq 0$ and so $mJ=M$ for some $m\in M$. Therefore $ma=-m$ for some $a\in J$. Since $a\circ b=0$ for some $b\in R$, we

have $-m = ma + (m + ma)b = m(a \circ b) = 0$ and therefore $M = 0$, contrary to the choice of M. Thus $J \subseteq J(R)$. □

The following characterization of the Jacobson radical of R is an immediate consequence of Lemma 2.1.

Corollary 2.2. *For every ring R, the Jacobson radical of R is the largest ideal of R which is contained in the adjoint group R° of R.*

By symmetry, Corollary 2.2 implies that $J(R)$ coincides with the intersection of the annihilators of all left simple R-modules.

The second important property of the Jacobson radical of R concerns the adjoint group of a factor ring of R.

Corollary 2.3. *If R is a ring and J is an ideal of R contained in $J(R)$, then J° is a normal subgroup of R° and $(R/J)^\circ = R^\circ/J \cong R^\circ/J^\circ$.*

Proof. Since $J \subseteq R^\circ$ by Lemma 2.1, this implies that $J = J^\circ$ and so $r + J \in (R/J)^\circ$ for some $r \in R$ if and only if $r \in R^\circ$. □

Note that if I is an ideal of R, then in general the adjoint group $(R/I)^\circ$ of the factor ring R/I need not be a factor group of R°. For instance, $(\mathbb{Z}/5\mathbb{Z})^\circ$ has order 4 and $\mathbb{Z}^\circ$ has order 2. On the other hand, it is clear that the image of R° in R/I is contained in $(R/I)^\circ$, so that $(J(R) + I)/I$ lies in $J(R/I)$.

Finally, the third property of $J(R)$ allows us to divide the class of all rings into two important subclasses: the class of rings over which irreducible modules exist and the class of rings whose all simple modules are trivial.

Corollary 2.4. *If R is a ring, then $J(R/J(R)) = 0$.*

Proof. Let J be the full preimage of $J(R/J(R))$ in R. Since $J/J(R) \subseteq (R/J(R))^\circ = R^\circ/J(R)$ by Lemma 2.1 and by Corollary 2.3, this implies $J \subseteq R^\circ$ and so $J = J(R)$. □

We shall say that a ring R is *semiprimitive* if there exists a collection of ideals of R modulo each of which the ring R is (right) primitive and the intersection of all ideals of this collection is zero. The following statement shows that in contrast to primitivity the definition of semiprimitivity is symmetrical.

Lemma 2.5. *A ring R is semiprimitive if and only if $J(R) = 0$.*

Proof. If I is an ideal of R modulo which R is primitive, then $J(R) \subseteq I$ because $J(R) + I/I \subseteq J(R/I) = 0$. Thus, $J(R) = 0$ provided that R is semiprimitive. Conversely, the annihilator of every irreducible right R-module is an ideal of R modulo which R is primitive. Since the intersection of the annihilators of all irreducible right R-modules is $J(R)$, this means that R is semiprimitive provided that $J(R) = 0$. □

The following theorem may be considered as one of the main tools of investigation of the structure of an arbitrary associative ring.

Theorem 2.6. *Every ring is semiprimitive modulo its Jacobson radical.*

Proof. It follows immediately from Corollary 2.4 and Lemma 2.5. □

3 Artinian rings.

Let R be a ring. A right R-module M is called *artinian* if M satisfies the minimal condition for the submodules of M. In other words, every set of submodules of M has a least element with respect to the set-theoretical inclusion.

Obviously submodules and factor modules of artinian modules are artinian. It is also clear that each direct sum of finitely many simple R-modules is an artinian R-module and that every semisimple artinian R-module has only finitely many irreducible direct summands. In particular, vector spaces over a division ring are artinian if and only if they are finite-dimensional.

A ring R is called (right) *artinian* if R, regarded as a right R-module, is artinian. The following description of artinian primitive rings is in fact obtained as an application of the density theorem of Jacobson.

Proposition 3.1. *A primitive ring R is artinian if and only if R is isomorphic to the ring $\mathrm{Mat}_n(D)$ of all $(n \times n)$-matrices over a division ring D for some $n \geq 1$.*

Proof. If R is isomorphic to $\mathrm{Mat}_n(D)$, then R can be considered as an n^2-dimensional vector space over D, so that R satisfies the minimal condition for the subspaces of R. Since every right ideal of R is simultaneously a subspace, the ring R is artinian.

Conversely, let R be an artinian ring, M a faithful irreducible R-module and $D = \mathrm{End}_R M$. Then D is a division ring and M can be viewed as a vector space over D. Show that M is finite-dimensional over D.

Suppose the contrary and let $\{m_1, \ldots, m_s, \ldots\}$ be a basis of M over D. For each $s \geq 1$, let $P_s = \bigcap_{i=1}^{s} \mathrm{Ann}_R(m_i)$ and let f_s denote the linear operator of $\mathrm{End}_D M$ defined by the rule $m_i^{f_s} = 0$ for $i \leq s$ and $m_i^{f_s} = m_i$ for $i > s$. Then $P_{s+1} \subseteq P_s$ and, by the density theorem, there exists some $r \in R$ such that $m_i^{f_s} = m_i r$ for all $1 \leq i \leq s+1$. Therefore $m_i r = o$ for $1 \leq i \leq s$ and $m_{s+1} r = m_{s+1}$, so that $r \in P_s \setminus P_{s+1}$. But then $P_1 \supset \ldots \supset P_s \supset P_{s+1} \supset \ldots$ is an infinite strongly descending chain of right ideals of R, contrary to the artinianity of R .

Hence $\dim_D M = n$ for some $n \geq 1$ and thus R is isomorphic to $\mathrm{End}_D M$ and so $\mathrm{Mat}_n(D)$ by Corollary 1.5. □

A ring R is called *simple* if R has no non-trivial ideals. Clearly every division ring is simple.

Lemma 3.2. *For each $n \geq 1$, the matrix ring $\mathrm{Mat}_n(D)$ over a division ring D is simple.*

Proof. For each pair (i,j) with $1 \leq i,j \leq n$, let E_{ij} denote the $(n \times n)$-matrix whose elements, excepting the unity 1 in the position (i,j), are zero. Define the element δ_{ij} of D by the rule $\delta_{ij} = 0$ if $i \neq j$ and $\delta_{ij} = 1$ if $i = j$. Then $E_{ij}E_{kl} = \delta_{jk}E_{il}$ for any pairs (i,j) and (k,l). Clearly the set $\{E_{ij} \mid 1 \leq i,j \leq n\}$ is a basis of $\mathrm{Mat}_n(D)$, regarded as a vector space over D. Thus a non-zero ideal I of $\mathrm{Mat}_n(D)$ contains a matrix $A = \sum_{i,j} a_{ij}E_{ij}$ with $a_{ij} \in D$ and $a_{pq} \neq 0$ for some pair (p,q). A routine calculation shows that $E_{kl} = a_{pq}^{-1}E_{kp}AE_{ql} \in I$ for each pair (k,l) and so $I = \mathrm{Mat}_n(D)$. □

As an application of Lemma 3.1 and Lemma 3.2, we can now obtain a complete description of the structure of semiprimitive artinian rings.

Theorem 3.3. (Wedderburn-Artin) *A semiprimitive ring R is artinian if and only if R is isomorphic to a direct sum of finitely many matrix rings over division rings.*

Proof. Since each direct sum of primitive rings is a semiprimitive ring, it follows from Proposition 3.1 that R is semiprimitive and artinian. To prove the converse, take a minimal ideal V of R. Clearly there exists an ideal I of R such that $I \cap V = 0$ and the factor ring R/I is primitive. It follows from Proposition 3.1 that R/I is isomorphic to a matrix ring $\mathrm{Mat}_n(D)$ for some division ring D and $n \geq 1$. As $(V+I)/I$ is a non-zero ideal of R/I, we have $R = I + V$ by Lemma 3.2 and so V, as a ring, is isomorphic to $\mathrm{Mat}_n(D)$. Furthermore, $VI = IV \subseteq I \cap V = 0$, so that every ideal of I is an ideal of R. In particular, the Jacobson radical $J(I)$ of I is an ideal of R. Since $J(I) \subseteq I^\circ \subseteq R^\circ$, this implies $J(I) \subseteq J(R) = 0$ by Lemma 2.1 and so I, regarded as a ring, is semiprimitive and artinian. Taking a minimal ideal of I which is simultaneously a minimal ideal of R and repeating the same arguments, we obtain after finitely many steps that R is a direct sum of finitely many ideals each of which, regarded as a ring, is isomorphic to a matrix ring over a division ring. □

If S and T are subrings of a ring R, then ST denotes the additive subgroup of R generated by the set $\{st \mid s \in S, t \in T\}$. It is easily verified that $(ST)U = S(TU)$ for every subring U of R. Furthermore, ST is even a (right) ideal of R provided that both S and T are (right) ideals of R. Finally, for each positive integer n we put $S^1 = S$ and $S^n = S^{n-1}S$. A subring S is called *nilpotent* if $S^n = 0$ for some $n \geq 1$.

Theorem 3.4. *The Jacobson radical of an artinian ring is nilpotent.*

Proof. Let R be an artinian ring and J the Jacobson radical of R. Then the descending chain $J \supseteq \ldots \supseteq J^n \supseteq J^{n+1} \supseteq \ldots$ of ideals of R cannot be infinite, so that $J^n = J^{n+1}$ for some $n \geq 1$. Assume that $J^n \neq 0$ and put $I = \{r \in R \mid rJ^n = 0\}$. Then I is an ideal of R and $J \not\subseteq I$ because otherwise $JJ^n = J^{n+1} = 0$. In particular, $R \neq I$ and so the factor ring R/I contains a minimal non-zero right ideal S/I. Since S/I is a simple right R-module, it must be annihilated by J, so that $SJ \subseteq I$. Therefore $(SJ)J^n = 0$ and hence

$SJ^{n+1} = SJ^n = 0$. This implies $S \subseteq I$, contrary to the choice of S. Thus $J^n = 0$. □

As it follows from Wedderburn-Artin theorem 3.3, every semiprimitive ring is right artinian if and only if it is left artinian. The following example shows that a corresponding statement is not true in general. In other words, there are rings satisfying the minimal condition for the right but not the left ideals.

Example. (Small) Let F be a field and $F(x)$ the field of rational functions over F. Consider the ring R of (2×2)-matrices of the form $\begin{pmatrix} F & F(x) \\ 0 & F(x) \end{pmatrix}$ over $F(x)$. Then $J(R) = \begin{pmatrix} 0 & F(x) \\ 0 & 0 \end{pmatrix}$, so that $J(R)^2 = 0$ and the factor ring $R/J(R)$ is isomorphic to the direct sum of two fields F and $F(x)$. It is easily seen that the right ideals of R contained in the Jacobson radical $J(R)$ are vector spaces over $F(x)$ and, on the other hand, the left ideals of that are vector spaces over F. Since $J(R)$, regarded as a vector space, is 1-dimensional over $F(x)$ and infinite-dimensional over F, it is a minimal right ideal of R and contains infinite strongly descending chains of left ideals of R. Thus the ring R is right artinian but not left artinian.

4 Division rings. Wedderburn's theorem on finite division rings.

4.1 Cyclotomic polynomials.

Let n be a positive integer and let $\epsilon_0, \ldots, \epsilon_{n-1}$ be all roots of degree n from 1 in the field $\mathbb{C}$ of complex numbers. We know that $\epsilon_m = \cos\frac{2\pi m}{n} + i\sin\frac{2\pi m}{n}$ for every $0 \leq m \leq n-1$. The root ϵ_m is primitive if $(m, n) = 1$. The number of all primitive roots of degree n is equal to $\varphi(n)$, Euler's function of n. Let $\delta_1, \ldots, \delta_{\varphi(n)}$ be the primitive roots of degree n. Put $\Phi_n(x) = (x - \delta_1)\ldots(x - \delta_{\varphi(n)})$. Then $\Phi_n(x)$ is the n-th cyclotomic polynomial on x. It is easy to see that

1) $\Phi_n(x) = \prod_{d|n} \Phi_d(x)$, where d runs all distinct divisors of n, and

2) $(\Phi_n(x), \Phi_m(x)) = 1$ provided that $m \neq n$.

Proposition 4.1. *For every positive integer n, the polynomial $\Phi_n(x)$ has integer coefficients, and if d is a proper divisor of n, then $\Phi_n(x)$ is a divisor of $\frac{x^n-1}{x^d-1}$.*

Proof. If $n = p$ is a prime, then $\Phi_p(x) = x^{p-1} + \ldots + x + 1$ and the proposition is proved. By induction on n, for every proper divisor d of n the polynomial $\Phi_d(x)$ has desired properties. Therefore the polynomial $f(x) = \prod_{d|n, d\neq n} \Phi_d(x)$ has integer coefficients and $f(x)$ is a divisor of $x^n - 1$ by 2). Since $x^n - 1 = \Phi_n(x)f(x)$ by 1), $\Phi_n(x)$ has the same property and clearly is coprime with $x^d - 1$ for every $d|n, d \neq n$. Hence $\Phi_n(x)$ is a divisor of $\frac{x^n-1}{x^d-1}$. □

4.2 Finite division rings.

The following important theorem was proved by Wedderburn almost hundred years ago. As usual, if D is a division ring, then D^* denotes the multiplicative group of D.

