Primality testing with Gaussian periods

H. W. Lenstra jr. and Carl Pomerance

Abstract. We exhibit a deterministic algorithm that, for some effectively computable real number c, decides whether a given integer $n>1$ is prime within time $(\log n)^6\cdot(2+\log\log n)^c$. The same result, with 21/2 in the place of 6, was recently proved by Agrawal, Kayal, and Saxena. Our algorithm follows the same pattern as theirs, performing computations in an auxiliary ring extension of $\mathbf{Z}/n\mathbf{Z}$. We allow our rings to be generated by Gaussian periods rather than by roots of unity, which leaves us greater freedom in the selection of the auxiliary parameters and enables us to obtain a better run time estimate. The proof depends on newly developed results in analytic number theory and on the following theorem from additive number theory, which was provided by D. Bleichenbacher: if t is a real number with $0<t\le 1$, and S is an open subset of the interval $(0,t)$ with $\int_S \mathrm{d}x/x>t$, then each real number greater than or equal to 1 is in the additive semigroup generated by S. A byproduct of our main result is an improved algorithm for constructing finite fields of given characteristic and approximately given degree.

Key words: primality testing, constructing finite fields, Frobenius problem.

1991 Mathematics subject classification: 11Y11, 11N13, 11P70, 12Y05.

Acknowledgements. We thank several people for their assistance, in particular Daniel Bernstein, Daniel Bleichenbacher, Andrew Granville, Vsevolod Lev, Willem Jan Palenstijn, Francesco Pappalardi, Jonathan Pila, Peter Stevenhagen, Mark Watkins, and Alessandro Zaccagnini. The second author was supported in part by NSF grants DMS-0401422 and DMS-0703850.

Contents

1. Introduction

Our main result reads as follows.

Theorem 1. *There exists, for some effectively computable real number c, a deterministic algorithm that, given an integer n with $n > 1$, decides whether or not n is prime, and does so in time at most $(\log n)^6 \cdot (2 + \log\log n)^c$.*

We shall actually exhibit an algorithm with the stated properties. Its run time is measured in bit operations. The constant c is effectively computable in the sense that our proof of the existence of c, combined with the proofs in the papers we refer to, implicitly contains an algorithm for computing c.

The same result, but with the run time exponent 6 replaced by 21/2, was obtained by Agrawal, Kayal, and Saxena [2]; they also prove a result with run time exponent 15/2 in which c is not effectively computable, and they argue that the true run time exponent of their algorithm may reasonably be conjectured to equal 6. We achieve our exponent 6 not by proving their conjecture, but by modifying their algorithm.

A fundamentally new idea would be required to obtain a deterministic primality testing algorithm with run time exponent smaller than 6. For *probabilistic* primality tests the situation is different: Bernstein [7], also elaborating upon [2], exhibited a probabilistic algorithm that, for some effectively computable constant c_1, has the following property for each integer $n > 1$: given n, it correctly decides whether or not n is prime, and it does so in expected time at most $(\log n)^4 \cdot (2 + \log\log n)^{c_1 \log\log\log(22 \log n)}$.

Like [2], the present paper has an algebraic and an analytic component, addressing the correctness and the efficiency of the algorithm, respectively. By working harder on the algebra, we leave the algorithm greater freedom in the selection of auxiliary parameters, thus simplifying the analytic problem of obtaining a good run time estimate. Specifically, both the algorithm of [2] and our own algorithm perform computations in a suitable ring extension of the ring $\mathbf{Z}/n\mathbf{Z}$ of integers modulo n; if d denotes the "degree" of the extension, then the run time estimate becomes $d^{3/2} \cdot (\log n)^3$ times a lower order factor, and the problem of obtaining a small run time exponent boils down to proving a good upper bound for the smallest d that can be used. Agrawal *et al.* use the ring $(\mathbf{Z}/n\mathbf{Z})[X]/(X^d - 1)$, and find that the problem of accurately estimating the least usable value for d leads to an unsolved problem in analytic number theory. We select our ring extension from a much wider class, for which estimating d becomes feasible.

The ring extensions of $\mathbf{Z}/n\mathbf{Z}$ that we use shall be referred to as *pseudofields*. If n

is a prime number, then these pseudofields are in fact *finite fields*, and our construction of pseudofields is inspired by a construction of finite fields proposed by Adleman and Lenstra [1]. They describe a deterministic algorithm that, for certain effectively computable constants c_2 and c_3, has the following properties: given a prime number p and a positive integer D, it computes an irreducible polynomial f in $(\mathbf{Z}/p\mathbf{Z})[X]$ satisfying $D \le \deg f \le c_2 D \log p$, and it does so within time $(D + \log p)^{c_3}$; the ring $(\mathbf{Z}/p\mathbf{Z})[X]/(f)$ is then a finite field of given characteristic p of degree "close" to a given number D. Our construction improves upon this result for large enough D, as follows.

Theorem 2. *There are effectively computable positive integers c_4, c_5 and a deterministic algorithm with the following properties.*

- *Given a prime number $p > c_5$ and an integer D with $p/2 > D > (\log p)^{46/25}$, the algorithm computes an irreducible polynomial f in $(\mathbf{Z}/p\mathbf{Z})[X]$ satisfying*

$$D \le \deg f < D/\exp((\log D)^{3/5}(\log\log(3D))^{-3/2}).$$

- *Given a positive integer D and a prime number $p \le \max\{c_5, 2D\}$, the algorithm computes an irreducible polynomial f in $(\mathbf{Z}/p\mathbf{Z})[X]$ with $D \le \deg f < 2D$.*

In either case, the run time of the algorithm is at most $(D \log p) \cdot (2 + \log D + \log\log p)^{c_4}$.

Note that the run time of our algorithm is essentially linear in terms of the length of the output. Note too that for each integer $D \ge 1$, we have $(\log D)^{3/5}(\log\log 3D)^{-3/2} \ge 0$, so that in the case that $p > c_5$ and $D > (\log p)^{46/25}$, the upper bound for $\deg f$ in the theorem is at most $(1 + 1/\mathrm{e})D < 2D$. Thus, for all primes p and all integers $D > (\log p)^{46/25}$, the algorithm produces an irreducible polynomial in $(\mathbf{Z}/p\mathbf{Z})[X]$ with degree in $[D, 2D)$.

