Abstract

The main focus of this thesis is Wedderburn's theorem that a finite division ring is a field. We present two proofs of this. The thesis also contains a proof of a theorem of Jacobson and a proof of a generalisation by Artin and Zorn that a finite alternative ring is associative, and therefore a field.

0 Introduction

It is well-known that a ring is a set with two operations in which certain axioms hold. Usually each one of the axioms seems independent of the other ones, at least at a first look. In this thesis, we will give two different proofs of a theorem that shows that that is not always the case. At first we make a formal definition of the special kind of ring that will be studied.

Definition 0.1. *A division ring is a ring D with identity 1_D satisfying the following axiom: For all $a \in D\backslash\{0_D\}$, the equation $ax = 1_D$ has a solution $x \in D$.*

The difference between the definition of a division ring and the definition of a field is that in a field we also assume multiplicative commutativity.

The main theorem of this thesis is Wedderburn's theorem.

Theorem 0.1 (Wedderburn's theorem)**.** *A finite division ring is a field.*

What Theorem 0.1 really tells us is that in the finite case, the general ring axioms together with existence of identity and inverses automatically imply multiplicative commutativity, i.e: $ab = ba$ for all $a, b \in D$.

In Section 1 we prove the so called class equation, which will be used in section two where we give a first proof of Wedderburn's theorem. Section 3 contains a second proof of Wedderburn's theorem. Section 4 contains a proof of a theorem of Jacobson. When nothing else is mentioned, [3] (mainly Section 7.2) is used for the definitions and proofs in Section 1-4.

In Section 5 we consider so-called alternative rings, which are rings that are not necessarily associative. We prove that finite alternative rings with identity in which the above inverse axiom hold are fields. Here [9] is used.

In Section 6, the end of the thesis, we give examples that shows that there exist non-commutative division rings and non-associative alternative rings. This is taken from [8] when nothing else is mentioned.

To understand this thesis, some basic knowledge of discrete mathematics and abstract algebra is needed.

1 The class equation

In the first proof of Wedderburn's theorem we need a formula in group theory called the Class equation. For that we need some theory about equivalence classes and a special equivalence relation called conjugacy. At first, we make formal definitions of equivalence relations and equivalence classes. In the two first definitions and the two first theorems, [5] has been used as a complement of [3].

Definition 1.1. *An equivalence relation on a set S is a relation $\sim$ that for all $a, b, c \in S$ satisfies the three following conditions:*
1: *$a \sim a$ (reflexivity)*
2: *$a \sim b \Rightarrow b \sim a$ (symmetry)*
3: *$a \sim b, b \sim c \Rightarrow a \sim c$ (transitivity)*

Definition 1.2. *Let S be a set, $a \in S$. Then the equivalence class of a is $[a] = \{x \in S \mid a \sim x\}$.*

We have two important results that are valid in any equivalence class.

Proposition 1.1. *Let S be a set, $a, b \in S$. Then $a \sim b \iff [a] = [b]$.*

Proof. $\Rightarrow$: Assume that $a \sim b$. We prove that $[a] = [b]$ by showing that both $[a] \subseteq [b]$ and $[b] \subseteq [a]$. Let $c \in [a]$. Then $c \sim a$. Now we have both $c \sim a$ and $a \sim b$, so by transitivity $c \sim b$, and hence $c \in [b]$. So we have that $c \in [a]$ implies that $c \in [b]$, which forces $[a] \subseteq [b]$. By symmetry, our assumption $a \sim b$ implies that $b \sim a$. So by reversing the roles of a and b in the above argument, we get that $[b] \subseteq [a]$, and thus we must have $[a] = [b]$, which is what we wanted.
$\Leftarrow$: Assume that $[a] = [b]$. Reflexivity gives that $a \in [a]$, which is the same as $a \in [b]$, which means that $a \sim b$, and the theorem is proved. □

Proposition 1.2. *Two equivalence classes of an equivalence relation on a set S are either equal or disjoint.*

Proof. Let $a, b \in S$, and let $[a]$ and $[b]$ be equivalence classes. If $[a] \cap [b] = \varnothing$, we are done. Otherwise, there exists an element $c \in [a] \cap [b]$, which implies that $c \in [a]$ but also $c \in [b]$. This gives that $c \sim a$ and $c \sim b$. By symmetry and transitivity, $a \sim b$. Proposition 1.1 now gives that $[a] = [b]$. □

Now we shall define a special relation called conjugacy and show that it is an equivalence relation.

Definition 1.3. *Let G be a group, $a, b \in G$, then b is a conjugate of a if $b = c^{-1}ac$ for some $c \in G$. We write this as $a \diamond b$, and the relation is called conjugacy.*

Proposition 1.3. *Conjugacy is an equivalence relation on a group G, which means that for all $a, b, c \in G$, the following conditions hold:*
1: *$a \diamond a$*
2: *$a \diamond b \Rightarrow b \diamond a$*
3: *$a \diamond b, b \diamond c \Rightarrow a \diamond c$.*

Proof. Let $a, b, c \in G$.
1: We let e be the identity element in G. Then $a = e^{-1}ae$, and hence $a \diamond a$.
2: Assume that $a \diamond b$. Then $b = x^{-1}ax$ for some $x \in G$. Multiplying with $(x^{-1})^{-1}$ from the left and with x^{-1} from the right we get $(x^{-1})^{-1}bx^{-1} = (x^{-1})^{-1}x^{-1}axx^{-1}$, which is the same as $a = (x^{-1})^{-1}bx^{-1}$. Now let $x^{-1} = y$. Then $a = y^{-1}by$, where $y \in G$, which tells that $a \sim b$.
3: Assume that $a \sim b$ and $b \sim c$. Then $b = x^{-1}ax$ and $c = y^{-1}by$, for some $x, y \in G$. Hence we can write $c = y^{-1}(x^{-1}ax)y = (y^{-1}x^{-1})a(xy) = (xy)^{-1}a(xy)$, where $xy \in G$, such that $a \sim c$, and the proof is done. □

The equivalence class $Cl(a) = \{x \in G \mid a \diamond x\}$ of a in G is called the conjugacy class of a in G.

Every $a \in G$ is contained in $Cl(a)$ since $a \diamond a$, and therefore a is contained in at least one conjugacy class. Since different equivalence classes are disjoint, every a is contained in exactly one conjugacy class. Hence we have that $|G| = \sum |Cl(a)|$ where each conjugacy class is represented exactly once in the sum.

Now we are going to define something else that, as we will see soon, has strong connections with the conjugacy classes.

Definition 1.4. *Let G be a group, $a \in G$. Then the centralizer of a in G is the set $C(a) = \{x \in G \mid xa = ax\}$.*

Proposition 1.4. *$C(a)$ is a subgroup of G.*

Proof. Since a clearly commutes with itself, we must have $a \in C(a)$, so that $C(a)$ is nonempty. Assume that $x, y \in C(a)$. We have to show closure and existence of inverse elements in $C(a)$.
Closure: We have that $xa = ax$ and $ya = ay$. Hence

$$(xy)a = x(ya) = x(ay) = (xa)y = (ax)y = a(xy),$$

which gives that $xy \in C(a)$.
Inverse: We have that

$$x^{-1}a = x^{-1}a(xx^{-1}) = x^{-1}(ax)x^{-1} = x^{-1}(xa)x^{-1} = (x^{-1}x)ax^{-1} = ax^{-1},$$

and therefore $x^{-1} \in C(a)$, and we are done. □

The strong connections between conjugacy classes and the centralizer will be clear in the following theorem. Here [4] has been used as a complement.

Theorem 1.5. *Let G be a finite group, then $|Cl(a)| = \frac{|G|}{|C(a)|}$, i.e. the number of elements in G that are conjugate to a equals the number of right cosets (the so-called index) of the centralizer of $a \in G$.*

Proof. We shall show that there is a one-to-one correspondence between conjugates of $a \in G$ and right cosets of $C(a)$. We say that x generates the conjugate α of a if $\alpha = x^{-1}ax$. We do this proof by showing that two elements are in

the same right coset of $C(a)$ if and only if they generate the same conjugate of $a \in G$.

Let $x, y \in G$, and assume that x and y lie in the same coset of $C(a)$, i.e. $C(a)x = C(a)y$. Then $x \in C(a)y$, which gives that $x = ky$ for some $k \in C(a)$. Therefore

$$x^{-1}ax = (ky)^{-1}a(ky) = (y^{-1}k^{-1})aky = y^{-1}k^{-1}(ak)y = \\ = y^{-1}k^{-1}(ka)y = y^{-1}(k^{-1}k)ay = y^{-1}ay,$$

which means that x and y generate the same conjugate of $a \in G$.

We now assume that x and y generate the same conjugate of $a \in G$. Thus $x^{-1}ax = y^{-1}ay$. If we multiply with x from the left and with y^{-1} from the right, we get $(xx^{-1})axy^{-1} = xy^{-1}a(yy^{-1})$, and hence $a(xy^{-1}) = (xy^{-1})a$. This implies that $xy^{-1} \in C(a)$, which means that $x \equiv y \pmod{C(a)}$, which is equivalent to $C(a)x = C(a)y$ since congruence is an equivalence relation. The proof is now complete. □

From this theorem we get the main result of this chapter, the so-called Class equation.

Corollary 1.6 (Class Equation)**.** *Let G be a group and $C(a)$ the conjugacy class of $a \in G$. Then $|G| = \sum \frac{|G|}{|C(a)|}$, where the sum runs over exactly one element a from each conjugacy class.*

Proof. The formula follows immediately from Theorem 1.5 and the fact that $|G| = \sum |Cl(a)|$. □

We end this section with a property that will be used in the next section.

Proposition 1.7. *Let G be a group, and let $Z(G) = Z = \{z \in G \mid zx = xz$ for all $x \in G\}$ be its center. Then $a \in Z \Leftrightarrow C(a) = G$.*

Proof. $\Rightarrow$: Assume that $a \in Z$. Then $xa = ax$ for all $x \in G$, and hence $C(a) = G$. $\Leftarrow$: Assume that $C(a) = G$. Then $xa = ax$ for all $x \in G$, which gives that $a \in Z$, and we are done. □

2 Wedderburn's theorem

We need some more lemmas and results before the actual proof of Wedderburn's theorem.

