Proceedings of the Edinburgh Mathematical Society (1999) 42, 1–16 ©

ON ADDITIVE POLYNOMIALS OVER A FINITE FIELD

by R. W. K. ODONI

(Received 20th March 1995, revised 1st August 1998)

This paper is based on the interpretation of the ring $\mathcal{A}_q$ of additive polynomials in one variable over a finite field $\mathbf{F}_q$ as a maximal R-order inside a certain skew-field D, R being a principal ideal domain isomorphic to $\mathbf{F}_p[T]$. The classical (1930's) structure theory of maximal orders in global fields is used to solve enumeration questions involving the iteration of members of $\mathcal{A}_q$.

1991 *Mathematics subject classification:* 12CO5, 16HO5.

0. Introduction

For h, p in $\mathbb{N}$, with p prime, let $\mathbb{F}_q$ be the finite field of order $q = p^h$, and let T, T' be independent (commuting) variables over $\mathbb{F}_q$. Consider the set

$$\mathcal{A}_q := \{f(T) \in \mathbb{F}_q[T]; f(T + T') = f(T) + f(T') \text{ in } \mathbb{F}_q[T, T']\} \tag{0.1}$$

The members of $\mathcal{A}_q$ are the *additive polynomials* over $\mathbb{F}_q$. It is classically well-known that the $f(T)$ in $\mathcal{A}_q$ are just the polynomials of the type $\sum_{j=0}^{d} a_j T^{p^j}$, for $d \geq 0$ and $a_0, \ldots, a_d \in \mathbb{F}_q$. Although closed under $+$, $\mathcal{A}_q$ (for $q \neq p$) is not a subring of $\mathbb{F}_q[T]$. However, as observed by O. Øre [8], $\mathcal{A}_q$ can be made into a (non-commutative) ring under the $+$ of $\mathbb{F}_q[T]$, and the product $\circ$ ("composition of maps"), i.e., $(f \circ g)(T) = f(g(T))$. The 1-element for $(\mathcal{A}_q, +, \circ)$ is $f(T) = T$. When speaking of $\mathcal{A}_q$ we shall henceforth assume that it is equipped with $(+, \circ)$. It is easily seen that the polynomials λT ($\lambda \in \mathbb{F}_p$) are central in $\mathcal{A}_q$, so that $\mathcal{A}_q$ may be regarded as an $\mathbb{F}_p$-algebra. Amongst the various results of [8], the following are both elementary and crucial for all that follows. Let deg be the standard degree-function on $\mathbb{F}_q[T]$. If $f(T), g(T) \in \mathcal{A}_q$ we have $\deg(f \circ g)(T) = \deg f(T).\deg g(T)$, and, in particular, $(f \circ g)(T) = 0$ if and only if $f(T)$ or $g(T) = 0$. Hence $\mathcal{A}_q$ is a (non-commutative) integral domain. Moreover deg induces "Euclidean algorithms" on $\mathcal{A}_q$ in the following sense. Let $g(T), f(T) \in \mathcal{A}_q$, with $f(T) \neq 0$. Then there exist unique $h_1(T), h_2(T), r_1(T), r_2(T)$ in $\mathcal{A}_q$, with $\deg r_1(T)$, $\deg r_2(T) < \deg f(T)$ and

$$\left.\begin{aligned} g(T) &= (f \circ h_1)(T) + r_1(T) \\ g(T) &= (h_2 \circ f)(T) + r_2(T) \end{aligned}\right\}. \tag{0.2}$$

(The first of these is obvious, while the second becomes clear once we recall that $\mathbb{F}_q$ is a *perfect* field.)

Consequently every left (resp. right-) ideal in $\mathcal{A}_q$ is left (resp. right-) principal. One easily checks that the units in $\mathcal{A}_q$ are just the $f(T) = \alpha T$, $(0 \neq \alpha \in \mathbb{F}_q)$, and that the centre C of $\mathcal{A}_q$ consists of the polynomials of the type $\sum_{j=0}^{d} b_j T^{q^j}$, $d \geq 0$, $b_j \in \mathbb{F}_p$. Further if $J = \mathcal{A}_q \circ f$ is a non-zero left ideal, then $\mathcal{A}_q/J$ is a finite-dimensional $\mathbb{F}_q$-vector space of dimension d, where $\deg f(T) = p^d$. The same is true of $\mathcal{A}_q/J'$, where $J' = f \circ \mathcal{A}_q$, a non-zero right ideal.

There are numerous problems about finite fields $\mathbb{F}_q$ which lead to questions about $\mathcal{A}_q$ – [**6**, **12**] for some examples.

In this paper we consider two apparently new problems about $\mathcal{A}_q$ which are not, in general, accessible to the classical (commutative) methods which have previously been used for $\mathbb{F}_q[T]$. In particular we consider:

Problem 1. Given $d \geq 1$, how many (distinct) irreducible elements π in $\mathcal{A}_q$ have degree p^d? (Here $(0 \neq)\pi \in \mathcal{A}_q$ is *irreducible* if and only if $\mathcal{A}_q \circ \pi$ is a maximal left-ideal; equivalently, π cannot be expressed as $\lambda \circ \mu$ with λ, μ non-units. This is also equivalent to saying that $\pi \circ \mathcal{A}_q$ is a maximal right-ideal.)

Problem 2. Let $f(T) \in \mathcal{A}_q$ with $\deg f(T) \geq p, f(T)$ not a power of T. For $n \in \mathbb{N}$, let $f^{(n)}(T) = (f \circ \ldots \circ f)(T)$ (n factors). For each $n \in \mathbb{N}$ let $s(n) = s(n, f)$ be the least $s \in \mathbb{N}$ such that $f^{(n)}(T)$ splits completely into linear factors in $\mathbb{F}_{q^s}[T]$. Determine the manner in which $s(n)$ varies with n.

In this paper we shall solve Problem 1 completely, and give very precise information about Problem 2. The exact statements of our results will be given later.

The new ideas introduced in this paper are as follows. In Section 1 we show that there is an $\mathbb{F}_p$-algebra isomorphism of $\mathcal{A}_q$ with the *skew-polynomial ring* $\mathbb{F}_q[X; \sigma]$, where $\sigma : \zeta \mapsto \zeta^p$ is the Frobenius automorphism of $\mathbb{F}_q$, i.e., the canonical generator of $\mathrm{Gal}(\mathbb{F}_q/\mathbb{F}_p)$. (This result is already implicit in [**9**].) Writing Λ for $\mathbb{F}_q[X; \sigma]$, we show that the centre R of Λ is the "ordinary" polynomial ring $\mathbb{F}_p[X^h]$, where $q = p^h$. (Thus R is a commutative P.I.D.)

We also show that, as an R-module, Λ is free of rank h^2. Let F be the field of fractions of R. We put $D = \Lambda \otimes_R F$, regarded initially as R-algebra, but, in the usual way, made into an F-algebra. We show that D is a skewfield, having F as centre, with $\dim_F(D) = h^2$, in which Λ (identified with $\Lambda \otimes_R R$) is a maximal R-order in the classical sense [**3**, **10**]. This makes available the full apparatus of "arithmetic" in maximal R-orders in simple-central algebras over global fields F, R being a Dedekind domain with field of fractions F; this beautiful theory is described in near-optimal fashion in [**10**].

As a result of our work in Section 1, the solution of Problem 1 is presented in Section 2.

Problem 2 needs some preliminary reductions before the results of Section 1 can be applied to it, and these reductions occupy Section 3. In Sections 4–5 we complete the

analysis of $s(n,f)$. The end-result of this work (expressed more precisely in Section 5 as Theorem 2) is that there exists a positive constant α_f such that $\liminf\left\{\frac{s(n,f)}{n}\right\} = \alpha_f$, $\limsup\left\{\frac{s(n,f)}{n}\right\} = p\alpha_f$. Theorem 2 actually gives more explicit descriptions of α_f and the behaviour of $s(n,f)$ for large n.