Adleman and Lenstra construct the finite field $(\mathbf{Z}/p\mathbf{Z})[X]/(f)$ by adjoining a system of *Gaussian periods* to $\mathbf{Z}/p\mathbf{Z}$. For Theorem 2, we use almost exactly the same construction, but we are much more careful in selecting the system of Gaussian periods, so that we are able to narrow the interval $[D, c_2 D \log p]$ for the degree down to $[D, 2D)$, and even narrower in a large range. The proof that an appropriate *period system* can be found, is the major technical hurdle we have to take; our desire that the constants in Theorem 1 and Theorem 2 be effectively computable has added to the difficulties. One of our auxiliary results is a weak but effective version of the Bombieri–Vinogradov inequality (see Lemma 10.2). Another one, which has independent interest, was provided by D. Bleichenbacher, who kindly allowed us to include his result and its proof. It reads as follows.

Theorem 3. *Suppose S is an open additively closed subset of the set of positive real numbers with $1 \notin S$. Then for each real number $t \in (0,1]$ one has $\int_{S\cap(0,t)} \mathrm{d}x/x \leq t$.*

A discussion of this theorem and its proof are found in Section 9.

In Section 2 we define pseudofields and period systems, and we state all properties of these concepts that go into our proofs. Taking these results for granted, we prove Theorem 1 in Section 3 and Theorem 2 in Section 4. In Sections 5–8 we prove the properties of pseudofields stated in Section 2. In Sections 10–12 we use analytic number theory to prove the existence result for period systems stated in Section 2.

In this paper, we write simply *ring* for *commutative ring*. As in [4, 19], a ring is required to have a unit element, a ring homomorphism is required to preserve the unit element, and a subring is required to contain the unit element. The ring of integers is denoted by $\mathbf{Z}$, and, for a prime number p, we write $\mathbf{F}_p$ for $\mathbf{Z}/p\mathbf{Z}$. For a ring R, we write R^* for the group of units of R, the *characteristic* $\operatorname{char} R$ is the non-negative integer n for which $n\mathbf{Z}$ is the kernel of the unique ring homomorphism $\mathbf{Z} \to R$, and we write $R[X]$ for the polynomial ring in one variable X over R; an element of $R[X]$ is *monic* if it has leading coefficient 1, the unit element of R.

Let S be a set, and let $f,\ g: S \to \mathbf{R}$ be two functions from S to the field $\mathbf{R}$ of real numbers such that for all $x \in S$ one has $g(x) \geq 0$. By the statement $f = O(g)$ we mean that there exists $c \in \mathbf{R}$ such that for all $x \in S$ one has $|f(x)| \leq c\cdot g(x)$, and by $f = \tilde{O}(g)$ we mean that there exists $c \in \mathbf{R}$ such that for all $x \in S$ one has $|f(x)| \leq g(x) \cdot \big(\log\max\{3, g(x)\}\big)^c$. We shall often apply this with S equal to a set of inputs to an algorithm, and $f(x)$ equal to the run time of the algorithm upon input x. For example, with the notation just introduced one expresses the run time estimates in Theorems 1 and 2 as $\tilde{O}\big((\log n)^6\big)$ and $\tilde{O}(D\log p)$, respectively.

Whenever we assert that a constant with certain properties exists, it will be effectively computable in the sense explained above; this is also valid for the constants implicit in our use of the O- and $\tilde{O}$-symbols. The same comment, *mutatis mutandis*, applies to the existence of algorithms. Except for Bernstein's algorithm mentioned above, all algorithms in the present paper are deterministic.

2. Pseudofields and period systems

Pseudofields. By a *pseudofield* we mean a pair (A, α) consisting of a ring A and an element $\alpha \in A$, such that for some integer $n > 1$, some integer $d > 0$, and some ring automorphism σ of A, the following conditions are satisfied:

$$\text{char } A = n, \tag{2.1}$$

$$\# A \leq n^d, \tag{2.2}$$

$$\sigma\alpha = \alpha^n, \tag{2.3}$$

$$\sigma^d \alpha = \alpha, \tag{2.4}$$

$$\sigma^{d/l}\alpha - \alpha \in A^* \text{ for each prime number } l \text{ dividing } d. \tag{2.5}$$

In Section 5 we shall prove the following result about pseudofields.

Proposition 2.6. *Let (A, α) be a pseudofield, and let n, d be as above. Then there is a unique monic polynomial $f \in (\mathbf{Z}/n\mathbf{Z})[X]$ with the property that there is a ring isomorphism $(\mathbf{Z}/n\mathbf{Z})[X]/(f) \cong A$ that maps the coset $(X \bmod f)$ to α. In addition, the degree of this polynomial equals d.*

The polynomial f from 2.6 and its degree d are called the *characteristic polynomial* and the *degree* of the pseudofield, respectively. The proposition implies that each element of A can in a unique way be written as $g(\alpha)$, where $g \in (\mathbf{Z}/n\mathbf{Z})[X]$ satisfies $\deg g < d$. This implies that equality holds in (2.2). It also implies that, as a ring, A is generated by α, so that the automorphism σ of A is uniquely determined by (2.3); we refer to it as the *Frobenius automorphism* of the pseudofield.

Example. If $n \in \mathbf{Z}$, $n > 1$, and $a \in \mathbf{Z}$, then the pair $(\mathbf{Z}/n\mathbf{Z}, a \bmod n)$ is a pseudofield if and only if one has $a^n \equiv a \bmod n$; for composite n, one often expresses this property by saying that n is a *pseudoprime to the base a*. In this example, the degree equals 1, the Frobenius automorphism is the identity, and the characteristic polynomial is $X - (a \bmod n)$.