Theorem 4.2. (Wedderburn) *Every finite division ring is commutative.*

Proof. Let D be a finite division ring and Z its centre. Then Z is a field, say of order q, and D is a vector space over Z. Therefore the order of D is q^n for some integer $n \geq 1$. Furthermore, for every $a \in D^*$ the centralizer $C_D(a)$ of a in D is a division subring of D containing Z, so that the order of $C_D(a)$ is equal to q^m for some divisor m of n. This means that the conjugacy class of D^* containing a has $\frac{q^n-1}{q^m-1}$ elements. Since D^* is a disjoint union of its conjugacy classes, we have

$$q^n - 1 = q - 1 + \sum_m \frac{q^n - 1}{q^m - 1},$$

where m runs some number of proper divisors of n. By Proposition 4.1, the integer $\Phi_n(q)$ is a divisor of $q^n - 1$ and each summand $\frac{q^n-1}{q^m-1}$, so that $q-1$ must also be divisible by $\Phi_n(q)$. On the other hand, $|\Phi_n(q)| = \prod_{j=1}^{\varphi(n)} |q - \delta_j| \geq (q-1)^{\varphi(n)} \geq q - 1$ and there is an equality only if $n = 1$. This implies $Z = D$ and hence D is commutative. □

5 The multiplicative group of division rings. Theorems of Cartan-Brauer-Hua, Scott and Jacobson.

We will use the parenthesis for group commutators of division rings: $(a, b) = a^{-1}b^{-1}ab$.

Let D be a division ring and $a, b \in D^*$ with $a \neq 1$. Then

$$b^{-1}(a-1)b = b^{-1}ab - 1.$$

Since $b^{-1}ab = a(a, b)$, we can rewrite this equality as

$$(a-1)(a-1, b) = a(a, b) - 1,$$

or also

$$a((a, b) - (a-1, b)) = 1 - (a-1, b). \tag{1}$$

Clearly if $ab \neq ba$, then both sides of (1) are non-zero, so that every non-central element of D lies in the division subring generated by all the group commutators.

Theorem 5.1. (Cartan-Brauer-Hua) *Let D be a division ring with centre Z and let K be a division subring of D such that K^* is a normal subgroup of D^*. Then either $K \subseteq Z$ or $K = D$.*

Proof. Suppose that $K \nsubseteq Z$ and let $b \in K \setminus Z$. Take an element $a \in D$ such that $ab \neq ba$. Since $1 \neq (a,b) = (a^{-1}b^{-1}a)b \in K$ and similarly $1 \neq (a-1,b) \in K$, both sides of (1) are non-zero and we have $a = (1-(a-1,b))((a,b)-(a-1,b))^{-1} \in K$. Thus, every element of D not commuting with b lies in K. But if c commutes with b and a does not, then both a and $a+c$ have to lie in K, so that $c \in K$. Therefore $K = D$. □

We show next that the multiplicative group of non-commutative division rings cannot be soluble.

Lemma 5.2. *Any finite subgroup of a division ring D of finite characteristic is cyclic.*

Proof. Let G be a finite subgroup of D^* and P be the prime subfield of D. Then P is finite and so the P-algebra F generated by G is finite because F is finite-dimensional over P. Since F is a division subring of D, it is commutative by Wedderburn's theorem and thus G is cyclic. □

Theorem 5.3. (Scott) *Let D be a division ring with centre Z. Then the factor group D^*/Z^* has no non-trivial abelian normal subgroups. In particular, if D^* is soluble, then D is commutative.*

Proof. Suppose the contrary and let A be a non-central normal subgroup of D^* such that $Z^* \subseteq A$ and the factor group A/Z^* is abelian. Then the division subring of D generated by A satisfies the hypothesis of Theorem 5.1 and so it must coincide with D. Therefore the subgroup A is non-abelian and hence there exist elements $a, b \in A$ such that $1 \neq (a,b) = c \in Z^*$. Show that Z^* is a subgroup of order 2 in D^* and $a^2, b^2 \in Z^*$.

Indeed, as

$$(1+a,b) = (1+a)^{-1}b^{-1}(1+a)b = (1+a)^{-1}(1+ac),$$

we have

$$(1+a,b,b) = ((1+a)^{-1}(1+ac),b) =$$

$$(1+ac)^{-1}(1+a)(1+ac)^{-1}(1+ac^2) = (1+a)(1+ac)^{-2}(1+ac^2),$$

and similarly

$$(1+a,b,b,b) = (1+a)^{-1}(1+ac)^3(1+ac^2)^{-3}(1+ac^3).$$

On the other hand, $(1+a,b) \in A$ so that $(1+a,b,b,b) = 1$. Hence

$$(1+ac)^3(1+ac^3) = (1+a)(1+ac^2)^3 = 0.$$

Reducing identical terms in this equality, we obtain

$$a(3c + c^3) + a^3(c^3 + 3c^5) = a(1 + 3c^2) + a^3(3c^4 + c^6)$$

or equivalently

$$(1 - c)^3 + a^2c^3(-1 + c)^3 = 0.$$

Since $1 - c \neq 0$, this implies $a^2 = c^{-3} \in Z^*$. By symmetry, also $b^2 \in Z^*$. Furthermore, if $z \in Z^*$, then likewise $(az)^2 = c^{-3}$ because $az \in A$ and $(az, b) = c$. Therefore $z^2 = 1$ and hence Z^* has exponent 2 and so order 2.

Thus Z is the prime subfield of order 3 in D and a, b are elements of finite order of D^*. Moreover, the subgroup of D^* generated by a, b is finite because it is nilpotent. Therefore this subgroup must be cyclic by Lemma 5.2 and hence $(a, b) = 1$, contrary to the choice of these elements. □

Finally, we generalize Wedderburn's theorem 4.2 by proving that a division ring D is commutative provided that D^* is periodic.

Lemma 5.4. (Herstein) *Let D be a division ring of prime characteristic p with centre Z and let a be a non-central element of D such that $a^n = 1$ for some integer n. Then there exists an element $b \in D^*$ and an integer m such that $b^{-1}ab = a^m \neq a$.*

Proof. Let P be a prime subfield of Z and $F = P(a)$. Since a is algebraic over P, the field F is finite and so has p^l elements for some positive integer $l > 1$. Therefore every element of F is a root of the polynomial $x^{p^l} - x$ over F, so that $x^{p^l} - x = \prod_{c \in F}(x - c)$.

Define a mapping $\alpha : D \to D$ by the rule $u^\alpha = [u, a] = ua - au$ for every $u \in D$. Then α is a linear transformation of D as a vector space over Z and it is easy to verify that $u^{\alpha^{p^s}} = [u, a^{p^s}]$ for every positive integer s. Since $a^{p^l} = a$, we have $\alpha^{p^l} = \alpha$. Now, for any $c \in F$ and the identity mapping ι of D, the linear transformation ιc is permutable with α. Indeed, $u^{((\iota c)\alpha)} = (uc)^\alpha = [uc, a] = [u, a]c = (u^\alpha)c = u^{(\alpha(\iota c))}$. Hence $0 = \alpha^{p^l} - \alpha = \prod_{c \in F}(\alpha - \iota c)$. This means that there exists a minimal non-negative integer k such that $\alpha(\alpha - \iota c_1) \dots (\alpha - \iota c_k) = 0$ for some non-zero distinct elements $c_1, \dots, c_k$ of F. Clearly $k \geq 1$ because $\alpha \neq 0$. Thus $\beta = \alpha(\alpha - \iota c_1) \dots (\alpha - \iota c_{k-1}) \neq 0$ and so there exists $u \in D$ such that $b = u^\beta \neq 0$. On the other hand, $b^{(\alpha - \iota c_k)} = u^{\beta(\alpha - \iota c_k)} = 0$ so that $bc_k = b^\alpha = ba - ab$. This implies that $b^{-1}ab = a - c_k \in F$ and $b^{-1}ab \neq a$ because $c_k \neq 0$. Since the multiplicative group F^* is cyclic and its elements a and $b^{-1}ab$ have the same order, there exists an integer m such that $b^{-1}ab = a^m \neq a$. □

Theorem 5.5. (Jacobson) *Let D be a division ring whose multiplicative group D^* is periodic. Then D is commutative.*

Proof. Since the multiplicative group of the field of rational numbers is non-periodic, the prime subfield of D must be finite and so D is of finite characteristic. Let Z be the centre of D and assume that $D \neq Z$. Take $a \in D \setminus Z$.

By Lemma 5.4, there exists $b \in D^*$ such that $b^{-1}ab = a^m \neq a$ for some integer m. The subgroup H generated by a, b is metacyclic and so finite. Therefore H must be cyclic by Lemma 5.2, contradicting the fact that $b^{-1}ab \neq a$. Thus $D = Z$. □

6 Radical rings.

A ring R is called *radical* if R coincides with its Jacobson radical. This means that R has no irreducible right R-modules or equivalently that every simple right R-module is trivial. It follows from Corollary 2.2 that R is radical if and only if $R = R^\circ$. For instance, every *null* ring, i.e. a ring R with $R^2 = 0$, is radical because its adjoint group R° coincides with the additive group R^+.

An (one-sided) ideal of a ring R will be called *radical* if it is radical as a ring.

Lemma 6.1. *Let R be a ring and P an one-sided ideal of R. Then P is radical if and only if P is contained in the Jacobson radical $J(R)$ of R. In particular, every ring generated by its radical one-sided ideals is radical.*

Proof. Clearly $P \subseteq R^\circ$ if $P = P^\circ$. Conversely, if $r \in P \subseteq R^\circ$ and $r \circ s = 0$, then $s = -r - rs \in P$ and so $P = P^\circ$. Thus, if P is a right ideal, it is radical if and only if $P \subseteq J(R)$ by Lemma 2.1. By symmetry, the same is true for the left ideal P. In particular, if R is generated by radical one-sided ideals, then $R = J(R)$. □

For each ring R, let R_u denote the ring obtained by adjoining a formal unity 1 to R when R has no unity, and $R_\mathrm{u} = R$ otherwise. Recall that R is an ideal of R_u such that every element of R_u can uniquely be written in the form $n + r$ for some $n \in \mathbb{Z}$ and $r \in R$. Moreover, the set $1 + R^\circ$ is a subgroup of the multiplicative group R_u^* of R_u and the mapping $r \mapsto 1 + r$ with $r \in R$ is an isomorphism from R° onto $1 + R^\circ$. In particular, if R itself is a ring with 1, then $1 + R^\circ = R^*$.

If R is a radical ring, then $1 + R$ is a subgroup of R_u^* because $R^\circ = R$. Clearly this property characterizes radical rings. It allows us to add a formal unity 1 to R and to consider the multiplicative group $1 + R$ instead of R°. Using this, it is easily verified that the direct limits, homomorphic images and cartesian products of radical rings are also radical. Furthermore, the one-sided ideals of radical rings are radical by Lemma 6.1. On the other hand, the subrings of a radical ring need not be radical.

Example 6.2. *Let p be a prime and $\mathbb{Q}_p$ the set of all rational numbers whose denominators are not divisible by p. Then $p\mathbb{Q}_p$ is a radical ring because $1+p\mathbb{Q}_p$ is a subgroup of the multiplicative group of rational numbers while the subring $p\mathbb{Z}$ of $p\mathbb{Q}_p$ is not radical.*

The following theorem indicates another important property of radical rings.

Theorem 6.3. *Let n be a positive integer and R a ring. The ring $\mathrm{Mat}_n(R)$ of all $(n \times n)$-matrices over R is radical if and only if R is radical.*

Proof. Let I denote the identity $(n \times n)$-matrix over R_{u} and E_{ij} the $(n \times n)$-matrix whose elements, excepting the unity 1 in the position (i, j), are zero. If the ring $\mathrm{Mat}_n(R)$ is radical, then for every $r \in R$ the matrix $I + rE_{11}$ is invertible in $\mathrm{Mat}_n(R_{\mathrm{u}})$. Therefore $1 + r$ is invertible in R_{u} and hence R is a radical ring.

Conversely, let the ring R be radical and for each $i = 1, \ldots, n$ let P_i denote the set of all matrices of $\mathrm{Mat}_n(R)$ whose rows other than i-th consist of zero elements. Then P_i is a right ideal of $\mathrm{Mat}_n(R)$. For each $j = 1, \ldots, n$, put next $P_{ij} = \{rE_{ij} \mid r \in R\}$. Then P_{ij} is a left ideal of P_i and $P_i = P_{1i} + \ldots + P_{ni}$. Furthermore, $P_{ij}^2 = 0$ if $i \neq j$ and P_{ii}, regarded as a ring, is isomorphic to R. Therefore every P_{ij} is a radical left ideal of P_i and so P_i is a radical right ideal of $\mathrm{Mat}_n(R)$ by Lemma 6.1. Since $\mathrm{Mat}_n(R) = P_1 + \ldots + P_n$, this implies that the ring $\mathrm{Mat}_n(R)$ is radical. □

A ring R is called *nil* if every element r of R is *nilpotent*, i.e. there exists a positive integer $n = n(r)$ such that $r^n = 0$. Obviously every nilpotent ring is nil. It is clear that each subring of a nil ring is also nil and that every nil ring is radical because the element $1 - r + r^2 - \ldots + (-1)^{n-1} r^{n-1}$ is an inverse for $1 + r$. The following proposition shows that, in contrast to the matrix rings, the polynomial rings over radical rings need not be radical.