Proposition 2.1. *A finite subring of a division ring is itself a division ring.*

Proof. Let D be a division ring and R a finite subring of D. We let $a \in R \backslash \{0_R\}$, and must show that then also $a^{-1} \in R$. Closure of R gives that all elements in the set $S = \{a, a^2, a^3, ...\}$ lies in R. Since R is finite, we must have $0_D \neq a^j = a^k \in S \subseteq R$ for some $j > k$. In D we have $0_D = a^j - a^k = (a^{j-k} - 1_D)a^k$. Since there are no zero divisors in D and $a^k \neq 0_D$, we must have $a^{j-k} = 1_D$. Now $a^{j-k} \in S \subseteq R$, and hence $1_D \in R$. We now have that $aa^{j-k-1} = a^{j-k} = 1_D$, which implies that $a^{-1} = a^{j-k-1} \in R$. □

Proposition 2.2. *Let R be a ring, and let Z be its center defined by*

$$Z = \{z \in R \mid zx = xz \text{ for all } x \in R\}.$$

Then Z is a subring of R.

Proof. Since $0_R \in Z$, $Z \neq \varnothing$. We have to show closure under subtraction and multiplication. Let $\alpha, \beta \in Z$.
Subtraction: We have that $\alpha x = x\alpha$ and $\beta x = x\beta$ for all $x \in D$. Hence $(\alpha - \beta)x = \alpha x - \beta x = x\alpha - x\beta = x(\alpha - \beta)$, and therefore $\alpha - \beta \in Z$.
Multiplication: We have that $\alpha x = x\alpha$ and $\beta x = x\beta$. Hence $(\alpha\beta)x = \alpha(\beta x) = \alpha(x\beta) = (\alpha x)\beta = (x\alpha)\beta = x(\alpha\beta)$, and therefore $\alpha\beta \in Z$. □

If D is a division ring, we know that $D \backslash \{0_D\}$ is a group under multiplication. So it is natural to define the centralizer $C(a) = \{x \in D \mid xa = ax\}$. Here the zero element 0_D in $C(a)$ is included.

Proposition 2.3. *Let R be a ring. Then the centralizer $C(a)$ of R is a subring of R.*

Proof. This is done the same way as in Proposition 2.2. Just change Z to $C(a)$ and the variable x to a fixed a. □

Proposition 2.4. *Let D be a finite division ring and let K be a subring of D that is also a division ring. If K contains q elements, D contains q^n elements, where n is the dimension of D as a vector space over K.*

Proof. Assume that n is the dimension of D over K. Then D has a basis of n vectors. Call them $e_1, e_2, ..., e_n$. Then every element $a \in D$ can be written as $a = \alpha_1 e_1 + \alpha_2 e_2 + ... + \alpha_n e_n$, where all $\alpha_i \in K$. Hence the number of elements in K is the number of different $\alpha_1 e_1 + \alpha_2 e_2 + ... + \alpha_n e_n$. Since K has q elements, there are q choices for each α_i. Therefore, by the multiplication principle, there are q^n elements in D. □

Lemma 2.5. *If $(x^m - 1) \mid (x^n - 1)$ in $\mathbb{Z}[x]$, then $m \mid n$.*

Proof. We assume that $m \nmid n$, and shall show that $(x^m-1) \nmid (x^n-1)$. We have $n = qm + r$, where $q \in \mathbb{N}$ and $0 < r < m$. We use long division of polynomials to get the relation

$$(x^n - 1) = (x^m - 1)(x^{n-m} + x^{n-2m} + ... + x^{n-qm}) + (x^r - 1),$$

which can easily be verified. Here $\deg(x^r - 1) = r < m = \deg(x^m - 1)$, so that the relation is exactly the division algorithm when $x^n - 1$ is divided by $x^m - 1$. Therefore $(x^m - 1) \nmid (x^n - 1)$ □

Corollary 2.6. *If $t \in \mathbb{N}\backslash\{0, 1\}$ and $(t^m - 1)|(t^n - 1)$, then $m \mid n$.*

Proof. Substituting x with $t \in \mathbb{N}\backslash\{0, 1\}$ in the above proof we get

$$(t^n - 1) = (t^m - 1)(t^{n-m} + t^{n-2m} + ... + t^{n-qm}) + (t^r - 1),$$

which, since $0 < r < m$ implies that $0 < t^r - 1 < t^m - 1$, is the division algorithm when the integer $t^n - 1$ is divided by the integer $t^m - 1$. Hence $(t^m - 1) \nmid (t^n - 1)$. □

For the proof of Wedderburn's theorem, we also need something called cyclotomic polynomials. Here [7] is used as a complement in Definition 2.1, Proposition 2.7 and Proposition 2.8. In $\mathbb{C}$ the solutions of $\alpha^k = 1$ are the k numbers on the form $\alpha = e^{2i\pi j/k}$, where $j \in \{0, 1, ..., k - 1\}$. These α's are the k roots of the polynomial $x^k - 1 \in \mathbb{C}[x]$. The α's with $\alpha^r \neq 1$ whenever $r < k$ are called primitive kth roots of unity. These are exactly the α's where $\gcd(j, k) = 1$. It follows from the fact that if j and k have a common nontrivial factor, say l, then $\alpha^{k/l} = (e^{2i\pi j/k})^{k/l} = e^{2i\pi \frac{j}{l}} = 1$, and $k/l < k$ so that the kth root α is not primitive.

Definition 2.1. *The polynomial*

$$\Phi_k(x) = \prod_{\substack{\gcd(j,k)=1 \\ 1 \leqslant j \leqslant k}} (x - e^{2i\pi j/k})$$

is called the kth cyclotomic polynomial.

Proposition 2.7. *We have that $x^k - 1 = \prod_{k' \mid k} \Phi_{k'}(x)$.*

Proof. From the factor theorem, we have

$$x^k - 1 = \prod_{1 \leqslant j \leqslant k} (x - e^{2i\pi j/k}).$$

We let $r = \gcd(j, k)$, $j' = j/r$ and $k' = k/r$. Then $e^{2i\pi j/k} = e^{2i\pi j'/k'}$, where $\gcd(j', k') = 1$. Now $(x - e^{2i\pi j'/k'})$ is a factor in $\Phi_{k'}(x)$. Since j runs over all integers from 1 to k, all the possible such fractions j'/k' in simplest form where $k'|k$ will be obtained this way. Therefore there is a one-to-one correspondence between the factors in $x^k - 1$ and $\prod_{k' \mid k} \Phi_{k'}(x)$, and we have equality. □

Proposition 2.8. *For all $k \in \mathbb{N}\backslash\{0\}$, $\Phi_k(x)$ is a monic polynomial with integer coefficients.*

Proof. We use induction on k. We have that $\Phi_1(x) = x - e^{2i\pi} = x - 1$, which is a monic polynomial with integer coefficients. We make the induction assumption that $\Phi_d(x)$ is monic and with integer coefficients when $d < k$ and $d|k$, and shall show that then also $\Phi_k(x)$ is monic and with integer coefficients. We now have that $x^k - 1 = \Phi_k(x)f(x)$, where $f(x)$ is a product of monic polynomials with integer coefficients, and hence itself monic and with integer coefficients. Then $\Phi_k(x) = \frac{x^k-1}{f(x)}$ has integer coefficients by an analogue of the division algorithm for monic polynomials. □

Proposition 2.9. *For all $d \mid k$ with $d < k$ we have that*

$$\Phi_k(x) \,\Big|\, \frac{x^k - 1}{x^d - 1},$$

and the quotient is a polynomial with integer coefficients.

Proof. We have that

$$x^k - 1 = \prod_{d \mid k} \Phi_d(x),$$

and therefore even

$$x^d - 1 = \prod_{r \mid d} \Phi_r(x).$$

We have that every divisor of d also divides k. Thus

$$x^k - 1 = (x^d - 1) \prod_{\substack{r \mid k \\ r \nmid d}} \Phi_r(x).$$

For a fixed d with $d < k$ we have that $\Phi_k(x)$ is not a factor in $x^d - 1$, and hence

$$\frac{x^k - 1}{x^d - 1} = \Phi_k(x) \prod_{\substack{r \mid k \\ r \neq k \\ r \nmid d}} \Phi_r(x),$$

which gives that

$$\Phi_k(x) \,\Big|\, \frac{x^k - 1}{x^d - 1},$$

and the quotient

$$\frac{\frac{x^k-1}{x^d-1}}{\Phi_k(x)} = \prod_{\substack{r \mid k \\ r < k \\ r \nmid d}} \Phi_r(x)$$

is a product of polynomials with integer coefficients, which is again a polynomial with integer coefficients. □

Corollary 2.10. *For all $t \in \mathbb{Z}$, $k \in \mathbb{N}\backslash\{0,1\}$ and $d \mid k$ with $d \neq k$ we have*

$$\Phi_k(t) \,\Big|\, \frac{t^k - 1}{t^d - 1}.$$

Proof. Since the product

$$\prod_{\substack{r \mid k \\ r<k \\ r \nmid d}} \Phi_r(x)$$

from the proof of Proposition 2.9 is a polynomial with integer coefficients, it follows immediately that the quotient

$$\frac{\frac{t^k-1}{t^d-1}}{\Phi_k(t)} = \prod_{\substack{r \mid k \\ r<k \\ r \nmid d}} \Phi_r(t)$$

is an integer. □

Lemma 2.11. *Let $\theta \in \mathbb{C}$ with $\theta \neq 1$ be a kth root of unity and $q \in \mathbb{N}\backslash\{0\}$. Then $|q - \theta| > q - 1$.*

Proof. Let $\theta = a + bi$, where $a, b \in \mathbb{R}$. Then $|\theta| = \sqrt{a^2 + b^2} = 1$. This and the fact that $\theta \neq 1$ give that $a < 1$. We now have

$$\begin{aligned} |q - \theta| = |q - (a + bi)| = |(q - a) - bi| &= \sqrt{(q-a)^2 + (-b)^2} \\ &= \sqrt{q^2 - 2qa + (a^2 + b^2)} = \sqrt{q^2 - 2aq + 1} \\ &> \sqrt{q^2 - 2q + 1} = \sqrt{(q-1)^2} = |q - 1| = q - 1. \end{aligned}$$

□

Now we are ready for the first proof of Wedderburn's theorem.