In Section 6 we illustrate Theorem 2 for a special choice of $f(T) \in \mathcal{A}_q$. For this particular f the sequence $s(n) = s(n,f)$ can be calculated explicitly via elementary linear algebra over $\mathbb{F}_p$. For general f this elementary approach does not usually work.

I am greatly indebted to my colleague Professor K. A. Brown for a number of useful discussions, and for providing reference [7] and to Dr. R. J. Chapman (Exeter) for pointing out an error in an earlier draft of this paper.

1. $\mathcal{A}_q$ as a skew-polynomial ring

Let K be any (commutative) perfect field and let $\sigma \in \mathrm{Aut}\,(K)$ have finite order. We recall Øre's classical construction [2, 9] of the *skew-polynomial ring* $K[X;\sigma]$. The underlying set of $K[X;\sigma]$ is the set of all sequences $\mathbf{a} = (a_o, a_1, \ldots, a_n, \ldots)$ where $a_n \in K$ for all $n \geq 0$, and only finitely many a_n are non-zero. The *sum* of two such sequences $\mathbf{a}$ and $\mathbf{b}$ is defined to be $\mathbf{c}$, where $c_n = a_n + b_n$ for all $n \geq 0$, while the *product* $\mathbf{a}.\mathbf{b}$ is defined to be $\mathbf{d}$, where, for all $n \geq 0$,

$$d_n = \sum_{i+j=n} a_i b_j^{\sigma^i}. \tag{1.1}$$

Here the calculations of c_n, d_n are performed within K (with its standard addition and multiplication), $b_j^{\sigma^i}$ being the image of b_j under σ^i. Under $(+,.)$, $K[X;\sigma]$ becomes a (non-commutative) ring with zero-element $\mathbf{0} = (0,\ldots,0,\ldots)$ and 1-element $\mathbf{1} = (1,0,\ldots,0,\ldots)$. The sequence $(0,1,0,\ldots,0,\ldots)$ is denoted by X, while $c \in K$ is identified with $(c,0,,\ldots,0,\ldots)$. The *degree* of $\mathbf{a} \in K[X;\sigma]$ is $\max\{n \geq 0; a_n \neq 0\}$. We have $\deg(\mathbf{a}.\mathbf{b}) = \deg(\mathbf{a}) + \deg(\mathbf{b})$ except if $\mathbf{a}$ or $\mathbf{b} = \mathbf{0}$. In particular, $\mathbf{a}.\mathbf{b} = \mathbf{0}$ if and only if $\mathbf{a}$ or $\mathbf{b} = \mathbf{0}$, so that $K[X;\sigma]$ is a (non-commutative) integral domain. The degree-function (defined above) yields Euclidean algorithms for both left- and right-division in $K[X;\sigma]$, so that every left/right ideal in $K[X;\sigma]$ is left/right principal. Finally, if $\mathbf{a}$ has degree d, then $\mathbf{a} = \sum_{i=0}^{d} a_i X^i$, but for $b \in K$ we have $X^j.b = b^{\sigma^j}.X^j$, i.e., in general, X "does not commute with the coefficients".

Now let $K = \mathbb{F}_q$, $q = p^h$, p prime, and let $\sigma : \zeta \mapsto \zeta^p$ be the Frobenius automorphism of $\mathbb{F}_q$, i.e., the canonical generator of $\mathrm{Gal}(\mathbb{F}_q/\mathbb{F}_p)$. We put $\Lambda = \mathbb{F}_q[X;\sigma]$; multiplication in Λ will be denoted by juxtaposition, suppressing the "dot" used above. Consider the (set-) map $\psi : \mathcal{A}_q \to \Lambda$, sending $\sum_{i=0}^{d} a_i T^{p^i}$ to $\sum_{i=0}^{d} a_i X^i$ whenever $a_0, \ldots, a_d \in \mathbb{F}_q$. As a set-map ψ is trivially bijective, but a simple calculation shows that it is also a ring-isomorphism from $(\mathcal{A}_q, +, \circ)$ to $(\Lambda, +, .)$, sending T to $\mathbf{1}$, and λT to $\lambda\mathbf{1}$ for $\lambda \in \mathbb{F}_p$. Moreover, ψ establishes an $\mathbb{F}_p$-algebra-isomorphism from $\mathcal{A}_q$ to Λ. Clearly ψ maps the centre C of $\mathcal{A}_q$ onto the centre R of Λ; by the results quoted in Section 0, $R = \psi(C)$ is the "ordinary" polynomial ring $\mathbb{F}_p[X^h]$ in one commutative variable X^h, a

(commutative) P.I.D. (That $\mathbb{F}_p[X^h]$ is the centre of Λ can also be verified directly, without reference to $\mathcal{A}_q$.)

Let $\omega \in \mathbb{F}_q$ be a *normal basis* for $\mathbb{F}_q$ over $\mathbb{F}_p$, i.e., the elements $\omega_0 = \omega, \omega_1 = \omega^p, \ldots, \omega_{h-1} = \omega^{p^{h-1}}$ form an $\mathbb{F}_p$-basis for $\mathbb{F}_q$. (The existence of such an ω is a standard fact from elementary Galois theory, since $\sigma : \xi \mapsto \xi^p$ generates $\mathrm{Gal}(\mathbb{F}_q/\mathbb{F}_p)$ – see[5].) It is simple to check that the $\omega_i X^j$ $(0 \leq i, j < h)$ *freely-generate* Λ as an R-module, i.e., Λ is a *free* R-module of rank h^2.

Let $F = \mathbb{F}_p(X^h)$ be the field of fractions of R. We now construct $D = \Lambda \otimes_R F$, which may be regarded as an F-algebra. We may identify Λ with $\Lambda \otimes_R R$, and do so whenever convenient. Certainly $\dim_F D = h^2$. We show that D is a skewfield with centre F (identifying F with $R \otimes_R F$).

Choose an R-basis for Λ ($b_1, \ldots, b_{h^2}$, say), and consider the R-linear map $y \mapsto y\lambda$ of Λ, where $\lambda \in \Lambda$. With respect to $b_1, \ldots, b_{h^2}$, $y \mapsto y\lambda$ is described by an $h^2 \times h^2$ R-matrix $\mathbf{M}(\lambda)$, and $\mathbf{M}(\lambda\lambda') = \mathbf{M}(\lambda)\mathbf{M}(\lambda')$ for all $\lambda, \lambda' \in \Lambda$. If $\lambda = 0, \mathbf{M}(\lambda) = \mathbf{0}$; if $\lambda \neq 0$ then $\det \mathbf{M}(\lambda) \neq 0$, since $\Lambda/\Lambda\lambda \cong \mathcal{A}_q/\mathcal{A}_q \circ f$, where $f = \psi^{-1}(\lambda) \neq 0$, and $\mathcal{A}_q/\mathcal{A}_q \circ f$ is a finite $\mathbb{F}_p$-module, (by Section 0), so that $rk_R(\Lambda\lambda) = rk_R(\Lambda) = h^2$.

