Remember! Not all rings are commutative!

1 Motivation

Definition. A ring R is called *Boolean* if all its elements are idempotents: for every $a \in R$, $a^2 = a$.

Proposition. *All Boolean rings are commutative.*

Proof. Suppose R is Boolean. We can write

$$1 + 1 = (1 + 1)^2 = 1^2 + 1^2 + 1^2 + 1^2 = 1 + 1 + 1 + 1$$

which implies that $1 = -1$ (i.e. R has characteristic 2). Now, fix $a, b \in R$. Then

$$a + b = (a + b)^2 = a^2 + ab + ba + b^2 = a + ab + ba + b$$

Subtracting $a + b$ from both sides gives $ab = -ba$, and since $1 = -1$, we get that $ab = ba$. Since this holds for all $a, b \in R$, we see that R is commutative. □

For an interesting extension of this theorem, call a ring R *troolean* if $a^3 = a$ for all $a \in R$. Then we have the following result:

Proposition. *All troolean rings are commutative.*

This can also be proved by similar elementary methods, though it's significantly trickier. However, both of these reults are special cases of the following massive generalization, which is both my favorite theorem and the subject of this talk:

Theorem (Jacobson's Commutativity Theorem)**.** *Let R be a ring, and suppose that it satisfies Jacobson's condition: for every $a \in R$, there exists some integer $n(a) \geq 2$ with $a^{n(a)} = a$. Then R is commutative.*

It's worth pondering this theorem—I find it deeply surprising, since it feels like powers should have very little to do with commutativity. Additionally, since we have no conditions on how $n(a)$ should behave, none of the elementary proofs, like the one above for Boolean rings, should be able to prove such a result. Indeed, in order to prove it, we will need to delve quite deeply into the structure theory of non-commutative rings, which we will do shortly.

2 Modules and Primitive Rings

Since we're dealing with non-commutative rings, we need to be careful about which side we multiply by scalars in our modules. For this talk, we will only be dealing with left modules, where scalars multiply on the left.

Definition. An R-module S is called *simple* if it has no non-trivial submodules, namely if $T \subseteq S$ is a sub-R-module, then $T = 0$ or $T = S$.

Simple modules have lots of nice properties, as exemplified by the following important lemma:

Lemma (Schur's Lemma)**.** *If S is a simple R-module, then $\mathrm{End}_R(S)$, the ring of R-linear endomorphisms of S, is a division ring, namely every non-zero element is invertible.*

Proof. Fix $0 \neq f \in \mathrm{End}_R(S)$. Then $\mathrm{im}(f)$ is a submodule of S, and since $f \neq 0$, we get that $0 \neq \mathrm{im}(f)$. Since S is simple, this implies that $\mathrm{im}(f) = S$, so f is surjective. Similarly, $\ker(f)$ is a submodule of S, and $\ker(f) \neq S$ since f is non-zero. So $\ker(f) = 0$, which means that f is injective. Thus, f is a bijection, and thus invertible. □

Definition. Given any R-module M, define the *annihilator* of M to be the collection of scalars that act trivially on M, namely

$$\mathrm{Ann}_R(M) = \{r \in R : rm = 0 \ \ \forall m \in M\}$$

A simple exercise is to prove that $\mathrm{Ann}_R(M)$ is a two-sided ideal in R. We call M a *faithful* module if $\mathrm{Ann}_R(M) = 0$, namely if all elements of R act non-trivally on M.

Definition. A ring R is called *primitive* if there is some R-module S that is both simple and faithful.

Example.

1. Any field K is primitive, since all K-modules (aka vector spaces) are faithful, and a one-dimensional vector space is also simple.
2. More generally, any division ring D is primitive, for exactly the same reason (in general, linear algebra over non-commutative division rings is very similar to linear algebra over fields).
3. If V is some vector space over a division ring D, then $R = \mathrm{End}_D(V)$ is a primitive ring. Indeed, V is an R-module, and it is straightforward to show that it is both simple and faithful. In the case where $\dim V = n < \infty$, this just says that the matrix ring $M_n(D)$ is primitive.
4. With a bit of effort, one can show that the free algebra $K\langle x, y\rangle$ is primitive for any field K.
5. If R is any ring and S is a simple module over R, then $R/\mathrm{Ann}_R(S)$ is a primitive ring, with S as the simple faithful module.

We will soon see that Example (3) is more or less the only example. Note that by Schur's Lemma, if R is primitive with a simple faithful module S, then $D = \mathrm{End}_R(S)$ is a division ring, and we can think of S as a D-vector space.

