CHAPTER 7

Selected Topics

In this final chapter we have set ourselves two objectives. Our first is to present some mathematical results which cut deeper than most of the material up to now, results which are more sophisticated, and are a little apart from the general development which we have followed. Our second goal is to pick results of this kind whose discussion, in addition, makes vital use of a large cross section of the ideas and theorems expounded earlier in the book. To this end we have decided on three items to serve as the focal points of this chapter.

The first of these is a celebrated theorem proved by Wedderburn in 1905 ("A Theorem on Finite Algebras," *Transactions of the American Mathematical Society*, Vol. 6 (1905), pages 349–352) which asserts that a division ring which has only a finite number of elements must be a commutative field. We shall give two proofs of this theorem, differing totally from each other. The first one will closely follow Wedderburn's original proof and will use a counting argument; it will lean heavily on results we developed in the chapter on group-theory. The second one will use a mixture of group-theoretic and field-theoretic arguments, and will draw incisively on the material we developed in both these directions. The second proof has the distinct advantage that in the course of executing the proof certain side-results will fall out which will enable us to proceed to the proof, in the division ring case, of a beautiful theorem due to Jacobson ("Structure Theory for Algebraic Algebras of Bounded Degree," *Annals of Mathematics*, Vol. 46 (1945), pages 695–707) which is a far-reaching generalization of Wedderburn's theorem.

Our second high-spot is a theorem due to Frobenius ("Über lineare Substitutionen und bilineare Formen," *Journal für die Reine und Angewandte Mathematik*, Vol. 84 (1877), especially pages 59–63) which states that the only division rings algebraic over the field of all real numbers are the field of real numbers, the field of complex numbers, and the division ring of real quaternions. The theorem points out a unique role for the quaternions, and makes it somewhat amazing that Hamilton should have discovered them in his somewhat ad hoc manner. Our proof of the Frobenius theorem, now quite elementary, is a variation of an approach laid out by Dickson and Albert; it will involve the theory of polynomials and fields.

Our third goal is the theorem that every positive integer can be represented as the sum of four squares. This famous result apparently was first conjectured by the early Greek mathematician Diophantos. Fermat grappled unsuccessfully with it and sadly announced his failure to solve it (in a paper where he did, however, solve the two-square theorem which we proved in Section 8 of Chapter 3). Euler made substantial inroads on the problem; basing his work on that of Euler, Lagrange in 1770 finally gave the first complete proof. Our approach will be entirely different from that of Lagrange. It is rooted in the work of Adolf Hurwitz and will involve a generalization of Euclidean rings. Using our ring theoretic techniques on a certain ring of quaternions, the Lagrange theorem will drop out as a consequence.

En route to establishing these theorems many ideas and results, interesting in their own right, will crop up. This is characteristic of a good theorem—its proof invariably leads to side-results of almost equal interest.

1. Finite Fields. Before we can enter into a discussion of Wedderburn's theorem and finite division rings it is essential that we investigate the nature of fields having only a finite number of elements. Such fields are called *finite fields*. Finite fields do exist, for the ring J_p of integers modulo any prime p, provides us with an example of such. In this section we shall determine all possible finite fields and many of the important properties which they possess.

We begin with

LEMMA 7.1. *Let F be a finite field with q elements and suppose that $F \subset K$ where K is also a finite field. Then K has q^n elements where $n = [K:F]$.*

Proof. K is a vector space over F and since K is finite it is certainly finite-dimensional as a vector space over F. Suppose that $[K:F] = n$; then K has a basis of n elements over F. Let such a basis be $v_1, v_2, \ldots, v_n$. Then every element in K has a unique representation in the form $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$ where $\alpha_1, \alpha_2, \ldots, \alpha_n$ are all in F. Thus the number of elements in K is the number of $\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_n v_n$ as the $\alpha_1, \alpha_2, \ldots, \alpha_n$ range over F. Since each coefficient can have q values K must clearly have q^n elements.

COROLLARY 1. *Let F be a finite field; then F has p^m elements where the prime number p is the characteristic of F.*

Proof. Since F has a finite number of elements, by Corollary 2 to Theorem 2.a, $f1 = 0$ where f is the number of elements in F. Thus F has characteristic p for some prime number p. Therefore F contains a field F_0 isomorphic to J_p. Since F_0 has p elements, F has p^m elements where $m = [F:F_0]$, by Lemma 7.1.

COROLLARY 2. *If the finite field F has p^m elements then every $a \in F$ satisfies $a^{p^m} = a$.*

Proof. If $a = 0$ the assertion of the corollary is trivially true.

On the other hand, the nonzero elements of F form a group under multiplication of order $p^m - 1$ thus by Corollary 2 to Theorem 2.a, $a^{p^m-1} = 1$ for all $a \neq 0$ in F. Multiplying this relation by a we obtain that $a^{p^m} = a$.

From this last corollary we can easily pass to

LEMMA 7.2. *If the finite field F has p^m elements then the polynomial $x^{p^m} - x$ in $F[x]$ factors in $F[x]$ as $x^{p^m} - x = \prod_{\lambda \in F} (x - \lambda)$.*

Proof. By Lemma 5.2 the polynomial $x^{p^m} - x$ has at most p^m roots in F. However, by Corollary 2 to Lemma 7.1 we know p^m such roots, namely all the elements of F. By the corollary to Lemma 5.1 we can conclude that $x^{p^m} - x = \prod_{\lambda \in F} (x - \lambda)$.

COROLLARY. *If the field F has p^m elements then F is the splitting field of the polynomial $x^{p^m} - x$.*

Proof. By Lemma 7.2, $x^{p^m} - x$ certainly splits in F. However, it cannot split in any smaller field for that field would have to have all the roots of this polynomial and so would have to have at least p^m elements. Thus F is the splitting field of $x^{p^m} - x$.

As we have seen in Chapter 5 (Theorem 5.j) any two splitting fields over a given field of a given polynomial are isomorphic. In light of the corollary to Lemma 7.2 we can state

LEMMA 7.3. *Any two finite fields having the same number of elements are isomorphic.*

Proof. If these fields have p^m elements, by the above corollary they are both splitting fields of the polynomial $x^{p^m} - x$, over J_p whence they are isomorphic.

Thus for any integer m and any prime number p there is, up to isomorphism, at most one field having p^m elements. The purpose of the next lemma is to demonstrate that for any prime number p and any integer m there is a field having p^m elements. When this is done we shall know that there is exactly one field having p^m elements where p is an arbitrary prime and m an arbitrary integer.

Lemma 7.4. *For every prime number p and every positive integer m there exists a field having p^m elements.*

Proof. Consider the polynomial $x^{p^m} - x$ in $J_p[x]$, the ring of polynomials in x over J_p, the field of integers mod p. Let K be the splitting field of this polynomial. In K let $F = \{a \in K \mid a^{p^m} = a\}$. The elements of F are thus the roots of $x^{p^m} - x$, which by Corollary 2 to Lemma 5.6 are distinct whence F has p^m elements. We now claim that F is a field. If $a, b \in F$ then $a^{p^m} = a$, $b^{p^m} = b$ and so $(ab)^{p^m} = a^{p^m}b^{p^m} = ab$; thus $ab \in F$. Also since the characteristic is p, $(a \pm b)^{p^m} = a^{p^m} \pm b^{p^m} = a \pm b$, whence $a \pm b \in F$. Consequently F is a subfield of K and so is a field. Having exhibited the field F having p^m elements we have proved Lemma 7.4.

Combining Lemmas 7.3 and 7.4 we have

Theorem 7.a. *For every prime number p and every positive integer m there is a unique field having p^m elements.*

We now return to group theory for a moment. The group-theoretic result we seek will determine the structure of any finite multiplicative subgroup of the group of nonzero elements of any field, and, in particular, it will determine the multiplicative structure of any finite field.

Lemma 7.5. *Let G be a finite abelian group enjoying the property that the relation $x^n = e$ is satisfied by at most n elements of G, for every integer n. Then G is a cyclic group.*

Proof. If the order of G is a power of some prime number q then the result is very easy. For suppose that $a \in G$ is an element whose order is as large as possible; its order must be q^r for some integer r. The elements $e, a, a^2, \ldots, a^{q^r-1}$ give us q^r distinct solutions of the equation $x^{q^r} = e$, which, by our hypothesis, implies that these are all the solutions of this equation. Now if $b \in G$ its order is q^s where $s \leq r$, hence $b^{q^r} = (b^{q^s})^{q^{r-s}} = e$. By the observation made above this forces $b = a^i$ for some i, and so G is cyclic.