Example. Let $n \in \mathbf{Z}$, $n > 1$, let r be a positive integer with $\gcd(r, n) = 1$, and denote by Φ_r the rth cyclotomic polynomial. Then the pair $\big((\mathbf{Z}/n\mathbf{Z})[X]/(\Phi_r), X \bmod \Phi_r\big)$ is a pseudofield if and only if $n \bmod r$ generates the group $(\mathbf{Z}/r\mathbf{Z})^*$. This pseudofield is closely related to the rings used in [2]. In this example, the degree equals $\varphi(r)$, where φ denotes Euler's function, the Frobenius automorphism maps each $(g \bmod \Phi_r)$ to $(g(X^n) \bmod \Phi_r)$, and Φ_r is the characteristic polynomial.

Finite fields yield pseudofields, as explained in the following result.

Proposition 2.7. *Let p be a prime number, let A be a ring of characteristic p, and let $\alpha \in A$. Then (A, α) is a pseudofield if and only if A is a finite field satisfying $A = \mathbf{F}_p(\alpha)$. In addition, if (A, α) is a pseudofield, and σ denotes its Frobenius automorphism, then for all $\beta \in A$ one has $\sigma\beta = \beta^p$.*

This proposition is proved in Section 5.

Primality testing with pseudofields. The following result shows that, for the purposes of primality testing, pseudofields can play the role that the rings $(\mathbf{Z}/n\mathbf{Z})[X]/(X^d - 1)$ play in [2].

Proposition 2.8. *Let (A, α) be a pseudofield of degree d with Frobenius automorphism σ, and let $n = \operatorname{char} A$. Suppose that for each $a = 1, 2, \ldots, \lfloor (d/3)^{1/2}(\log n)/\log 2 \rfloor$ one has $\alpha^n + a = (\alpha + a)^n$. Suppose also that one has $d > (\log n)^2/\big(3 \cdot (\log 2)^2\big)$, and that n has a prime factor greater than $(d/3)^{1/2}(\log n)/\log 2$. Then n is a power of a prime number.*

The proof of Proposition 2.8 in given in Section 6.

Algorithmic aspects of pseudofields. Proposition 2.6 shows that a pseudofield is, up to isomorphism, determined by its characteristic n and its characteristic polynomial f. We shall for algorithmic purposes always assume a pseudofield to be specified by the pair (n, f), the polynomial f being represented by its vector of coefficients; this applies in particular when a pseudofield forms part of the input or output of an algorithm. The pseudofield represented by (n, f) equals $\big((\mathbf{Z}/n\mathbf{Z})[X]/(f), X \bmod f\big)$, and its elements are represented as polynomials in $(\mathbf{Z}/n\mathbf{Z})[X]$ of degree smaller than the degree d of the pseudofield. It is well-known that there are algorithms that, given n, f, and two elements of $(\mathbf{Z}/n\mathbf{Z})[X]/(f)$, compute the sum and the product of these two elements within time $\tilde{O}(d \log n)$ (see [6]). As a consequence, testing the equality $\alpha^n + a = (\alpha + a)^n$ from 2.8 for a single value of a in $\mathbf{Z}/n\mathbf{Z}$ can be done in time $\tilde{O}\big(d(\log n)^2\big)$, and for about $(d/3)^{1/2}(\log n)/\log 2$ values of a in time $\tilde{O}\big((d^{1/2} \log n)^3\big)$. This time bound will equal the time bound $\tilde{O}\big((\log n)^6\big)$ from Theorem 1 if we use a pseudofield for which the degree d is, as a function of n, not too much larger than the lower bound $(\log n)^2/\big(3 \cdot (\log 2)^2\big)$ from 2.8. Thus, we are faced with the problem of constructing a pseudofield of given characteristic and approximately given degree.

The techniques that we develop for constructing pseudofields culminate in the following result. Let $n \in \mathbf{Z}$, $n > 1$. By a *period pair* for n we mean a pair (r, q) of integers with the properties

$$r \text{ is a prime number not dividing } n, \tag{2.9}$$

$$q \text{ divides } r-1 \text{ and } q > 1, \tag{2.10}$$

$$\text{the multiplicative order of } n^{(r-1)/q} \text{ modulo } r \text{ equals } q. \tag{2.11}$$

Further, a *period system* for n is a finite set $\mathcal{P}$ of period pairs for n such that

$$\gcd(q, q') = 1 \text{ whenever } (r, q), (r', q') \in \mathcal{P},\ (r, q) \neq (r', q'), \tag{2.12}$$

and the *degree* of $\mathcal{P}$ is $\prod_{(r,q)\in\mathcal{P}} q$.

Proposition 2.13. *There is an algorithm that, given an integer n with $n > 1$ and a period system $\mathcal{P}$ for n satisfying $n > \prod_{(r,q)\in\mathcal{P}} q$, either correctly declares n composite or constructs a pseudofield of characteristic n and degree $\prod_{(r,q)\in\mathcal{P}} q$, and that runs in time*

$$\tilde{O}\Big(\Big(\prod_{(r,q)\in\mathcal{P}} q + \sum_{(r,q)\in\mathcal{P}} q(r + \log n)\Big) \log n\Big).$$

The proof of Proposition 2.13 is given in Section 8.

If n is known to be prime, then the algorithm of Proposition 2.13 simplifies somewhat, and the term involving $(\log n)^2$ in the run time estimate may be omitted; in view of Proposition 2.7, this leads to the following result.

Proposition 2.14. *There is an algorithm that, given a prime number p and a period system $\mathcal{P}$ for p satisfying $p > \prod_{(r,q)\in\mathcal{P}} q$, constructs a monic irreducible polynomial $f \in \mathbf{F}_p[X]$ with $\deg f = \prod_{(r,q)\in\mathcal{P}} q$, and that runs in time*

$$\tilde{O}\Big(\Big(\prod_{(r,q)\in\mathcal{P}} q + \sum_{(r,q)\in\mathcal{P}} qr\Big) \log p\Big).$$

The proof of Proposition 2.14 is also given in Section 8.

The existence of period systems. Our final auxiliary result reads as follows.

Proposition 2.15. *There is an effectively computable positive integer c_5 such that, for each integer $n > c_5$ and each integer $D > (\log n)^{46/25}$, there exists a period system $\mathcal{P}$ for n consisting of pairs (r, q) with*

$$r < D^{6/11}, \quad q < D^{3/11}, \quad q \text{ prime},$$

and with degree d satisfying $D \le d < D + D/\exp((\log D)^{3/5}(\log\log(3D))^{-3/2})$. In particular, $d \in [D, 2D)$.