Let $\lambda \neq 0$ lie in Λ. Then $\det \mathbf{M}(\lambda)$ is a non-zero element of R; we write $\det(\lambda)$ rather than $\det \mathbf{M}(\lambda)$. Thus $0 \neq \det(\lambda) \in R$, and clearly $\det(\lambda) = \lambda\lambda' = \lambda'\lambda$ for some $\lambda' \in \Lambda$, $\lambda' \neq 0$. We are now ready to prove:

Lemma 1.1. *$D = \Lambda \otimes_R F$ is a skewfield, with $R \otimes_R F$ as its centre.*

Proof. Let $\delta \in D$ be non-zero. Writing $\otimes$ for $\otimes_R$, we have $\delta = \sum_{i=1}^n \lambda_i \otimes f_i$ for some $n \geq 1$, where $\lambda_i \in \Lambda$, $f_i \in F$. We may assume that no $\lambda_i = 0$ and no $f_i = 0$. Choose some non-zero $r \in R$ such that all rf_i lie in R, i.e., $rf_i = r_i \neq 0$ $(i = 1, \ldots, n)$. Then

$$\delta(1 \otimes r) = \sum_{i=1}^{n} \lambda_i \otimes r_i \tag{1.2}$$

Here $1 \otimes r$ is clearly a (central) unit in D. Hence $0 \neq \delta(1 \otimes r) = \sum_{i=1}^n (r_i\lambda_i \otimes 1) = \lambda \otimes 1$, where $\lambda \in \Lambda$ is non-zero. From the foregoing, we have $0 \neq \det \lambda = \lambda\lambda' = \lambda'\lambda \in R$, for some $\lambda' \neq 0$ in Λ. Thus $\delta(1 \otimes r)(\lambda' \otimes 1) = (1 \otimes \det \lambda)$ is a (central) unit in D. Hence δ is a unit in D, since $(1 \otimes r)$ is a (central) unit. This shows that D is, indeed, a skewfield.

Finally, let $\delta \in D$ be central. Then δ commutes with $y \otimes 1$ for all $y \in \Lambda$. As above, let $\delta(1 \otimes r) = \lambda \otimes 1$ where $0 \neq r \in R$. Since $1 \otimes r$ is central, so is $\lambda \otimes 1$; hence $(\lambda \otimes 1)(y \otimes 1) = (y \otimes 1)(\lambda \otimes 1)$ for all $y \in \Lambda$, so that $\lambda y = y\lambda$; thus $\lambda \in R$. This gives $R \otimes F =$ centre of D, proving the lemma.

To summarise, D is a simple central F-algebra of dimension h^2, and is a skewfield, while Λ is an R-order in D. It is an easy exercise to show that D may also be regarded as the classical (Øre) skewfield of fractions of the domain Λ.

Since R is a commutative P.I.D., the results of [7] show that Λ is a maximal R-order in D. Since $F \cong \mathbb{F}_p(T)$ is a global field, and $R \cong \mathbb{F}_p[T]$ is a Dedekind domain (indeed,

a P.I.D.), having F as field of fractions, the whole "classical theory of maximal orders" given in [**3**, **10**] may be applied to study D, Λ and $(\mathcal{A}_q, +, \circ)$. In particular, we are now well-placed to solve Problems 1 and 2 of Section 0.

2. Solution of problem 1

In Problem 1 we are given $d \geq 1$, and seek to enumerate the $\hat{\pi} \in \mathcal{A}_q$ of degree p^d such that $\mathcal{A}_q \circ \hat{\pi}$ is a maximal left ideal in $(\mathcal{A}_q, +, \circ)$. Using the isomorphism $\psi : \mathcal{A}_q \to \Lambda$ of Section 1, Problem 1 is equivalent to enumerating the π in Λ of degree d such that $\Lambda\pi$ is a maximal left ideal in Λ.

For $d = 1$, the enumeration is trivial, since the $\hat{\pi}$ in $\mathcal{A}_q$ of degree p are just the $aT^p + bT$, $a, b \in \mathbb{F}_q$, $a \neq 0$. Putting

$$N_d := \#\{\hat{\pi} \in \mathcal{A}_q; \text{ irreducible, } \deg \hat{\pi} = p^d\}, \tag{2.1}$$

we thus have

$$N_1 = q(q-1). \tag{2.2}$$

We assume that $d \geq 2$ from now onwards. We shall evaluate N_d by working in Λ. Let $\pi \in \Lambda$ be irreducible of degree d, so that $\Lambda\pi$ is a maximal left Λ-ideal. By Sections 0, 1,

$$\#(\mathcal{A}_q/\mathcal{A}_q \circ \hat{\pi}) = q^d = (\Lambda : \Lambda\pi), \tag{2.3}$$

where $\hat{\pi} = \psi^{-1}(\pi)$, $\deg \hat{\pi} = p^d$, $\deg \pi = d$. Here, and throughout, $(\Gamma : \Delta)$ denotes the index of the subgroup Δ of an additive abelian group Γ such that Γ/Δ is finite.

Let $\mathfrak{p} = \Lambda\pi \cap R$; clearly $\mathfrak{p}$ is a maximal ideal of R. By [**10**, pp. 195–6], $\Lambda/\Lambda\pi$ is a simple left Λ-module, while $\mathbf{P} := \operatorname{ann}_\Lambda(\Lambda/\Lambda\pi)$ is a maximal two-sided Λ-ideal with $\mathbf{P} \subset \Lambda\pi$ and $\mathbf{P} \cap R = \mathfrak{p}$.

Further $\Lambda/\mathbf{P}$ is a finite simple ring, hence a total matrix algebra $M_K(\mathbb{F}_Q)$ over a finite field $\mathbb{F}_Q$ with Q elements.

Moreover $\Lambda/\Lambda\pi$ is a simple left $\Lambda/\mathbf{P}$-module, and so, as a left $\Lambda/\mathbf{P}$-module, $\Lambda/\mathbf{P}$ itself is isomorphic to a direct sum of K copies of $\Lambda/\Lambda\pi$. This gives

$$(\Lambda : \mathbf{P}) = (\Lambda : \Lambda\pi)^K = q^{dK} = Q^{K^2}. \tag{2.4}$$

Next, let $\overline{R}$ be the finite field $R/\mathfrak{p}$. Then, for some $f \in \mathbb{N}$, $\Lambda/\mathbf{P}$ is of dimension f over $\overline{R}$, so that

$$(\Lambda : \mathbf{P}) = (R : \mathfrak{p})^f. \tag{2.5}$$

Using [**10**, p. 213], we have

https://doi.org/10.1017/S0013091500019970 Published online by Cambridge University Press

$$f = Kh \quad (q = p^h). \tag{2.6}$$

Also, by [**10**, p. 222] we have $\mathfrak{p}\Lambda = \mathbf{P}^e$ for some $e \in \mathbb{N}$, where $e > 1$ only if $\mathfrak{p}$ divides the discriminant $\mathrm{dis}(\Lambda/R)$. Using the R-basis $\omega_i X^j$ $(0 \le i, j < h)$ for Λ, described just before Lemma 1.1, it is easily seen that $e = 1$ unless $\mathfrak{p} = (X^h) = X^h R$, corresponding to $\Lambda\pi = \Lambda X$, $\deg \pi = 1$. Hence, for $\deg \pi = d \ge 2$, we have $e = 1$, and $p\Lambda = \mathbf{P}$. This gives

$$(\Lambda : \mathfrak{p}\Lambda) = (R : \mathfrak{p})^{h^2} = (\Lambda : \mathbf{P}). \tag{2.7}$$

since Λ is R-free of rank h^2. From (2.5)–(2.7) we see that $f = h^2$, $K = h$, and then, by (2.4), $Q = p^d = (R : \mathfrak{p})$.

To summarise, for $d \ge 2$, and $\pi \in \Lambda$ irreducible of degree d, $\mathfrak{p} = \Lambda\pi \cap R$ is a maximal ideal of R with $(R : \mathfrak{p}) = p^d$, while $\mathfrak{p}\Lambda = \mathbf{P}$ is $\mathrm{ann}_\Lambda \Lambda/\Lambda\pi$, $\mathbf{P} \cap R = \mathfrak{p}$ and $\Lambda/\mathbf{P} \cong M_h(\mathbb{F}_{p^d})$. Conversely, given a maximal ideal $\mathfrak{p}$ of R with $(R : \mathfrak{p}) = p^d$, the maximal left-Λ ideals $\Lambda\pi$ which contain $\mathbf{P} = \mathfrak{p}\Lambda$ satisfy $\deg \pi = d$. This immediately gives (via (2.1)) that, for $d \ge 2$,

$$N_d = \sum_{(R:\mathfrak{p}) = p^d} N_{d,\mathfrak{p}}, \tag{2.8}$$

where $N_{d,\mathfrak{p}}$ is $\#\{\text{irreducible } \pi \text{ in } \Lambda;\ \deg \pi = d,\ R \cap \Lambda\pi = \mathfrak{p}\}$.