The general finite abelian group G can be realized as $G = S_{q_1}S_{q_2}, \ldots, S_{q_k}$ where the q_i are the distinct prime divisors of $o(G)$ and where the S_{q_i} are the Sylow subgroups of G. Moreover, every element $g \in G$ can be written in a *unique* way as $g = s_1s_2, \ldots, s_k$ where $s_i \in S_{q_i}$ (see Section 7, Chapter 2). Any solution of $x^n = e$ in S_{q_i} is one of $x^n = e$ in G so that each S_{q_i} inherits the hypothesis we have imposed on G. By the remarks of the first paragraph of the proof each S_{q_i} is a cyclic group; let a_i be a generator of S_{q_i}. We claim that $c = a_1a_2, \ldots, a_k$ is a cyclic generator of G. To verify this all we must do is prove that $o(G)$ divides m, the order of c. Since $c^m = e$, we have that $a_1{}^m a_2{}^m, \ldots, a_k{}^m = e$. By the uniqueness of representation of an element of G as a product of elements in the S_{q_i}, we conclude that each $a_i^m = e$. Thus $o(S_{q_i}) \mid m$ for every i. Thus $o(G) = o(S_{q_1})o(S_{q_2}) \ldots o(S_{q_k}) \mid m$. However, $m \mid o(G)$ and so $o(G) = m$. This proves that G is cyclic.

Lemma 7.5 has as an important consequence

LEMMA 7.6. *Let K be a field and let G be a finite subgroup of the multiplicative group of nonzero elements of K. Then G is a cyclic group.*

Proof. Since K is a field, any polynomial of degree n in $K[x]$ has at most n roots in K. Thus in particular, for any integer n, the polynomial $x^n - 1$ has at most n roots in K, and all the more so, at most n roots in G. The hypothesis of Lemma 7.5 is satisfied, so G is cyclic.

Even though the situation of a finite field is merely a special case of Lemma 7.6, it is of such wide-spread interest that we single it out as

THEOREM 7.B. *The multiplicative group of nonzero elements of a finite field is cyclic.*

Proof. Let F be a finite field. By merely applying Lemma 7.6 with $F = K$ and $G =$ the group of nonzero elements of F, the result drops out.

We conclude this section by using a counting argument to prove the existence of solutions of certain equations in a finite field. We shall need the result in one proof of the Wedderburn theorem.

LEMMA 7.7. *If F is a finite field and $\alpha \neq 0$, $\beta \neq 0$ are two elements of F then we can find elements a and b in F such that $1 + \alpha a^2 + \beta b^2 = 0$.*

Proof. If the characteristic of F is 2, F has 2^n elements and every element x in F satisfies $x^{2^n} = x$. Thus every element in F is a square. In particular $\alpha^{-1} = a^2$ for some $a \in F$. Using this a and $b = 0$ we have $1 + \alpha a^2 + \beta b^2 = 1 + \alpha\alpha^{-1} + 0 = 1 + 1 = 0$, the last equality being a consequence of the fact that the characteristic of F is 2.

If the characteristic of F is an odd prime p, F has p^n elements. Let $W_\alpha = \{1 + \alpha x^2 \mid x \in F\}$. How many elements are there in W_α? We must check how often $1 + \alpha x^2 = 1 + \alpha y^2$. But this relation forces $\alpha x^2 = \alpha y^2$ and so, since $\alpha \neq 0$, $x^2 = y^2$. Finally this leads to $x = \pm y$. Thus for $x \neq 0$ we get from each pair x and $-x$ one element in W_α, and for $x = 0$ we get $1 \in W_\alpha$. Thus W_α has $1 + \dfrac{p^n - 1}{2} = \dfrac{p^n + 1}{2}$ elements. Similarly $W_\beta = \{-\beta x^2 \mid x \in F\}$ has $\dfrac{p^n + 1}{2}$ elements. Since each of W_α and W_β has more than half the elements of F they must have a nonempty intersection. Let $c \in W_\alpha \cap W_\beta$. Since $c \in W_\alpha$, $c = 1 + \alpha a^2$ for some $a \in F$; since $c \in W_\beta$, $c = -\beta b^2$ for some $b \in F$. Therefore $1 + \alpha a^2 = -\beta b^2$, which, on transposing yields the desired result $1 + \alpha a^2 + \beta b^2 = 0$.

PROBLEMS

1. By Theorem 7.b the nonzero elements of J_p form a cyclic group under multiplication. Any generator of this group is called a *primitive root* of p.

(a) Find primitive roots of: 17, 23, 31.

(b) How many primitive roots does a prime p have?

2. Using Theorem 7.b prove that $x^2 \equiv -1 \bmod p$ is solvable if and only if the prime p is of the form $4n + 1$.

3. If a is an integer not divisible by the odd prime p, prove that $x^2 \equiv a \bmod p$ is solvable for some integer x if and only if $a^{(p-1)/2} \equiv 1 \bmod p$. (This is called the *Euler criterion* that a be a quadratic residue mod p.)

4. Using the result of Problem 3 determine if:

(a) 3 is a square mod 17.

(b) 10 is a square mod 13.

5. If the field F has p^n elements prove that the automorphisms of F form a cyclic group of order n.

6. If F is a finite field, by the quaternions over F we shall mean the set of all $\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$ where $\alpha_0, \alpha_1, \alpha_2, \alpha_3 \in F$ and where addition and multiplication are carried out as in the real quaternions (i.e., $i^2 = j^2 = k^2 = ijk = -1$, etc.). Prove that the quaternions over a finite field *do not* form a division ring.

2. Wedderburn's Theorem on Finite Division Rings. In 1905 Wedderburn proved the theorem, now considered a classic, that a finite division ring must be a commutative field. This result has caught the imagination of most mathematicians because it is so unexpected, interrelating two seemingly unrelated things, namely the number of elements in a certain algebraic system and the multiplication of that system. Aside from its intrinsic beauty the result has been very important and useful since it arises in so many contexts. To cite just one instance, the only known proof of the purely geometric fact that in a finite geometry the Desargues configuration implies that of Pappus (for the definition of these terms look in any good book on projective geometry) is to reduce the geometric problem to an algebraic one, and this algebraic question is then answered by invoking the Wedderburn theorem. For algebraists the Wedderburn theorem has served as a jumping-off point for a large area of research, in the 1940's and 1950's, concerned with the commutativity of rings.

THEOREM 7.C (WEDDERBURN). *A finite division ring is necessarily a commutative field.*

First Proof. Let K be a finite division ring and let $Z = \{z \in K \mid zx = xz$ for all $x \in K\}$ be its center. If Z has q elements then, as in the proof of Lemma 7.1, it follows that K has q^n elements. Our aim is to prove that $Z = K$, or, equivalently, that $n = 1$.

If $a \in K$ let $N(a) = \{x \in K \mid xa = ax\}$. $N(a)$ clearly contains Z, and, as a simple check reveals, $N(a)$ is a subdivision ring of K. Thus $N(a)$ contains $q^{n(a)}$ elements for some integer $n(a)$. We claim that $n(a) \mid n$. For, the nonzero elements of $N(a)$ form a subgroup of order $q^{n(a)} - 1$ of the group of nonzero elements, under multiplication, of K which has $q^n - 1$ elements. By Lagrange's theorem (Theorem 2.a) $q^{n(a)} - 1$ is a divisor of $q^n - 1$; but this forces $n(a)$ to be a divisor of n (see Problem 1 at the end of this section).

In the group of nonzero elements of K we have the conjugacy relation used in Chapter 2, namely a is a conjugate of b if $a = x^{-1}bx$ for some $x \neq 0$ in K.

By Theorem 2.h the number of elements in K conjugate to a is the index of the normalizer of a in the group of nonzero elements of K. Therefore the number of conjugates of a in K is $\dfrac{q^n - 1}{q^{n(a)} - 1}$. Now $a \in Z$ if and only if $n(a) = n$, thus by the class equation (see the corollary to Theorem 2.h)

$$q^n - 1 = q - 1 + \sum_{\substack{n(a) \mid n \\ n(a) \neq n}} \frac{q^n - 1}{q^{n(a)} - 1} \tag{1}$$

where the sum is carried out over one a in each conjugate class for a's *not* in the center.