Proposition 6.4. *Let R be a ring and x an indeterminate. If the ring $R[x]$ of all polynomials in x over R is radical, then R is a nil ring.*

Proof. If $R[x]$ is radical, then for each $r \in R$ there exists a polynomial $f(x) \in R[x]$ such that $(1 + f(x))(1 - rx) = 1$. Let n be the degree of $f(x)$. Since $1 - (rx)^{n+1} = (1 - rx)(1 + rx + \ldots + (rx)^n)$, it follows that $(1 + f(x))(1 - (rx)^{n+1}) = 1 + rx + \ldots + (rx)^n$ and hence $f(x) = rx + \ldots + (rx)^n + (rx)^{n+1} + f(x)(rx)^{n+1}$. This implies $f(x) = rx + \ldots + (rx)^n$ and therefore $(rx)^{n+1} = 0$, so that $r^{n+1} = 0$. □

It is unknown at present whether the polynomial ring $R[x]$ over a nil ring R is radical. However, as it has recently been shown by Smoktunowicz (2000), polynomial rings over nil rings need not be nil. On the other hand, we show next that the polynomial rings over radical rings can even be semiprimitive. But first we give a description of the polynomials over a commutative ring A which belong to the adjoint group $A[x]^\circ$ of the polynomial ring $A[x]$.

Recall that A is an *integral domain* if every product of two non-zero elements of A is non-zero.

Lemma 6.5. *Let A be a commutative ring. Then the following statements hold.*

1) *If a is a non-nilpotent element of A, then there exists an ideal P of A such that $a \in P$ and the factor ring A/P is an integral domain.*

2) A polynomial $f(x) = a_0 + a_1x + \ldots + a_nx^n \in A[x]$ *belongs to the adjoint group* $A[x]^\circ$ *of* $A[x]$ *if and only if* $a_0 \in A^\circ$ *and the coefficients* $a_1, \ldots, a_n$ *are nilpotent.*

Proof. 1) Since $a^n \neq 0$ for every positive integer n, in view of Zorn's lemma there exists an ideal P of A which is maximal with respect to the condition $P \cap \{a^n \mid n \geq 1\} = \emptyset$. Show that factor ring A/P is an integral domain.

Indeed, otherwise there exist elements $b, c \in A \setminus P$ such that $bc \in P$. If B and C are the ideals of A generated by b and c, respectively, then P is properly contained in $B + P$ and $C + P$, so that $a^l \in B + P$ and $a^m \in C + P$ for some positive integers l, m. But then $a^{l+m} \in (B + P)(C + P) \subseteq P$, contrary to the choice of P.

2) Let $f(x) \in A[x]^\circ$, so that there exists a polynomial $g(x) = b_0 + b_1x + \ldots + b_mx^m \in A[x]$ with $b_m \neq 0$ such that $f(x) + g(x) + f(x)g(x) = 0$. Then $a_0 + b_0 + a_0b_0 = 0$ and so $a_0 \in A^\circ$. For $n > 0$ and $a_n \neq 0$, it follows that $a_nb_m = 0$ for $m > 0$ and $a_nb_0 + a_n = 0$ for $m = 0$, so that $a_1 = \ldots = a_n = 0$, provided that A is an integral domain.

Assume that the coefficient a_i is not nilpotent for some $i > 0$. Then there exists an ideal P of A such that $a_i \notin P$ and the factor ring $\bar{A} = A/P$ is an integral domain by 1). Clearly $P[x]$ is an ideal of $A[x]$ and the factor ring $A[x]/P[x]$ is isomorphic to the polynomial ring $\bar{A}[x]$. But then the coefficients $a_1, \ldots, a_n$ are zero modulo P by the above and this means in particular that $a_i \in P$, contrary to the choice of P. Thus these coefficients $a_1, \ldots, a_n$ must be nilpotent.

Conversely, if $a_0 \in A^\circ$ and $a_1, \ldots, a_n$ are nilpotent elements of A, then $a_0 \in A[x]^\circ$ and the polynomial $h(x) = a_1x + \ldots + a_nx^n$ is nilpotent. Furthermore, there exists $b_0 \in A$ such that $a_0 + b_0 + a_0b_0 = 0$ and so $f(x) = a_0 + a_0(1 + b_0)h(x) + (1 + b_0)h(x) = (a_0) \circ ((1 + b_0)h(x))$. Since the polynomial $(1 + b_0)h(x)$ is also nilpotent, it belongs to $A[x]^\circ$ and so $f(x) \in A[x]^\circ$. $\square$

Theorem 6.6. *Let* R *be a ring without non-zero nil ideals. Then the Jacobson radical of* $R[x]$ *is trivial, i.e.* $R[x]$ *is a semiprimitive ring.*

Proof. Assume the contrary and let $f(x) = a_0 + a_1x + \ldots + a_nx^n$ be a non-zero polynomial of the Jacobson radical $J(R[x])$ such that the number of its non-zero coefficients is minimal and $a_n \neq 0$. Then each polynomial $a_if(x) - f(x)a_i$ belongs to $J(R[x])$ and the number of its non-zero coefficients is less than in $f(x)$. Therefore $a_if(x) = f(x)a_i$ for every a_i and hence $a_ia_j = a_ja_i$ for all i, j. Thus the subring A of R generated by all coefficients of $f(x)$ is commutative and so for each positive integer m the identity $1 - (xf(x))^{m+1} = (1 - xf(x))(1 + xf(x) + \ldots + (xf(x))^m)$ holds.

Since $xf(x) \in J(R[x])$, there exists a polynomial $g(x) \in J(R[x])$, say of a degree m, such that $(1 + g(x))(1 - xf(x)) = 1$. Multiplying the above identity for $xf(x)$ from the left on $1 + g(x)$, we have $(1 + g(x))(1 - (xf(x))^{m+1}) = (1 + xf(x) + \ldots + (xf(x))^m)$ and therefore $g(x) = xf(x) + \ldots + (xf(x))^m + (xf(x))^{m+1} + g(x)(xf(x))^{m+1}$. Comparing the degrees of both sides, we obtain that the coefficients of the polynomial $g(x)$ are contained among the coefficients

of the polynomial $xf(x)+\ldots+(xf(x))^m$ and so in A. Therefore the polynomial $-xf(x)$ belongs to the adjoint group $A[x]^\circ$ of the polynomial ring $A[x]$ and hence its coefficients $a_0, \ldots, a_n$ are nilpotent by Lemma 6.5.2).

Clearly, for each finite set $\{r_1, s_1, \ldots, r_k, s_k\}$ of elements of R, the polynomial $\sum_{i=1}^{k} r_i f(x) s_i$ belongs to $J(R[x])$. By the choice of $f(x)$, this polynomial can be either zero or such that the number of its non-zero coefficients is the same as in $f(x)$. By proved above, this means that the coefficient $\sum_{i=1}^{k} r_i a_n s_i$ is nilpotent and so the non-zero ideal of R generated by a_n is nil, contrary to the hypothesis of the theorem. Thus $J(R([x]) = 0$, as desired. □

Since subrings of a radical ring R need not be radical, for a subset S of R we define the *radical join* $< S >_{rad}$ of S in R to be the intersection of all radical subrings of R containing S. If $R =< S >_{rad}$, we will say that R is the *radical ring on* S as a set of generators. In particular, if $S = \{r\}$, then $R =< r >_{rad}$ is the radical ring on the single generator r. Note that the radical ring $p\mathbb{Q}_p$ from Example 6.2 has this property with the prime p as a generator. On the other hand, $p\mathbb{Q}_p$ is not finitely generated as a ring because each non-zero subring of $\mathbb{Q}$ generated by finitely many elements is in fact isomorphic to a non-zero subring generated by one element and so it cannot be radical. Indeed, the subring generated by an irreducible fraction $\frac{m}{n}$ cannot contain the number $(1 + \frac{m}{n})^{-1} = \frac{n}{m+n}$.

It is easy to see that every ring R generated by one element is a homomorphic image of the ring $x\mathbb{Z}[x]$. Indeed, if r is a generator of R, then the mapping $xf(x) \mapsto rf(r)$ with $f(x) \in \mathbb{Z}[x]$ is a ring homomorphism from $x\mathbb{Z}[x]$ onto R. The following lemma shows that every radical homomorphic image of the ring $x\mathbb{Z}[x]$ is nilpotent.

Lemma 6.7. *Let R be a ring generated by one element. If R is radical, then R is nilpotent.*

Proof. Assume the contrary and let r be a generator of R. Then $r^n \neq 0$ for every positive integer n because r^n is a generator of the subring R^n. Therefore there exists an ideal I of R such that $r \notin I$ and the factor ring R/I is an integral domain by Lemma 6.5.1). Passing to the factor ring R/I, we may restrict ourselves to the case that R is an integral domain and so a subring of its field of quetients F. Since R is radical, $1 + R$ is a subgroup of the multiplicative group of F, so that $(1 + r)^{-1} = 1 + rf(r)$ for some non-zero polynomial $f(x) \in \mathbb{Z}[x]$. Therefore $(1 + r)(1 + rf(r)) = 1$ which means that r is an algebraic element over the prime subfield Q of F. Clearly $R \subseteq Q[r]$, so that R is finite if Q is. But then R is nilpotent by Theorem 3.4, contrary to the assumption. Thus Q is the field of rational numbers and the intersection $Q \cap R$ is a radical subring of Q because for every $z \in Q \cap R$ we have $(1 + z)^{-1} \in Q \cap (1 + R) = 1 + (Q \cap R)$.

Let $g(x) = z_0 + z_1 x + \ldots + z_n x^n$ be a polynomial of the least degree with integers coefficients $z_0, \ldots, z_n$ such that $g(r) = 0$ and so $z_n \neq 0$. If S is the subring of Q generated by z_n^{-1}, then $R \subseteq S + Sr + \ldots + Sr^{n-1}$ because

$r^n = z_n^{-1}(z_0 + z_1 r + \ldots + z_{n-1} r^{n-1})$. Therefore, for each $z \in Q \cap R$, there exist elements $s_1, \ldots, s_n$ such that $z = s_1 + s_2 r + \ldots + s_n r^{n-1}$. Since the elements $1, r, \ldots, r^{n-1}$ are linear independent over Q, this implies $z = s_1 \in S$ and hence $Q \cap R \subseteq S$. However, it is easy to see that the ring S has no non-zero radical subrings. Therefore $Q \cap R = 0$ and hence $z_0 = 0$ because $z_0 = -z_1 r - \ldots - z_n r^n \in Q \cap R$. But then $z_1 + z_2 r + \ldots + z_n r^{n-1} = 0$ and thus $z_1 = \ldots = z_n = 0$. This contradiction completes the proof. □

The following assertion is a direct consequence of Lemma 6.7.

Corollary 6.8. *A radical ring R is nil if and only if every subring of R is radical.*

7 Nil rings and Golod's construction.

A ring R is said to be *locally nilpotent* if every finitely generated subring of R is nilpotent. It is easy to see that the class of locally nilpotent rings properly contains the class of nilpotent rings and is a subclass of the class of nil rings. For instance, the direct sum $\bigoplus_{n=1}^{\infty} p\mathbb{Z}/p^{n+1}\mathbb{Z}$ is a locally nilpotent ring which is non-nilpotent. The question on existence of nil rings which are not locally nilpotent is more complicated. The first example of such a ring was given by Golod. A simplified proof of Golod's theorem is due to Ol'shanskii and below we produce this proof. Recall that R is an *algebra* over a commutative ring K with unity or more shortly a K-algebra if R is simultaneously a ring and a module over K such that $(kr)s) = r(ks) = k(rs)$ for all elements $k \in K$ and $r, s \in R$.

Let F be a field and let $F\langle x, y\rangle$ denote the algebra of all polynomials in two non-commuting indeterminates x, y over F. Denote by A the subalgebra of the polynomials of $F\langle x, y\rangle$ without constant terms. It is clear that A is generated as an F-algebra by x, y. Furthermore A, regarded as a vector space over F, is a direct sum $A = A_1 \oplus A_2 \oplus \ldots$ of its *homogeneous components* A_i whose elements are the homogeneous polynomials of degrees $i = 1, 2, \ldots$, so that $A_1 = Fx \oplus Fy$, $A_2 = Fx^2 \oplus Fxy \oplus Fyx \oplus Fy^2$ and so on. An ideal I of A is called *homogeneous* if $I = I_1 \oplus I_2 \oplus \ldots$, where $I_n = A_n \cap I$ for every $n \geq 1$. It is easy to see that the factor algebra A/I is nilpotent if and only if there exists a positive integer m such that $I_n = A_n$ for all $n \geq m$. It is also obvious that in this case $\dim_F A/I$ is finite.