Since Λ has $q - 1$ units, we see from the above that

$$N_{d,\mathfrak{p}} = (q - 1)\#\{\text{maximal left ideals in } M_h(\mathbb{F}_{p^d})\}. \tag{2.9}$$

In particular $N_{d,\mathfrak{p}}$ is independent of the choice of $\mathfrak{p}$ with $(R : \mathfrak{p}) = p^d$; writing M_d for this common value of the $N_{d,\mathfrak{p}}$ in (2.9), and G_d for $\#\{\text{maximal } \mathfrak{p} \text{ in } R;\ (R : \mathfrak{p}) = d\}$, we have, from (2.8) and (2.9), that

$$N_d = G_d M_d. \tag{2.10}$$

By a celebrated result of Dedekind [**4**], we have

$$dG_d = \sum_{e/d} \mu(d/e) p^e, \tag{2.11}$$

where $\mu(\ldots)$ is the classical Möbius function on $\mathbb{N}$. It now only remains to evaluate M_d of (2.10).

For any field K, $\Gamma := M_h(K)$ is a simple central K-algebra, and, by a trivial special case of "Morita equivalence", the maximal left ideals of Γ are in bijective correspondence with codimension-one K-vector subspaces of K^h. Specialising to $K = \mathbb{F}_{p^d}$, we thus see that $M_d = (q^d - 1)/(p^d - 1)$, the number of codimension-one $\mathbb{F}_{p^d}$ subspaces of $(\mathbb{F}_{p^d})^h$.

To summarize, with N_d as in (2.1), we have proved

https://doi.org/10.1017/S0013091500019970 Published online by Cambridge University Press

Theorem 1. (i) *For* $d = 1$, $N_d = N_1 = q(q-1)$.

(ii) *For* $d \geq 2$,

$$N_d = d^{-1}(p^d - 1)^{-1}(q-1)(q^d - 1)\sum_{e|d} \mu(e/d)p^e.$$

Clearly Theorem 1 yields the complete solution to Problem 1.

3. Problem 2 – preliminary reductions

Let $f \in \mathcal{A}_q$, $\deg f \geq p$. We write

$$f(T) = \sum_{i=c}^{d} a_i T^{p^i}, \tag{3.1}$$

where the $a_i \in \mathbb{F}_q$ and $a_c a_d \neq 0$. (Here $c \leq d$.) A simple induction on n shows that the nth iterate $f^{(n)}$ of f has the form

$$f^{(n)}(T) = \sum_{i=nc}^{nd} a_i^{(n)} T^{p^i}, \tag{3.2}$$

where $a_i^{(n)} \in \mathbb{F}_q$ for all i and $a_{nc}^{(n)} a_{nd}^{(n)} \neq 0$.

Now let $\overline{\mathbb{F}}_q$ be the algebraic closure of $\mathbb{F}_q$. We regard $\mathcal{A}_q$ as a subset of $\mathbb{F}_q[T]$ and $\mathbb{F}_q[T]$ as a subset of $\overline{\mathbb{F}}_q[T]$. Problem 2 asks for the least $s = s(n,f)$ such that $f^{(n)}(T)$ splits completely into linear factors in $\mathbb{F}_{q^s}[T]$.

From (3.2) there is, for all $n \in \mathbb{N}$, a finite subset $J(n) = J(n,f)$ of $\mathbb{N}$, such that, in $\overline{\mathbb{F}}_q[T]$,

$$f^{(n)}(T) = b_n \prod_{j \in J(n)} \left(T - \alpha_j^{(n)}\right)^{p^{nc}}; \tag{3.3}$$

here $b_n \neq 0$ is in $\mathbb{F}_q$, and $\#J(n) = p^{n(d-c)}$, the $\alpha_j^{(n)}$ $(j \in J(n))$ being distinct elements of $\overline{\mathbb{F}}_q$ (one of which is 0).

Lemma 3.1. *Let* $f \in \mathcal{A}_q$ *be as above, and let* $s \in \mathbb{N}, n \in \mathbb{N}$. *The following are equivalent:*

(i) *all zeros of* $f^{(n)}(T)$ *lie in* $\mathbb{F}_{q^s}$.

(ii) $(T^{q^s} - T)^{p^{nc}}$ *belongs to* $\mathcal{A}_q \circ f^{(n)}$.

Proof. Let (i) hold. Then, by (3.3), all $\alpha_j^{(n)}$ $(j \in J(n))$ lie in $\mathbb{F}_{q^s}$. Since the $\alpha_j^{(n)}$ are distinct, $\prod_{j \in J(n)}(T - \alpha_j^{(n)})$ divides $T^{q^s} - T$ in $\mathbb{F}_{q^s}[T]$, so that $f^{(n)}(T)$ divides $(T^{q^s} - T)^{p^{nc}}$ in $\mathbb{F}_{q^s}[T]$. Both these polynomials lie in $\mathbb{F}_Q[T]$, so that $(T^{q^s} - T)^{p^{nc}} = g(T)f^{(n)}(T)$ in $\mathbb{F}_q[T]$

https://doi.org/10.1017/S0013091500019970 Published online by Cambridge University Press

for some $g(T)$ in $\mathbb{F}_q[T]$. It is clear that $(T^{q^s} - T)^{p^{nc}}$ lies in $\mathcal{A}_q$. From the Euclidean algorithm in $\mathcal{A}_q$, there exist unique $h(T)$, $r(T)$ in $\mathcal{A}_q$ such that $\deg r(T) < p^{nd}$ and

$$(T^{q^s} - T)^{p^{nc}} = (h \circ f^{(n)})(T) + r(T) \tag{3.4}$$

We now regard (3.4) as an equation in $\overline{\mathbb{F}}_q[T]$. Since $h \in \mathcal{A}_q$, we have $h(0) = 0$, and so $r(T)$ vanishes with multiplicity at least p^{nc} at each $\alpha_j^{(n)}$, $(j \in J(n))$. Since $\deg r(T) < p^{nd} = \deg f^{(n)}(T)$, we see that $r(T) \equiv 0$, and so

$$(T^{q^s} - T)^{p^{nc}} = (h \circ f^{(n)})(T) \in \mathcal{A}_q \circ f^{(n)}. \tag{3.5}$$

Conversely, if (3.5) holds, then

$$(\alpha_j^{(n)q^s} - \alpha_j^{(n)})^{p^{nc}} = h(f^{(n)}(\alpha_j^{(n)})) = 0,$$

for all $j \in J(n)$ (since $h(0) = 0$), and so all $\alpha_j^{(n)}$ lie in $\mathbb{F}_{q^s}$.

Lemma 3.1 reduces Problem 2 to a congruence question in $\mathcal{A}_q$. Writing $g \in \Lambda$ for $\psi(f)$ (ψ as in Section 1), we thus seek (in Problem 2) the least $s = s(n, f)$ such that

$$X^{nc}(X^{sh} - 1) \in \Lambda g^n \tag{3.6}$$

Our plan will be to evaluate $s = s(n, f)$ in the first place along a certain arithmetic progression of n, and then to deal with the general case. Note that for $nc \equiv 0 \pmod{h}$, (3.6) is equivalent to

$$X^{nc}(X^{sh} - 1) \in I_n := R \cap \Lambda g^n. \tag{3.7}$$

In the next section we investigate the variation of I_n with n; in Section 5 this will yield our main result (Theorem 2) about Problem 2.