The problem has been reduced to proving that no equation such as (1) can hold in the integers. Up to this point we have followed the proof in Wedderburn's original paper quite closely. He went on to rule out the possibility of equation (1) by making use of the following number-theoretic result due to Birkhoff and Vandiver: for $n > 1$ there exists a prime number which is a divisor of $q^n - 1$ but is not a divisor of *any* $q^m - 1$ where m is a proper divisor of n, with the exceptions of $2^6 - 1 = 63$ whose prime factors already occur as divisors of $2^2 - 1$ and $2^3 - 1$, and $n = 2$, and p a prime of the form $2^k - 1$. If we grant this result, how would we finish the proof? This prime number would be a divisor of the left-hand side of (1) and also a divisor of each term in the sum occurring on the right-hand side since it divides $q^n - 1$ but not $q^{n(a)} - 1$; thus this prime would then divide $q - 1$ giving us a contradiction. The case $2^6 - 1$ still would need ruling out but that is simple. In case $n = 2$, the other possibility not covered by the above argument, there can be no subfield between Z and K and this forces $Z = K$. (Prove!—See Problem 2.)

However, we do not want to invoke the result of Birkhoff and Vandiver without proving it, and its proof would be too large a digression here. So we look for another artifice. Our aim is to find an integer which divides $\dfrac{q^n - 1}{q^{n(a)} - 1}$, for all divisors $n(a)$ of n except $n(a) = n$, but does not divide $q - 1$. Once this is done, Equation (1) will be impossible unless $n = 1$ and, therefore, Wedderburn's theorem will have been proved. The means to this end is the theory of cyclotomic polynomials. (These have been mentioned in the problems at the end of Section 6, Chapter 5.)

Consider the polynomial $x^n - 1$ considered as an element of $C[x]$ where C is the field of complex numbers. In $C[x]$

$$x^n - 1 = \Pi(x - \lambda), \tag{2}$$

where this product is taken over all λ satisfying $\lambda^n = 1$.

A complex number θ is said to be a *primitive nth root of unity* if $\theta^n = 1$ but $\theta^m \neq 1$ for any positive integer $m < n$. The complex numbers satisfying $x^n = 1$ form a finite subgroup, under multiplication, of the complex numbers, so by Theorem 7.b this group is cyclic. Any cyclic generator of this group must then be a primitive nth root of unity, so we know that such primitive roots exist. (Alternately, $\theta = e^{2\pi i/n}$ yields us a primitive nth root of unity.)

Let $\Phi_n(x) = \Pi(x - \theta)$ where this product is taken over all the primitive nth roots of unity. This polynomial is called a *cyclotomic* polynomial. We list the first few cyclotomic polynomials: $\Phi_1(x) = x - 1$, $\Phi_2(x) = x + 1$, $\Phi_3(x) = x^2 + x + 1$, $\Phi_4(x) = x^2 + 1$, $\Phi_5(x) = x^4 + x^3 + x^2 + x + 1$, $\Phi_6(x) = x^2 - x + 1$. Notice that these are all monic polynomials with integer coefficients.

Our first aim is to prove that in general $\Phi_n(x)$ is a monic polynomial with integer coefficients. We regroup the factored form of $x^n - 1$ as given in (2), and obtain

$$x^n - 1 = \prod_{d|n} \Phi_d(x). \tag{3}$$

By induction we assume that $\Phi_d(x)$ is a monic polynomial with integer coefficients for $d|n$, $d \neq n$. Thus $x^n - 1 = \Phi_n(x)g(x)$ where $g(x)$ is a monic polynomial with integer coefficients. Therefore,

$$\Phi_n(x) = \frac{x^n - 1}{g(x)},$$

which, on actual division (or by comparing coefficients), tells us that $\Phi_n(x)$ is a monic polynomial with integer coefficients.

We now claim that for any divisor d of n, where $d \neq n$,

$$\Phi_n(x) \,\Big|\, \frac{x^n - 1}{x^d - 1}$$

in the sense that the quotient is a polynomial with integer coefficients. To see this first note that $x^d - 1 = \prod_{k \mid d} \Phi_k(x)$, and since every divisor of d is also a divisor of n, by regrouping terms on the right-hand side of (3) we obtain $x^d - 1$ on the right-hand side; also since $d < n$, $x^d - 1$ does not involve $\Phi_n(x)$. Therefore, $x^n - 1 = \Phi_n(x)(x^d - 1)f(x)$ where $f(x) = \prod_{\substack{k \mid n \\ k \nmid d}} \Phi_k(x)$ has integer coefficients, and so

$$\Phi_n(x) \,\Big|\, \frac{x^n - 1}{x^d - 1}$$

in the sense that the quotient is a polynomial with integer coefficients. This establishes our claim.

For any integer t, $\Phi_n(t)$ is an integer and from the above as an integer divides $(t^n - 1)/(t^d - 1)$. In particular, returning to equation (1),

$$\Phi_n(q) \,\Big|\, \frac{q^n - 1}{q^{n(a)} - 1}$$

and $\Phi_n(q) \mid (q^n - 1)$; thus by (1), $\Phi_n(q) \mid (q - 1)$. We claim, however, that if $n > 1$ then $|\Phi_n(q)| > q - 1$. For $\Phi_n(q) = \Pi(q - \theta)$ where θ runs over all primitive nth roots of unity and $|q - \theta| > q - 1$ for all $\theta \neq 1$ a root of unity (Prove!) whence $|\Phi_n(q)| = \Pi|q - \theta| > q - 1$. Clearly, then $\Phi_n(q)$ cannot divide $q - 1$, leading us to a contradiction. We must, therefore, assume that $n = 1$, forcing the truth of the Wedderburn theorem.

Second Proof. Before explicitly examining finite division rings again, we prove some preliminary lemmas.

LEMMA 7.8. *Let R be a ring and let $a \in R$. Let T_a be the mapping of R into itself defined by $xT_a = xa - ax$. Then*

$$xT_a{}^m = xa^m - maxa^{m-1} + \frac{m(m-1)}{2}a^2xa^{m-2} - \frac{m(m-1)(m-2)}{3!}a^3xa^{m-3} + \cdots.$$

Proof. What is $xT_a{}^2$? $xT_a{}^2 = (xT_a)T_a = (xa - ax)T_a = (xa - ax)a - a(xa - ax) = xa^2 - 2axa + a^2x$. What about $xT_a{}^3$? $xT_a{}^3 = (xT_a{}^2)T_a = (xa^2 - 2axa + a^2x)a - a(xa^2 - 2axa + a^2x) = xa^3 - 3axa^2 + 3a^2xa - a^3x$. Continuing in this way, or by the use of induction, we get the result of Lemma 7.8.

COROLLARY. *If R is a ring in which $px = 0$ for all $x \in R$, where p is a prime number, then $xT_a^{p^m} = xa^{p^m} - a^{p^m}x$.*

Proof. By the formula of Lemma 7.8, if $p = 2$, $xT_a^2 = xa^2 - a^2x$, since $2axa = 0$. Thus, $xT_a^4 = (xa^2 - a^2x)a^2 - a^2(xa^2 - a^2x) = xa^4 - a^4x$, and so on for $xT_a^{2^m}$.

If p is an odd prime, again by the formula of Lemma 7.8,

$$xT_a^p = xa^p - paxa^{p-1} + \frac{p(p-1)}{2}a^2xa^{p-2} + \cdots - a^px,$$

and since

$$p \left| \frac{p(p-1)\ldots(p-i+1)}{i!} \right.$$

for $i < p$, all the middle terms drop out and we are left with $xT_a^p = xa^p - a^px = xT_{a^p}$. Now $xT_a^{p^2} = x(T_{a^p})^p = xT_{a^{p^2}}$, and so on for the higher powers of p.

LEMMA 7.9. *Let D be a division ring of characteristic $p > 0$ with center Z, and let $P = \{0, 1, 2, \ldots, (p-1)\}$ be the subfield of Z isomorphic to J_p. Suppose that $a \in D$, $a \notin Z$ is such that $a^{p^n} = a$ for some $n > 1$. Then there exists an $x \in D$ such that*

(1) $xax^{-1} \neq a$.

(2) $xax^{-1} \in P(a)$ the field obtained by adjoining a to P.

Proof. Define the mapping T_a of D into itself by $yT_a = ya - ay$ for every $y \in D$.

$P(a)$ is a finite field, since a is algebraic over P and has, say, p^m elements. These all satisfy $u^{p^m} = u$. By the corollary to Lemma 7.8, $yT_a^{p^m} = ya^{p^m} - a^{p^m}y = ya - ay = yT_a$, and so $T_a^{p^m} = T_a$.

Now, if $\lambda \in P(a)$, $(\lambda x)T_a = (\lambda x)a - a(\lambda x) = \lambda xa - \lambda ax = \lambda(xa - ax) = \lambda(xT_a)$, since λ commutes with a. Thus the mapping λI of D into itself defined by $\lambda I : y \rightarrow \lambda y$ commutes with T_a for every $\lambda \in P(a)$. Now the polynomial $u^{p^m} - u = \prod_{\lambda \in P(a)} (u - \lambda)$ by Lemma 7.2. *Since T_a commutes with λI for every $\lambda \in P(a)$*, and since $T_a^{p^m} = T_a$, we have that $0 = T_a^{p^m} - T_a = \prod_{\lambda \in P(a)} (T_a - \lambda I)$.