Proof. Let D be a finite division ring. We will show that the multiplicative commutativity axiom holds in D by showing that its center $Z = \{z \in D | zx = xz$ for all $x \in D\}$ has the same number of elements as the whole of D. Because that would imply that $Z = D$, and therefore the axiom would hold in D.

We assume that Z has q elements. Then, by Proposition 2.4, D has q^n elements for some $n \in \mathbb{N}\backslash\{0\}$. So we want to show that we must have $n = 1$. We define the centralizer $C(a) = \{x \in G \mid xa = ax\}$ for $a \in D$. Then Z is contained in $C(a)$, and since Z is a subring of D, it is a subring of $C(a)$. Thus, by Proposition 2.4 again, $C(a)$ contains $q^{m(a)}$ elements, where $m(a) \in \mathbb{N}\backslash\{0\}$ is depending on $a \in D$.

Now we have that the groups $D\backslash\{0_D\}$, $Z\backslash\{0_D\}$, and $C(a)\backslash\{0_D\}$ (under multiplication) have orders $q^n - 1$, $q - 1$ and $q^{m(a)} - 1$ respectively. Since we know from Proposition 2.3 that $C(a)$ is a subring of D, it is clear that $C(a)\backslash\{0_D\}$ is a subgroup of $D\backslash\{0_D\}$. Lagrange's theorem in group theory therefore gives that

$(q^{m(a)}-1) \mid (q^n-1)$. Hence we have that $m(a) \mid n$ by Corollary 2.6. The number of elements in the conjugacy class $Cl(a)$ for $a \in D\backslash\{0_D\}$ is $(q^n-1)/(q^{m(a)}-1)$. So by the class equation we have

$$q^n - 1 = \sum_{m(a)\,|\,n} \frac{q^n-1}{q^{m(a)}-1},$$

where the sum runs over exactly one a from each conjugacy class. Proposition 1.7 gives that $a \in Z$ if and only if $C(a) = D$, but $C(a) = D$ if and only if $m(a) = n$. Hence we can rewrite the equation as

$$q^n - 1 = (q-1) + \sum_{\substack{m(a)\,|\,n \\ m(a)\neq n}} \frac{q^n-1}{q^{m(a)}-1},$$

where the sum now only runs over a's that are not contained in Z.

We know that this equation holds under our assumptions and definitions and shall show that the equality is impossible unless $n = 1$. We assume for contradiction that $n > 1$ and shall find an integer which divides all the terms in the equation except $q-1$, which leads to the contradiction that the integer divides the left-hand side but not the right-hand side of the equation.

If we in Corollary 2.10 let the t be our q and k be our n, we have that

$$\phi_n(q) \,\Big|\, \frac{q^n-1}{q^d-1},$$

when $d \mid n$ and $d < n$. Then it is obvious that we also have $\phi_n(q) \mid q^n - 1$. Therefore we have found an integer which divides the left-hand side and all the terms in the sum. It only remains to show that $\phi_n(q) \nmid (q-1)$. By Lemma 2.11, $|q-\theta| > q-1$ when θ is a root of unity. We must have $q \geqslant 2$ since $0_D, 1_D \in Z$. Therefore $|\phi_n(q)| = \prod |q-\theta| > q-1$, and $\phi_n(q) \nmid (q-1)$. Therefore we must have $n = 1$, which forces that a finite division ring is a field! $\square$

3 Second proof of Wedderburn's theorem

Now we shall give a second proof of Wedderburn's Theorem. We need several definitions and results before the actual proof. For Definition 3.1, Definition 3.2 and Proposition 3.1 [2] is used.

Definition 3.1. *A group endomorfism is a homomorphism from a group to itself.*

Definition 3.2. *Let G be an abelian group under addition, then we let End(G) be the set of endomorphisms in G, that is, End(G)=* $\{f : G \to G \mid f(a+b) = f(a) + f(b)\}$.

Proposition 3.1. *Define addition and multiplication in End(G) as* $(f+g)(a) = f(a) + g(a)$ *and* $fg(a) = f(g(a))$. *Then it is a ring.*

Proof. We have to check all the ring axioms. Let $f, g, h \in \mathrm{End}(G)$ and $a, b \in G$. At first we show closure under addition. We have to show that $f + g$ is again an endomorphism. We have

$$\begin{aligned}(f+g)(a+b) &= f(a+b) + g(a+b) = f(a) + f(b) + g(a) + g(b) \\ &= (f(a) + g(a)) + (f(b) + g(b)) = (f+g)(a) + (f+g)(b)\end{aligned}$$

and therefore it is true.

Now we show additive commutativity. Since G is abelian we have

$$(f+g)(a) = f(a) + g(a) = g(a) + f(a) = (g+f)(a).$$

Additive associativity follows from

$$\begin{aligned}\big(f + (g+h)\big)(a) &= f(a) + \big((g+h)(a)\big) \\ &= f(a) + \big(g(a) + h(a)\big) \\ &= \big(f(a) + g(a)\big) + h(a) \\ &= (f+g)(a) + h(a) \\ &= \big((f+g) + h\big)(a).\end{aligned}$$

It is easily checked that the endomorphism k defined by $k(a) = e_G$ for all $a \in G$ works as $0_{End(G)}$, and that the identity map from G to G works as $1_{End(G)}$.

We define the element $-f$ by $(-f)(a) = -f(a)$. It follows from

$$\begin{aligned}(-f)(a+b) &= -f(a+b) = -\big(f(a) + f(b)\big) \\ &= -f(a) - f(b) = (-f)(a) + (-f)(b)\end{aligned}$$

that $-f \in \mathrm{End}(G)$. It is also clear that $f + (-f) = 0_{End(G)}$.

The multiplication is clearly associative, since both $(fg)h(a)$ and $f(gh)(a)$ means $f(g(h(a)))$.

Closure under multiplication follows from

$$fg(a+b) = f(g(a)+g(b)) = f(a'+b'),$$

where $a'+b' \in G$.

At last we must show the distributive laws. We have

$$\begin{aligned}(f(g+h))(a) &= f((g+h)(a)) = f(g(a)+h(a)) \\ &= f(g(a)) + f(h(a)) = fg(a) + fh(a) \\ &= (fg+fh)(a),\end{aligned}$$

so that $f(g+h) = fg+fh$. We also have

$$\begin{aligned}((f+g)h)(a) &= (f+g)(h(a)) = f(h(a)) + g(h(a)) = (fh)(a) + gh(a) \\ &= (fh+gh)(a),\end{aligned}$$

so that $(f+g)h = fh+gh$. This finishes the proof. □

Now we shall define a special endomorphism that will be used. The proof of the binomial theorem in [1] has been used as help for the proof of Lemma 3.2.

Lemma 3.2. *Let R be a ring and $a \in R$. Let $T_a : R \to R$ be the endomorphism (Here R is viewed as an abelian group under addition) with $T_a(x) = xa - ax$. For $m \in \mathbb{N}$ we define $T_a^m(x)$ such that $T_a^2(x) = T_a(T_a(x))$, $T_a^3(x) = T_a(T_a(T_a(x)))$ and so on.*

$$T_a^m(x) = \sum_{k=0}^{m} (-1)^k \binom{m}{k} a^k x a^{m-k}.$$

Proof. We prove this formula by induction on m. For $m = 1$, the formula gives $T_a = \sum_{k=0}^{1} (-1)^k \binom{1}{k} a^k x a^{1-k} = xa - ax$, which is true by the definition of T_a.

Now we assume that the formula is true for $m = n$, and show that it is then even true for $m = n+1$. We have

$$
\begin{aligned}
T_a^{n+1}(x) &= T_a(T_a^n(x)) = \\
&= \left(\sum_{k=0}^{n}(-1)^k\binom{n}{k}a^k x a^{n-k}\right)a - a\left(\sum_{j=0}^{n}(-1)^j\binom{n}{j}a^j x a^{n-j}\right) \\
&= \sum_{k=0}^{n}(-1)^k\binom{n}{k}a^k x a^{n-k+1} + \sum_{j=0}^{n}(-1)^{j+1}\binom{n}{j}a^{j+1}x a^{n-j} \\
&= \sum_{k=1}^{n}(-1)^k\binom{n}{k}a^k x a^{n-k+1} + xa^{n+1} \\
&\quad + \sum_{k=1}^{n}(-1)^k\binom{n}{k-1}a^k x a^{n-k+1} + (-1)^{n+1}a^{n+1}x \\
&= \sum_{k=1}^{n}(-1)^k\left[\binom{n}{k} + \binom{n}{k-1}\right]a^k x a^{n-k+1} + xa^{n+1} + (-1)^{n+1}a^{n+1}x \\
&= \sum_{k=1}^{n}(-1)^k\binom{n+1}{k}a^k x a^{n-k+1} + xa^{n+1} + (-1)^{n+1}a^{n+1}x \\
&= \sum_{k=0}^{n+1}(-1)^k\binom{n+1}{k}a^k x a^{n-k+1},
\end{aligned}
$$

which is exactly the form the formula should take when $m = n+1$. Hence the formula follows by induction. In the calculations we used the so-called Pascal's relation

$$\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k},$$

which is very easy to prove by just writing the fractions on the left-hand side on a common denominator and then cancelling some common factors. □

Corollary 3.3. *Let R be a ring, and let p be a prime with $px = 0_R$ for all $x \in R$. Then $T_a^{p^m}(x) = xa^{p^m} - a^{p^m}x$ for all $m \in \mathbb{N}\backslash\{0\}$.*

Proof. Lemma 3.2 gives

$$T_a^{p^m}(x) = \sum_{k=0}^{p^m}(-1)^k\binom{p^m}{k}a^{p^m}xa^{p^m-k}.$$

It is shown in [6] that p is a factor in $\binom{p^m}{k}$ except when $k = 0$ or $k = p^m$. Hence all terms in the sum except the first and the last vanish, which means that

$$T_a^{p^m}(x) = xa^{p^m} + (-1)^{p^m}a^{p^m}x.$$

When p is odd, we immediately get $T_a^{p^m}(x) = xa^{p^m} - a^{p^m}x$, which is what we wanted. When $p = 2$ we have

$$T_a^{2^m}(x) = xa^{2^m} + a^{2^m}x - 0_R = xa^{2^m} + a^{2^m}x - 2a^{2^m}x = xa^{2^m} - a^{2^m}x,$$

and the proof is done. □

Lemma 3.4. *Let F be a field containing p^n elements. In $F[x]$ we then have*

$$x^{p^n} - x = \prod_{\lambda \in F}(x - \lambda).$$

Proof. Let $g(x) = x^{p^n} - x$. Then $\deg(g) = p^n$. Therefore we know that $g(x)$ has at most p^n roots in F. Since $\lambda^{p^n-1} = 1_F$ for all $\lambda \in F\backslash\{0_F\}$, we have that $\lambda^{p^n} = \lambda$, which is the same as $\lambda^{p^n} - \lambda = 0_F$, which obviously is true even for $\lambda = 0_F$. This gives that every $\lambda \in F$ is a root of $g(x)$, and the formula follows from the factor theorem. □

We need a theorem that we state without proof. It can be proved the same way as Theorem 22.4 in [2].