Lemma 7.1. *Let I be the homogeneous ideal of A generated by a set $\{r_6, r_7, \ldots\}$ of elements $r_i \in A_i$ for all $i \geq 6$. Then $I_n \neq A_n$ for every $n \geq 1$.*

Proof. Since $I = \sum_{i=6}^{\infty} (F + A) r_i (F + A)$, every component I_n is a linear envelope of all products $ur_j v$ where $6 \leq j \leq n$ and u, v are monomials. If $\deg v > 0$, then $v \in {A_1}^s$ for some positive integer s and so $ur_j v \in I_{n-1} A_1$. If $\deg v = 0$, i.e. v is an empty word, then $ur_j v = ur_j \in A_{n-j} r_j$. In both cases

we obtain

$$I_n \subseteq I_{n-1}A_1 + \sum_{j=6}^{n} A_{n-j}r_j. \tag{2}$$

Choose a subspace C_{n-j} in A_{n-j} such that $A_{n-j} = I_{n-j} \oplus C_{n-j}$. Then

$$A_{n-j}r_j = I_{n-j}r_j + C_{n-j}r_j. \tag{3}$$

Equality (3) and the obvious inclusion $I_{n-j}r_j \subseteq I_{n-1}A_1$ allow us to rewrite (2) as follows

$$I_n \subseteq I_{n-1}A_1 + \sum_{j=6}^{n} C_{n-j}r_j. \tag{4}$$

Put $c_n = \dim_F A_n - \dim_F I_n = \dim_F C_n$ and show that $c_n > 0$ for every $n \geq 1$. Clearly, $c_n = \dim_F A_n = 2^n$ for $1 \leq n \leq 5$. By induction on n, we will prove that $c_n \geq \frac{3}{2}c_{n-1}$ for all $n \geq 6$. Let $d_n = \dim_F I_n$. Then from (4) it follows that $d_n \leq 2d_{n-1} + \sum_{j=6}^{n} c_{n-j}$. Next $d_n = \dim_F A_n - c_n = 2^n - c_n$. Therefore

$$2^n - c_n \leq 2(2^{n-1} - c_{n-1}) + \sum_{j=6}^{n} c_{n-j}. \tag{5}$$

By induction hypothesis, $c_{n-6} \leq (\frac{2}{3})^5 c_{n-1}$, $c_{n-7} \leq (\frac{2}{3})^6 c_{n-1}$ and so on. Hence inequality (5) implies

$$c_n \geq 2c_{n-1} - \sum_{j=6}^{n} c_{n-j} \geq 2c_{n-1} - c_{n-1}((\frac{2}{3})^5 + (\frac{2}{3})^6 + \ldots) \geq$$

$$\geq c_{n-1}(2 - \frac{(\frac{2}{3})^5}{1 - \frac{2}{3}}) = c_{n-1}(2 - \frac{32}{81}) \geq \frac{3}{2}c_{n-1}.$$

Thus, $c_n > 0$ for every $n \geq 1$ and the lemma is proved. □

A ring (an algebra) R is called *residually nilpotent* if there exists a collection of ideals of R modulo each of which R is nilpotent and the intersection of all ideals of this collection is zero.

Theorem 7.2. (Golod) *Let F be a countable field. Then there exists a two generated nil algebra R over F which is residually nilpotent and whose F-dimension $\dim_F R$ is infinite. In particular, the algebra R is not nilpotent.*

Proof. Since the algebra A is countable, we can enumerate all elements of A. Let for instance $A = \{a_1, a_2, \ldots\}$. Choose a positive integer $n_1 \geq 6$ and write the element ${a_1}^{n_1}$ in the form of a sum of its homogeneous components. Then ${a_1}^{n_1} = r_{1\,6} + r_{1\,7} + \ldots + r_{1\,m_1}$ and each $r_{1\,i}$ is an element of A_i where $6 \leq i \leq m_1$. Next let $n_2 > m_1$. Then ${a_2}^{n_2} = r_{2\,m_1+1} + \ldots + r_{2\,m_2}$ and each $r_{2\,i}$ is an element of A_i where $m_1+1 \leq i \leq m_2$, and so on. By the choice of the integers n_j, each component A_i of A contains at most one element of the set $S = \{r_{1\,6}, \ldots, r_{1\,m_1}, r_{2\,m_1+1}, \ldots, r_{2\,m_2}, \ldots\}$. Therefore the homogeneous ideal

I generated by S satisfies the condition of Lemma 7.1 and hence $I_n \neq A_n$ for every $n \geq 1$. The factor algebra $R = A/I$ is a nil algebra over F because every element of R is nilpotent by the choice of S. Moreover, R is generated by two elements $x+I, y+I$ and $R = R_1 \oplus \ldots \oplus R_n \oplus \ldots$ where each $R_n = (A_n + I)/I$ as a vector space over F is isomorphic to A_n/I_n and so is non-trivial. Thus $\dim_F R$ is infinite and so the algebra R cannot be nilpotent. Furthermore, if $P_n = R_n \oplus R_{n+1} \oplus \ldots$ for each $n \geq 1$, then P_n is an ideal of R such that the factor algebra R/P_n is nilpotent and $\bigcap_{n=1}^{\infty} P_n = 0$. This means that R is residually nilpotent and the theorem is proved. □

A similar approach can be applied for a proof of the original theorem of Golod on the existence of an infinite dimensional nil algebra R over F with $d \geq 2$ generators such that every $(d-1)$-generated subalgebra of R is nilpotent.

The following two lemmas show that the nilpotent subrings of rings have some special properties which are useful in particular in studying nil rings.

Lemma 7.3. *Let R be a ring and S a nilpotent proper subring of R. Then* $\mathrm{Id}_R S \neq S$.

Proof. Show first that there exists a subring T of R containing S such that S is a proper right ideal of T. Indeed, if $SR \subseteq S$, then $T = R$. Otherwise there exists a maximal positive integer n such that $S^n R \not\subseteq S$. Therefore $S^{n+1}R \subseteq S$. Hence the set $T = S + S^n R$ is a subring of R properly containing S and S is a right ideal of T. Repeating the same arguments from the left side, we obtain that there exists a subring U of T which contains S as a proper left ideal of U. But then S is an ideal of U and so $S \subsetneqq U \subseteq \mathrm{Id}_R S$. □

Lemma 7.4. *Let R be a ring and P a (locally) nilpotent right ideal of R. Then P is contained in a (locally) nilpotent ideal of R.*

Proof. Consider R as an ideal in R_{u}. Then $R_{\mathrm{u}}P$ is an ideal of R and $(R_{\mathrm{u}}P)^n \subseteq R_{\mathrm{u}}P^n$ for each positive integer n. In particular, $(R_{\mathrm{u}}P)^n = 0$ if $P^n = 0$, so that the ideal $R_{\mathrm{u}}P$ is nilpotent if P is.

Let P be locally nilpotent and let S be a finite subset in $R_{\mathrm{u}}P$. Since every element of $R_{\mathrm{u}}P$ can be written as a finite sum of elements of the form ab with $a \in R_{\mathrm{u}}$ and $b \in P$, there exist finitely generated subrings A of R_{u} and B of P such that $S \subseteq AB$. The product BA is contained in P and so the subring generated by BA is nilpotent because it is finitely generated. Therefore $(BA)^n = 0$ for some $n \geq 1$ and hence $(AB)^{n+1} = A(BA)^n B = 0$. This implies that the subring generated by AB and so by S is nilpotent and thus the ideal $R_{\mathrm{u}}P$ is locally nilpotent □

Note that the question whether every nil one-sided ideal of a ring is contained in a nil ideal of this ring is still open. This is a famous problem of Koethe. As was proved by Krempa, there are many equivalent statements to this problem. For instance, whether the polynomial ring $R[x]$ is radical or the matrix ring $\mathrm{Mat}_2(R)$ is nil if R is nil. On the other hand, it is easy to check that a ring R

is nilpotent or locally nilpotent if and only if the matrix ring $\mathrm{Mat}_n(R)$ or the polynomial ring $R[x]$ have the same property, respectively.

The only maximal locally nilpotent ideal of a ring R is called the *Levitzki radical* of R. The following theorem shows that this radical exists in every ring and notes some of its properties.

Theorem 7.5. (Levitzki) *Let R be a ring and $L(R)$ the sum of all locally nilpotent right ideals of R. Then $L(R)$ is the only maximal locally nilpotent ideal of R and the factor ring $R/L(R)$ has no non-trivial locally nilpotent right ideals. In particular, $L(R/L(R)) = 0$.*

Proof. Note first that a subring S of R is locally nilpotent if S is the same modulo a locally nilpotent ideal N of R. Indeed, if K is a finitely generated subring of S, then the factor ring $K/K \cap N \simeq (K+N)/N$ is nilpotent and so $K^n \subseteq N$ for some positive integer n. But K^n is finitely generated subring of N, so that K^n and so K is nilpotent. Next, by Lemma 7.4, $L(R)$ coincides with the sum of all locally nilpotent ideals of R. By proved above, each sum of two and so finitely many locally nilpotent ideals of R is locally nilpotent and thus $L(R)$ is the only maximal locally nilpotent ideal of R. By the same reason, the factor ring $R/L(R)$ has no non-trivial locally nilpotent ideals and therefore such right ideals by Lemma 7.4. □

8 The adjoint group of radical rings.

Throughout this section R is a radical ring regarded as an ideal of the ring R_{u} with unity 1, so that $1+R$ is a normal subgroup of the multiplicative group of R_{u} isomorphic to the adjoint group R°.

We start from several lemmas about elements of finite order in the adjoint group of a radical ring.

Lemma 8.1. *Let R be a radical ring.*

(1) If r is an element of finite order n of R° such thar $r \in nR$, then $nr = 0$.

(2) If the additive group R^{+} of R has no elements of a prime order p, then the adjoint group of pR has also no such elements.

Proof. Since $(1+ns)^n = 1 + \sum_{i=1}^{n} \binom{n}{i}(ns)^i$ for every $s \in R$ and every positive integer n, the equality $(1+ns)^n = 1$ implies that $n^2 s(1 + \sum_{i=2}^{n} n^{i-2}s^{i-1} = 0$ and so $n^2 s = 0$. Therefore for $r = ns$ it follows that $nr = n^2 s = 0$ and statement (1)is proved. Statement (2) is an immediate concequence from (1). □

Let π be a set of primes. An element of a group is called π*-element* if its order is a product of powers of primes of π. We denote by π' the complement to π in the set of all primes and write $\{p\}' = p'$.

Corollary 8.2. *The adjoint group of a radical algebra over a field F is torsion-free, if the characteristic of F is zero, and does not contain non-trivial p'-elements, if the characteristic of F is the prime p.*

Recall that a group G is called *π-radicable* for a set of primes π if for each $g \in G$ there exists an element $h \in G$ such that $g = h^p$ for every prime $p \in \pi$. The additive group with this property is usually called *π-divisible*.

Lemma 8.3. *Let R be a radical ring whose additive group is π-divisible for some set of primes π. Then the set Q of all π-elements of R° forms a radicable π-subgroup contained in the annihilator $\mathrm{Ann}(R)$ of R. In particular, Q is a central subgroup of R°.*

Proof. Let g be an element of Q and n the order of g in R°. Since R^+ is π-divisible, we have $nR = R$ and so $ng = 0$ by Lemma 8.1. Therefore $gR + Rg = g(nR) + (nR)g = 0$ and thus $g \in \mathrm{Ann}(R)$. In particular, g is central in R°. Hence Q is a central subgroup of R° contained in $\mathrm{Ann}(R)$. Furthermore, as $pR = R$ for every prime $p \in \pi$, this implies $g = ph$ for some element $h \in R$ and from $g \in \mathrm{Ann}(R)$ it follows that $h \in \mathrm{Ann}(R)$. But then $(1+h)^p = 1 + \sum_{i=1}^{p} \binom{p}{i} h^i = ph = g$, so that $h \in Q$ and thus Q is a radicable π-subgroup of R°. □

Lemma 8.4. *Let R be a radical ring and G a subgroup of R° generated by π-elements for some set of primes π. Then G has a central π'-subgroup C such that G/C does not contain non-trivial π'-elements. Moreover, if the additive group of R has no non-trivial divisible π'-subgroup, then $C = 0$.*

Proof. Clearly we may assume that R is the radical join of G. Let q be a prime not contained in π. Since the factor group $G/G \cap qR$ is isomorphic to a subgroup of the adjoint group $(R/qR)^\circ$ of the factor ring R/qR, this factor group has no non-trivial π-elements by Lemma 8.1. Therefore $G = G \cap qR$ because G is generated by π-elements. Hence $G \subseteq qR$ and so $qR = R$. Thus the additive group R^+ is π'-divisible. By Lemma 8.3, the set Q of all π'-elements of R° forms a central radicable π'-subgroup of R° contained in $\mathrm{Ann}(R)$ and so the intersection $C = G \cap Q$ is a central π'-subgroup of G such that the factor group G/C has no non-trivial π'-elements. Obviously, if R^+ does not contain non-trivial divisible π'-subgroups, then $Q = 0$ and so $C = 0$. □

Note that if in Lemma 8.4 the group G is periodic, then the factor group G/C is a π-group. This implies the following result.

Theorem 8.5. *Every finite subgroup of the adjoint group of a radical ring is nilpotent.*

Proof. Let G a finite subgroup of R°. If G_p denotes the set of all p-elements of G for a prime p, then the subgroup $< G_p >$ generated by G_p must be nilpotent by Lemma 8.4. Therefore G_p is the normal Sylow p-subgroup of G and so G is nilpotent. □

The following theorem gives some information on the structure of the adjoint group of an arbitrary radical ring. We shall say that a group G is *generalized soluble* if the derived subgroup of every non-trivial finitely generated subgroup H of G is properly contained in H.