If for every $\lambda \neq 0$ in $P(a)$, $T_a - \lambda I$ annihilates no nonzero element in D (if $y(T_a - \lambda I) = 0$ implies $y = 0$), since $T_a(T_a - \lambda_1 I) \ldots (T_a - \lambda_k I) = 0$, where $\lambda_1, \ldots, \lambda_k$ are the nonzero elements of $P(a)$, we would get $T_a = 0$. That is, $0 = yT_a = ya - ay$ for every $y \in D$ forcing $a \in Z$ contrary to hypothesis. Thus there is a $\lambda \neq 0$ in $P(a)$ and an $x \neq 0$ in D such that $x(T_a - \lambda I) = 0$. Writing this out explicitly, $xa - ax - \lambda x = 0$; hence,

$xax^{-1} = a + \lambda$ is in $P(a)$ and is not equal to a since $\lambda \neq 0$. This proves the lemma.

COROLLARY. *In Lemma* 7.9, $xax^{-1} = a^i \neq a$ *for some integer* i.

Proof. Let a be of order s; then in the field $P(a)$ all the roots of the polynomial $u^s - 1$ are $1, a, a^2, \ldots, a^{s-1}$ since these are all distinct roots and they are s in number. Since $(xax^{-1})^s = xa^sx^{-1} = 1$, and since $xax^{-1} \in P(a)$, xax^{-1} is a root in $P(a)$ of $u^s - 1$, whence $xax^{-1} = a^i$.

We now have all the pieces that we need to carry out our second proof of Wedderburn's theorem.

Let D be a finite division ring and let Z be its center. By induction we may assume that any division ring having fewer elements than D is a commutative field.

We first remark that if $a, b \in D$ are such that $b^ta = ab^t$ but $ba \neq ab$ then $b^t \in Z$. For, consider $N(b^t) = \{x \in D \mid b^tx = xb^t\}$. $N(b^t)$ is a subdivision ring of D; if it were not D, by our induction hypothesis, it would be commutative. However, both a and b are in $N(b^t)$ and these do not commute; consequently, $N(b^t)$ is not commutative so must be all of D. Thus $b^t \in Z$.

Every nonzero element in D has finite order, so some positive power of it falls in Z. Given $w \in D$ let the *order of w relative to Z* be the smallest positive integer $m(w)$ such that $w^{m(w)} \in Z$. Pick an element a in D but not in Z having minimal possible order relative to Z, and let this order be r. *We claim that r is a prime number* for if $r = r_1r_2$ with $1 < r_1 < r$ then a^{r_1} is not in Z. Yet $(a^{r_1})^{r_2} = a^r \in Z$, implying that a^{r_1} has an order relative to Z smaller than that of a.

By the corollary to Lemma 7.9 there is an $x \in D$ such that $xax^{-1} = a^i \neq a$; thus $x^2ax^{-2} = x(xax^{-1})x^{-1} = xa^ix^{-1} = (xax^{-1})^i = (a^i)^i = a^{i^2}$. Similarly, we get $x^{r-1}ax^{-(r-1)} = a^{i^{r-1}}$. However, r is a prime number thus by the little Fermat theorem (corollary to Theorem 2.a), $i^{r-1} = 1 + u_0r$, hence $a^{i^{r-1}} = a^{1+u_0r} = aa^{u_0r} = \lambda a$ where $\lambda = a^{u_0r} \in Z$. Thus $x^{r-1}a = \lambda azx^{r-1}$. Since $x \not\in Z$, by the minimal nature of r, x^{r-1} cannot be in Z. By the remark of the earlier paragraph since $xa \neq ax$, $x^{r-1}a \neq ax^{r-1}$ and so $\lambda \neq 1$. Let $b = x^{r-1}$; thus $bab^{-1} = \lambda a$; consequently, $\lambda^ra^r = (bab^{-1})^r = ba^rb^{-1} = a^r$ since $a^r \in Z$. This relation forces $\lambda^r = 1$.

We claim that if $y \in D$ then whenever $y^r = 1$, then $y = \lambda^i$ for some i, for in the *field* $Z(y)$ there are at most r roots of the polynomial $u^r - 1$; the elements $1, \lambda, \lambda^2, \ldots, \lambda^{r-1}$ in Z are all distinct since λ is of the prime order r and they already account for r roots of $u^r - 1$ in $Z(y)$, in consequence of which $y = \lambda^i$.

Since $\lambda^r = 1$, $b^r = \lambda^rb^r = (\lambda b)^r = (a^{-1}ba)^r = a^{-1}b^ra$ from which we get $ab^r = b^ra$. Since a commutes with b^r but does not commute with b, by the remark made earlier, b^r must be in Z. By Theorem 7.b the multiplicative group of nonzero elements of Z is cyclic; let $\gamma \in Z$ by a generator. Thus $a^r = \gamma^j$, $b^r = \gamma^k$; if $j = sr$ then $a^r = \gamma^{sr}$, whence $(a/\gamma^s)^r = 1$; this would imply that $a/\gamma^s = \lambda^i$, leading to $a \in Z$, contrary to $a \notin Z$. Hence, $r \nmid j$; similarly $r \nmid k$. Let $a_1 = a^k$ and $b_1 = b^j$; a direct computation from $ba = \lambda ab$ leads to $a_1b_1 = \mu b_1a_1$ where $\mu = \lambda^{-jk} \in Z$. Since the prime number r which is the order of λ does not divide j or k, $\lambda^{jk} \neq 1$ whence $\mu \neq 1$. Note that $\mu^r = 1$.

Let us see where we are. We have produced two elements a_1, b_1 such that:

(1) $a_1{}^r = b_1{}^r = \alpha \in Z$.
(2) $a_1b_1 = \mu b_1a_1$ with $\mu \neq 1$ in Z.
(3) $\mu^r = 1$.

We compute $(a_1{}^{-1}b_1)^r$; $(a_1{}^{-1}b_1)^2 = a_1{}^{-1}b_1a_1{}^{-1}b_1 = a_1{}^{-1}(b_1a_1{}^{-1})b_1 = a_1{}^{-1}(\mu a_1{}^{-1}b_1)b_1 = \mu a_1{}^{-2}b_1{}^2$. If we compute $(a_1{}^{-1}b_1)^3$ we find it equal to $\mu^{1+2}a_1{}^{-3}b_1{}^3$. Continuing we obtain $(a_1{}^{-1}b_1)^r = \mu^{1+2+\cdots+(r-1)}a_1{}^{-r}b_1{}^r = \mu^{1+2+\cdots+(r-1)} = \mu^{r(r-1)/2}$. If r is an odd prime, since $\mu^r = 1$, we get $\mu^{r(r-1)/2} = 1$, whence $(a_1{}^{-1}b_1)^r = 1$. Being a solution of $y^r = 1$, $a_1{}^{-1}b_1 = \lambda^i$ so that $b_1 = \lambda^i a_1$; but then $\mu b_1a_1 = a_1b_1 = b_1a_1$, contradicting $\mu \neq 1$. Thus if r is an odd prime number, the theorem is proved.

We must now rule out the case $r = 2$. In that special situation we have two elements $a_1, b_1 \in D$ such that $a_1{}^2 = b_1{}^2 = \alpha \in Z$, $a_1b_1 = \mu b_1a_1$ where $\mu^2 = 1$ and $\mu \neq 1$. Thus $\mu = -1$ and $a_1b_1 = -b_1a_1 \neq b_1a_1$; in consequence, the characteristic of D is *not* 2. By Lemma 7.7 we can find elements $\zeta, \eta \in Z$ such that $1 + \zeta^2 - \alpha\eta^2 = 0$. Consider $(a_1 + \zeta b_1 + \eta a_1b_1)^2$; on computing this out we find that $(a_1 + \zeta b_1 + \eta a_1b_1)^2 = \alpha(1 + \zeta^2 - \alpha\eta^2) = 0$. Being in a division ring this yields that $a_1 + \zeta b_1 + \eta a_1b_1 = 0$; thus $0 \neq 2a_1{}^2 = a_1(a_1 + \zeta b_1 + \eta a_1b_1) + (a_1 + \zeta b_1 + \eta a_1b_1)a_1 = 0$. This contradiction finishes the proof and Wedderburn's theorem is established.