Proposition 2.15 is proved in Section 12 using Theorem 3 and some tools from analytic number theory which are developed in Sections 10 and 11. Proposition 2.15 is a key tool for showing that the algorithms of Theorems 1 and 2 perform as stated. In particular the number c_5 of Theorem 2 is the same as in 2.15.

In Section 10 we review some results concerning the distribution of primes in residue classes, and give a somewhat weaker, but effective version of the Bombieri–Vinogradov inequality. (See [25] for a similar result.) We also introduce our major tool, a theorem of Deshouillers and Iwaniec [16]. This result precedes Fouvry's theorem, and is interesting to us not only for its strength, but because it is effective in principle.

In Section 11 we show that there are many primes r with certain constraints on the primes in $r - 1$. For this we follow closely a paper of Balog [5]. This paper uses the same theorem of Fouvry as in [2], and also the Bombieri–Vinogradov theorem. To achieve effectively computable estimates, we use instead the Deshouillers–Iwaniec result and the effective Bombieri–Vinogradov inequality from Section 10.

3. The primality test

In this section we deduce Theorem 1 from the results stated in Section 2. We begin with a straightforward transformation of 2.15 into an algorithm for constructing period systems.

Algorithm 3.1. We describe an algorithm that takes as input an integer $n > 1$ and an integer $D > 0$, and that searches for a period system $\mathcal{P}$ for n with the properties listed in 2.15.

Step 1. Using a modified version of the sieve of Eratosthenes, sieving with prime powers rather than just with primes, compute the prime factorizations of all integers in $[1, 2D)$.

Step 2. For each prime number $r < D^{6/11}$ not dividing n, in increasing order, determine the set $\mathcal{Q}(r)$ of prime factors q of $r - 1$ that satisfy

$$q < D^{3/11}, \qquad n^{(r-1)/q} \not\equiv 1 \bmod r, \qquad q \notin \bigcup_{r' < r} \mathcal{Q}(r').$$

Put $\mathcal{Q} = \bigcup_r \mathcal{Q}(r)$ and, for each $q \in \mathcal{Q}$, put $r_q = r$ if $q \in \mathcal{Q}(r)$.

Step 3. If there is some integer in $[D, 2D)$ that is squarefree and composed solely of primes from $\mathcal{Q}$, let d be the least such integer, let $\mathcal{P}$ be the set of all pairs (r_q, q), with q

ranging over the prime factors of d, return $\mathcal{P}$, and halt. If no such integer exists, pronounce failure and halt.

This completes the description of Algorithm 3.1.

The constant c_5 in the following result is as in 2.15.

Proposition 3.2. *Algorithm* 3.1, *on input integers $n > 1$ and $D > 0$, successfully computes a period system for n with the properties listed in* 2.15 *if and only if such a period system exists, which is the case if $n > c_5$ and $D > (\log n)^{46/25}$; the runtime of the algorithm is $\tilde{O}(D + D^{6/11} \log n)$.*

Proof. The "if and only if" statement is clear from the algorithm, the second assertion is immediate from 2.15, and proof of the run time estimate is entirely straightforward. This proves 3.2.

Primality testing. We describe an algorithm that has the properties stated in Theorem 1. We let c_5 again be as in 2.15.

Algorithm 3.3. Given an integer $n > 1$, this algorithm decides whether or not n is prime.

Step 1. If $n \leq c_5$, find by trial division the least prime p dividing n, declare n prime or composite according as $n = p$ or $n \neq p$, and halt.

Step 2. Using the algorithm of [8], determine the largest $k \in \mathbf{Z}$ for which there exists $m \in \mathbf{Z}$ with $n = m^k$. If $k > 1$, declare n composite and halt.

Step 3. Using standard algorithms for computing elementary functions (cf. [6, 11]), compute an integer D satisfying

$$D - 2 < \max\{(\log n)^2 / \big(3 \cdot (\log 2)^2\big), (\log n)^{46/25}\} < D.$$

Next, using Algorithm 3.1, construct a period system $\mathcal{P}$ for n with the properties listed in 2.15. Put $d = \prod_{(r,q) \in \mathcal{P}} q$.

Step 4. Using standard algorithms for computing elementary functions (cf. [6, 11]), compute an integer b satisfying

$$b - 1 < (d/3)^{1/2} (\log n) / \log 2 < b + 1,$$

and test by trial division whether n has a divisor among 2, 3, $\ldots$, $\max\{d, b\}$. If it does, let p be the least such divisor, declare n prime or composite according as $n = p$ or $n \neq p$, and halt.

Step 5. Using the algorithm of 2.13, either declare n composite and halt, or construct a pseudofield (A,α) of characteristic n and degree d.

Step 6. For $a = 1, 2, \ldots, b$, test the equality $\alpha^n + a = (\alpha + a)^n$ in A. If all of these are valid, declare n prime and halt. If at least one fails to be valid, declare n composite and halt.

This completes the description of Algorithm 3.3.