Definition. Suppose D is a division ring, and R is some ring acting on a D-vector space V (equivalently, R is a subring of $\mathrm{End}_D(V)$). Then we say that R acts *k-transitively*, for $k \in \mathbb{N}$, if for any collection $v_1, \dots, v_k \in V$ which are D-linearly independent, and for any $w_1, \dots, w_k \in V$, there is some $r \in R$ with $r(v_i) = w_i$ for all $1 \leq i \leq k$.

R is said to act *densely* or to be a *dense subring* of $\mathrm{End}_D(V)$ if it is k-transitive for all $k \in \mathbb{N}$.

Remark. There is a good reason for the "dense" terminology: suppose we give V the discrete topology, the function space V^V the product topology, and finally endow $\mathrm{End}_D(V) \subset V^V$ with the subspace topology. Then one can check that $R \subseteq \mathrm{End}_D(V)$ is dense (as above) if and only if it is topologically dense as a subspace.

With all of this groundwork, we are now able to state the first important structural theorem we will use:

Theorem (Jacobson-Chevalley Density Theorem)**.** *Suppose R is a primitive ring, with a simple faithful module S. Let $D = \mathrm{End}_R(S)$ be a division ring, think of S as a D-vector space, and of R as a subring of $\mathrm{End}_D(S)$. Then R is dense.*

The proof proves k-transitivity by induction on k, with the base case coming from the simplicity of S over R and the inductive step using faithfulness. Though the density theorem is very powerful on its own, we will more frequently use the following corollary of it:

Corollary. *Suppose R is a primitive ring. Then there is a division ring D such that one of the following happens:*

1. *R is isomorphic to the matrix ring $M_n(D)$ for some $n \in \mathbb{N}$*
2. *For all $n \in \mathbb{N}$, there is a subring $Q \subseteq R$ and a two-sided ideal $I \trianglelefteq Q$ such that $Q/I \cong M_n(D)$. Such a quotient of a subring is called a* slice*, so this says that R has $M_n(D)$ as a slice for all $n \in \mathbb{N}$.*

Proof. Let S be a simple faithful R-module, and let $D = \mathrm{End}_R(S)$. Then S is a D-vector space. If $\dim_D S = n < \infty$, then the n-transitivity of R implies that $R = \mathrm{End}_D(S) \cong M_n(D)$. This is because n-transitivity on an n-dimensional vector space is equivalent to saying that all endomorphisms can be found in R.

The other case is when $\dim_D S = \infty$. In that case, for any $n \in \mathbb{N}$, let $v_1, \dots, v_n$ be D-linearly independent in S. Let Q be the subring of R that fixes $\mathrm{span}(v_1, \dots, v_n)$, namely

$$Q = \{r \in R : rv_i \in \mathrm{span}(v_1, \dots, v_n) \ \ \forall i\}$$

and let

$$I = \{r \in Q : rv_i = 0 \ \ \forall i\}$$

Then by transitivity, we get that Q/I is a primitive ring with $\mathrm{span}(v_1, \dots, v_n)$ as a simple faithful module. By the case previously considered, we get that $Q/I \cong M_n(D)$. □

As a simple example of the power of this structure theorem, we can prove the following classification of commutative primitive rings:

Theorem. *If R is primitive and commutative, then it is a field.*

Proof. By the above, there is a division ring D such that either $R \cong M_n(D)$ for some n or else R has $M_n(D)$ as a slice for all n. For any $n \geq 2$, $M_n(D)$ is non-commutative, so $R \ncong M_n(D)$ if $n \geq 2$. Additionally, all slices of commutative rings are commutative, so R can't have $M_n(D)$ as a slice for $n \geq 2$. This only leaves the case where $R \cong M_1(D) \cong D$, so R is a division ring. Since it is additionally commutative, it is a field. □

3 The Jacobson Radical

The Jacobson radical is an important two-sided ideal inside any ring, which somehow collects all of the "bad stuff" that exists in the ring. It can be thought of as a replacement for the nilradical familiar from commutative algebra, since in general rings, the nilpotent elements need not form an ideal (or even an additive subgroup).