Theorem 3.5. *Let R and S be rings and let $f : R \to S$ be a homomorphism. Let $a \in S$ be an element that commutes with $f(x)$ for all $x \in R$. Then there exists a homomorphism $\bar{f} : R[u] \to S$ such that $\bar{f} \restriction_R = f$ and $\bar{f}(u) = a$. Here $R[u]$ denotes a polynomial ring.*

Lemma 3.6. *Let D be a division ring of characteristic p (the smallest $p \in \mathbb{N}\backslash\{0\}$ with $p1_D = 0_D$). Let Z be the center of D and let $P = \{0_F, 1_F, 2 \cdot 1_F, ..., (p-1)1_F\}$ be the subfield of D isomorphic to $\mathbb{Z}_p$. Let $a \in D\backslash Z$ be an element satisfying $a^{p^n} = a$ for some $n \in \mathbb{N}\backslash\{0\}$. Then there exists an $x \in D$ with $xax^{-1} \neq a$ but $xax^{-1} \in P(a)$, where $P(a)$ is the field obtained by adjoining a to P.*

Proof. The relation $a^{p^n} = a$ implies that $a \in D$ is a root of the polynomial $x^{p^n} - x$ in $P[x]$. Thus a is algebraic over P, which implies that $P(a)$ is a finite field containing p^m elements for some $m \in \mathbb{N}\backslash\{0\}$. All these elements $y \in P(a)$ satisfy $y^{p^m} = y$ (as in the proof of Lemma 3.4). Defining the function $T_a : D \to D$ as before, we now have $T_a^{p^m}(z) = za^{p^m} - a^{p^m}z = za - az = T_a(z)$ for all $z \in D$. Hence $T_a^{p^m} = T_a$. Let $\alpha \in P(a)$ and $x \in D$. Then α commutes with a since both are contained in $P(a)$, and

$$\begin{aligned} T_a(\alpha x) &= (\alpha x)a - a(\alpha x) = \alpha xa - (a\alpha)x = \alpha xa - (\alpha a)x \\ &= \alpha(xa - ax) = \alpha T_a(x). \end{aligned}$$

We can therefore say that the endomorphism I_α satisfying $I_\alpha(x) = \alpha x$ commutes with T_a in $\mathrm{End}(D)$.

Now we use Lemma 3.4 to write

$$u^{p^m} - u = \prod_{\alpha \in P(a)}(u - \alpha).$$

In Theorem 3.5 we let R be our $P(a)$, S our End(D), a our T_a, $R[u]$ the polynomial ring $P(a)[u]$ and f the homomorphism with $f(\alpha) = I_\alpha$. Then we can use $\bar{f}$ on both the left-hand-side and the right-hand-side of the above equality to get

$$T_a^{p^m} - T_a = \prod_{\alpha \in P(a)} (T_a - I_\alpha).$$

Since $T_a^{p^m} - T_a = 0_D$ we get

$$\prod_{\alpha \in P(a)} (T_a - I_\alpha) = 0_D.$$

Now, for all $\alpha \in P(a)\backslash\{0_{p(a)}\}$, we assume that $y = 0_D$ whenever $(T_a - I_\alpha)(y) = 0_D$. Then the relation forces that $T_a = 0_{End(D)}$, so that $T_a(y) = 0_D$ for all $y \in D$. This means that $ya - ay = 0_D$ for all $y \in D$, so that $a \in Z$, which contradicts the assumption that $a \in D\backslash Z$.

There must therefore exist some $\alpha \in P(a)$ and $x \in D$ with $\alpha, x \neq 0_D$ satisfying $(T_a - I_\alpha)(x) = 0_D$, which means that $xa - ax - \alpha x = 0_D$. Moving ax and αx to the right-hand-side and multiplying with x^{-1} from the right we get $xax^{-1} = a + \alpha$. Since both a and α are in $P(a)$, closure gives that $xax^{-1} \in P(a)$, but $a + \alpha \neq a$ since $\alpha \neq 0_D$. Therefore the proof is complete. □

Corollary 3.7. *The element $xax^{-1} \in P(a)$ in the above lemma satisfies $xax^{-1} = a^i$ for some $i \in \mathbb{N}\backslash\{1\}$.*

Proof. Let ord$(a) = k$ in the field $P(a)$. Then the set $S = \{1, a, a^2, ..., a^{k-1}\}$ is the set of all the k distinct roots of the polynomial $u^k - 1_{P(a)} \in P(a)[u]$. We have that

$$\begin{aligned}(xax^{-1})^k &= (xax^{-1})(xax^{-1})...(xax^{-1}) = xa(x^{-1}x)ax^{-1}...xax^{-1} \\ &= xa^kx^{-1} = xx^{-1} = 1_{P(a)},\end{aligned}$$

such that $xax^{-1} \in S$. Hence $xax^{-1} = a^i$ for some $i \in \{0, 2, 3, ..., k-1\}$. Here we exclude 1, since we know that $xax^{-1} \neq a$. More generally we have $i \in \mathbb{N}\backslash\{1\}$, which was to be proved. □

Lemma 3.8. *Let F be a finite field, and let $\alpha, \beta \in F\backslash\{0_F\}$. Then there exist $a, b \in F$ satisfying $1_F + \alpha a^2 + \beta b^2 = 0_F$.*

Proof. Assume that F has characteristic 2. Then F contains 2^n elements, where $n \in \mathbb{N}\backslash\{0\}$. Then for all $x \in F$ we have $x^{2^n} = x$, which gives that any $x \in F$ is a square. This implies that we can let $\alpha^{-1} = a^2$ for some $a \in F$. If we let $b = 0_F$, we can use this a and b to get

$$1_F + \alpha a^2 + \beta b^2 = 1_F + \alpha\alpha^{-1} + \beta 0_F = 1_F + 1_F + 0_F = 2(1_F) = 0_F,$$

which is what we wanted.

Assume that F has characteristic p, where p is an odd prime. Then F has p^n elements for some $n \in \mathbb{N}\backslash\{0\}$. We define the set $S_\alpha = \{1_F + \alpha x^2 \mid x \in F\}$ and

shall calculate the number of elements in S_α. We do this by checking how often $1_F + \alpha x^2 = 1_F + \alpha y^2$. The relation is equivalent to that $\alpha x^2 - \alpha y^2 = 0_F$. Since $\alpha \neq 0_F$ we can multiply by α^{-1} to get $x^2 - y^2 = (x+y)(x-y) = 0$, leading to $x = \pm y$. This gives that for all $w \in F$, $x = w$ and $x = -w$ generate the same element of S_α, but for two elements $i, j \in F$ with $i \neq \pm j$ we have that $x = i$ and $x = j$ generate different elements of S_α. Therefore, when $x \neq 0_F$, each of the $\frac{p^n-1}{2}$ pairs $x, -x$ generate one element in S_α each, and $x = \pm 0$ generate only one element. Thus the number of elements in S_α are

$$\frac{p^n - 1}{2} + 1 = \frac{p^n - 1 + 2}{2} = \frac{p^n + 1}{2}.$$

The same reasoning on the set $S_\beta = \{-\beta x^2 \mid x \in F\}$ gives that S_β also contains $\frac{p^n+1}{2}$ elements. Together S_α and S_β contain $p^n + 1 > p^n$ elements, which means that $S_\alpha \cap S_\beta \neq \varnothing$, so there exists an element $r \in S_\alpha \cap S_\beta$. Then $r = 1_F + \alpha a^2$ for some $a \in F$, but also $r = -\beta b^2$ for some $b \in F$. Using these a and b we get $0_F = r - r = 1_F + \alpha a^2 + \beta b^2$, and we are done. □

Now it is time for the second proof of Wedderburn's theorem.

Proof. Let D be a finite division ring and $Z = \{z \in D \mid z\gamma = \gamma z \text{ for all } \gamma \in D\}$ its center. Our goal in this proof is to reach a contradiction when assuming that $D \backslash Z \neq \varnothing$. So we assume for contradiction that there exists an element $w \in D \backslash Z$. In order to reach the contradiction, we need several technicalities, and will therefore divide the proof into some steps.

We let D be a division ring and make the induction assumption that any division ring having fewer elements than D is a field. Using this assumption we make a first claim.

Claim 3.9. *Let $\alpha, \beta \in D$ such that $\alpha\beta^t = \beta^t\alpha$ for some $t \in \mathbb{N}\backslash\{0,1\}$, but $\alpha\beta \neq \beta\alpha$. Then $\beta^t \in Z$.*

Proof. We have that $C(\beta^t) = \{\gamma \in D \mid \beta^t\gamma = \gamma\beta^t\}$ is a subring of D that is also a division ring. Now $\alpha, \beta \in C(\beta^t)$ but α and β do not commute, which implies that $C(\beta^t)$ is not commutative. Therefore, by the induction assumption, $C(\beta^t) = D$. By Proposition 1.7, $\beta^t \in Z$. □

Since D is finite, $D\backslash\{0_D\}$ is a finite group under multiplication, and hence every element in $D\backslash\{0_D\}$ (and specifically in $D\backslash Z$ since $0_D \in Z$) has finite order. Therefore $w^k = 1_D$ for some $k \in \mathbb{N}\backslash\{0,1\}$. Since $1_D \in Z$ it follows that when $w \in D\backslash Z$, $w^s \in Z$ for some $s \in \mathbb{N}\backslash\{0,1\}$. The smallest such s we call the order of w relative to Z. Clearly $s \leqslant k$. We now let a be an element in $D\backslash Z$ with the least order relative to Z. This least order, call it r, must be a prime number. We prove this claim by assuming for contradiction that r is composite. Then $r = r_1 r_2$ for some r_1, r_2 with $1 < r_1, r_2 < r$. Since $a^r \in Z$ we now have $a^{r_1 r_2} = (a^{r_1})^{r_2} \in Z$, which implies that a^{r_1} has order $\leqslant r_2 < r$ contradicting the fact that r is the least order.