This second proof has some advantages in that we can use parts of it to proceed to a remarkable result due to Jacobson, namely,

THEOREM 7.D (JACOBSON). *Let D be a division ring such that for every $a \in D$ there exists a positive integer $n(a) > 1$, depending on a, such that $a^{n(a)} = a$. Then D is a commutative field.*

Proof. If $a \neq 0$ is in D then $a^n = a$ and $(2a)^m = 2a$ for some integers $n, m > 1$. Let $s = (n-1)(m-1) + 1$; $s > 1$ and a simple calculation shows that $a^s = a$ and $(2a)^s = 2a$. But $(2a)^s = 2^sa^s = 2^sa$, whence $2^sa = 2a$ from which we get $(2^s - 2)a = 0$. Thus D has characteristic $p > 0$. If $P \subset Z$ is the field having p elements (isomorphic to J_p), since a is algebraic over P, $P(a)$ has a finite number of elements, in fact, p^h elements for some integer h. Thus, since $a \in P(a)$, $a^{p^h} = a$. Therefore, if $a \notin Z$ all the conditions of Lemma 7.9 are satisfied, hence there exists a $b \in D$ such that

$$bab^{-1} = a^{\mu} \neq a. \tag{1}$$

By the same argument, $b^{p^k} = b$ for some integer $k > 1$. Let $W = \{x \in D \mid \sum_{i=1}^{p^h} \sum_{j=1}^{p^k} p_{ij}a^ib^j$ where $p_{ij} \in P\}$. W is finite and is closed under addition. By virtue of (1) it is also closed under multiplication. (Verify!) Thus W is a finite ring, and being a subring of the division ring D, it itself must be a division ring (Problem 3). Thus W is a finite division ring; by Wedderburn's theorem it is commutative. But a and b are both in W; therefore, $ab = ba$ contrary to $a^{\mu}b = ba$. This proves the theorem.

Jacobson's theorem actually holds for *any* ring R satisfying $a^{n(a)} = a$ for every $a \in R$, not just for division rings. The transition from the division ring case to the general case while not difficult involves the axiom of choice, and to discuss it would take us too far afield.

PROBLEMS

1. If $t > 1$ is an integer and $(t^m - 1) \mid (t^n - 1)$, prove that $m \mid n$.

2. If D is a division ring, prove that its dimension (as a vector space) over its center cannot be 2.

3. Show that any finite subring of a division ring is a division ring.

4. (a) Let D be a division ring of characteristic $p \neq 0$ and let G be a finite subgroup of the group of nonzero elements of D under multiplication. Prove that G is abelian. (*Hint:* consider the subset $\{x \in D \mid x = \Sigma\lambda_i g_i,\ \lambda_i \in P,\ g_i \in G\}$.)

(b) In part (a) prove that G is actually cyclic.

*5. (a) If R is a finite ring in which $x^n = x$, for all $x \in R$ where $n > 1$ prove that R is commutative.

(b) If R is a finite ring in which $x^2 = 0$ implies that $x = 0$, prove that R is commutative.

*6. Let D be a division ring and suppose that $a \in D$ only has a finite number of conjugates (i.e., only a finite number of distinct $x^{-1}ax$). Prove that a has only one conjugate and must be in the center of D.

7. Use the result of Problem 6 to prove that if a polynomial of degree n having coefficients in the center of a division ring has $n + 1$ roots in the division ring then it has an infinite number of roots in that division ring.

*8. Let D be a division ring and K a subdivision ring of D such that $xKx^{-1} \subset K$ for every $x \neq 0$ in D. Prove that either $K \subset Z$, the center of D or $K = D$. (This result is known as the *Brauer-Cartan-Hua theorem.*)

***9.** Let D be a division ring and K a subdivision ring of D. Suppose that the group of nonzero elements of K is a subgroup of finite index in the group (under multiplication) of nonzero elements of D. Prove that either D is finite or $K = D$.

10. If $\theta \neq 1$ is a root of unity and if q is a positive integer, prove that $|q - \theta| > q - 1$.

3. A Theorem of Frobenius.

In 1877 Frobenius classified all division rings having the field of real numbers in their center and satisfying, in addition, one other condition to be described below. The aim of this section is to present this result of Frobenius.

In Chapter 6 we brought attention to two important facts about the field of complex numbers. We recall them here:

Fact 1. Every polynomial of degree n over the field of complex numbers has all its n roots in the field of complex numbers.

Fact 2. The only irreducible polynomials over the field of real numbers are of degree 1 or 2.

Definition. A division algebra D is said to be *algebraic over a field F* if:

(1) F is contained in the center of D;

(2) every $a \in D$ satisfies a nontrivial polynomial with coefficients in F.

If D, as a vector space, is finite-dimensional over the field F which is contained in its center, it can easily be shown that D is algebraic over F (see Problem 1, end of this section). However, it can happen that D is algebraic over F yet is not finite-dimensional over F.

We start our investigation of division rings algebraic over the real field by first finding those algebraic over the complex field.

Lemma 7.10. *Let C be the field of complex numbers and suppose that the division ring D is algebraic over C. Then $D = C$.*

Proof. Suppose that $a \in D$. Since D is algebraic over C, $a^n + \alpha_1 a^{n-1} + \cdots + \alpha_{n-1}a + \alpha_n = 0$ for some $\alpha_1, \alpha_2, \ldots, \alpha_n$ in C.

Now the polynomial $p(x) = x^n + \alpha_1 x^{n-1} + \cdots + \alpha_{n-1}x + \alpha_n$ in $C[x]$, by Fact 1, can be factored, in $C[x]$, into a product of linear factors; that is, $p(x) = (x - \lambda_1)(x - \lambda_2) \ldots (x - \lambda_n)$, where $\lambda_1, \lambda_2, \ldots, \lambda_n$ are all in C. Since C is in the center of D, every element of C commutes with a, hence $p(a) = (a - \lambda_1)(a - \lambda_2) \ldots (a - \lambda_n)$. But, by assumption, $p(a) = 0$, thus $(a - \lambda_1)(a - \lambda_2) \ldots (a - \lambda_n) = 0$. Since a product in a division ring is zero only if one of the terms of the product is zero, we conclude that

$a - \lambda_k = 0$ for some k, hence $a = \lambda_k$, from which we get that $a \in C$. Therefore, every element of D is in C; since $C \subset D$, we obtain $D = C$.

We are now in a position to prove the classic result of Frobenius, namely,

THEOREM 7.E (FROBENIUS). *Let D be a division ring algebraic over F, the field of real numbers. Then D is isomorphic to one of: the field of real numbers, the field of complex numbers, or the division ring of real quaternions.*

Proof. The proof consists of three parts. In the first, and easiest, we dispose of the commutative case; in the second, assuming that D is not commutative, we construct a replica of the real quaternions in D; in the third part we show that this replica of the quaternions fills out all of D.

Suppose that $D \neq F$ and that a is in D but not in F. By our assumptions, a satisfies some polynomial over F, hence some irreducible polynomial over F. In consequence of Fact 2, a satisfies either a linear or quadratic equation over F. If this equation is linear, a must be in F contrary to assumption. So we may suppose that $a^2 + 2\alpha a + \beta = 0$ where $\alpha, \beta \in F$. Thus $(a - \alpha)^2 = \alpha^2 - \beta$; we claim that $\alpha^2 - \beta < 0$ for, otherwise, it would have a real square root δ and we would have $a - \alpha = \pm\delta$ and so a would be in F. Since $\alpha^2 - \beta < 0$ it can be written as $-\gamma^2$ where $\gamma \in F$. Consequently $(a - \alpha)^2 = -\gamma^2$, whence $\left(\frac{a - \alpha}{\gamma}\right)^2 = -1$. *Thus if $a \in D$, $a \notin F$ we can find real α, γ such that $\left(\frac{a - \alpha}{\gamma}\right)^2 = -1$.*

If D is commutative, pick $a \in D$, $a \notin F$ and let $i = \frac{a - \alpha}{\gamma}$ where α, γ in F are chosen so as to make $i^2 = -1$. Therefore D contains $F(i)$, a field isomorphic to the field of complex numbers. Since D is commutative and algebraic over F it is, all the more so, algebraic over $F(i)$. By Lemma 7.10 we conclude that $D = F(i)$. Thus if D is commutative it is either F or $F(i)$.

Assume, then, that D is *not* commutative. We claim that the center of D must be exactly F. If not there is an a in the center, a not in F. But then for some $\alpha, \gamma \in F$, $\left(\frac{a - \alpha}{\gamma}\right)^2 = -1$ so that the center contains a field isomorphic to the complex numbers. However, by Lemma 7.10 if the complex numbers (or an isomorph of them) were in the center of D then $D = C$ forcing D to be commutative. Hence F is the center of D.