4. The sequence $I_n = R \cap \Lambda g^n$

Throughout $g \in \Lambda$, g neither 0 nor a unit. For $n \geq 0$ let $I_n = R \cap \Lambda g^n$. Then $I_n \lhd R$. We have $I_0 = R$, and $I_{n+1} \subseteq I_n$ for all n. Moreover, as $R =$ centre of Λ, we have $I_m I_n \subseteq I_{m+n}$ for all $n, m \geq 0$. By Sections 1–2, $\Lambda/\Lambda g^n$ is a finite set, hence is torsion as a left R-module. It is easily checked that $I_n = \operatorname{ann}_R(\Lambda/\Lambda g^n)$ for all $n \geq 0$. To summarise, we have

$$I_{n+1} \subseteq I_n,\ I_m I_n \subseteq I_{m+n},\ I_n = \operatorname{ann}_R(\Lambda/\Lambda g^n). \tag{4.1}$$

To these relations we may also append

https://doi.org/10.1017/S0013091500019970 Published online by Cambridge University Press

$$I_n \neq 0 \quad \text{and} \quad \bigcap_{n \in \mathbf{N}} I_n = 0. \tag{4.2}$$

For certainly $\det(g) \in I_1$ and $\det(g^n) = (\det g)^n \in I_n$, while $\det(g) \neq 0$; here $\det(g)$ is as in Section 1. Also if $r \neq 0$ lies in I_n then $\deg r \geq n \deg(g) \geq n$; hence only 0 lies in $\bigcap_{n \in \mathbf{N}} I_n$.

We now analyse the sequence I_n in more detail, using (4.1). Choose any free R-module basis $\beta_1, \ldots, \beta_{h^2}$ of Λ, and let $\mathbf{G}$ be the matrix describing the R-linear map $\lambda \mapsto \lambda g$ on Λ with respect to $\beta_1, \ldots, \beta_{h^2}$. Then, with respect to this basis, $\mathbf{G}^n$ describes $\lambda \mapsto \lambda g^n$. Now let F be the field $\mathbb{F}_p(X^h)$ of fractions of R, with algebraic closure $\overline{F}$. Suppose that, over $\overline{F}$, the characteristic polynomial $\det(x\mathbf{I} - \mathbf{G})$ of $\mathbf{G}$ factorises as $\prod_{j \leq h^2}(x - \mu_j)$; we put $\hat{F} = F(\mu_1, \ldots, \mu_{h^2})$, a finite extension of F. Since R is integrally closed in F, the eigenvalues $\mu_1, \ldots, \mu_{h^2}$ of $\mathbf{G}$ lie in $\hat{R}$, the integral closure of R in $\hat{F}$. Also, by the Akizuki-Krull theorem [5], $\hat{R}$ is a Dedekind domain, since R is a P.I.D.

Now let $\mathbf{J}$ be the Jordan canonical form of $\mathbf{G}$. There is an element $\mathbf{U}$ in $\mathrm{GL}(h^2, \hat{F})$ such that

$$\mathbf{U}^{-1}\mathbf{G}\mathbf{U} = \mathbf{J}, \mathbf{U}^{-1}\mathbf{G}^n\mathbf{U} = \mathbf{J}^n, \quad \text{for all } n \geq 0. \tag{4.3}$$

Now there is some $\rho \in \hat{R}, \rho \neq 0$, such that $\mathbf{W} := \rho\mathbf{U}$ has all entries in $\hat{R}$. We thus have

$$\mathbf{G}^n\mathbf{W} = \mathbf{W}\mathbf{J}^n \qquad (\forall n \geq 0) \tag{4.4}$$

in $M_{h^2}(\hat{R})$, the $h^2 \times h^2$-matrix ring over $\hat{R}$. Here $\mathbf{G}^n, \mathbf{W}, \mathbf{J}^n$ lie in $M_{h^2}(\hat{R})$, none of them having zero determinant. Recall that the object of this chapter is to evaluate the ideals I_n in R. It is clear that $r \in I_n$ if and only if

$$r\mathbf{I} = \mathbf{X}\mathbf{G}^n \tag{4.5}$$

for some $\mathbf{X} \in M_{h^2}(R)$, $\mathbf{I}$ being the identity element of $M_{h^2}(R)$. Using (4.4) and (4.5), and working in $M_{h^2}(\hat{R})$, we see that, for $r \in R$, we have $r \in I_n$ if and only if there is an equation

$$r\mathbf{W} = \mathbf{Y}\mathbf{W}\mathbf{J}^n \qquad \text{in } M_{h^2}(\hat{R}). \tag{4.6}$$

For, if $\mathbf{Y}$ satisfies (4.6) then $r\mathbf{I} = \mathbf{Y}\mathbf{G}^n$, so that the entries of $\mathbf{Y}$ lie in $F \cap \hat{R} = R$, while if $\mathbf{X} \in M_{h^2}(R)$ satisfies (4.5), then $\mathbf{Y} = \mathbf{X}$ satisfies (4.6).

Since $\hat{F}$ and F have characteristic p, we have

$$\mathbf{J}^n = \Delta^n \qquad \text{whenever } p^{h^2} | n, \tag{4.7}$$

where $\Delta = \mathrm{diag}(\mu_1, \ldots, \mu_{h^2})$. This observation allows us to evaluate I_n for $p^{h^2} | n$, by means of (4.6). Indeed, for such n, (4.6) shows that $r \in I_n$ if and only if

https://doi.org/10.1017/S0013091500019970 Published online by Cambridge University Press

$$r\mathbf{W} = \mathbf{Y}\mathbf{W}\Delta^n \qquad \text{holds in } M_{h^2}(\hat{R}). \tag{4.8}$$

We now work with the entries of the matrices in (4.8). Given r, n, (4.8) is soluble if and only if

$$rw_{ik} = \sum_j y_{ij} w_{jk} \mu_k^n \tag{4.9}$$

for all $i, k \leq h^2$, for some $y_{ij} \in \hat{R}$ $(i, j \leq h^2)$. Let Γ_k be the $\hat{R}$-ideal generated by the kth column of $\mathbf{W}$; then $\Gamma_k \neq 0$ as $\det(\mathbf{W}) \neq 0$. Clearly (4.9) holds if and only if

$$rw_{ik} \in \mu_k^n \Gamma_k, \qquad \forall i, k \leq h^2 \tag{4.10}$$

For any fixed k we first vary i in (4.10), and then vary k, obtaining

$$r \in I_n \Leftrightarrow r \in R \cap \bigcap_{k \leq h^2} (\mu_k^n \hat{R}). \tag{4.11}$$

Hence, provided that $p^{h^2} | n$, we have

$$I_n = R \cap \bigcap_{k \leq h^2} (\mu_k^n \hat{R}). \tag{4.12}$$

Equation (4.12) allows us to obtain, for $n \in p^{h^2}\mathbb{N}$, the factorisation of I_n in terms of maximal ideals in R. For $\hat{F}$ is a finite normal extension of F, being the splitting field over F of the characteristic polynomial of $\mathbf{G}$. Possibly $\hat{F}/F$ is inseparable; this does not matter, since, by [11], every maximal ideal $\mathfrak{p}$ of R decomposes in $\hat{R}$ as

$$\mathfrak{p}\hat{R} = (\mathfrak{P}_1 \ldots \mathfrak{P}_m)^{e(\mathfrak{p})}, \tag{4.13}$$

where $m = m(\mathfrak{p}) \in \mathbb{N}$, $e(\mathfrak{p}) \in \mathbb{N}$ and $\mathfrak{P}, \ldots, \mathfrak{P}_m$ are distinct maximal ideals in $\hat{R}$. (Here the residue field $R/\mathfrak{p}$ is finite, hence a perfect field.) Because of this, the maximal ideals $\mathfrak{p}$ occurring as factors of I_n for any $n \geq 1$, are precisely those which divide $\det(g)$. Let $S \neq \varnothing$ be the latter finite set of maximal ideals of R. Then, for $k \leq h^2$ we have

$$\mu_k \hat{R} = \prod_{\mathfrak{p} \in S} \prod_{\mathfrak{P} | \mathfrak{p}\hat{R}} \mathfrak{P}^{a(k, \mathfrak{P})}, \tag{4.14}$$

where the $a(k, \mathfrak{P}) \geq 0$ lie in $\mathbb{Z}$. Thus, for $n \in \mathbb{N}$, we have

$$\bigcap_{k \leq h^2} \mu_k^n \hat{R} = \prod_{\mathfrak{p} \in S} \prod_{\mathfrak{P} | \mathfrak{p}\hat{R}} \mathfrak{P}^{\max_k \{n a(k, \mathfrak{P})\}}. \tag{4.15}$$