Now we shall use this $a \in D \backslash Z$ with prime order r relative to Z to produce two elements a_1 and b_1 satisfying these two conditions:
1: $a_1^r = b_1^r \in Z$.
2: $a_1 b_1 = \mu b_1 a_1$, where $\mu \in Z$ with $\mu \neq 1_d$ but $\mu^r = 1_D$.

We know from Corollary 3.7 that there exists an $x \in D$ with $xax^{-1} = a^i \neq a$, where $i \in \mathbb{N} \backslash \{0, 1\}$. Using this, we have

$$x^2 a x^{-2} = x(xax^{-1})x^{-1} = xa^i x^{-1} = (xax^{-1})^i = (a^i)^i = a^{i^2},$$
$$x^3 a x^{-3} = x(x^2 a x^{-2})x^{-1} = xa^{i^2} x^{-1} = (xax^{-1})^{i^2} = (a^{i^2})^i = a^{i^3},$$

and so on. Specifically $x^{r-1} a x^{-(r-1)} = a^{i^{r-1}}$. Since r is a prime, Fermat's little theorem gives that $i^{r-1} \equiv 1 \pmod{r}$, which means that $i^{r-1} = qr + 1$ for some $q \in \mathbb{Z}$. Therefore we have $a^{i^{r-1}} = a^{qr+1} = a^{qr} a$, which we rewrite as λa where $\lambda = a^{qr} = (a^r)^q$, so that $\lambda \in Z$ by closure of Z. This implies that $x^{r-1} a x^{-(r-1)} = \lambda a$, which by multiplying with x^{r-1} from the right can be rewritten as $x^{r-1} a = \lambda a x^{r-1}$.

We have that $xa \neq ax$. This follows from the fact that if we assume $xa = ax$ we would reach the contradiction $xax^{-1} = a$ by just multiplying with x^{-1} from the right in the relation. Therefore $x \notin Z$, implying that $x^{r-1} \notin Z$ by the definition of r. We have that $x^{r-1} a \neq a x^{r-1}$. Otherwise Claim 3.9 would lead us to the contradiction that $x^{r-1} \in Z$. Hence we have that $\lambda \neq 1_D$. If we let $b = x^{r-1}$ we have $ba = \lambda ab$, or $bab^{-1} = \lambda a$ when just multiplying with b^{-1} from the right. Since $a^r \in Z$ we now have

$$\lambda^r a^r = (\lambda a)^r = (bab^{-1})^r = ba^r b^{-1} = a^r (bb^{-1}) = a^r,$$

forcing $\lambda^r = 1_D \in Z$. Hence the order $|\lambda| = r$.

Claim 3.10. *Let $y \in D$ such that $y^r = 1_D$. Then $y = \lambda^i$ for some $i \in \mathbb{N}$.*

Proof. In the field $Z(y)$ obtained by adjoining Z and y, there are at most r roots of the polynomial $u^r - 1_D \in Z(y)[u]$. The set $S = \{1_D, \lambda, \lambda^2, ..., \lambda^{r-1}\}$ is the set of r different roots of the polynomial. Hence we must have $y \in S$, which gives that $y = \lambda^i$ for some $i \in \{0, 1, ..., r-1\}$, or more generally $i \in \mathbb{N}$. □

Since $\lambda^r = 1_D$ and $\lambda \in Z$ we have that

$$b^r = 1_D b^r = \lambda^r b^r = (\lambda b)^r = (a^{-1} ba)^r = a^{-1} b^r a,$$

which, by multiplying with a from the left, gives $ab^r = b^r a$. Claim 3.9 and the fact that $ab \neq ba$ now gives that $b^r \in Z$.

Since Z is a field, it contains a primitive element σ, which works as a cyclic generator of the multiplicative group $Z \backslash \{0_D\}$. Therefore $a^r = \sigma^m$ and $b^r = \sigma^n$, for some $m, n \in \mathbb{N}$.

Claim 3.11. *We have that $r \nmid m$ and $r \nmid n$.*

Proof. Assume for contradiction that $m = kr$ for some $k \in \mathbb{N}$. Then $a^r = \sigma^{kr}$, and multiplying with σ^{-kr} gives that $a^r\sigma^{-kr} = 1_D$, or equivalently $(a\sigma^{-k})^r = 1_D$. Then, by Claim 3.10, $a\sigma^{-k} = \lambda^i$ for some $i \in \mathbb{N}$. Multiplying with σ^k, we get $a = \lambda^i\sigma^k$, and closure of Z hence implies that $a \in Z$, which is a contradiction. Therefore $r \nmid m$, and by the same argument we also have $r \nmid n$. □

Now let $a_1 = a^m$ and $b_1 = b^n$.

Claim 3.12. *We have that* $a_1b_1 = \lambda^{-mn}b_1a_1$.

Proof. From the relation $bab^{-1} = \lambda a$ from above we get $\lambda^{-1}ba = ab$ by multiplying with b from the right and with λ^{-1} from the left. We now have that

$$\begin{aligned}
a_1b_1 &= a^mb^n = a^{m-1}(ab)b^{n-1} = a^{m-1}(\lambda^{-1}ba)b^{n-1} = \lambda^{-1}a^{m-2}(ab)ab^{n-1} \\
&= \lambda^{-1}a^{m-2}(\lambda^{-1}ba)ab^{n-1} = \lambda^{-2}a^{m-3}(ab)a^2b^{n-1} = ... = \lambda^{-m}ba^mb^{n-1} \\
&= \lambda^{-m}ba^{m-1}(ab)b^{n-2} = \lambda^{-m}ba^{m-1}(\lambda^{-1}ba)b^{n-2} = \lambda^{-(m+1)}ba^{m-2}(ab)ab^{n-2} \\
&= ... = \lambda^{-2m}b^2a^mb^{n-2} = ... = \lambda^{-mn}b^na^m = \lambda^{-mn}b_1a_1.
\end{aligned}$$

□

Now let $\mu = \lambda^{-mn} \in Z$. Since $r{=}|\lambda|$ is a prime and $r \nmid m$, $r \nmid n$, we have $r \nmid mn$. Hence $\mu = \lambda^{-mn} \neq 1_D$, but we have

$$\mu^r = (\lambda^{-mn})^r = \lambda^{-(mn)r} = (\lambda^r)^{-mn} = 1_D^{-mn} = 1_D.$$

Now we have reached our first goal, i.e. we have produced elements a_1 and b_1 satisfying
1: $a_1^r = b_1^r \in Z$.
2: $a_1b_1 = \mu b_1a_1$, where $\mu \in Z$ with $\mu \neq 1_d$ but $\mu^r = 1_D$.

Now we shall proceed by using these elements a_1 and b_1 to reach the contradiction we want.

Claim 3.13. *We have that*

$$(a_1^{-1}b_1)^r = \mu^{\frac{r(r-1)}{2}}.$$

Proof. From $a_1b_1 = \mu b_1a_1$ we get $b_1a_1^{-1} = (a^{-1}\mu)b_1 = \mu a_1^{-1}b_1$ by multiplying with a_1^{-1} both from the left and from the right, and hence

$$\begin{aligned}
(a_1^{-1}b_1)^2 &= a_1^{-1}(b_1a_1^{-1})b_1 = a_1^{-1}(\mu a_1^{-1}b_1)b_1 = \mu a_1^{-2}b_1^2, \\
(a_1^{-1}b_1)^3 &= a_1^{-1}(b_1a_1^{-1})^2b_1 = a_1^{-1}(\mu a_1^{-1}b_1)^2b_1 = a_1^{-1}\mu^2(a_1^{-1}b_1)^2b_1 \\
&= a_1^{-1}\mu^2(\mu a_1^{-2}b_1^2)b_1 = \mu^{1+2}a_1^{-3}b_1^3, \\
(a_1^{-1}b_1)^4 &= a_1^{-1}(b_1a_1^{-1})^3b_1 = a_1^{-1}(\mu a_1^{-1}b_1)^3b_1 = a_1^{-1}\mu^3(a_1^{-1}b_1)^3b_1 \\
&= a_1^{-1}\mu^3(\mu^{1+2}a_1^{-3}b_1^3)b_1 = \mu^{1+2+3}a_1^{-4}b_1^4,
\end{aligned}$$

and so on. Specifically

$$\begin{aligned}(a_1^{-1}b_1)^r &= \mu^{1+2+\ldots+(r-1)}a_1^{-r}b_1^r = \mu^{1+2+\ldots+(r-1)}(a^r)^{-1}a^r \\ &= \mu^{1+2+\ldots+(r-1)} = \mu^{\frac{r(r-1)}{2}},\end{aligned}$$

which was to be proved. □

Claim 3.14. *The prime r must be odd.*

Proof. Assume for contradiction that $r = 2$. Then $a_1^2 = b_1^2 = \alpha$ for some $\alpha \in Z$, and $a_1b_1 = \mu b_1a_1$ with $\mu \neq 1$, but $\mu^2 = 1_D$. Hence μ satisfies $0_D = \mu^2 - 1_D = (\mu - 1_D)(\mu + 1_D)$. Since Z contain no zero divisors and $\mu \neq 1_D$, we must have $\mu = -1_D$. Therefore $a_1b_1 = -b_1a_1 \neq b_1a_1$. The last inequality implies that $0_D \neq 2b_1a_1 = 2 \cdot 1_D b_1a_1$, and therefore D has not characteristic 2.