Let $a \in D$, $a \notin F$; for some $\alpha, \gamma \in F$, $i = \frac{a - \alpha}{\gamma}$ satisfies $i^2 = -1$.

Since $i \not\in F$, i is not in the center of F. Therefore there is an element $b \in D$ such that $c = bi - ib \neq 0$. We compute $ic + ci$; $ic + ci = i(bi - ib) + (bi - ib)i = ibi - i^2b + bi^2 - ibi = 0$ since $i^2 = -1$. Thus $ic = -ci$; from this we get $ic^2 = -c(ic) = -c(-ci) = c^2i$, and so c^2 commutes with i. Now c satisfies some quadratic equation over F, $c^2 + \lambda c + \mu = 0$. Since c^2 and μ commute with i, λc must commute with i; that is, $\lambda ci = i\lambda c = \lambda ic = -\lambda ci$, hence $2\lambda ci = 0$, and since $2ci \neq 0$ we have that $\lambda = 0$. Thus $c^2 = -\mu$; since $c \not\in F$ (for $ci = -ic \neq ic$) we can say, as we have before, that μ is positive and so $\mu = \nu^2$ where $\nu \in F$. Therefore $c^2 = -\nu^2$; let $j = \dfrac{c}{\nu}\cdot$ Then j satisfies:

(1) $j^2 = \dfrac{c^2}{\nu^2} = -1.$

(2) $ji + ij = \dfrac{c}{\nu}i + i\dfrac{c}{\nu} = \dfrac{ci + ic}{\nu} = 0.$

Let $k = ij$. The i, j, k we have constructed behave like those for the quaternions, whence $T = \{\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k \,|\, \alpha_0, \alpha_1, \alpha_2, \alpha_3 \in F\}$ forms a subdivision ring of D isomorphic to the real quaternions. We have produced a replica, T, of the division ring of real quaternions in D!

Our last objective is to demonstrate that $T = D$.

If $r \in D$ satisfies $r^2 = -1$ let $N(r) = \{x \in D \,|\, xr = rx\}$. $N(r)$ is a subdivision ring of D; moreover r, and so all $\alpha_0 + \alpha_1 r$, α_0, $\alpha_1 \in F$, are in the center of $N(r)$. By Lemma 7.10 it follows that $N(r) = \{\alpha_0 + \alpha_1 r \,|\, \alpha_0, \alpha_1 \in F\}$. Thus if $xr = rx$ then $x = \alpha_0 + \alpha_1 r$ for some α_0, α_1 in F.

Suppose that $u \in D$, $u \not\in F$. For some α, $\beta \in F$, $w = \dfrac{u - \alpha}{\beta}$ satisfies $w^2 = -1$. We claim that $wi + iw$ commutes with both i and w; for $i(wi + iw) = iwi + i^2w = iwi + wi^2 = (iw + wi)i$ since $i^2 = -1$. Similarly $w(wi + iw) = (wi + iw)w$. By the remark of the preceding paragraph, $wi + iw = \alpha_0' + \alpha_1' i = \alpha_0 + \alpha_1 w$. If $w \not\in T$ this last relation forces $\alpha_1 = 0$ (for otherwise we could solve for w in terms of i). Thus $wi + iw = \alpha_0 \in F$. Similarly $wj + jw = \beta_0 \in F$ and $wk + kw = \gamma_0 \in F$. Let

$$z = w + \frac{\alpha_0}{2}i + \frac{\beta_0}{2}j + \frac{\gamma_0}{2}k.$$

Then

$$\begin{aligned} zi + iz &= wi + iw + \frac{\alpha_0}{2}(i^2 + i^2) + \frac{\beta_0}{2}(ji + ij) + \frac{\gamma_0}{2}(ki + ik) \\ &= \alpha_0 - \alpha_0 = 0; \end{aligned}$$

similarly $zj + jz = 0$ and $zk + kz = 0$. We claim these relations force z to be 0. For $0 = zk + kz = zij + ijz = (zi + iz)j + i(jz - zj) = i(jz - zj)$ since $zi + iz = 0$. However $i \neq 0$, and since we are in a division ring, it follows that $jz - zj = 0$. But $jz + zj = 0$. Thus $2jz = 0$, and since $2j \neq 0$ we have that $z = 0$. Going back to the expression for z we get

$$w + \frac{\alpha_0}{2}i + \frac{\beta_0}{2}j + \frac{\gamma_0}{2}k = 0,$$

whence $w \in T$, contradicting $w \not\in T$. Thus, indeed, $w \in T$. Since $w = \dfrac{u - \alpha}{\beta}$, $u = \beta w + \alpha$ and so $u \in T$. We have proved that any element in D is in T. Since $T \subset D$ we conclude that $D = T$; because T is isomorphic to the real quaternions we now get that D is isomorphic to the division ring of real quaternions. This, however, is just the statement of the theorem.

PROBLEMS

1. If the division ring D is finite-dimensional, as a vector space, over the field F contained in the center of D, prove that D is algebraic over F.

2. Given an example of a field K algebraic over another field F but not finite-dimensional over F.

3. If A is a ring algebraic over a field F and A has no zero divisors prove that A is a division ring.

4. Integral Quaternions and the Four-Square Theorem. In Chapter 3 we considered a certain special class of integral domains called Euclidean rings. When the results about this class of rings were applied to the ring of Gaussian integers we obtained, as a consequence, the famous result of Fermat that every prime number of the form $4n + 1$ is the sum of two squares.

We shall now consider a particular subring of the quaternions which, in all ways except for its lack of commutativity, will look like a Euclidean ring. Because of this it will be possible to explicitly characterize all its left-ideals. This characterization of the left-ideals will lead us quickly to a proof of the classic theorem of Lagrange that every positive integer is a sum of four squares.

Let Q be the division ring of real quaternions. In Q we now proceed to introduce an adjoint operation, $*$, by making the

DEFINITION. For $x = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$ in Q the *adjoint* of x, denoted by x^*, is defined by $x^* = \alpha_0 - \alpha_1 i - \alpha_2 j - \alpha_3 k$.

LEMMA 7.11. *The adjoint in Q satisfies*

(1) $x^{**} = x$
(2) $(\delta x + \gamma y)^* = \delta x^* + \gamma y^*$
(3) $(xy)^* = y^* x^*$

for all x, y in Q and all real δ and γ.

Proof. If $x = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$ then $x^* = \alpha_0 - \alpha_1 i - \alpha_2 j - \alpha_3 k$, whence $x^{**} = (x^*)^* = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$, proving (1).

Let $x = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$ and $y = \beta_0 + \beta_1 i + \beta_2 j + \beta_3 k$ be in Q and let δ and γ be arbitrary real numbers. Thus $\delta x + \gamma y = (\delta\alpha_0 + \gamma\beta_0) + (\delta\alpha_1 + \gamma\beta_1)i + (\delta\alpha_2 + \gamma\beta_2)j + (\delta\alpha_3 + \gamma\beta_3)k$, therefore by the definition of the $*$, $(\delta x + \gamma y)^* = (\delta\alpha_0 + \gamma\beta_0) - (\delta\alpha_1 + \gamma\beta_1)i - (\delta\alpha_2 + \gamma\beta_2)j - (\delta\alpha_3 + \gamma\beta_3)k = \delta(\alpha_0 - \alpha_1 i - \alpha_2 j - \alpha_3 k) + \gamma(\beta_0 - \beta_1 i - \beta_2 j - \beta_3 k) = \delta x^* + \gamma y^*$. This, of course, proves (2).

In light of (2), to prove (3) it is enough to do so for a basis of Q over the reals. We prove it for the particular basis $1, i, j, k$. Now $ij = k$ hence $(ij)^* = k^* = -k = ji = (-j)(-i) = j^*i^*$. Similarly $(ik)^* = k^*i^*$, $(jk)^* = k^*j^*$. Also $(i^2)^* = (-1)^* = -1 = (i^*)^2$, and similarly for j and k. Since (3) is true for the basis elements and (2) holds, (3) is true for all linear combinations of the basis elements with real coefficients, hence (3) holds for arbitary x and y in Q.

DEFINITION. If $x \in Q$ then the *norm* of x, denoted by $N(x)$, is defined by $N(x) = xx^*$.

Note that if $x = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$ then $N(x) = xx^* = (\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k)(\alpha_0 - \alpha_1 i - \alpha_2 j - \alpha_3 k) = \alpha_0^{\,2} + \alpha_1^{\,2} + \alpha_2^{\,2} + \alpha_3^{\,2}$; therefore $N(0) = 0$ and $N(x)$ is a *positive* real number for $x \neq 0$ in Q. In particular, for any real number α, $N(\alpha) = \alpha^2$. If $x \neq 0$ note that $x^{-1} = \dfrac{1}{N(x)}\,x^*$.