Definition. For a ring R, its Jacobson radical $J(R)$ is defined as

$$J(R) = \bigcap_{\substack{N \subseteq R \\ N \text{ maximal left ideal}}} N$$

This is kind of a weird definition, but here's a proposition that hopefully justifies the idea that $J(R)$ contains all the "bad" elements of R:

Proposition. *For $y \in R$, the following are equivalent:*

1. *$y \in J(R)$*
2. *$1 - xy$ is left invertible for all $x \in R$*
3. *For any simple R-module S, $y \in \mathrm{Ann}_R(S)$.*

Proof. First, suppose that $y \in J(R)$. Then if $1 - xy$ is not left-invertible, then $R(1 - xy)$ is a left ideal of R which is not the whole ring, and thus there is some maximal left ideal N that contains $1 - xy$. Additionally, since $y \in J(R) \subseteq N$, we have that $xy \in N$ as well. Thus, $1 \in N$, a contradiction. So (1) $\Rightarrow$ (2).

Now, suppose that $1 - xy$ is left invertible for all x, and suppose for contradiction that there is some simple module S and some $s \in S$ with $ys \neq 0$. Then consider the submodule Rys, which can't be the zero submodule since $ys \neq 0$. Since S is simple, we get that $Rys = S$, so there is some $x \in R$ with $xys = s$.

This gives $(1-xy)s = 0$, which, by the left invertibility of $1-xy$, implies that $s = 0$, a contradiction. So (2) $\Rightarrow$ (3).

Finally, suppose that $y \in \mathrm{Ann}_R(S)$ for all S, and let N be any maximal left ideal. Then R/N is a simple module, so $y \in \mathrm{Ann}_R(R/N)$, or equivalently $y \in N$. Since this holds for all maximal left ideals N, we get that $y \in J(R)$, and thus (3) $\Rightarrow$ (1). $\square$

An immediate corollary of this proposition is the following:

Corollary.

$$J(R) = \bigcap_{S \text{ a simple } R\text{-module}} \mathrm{Ann}_R(S)$$

In particular, $J(R)$ is a two-sided ideal, as each $\mathrm{Ann}_R(S)$ is a two-sided ideal.

We need one more structural definition.

Definition. Suppose that $R, \{R_i\}_{i \in I}$ are rings, for some index set I. We say that R is a *subdirect product* of the R_i if there is some injective ring homomorphism

$$\iota : R \hookrightarrow \prod_{i \in I} R_i$$

such that each of the compositions $\pi_i \circ \iota : R \to R_i$ is surjective.

Remark. Since direct products are dual to direct sums and injections are dual to surjections, one can think of subdirect products as dual to presentations, which is when we present R as a quotient of a direct sum. This analogy doesn't make a ton of sense in the category of rings, since direct sums aren't really well defined, but it does actually work in other categories (e.g. abelian groups).

The reason the notion of subdirect products is useful is that we can often express a ring we're interested in as a subdirect product of "simpler" rings. Since we get surjections $R \twoheadrightarrow R_i$, we can often push forward information from R to information about the R_i, while at the same time using the injection $R \hookrightarrow \prod R_i$ to get information about R from properties of the R_i. This is the basic proof technique we will use to prove the Jacobson commutativity theorem. In order to do this, we first need the following fact:

Theorem. *If R is a ring with $J(R) = 0$, then R is the subdirect product of primitive rings $\{R_i\}_{i \in I}$.*

Rings with zero Jacobson radical are often called semiprimitive*, and this theorem explains why.*

Proof. Let $\{S_i\}_{i \in I}$ be the collection of all simple modules of R (up to isomorphism). Set $R_i = R/\mathrm{Ann}_R(S_i)$; since R_i is the quotient of R by an annihilator of a simple module, we know that R_i is primitive. Since we have natural maps $R \twoheadrightarrow R_i$, we get a map to the product,

$$\iota : R \to \prod_{i \in I} R_i$$

All that remains to prove is that ι is injective, namely that $\ker(\iota) = 0$. However, by definition,

$$\ker(\iota) = \bigcap_{i \in I} \ker(R \twoheadrightarrow R_i) = \bigcap_{i \in I} \mathrm{Ann}_R(S_i) = J(R) = 0$$

$\square$

Remark. One can be worried in this proof that the collection of all simple modules of R is a proper class, rather than a set. This doesn't actually pose a problem, though, since we only actually use the collection of all annihilators of simple modules, and there are at most $2^{|R|}$ such annihilators, as each is a subset of R.