Proof of Theorem 1. We prove that Algorithm 3.3 has the properties claimed in Theorem 1; that is, it terminates within time $\tilde{O}\big((\log n)^6\big)$, correctly declaring n prime or composite. Step 1 runs in time $O(1)$, and by [8], Step 2 runs in time $\tilde{O}(\log n)$. If the algorithm halts during one of these two steps, it is clearly correct. Assume otherwise, so that one has $n > c_5$ and n is not a proper power. The first part of Step 3 runs in time $O(\log n)$, and from $D > (\log n)^{46/25}$ and $D = O\big((\log n)^2\big)$ it follows, by 3.2, that the second part of Step 3 successfully computes a period system in time $\tilde{O}\big((\log n)^{23/11}\big)$. We have $d = O\big((\log n)^2\big)$, and from $d \geq 2^{\#\mathcal{P}}$ one obtains $\#\mathcal{P} = O\big(\log(2\log n)\big)$. Step 4 runs in time $\tilde{O}\big((\log n)^3\big)$ because $b = O\big((\log n)^2\big)$. If the algorithm halts in Step 4, it is clearly correct. Suppose otherwise. Then we have $n > d$, so by 2.13 and the inequalities in 2.15, Step 5 runs in time $\tilde{O}\big((\log n)^3\big)$. As we argued in Section 2, the test in Step 6 can be done in time $\tilde{O}\big((d^{1/2}\log n)^3\big)$, which is $\tilde{O}\big((\log n)^6\big)$. Since n passed Step 4, it has a prime divisor greater than $(d/3)^{1/2}(\log n)/\log 2$, so 2.8 implies that, if n passes the test in Step 6, it is a prime power; not being a proper power, it must be prime. If n does not pass the test in Step 6, then by 2.7 (with n in the role of p and $\alpha + a$ in the role of β) it cannot be a prime number. This concludes the proof of Theorem 1.

4. Constructing finite fields

In this section we prove Theorem 2. We begin with two lemmas that are used to deal with certain exceptional cases.

Lemma 4.1. *Let k be a finite field, r a prime number, h a non-negative integer, and $b \in k^*$ an element that is not an rth power in k^*. Assume that one has $\#k \equiv 1 \bmod 4$ if $r^h \equiv 0 \bmod 4$. Then $X^{r^h} - b$ is irreducible in $k[X]$.*

Proof. See [21, Theorem 3.75].

Lemma 4.2. *For any non-negative integer h, the polynomials $X^{2\cdot 3^h} + X^{3^h} + 1$ and $X^{10\cdot 3^h} + X^{5\cdot 3^h} + 1$ are irreducible in $\mathbf{F}_2[X]$. For any prime number p with $p \equiv 1 \bmod 4$, any non-negative integer h, and any $a \in \mathbf{F}_p$ satisfying $\left(\frac{a}{p}\right) = -1$, the polynomial $X^{2^h} - a$ is* *irreducible in* $\mathbf{F}_p[X]$*. For any prime number* p *with* $p \equiv -1 \bmod 4$ *there exists* $a \in \mathbf{F}_p$ *with* $\left(\frac{a^2+4}{p}\right) = -1$*, and for any such* a *and any non-negative integer* h *the polynomial* $X^{2^{h+1}} - aX^{2^h} - 1$ *is irreducible in* $\mathbf{F}_p[X]$*.*

Proof. In this proof, we denote algebraic closures by an overhead bar.

First let $p = 2$. Let b, $\alpha \in \bar{\mathbf{F}}_2$ satisfy $b^2 + b + 1 = 0$ and $\alpha^{3^h} = b$. Then $\mathbf{F}_2(\alpha)$ contains $\mathbf{F}_2(b)$, and the latter field has degree 2 over $\mathbf{F}_2$. By 4.1, one has $[\mathbf{F}_2(\alpha) : \mathbf{F}_2(b)] = 3^h$, and therefore $[\mathbf{F}_2(\alpha) : \mathbf{F}_2] = 2 \cdot 3^h$. Since α is a zero of $X^{2\cdot 3^h} + X^{3^h} + 1$, this polynomial is irreducible in $\mathbf{F}_2[X]$. Now let $\beta \in \bar{\mathbf{F}}_2$ satisfy $\beta^5 = \alpha$. Since $2^{2\cdot 3^h} \not\equiv 1 \bmod 5$, $\mathbf{F}_2(\alpha)^*$ contains no element of order 5, so that by 4.1, $[\mathbf{F}_2(\beta) : \mathbf{F}_2(\alpha)] = 5$. Thus, $[\mathbf{F}_2(\beta) : \mathbf{F}_2] = 10 \cdot 3^h$, and since β is a root of $X^{10\cdot 3^h} + X^{5\cdot 3^h} + 1$, this polynomial is irreducible in $\mathbf{F}_2[X]$.

Next let $p \equiv 1 \bmod 4$. In this case the Lemma is immediate from 4.1.

Finally suppose $p \equiv -1 \bmod 4$. If c is the least positive integer with $\left(\frac{c}{p}\right) = -1$, then one can write $(c - 1 \bmod p) = e^2$ with $e \in \mathbf{F}_p$, and $a = 2e$ then satisfies $\left(\frac{a^2+4}{p}\right) = \left(\frac{4}{p}\right) \cdot \left(\frac{c}{p}\right) = -1$. Next let b, $\alpha \in \bar{\mathbf{F}}_p$ satisfy $b^2 - ab - 1 = 0$ and $\alpha^{2^h} = b$. From $\left(\frac{a^2+4}{p}\right) = -1$ it follows that $X^2 - aX - 1$ is irreducible in $\mathbf{F}_p[X]$, so the field $\mathbf{F}_p(b)$, which is a subfield of $\mathbf{F}_p(\alpha)$, has degree 2 over $\mathbf{F}_p$. The product of b and its conjugate equals -1, which is not a square in $\mathbf{F}_p$, so b is not a square in $\mathbf{F}_p(b)$. Since one also has $\#\mathbf{F}_p(b) = p^2 \equiv 1 \bmod 4$, Lemma 4.1 implies $[\mathbf{F}_p(\alpha) : \mathbf{F}_p(b)] = 2^h$ and therefore $[\mathbf{F}_p(\alpha) : \mathbf{F}_p] = 2^{h+1}$. Since α is a zero of $X^{2^{h+1}} - aX^{2^h} - 1$, the latter polynomial is irreducible in $\mathbf{F}_p[X]$.

This proves 4.2.

We describe an algorithm that has the properties stated in Theorem 2.

Algorithm 4.3. Given a prime number p and a positive integer D, this algorithm attempts to construct an irreducible polynomial $f \in \mathbf{F}_p[X]$ with $D \leq \deg f < 2D$. We let c_5 be as in 2.15.