From this, when $n \in p^{h^2}\mathbb{N}$, we have by (4.12),

https://doi.org/10.1017/S0013091500019970 Published online by Cambridge University Press

$$I_n = \prod_{\mathfrak{p} \in S} \mathfrak{p}^{b(n,\mathfrak{p})}, \tag{4.16}$$

where

$$b(n, \mathfrak{p}) = \max\{[na(k, \mathfrak{P})/e(\mathfrak{p})]^+; k \leq h^2, \mathfrak{P} | \mathfrak{p}\hat{R}\}. \tag{4.17}$$

Here, for $y \in \mathbb{R}$, $[y]^+ = \min\{m \in \mathbb{Z}; m \geq y\}$.

When $n \in \mathbb{N}$ is not in $p^{h^2}\mathbb{N}$ we *define* $b(n, \mathfrak{p})$ via (4.16). For such n, (4.17) no longer holds, but we still have, for all $\mathfrak{p} \in S$, and all $n \in \mathbb{N}$,

$$b(n+1, \mathfrak{p}) \geq b(n, \mathfrak{p}) \geq 1, b(m+n, \mathfrak{p}) \leq b(m, \mathfrak{p}) + b(n, \mathfrak{p}) \tag{4.18}$$

for all $m, n \geq 1$, while $b(n, \mathfrak{p}) \to \infty$ as $n \to \infty$, by (4.17).

From (4.18) an elementary argument shows that $n^{-1}b(n, \mathfrak{p})$ tends to a finite limit $l(\mathfrak{p})$ as $n \to \infty$. Indeed,

$$l(\mathfrak{p}) = \inf\{n^{-1}b(n, \mathfrak{p}); n \in \mathbb{N}\}. \tag{4.19}$$

Letting $n \to \infty$ through $p^{h^2}\mathbb{N}$, we see from (4.17) that

$$l(\mathfrak{p}) = \max\{a(k, \mathfrak{P})/e(\mathfrak{p}); k \leq h^2, \mathfrak{P} | \mathfrak{p}\hat{R}\} \tag{4.20}$$

In more precise terms, let $n \in \mathbb{N}$, and write $n = dp^{h^2} + r, d \geq 0, 0 \leq r < p^{h^2}$. Then, by (4.18), for $\mathfrak{p} \in S$,

$$\begin{aligned} b(dp^{h^2}, \mathfrak{p}) \leq b(n, \mathfrak{p}) &\leq b(dp^{h^2}, \mathfrak{p}) + b(r, \mathfrak{p}) \\ &\leq b(dp^{h^2}, \mathfrak{p}) + p^{h^2} b(1, \mathfrak{p}), \end{aligned}$$

while $b(dp^{h^2}, \mathfrak{p})$ can be calculated via (4.17). This gives

$$l(\mathfrak{p}) \leq \frac{b(n, \mathfrak{p})}{n} \leq l(\mathfrak{p}) + n^{-1}L(\mathfrak{p}) \tag{4.21}$$

for all $n \in \mathbb{N}, \mathfrak{p} \in S$, where $L(\mathfrak{p}) \geq 0$ depends only on $\mathfrak{p} \in S$.

5. Growth of $s(n, f)$ with n

In connection with Problem 2 we study the function $s(n, f)$ for any fixed $f \in \mathcal{A}_q$, $\deg(f) \geq p$. We put $s(n) = s(n, f)$ for $n \in \mathbb{N}$. Thus $s(n)$ is the least $s \in \mathbb{N}$ such that $f^{(n)}(T)$ splits completely in $\mathbb{F}_{q^s}[T]$. Since $f^{(n)}(0) = 0$ we clearly have $s(n) | s(n+1)$ for all n. In particular $s(n)$ is non-decreasing. We prove

https://doi.org/10.1017/S0013091500019970 Published online by Cambridge University Press

Theorem 2. (a) *There is a positive constant* α_f *such that* $\liminf_{n\to\infty}\{\frac{s(n)}{n}\} = \alpha_f$ *and* $\limsup_{n\to\infty}\{\frac{s(n)}{n}\} = p\alpha_f$.

(b) *For* $n \geq n_o(f)$ *we have* $s(n+1) = s(n)$ *or* $ps(n)$.

Remark. Since $s(n) \uparrow$, it suffices to prove Theorem 2(a) as $n \to \infty$ through $\nu\mathbb{N}$, for some $\nu \in \mathbb{N}$, arbitrarily chosen but fixed. We adopt this approach; to begin with we merely assume $\nu \in h\mathbb{N}$ where $q = p^h$, but later impose a stronger condition on ν which suffices to obtain Theorem 2(a); Theorem 2(b) will then be a simple corollary. Subject to later refinement, we now assume that $\nu \in h\mathbb{N}$ and $n \in \nu\mathbb{N}$. By (3.7), with $g = \psi(f)$, (ψ as in Section 1), we have: $s(n)$ is the least $s \in \mathbb{N}$ such that $X^{nc}(X^{hs} - 1) \in I_n$, which, by (4.16) is equivalent to

$$X^{nc}(X^{hs} - 1) \equiv 0 (\mathrm{mod}\ \mathfrak{p}^{b(n,\mathfrak{p})})\ (\forall \mathfrak{p} \in \Gamma^*), \tag{5.1}$$

where $\Gamma^* = \{\text{maximal } \mathfrak{p} \lhd R;\ \mathfrak{p} | \det(g)\}$. Let $\Gamma = \Gamma^* \backslash \{\mathfrak{p}_0\}$, where $\mathfrak{p}_0 = X^h R$. By hypothesis, $\Gamma \neq \varnothing$. Then (5.1) is equivalent to

$$\begin{cases} X^{nc} \in \mathfrak{p}_0^{b(n,\mathfrak{p}_0)} \\ X^{hs} \equiv 1 \bmod \mathfrak{p}^{b(n,\mathfrak{p})}, \quad (\forall \mathfrak{p} \in \Gamma). \end{cases} \tag{5.2}$$

It is easy to see that the condition $X^{nc} \in \mathfrak{p}_0^{b(n,\mathfrak{p}_0)}$ holds automatically for $n \in h\mathbb{N}$, imposing no condition on s. Hence, by the Chinese remainder theorem, it follows from (5.2) that

$$s(n) = l.c.m.\{\sigma(n, \mathfrak{p});\ \mathfrak{p} \in \Gamma\}, \tag{5.3}$$

where

$$\sigma(n, \mathfrak{p}) = \min\{s \in \mathbb{N};\ X^{sh} \equiv 1 (\mathrm{mod}\ \mathfrak{p}^{b(n,\mathfrak{p})})\}. \tag{5.4}$$

The next, elementary, result leads simply to the evaluation of $\sigma(n, \mathfrak{p})$ for $n \in h\mathbb{N}$.