Example. Since the maximal ideals of $\mathbb{Z}$ are $p\mathbb{Z}$ for primes p, we see that $J(\mathbb{Z}) = 0$, so we should be able to represent $\mathbb{Z}$ as a subdirect product of primitive rings. Indeed, if we go through the process described in this theorem, we see that we get

$$\mathbb{Z} \hookrightarrow \prod_{p \text{ prime}} \mathbb{Z}/p\mathbb{Z}$$

with the map given by

$$a \mapsto (a \mod p)_{p \text{ prime}}$$

which is indeed injective to the product but surjective onto each factor. Moreover, since $\mathbb{Z}/p\mathbb{Z}$ is a field for each prime p, the rings in the subdirect product are indeed all primitive.

4 Proving the Jacobson Commutativity Theorem

We are now ready to prove the Jacobson commutativity theorem. Remember the statement: if a ring R satisfies the Jacobson condition, namely that each $a \in R$ has some $n(a) \in \mathbb{N}$ such that $a^{n(a)} = a$, then R is commutative.

Lemma. *If R satisfies the Jacobson condition, then $J(R) = 0$.*

Proof. Suppose $a \in J(R)$. Then by our second characterization of the Jacobson radical, we know that $1 - a^{n(a)-1}$ is left invertible; let its left inverse be b. Then

$$1 = b(1 - a^{n(a)-1})$$

Multiplying by a on the right gives

$$a = b(a - a^{n(a)}) = b \cdot 0 = 0$$

so $a = 0$, so $J(R) = 0$. □

Therefore, by the theorem we proved earlier, we know that we can represent R as a subdirect product of primitive rings $\{R_i\}$.

Lemma. *Each R_i satisfies the Jacobson condition.*

Proof. We know that we have a surjective homomorphism $s : R \twoheadrightarrow R_i$. So for each $c \in R_i$, we can find some $a \in R$ that maps to c. But therefore

$$c^{n(a)} = s(a)^{n(a)} = s(a^{n(a)}) = s(a) = c$$

So setting $n(c) = n(a)$, we get that $c^{n(c)} = c$, so R_i satisfies the Jacobson condition. □

Lemma. *If a primitive ring R_i satisfies the Jacobson condition, then it is a division ring.*

Proof. By the corollary of the density theorem, we know that there is some division ring D such that either $R_i \cong M_n(D)$ for some n, or else R_i has a slice isomorphic to $M_n(D)$ for all n. Next, observe that if $n \geq 2$, then no $M_n(D)$ can satisfy the Jacobson condition; this is because $M_n(D)$ has nilpotent elements for $n \geq 2$ (e.g. upper triangular matrices with zeroes on the diagonal). No non-zero nilpotent element can exist in a ring satisfying the Jacobson condition, so we see that $M_n(D)$ cannot satisfy the Jacobson condition for $n \geq 2$. Thus, if $R_i \cong M_n(D)$, then we must have $n = 1$ and $R_i \cong D$.

The remaining case is where R_i has a slice isomorphic to $M_n(D)$ for all $n \in \mathbb{N}$. But the same argument holds: any subring Q of R_i must also satisfy the Jacobson condition, as must any quotient of Q, by the argument above. So R_i cannot have $M_n(D)$ as a slice for any $n \geq 2$, so we are not in this situation, and $R_i \cong D$. □

Now all that remains is proving the following theorem:

Theorem. *If D is a division ring satisfying the Jacobson condition, then it is commutative (i.e. D is a field).*

Why does this suffice? By the above, if R satisfies the Jacobson condition, then it is a subdirect product of a bunch of division rings that each satisfy the Jacobson condition. So once we prove this theorem, we will get that R is a subdirect product of a bunch of fields, and in particular, it is a subring of a commutative ring, and is thus itself commutative.

5 The Division Ring Case

One important theorem we need is the following:

Theorem (Wedderburn's Little Theorem)**.** *If R is a finite domain (i.e. it has no zero divisors), then it's a field.*

The fact that all finite domains are division rings is quite straightforward, and the fact that they are commutative follows from some clever applications of the theory of cyclotomic polynomials. A consequence of Wedderburn's Little Theorem is the following:

Corollary. *If D is a division ring of positive characteristic and G is a finite subgroup of D^*, then G is abelian.*

Proof. Let P be the prime field of D, namely the subfield generated by 1. Consider the group algebra

$$P[G] = \left\{ \sum_{g \in G} p_g g \;\middle|\; p_g \in P \right\}$$

We can view $P[G]$ as a subring of D. Since D is a division ring, $P[G]$ has no zero divisors. Additionally, since both P and G are finite, so is $P[G]$. So by Wedderburn's Little Theorem, $P[G]$ is a field, so in particular G must be abelian. □

We need one final technical lemma, which is essentially a group theory lemma.