In Lemma 3.8, we now let the α be our $-\alpha$ and $\beta = 1_Z$. Then there are $\tau, \omega \in Z$ satisfying $1_D + \tau^2 - \alpha\omega^2 = 0_D$. We now consider the expression $(a_1 + \tau b_1 + \omega a_1b_1)^2$. We have

$$\begin{aligned}(a_1 + \tau b_1 + \omega a_1b_1)^2 &= a_1^2 + \tau a_1b_1 + \omega a_1^2b_1 + \tau b_1a_1 + \tau^2 b_1^2 + \tau\omega b_1a_1b_1 \\ &\quad + \omega a_1b_1a_1 + \tau\omega a_1b_1^2 + \omega^2 a_1b_1a_1b_1 \\ &= a_1^2 + \tau a_1b_1 + \omega a_1^2b_1 - \tau a_1b_1 + \tau^2 b_1^2 - \tau\omega a_1b_1^2 \\ &\quad - \omega b_1a_1^2 + \tau\omega a_1b_1^2 - \omega^2 a_1^2b_1^2 \\ &= \alpha + \tau^2\alpha - \omega^2\alpha^2 = \alpha(1_D + \tau^2 - \alpha\omega^2) = \alpha 0_D = 0_D\end{aligned}$$

Since D contains no zero divisors we must have that $a_1 + \tau b_1 + \omega a_1b_1 = 0_D$. Since the characteristic is not 2 and $a_1 \neq 0_D$, we have that $2a^2 \neq 0_D$. But

$$\begin{aligned}2a_1^2 &= 2a_1^2 + 0_D + 0_D = 2a_1^2 + (\tau a_1b_1 - \tau a_1b_1) + (\omega a_1^2b_1 - \omega a_1^2b_1) \\ &= (a_1^2 + \tau a_1b_1 + \omega a_1^2b_1) + (a_1^2 + \tau b_1a_1 + \omega a_1b_1a_1) \\ &= a_1(a_1 + \tau b_1 + \omega a_1b_1) + (a_1 + \tau b_1 + \omega a_1b_1)a_1 \\ &= a_1 0_D + 0_D a_1 = 0_D\end{aligned}$$

which is a contradiction, so the claim is proved. □

Since r is odd we have

$$(a_1^{-1}b_1)^r = \mu^{\frac{r(r-1)}{2}} = (\mu^r)^{\frac{r-1}{2}} = 1_D^{\frac{r-1}{2}} = 1_D,$$

so that $(a_1^{-1}b_1)$ solves $y^r = 1_D$. Therefore, by Claim 3.10, we have that $a_1^{-1}b_1 = \lambda^j$ for some $j \in \mathbb{N}$. Multiplying with a_1 from the left we get $b_1 = a_1\lambda^j = \lambda^j a_1$. Then $\mu b_1a_1 = a_1b_1 = a_1(\lambda^j a_1) = (\lambda^j a_1)a_1 = b_1a_1$, which contradicts the fact that $\mu \neq 1_D$. Therefore the assumption that there is an element $w \in D\backslash Z$ cannot be true, and thus we must have that $Z = D$, and Wedderburn's theorem is again proved! □

4 Jacobson's theorem

Now we focus on a generalisation of Wedderburn's theorem.

Theorem 4.1 (Jacobson's theorem)**.** *Let D be a division ring such that for all $a \in D$, there is a $h(a) \in \mathbb{N}\backslash\{0,1\}$ depending on a, satisfying $a^{h(a)} = a$. Then D is a field.*

Proof. If $a \in D\backslash\{0_D\}$ we have that $a^n = a$ and $(2a)^m = 2a$ for some $m, n \in \mathbb{N}\backslash\{0,1\}$. Let $s = (n-1)(m-1)+1$. Then $s > 1$, and by construction of s we have

$$a^s = a^{(n-1)(m-1)+1} = (a^{n-1})^{m-1}a = (a^n a^{-1})^{m-1}a = (aa^{-1})^{m-1}a = 1_D a = a,$$

and also

$$\begin{aligned}(2a)^s &= (2a)^{(n-1)(m-1)+1} = ((2a)^{m-1})^{n-1}(2a) = ((2a)^m(2a)^{-1})^{n-1}(2a)\\ &= ((2a)(2a)^{-1})^{n-1}(2a) = 2a.\end{aligned}$$

We also have $(2a)^s = 2^s a^s = 2^s a$. Combining these relations we get $2^s a = 2a$, or equivalently $(2^s - 2)a = 0_D$. This gives that D has characteristic $p > 0$. We let P be the subfield of Z isomorphic to $\mathbb{Z}_p$. Since a is a root of the polynomial $u^n - u \in P[u]$, a is algebraic over P. Therefore $P(a)$, the field obtained by adjoining a to P, is finite and contains p^h elements for some $h \in \mathbb{N}\backslash\{0\}$. Hence $a^{p^h-1} = 1_{P(a)}$, so that $a^{p^h} = a$. We now fix our a and assume for contradiction that $a \in D\backslash Z$. Then the conditions in Lemma 3.6 and Corollary 3.7 are satisfied, so that there is a $b \in D$ with $bab^{-1} = a^\mu$, for some $\mu \in \mathbb{N}\backslash\{1\}$. Since $b \in D$, the same argument as above gives that $b^{p^k} = b$ for some $k \in \mathbb{N}\backslash\{0,1\}$. Now let

$$S = \left\{ x \in D \;\middle|\; x = \sum_{i=1}^{p^h}\sum_{j=1}^{p^k} \rho_{ij} a^i b^j \text{ where } \rho_{i,j} \in P \right\}.$$

The set S is clearly finite but nonempty. We show that it is a subring of D. We have to show closure under subtraction and multiplication. Closure under subtraction is almost immediate. The elements in S can be seen as linear combinations of the vectors $a^i b^j$, and we know that a linear combination of some vectors minus another linear combination of the same vectors is again a linear combination of the vectors. To show closure under multiplication we let $y, z \in S$. Then $y = \sum_{i=1}^{p^h}\sum_{j=1}^{p^k} \rho_{ij} a^i b^j$ and $z = \sum_{r=1}^{p^h}\sum_{s=1}^{p^k} \omega_{rs} a^r b^s$. We now have that

$$yz = \left(\sum_{i=1}^{p^h}\sum_{j=1}^{p^k} \rho_{ij} a^i b^j\right)\left(\sum_{r=1}^{p^h}\sum_{s=1}^{p^k} \omega_{rs} a^r b^s\right) = \sum_{i=1}^{p^h}\sum_{j=1}^{p^k}\sum_{r=1}^{p^h}\sum_{s=1}^{p^k} \rho_{ij}\omega_{rs} a^i b^j a^r b^s.$$

From $bab^{-1} = a^\mu$ we get $ba = a^\mu b$ by multiplying with b from the right. Therefore we have

$$a^i b^j a^r b^s = a^i b^{j-1}(ba)a^{r-1}b^s = a^i b^{j-1}(a^\mu b)a^{r-1}b^s = \ldots$$

Continuing substituting bas with $a^\mu b$s in this way we will end up with $a^w b^{j+s}$, where w is a very large integer, but we can use the relations $a^{h(a)} = a$ and $b^{h(b)} = b$ to reduce it to $a^i b^j a^r b^s = a^t b^u =$ with $1 \leqslant t \leqslant p^h$ and $1 \leqslant u \leqslant p^k$. Hence $a^i b^j a^r b^s = 1_p 1_p a^t b^u \in S$. Since $a^i b^j a^r b^s \in S$ and $\rho_{ij}\rho_{rs} \in P$, we have that $yz \in S$.

Proposition 2.1 now gives that S is a division ring, and therefore Wedderburn's theorem gives that it is a field. But $a, b \in S$, so that $ba = ab$, which is a contradiction since the fact that $bab^{-1} \neq a$ gives that $ba \neq ab$ if we multiply with b from the right. This contradiction proves the theorem. □

5 Alternative rings and the Artin-Zorn theorem

Now we are going to discuss a generalisation of Wedderburn's theorem to so-called Alternative rings.

Definition 5.1. *An alternative ring is a ring Ψ in which the multiplication associativity axiom is replaced by the weaker so-called alternative axiom: For all $a, b \in \Psi$, $(aa)b = a(ab)$, $(ab)a = a(ba)$ and $(ba)a = b(aa)$.*

In the rest of this section we let Ψ denote an alternative ring. Generally, when associativity does not hold, for all $a, b, c \in \Psi$ we can use the so-called associator

$$[a, b, c] = (ab)c - a(bc)$$

as some kind of "measurement" of how close it is to be associative. It is straightforward that the associator is linear in the three arguments.

Proposition 5.1. *let Ψ be an alternative ring and $a, b, c, d \in \Psi$. Then*

$$[ab, c, d] - [a, bc, d] + [a, b, cd] = [a, b, c]d + a[b, c, d].$$

Proof. Using the definition of the associator and distributivity we have

$$\begin{aligned}
[ab, c, d] - [a, bc, d] + [a, b, cd] &= \Big(\big((ab)c\big)d - (ab)(cd)\Big) \\
&\quad - \Big(\big(a(bc)\big)d - a\big((bc)d\big)\Big) + \Big((ab)(cd) - a\big(b(cd)\big)\Big) \\
&= \Big(\big((ab)c\big)d - \big(a(bc)\big)d\Big) + \Big(a\big((bc)d\big) - a\big(b(cd)\big)\Big) \\
&= \Big((ab)c - a(bc)\Big)d + a\Big((bc)d - b(cd)\Big) = [a, b, c]d + a[b, c, d].
\end{aligned}$$

□

In Ψ, it follows from the definition of the associator that the alternative axiom is equivalent to the claim that for all $a, b \in \Psi$ we have

$$[a, a, b] = [a, b, a] = [b, a, a] = 0_\Psi.$$

Proposition 5.2. *Let $a, b, c \in \Psi$. Then the associator $[a, b, c]$ alternates in the arguments, that is, if two arguments changes place, only the sign changes.*

Proof. We have

$$\begin{aligned}
0_\Psi &= [a + b, a + b, c] = [a, a, c] + [a, b, c] + [b, a, c] + [b, b, c] \\
&= [a, b, c] + [b, a, c],
\end{aligned}$$

so that $[a, b, c] = -[b, a, c]$. Starting with $0_\Psi = [a, b + c, b + c]$ instead, similar reasoning gives us that $[a, b, c] = -[a, c, b]$, and from $0_A = [a + c, b, a + c]$ we get $[a, b, c] = -[c, b, a]$, and the result follows. □

It follows from the proposition that if the associator $[a,b,c] = 0_\Psi$, then the associator $= 0_\Psi$ for any permutation of a, b and c. Hence we have that if 3 fixed elements in Ψ associate in some order, they associate in every order.