LEMMA 7.12. *For all $x, y \in Q$, $N(xy) = N(x)N(y)$.*

Proof. By the very definition of norm, $N(xy) = (xy)(xy)^*$; by part (3) of Lemma 7.11, $(xy)^* = y^*x^*$ and so $N(xy) = xyy^*x^*$. However, $yy^* = N(y)$ is a real number, and thereby it is in the center of Q; in particular it must commute with x^*. Consequently $N(xy) = x(yy^*)x^* = (xx^*)(yy^*) = N(x)N(y)$.

As an immediate consequence of Lemma 7.12 we obtain

LEMMA 7.13 (LAGRANGE IDENTITY). *If $\alpha_0, \alpha_1, \alpha_2, \alpha_3$ and $\beta_0, \beta_1, \beta_2, \beta_3$ are real numbers then* $(\alpha_0^{\,2} + \alpha_1^{\,2} + \alpha_2^{\,2} + \alpha_3^{\,2})(\beta_0^{\,2} + \beta_1^{\,2} + \beta_2^{\,2} + \beta_3^{\,2}) = (\alpha_0\beta_0 - \alpha_1\beta_1 - \alpha_2\beta_2 - \alpha_3\beta_3)^2 + (\alpha_0\beta_1 + \alpha_1\beta_0 + \alpha_2\beta_3 - \alpha_3\beta_2)^2 + (\alpha_0\beta_2 - \alpha_1\beta_3 + \alpha_2\beta_0 + \alpha_3\beta_1)^2 + (\alpha_0\beta_3 + \alpha_1\beta_2 - \alpha_2\beta_1 + \alpha_3\beta_0)^2$.

Proof. Of course there is one obvious proof of this result, namely, multiply everything out and compare terms.

However, an easier way both to reconstruct the result at will and, at the same time, to prove it is to notice that the left-hand side is $N(x)N(y)$ while the right-hand side is $N(xy)$ where $x = \alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k$ and $y = \beta_0 + \beta_1 i + \beta_2 j + \beta_3 k$. By Lemma 7.12 $N(x)N(y) = N(xy)$, ergo the Lagrange identity!

The Lagrange identity says that the sum of four squares times the sum of four squares is again, in a very specific way, the sum of four squares. A very striking result of Adolf Hurwitz says that if the sum of n squares times the sum of n squares is again a sum of n squares, where this last sum has terms computed bilinearly from the other two sums, then $n = 1, 2, 4$, or 8. There is, in fact, an identity for the product of sums of eight squares but it is too long and cumbersome to write down here.

Now is the appropriate time to introduce the Hurwitz ring of integral quaternions. Let $\zeta = \frac{1}{2}(1 + i + j + k)$ and let

$$H = \{m_0\zeta + m_1 i + m_2 j + m_3 k \mid m_0, m_1, m_2, m_3 \text{ integers}\}.$$

LEMMA 7.14. *H is a subring of Q. If $x \in H$ then $x^* \in H$ and $N(x)$ is a* positive integer *for every nonzero x in H.*

We leave the proof of Lemma 7.14 to the reader. It should offer no difficulties.

In some ways H might appear to be a rather contrived ring. Why use the quaternions ζ? Why not merely consider the more natural ring $Q_0 = \{m_0 + m_1 i + m_2 j + m_3 k \mid m_0, m_1, m_2, m_3$ are integers$\}$? The answer is that Q_0 is not large enough, whereas H is, for the key lemma which follows to hold in it. But we want this next lemma to be true in the ring at our disposal for it allows us to characterize its left-ideals. This, perhaps, indicates why we (or rather Hurwitz) chose to work in H rather than in Q_0.

LEMMA 7.15 (LEFT-DIVISION ALGORITHM). *Let a and b be in H with $b \neq 0$. Then there exist two elements c and d in H such that $a = cb + d$ and $N(d) < N(b)$.*

Proof. Before proving the lemma, let's see what it tells us. If we look back in the section in Chapter 3 which deals with Euclidean rings, we can see that Lemma 7.15 assures us that except for its lack of commutativity H has all the properties of a Euclidean ring. The fact that elements in H may fail to commute will not bother us. True, we must be a little careful not to jump to erroneous conclusions; for instance $a = cb + d$ but we have no right to assume that a is also equal to $bc + d$ for b and c might not commute. But this will not influence any argument that we shall use.

In order to prove the lemma we first do so for a very special case, namely, that one in which a is an arbitrary element of H but b is a positive integer n. Suppose that $a = t_0\zeta + t_1 i + t_2 j + t_3 k$ where t_0, t_1, t_2, t_3 are integers and that $b = n$ where n is a positive integer. Let $c = x_0\zeta + x_1 i + x_2 j + x_3 k$ where x_0, x_1, x_2, x_3 are integers yet to be determined. We want to choose them in such a manner as to force $N(a - cn) < N(n) = n^2$. But

$$\begin{aligned} a - cn = {} & \left(t_0\left(\frac{1+i+j+k}{2}\right) + t_1 i + t_2 j + t_3 k\right) - nx_0\left(\frac{1+i+j+k}{2}\right) \\ & - nx_1 i - nx_2 j - nx_3 k \\ & = \tfrac{1}{2}(t_0 - nx_0) + \tfrac{1}{2}(t_0 + 2t_1 - n(t_0 + 2x_1))i \\ & \quad + \tfrac{1}{2}(t_0 + 2t_1 - n(t_0 + 2x_2))j + \tfrac{1}{2}(t_0 + 2t_3 - n(t_0 + 2x_3))k. \end{aligned}$$

If we could choose the integers x_0, x_1, x_2, x_3 in such a way as to make $|t_0 - nx_0| \le \frac{1}{2}n$, $|t_0 + 2t_1 - n(t_0 + 2x_1)| \le n$, $|t_0 + 2t_2 - n(t_0 + 2x_2)| \le n$ and $|t_0 + 2t_3 - n(t_0 + 2x_3)| \le n$ then we would have

$$\begin{aligned} N(a - cn) = {} & \frac{(t_0 - nx_0)^2}{4} + \frac{(t_0 + 2t_1 - n(t_0 + 2x_1))^2}{4} + \cdots \\ & \le \tfrac{1}{16}n^2 + \tfrac{1}{4}n^2 + \tfrac{1}{4}n^2 + \tfrac{1}{4}n^2 < n^2 = N(n), \end{aligned}$$

which is the desired result. But now we claim this can always be done:

(1) There is an integer x_0 such that $t_0 = x_0 n + r$ where $-\dfrac{n}{2} \le r \le \dfrac{n}{2}$; for this x_0, $|t_0 - x_0 n| = |r| \le \dfrac{n}{2}$.

(2) There is an integer k such that $t_0 + 2t_1 = kn + r$ and $0 \le r < n$. If $k - t_0$ is even, put $2x_1 = k - t_0$; then $t_0 + 2t_1 = (2x_1 + t_0)n + r$ and $|t_0 + 2t_1 - (2x_1 + t_0)n| = r < n$. If, on the other hand, $k - t_0$ is odd, put $2x_1 = k - t_0 + 1$; thus $t_0 + 2t_1 = (2x_1 + t_0 - 1)n + r = (2x_1 + t_0)n + r - n$, whence $|t_0 + 2t_1 - (2x_1 + t_0)n| = |r - n| \le n$ since $0 \le r < n$. Therefore we can find an integer x_1 satisfying $|t_0 + 2t_1 - (2x_1 + t_0)n| \le n$.

(3) As in (2) we can find integers x_2 and x_3 which satisfy $|t_0 + 2t_2 - (2x_2 + t_0)n| \le n$ and $|t_0 + 2t_3 - (2x_3 + t_0)n| \le n$ respectively.

In the special case in which a is an arbitrary element of H and b is a positive integer we have now shown the lemma to be true.

We go to the general case wherein a and b are arbitrary elements of H and $b \neq 0$. By Lemma 7.14 $n = bb^*$ is a positive integer thus there exists a $c \in H$ such that $ab^* = cn + d_1$ where $N(d_1) < N(n)$. Thus $N(ab^* - cn) < N(n)$; but $n = bb^*$ whence we get $N(ab^* - cbb^*) < N(n)$, and so $N((a - cb)b^*) < N(n) = N(bb^*)$. By Lemma 7.12 this reduces to $N(a - cb)N(b^*) < N(b)N(b^*)$; since $N(b^*) > 0$ we get $N(a - cb) < N(b)$.