Theorem 8.6. *The adjoint group R° of every radical ring R is generalized soluble and every periodic subgroup of R° generated by two elements of coprime orders is abelian.*

Proof. Let G be a finitely generated subgroup of R°. It follows from Zorn's Lemma that there exists a right ideal S of R which does not contain G and is maximal with this property. If T is the right ideal of R generated by G, then the right R-module $T/S \cap T$ is simple and so annihilated by R. Therefore $TR \subseteq S$ and, in particular, TS as well as ST are contained in S. Hence $S \cap T$ is a (two-sided) ideal of T. The right ideals S and T are both radical subrings of R. Since $TT \subseteq S \cap T$, the factor ring $T/S \cap T$ has trivial multiplication and so the factor group $T^\circ/S^\circ \simeq (T/S)^\circ$ is abelian. As $G \subseteq T^\circ$, the derived subgroup G' of G is contained in S° and thus $G' \neq G$.

Now let G be the periodic subgroup of R° generated by two its elements g, h of coprime orders. The maximal periodic divisible subgroup Q of R^+ forms an ideal of R and the additive group of the factor ring $\bar{R} = R/Q$ does not contain non-trivial divisible subgroups. Therefore the image $\bar{G}$ of G in $\bar{R}$ is a subgroup of the adjoint group of $\bar{R}$ isomorphic to the factor group $G/G \cap Q$. Furthermore, for each prime p, the set of all p-elements of $\bar{G}$ is a subgroup of $\bar{G}$ by Lemma 8.4. Hence $\bar{G}$ is the direct product of its Sylow p-subgroups and so $\bar{G}$ is a finite group. Since $Q^2 = 0$ by Lemma 8.3, the intersection $G \cap Q$ is an abelian subgroup of finite index in G and so G is finite. Therefore G is nilpotent by Theorem 8.5 and thus commutative because $gh = hg$. □

It is not clear whether every periodic subgroup of the adjoint group of a radical ring is the direct product of its primary components. It follows from the above considerations that this is in fact the case if the annihilator $\mathrm{Ann}(R)$ of R is trivial or R° has no central Prüfer subgroups. The matter in question can be reduced to the following more general problem which is of independent interest. Does every central extension G of a Prüfer p-group by a periodic p'-group split if G is a generalized soluble group? Note that if G is not generalized soluble, then an example of Adjan shows that there exists such a non-split extension.

9 Free algebras and power series rings.

For a set A and a positive integer n, let $\underbrace{A \times \ldots \times A}_{n}$ be the n-th cartesian power of A which is an empty set under $n = 0$. An n-ary operation ω on A is a mapping $\omega : \underbrace{A \times \ldots \times A}_{n} \to A$. A non-empty set A is said to be a

universal algebra with a set Ω of operations on A if for each $\omega \in \Omega$ there exists a non-negative integer $n = n(\omega)$ such that ω is an n-ary operation on A. A non-empty subset B of A is called a *subalgebra* of A if every operation $\omega \in \Omega$ with $n = n(\omega)$ maps $\underbrace{B \times \ldots \times B}_{n}$ into B. The subalgebra B is generated by its subset C if B is the least subalgebra of A containing C. Typical examples of universal algebras are semigroups, groups, rings, modules over a ring, algebras over a commutative ring with unity, radical algebras and so on.

If A_1 and A_2 are universal algebras with the same set Ω of operations, then a *universal algebra homomorphism* from A_1 into A_2 is a mapping $\varphi : A_1 \to A_2$ such that for every operation $\omega \in \Omega$ with $n = n(\omega)$ and every n-tuple $(a_1, \ldots, a_n)$ of elements of A_1 the equality $(\omega(a_1, \ldots, a_n))^{\varphi} = \omega(a_1^{\varphi}, \ldots, a_n^{\varphi})$ holds.

Let X be a non-empty set, $\mathcal{K}$ a class of universal algebras with a set of operations and F_X a universal algebra contained in $\mathcal{K}$. If there exists a mapping $\varepsilon : X \to F_X$ such that for each universal algebra $A \in \mathcal{K}$ and every mapping $\alpha : X \to A$ there exists a unique universal algebra homomorphism $\beta : F_X \to A$ with $x^{\alpha} = (x^{\varepsilon})^{\beta}$ for every $x \in X$, then F_X is called a *$\mathcal{K}$-free universal algebra on X* with mapping $\varepsilon : X \to F_X$. In other words, F_X is $\mathcal{K}$-free if the diagram

$$\begin{array}{ccc} X & \xrightarrow{\alpha} & A \\ & {\scriptstyle\varepsilon}\searrow \quad \nearrow{\scriptstyle\beta} & \\ & F_X & \end{array}$$

is commutative. It is easy to see that in this case every universal algebra $A \in \mathcal{K}$ generated by a set whose cardinal number does not exceed the cardinal number of X is a homomorphic image of F_X.

Theorem 9.1. *Let $\mathcal{K}$ be a class of universal algebras with the same set of operations closed under the subalgebras and containing universal algebras with at least two elements. If F_X is a $\mathcal{K}$-free universal algebra on X with mapping $\varepsilon : X \to F_X$, then ε is injective and its image X^{ε} generates F_X. Moreover, if F'_X is another $\mathcal{K}$-free universal algebra on the same set X with mapping $\varepsilon' : X \to F'_X$, then there exists a unique universal algebra isomorphism $\gamma : F_X \to F'_X$ such that $(x^{\varepsilon})^{\gamma} = x^{\varepsilon'}$ for every $x \in X$.*

Proof. If x, y are two distinct elements of X and A is a universal algebra containing at least two elements, then there exists a mapping $\alpha : X \to A$ with $x^{\alpha} \neq y^{\alpha}$. Since $(x^{\varepsilon})^{\beta} = x^{\alpha}$ and $(y^{\varepsilon})^{\beta} = y^{\alpha}$, this implies $x^{\varepsilon} \neq y^{\varepsilon}$. Therefore ε is injective.

Now let E be the subalgebra of F_X generated by X^{ε} and let $\alpha : X \to E$ be given by $x^{\alpha} = x^{\varepsilon}$ for every $x \in X$. Then there exists a unique universal algebra homomorphism $\gamma : F_X \to E$ such that $x^{\varepsilon} = (x^{\varepsilon})^{\gamma}$ for every $x \in X$. Let $\eta : E \to F_X$ be the identity embedding and $\iota : F_X \to F_X$ the identity mapping. Since $\gamma\eta$ and ι are homomorphisms from F_X into F_X

such that $((x^\varepsilon)^\gamma)^\eta = x^\varepsilon = (x^\varepsilon)^\iota$, it follows from the uniqueness of an extended homomorphism that $\gamma\eta = \iota$. Hence $\eta = \iota$ and so $E = F_X$.

Finally, if F'_X is another free universal algebra on X with mapping $\epsilon' : X \to F'_X$, then there exists a unique homomorphism $\gamma : F_X \to F'_X$ such that $x^{\varepsilon'} = (x^\varepsilon)^\gamma$ for every $x \in X$. Similarly, there exists a unique homomorphism $\delta : F'_X \to F_X$ such that $x^\varepsilon = (x^{\varepsilon'})^\delta$. Then the composition $\gamma\delta$ is a homomorphism from F_X into F_X such that $x^\varepsilon = ((x^\varepsilon)^\gamma)^\delta$. The uniqueness of an extended homomorphism implies that $\gamma\delta$ is the identity mapping on F_X. Therefore γ is injective and δ is surjective. By the same reason, $\delta\gamma$ is the identity mapping of F'_X, so that γ is also surjective and so an isomorphism from F_X onto F'_X. □

Let W be the set of all words in X, regarded as an alphabet. If u, v are two words, their product uv is defined as a word obtained by ascribing v to u in the stated order. It is clear that W forms a semigroup under this multiplication of the words. For each word $w \in W$, its length $|w|$ is defined as the number of all letters which are contained in w, including their recurrence. Clearly $|uv| = |u| + |v|$ for every words $u, v \in W$.

Proposition 9.2. *The semigroup W is the free semigroup on X.*

Proof. We have to show that there exists a mapping $\varepsilon : X \to W$ such that for each semigroup S and every mapping $\alpha : X \to S$ there exists a unique semigroup homomorphism $\beta : W \to S$ with $x^\alpha = (x^\varepsilon)^\beta$ for every $x \in X$.

Indeed, for each $x \in X$, let x^ε be the word consisting of the single letter x. By induction on the length of words, define a mapping $\beta : W \to S$ by taking $u^\beta = u^\alpha$ if $|u| = 1$, i.e. $u \in X$, and $(uv)^\beta = u^\beta v^\beta$ if $|uv| \geq 2$. Then β is a semigroup homomorphism from W into S with desired properties. □

If we add to W an empty word denoted by 1, then the union $W_\mathrm{u} = W \cup \{1\}$ is a semigroup with unity 1 generated by the set $X \cup \{1\}$. This semigroup is in fact the *free semigroup with unity* on X, so that every semigroup with unity can be obtained as a homomorphic image of W_u for a suitable set X.

Let R be a ring, viewed as an ideal of a ring R_u with unity 1, and M a right R-module. Then M is also a unital right R_u-module. A subset $\{m_i \mid i \in I\}$ of M is said to be a *basis* of M over R if every element $m \in M$ can uniquely be written in the form $m = \sum_{i \in I} m_i r_i$ with finitely many non-zero elements $r_i \in R_\mathrm{u}$. This means that $\sum_{i \in I} m_i r_i = 0$ only if $r_i = 0$ for all $i \in I$. In particular, $m_i r = 0$ for some $r \in R_\mathrm{u}$ only if $r = 0$, so that the right R-modules $m_i R$ and R as well as $m_i R_\mathrm{u}$ and R_u are isomorphic. It is easy to see that $\{m_i \mid i \in I\}$ is a basis of M if and only if $M = \bigoplus_{i \in I} m_i R_\mathrm{u}$ is a direct sum of its R-submodules $m_i R_\mathrm{u}$.

Proposition 9.3. *A (right) R-module M is the free (right) R-module on a set X if and only if there exists a basis $\{m_x \mid x \in X\}$ of M over R.*

Proof. If X is a set and M is a free R-module on X, then there exists a mapping $\varepsilon : X \to M$ such that for each R-module N and every mapping

$\alpha : X \to N$ there exists a unique module homomorphism $\beta : M \to N$ with $x^\alpha = (x^\varepsilon)^\beta$ for every $x \in X$. Take N to be a direct sum $\bigoplus_{x\in X} N_x$ of modules N_x each of which is isomorphic to $R_\mathfrak{u}$ as a right R-module and let $x^\alpha = e_x$ be a unity of N_x. If $x^\varepsilon = m_x \in M$ and $\sum_{x\in X} m_x r_x = 0$ for some elements $r_x \in R_\mathfrak{u}$ among which only finitely many are non-zero, then $e_x = m_x^\beta$ and from $0 = (\sum_{x\in X} m_x r_x)^\beta = \sum_{x\in X} e_x r_x$ it follows that $r_x = 0$ for every $x \in X$. Since M is generated by the set $\{m_x \mid x \in X\}$ by Theorem 9.1, this implies that this set is a basis of M.

Conversely, suppose that there exists a basis $\{m_x \mid x \in X\}$ of M and the mapping $\varepsilon : X \to M$ is defined by $x^\varepsilon = m_x$ for every $x \in X$. Then for each R-module N and every mapping $\alpha : X \to N$ the mapping $\beta : M \to N$ given by $(\sum_{x\in X} m_x r_x)^\beta = \sum_{x\in X} x^\alpha r_x$ with finitely many non-zero elements $r_x \in R_\mathfrak{u}$ is a unique module homomorphism satisfying the condition $m_x^\beta = x^\alpha$. Therefore M is a free R-module on X. □

Let K be a commutative ring with 1 and $K\langle\langle X\rangle\rangle$ the set of all formal sums of the form $r = \sum_w k_w w$ with $k_w \in K$ and $w \in W_\mathfrak{u}$. If $k \in K$ and $s = \sum_w l_w w \in K\langle\langle X\rangle\rangle$, we define $kr = \sum_w (kk_w)w$, $r + s = \sum_w (k_w + l_w)w$ and $rs = \sum_w (\sum_{w=uv} k_u l_v)w$. It is easy to see that under these operations $K\langle\langle X\rangle\rangle$ is an algebra over K whose unity is an empty word with coefficient 1. This algebra is called the *algebra of all formal power series over K in non-commuting indeterminates* of X. The subalgebra $K\langle X\rangle$ of $K\langle\langle X\rangle\rangle$ generated by the set $X \cup \{1\}$ is the algebra of all polynomials over K in non-commuting indeterminates of X.

The set $A[[X]]$ of all formal power series over K without constant terms is an ideal of $K\langle\langle X\rangle\rangle$ such that the factor algebra $K\langle\langle X\rangle\rangle/A[[X]]$ is isomorphic to K. The intersection $A[[X]] \cap K\langle X\rangle$ consists of all polynomials over K without constant terms and is denoted by $A[X]$. Clearly it is an ideal of $K\langle X\rangle$ and each element of $A[x]$ can be written as a finite sum of the form $r = \sum_w k_w w$ with $k_w \in K$ and $w \in W$.