Before stating and proving the last results, we introduce some notation that will be used to simplify the discussions.

Definition 5.2. *If M_1, M_2 and M_3 are subsets of Ψ we can write*

$$[M_1, M_2, M_3] = 0_\Psi$$

if

$$[m_1, m_2, m_3] = 0_\Psi$$

whenever $m_i \in M_i$. We can also use a mixed notation where one or two arguments are fixed elements in Ψ, and the others subsets of Ψ.

Definition 5.3. *Let C and D be subsets of a ring X and let $x \in X$ and $n \in \mathbb{N}$. Then we let*

$$\begin{aligned} C + D &= \{c + d \mid c \in C, d \in D\} \\ CD &= \{cd \mid c \in C, d \in D\} \\ -C &= \{-c \mid c \in C\} \\ xC &= \{xc \mid c \in C\} \\ Cx &= \{cx \mid c \in C\} \\ nC &= C + C + ... + C \quad \textit{(n terms)} \\ (-n)C &= -C - C - ... - C \quad \textit{(n terms)} \end{aligned}$$

Definition 5.4. *If $M = \{a, b, c, ...\}$ is a subset of Ψ we let $\langle M \rangle = \langle a, b, c, ... \rangle$ be the set generated from M by taking the possible sums of plus or minus the products of the elements in M.*

Proposition 5.3. *Let M be a nonempty subset of an alternative ring Ψ. Then $\langle M \rangle$ is an alternative subring of Ψ.*

Proof. Since the elements in M are in $\langle M \rangle$, it is clearly nonempty. We have to show closure under subtraction and multiplication. It is almost obvious. A difference of two sums of plus or minus some products of elements in M is again that kind of sum, and a product of two sums of that kind is again that kind of sum. □

Definition 5.5. *A subset Γ of Ψ satisfying $[\Gamma, \Gamma, \Psi] = 0_\Psi$ is called an A-set. If Γ is a ring, we call it an A-ring.*

Theorem 5.4. *Let Γ be an A-set in Ψ. Then the ring $\langle \Gamma \rangle$ is an A-ring.*

Proof. We have to show that the ring $\langle\Gamma\rangle$ satisfies $[\langle\Gamma\rangle, \langle\Gamma\rangle, \Psi] = 0_\Psi$. We let

$$\begin{aligned} A_1 &= \Gamma, \\ A_2 &= A_1 A_1, \\ A_3 &= A_1 A_2 \cup A_2 A_1, \\ A_4 &= A_1 A_3 \cup A_3 A_1 \cup A_2 A_2, \\ A_n &= \bigcup_{i+j=n} A_i A_j. \end{aligned}$$

We now let $B_n = \{\sum_{i=1}^{k} z_i a_i \mid z_i \in \mathbb{Z}, a_i \in A_n\}$ Then $\langle\Gamma\rangle = \sum_{n=1}^{\infty} B_n$. Therefore, by linearity of the associator, it is enough to show that $[B_n, B_m, \Psi] = 0_\Psi$ for all $n, m \in \mathbb{N}\backslash\{0\}$, and by linearity again, we only have to show $[A_n, A_m, \Psi] = 0_\Psi$. We use induction over $m + n$. If $m + n = 2$, we must have $m = n = 1$, and $[A_1, A_1, \Psi] = [\Gamma, \Gamma, \Psi] = 0_\Psi$ by the definition of Γ. We now assume that $[A_j, A_k, \Psi] = 0_\Psi$ whenever $2 \leqslant j+k < m+n$ for some fixed positive integer $m+n$. We must show that then also $[A_m, A_n, \Psi] = 0_\Psi$. Permuting the arguments A_m and A_n in the associator only changes the sign, and therefore we can assume that $m \geqslant n$ which forces $m \geqslant 2$. Let $a_m \in A_m$, $a_n \in A_n$ and $x \in \Psi$. Then we can write $a_m = a'a''$ where $a' \in A_r$ and $a'' \in A_s$ for some r and s with $r+s = m$. Using Proposition 5.1 we have

$$[a'a'', x, a_n] - [a', a''x, a_n] + [a', a'', xa_n] = a'[a'', x, a_n] + [a', a'', x]a_n.$$

This implies that

$$[A_r A_s, \Psi, A_n] - [A_r, A_s\Psi, A_n] + [A_r, A_s, \Psi A_n] = A_r[A_s, \Psi, A_n] + [A_r, A_s, \Psi]A_n.$$

Since $r + n < m + n$, $r + s < m + n$ and $s + n < m + n$, our assumption gives that that all the terms except the first one vanish, which leaves us with

$$[A_r A_s, \Psi, A_n] = 0_\Psi.$$

From this we get that $[A_m, A_n, \Psi] = 0_\Psi$, which is what we wanted. □

Definition 5.6. *The opposite of a ring R is the set R' that contains the same elements of R, but with the multiplication $*$ in R' of two elements a and b is redefined as $a * b = b \cdot a$, where $\cdot$ is the multiplication in R.*

Proposition 5.5. *Let Ψ be an alternative ring. Then its opposite Ψ' is also an alternative ring.*

Proof. It is easy to show that the ring axioms still hold, since we are able to construct exactly the sums and products as in Ψ, by just interchanging the factors. We only show the alternative axiom. Let $\cdot$ be the multiplication in Ψ and $*$ the multiplication in Ψ'. For $a, b \in \Psi'$ we have

$$\begin{aligned} [a, a, b]_{\Psi'} &= (a * a) * b - a * (a * b) = b \cdot (a \cdot a) - (b \cdot a) \cdot a = -[b, a, a]_\Psi = 0_\Psi, \\ [a, b, b]_{\Psi'} &= (a * b) * b - a * (b * b) = b \cdot (b \cdot a) - (b \cdot b) \cdot a = -[b, b, a]_\Psi = 0_\Psi, \\ [a, b, a]_{\Psi'} &= (a * b) * a - a * (b * a) = a \cdot (b \cdot a) - (a \cdot b) \cdot a = -[a, b, a]_\Psi = 0_\Psi. \end{aligned}$$

□

Theorem 5.6. *Let A and B be two A-rings in Ψ. Then the ring $C = \langle A + B\rangle$ is associative.*

Proof. We have to show that when we assume that $[A, A, \Psi] = 0_\Psi = [B, B, \Psi]$, it follows that $[\langle A + B\rangle, \langle A + B\rangle, \langle A + B\rangle] = [C, C, C] = 0_\Psi$. From the last theorem we have that if $A + B$ is an A-ring, $\langle A + B\rangle$ is also an A-ring. So if we show that $[A + B, A + B, C] = 0_\Psi$, we are done. A and B are obviously subsets of the ring C, which gives that they can be seen as A-sets. The linearity of the associator and the definition of a sum of two subgroups under addition gives $[A + B, A + B, C] = [A, A, C] + [B, B, C] + [A, B, C] + [B, A, C] = [A, B, C] - [A, B, C]$. Hence it is enough to show $[A, B, C] = 0_\Psi$.

Claim 5.7. *Whenever we have $[A, B, x] = 0_\Psi$ for some fixed $x \in \Psi$, we also have*

$$[A, B, xA] = [A, B, Ax] = [A, B, xB] = [A, B, Bx] = 0_\Psi.$$

Proof. Since the associator alternates in the arguments it is clear that $[A, B, Bx] = 0_\Psi$ if we know that $[A, B, Ax] = 0_\Psi$. Then the rest follows if we consider the opposite alternative ring of Ψ. Therefore it is enough to show $[A, B, Ax] = 0_\Psi$.

If $a_1, a_2 \in A, b \in B$, Proposition 5.1 gives that

$$[a_1a_2, x, b] - [a_1, a_2x, b] + [a_1, a_2, xb] = a_1[a_2, x, b] + [a_1, a_2, x]b.$$

Proposition 5.2 and the fact that $[A, A, \Psi] = [A, B, x] = 0_\Psi$ give that the right-hand-side and the first and third term in the left-hand-side are 0_Ψ. This leaves us with $[a_1, a_2x, b] = 0_\Psi$, which means that $[A, Ax, B] = 0_\Psi$ and hence even $[A, B, Ax] = 0_\Psi$. □

We now let

$$\begin{aligned} N_1 &= A \cup B, \\ N_2 &= \big((A \cup B)N_1\big) \cup \big(N_1(A \cup B)\big), \\ N_m &= \big((A \cup B)N_{m-1}\big) \cup \big(N_{m-1}(A \cup B)\big). \end{aligned}$$

We call the elements in these sets normal products. We shall show that $[A, B, N_m] = 0_\Psi$ for all $m \in \mathbb{N}\backslash\{0\}$. We use induction over m. We have

$$[A, B, N_1] = [A, B, A \cup B] = 0_\Psi.$$

We now assume that $[A, B, N_{m-1}] = 0_\Psi$, and shall show that then also $[A, B, N_m] = 0_\Psi$. Every element in N_m can be written as an_{m-1}, $n_{m-1}a$, bn_{m-1} or $n_{m-1}b$, where $a \in A$, $b \in B$, and $n_{m-1} \in N_{m-1}$. Therefore $[A, B, N_m] = 0_\Psi$ by Claim 5.7, and the induction proof is done.

We now let $M_m = \{\sum_{i=1}^{k} a_i n_i \mid a_i \in \mathbb{Z}, n_i \in N_m\}$. The linearity of the associator gives that $[A, B, M_m] = 0_\Psi$ for all $m \in \mathbb{N}\backslash\{0\}$. It is now clear that $[A, B, M] = 0_\Psi$, where $M = \sum_{m=1}^{\infty} M_m$. We call the elements in M normal elements.