Step 1. [This step takes care of the case in which p is too small for 2.14 or for 3.2 to apply.] If $p = 2$, determine the least non-negative integer h with $2 \cdot 3^h \geq D$; if $2 \cdot 3^h < 2D$, return $f = X^{2\cdot 3^h} + X^{3^h} + 1$, and halt. Else, if $h \geq 2$, return $f = X^{10\cdot 3^{h-2}} + X^{5\cdot 3^{h-2}} + 1$, and halt. If $2 \cdot 3^h \geq 2D$ and $h < 2$, return $f = X + 1$ if $D = 1$, return $f = X^3 + X + 1$ if $D = 3$, and return $f = X^{10} + X^5 + 1$ if $7 \leq D \leq 9$, halting in each case. If $p \equiv 1 \bmod 4$ and $p \leq \max\{c_5, 2D\}$, determine the least positive integer a with $\left(\frac{a}{p}\right) = -1$ and the least non-negative integer h with $2^h \geq D$, return $f = X^{2^h} - a$, and halt. If $p \equiv -1 \bmod 4$ and $p \leq \max\{c_5, 2D\}$, determine the least positive integer a with $\left(\frac{a^2+4}{p}\right) = -1$ and the least non-negative integer h with $2^{h+1} \geq D$, return $f = X^{2^{h+1}} - aX^{2^h} - 1$, and halt.

Step 2. [In this case we have $p > c_5$ and $p > 2D$.] Apply Algorithm 3.1 to $n = p$ and D; if that algorithm pronounces failure, pronounce failure and halt. Otherwise, let $\mathcal{P}$ be the period system for p produced by Algorithm 3.1, apply the algorithm of 2.14 to $\mathcal{P}$, return the polynomial produced by the latter algorithm, and halt.

This completes the description of Algorithm 4.3.

Theorem 2 is now an immediate consequence of the following result.

Theorem 4.4. *Algorithm* 4.3*, on input of a prime number p and a positive integer D, runs in time $\tilde{O}(D \log p)$, and if it does not pronounce failure then it computes a monic irreducible polynomial $f \in \mathbf{F}_p[X]$ satisfying $D \leq \deg f < 2D$; in addition, it does not pronounce failure if $p \leq \max\{c_5, 2D\}$ or $D > (\log p)^{46/25}$. Further, in the case $p > c_5$ and $p/2 > D > (\log p)^{46/25}$, one has $D \leq \deg f < D + D/\exp((\log D)^{3/5}(\log\log(3D))^{-3/2})$.*

Proof. First suppose $p \leq \max\{c_5, 2D\}$. Then by 4.2 the algorithm halts in Step 1 and returns a polynomial f that is irreducible over $\mathbf{F}_p$ and that satisfies $D \leq \deg f < 2D$. From $p = O(D)$ one readily deduces that the computation of h in Step 1 and, if p is odd, a in Step 1 can be done in time $\tilde{O}(D)$. Next assume $p > \max\{c_5, 2D\}$. If $D > (\log p)^{46/25}$, then by 3.2 the algorithm successfully computes a period system for p, and if it successfully computes a period system, then by 2.14 it computes a polynomial f with the stated properties. The run time estimate for Step 2 is obtained in a routine manner from 3.2 and 2.14; note that the sum $\sum_{(r,q)\in\mathcal{P}} qr$ occurring in 2.14 is $\tilde{O}(D^{9/11})$, by the inequalities in 2.15. This proves 4.4.

5. Algebraic properties of pseudofields

In Section 2 we defined pseudofields, and the present section is devoted to their basic algebraic properties.

For a ring A, an element $\alpha \in A$, and a ring automorphism σ of A, we will have occasion to refer to the condition

(5.1) $\sigma\alpha$ belongs to the subring of A generated by α.

This condition is implied by condition (2.3), if n is a positive integer.

Proposition 5.2. *Let A be a ring, let $\alpha \in A$, let $d \in \mathbf{Z}_{>0}$, and let σ be a ring automorphism of A such that* (2.4), (2.5), *and* (5.1) *are satisfied. Then for any i, $j \in \mathbf{Z}$ with $i \not\equiv j \bmod d$ one has $\sigma^i\alpha - \sigma^j\alpha \in A^*$.*

Proof. Let $h \in \mathbf{Z}$, $h \notin d\mathbf{Z}$, and let $I = (\sigma^h\alpha - \alpha)$ be the A-ideal generated by $\sigma^h\alpha - \alpha$. The set $\{\beta \in A : \sigma^h\beta \equiv \beta \bmod I\}$ is a subring of A that contains α, so by (5.1) it contains $\sigma\alpha$; that is, one has $\sigma^{h+1}\alpha \equiv \sigma\alpha \bmod I$, so $\sigma(\sigma^h\alpha - \alpha)$ belongs to I, and therefore one has $\sigma I \subset I$. Since σ^d maps $\sigma^h\alpha - \alpha$ to itself, we actually have $\sigma I = I$, so for all $m \in \mathbf{Z}$ one has $\sigma^m I = I$. It follows that the set $H = \{m \in \mathbf{Z} : \sigma^m\alpha \equiv \alpha \bmod I\}$ is a subgroup of $\mathbf{Z}$. It contains d and h, where $h \notin d\mathbf{Z}$, so one has $H = d'\mathbf{Z}$ where d' is a divisor of d with $1 \le d' < d$. Choose a prime number l that divides d/d'. Then $d/l \in d'\mathbf{Z} = H$, so $\sigma^{d/l}\alpha - \alpha \in I$. Thus, by (2.5) the ideal I contains a unit, and therefore $I = A$. This implies $\sigma^h\alpha - \alpha \in A^*$. Now let $i, j \in \mathbf{Z}$, $i \not\equiv j \bmod d$. Then the integer $i - j$ does not belong to $d\mathbf{Z}$, so by the result just proved we have $\sigma^{i-j}\alpha - \alpha \in A^*$. Applying σ^j we find $\sigma^i\alpha - \sigma^j\alpha \in A^*$, as required. This proves 5.2.