Putting $d = a - cb$ we have $a = cb + d$ where $N(d) < N(b)$. This completely proves the lemma.

As in the commutative case we are able to deduce from Lemma 7.15

LEMMA 7.16. *Let L be a left-ideal of H. Then there exists an element $u \in L$ such that every element in L is a left-multiple of u; in other words, there exists a $u \in L$ such that every $x \in L$ is of the form $x = ru$ where $r \in H$.*

Proof. If $L = (0)$ there is nothing to prove, merely put $u = 0$.

Therefore we may assume that L has nonzero elements. The norms of the nonzero elements are positive integers (Lemma 7.14) whence there is an element $u \neq 0$ in L whose norm is minimal over the nonzero elements of L. If $x \in L$, by Lemma 7.15, $x = cu + d$ where $N(d) < N(u)$. However d is in L because both x and u, and so cu, are in L which is a left-ideal. Thus $N(d) = 0$ and so $d = 0$. From this $x = cu$ is a consequence.

Before we can prove the four-square theorem, which is the goal of this section, we need one more lemma, namely

LEMMA 7.17. *If $a \in H$ then $a^{-1} \in H$ if and only if $N(a) = 1$.*

Proof. If both a and a^{-1} are in H, then by Lemma 7.14 both $N(a)$ and $N(a^{-1})$ are positive integers. However, $aa^{-1} = 1$, whence, by Lemma 7.12, $N(a)N(a^{-1}) = N(aa^{-1}) = N(1) = 1$. This forces $N(a) = 1$.

On the other hand, if $a \in H$ and $N(a) = 1$, then $aa^* = N(a) = 1$ and so $a^{-1} = a^*$. But, by Lemma 7.14, since $a \in H$ we have that $a^* \in H$, and so $a^{-1} = a^*$ is also in H.

We now have determined enough of the structure of H to use it effectively to study properties of the integers. We prove the famous, classical theorem of Lagrange,

THEOREM 7.F. *Every positive integer can be expressed as the sum of squares of four integers.*

Proof. Given a positive integer n we claim in the theorem that $n = x_0^2 + x_1^2 + x_2^2 + x_3^2$ for four integers x_0, x_1, x_2, x_3. Since every integer factors into a product of prime numbers, if every prime number were realizable as a sum of four squares, in view of Lagrange's identity (Lemma 7.13) every integer would be expressible as a sum of four squares. We have reduced the problem to consider only prime numbers n. Certainly the prime number 2 can be written as $1^2 + 1^2 + 0^2 + 0^2$ as a sum of four squares.

Thus, without loss of generality, we may assume that n is an *odd prime number*. As is customary we denote it by p.

Consider the quaternions W_p over J_p, the integers mod p; $W_p = \{\alpha_0 + \alpha_1 i + \alpha_2 j + \alpha_3 k \mid \alpha_0, \alpha_1, \alpha_2, \alpha_3 \in J_p\}$. W_p is a finite ring; moreover, since $p \neq 2$ it is not commutative for $ij = -ji \neq ji$. Thus, by Wedderburn's theorem it cannot be a division ring, whence by Problem 1 at the end of Section 5 in Chapter 3, it must have a left-ideal which is neither (0) nor W_p.

But then the two-sided ideal V in H defined by $V = \{x_0\zeta + x_1 i + x_2 j + x_3 k \mid p$ divides all of $x_0, x_1, x_2, x_3\}$ cannot be a maximal left-ideal of H, since H/V is isomorphic to W_p. (Prove!) (If V were a maximal left-ideal in H, H/V, and so W_p, would have no left-ideals other than (0) and H/V).

Thus there is a left-ideal L of H satisfying: $L \neq H$, $L \neq V$, and $L \supset V$. By Lemma 7.16, there is an element $u \in L$ such that every element in L is a left-multiple of u. Since $p \in V$, $p \in L$, whence $p = cu$ for some $c \in H$. Since $u \not\in V$, c cannot have an inverse in H, otherwise $u = c^{-1}p$ would be in V. Thus $N(c) > 1$ by Lemma 7.17. Since $L \neq H$, u cannot have an inverse in H, whence $N(u) > 1$. Since $p = cu$, $p^2 = N(p) = N(cu) = N(c)N(u)$. But $N(c)$ and $N(u)$ are integers, since both c and u are in H, both are larger than 1 and both divide p^2. The only way this is possible is that $N(c) = N(u) = p$.

Since $u \in H$, $u = m_0\zeta + m_1 i + m_2 j + m_3 k$ where m_0, m_1, m_2, m_3 are integers; thus $2u = 2m_0\zeta + 2m_1 i + 2m_2 j + 2m_3 k = (m_0 + m_0 i + m_0 j + m_0 k) + 2m_1 i + 2m_2 j + 2m_3 k = m_0 + (2m_1 + m_0)i + (2m_2 + m_0)j + (2m_3 + m_0)k$. Therefore $N(2u) = m_0^2 + (2m_1 + m_0)^2 + (2m_2 + m_0)^2 + (2m_3 + m_0)^2$. But $N(2u) = N(2)N(u) = 4p$ since $N(2) = 4$ and $N(u) = p$. We have shown that $4p = m_0^2 + (2m_1 + m_0)^2 + (2m_2 + m_0)^2 + (2m_3 + m_0)^2$. We are almost done.

To finish the proof we introduce an old trick of Euler's: If $2a = x_0^2 + x_1^2 + x_2^2 + x_3^2$ where a, x_0, x_1, x_2 and x_3 are integers, then $a = y_0^2 + y_1^2 + y_2^2 + y_3^2$ for some integers y_0, y_1, y_2, y_3. To see this note that, since $2a$ is even, the x's are all even, all odd or two are even and two are odd. At any rate in all three cases we can renumber the x's and pair them in such a way that

$$y_0 = \frac{x_0 + x_1}{2},\ y_1 = \frac{x_0 - x_1}{2},\ y_2 = \frac{x_2 + x_3}{2}, \quad \text{and} \quad y_3 = \frac{x_2 - x_3}{2}$$

are all integers. But

$$\begin{aligned} y_0^2 + y_1^2 + y_2^2 + y_3^2 & \\ &= \left(\frac{x_0 + x_1}{2}\right)^2 + \left(\frac{x_0 - x_1}{2}\right)^2 + \left(\frac{x_2 + x_3}{2}\right)^2 + \left(\frac{x_2 - x_3}{2}\right)^2 \\ &= \tfrac{1}{2}(x_0^2 + x_1^2 + x_2^2 + x_3^2) \\ &= \tfrac{1}{2}(2a) \\ &= a. \end{aligned}$$

Since $4p$ is a sum of four squares, by the remark just made $2p$ also is; since $2p$ is a sum of four squares, p also must be such a sum. Thus $p = a_0^2 + a_1^2 + a_2^2 + a_3^2$ for some integers a_0, a_1, a_2, a_3 and Lagrange's theorem is established.

This theorem itself is the starting point of a large research area in number theory, the so-called *Waring problem.* This asks if every integer can be written as a sum of a fixed number of kth powers. For instance it can be shown that every integer is a sum of nine cubes, nineteen fourth powers, etc. The Waring problem was shown to have an affirmative answer, in this century, by the great mathematician Hilbert.

PROBLEMS

1. Prove Lemma 7.14.

2. Find all the elements a in Q_0 such that a_0^{-1} is also in Q_0.

3. Prove that there are exactly 24 elements a in H such that a^{-1} is also in H. Determine all of them.

4. Give an example of an a and b, $b \neq 0$, in Q_0 such that it is impossible to find c and d in Q_0 satisfying $a = cb + d$ where $N(d) < N(b)$.

5. Prove that if $a \in H$ then there exist integers α, β such that $a^2 + \alpha a + \beta = 0$.

6. Prove that there is a positive integer which cannot be written as the sum of three squares.

***7.** Exhibit an infinite number of positive integers which cannot be written as the sum of three squares.

Supplementary Reading

For a deeper discussion of finite fields: ALBERT, A. A., *Fundamental Concepts of Higher Algebra*. University of Chicago Press, Chicago, 1956.

For many proofs of the four-square theorem and a discussion of the Waring problem: HARDY, G. H., and WRIGHT, E. M., *An Introduction to the Theory of Numbers,* second edition. Clarendon Press, Oxford, England, 1945.

For another proof of the Wedderburn theorem: ARTIN, E., "Uber einen Satz von Herrn J. H. M. Wedderburn," *Abhandlungen, Hamburg Mathematisches Seminar*, Vol. 5 (1928), pages 245–50.