Proposition 9.4. *The algebra $A[X]$ is a free K-algebra on X.*

Proof. Since W is a free semigroup on X, for every K-algebra R, regarded as a multiplicative semigroup, and every mapping $\alpha : X \to R$ there exists a unique semigroup homomorphism $\beta : W \to R$ such that $x^\alpha = x^\beta$ for every $x \in X$. Clearly $A[X]$ is a free K-module on W as a basis. Therefore there exists a unique module homomorphism $\gamma : A[X] \to R$ such that $w^\beta = w^\gamma$ for every $w \in W$. Hence γ is an algebra homomorphism from $A[X]$ into R which is unique with the property $x^\alpha = x^\gamma$. Thus the K-algebra $A[X]$ is free. □

For each non-negative integer n and every $r = \sum_w k_w w \in K\langle\langle X\rangle\rangle$, we put $r_n = \sum_{|w|=n} k_w w$, where $|w| = 0$ if and only if w is an empty word. Then $r = \sum_{n=0}^\infty r_n$ and $(rs)_n = \sum_{m=0}^n r_m s_{n-m}$. Moreover, if $r \in A[[X]]^n$ for some $n \geq 1$, then it easy to see that $r = \sum_{i=n}^\infty r_i$, so that the factor algebra $K\langle\langle X\rangle\rangle/A[[X]]^n$ is finite-dimensional over K if and only if the set X

is finite. Put $G_n(X) = 1 + A[[X]]^n$ for each $n \geq 1$ and write $G(X) = G_1(X)$. Since $A[[X]]^{n+1} \subset A[[X]]^n$ for each $n \geq 1$ and $\bigcap_{n=1}^{\infty} A[[X]]^n = 0$, we have $G(X) = G_1(X) \supset \ldots \supset G_n(X) \supset \ldots$ and $\bigcap_{n=1}^{\infty} G_n(X) = 1$.

Lemma 9.5. *The set $G(X)$ is a subgroup of the multiplicative group of $K\langle\langle X\rangle\rangle$ and, for each $n \geq 1$, $G_n(X)$ is a normal subgroup of $G(X)$.*

Proof. Clearly $G(X)$ is closed under the multiplication in $K\langle\langle X\rangle\rangle$. Furthermore, for every $r \in A[[X]]$ and each integer $m \geq 1$, the series $s = \sum_{n=1}^{\infty}(-1)^n r^n$ has only finitely many words of length m, so that s is also an element of $A[[X]]$. Since $(1+r)(1+s) = (1+s)(1+r) = 1$, every element of $G(X)$ is invertible and so $G(X)$ is a group. Since $A[[X]]^n$ is an ideal in $K\langle\langle X\rangle\rangle$ and $G_n(X) = G(X) \cap (1 + A[[X]]^n)$, the lemma is proved. □

In particular, Lemma 9.5 and Proposition 2.2 imply the following.

Corollary 9.6. *The ideal $A[[X]]$ is contained in the Jacobson radical of the algebra $K\langle\langle X\rangle\rangle$. In particular, the algebra $A[[X]]$ is radical.*

Let $F(X)$ be the subgroup of $G(X)$ generated by all elements of the form $1 + x$ with $x \in X$.

Proposition 9.7. *If K is an integral domain, then the group $F(X)$ is a free group on the set $\{1 + x \mid x \in X\}$.*

Proof. Let $x_1, \ldots, x_n$ be elements of X with $x_i \neq x_{i+1}$ for any $1 \leq i \leq n-1$. It suffices to prove that for each positive integer n and any non-zero integers $m_1, \ldots, m_n$ the equality

$$(1 + x_1)^{m_1} \ldots (1 + x_n)^{m_n} = 1$$

does not hold. By Binomial theorem,

$$(1 + x_i)^{m_i} = 1 + m_i x_i + \binom{m_i}{2} x_i^2 + \ldots$$

and obviously at least one of coefficients, say $\binom{m_i}{l_i}$, must be non-zero. Then the coefficient under the word ${x_1}^{l_1} \ldots {x_n}^{l_n}$ is $\binom{m_1}{l_1} \ldots \binom{m_n}{l_n} \neq 0$. □

Put $F_n(X) = F(X) \cap G_n(X)$ for each $n \geq 1$. Then $F_n(X)$ is a normal subgroup of $F(X)$ and $\bigcap_{n=1}^{\infty} F_n(X) = 1$.

Lemma 9.8. *The series $F(X) = F_1(X) \supset \ldots \supset F_n(X) \supset \ldots$ is a lower central series of $F(X)$.*

Proof. Since $(1 + r, 1 + s) = 1 + (1 + r)^{-1}(1 + s)^{-1}[r, s]$ for any $r, s \in A[[X]]$ and since $[r, s] = rs - sr \in A[[X]]^{m+n}$ if $r \in A[[X]]^m$ and $s \in A[[X]]^n$, we have $(G_m(X), G_n(X)) \subseteq G_{m+n}(X)$ for any positive integers m, n. Therefore $(F_m(X), F_n(X)) \subseteq F_{m+n}(X)$, as desired. □

Since every radical K-algebra can be regarded as an algebra over K with an additional unary operation $r \to r'$ satisfying the laws $r + r' = -rr' =$

$-r'r$, the class of radical K-algebras forms a *variety* of universal algebras. This implies that for each set X there exists a free radical K-algebra on X. Moreover, it was proved by Cohn that the radical join $A\langle X\rangle_{rad}$ of X in the radical algebra $A[[X]]$ is such a K-algebra on X. Here we restrict ourselves to the simplest case of a free radical algebra on a single generator.

It is easy to see that, for $X = \{x\}$, the radical join $A\langle x\rangle_{rad}$ of x in $A[[x]]$ consists of all power series in x each of which can be written as a fraction $xf(x)/(1 + xg(x))$ with some polynomials $f(x), g(x) \in K[x]$.

Proposition 9.9. *The algebra $A\langle x\rangle_{rad}$ is a free radical K-algebra on x.*

Proof. If R is a radical K-algebra and $x^\alpha = r \in R$, then it is immediately verified that the mapping $\beta : xf(x)/(1 + xg(x)) \mapsto rf(r)/(1 + rg(r))$ with $f(x), g(x) \in K[x]$ determines a unique ring homomorphism from $A\langle x\rangle_{rad}$ into R such that $x^\alpha = x^\beta$. □

Free radical rings on more than one generator are much more complicated, and no simple normal form for its elements is known.

10 Tensor products.

Let R be a ring with 1 and suppose that A is a right R-module and B is a left R-module. Consider a free abelian group $G = G(A, B)$ whose set of free generators is the cartesian product $A \times B$ of the sets A and B. In other words, every element of G is uniquely expressed as a finite sum of the form $\sum_{(a,b)\in A\times B} z_{a,b}(a, b)$ with $z_{a,b} \in \mathbb{Z}$. Let G_0 be the subgroup of G generated by all elements of the form

T1: $(ar, b) - (a, rb)$,

T2: $(a_1 + a_2, b) - (a_1, b) - (a_2, b)$ and

T3: $(a, b_1 + b_2) - (a, b_1) - (a, b_2)$

with $a, a_1, a_2 \in A$, $b, b_1, b_2 \in B$ and $r \in R$. The abelian factor group G/G_0 is called a *tensor product* of A and B over R and denoted by $A\otimes_R B$. For every $(a, b) \in A \times B$, its image under the natural homomorphism $\tau : G \to G/G_0$ is denoted by $a \otimes b$, so that $a \otimes b = (a, b) + G_0$. It follows from T1 - T3 that

$$ar \otimes b = a \otimes rb,$$

$$(a_1 + a_2) \otimes b = a_1 \otimes b + a_2 \otimes b \qquad \text{and}$$

$$a \otimes (b_1 + b_2) = a \otimes b_1 + a \otimes b_2$$

for all appropriate elements in A, B and R. Thus every element u of $A\otimes_R B$ can be written in the form

$$u = \sum_{j=1}^{n} a_j \otimes b_j$$

for some elements $a_1, \ldots, a_n \in A$ and $b_1, \ldots, b_n \in B$.

Let R and S be two arbitrary rings with 1. An abelian group M is called an (R, S)-*bimodule* if M simultaneously is a left R-module and a right S-module with $(rm)s = r(ms)$ for all $r \in R$, $m \in M$ and $s \in S$. Now, if Q is a third ring with 1 such that A is a (Q, R)-bimodule and B is an (R, S)-bimodule, then the tensor product $A \otimes_R B$ can also be viewed as a (Q, S)-bimodule by the rule $q(a \otimes b)s = (qa) \otimes (bs)$ for all $q \in Q$ and $s \in S$. Clearly, for a commutative ring K, every right K-module M can be turned into a left K-module and conversely by setting $km = mk$ for any $k \in K$ and $m \in M$. In this case $A \otimes_K B$ can be regarded as an (K, K)-bimodule.

In what follows let K be a commutative ring with 1 and let all K-modules be considered as (K, K)-bimodules.

Proposition 10.1. *If C is a K-module and $\Phi : A \times B \to C$ is a K-bilinear mapping, then there exists a uniquely determined K-homomorphism $\varphi : A \otimes_K B \to C$ such that $(a \otimes b)^\varphi = \Phi(a, b)$ for every $a \in A$ and $b \in B$.*

Proof. There exists a uniquely determined $\mathbb{Z}$-homomorphism $\phi : G \to C$ from the free abelian group $G = G(A, B)$ into C such that $(a, b)^\phi = \Phi(a, b)$. Since Φ is K-bilinear, the subgroup G_0 is contained in $\operatorname{Ker}\phi$. Thus, there exists a unique $\mathbb{Z}$-homomorphism $\varphi : G/G_0 = A \otimes_K B \to C$ such that $\tau\varphi = \phi$. In fact, φ is an K-homomorphism because $(r(a \otimes b))^\varphi = (ra \otimes b)^\varphi = \Phi(ra, b) = r\Phi(a, b) = r(a \otimes b)^\varphi$ for all $a \in A$ and $b \in B$. □

It turns out that the tensor product $A \otimes_K B$ is even characterized by the above property. In other words, a K-module M is isomorphic to $A \otimes_K B$ if there exists a K-bilinear mapping $\Psi : A \times B \to M$ such that for each K-module C and every K-bilinear mapping $\Phi : A \times B \to C$ there exists a unique K-homomorphism $\varphi : M \to C$ satisfying the condition $(\Psi(a, b))^\varphi = \Phi(a, b)$ for all $a \in A$ and $b \in B$ which means that the diagram

$$\begin{array}{ccc} A \times B & \xrightarrow{\Phi} & C \\ & \Psi\searrow \quad \nearrow \beta & \\ & M & \end{array}$$

is commutative.

In general, a tensor product of two modules is a very intricate operation and, for instance, tensor products of some non-trivial $\mathbb{Z}$-modules can be trivial.

Example 10.2. *If π is a set of primes, A is a π-divisible abelian group and B is a periodic abelian π-group, then $A \otimes_{\mathbb{Z}} B = 0$.*

Proof. Indeed, if $b \in B$ is an element of order n and $a \in A$, then $a = na_1$ for some element $a_1 \in A$, so that $a \otimes b = (na_1) \otimes b = n(a_1 \otimes b) = a_1 \otimes (nb) = 0$. □

However, the situation is much better for tensor products of two free K-modules.

Proposition 10.3. *If A and B are free K-modules with bases $\{a_i \mid i \in \mathcal{I}\}$ and $\{b_j \mid j \in \mathcal{J}\}$ respectively, then $A \otimes_K B$ is also a free K-module with the basis $\{a_i \otimes b_j \mid i \in \mathcal{I}, j \in \mathcal{J}\}$.*

Furthermore, if K is a subring of a commutative ring R with the same unity 1, then $A \otimes_K R$ is a free R-module with the basis $\{a_i \mid i \in \mathcal{I}\}$.

Proof. We shall prove only the first part of the proposition. The second can be proved by similar arguments.

If $a = \sum_i a_i r_i$ and $b = \sum_j s_j b_j$, then $a \otimes b = (\sum_i a_i r_i) \otimes (\sum_j s_j b_j) = \sum_{i,j} r_i s_j (a_i \otimes b_j)$, so that $A \otimes_K B$ is generated by all $a_i \otimes b_j$. Assume that $\sum_{k,l} q_{kl}(a_k \otimes b_l) = 0$ for some elements $q_{kl} \in K$ with $k \in \mathcal{I}$ and $l \in \mathcal{J}$. For every pair $(i,j) \in \mathcal{I} \times \mathcal{J}$ we define a K-bilinear mapping $\Phi_{ij} : A \times B \to K$ by the rule

$$\Phi_{ij}(a_k, b_l) = \delta_{ik}\delta_{jl},$$

where $\delta_{kk} = 1$ and $\delta_{kl} = 0$ for $k \neq l$. This mapping can be extended to a K-homomorphism $\varphi_{ij} : A \otimes_K B \to K$, so that

$$0 = 0^{\varphi_{ij}} = \sum_{i,j} q_{kl}(a_k \otimes b_l)^{\varphi_{ij}} = q_{ij},$$

as desired. □

The next statement concerns properties of commutativity and associativity of tensor products.

Proposition 10.4. *Let A, B and C be K-modules. Then the mappings $a \otimes b \mapsto b \otimes a$ and $a \otimes (b \otimes c) \mapsto (a \otimes b) \otimes c$ with $a \in A$, $b \in B$ and $c \in C$ determine natural isomorphisms from $A \otimes_K B$ onto $B \otimes_K A$ and from $A \otimes_K (B \otimes_K C)$ onto $(A \otimes_K B) \otimes_K C$.*

Proof. We shall restrict ourselves to the case of associativity of the tensor products, leaving the proof of commutativity to the reader.