Lemma 5.1. *Let* $\mathfrak{p}$ *be a maximal ideal of* R, $p \neq X^h R$. *Then*

(i) *the least* $S = S_1(\mathfrak{p})$ *in* $\mathbb{N}$ *with* $X^{hS_1(\mathfrak{p})} \equiv 1 (\mathrm{mod}\ \mathfrak{p})$ *is not divisible by* p.

(ii) *With* $S_1(\mathfrak{p})$ *as in* (i), *let* $\mathfrak{p}^{w(\mathfrak{p})} \| X^{hS_1(\mathfrak{p})} - 1$. *(Thus* $w(\mathfrak{p}) \in \mathbb{N}$.*) If* $u \in \mathbb{N} \backslash p\mathbb{N}$ *and* $l \geq 0$, *then* $\mathfrak{p}^{p^l w(\mathfrak{p})} \| X^{hup^l S_1(\mathfrak{p})} - 1$.

Proof. (i) $R/\mathfrak{p}$ is a finite field, of order p^a, say, $a \in \mathbb{N}$, while $X^h (\mathrm{mod}\ \mathfrak{p})$ is non-zero in $R/\mathfrak{p}$. The order e of $X^h (\mathrm{mod}\ \mathfrak{p})$ in $(R/\mathfrak{p})^*$ divides $p^a - 1$, and so is not a multiple of p. Clearly $e = S_1(\mathfrak{p})$.

(ii) Put $Y = X^h$, $S_1 = S_1(\mathfrak{p})$, $w = w(\mathfrak{p})$. By hypothesis, $\mathfrak{p}^w \| Y^{S_1} - 1$. Put $\mathfrak{p} = \pi R$. Then, in R, $Y^{S_1} = 1 + \epsilon\pi^w$, where $\epsilon \in R \backslash \mathfrak{p}$.

https://doi.org/10.1017/S0013091500019970 Published online by Cambridge University Press

If $u \in \mathbb{N}\backslash p\mathbb{N}$ we have

$$Y^{uS_1} = (1+\epsilon\pi^w)^u = 1 + u\epsilon\pi^w + v\pi^{2w}, \qquad (v \in R)$$

and so has the form $1+\epsilon_*\pi^w$, $\epsilon_* \in R\backslash\mathfrak{p}$. Now let $l \geq 0$. Since R has characteristic p, we have $Y^{p^l uS_1} = (1+\epsilon_*\pi^w)^{p^l} = 1+\epsilon_l\pi^{wp^l}$, with $\epsilon_l \in R\backslash\mathfrak{p}$. This proves the lemma.

We now apply Lemma 5.1 with $\mathfrak{p} \in \Gamma$, in order to find $\sigma(n, \mathfrak{p})$ of (5.4). The order of $X^h \pmod{\mathfrak{p}^r}$ for $r \in \mathbb{N}$, $\mathfrak{p} \in \Gamma$, is clearly $S_1(\mathfrak{p})p^{t(r,\mathfrak{p})}$, where $t(r, \mathfrak{p}) = [\log_p(\frac{r}{w(\mathfrak{p})})]^+$. Hence, for $\mathfrak{p} \in \Gamma$, $n \in h\mathbb{N}$, $\sigma(n, \mathfrak{p}) = S_1(\mathfrak{p})p^{t(b(n,\mathfrak{p}),\mathfrak{p})}$, which, by (5.2), yields

$$s(n) = \beta_f p^{H(n)} \tag{5.5}$$

where $\beta_f = l.c.m\{S_1(\mathfrak{p}); \mathfrak{p} \in \Gamma\}$, and

$$H(n) = \left[\log_p\left(\max_{\mathfrak{p}\in\Gamma}\left\{\frac{b(n,\mathfrak{p})}{w(\mathfrak{p})}\right\}\right)\right]^+. \tag{5.6}$$

We recall that (5.5) and (5.6) hold for $n \in h\mathbb{N}$. If we now further assume that $n \in hp^{h^2}\mathbb{N}$ we may also use (4.17) to evaluate $b(n, \mathfrak{p})$. Now suppose, additionally, that

$$v = hp^{h^2}\prod_{\mathfrak{p}\in\Gamma}(e(\mathfrak{p}).w(\mathfrak{p})) \tag{5.7}$$

Then for $n \in v\mathbb{N}$ and $p \in \Gamma$, (4.17) gives

$$\frac{b(n,\mathfrak{p})}{w(\mathfrak{p})} = \max\left\{\frac{na(k,\mathfrak{P})}{w(\mathfrak{p})e(\mathfrak{p})}; k \leq h^2, \mathfrak{P}/\mathfrak{p}\hat{R}\right\}, \tag{5.8}$$

so that $H(n)$ of (5.6) is given by

$$H(n) = \left[\log_p\left(\max\left\{\frac{na(k,\mathfrak{P})}{w(\mathfrak{p})e(\mathfrak{p})}; \mathfrak{p} \in \Gamma, k \leq h^2, \mathfrak{P}|\mathfrak{p}\hat{R}\right\}\right)\right]^+. \tag{5.9}$$

Putting

$$\gamma_f = v\max\left\{\frac{a(k,\mathfrak{P})}{w(\mathfrak{p})e(\mathfrak{p})}; \mathfrak{p} \in \Gamma, \mathfrak{P}|\mathfrak{p}\hat{R}, k \leq h^2\right\}, \tag{5.10}$$

we have, for $n \in v\mathbb{N}$,

$$H(n) = \left[\log_p\left(\frac{n}{v}\gamma_f\right)\right]^+. \tag{5.11}$$

https://doi.org/10.1017/S0013091500019970 Published online by Cambridge University Press

For $r \in \mathbb{R}$ let $((r)) = r - \max\{m \in \mathbb{Z};\ m \leq r\}$, so that $0 \leq ((r)) < 1$. Clearly

$$[r]^+ = r + ((-r)). \tag{5.12}$$

Thus, in (5.11) we have

$$p^{H(n)} = n\gamma_f v^{-1} p^{((-\log_p(n\gamma_f v^{-1})))}, \tag{5.13}$$

and so, for $n \in v\mathbb{N}$, (5.5) gives

$$s(n) = n\alpha_f p^{((-\log_p(n\gamma_f v^{-1})))}, \tag{5.14}$$

where $\alpha_f = \beta_f \gamma_f v^{-1}$. We now show that *this* value of α_f is the α_f required for Theorem 2(a). Let $t \in \mathbb{N}$ be prime, $t \neq p$. Then $\log_p(t)$ is irrational, so that, by [1], the sequence $u_m = ((m \log_p(t)))$, $(m \in \mathbb{N})$ is uniformly distributed in $[0, 1)$. This is certainly enough to show that the numbers $((-\log_p(n\gamma_f v^{-1})))$, $n \in v\mathbb{N}$, are everywhere dense in $[0, 1)$. In view of the remarks following the statement of Theorem 2, this suffices to prove Theorem 2(a).

Finally, we turn to the proof of Theorem 2(b). In fact, the preceding analysis permits the proof of a rather stronger result, which we formulate as follows.

Lemma 5.2. *For $n \in \mathbb{N}$, let a_n be a sequence in $\mathbb{N}$ with $n^{-1}a_n \to 0$. Then, for sufficiently large n, either $s(n + a_n) = s(n)$ or $s(n + a_n) = ps(n)$.*

Proof. Let v satisfy (5.7). We assume that $n > v$ and put $k_n = [nv^{-1}]^+ - 1$, $l_n = [(n + a_n)v^{-1}]^+$. By hypothesis $l_n k_n^{-1} \to 1$ as $n \to \infty$. We clearly have $s(vk_n) | s(n)$ | $s(n + a_n) | s(vl_n)$ in $\mathbb{N}$ whenever $n > v$. It therefore suffices to prove that $\lambda_n := s(vl_n)/s(vk_n) = 1$ or p for all large n.