Lemma 5.3. *Let A be a ring, let $k \in \mathbf{Z}_{\ge 0}$, and let $\alpha_1, \alpha_2, \ldots, \alpha_k \in A$ be such that $\alpha_i - \alpha_j \in A^*$ whenever $1 \le i < j \le k$. Then for each $g \in A[X]$ which vanishes at $\alpha_1, \alpha_2, \ldots, \alpha_k$, one has $g \in A[X] \cdot \prod_{i=1}^k (X - \alpha_k)$.*

Proof. Let $I_i = A[X] \cdot (X - \alpha_i)$, for $1 \le i \le k$. For $i \ne j$, the unit $\alpha_i - \alpha_j$ can be written as $-(X - \alpha_i) + (X - \alpha_j)$, so $I_i + I_j = A[X]$. This implies $\prod_{i=1}^k I_i = \bigcap_{i=1}^k I_i$, by [4, Proposition 1.10(i)]. From $X \equiv \alpha_i \bmod I_i$ one obtains $g \equiv g(\alpha_i) \bmod I_i$ for each $g \in A[X]$, so if each $g(\alpha_i)$ vanishes then one has $g \in \bigcap_{i=1}^k I_i = \prod_{i=1}^k I_i = A[X] \cdot \prod_{i=1}^k (X - \alpha_k)$, as required. This proves 5.3.

The following result summarizes the technical information on pseudofields that we shall need.

Proposition 5.4. *Let A be a ring, let $\alpha \in A$, and let the integers $n \in \mathbf{Z}_{>0}$, $d \in \mathbf{Z}_{>0}$ and the ring automorphism σ of A satisfy* (2.1), (2.2), (2.4), (2.5), *and* (5.1). *Then one has:*

(a) *for each $\beta \in A$ there are unique $a_0, a_1, \ldots, a_{d-1} \in \mathbf{Z}/n\mathbf{Z}$ with $\beta = \sum_{i=0}^{d-1} a_i\alpha^i$;*

(b) *one has $\#A = n^d$, and σ^d equals the identity;*

(c) *the polynomial $f = \prod_{i=0}^{d-1}(X - \sigma^i\alpha)$ belongs to the subring $(\mathbf{Z}/n\mathbf{Z})[X]$ of $A[X]$;*

(d) *the ring homomorphism $(\mathbf{Z}/n\mathbf{Z})[X] \to A$ sending X to α is surjective, and its kernel is generated by the polynomial f from* (c)*;*

(e) *if $I \subset A$ is an ideal, then one has $\sigma I \subset I$ if and only if there exists a divisor m of n such that $I = mA$;*

(f) *for each prime factor p of n there exists a unique residue class $(i \bmod d)$ such that for all $\beta \in A$ one has $\beta^p \equiv \sigma^i\beta \bmod pA$.*

Proof. It is clear that there is a unique ring homomorphism $\psi\colon(\mathbf{Z}/n\mathbf{Z})[X]\to A$ as in (d), and that it maps each $g\in(\mathbf{Z}/n\mathbf{Z})[X]$ to $g(\alpha)$. If $g\in\ker\psi$, then for each $i\in\mathbf{Z}$ one has $g(\sigma^i\alpha)=\sigma^i(g(\alpha))=\sigma^i(\psi(g))=0$, so by 5.2 and 5.3 one has $g\in A[X]f$, where f is as in (c). Since each non-zero $g\in A[X]f$ has degree at least d, this implies

$$\ker\psi\cap\big((\mathbf{Z}/n\mathbf{Z})+(\mathbf{Z}/n\mathbf{Z})X+\ldots+(\mathbf{Z}/n\mathbf{Z})X^{d-1}\big)=\{0\},$$

so that the restriction of ψ to $(\mathbf{Z}/n\mathbf{Z})+(\mathbf{Z}/n\mathbf{Z})X+\ldots+(\mathbf{Z}/n\mathbf{Z})X^{d-1}$ is injective. From (2.2) one now sees that it is surjective as well, which proves (a), the first statement of (b), and the surjectivity in (d). Since each element of A can be expressed in α, the second statement of (b) follows from (2.4). Applying (a) to $\beta=\alpha^d$, one finds $a_0, a_1, \ldots, a_{d-1}\in\mathbf{Z}/n\mathbf{Z}$ for which the polynomial $g=X^d-\sum_{i=0}^{d-1}a_iX^i$ belongs to $\ker\psi$; hence $g\in A[X]f$, and comparing degrees and leading coefficients one finds $g=f$. This implies (c). We have $\ker\psi=A[X]f\cap(\mathbf{Z}/n\mathbf{Z})[X]=(\mathbf{Z}/n\mathbf{Z})[X]f$, the latter equality because f is a monic polynomial in $(\mathbf{Z}/n\mathbf{Z})[X]$. This proves the remaining assertion of (d).

The "if"-part of (e) is clear. For the "only if"-part, let I be an ideal of A with $\sigma I\subset I$, and let $\bar{A}$ be the ring A/I. From $\sigma I\subset I$ it follows that σ induces a ring homomorphism $\bar\sigma\colon\bar{A}\to\bar{A}$. From (b) one sees that $\bar\sigma^d$ is the identity on $\bar{A}$, so $\bar\sigma$ is an automorphism of $\bar{A}$. Put $m=\operatorname{char}\bar{A}$. Then m divides n, and we have $mA\subset I$, so from (a) we see $\#\bar{A}\le\#A/mA=m^d$, with equality if and only if $mA=I$. We claim that (2.1), (2.2), (2.4), (2.5), and (5.1), with $\bar{A}$, m, d, $\bar\sigma$, and $\bar\alpha=(\alpha \bmod I)$ in the roles of A, n, d, σ, and α, are satisfied. For (2.2) we just proved this, (2.1) is true by definition, and (2.4), (2.5), and (5.1) follow from the corresponding properties of A, n, d, σ, and α. Hence, applying (b) to this new situation, we find $\#\bar{A}=m^d$, so that $mA=I$. This proves (e).