If we can show that $M = C$, the proof is done. At first we show that M is a subring of Ψ. We must show closure under subtraction and multiplication. Closure under subtraction follows immediately from the construction. To show closure under multiplication, we show that $N_m M \subseteq M$ for all $m \in \mathbb{N}\backslash\{0\}$. We use induction over m. We have that $N_1 M = (A \cup B)M$. $(A \cup B)M$ consists of sums of products $a_i N_j = N_{j+1}$, where different a_i's$\in A \cup B$. So $(A \cup B)M \subseteq M$. Now we assume that $N_{m-1}M \subseteq M$ and shall show that then also $N_m M \subseteq M$. We have that $N_m = N_{m-1}a$ or $N_m = aN_{m-1}$ for some $a \in A \cup B$. If $N_m = N_{m-1}a$ we have

$$\begin{aligned} N_m M = (N_{m-1}a)M &= N_{m-1}(aM) + [N_{m-1}, a, M] \\ &= N_{m-1}(aM) - [N_{m-1}, M, a] \\ &= N_{m-1}(aM) - (N_{m-1}M)a + N_{m-1}(Ma). \end{aligned}$$

We know that $aM \subseteq M$ and $Ma \subseteq M$. This and the assumption $N_{m-1}M \subseteq M$ give that all the three last terms are subsets of M, and hence $N_m M \subseteq M$. If instead $N_m = aN_{m-1}$ we do the same reasoning, only changing the order in the products. The distributive law now gives that $MM \subseteq M$.

Therefore M is a ring. Since M is constructed from the elements in $A \cup B$, it must be a subset of $\langle A \cup B\rangle$, but $\langle A \cup B\rangle$ is the least ring containing all the elements in $A \cup B$. This forces $M = \langle A \cup B\rangle$. It is clear that $A + B \subseteq \langle A \cup B\rangle$, and thus $\langle A + B\rangle \subseteq \langle A \cup B\rangle$. It is also clear that $A \cup B \subseteq A + B$, so that $\langle A \cup B\rangle \subseteq \langle A + B\rangle$. This forces $\langle A \cup B\rangle = \langle A + B\rangle = C$

Therefore $M = C$ so that $[A, B, C] = 0_\Psi$, and the proof is done. □

Corollary 5.8 (Artin's theorem)**.** *Let $a, b \in \Psi$. Then the ring $\langle a, b\rangle$ is associative.*

Proof. The single element sets $\{a\}$ and $\{b\}$ are clearly A-sets. Hence, by Theorem 5.4, $\langle a\rangle$ and $\langle b\rangle$ are A-rings. Theorem 5.6 now gives that the ring $\langle\langle a\rangle+\langle b\rangle\rangle$ is associative. It is clear that $\langle a, b\rangle \subseteq \langle\langle a\rangle+\langle b\rangle\rangle$, so that $\langle a, b\rangle$ is associative. □

Now we are ready to state and prove the main result of this section.

Theorem 5.9 (the Artin-Zorn theorem)**.** *Let Λ be a finite alternative ring with identity 1_Λ in which every nonzero element has a multiplicative inverse. Then Λ is a field.*

Proof. Let $a, b, c \in \Lambda$. Then the ring $\langle a, b\rangle$ is associative by Artin's theorem. Therefore it is a finite division ring, which is a field by Wedderburn's theorem. The field $F = \langle a, b\rangle$ contains a primitive element x which works as a generator of $F\backslash\{0_\Lambda\}$. It follows that $\langle a, b\rangle = \langle x\rangle$. Therefore $\langle a, b, c\rangle = \langle x, c\rangle$, so that $\langle a, b, c\rangle$ is a field. This forces $[a, b, c] = 0_\Lambda$ for all $a, b, c \in \Lambda$, which means that Λ is an associative ring. Hence it is a finite division ring, which is a field by Wedderburn's theorem! □

6 Examples

We shall give examples of infinite division rings that are not fields and infinite alternative rings that are not associative. We will use so-called Cayley-Dickson process. At first we need a definition.

Definition 6.1. *Let F be a field. Then K is a nonassociative F-algebra over F if K is a vector space over F and there is a multiplication $K \times K \to K$ with $(a, b) \mapsto ab$ which is F-bilinear, which includes that $\alpha(ab) = (\alpha a)b = a(\alpha b)$ for all $\alpha \in F$ and $a, b \in K$.*

We now let K be an nonassociative F-algebra with dimension n. An involution in K is a linear operator, denoted by overline on the elements, that satisfies $\overline{xy} = \overline{y}\ \overline{x}$ and $\overline{\overline{x}} = x$ for all $x, y \in K$. In this case we also assume that $x + \overline{x} \in F$ and $x\overline{x} \in F$ for all $x \in K$. We fix such an involution.

Now we construct an F-algebra L of dimension $2n$ as the set of ordered pairs (a, b) where $a, b \in K$, where addition is defined as $(a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2)$, and multiplication as

$$(a_1, b_1)(a_2, b_2) = (a_1a_2 + \mu b_2\overline{b}_1, \overline{a}_1b_2 + a_2b_1)$$

for all $a_1, a_2, b_1, b_2 \in K$, and where $\mu \in F\backslash\{0_F\}$ is fixed. It is easy to show that $1_L = (1_K, 0)$ and that $K' = \{(a, 0_K) \mid a \in K\}$ is isomorphic to K. The element $v = (0, 1) \in L$ satisfies $v^2 = (\mu, 0_K)$, which can be viewed as $\mu \in F$. With this defined, we can think of L as a direct sum and write $L = K' + vK'$. Identifying K' with K we can write $(a, b) \in L$ as $x = a + vb$, and then the multiplication can be written as

$$(a_1 + vb_1)(a_2 + vb_2) = (a_1a_2 + \mu b_2\overline{b}_1) + v(\overline{a}_1b_2 + a_2b_1).$$

When $x \in L$ satisfies $x = (a, b) = a + vb$ we can define $\overline{x} = \overline{a} - vb$. It is easy to show that this now will be an involution in L.

When we in this construction let $n = 1$, $K = F = \mathbb{R}$, $\mu = -1$, $v = i$ and $\overline{x} = x$ for $x \in \mathbb{R}$ it follows that we get $L = \mathbb{C}$ where the involution in $\mathbb{C}$ is the normal conjugate. Since this can be viewed as a construction of $\mathbb{C}$, we write $\mu_{\mathbb{C}} = -1$ and $v_{\mathbb{C}} = i$.

If we now let $K = \mathbb{C}$, but still $F = \mathbb{R}$, we will get $L = \mathbb{H}$ (The quaternions defined in Section 24 in [2]) by this construction, when we let $\mu_{\mathbb{H}} = -1$, $v_{\mathbb{H}} = j$ and $v_{\mathbb{H}}(-i) = j(-i) = k$. We have that $ij = iv_{\mathbb{H}} = (i + v_{\mathbb{H}}0)(0 + v_{\mathbb{H}}1) = v(-i) = k$. We also have that $ji = v_{\mathbb{H}}i = -v_{\mathbb{H}}(-i) = -k$. By similar reasoning we get all the relations in the definition of $\mathbb{H}$.

Since $ij = k \neq -k = ji$, we have that $\mathbb{H}$ is not commutative. It is shown in Section 24 in [2] that $\mathbb{H}$ is a division ring. Therefore the quaternions is an example of a division ring that is not a field. Therefore Wedderburn's theorem is only valid in the finite case, and it really makes sense to talk about fields and division rings as different mathematical structures in general.

If we now let $K = \mathbb{H}$, the construction gives us $L = \mathbb{O}$, the so-called octonions or Cayley numbers.

On page 44 in [8] there is a proof of the next proposition.

Proposition 6.1. *Let K and L be the F-algebras defined as before. Then K is associative if and only if L is alternative.*

Since $\mathbb{H}$ is associative, Proposition 6.1 gives that $\mathbb{O}$ is at least alternative.

Proposition 6.2. *The octonions $\mathbb{O}$ is not associative.*

Proof. Since $\mathbb{H}$ is not commutative, there exist some $q_1, q_2 \in \mathbb{H}$ so that $q_1q_2 - q_2q_1 \neq 0$. Therefore

$$\begin{aligned}
[v_{\mathbb{O}}, q_2, q_1] = (v_{\mathbb{O}}q_2)q_1 - v_{\mathbb{O}}(q_2q_1) &= \big(0 + v_{\mathbb{O}}1)(q_2 + v_{\mathbb{O}}0)\big)(q_1 + v_{\mathbb{O}}0) - v_{\mathbb{O}}(q_2q_1) \\
&= (0 + q_2v_{\mathbb{O}})(q_1 + v_{\mathbb{O}}0) - v_{\mathbb{O}}(q_2q_1) \\
&= v_{\mathbb{O}}(q_1q_2) - v_{\mathbb{O}}(q_2q_1) \\
&= v_{\mathbb{O}}(q_1q_2 - q_2q_1) \\
&\neq 0.
\end{aligned}$$

□

This gives that there exist infinite alternative rings that are not associative, so that it surely makes sense to distinguish between alternative rings and normal rings in general.

References

[1] D. M. Burton, *Elementary Number Theory*, 7th ed, McGraw-Hill, 2011.

[2] J. B. Fraleigh, *A First Course In Abstract Algebra*, 7th ed., Addison Wesley, 2003.

[3] I. N. Herstein, *Topics in Algebra*, Xerox College Publishing, 1964.

[4] L. Howe, *The Class Equation*, `http://www.rowan.edu/open/depts/math/howe/BookChapters/The%20Class%20Equation.pdf`, 2000.

[5] T. W. Hungerford, *Abstract Algebra, An Introduction*, Brooks/Cole 3 ed., 2014.

[6] `http://math.stackexchange.com/questions/751908/proof-that-p-pn-choose-k-for-any-prime-p-and-k-pn`, Mathematics Stack Exchange, 2014.

[7] MIT OpenCourseWare, `http://ocw.mit.edu/courses/mathematics/18-781-theory-of-numbers-spring-2012/lecture-notes/MIT18_781S12_lec12.pdf`, 18.781 Theory of Numbers, 2012.

[8] R. D. Schafer, *An Introduction to Nonassociative Algebras*, Academic Press, 1966.

[9] M. Zorn, *Theorie der alternativen Ringe*, Abh. Math. Sem. Univ. Hamburg **8** no. 1, 1931, 123-147.