In view of (5.5), λ_n is a power of p and lies in $\mathbb{N}$, i.e., $\lambda_n = p^{\mu_n}(\mu_n \geq 0$ in $\mathbb{Z})$.

Also, by (5.14),

$$\lambda_n = l_n k_n^{-1} p^{v_n - u_n}, \quad \text{where } v_n, u_n \in [0, 1).$$

Since $l_n k_n^{-1} \to 1$ we certainly have $\lambda_n < p + 1$ for all large n. For such n, μ_n is 0 or 1, and this proves the lemma.

6. A special case of Theorem 2

We take $q = p^h$ and $f(T) = T^p - T \in \mathcal{A}_q$. For this f, we shall calculate all terms $s(n, f) = s(n)$ of the sequence occurring in Theorem 2. For this purpose we shall use elementary linear algebra over $\mathbb{F}_p$; the special arguments used here will not work for the general f in $\mathcal{A}_q$. We define maps $\mathbf{L}, \mathbf{P}, \mathbf{Q}$ on $\overline{\mathbb{F}}_p$ via

https://doi.org/10.1017/S0013091500019970 Published online by Cambridge University Press

$$\mathbf{L}: \xi \mapsto f(\xi), \mathbf{P}: \xi \mapsto \xi^p, \mathbf{Q}: \xi \mapsto \xi^q \tag{6.1}$$

All three are $\mathbb{F}_p$-linear, i.e., lie in $\mathcal{H} = \mathrm{Hom}_{\mathbb{F}_p}(\overline{\mathbb{F}_p}, \overline{\mathbb{F}_p})$, while $\mathbf{Q} = \mathbf{P}^h$, $\mathbf{P}$ and $\mathbf{Q}$ are both units in $\mathcal{H}$, and $\mathbf{L}$ is surjective, but $\ker \mathbf{L} = \mathbb{F}_p$. For $n \in \mathbb{N}$ let $V_n = \ker \mathbf{L}^n$. Then $V_n = \{\text{zeros of } f^{(n)}(T)\}$ and $\#V_n = p^n$, so that $\dim_{\mathbb{F}_p} V_n = n$. Also the splitting field $\sum_n$ of $f^{(n)}(T)$ over $\mathbb{F}_p$ is $\mathbb{F}_q(V_n)$, so that

$$s(n) = [\mathbb{F}_q(V_n) : \mathbb{F}_q]. \tag{6.2}$$

Now, in $\mathcal{H}$, $\mathbf{P} = \mathbf{I} + \mathbf{L}$, where $\mathbf{I}$ is the identity map. Writing $h = p^t m$ with $t \geq 0$ and $m \in \mathbb{N} \backslash p\mathbb{N}$, we have, for $r \geq 0$,

$$\mathbf{Q}^{p^r} = (\mathbf{I} + \mathbf{L}^{p^{r+t}})^m. \tag{6.3}$$

Now define a sequence v_n $(n \in \mathbb{N})$ in $\overline{\mathbb{F}_p}$ as follows. Take $v_1 = 1$, and for every $n \in \mathbb{N}$ let v_{n+1} be arbitrary, subject to $\mathbf{L}v_{n+1} = v_n$. Then for every $n \in \mathbb{N}$ it is clear that $v_1, \ldots, v_n$ is an $\mathbb{F}_p$-basis for $V_n = \ker \mathbf{L}^n$. It is convenient to put $v_j = 0$ for $j \leq 0$. Then $\mathbf{L}^a(v_b) = 0$ whenever $a \geq 0$ and $a \geq b$ in $\mathbb{Z}$.

Now let $n \in \mathbb{N}$ and suppose that $p^{r+t} \geq n$. Then clearly $\mathbf{Q}^{p^r} v_j = v_j$ for all $j \leq n$. Hence $\mathbf{Q}^{p^r}$ fixes $\sum_n = \mathbb{F}_q(V_n)$ pointwise; it follows that $s(n) | p^r$, i.e., $s(n)$ is a p-power. Suppose now that $p^{r+t} < n$. Write $c = p^{r+t}$. Then $\mathbf{Q}^{p^r} = (\mathbf{I} + \mathbf{L}^c)^m = \mathbf{I} + \sum_{j=1}^m a_j \mathbf{L}^{cj}$, with $a_j \in \mathbb{F}_p$ and $a_1 = m (\mathrm{mod}\ p) \neq 0$.

Then $\mathbf{Q}^{p^r} v_n = v_n + \sum_{j \geq 1} a_j v_{n-jc} \neq v_n$, since $a_1 \neq 0$ and $n > n - c > 0$, while $v_n, \ldots, v_1$ are linearly independent over $\mathbb{F}_p$. Since $\mathbf{Q}^{p^r}$ does not fix v_n it cannot fix $\sum_n$ pointwise. Hence $s(n)$ cannot divide p^r if $p^{r+t} < n$.

From this we deduce that

$$s(n) = p^r, \quad r = \min\{l \geq 0; p^{t+l} \geq n\}. \tag{6.4}$$

Hence $s(n) = 1$ if $n \leq p^t$, while, for $n > p^t$, $p^t s(n) = p^{[\log_p(n)]^+}$.

Suppose now that $n > p^t$. Then if $l \in \mathbb{N}$ and $p^{l-1} < n \leq p^l$ we have $p^t s(n) = p^l$. This immediately gives

$$\alpha_f := \liminf \left\{ \frac{s(n)}{n} \right\} = p^{-t}$$

and

$$\limsup \left\{ \frac{s(n)}{n} \right\} = p^{1-t},$$

hence confirming Theorem 2 for f.

The reader will note that this easy verification of Theorem 2 for f is due to the particularly simple relation $\mathbf{L} + \mathbf{I} = \mathbf{P}$ for the f in question. For more general f in $\mathcal{A}_q$ there is no simple analogue of the formula $\mathbf{L} + \mathbf{I} = \mathbf{P}$, and thus no easy proof of Theorem 2. Nonetheless, this special example suggests that, in some sense, Theorem 2 is best possible.

REFERENCES

1. K. CHANDRASEKHARAN, *Introduction to analytic number theory* (Grundlehren, Band **148**, Springer Verlag, Berlin, 1968), Ch. 8.

2. P. M. COHN, *Algebra* (volume 3) (second edition, J. Wiley & Sons, 1991), Ch. 9.

3. M. DEURING, *Algebra* (Ergebnisse der Mathematik, Band **4**, Springer Verlag, Berlin, 1935).

4. K. IRELAND and M. ROSEN, *A classical introduction to modern number theory* (Springer Verlag, New York, 1982), Ch. 7.

5. N. JACOBSON, *Basic Algebra*, (I) (W. H. Freeman and Co., U.S.A., 1974), 282–284.

6. R. LIDL and H. NIEDERREITER, *Finite fields* (Encyclopaedia of Mathematics and its Applications, vol **20**; Addison-Wesley, Reading, Mass., 1983), Ch. 3.

7. G. MAURY and J. REYNAUD, *Ordres maximaux au sens de K. Asano* (Springer Lecture Notes in Maths., **808**, 1980), Ch. V.

8. O. ØRE, On a special class of polynomials, *Trans. Amer. Math. Soc.* **35** (1933), 559–584; corrigendum, **36** (1934), 275.

9. O. ØRE, Theory of non-commutative polynomials, *Ann. of Maths.* **34** (1933), 480–508.

10. I. REINER, *Maximal Orders* (Academic Press, London, 1975).

11. O. F. G. SCHILLING, *The Theory of Valuations* (Amer. Math. Soc., New York, 1950), 57, 101.

12. C. WILKERSON, A primer on the Dickson invariants, *Contemp. Math.* **19** (1983), 421–434.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF GLASGOW
GLASGOW G12 8QW
SCOTLAND