Show first that the mapping $(a \otimes b, c) \mapsto a \otimes (b \otimes c)$ determines a K-bilinear mapping

$$\Phi : (A \otimes_K B) \times C \to A \otimes_K (B \otimes_K C).$$

Indeed, it is easy to see that the mapping $(a, b, c) \mapsto a \otimes (b \otimes c)$ is K-linear in every variable. Thus, for the free abelian group $G(A, B)$, there exists a uniquely determined $\mathbb{Z}$-bilinear mapping

$$\Upsilon : G(A, B) \times C \to A \otimes_K (B \otimes_K C)$$

such that $\Upsilon((a, b), c) = a \otimes (b \otimes c)$. Moreover, it can easily be verified that $\Upsilon(G_0, C) = 0$, so that there exists a unique $\mathbb{Z}$-bilinear mapping

$$\Phi : (A \otimes_K B) \times C \to A \otimes_K (B \otimes_K C)$$

with $\Phi(a \otimes b, c) = a \otimes (b \otimes c)$. In fact, Φ is a K-bilinear mapping because $\Phi(r(a \otimes b), c) = \Phi(ra \otimes b, c) = ra \otimes (b \otimes c) = r(a \otimes (b \otimes c)) = r\Phi(a \otimes b, c)$ for all $a \in A$, $b \in B$ and $c \in C$.

By Proposition 10.1, there exists a uniquely determined K-homomorphism

$$\varphi : (A \otimes_K B) \otimes_K C \to A \otimes_K (B \otimes_K C)$$

such that $((a \otimes b) \otimes c)^{\varphi} = a \otimes (b \otimes c)$. Similarly, the mapping $(a, b, c) \mapsto (a \otimes b) \otimes c$ determines a K-bilinear mapping

$$\Psi : A \times_K (B \otimes_K C) \to (A \otimes_K B) \otimes C,$$

so that there exists a uniquely determined K-homomorphism

$$\psi : A \otimes_K (B \otimes_K C) \to (A \otimes_K B) \otimes_K C$$

such that $(a \otimes (b \otimes c))^{\psi} = (a \otimes b) \otimes c$. Hence both compositions $\varphi\psi$ and $\psi\varphi$ determine the identity mappings on $(A \otimes_K B) \otimes_K C$ and on $A \otimes_K (B \otimes_K C)$, respectively, and so φ and ψ are K-isomorphisms. □

Thus, for each integer $n \geq 1$ and K-modules $A_1, \ldots, A_n$, the tensor product $A_1 \otimes_K \ldots \otimes_K A_n$ can uniquely be determined by

$$A_1 \otimes_K \ldots \otimes_K A_n = (A_1 \otimes_K \ldots \otimes_K A_{n-1}) \otimes_K A_n.$$

In particular, for any K-module A we put $T_0(A) = K$ and

$$T_n(A) = \underbrace{A \otimes_K \ldots \otimes_K A}_{n}$$

and say that $T_n(A)$ is the n-th tensor degree of the K-module M. Consider next the direct sum

$$T(A) = T_0(A) \oplus T_1(A) \oplus \ldots \oplus T_n(A) \oplus \ldots$$

of all the K-modules $T_n(A), n \geq 0$. For any elements $u \in T_m(A)$ and $v \in T_n(A)$ with $m, n \geq 0$, we put $uv = u \otimes v$. Clearly this determines an associative multiplication in $T(A)$, so that $T(A)$ turns into an associative K-algebra with 1 which is called the *tensor algebra* of A over K. If $\{a_i \mid i \in \mathcal{I}\}$ is a subset of A which generates A as a K-module, then it is easy to see that the set $\{1, a_{i_1} \otimes \ldots \otimes a_{i_m} \mid i_1, \ldots, i_m \in \mathcal{I}, m \geq 1\}$ generates $T(A)$ as a K-module.

Example 10.5. *Let A be a free K-module with a basis $\{a_i \mid i \in I\}$. Then the set $\{1, a_{i_1} \otimes \ldots \otimes a_{i_m} \mid i_1, \ldots, i_m \in \mathcal{I}, m \geq 1\}$ is a basis of the tensor algebra $T(A)$, so that $T(A)$ is isomorphic to the algebra $K\langle x_i \mid i \in \mathcal{I}\rangle$ of all polynomials over K in non-commuting indeterminates $\{x_i \mid i \in \mathcal{I}\}$.*

In particular, if A is a free K-module with a single generator, then the tensor algebra $T(A)$ is isomorphic to the polynomial algebra $K[x]$ in one indeterminate x.

Proposition 10.6. *Let A be a K-module, R an associative K-algebra with 1 and $\varphi : A \to R$ a K-module homomorphism. Then there exists a uniquely determined algebra homomorphism $\hat{\varphi} : T(A) \to R$ from the tensor algebra $T(A)$ into R such that $a^{\hat{\varphi}} = a^{\varphi}$ for all $a \in A$ and $1^{\hat{\varphi}} = 1$.*

Proof. Show first that for each $n \geq 1$ there exists a uniquely determined K-module homomorphism $\phi_n : T_n(A) \to R$ such that $(a_1 \otimes \ldots \otimes a_n)^{\phi_n} = a_1^{\varphi} \ldots a_n^{\varphi}$ for any $a_1, \ldots, a_n \in A$.

Since $T_1(A) = A$, we have $\phi_1 = \varphi$. By induction on n, we may assume that ϕ_{n-1} is constructed. For any $u \in T_{n-1}(A)$ and $a \in A$, we put $\Phi(u, a) = u^{\phi_{n-1}} a^{\varphi}$. Then $\Phi : (T_{n-1} \times A) \to R$ is a K-bilinear mapping. By Proposition 10.1, this mapping can uniquely be extended to a K-homomorphism $\phi_n : (T_{n-1} \otimes_K A = T_n(A)) \to R$, as required.

Next, to define a K-module homomorphism $\hat{\varphi} : T(A) \to R$ it suffices to establish its value on $T_n(A)$ for each $n \geq 0$ and to extend by linearity. Thus, for any $r \in K = T_0(A)$ and $u \in T_n(A)$ with $n \geq 1$, we put $r^{\hat{\varphi}} = r$ and $u^{\hat{\varphi}} = u^{\phi_n}$. Now, if $u_m \in T_m(A)$ and $u_n \in T_n(A)$ with $m, n \geq 1$, it follows from the construction of ϕ_{m+n} that $(u_m \otimes u_n)^{\phi_{m+n}} = u_m^{\phi_m} u_n^{\phi_n}$. Therefore, for any $v, w \in T(A)$, we have $(vw)^{\hat{\varphi}} = v^{\hat{\varphi}} w^{\hat{\varphi}}$ and hence $\hat{\varphi}$ is an algebra homomorphism, as desired. □

11 Lie algebras and their universal enveloping.

Let K be a commutative ring with 1. A *Lie algebra* over K is a K-module L with a K-bilinear mapping $L \times L \to L$ which is called the Lie-multiplication on L and usually denoted by brackets $[\,,\,]$ such that

1) $[a, a] = 0$ for every $a \in L$, and

2) $[[a, b], c] + [[c, a], b] + [[b, c], a] = 0$ for all $a, b, c \in L$ (the Jacobi identity).

In particular, from 1) it follows that $[a, b] = -[b, a]$ for every $a, b \in L$.

Example. Let R be an associative algebra over K and $[r, s] = rs - sr$ for all $r, s \in R$. Then the mapping $R \times R \to R$ given by $(r, s) \mapsto [r, s]$ is K-bilinear and conditions 1) and 2) hold, so that R is a Lie-algebra over K under the Lie-multiplication $[r, s]$. This algebra is called the *associated Lie algebra* of R. For the K-algebra $\mathrm{End}_K M$ of all K-endomorphisms of a K-module M, the associated Lie algebra of $\mathrm{End}_K M$ is usually denoted by $\mathfrak{l}_K(M)$ or if K is a field by $\mathfrak{gl}_K(M)$.

If V and W are two non-empty subsets of a Lie algebra L over K, then $[V, W]$ denotes the K-submodule of L generated by all Lie-commutators $[v, w]$ with $v \in V$ and $w \in W$. A K-submodule V of L is called a *subalgebra* of L if $[V, V] \subseteq V$ and an *ideal* of L if $[V, L] \subseteq V$. In particular, the *centre* $Z(L) = \{a \in L \mid [a, L] = 0\}$ and the *commutator subalgebra* $[L, L]$ of L are ideals in L. As usual, for each ideal V of L, the factor module L/V is turned into a Lie algebra over K if the Lie-multiplication on L/V is defined

by $[a+V, b+V] = [a,b]+V$ for all $a, b \in L$. The Lie algebra L/V is called the *factor algebra* of L modulo V.

For two Lie algebras L_1 and L_2 over K, a K-module homomorphism $\alpha : L_1 \to L_2$ is called a *Lie-homomorphism* from L_1 into L_2 if $[a,b]^\alpha = [a^\alpha, b^\alpha]$ for all $a, b \in L_1$. It is clear that the image $\operatorname{Im}\alpha$ is a subalgebra of L_2, the kernel $\operatorname{Ker}\alpha$ is an ideal of L_1 and the Lie algebras $L_1/\operatorname{Ker}\alpha$ and $\operatorname{Im}\alpha$ are isomorphic.

If L is a Lie algebra over K and $a \in L$, then the mapping $\operatorname{ad} a : L \to L$ with $(\operatorname{ad} a)(b) = [a,b]$ for every $b \in L$ is a K-endomorphism of L satisfying the condition

$$(\operatorname{ad} a)([b,c]) = [(\operatorname{ad} a)(b), c] + [b, (\operatorname{ad} a)(c)]$$

for all $b, c \in L$. This mapping is called a *derivation* of L. The set of all derivations of L forms a K-submodule $\operatorname{Der} L$ of $\operatorname{End}_K L$. Since

$$(\operatorname{ad} a)(\operatorname{ad} b) - (\operatorname{ad} b)(\operatorname{ad} a) = \operatorname{ad}[a,b]$$

for every $a, b \in L$, the K-module $\operatorname{Der} L$ is a Lie algebra under the Lie-multiplication $[\operatorname{ad} a, \operatorname{ad} b] = (\operatorname{ad} a)(\operatorname{ad} b) - (\operatorname{ad} b)(\operatorname{ad} a)$ which is called the *Lie algebra of all derivations* of L. Furthermore, if R is an associative algebra over K, then $\operatorname{Der} R$ denotes the Lie algebra of all derivations of the associated Lie algebra of R. For each $a \in L$, the mapping $a \mapsto \operatorname{ad} a$ determines a Lie-homomorphism $\operatorname{ad} : L \to \mathfrak{l}_K(L)$ whose image is $\operatorname{Der} L$ and whose kernel coincides with the centre $Z(L)$ of L, so that $\operatorname{Der} L$ is isomorphic to the factor algebra $L/Z(L)$. The Lie-homomorphism $\operatorname{ad} : L \to \mathfrak{l}_K(L)$ is called an *adjoint representation* of L in $\operatorname{End}_K L$. In general, if M is a K-module, then every Lie-homomorphism from L into $\mathfrak{l}_K(M)$ is called a *representation* of L in $\operatorname{End}_K M$.

If R is an associative K-algebra and L is a Lie algebra over K, then a Lie-homomorphism from L into R is one from L into the associated Lie algebra of R. An associative K-algebra $U = U(L)$ with unity 1 is called a *universal enveloping algebra* of L over K if the following statements hold:

i) there exists a Lie-homomorphism $\epsilon : L \to U$ and

ii) for every associative K-algebra R with unity 1, and any Lie-homomorphism $\alpha : L \to R$ there exists a unique homomorphism $\phi : U \to A$ of associative K-algebras with unity such that $a^\alpha = (a^\epsilon)^\phi$ for every $a \in L$. In other words, the diagram

$$\begin{array}{ccc} L & \xrightarrow{\alpha} & R \\ & {\scriptstyle\epsilon}\searrow \quad \nearrow{\scriptstyle\phi} & \\ & U & \end{array}$$

is commutative.

To prove that for any Lie algebra L over K there exists a universal enveloping algebra $U(L)$ of L, we consider the tensor algebra $T(L)$ of L over

K and take the ideal I of $T(L)$ generated by all elements of the form

$$a \otimes b - b \otimes a - [a, b]$$

with $a, b \in L$. In what follows let $U(L)$ denote the factor algebra $T(L)/I$ and let $\epsilon : L \to U(L)$ be the mapping given by $a^\epsilon = a + I$ for every $a \in L$.

Theorem 11.1. *The algebra $U(L)$ is a universal enveloping algebra of L over K and the mapping $\epsilon : L \to U(L)$ is a Lie-homomorphism from L into $U(L)$.*

Proof. Show first that the mapping $a^\epsilon = a+I$ determines a Lie-homomorphism from L into $U(L)$.

Indeed, it is obvious that $\epsilon : L \to U(L)$ is a K-homomorphism. Furthermore, for any $a, b \in L$, we have

$$[a, b]^\epsilon = [a, b] + I = a \otimes b - b \otimes a + I = \\ (a + I)(b + I) - (b + I)(a + I) = a^\epsilon b^\epsilon - b^\epsilon a^\epsilon = [a^\epsilon, b^\epsilon].$$

Next, if V is an associative K-algebra with unity 1 and $\phi : L \to V$ is a Lie-homomorphism, in particular, a K-module homomorphism, then there exists a uniquely determined homomorphism $\hat{\phi} : T(L) \to V$ of associative algebras with unity such that $a^{\hat{\phi}} = a^\phi$ for every $a \in L$ (see Proposition 10.6). Since

$$(a \otimes b - b \otimes a - [a, b])^{\hat{\phi}} = \\ a^\phi b^\phi - b^\phi a^\phi - [a, b]^\phi = [a^\phi, b^\phi] - [a, b]^\phi = 0$$

for all $a, b \in L$, we have $I \subseteq \operatorname{Ker} \hat{\phi}$ and thus there exists a unique algebra homomorphism $\alpha : U(L) \to V$, namely $(u + I)^\alpha = u^{\hat{\phi}}$ for all $u \in T(L)$, such that $\alpha = \epsilon\phi$. □

The above Lie-homomorphism $\epsilon : L \to U(L)$ will be called the *canonical Lie-homomorphism* from the Lie algebra L into its universal enveloping algebra $U(L)$.