Lemma. *Let D be a division ring with positive characteristic, and let Z be the center of D, namely the set of all elements that commute with everything in D. Let $a \in D \setminus Z$, and suppose that a is algebraic over the prime field $P \subseteq D$. Let A be the subgroup generated by a. Then the normalizer of A is strictly bigger than the centralizer of A; namely, there is some $w \in D^*$ with $waw^{-1} = a^i \neq a$, for some exponent i.*

Proof. The element a defines a derivation $\delta : D \to D$, defined by

$$\delta(x) = xa - ax$$

One can check that δ satisfies the Leibniz rule, which is why it's called a derivation. From the Leibniz rule, we get that

$$\delta^k(x) = \sum_{i=0}^{k} \binom{k}{i} (-1)^k a^k x a^{k-i}$$

If the characteristic of D is $p > 0$, then this implies that $\delta^{p^k}(x) = xa^{p^k} - a^{p^k}x$ for all $k \in \mathbb{N}$. Since a is algebraic over P, we get that $P(a)$ is a finite field, so there is some m with $a^{p^m} = a$. Therefore, by the above, δ satisfies the same relation, $\delta^{p^m} = \delta$. Additionally, $P(a)$ is the splitting field for the polynomial $x^{p^m} - x$, so we may write

$$x^{p^m} - x = (x - \lambda_{p^m}) \cdots (x - \lambda_2)x$$

for some non-zero $\lambda_i \in P(a)$. Moreover, the Leibniz rule implies that δ commutes with multiplication by $\lambda_i \in P(a)$, so we can substitute δ into this equation to get that

$$0 = \delta^{p^m} - \delta = (\delta - \lambda_{p^m} \operatorname{id}_D) \cdots (\delta - \lambda_2 \operatorname{id}_D)\delta$$

Since $a \notin Z$, it doesn't commute with everything, so we get that $\delta \neq 0$. Thus, there is some $k \geq 2$ such that

$$(\delta - \lambda_{k-1} \operatorname{id}_D) \cdots (\delta - \lambda_2 \operatorname{id}_D)\delta \neq 0$$

while

$$(\delta - \lambda_k \operatorname{id}_D) \cdots (\delta - \lambda_2 \operatorname{id}_D)\delta = 0$$

Therefore, we get some $w \neq 0$ with $(\delta - \lambda_k \operatorname{id}_D)(w) = 0$, or equivalently $wa - aw = \lambda_k w$. This means that $waw^{-1} = \lambda_k + a$, and $\lambda_k \neq 0$. In particular, $waw^{-1} \neq a$. On the other hand, waw^{-1} and a have the same order as elements of the group $P(a)^*$, and since $P(a)^*$ is cyclic, they must generate the same cyclic subgroup, namely A. Thus, $waw^{-1} = a^i$ for some i, while $waw^{-1} \neq a$, as desired. □

Finally, we can prove what we want, namely that all division rings satisfying the Jacobson condition are commutative.

Proof. With $2 = 1 + 1$, we know that $2^{n(2)} = 2$, which implies that D has some finite characteristic. Let Z be the center of D and P the prime field of D. Every $a \in D$ satisfies the polynomial $x^{n(a)} - x$, so it is algebraic over P. If D is not commutative, then $Z \neq D$, so there is some $a \in D \setminus Z$, which must be algebraic over P. So by the previous lemma, we get some $w \in D^*$ with $waw^{-1} = a^i \neq a$. Let G be the subgroup of D^* generated by a and w. We know that both a and w have finite order, namely $n(a) - 1$ and $n(w) - 1$, respectively. Additionally, since waw^{-1} is a power of a, we get that w normalizes a, and thus the subgroup generated by a is normal in G. So G has a finite normal subgroup of finite index, so G is finite. By the corollary of Wedderburn's Little Theorem, we get that G is abelian. But that contradicts the fact that $waw^{-1} \neq a$, so we get that D is indeed commutative. Combining this with all our previous work, we get Jacobson's Commutativity Theorem. □

Remark. One disadvantage of the Jacobson Commutativity Theorem is that its converse is blatantly false (e.g. $\mathbb{Z}$ is commutative, but does not satisfy the Jacobson condition). Herstein later proved the following generalization:

Theorem. *R is commutative if and only if for all $x, y \in R$, there is some $n(x, y) > 1$ such that*

$$(xy - yx)^{n(x,y)} = xy - yx$$

In other words, requiring the Jacobson condition to hold only for commutators gives us an if and only if statement.