Lemma 11.2. *If L, regarded as a K-module, is generated by its subset $\{a_i \mid i \in \mathcal{I}\}$, then, for each linear ordering $\leq$ on $\mathcal{I}$, the set $S = \{1, a_{i_1}^\epsilon \dots a_{i_n}^\epsilon \mid i_1 \leq \dots \leq i_n;\ i_1, \dots, i_n \in \mathcal{I}\}$ generates $U(L)$ as a K-module.*

Proof. Let M be the submodule of $U(L)$ generated by S. Since the tensor algebra $T(L)$ is generated as a K-module by the set $\{1, a_{i_1} \otimes \dots \otimes a_{i_n} \mid i_1, \dots, i_n \in \mathcal{I}, n \geq 1\}$, the set $\{1, a_{i_1}^\epsilon \dots a_{i_n}^\epsilon \mid i_1, \dots, i_n \in \mathcal{I}, n \geq 1\}$ generates $U(L)$ as a K-module. Hence, if $M \neq U(L)$, then there exist a least positive integers n and a permutation $(i_1, \dots, i_m, i_{m+1}, \dots, i_n)$ with minimal number of inversions such that the element $a_{i_1}^\epsilon \dots a_{i_m}^\epsilon a_{i_{m+1}}^\epsilon \dots a_{i_n}^\epsilon$ does not belong to M and $i_m > i_{m+1}$. As $[a_{i_m}, a_{i_{m+1}}] = \sum_{j \in \mathcal{I}} k_j a_j$ for some elements $k_j \in K$ only finitely many of which are non-zero, we have $a_{i_1}^\epsilon \dots a_{i_m}^\epsilon a_{i_{m+1}}^\epsilon \dots a_{i_n}^\epsilon = a_{i_1}^\epsilon \dots a_{i_{m+1}}^\epsilon a_{i_m}^\epsilon \dots a_{i_n}^\epsilon + \sum_{j \in \mathcal{I}} k_j a_{i_1}^\epsilon \dots a_{i_{m-1}}^\epsilon a_j^\epsilon a_{i_{m+2}}^\epsilon \dots a_{i_n}^\epsilon$. By the choice of n and $(i_1, \dots, i_m, i_{m+1}, \dots, i_n)$, every summand from the right side of this equality belongs to M. Therefore the element $a_{i_1}^\epsilon \dots a_{i_m}^\epsilon a_{i_{m+1}}^\epsilon \dots a_{i_n}^\epsilon$ must also belong to M, contrary to its choice, so that $M = U(L)$. □

Theorem 11.3. (Poincare-Birkhoff-Witt) *If L is a Lie algebra over K which is free as a K-module with a basis $\{a_i \mid i \in \mathcal{I}\}$, then for a linear ordering $\leq$ on $\mathcal{I}$ the set $S = \{1, a_{i_1}^{\epsilon} \dots a_{i_n}^{\epsilon} \mid i_1 \leq \dots \leq i_n;\ i_1, \dots, i_n \in \mathcal{I}, n \geq 1\}$ is a basis of the universal enveloping algebra $U(L)$ of L over K. In particular, the canonical Lie homomorphism $\epsilon : L \to U(L)$ is injective.*

Proof. By Lemma 11.2, the set S generates $U(L)$, so that it remains to prove that S is linearly independent over K. For this purpose, it suffices to construct a K-module V and a representation $\rho : L \to \mathrm{End}_K V$ such that the set $\{a_{i_1}^{\rho} \dots a_{i_n}^{\rho} \mid i_1 \leq \dots \leq i_n;\ i_1, \dots, i_n \in \mathcal{I}, n \geq 1\}$ is linearly independent over K. Indeed, since there exists a unique homomorphism $\phi : U(L) \to \mathrm{End}_K V$ of the associative algebras with unity such that $a^{\rho} = (a^{\epsilon})^{\phi}$ for every $a \in L$, the S must also be linearly independent over K.

Let $\mathcal{J}$ be the set of all finite ascending sequences of $\mathcal{I}$ including an empty sequence and let V be a free K-module with basis $\{v_j \mid j \in \mathcal{J}$. We want to define such a representation $\rho : L \to \mathrm{End}_K V$ with the above property. It is clear that ρ as a K-homomorphism is determined by its values on the set $\{a_i \mid i \in \mathcal{I}\}$. Moreover, each element $a_i^{\rho} \in \mathrm{End}_K V$ is determined by its values on the basic elements v_j of V. For brevity, we shall write $a_i v_j$ for $a_i^{\rho}(v_j)$.

Take any $i \in \mathcal{I}$ and $j \in \mathcal{J}$, so that $j = \{i_1, \dots, i_n\}$ for some $i_1, \dots, i_n \in \mathcal{I}$ with $i_1 \leq \dots \leq i_n$. We put $|j| = n$ and write $i \leq j$ if $i \leq i_1$. Furthermore, let $j' = \{i_2, \dots, i_n\}$ and let ij be the union $\{i\} \cup j$ ordered under $\leq$. Define the element $a_i v_j$ by induction on $|j|$ as follows:

$$a_i v_{\emptyset} = v_i\,,$$

$$a_i v_j = v_{ij} \qquad \text{if } i \leq j \text{ and}$$

$$(1) \qquad a_i v_j = [a_i, a_{i_1}] v_{j'} + a_{i_1}(a_i v_{j'}) \qquad \text{if } i \nleq j\,.$$

Since the Lie-commutator $[a_i, a_{i_1}]$ is a linear combination over K of basic elements $\{a_s \mid s \in \mathcal{I}\}$ and $|j'| = n - 1 < |j|$, the first summand of the right side of (1) is well defined. To show that the second summand is also defined we prove by induction on $|j|$ that

(2) every element $u_{i,j} = a_i v_j - v_{ij}$ is a linear combination over K of basic elements v_k of V with $|k| \leq |j|$.

Indeed, if $i \leq j$, then $a_i v_j = v_{ij}$ and so $u_{i,j} = 0$. Let $i \nleq j$. By induction assumption, equality (1) is defined if j is replaced by j' because $|j'| < |j|$, so that $a_i v_{j'} = [a_i, a_{i_2}] v_{(j')'} + a_{i_2}(a_i v_{(j')'})$. By the same reason, the element $u_{i,j'} = a_i v_{j'} - v_{ij'}$ is a linear combination over K of basic elements v_k of V with $|k| \leq |j'|$. Furthermore, $a_{i_2}(a_i v_{(j')'}) = a_{i_2}(v_{i(j')'} + u_{i,(j')'}) = v_{ij'} + a_{i_2} u_{i,(j')'}$. Therefore $a_i v_{j'} = v_{ij'} + [a_i, a_{i_2}] v_{(j')'} + a_{i_2} u_{i,(j')'}$ and so $u_{i,j'} = [a_i, a_{i_2}] v_{(j')'} + a_{i_2} u_{i,(j')'}$. Hence $a_{i_1}(a_i v_{j'}) = a_{i_1} v_{ij'} + a_{i_1} u_{i,j'} = v_{ij} + a_{i_1} u_{i,j'}$ and the element $a_{i_1} u_{i,j'}$ is a linear combination over K of elements v_k with $|k| \leq |j|$ by induction assumption. Thus the element $u_{i,j} = a_i v_j - v_{ij} = [a_i, a_{i_1}] v_{j'} + a_{i_1}(a_i v_{j'}) - v_{ij} = [a_i, a_{i_1}] v_{j'} + a_{i_1} u_{i,j'}$ is also such a linear combination, as desired.

To show that $\rho : L \to \mathrm{End}_K V$ is a Lie-homomorphism it suffices to prove that

$$(3) \qquad [a_i, a_s]v_j = a_i(a_s v_j) - a_s(a_i v_j)$$

for every $i, s \in \mathcal{I}$ with $i \le s$ and for each $j \in \mathcal{J}$. We argue by induction on $|j|$,.

If $i \le j$, then $v_{ij} = a_i v_j$ and so $a_s(a_i v_j) = [a_s, a_i]v_j + a_i(a_s v_j)$ by (1), as required. Let $i \not\le j$ and let k be the least element in the sequence j, so that $k < i \le s$. Then $j = kj'$ and therefore $v_j = a_k v_{j'}$. Hence we can rewrite (3) as

$$[a_i, a_s](a_k v_{j'}) = a_i(a_s(a_k v_{j'})) - a_s(a_i(a_k v_{j'})).$$

By induction assumption, $[[a_i, a_s], a_k]v_{j'} = [a_i, a_s](a_k v_{j'}) - a_k([a_i, a_s]v_{j'})$ because $|j'| < |j|$. Therefore $[a_i, a_s](a_k v_{j'}) = [[a_i, a_s], a_k]v_{j'} + a_k([a_i, a_s]v_{j'}) = [[a_i, a_s], a_k]v_{j'} + a_k(a_i(a_s v_{j'})) - a_k(a_s(a_i v_{j'}))$. Applying the Jacobi identity to the latter part of this equality, we have

$$\begin{aligned}
[a_i, a_s](a_k v_{j'}) = -[[a_s, a_k], a_i]v_{j'} - [[a_k, a_i], a_s]v_{j'} + a_k(a_i(a_s v_{j'})) - \\
a_k(a_s(a_i v_{j'}))) = -[a_s, a_k](a_i v_{j'}) + a_i([a_s, a_k]v_{j'}) - [a_k, a_i](a_s v_{j'}) + \\
a_s([a_k, a_i]v_{j'}) + a_k(a_i(a_s v_{j'})) - a_k(a_s(a_i v_{j'}))) = [a_k, a_s](a_i v_{j'}) + \\
a_i(a_s(a_k v_{j'})) - a_i(a_k(a_s v_{j'}))) + [a_i, a_k](a_s v_{j'}) + a_s(a_k(a_i v_{j'})) - \\
a_s(a_i(a_k v_{j'}))) + a_k(a_i(a_s v_{j'})) - a_k(a_s(a_i v_{j'}))) = \{a_i(a_s(a_k v_{j'})) - \\
a_s(a_i(a_k v_{j'}))\} + \{[a_k, a_s](a_i v_{j'}) + a_s(a_k(a_i v_{j'})) - a_k(a_s(a_i v_{j'}))\} + \\
\{[a_i, a_k](a_s v_{j'}) + a_k(a_i(a_s v_{j'})) - a_i(a_k(a_s v_{j'}))\}
\end{aligned}$$

Thus, to prove (3) it remains to show that the second and third braces are zero.

Indeed, since $a_i v_{j'} = v_{ij'} + u_{i,j'}$ by (2), we have $a_s(a_k(a_i v_{j'})) = a_s(a_k v_{ij'}) + a_s(a_k u_{i,j'})$ and $a_k(a_s(a_i v_{j'})) = a_k(a_s v_{ij'}) + a_k(a_s u_{i,j'})$. By the choice of k, it holds $a_k v_{ij'} = v_{k(ij')} = v_{ij}$. Furthermore, $a_s v_{ij} = [a_s, a_k]v_{ij'} + a_k(a_s v_{ij'})$ by (1) and $a_s(a_k u_{i,j'}) = a_k(a_s u_{i,j'}) + [a_s, a_k]u_{i,j'}$ by induction assumption. Therefore

$$\begin{aligned}
a_s(a_k(a_i v_{j'})) - a_k(a_s(a_i v_{j'})) = [a_s, a_k]v_{ij'} + a_k(a_s v_{ij'}) + \\
a_k(a_s u_{i,j'}) + [a_s, a_k]u_{i,j'} - a_k(a_s v_{ij'}) - a_k(a_s u_{i,j'}) = \\
[a_s, a_k]v_{ij'} + [a_s, a_k]u_{i,j'} = [a_s, a_k](a_i v_{j'}) = -[a_k, a_s](a_i v_{j'})
\end{aligned}$$

and hence the second brace is zero. By similar arguments, the third brace is also zero, so that equality (3) is proved.

Suppose now that a finite linear combination $\alpha = \sum_{j \in \mathcal{J}} c_j a^\rho_{i_1} \dots a^\rho_{i_n}$ with $j = \{i_1, \dots, i_n\}$ and $c_j \in K$ is equal to zero in $\mathrm{End}_K V$. Then $0 = \alpha(v_\emptyset) = \sum_j c_j a_{i_1}(\dots(a_{i_n} v_\emptyset)\dots) = \sum_j c_j v_j$ and so $c_j = 0$ for each j because $\{v_j \mid j \in \mathcal{J}\}$ is a basis of V. Thus the set $\{a^\rho_{i_1} \dots a^\rho_{i_n} \mid i_1 \le \dots \le i_n;\, i_1, \dots, i_n \in \mathcal{I}, \quad n \ge 1\}$ is linearly independent over K and the theorem is proved. $\square$