To prove (f), we replace, for notational convenience, n and A by p and A/pA, so that we may assume $n=p$. Let $\phi\colon A\to A$ be the ring homomorphism that maps each $\beta\in A$ to β^p, and let $g\in(\mathbf{Z}/n\mathbf{Z})[X]$ be such that $\sigma\alpha=g(\alpha)$. If $\rho\colon A\to A$ is any ring homomorphism with $\sigma\rho=\rho\sigma$, then one has $\sigma(\rho\alpha)=\rho(\sigma\alpha)=\rho(g(\alpha))=g(\rho\alpha)$. Applying this to $\rho=\phi$ and to $\rho=\sigma^i$, where $i\in\mathbf{Z}$, we obtain $\sigma(\phi\alpha)=g(\phi\alpha)$ and $\sigma(\sigma^i\alpha)=g(\sigma^i\alpha)$ and therefore $\sigma(\phi\alpha)\equiv\sigma(\sigma^i\alpha) \bmod (\phi\alpha-\sigma^i\alpha)A$. Hence, for any $i\in\mathbf{Z}$, the ideal $I=(\phi\alpha-\sigma^i\alpha)A$ satisfies $\sigma I\subset I$, so by (e) and the fact that n is prime one has $I=A$ or $I=nA=0$, so that $\phi\alpha-\sigma^i\alpha$ is either a unit or 0. From $\prod_{i=0}^{d-1}(\phi\alpha-\sigma^i\alpha)=f(\phi\alpha)=\phi(f(\alpha))=0^p=0$ we see that not all $\phi\alpha-\sigma^i\alpha$ can be units, so at least one of them is 0. Then one has $\phi\alpha=\sigma^i\alpha$, so $\phi=\sigma^i$ by (a). The uniqueness of $i \bmod d$ follows from 5.2. This proves 5.4.

We can now prove two propositions stated in Section 2.

Proof of Proposition 2.6. Let the notation and hypotheses be as in 2.6. Since (2.3) implies (5.1), Proposition 5.4 applies. The existence of f as in 2.6 follows from 5.4(d). No two distinct monic polynomials in $(\mathbf{Z}/n\mathbf{Z})[X]$ generate the same ideal, so f is unique. From 5.4(c) one sees $\deg f = d$. This proves 2.6.

Proof of Proposition 2.7. Let p, A, and α be as in 2.7. For the "if"-part, assume that A is a finite field with $A = \mathbf{F}_p(\alpha)$. Write $d = [A : \mathbf{F}_p]$ and define $\sigma : A \to A$ by putting $\sigma\beta = \beta^p$ for every $\beta \in A$. It is a standard property of finite fields that σ is a field automorphism of A of order d. Now (2.1)–(2.4) are obvious. If l is a prime number dividing d, then $\sigma^{d/l}$ is not the identity, so by $A = \mathbf{F}_p(\alpha)$ we have $\sigma^{d/l}\alpha \neq \alpha$; since A is a field, this implies (2.5).

To prove the "only if"-part and the last statement of 2.7, assume that (A, α) is a pseudofield. Write d for the degree and σ for the Frobenius automorphism. Since p is prime, the map $A \to A$ sending each β to β^p is a ring homomorphism. It agrees with σ on α, so by 5.4(a) on all of A, which is the last statement of 2.7. To prove that A is a field, we let $\beta \in A$, and we prove that β equals 0 or is a unit. Put $I = A\beta$. From $\sigma\beta = \beta^p$ one sees $\sigma I \subset I$, so by 5.4(e) and the fact that p is prime we have $I = A$ or $I = pA = 0$. In the first case, β is a unit, in the second case it equals 0. Thus, A is a field. By 5.4(a), it is finite, and one has $A = \mathbf{F}_p(\alpha)$. This completes the proof of 2.7.

6. Primality testing with pseudofields

In this section we prove 2.8. We begin with an elegant lemma.

Lemma 6.1. *Let R be a ring, and let G be a finite subgroup of R^* such that for each $\beta \in G$, $\beta \neq 1$, one has $\beta - 1 \in R^*$. Then G is cyclic.*

Proof. We may clearly assume $R \neq \{0\}$, so that we can choose a maximal ideal M of R. For each $\beta \in G$, $\beta \neq 1$, the unit $\beta - 1$ does not belong to M, so that β is not in the kernel of the natural group homomorphism $R^* \to (R/M)^*$. Hence the restriction of the latter map to G is injective, and G is isomorphic to its image in $(R/M)^*$. Since any finite subgroup of the multiplicative group of a field is cyclic, this implies 6.1.

The reader may enjoy proving 6.1 without using maximal ideals, for example by applying 5.3.

Let (A, α) be a pseudofield, and denote by n, d, and σ its characteristic, its degree, and its Frobenius automorphism, respectively. We let p be a prime divisor of n, and put

$R = A/pA$. We shall simply write α for the image of α in R, and σ for the automorphism of R induced by σ. Note that the conditions of Proposition 5.4, with R, α, p, d, σ in the roles of A, α, n, d, σ, are satisfied. As in the proof of 2.7, we have

$$\text{(6.2)} \qquad \text{if } \beta \in R \text{ satisfies } \sigma\beta \in R\beta, \text{ then } \beta = 0 \text{ or } \beta \in R^*,$$

by 5.4(e) applied to $I = R\beta$. We put

$$G = \{\beta \in R : \beta \neq 0, \, \sigma\beta = \beta^n\}.$$

For any $\beta \in G$, one has $\sigma\beta = \beta^n \in R\beta$, so $\beta \in R^*$ by (6.2). Since G is finite and closed under multiplication, and contains 1, it is a subgroup of R^*. Also, for any $\beta \in G$, $\beta \neq 1$, one has $\sigma\beta = \beta^n \equiv 1 \bmod R \cdot (\beta - 1)$, so $\sigma(\beta - 1) \in R \cdot (\beta - 1)$ and $\beta - 1 \in R^*$, again by (6.2). Thus, Lemma 6.1 implies

$$\text{(6.3)} \qquad G \text{ is a cyclic subgroup of } R^*.$$

Lemma 6.4. *If* $\#G > n^{\sqrt{d/3}} - 1$, *then* n *is a power of* p.

Proof. If $n = p$ the lemma is true, so assume $n > p$. We let ϕ be the ring homomorphism $R \to R$ that sends each $\beta \in R$ to β^p. By 5.4(f), this map is a power of σ; in particular, it is an *automorphism* of R. The definitions of ϕ and G then imply that for all $\beta \in G$ one has $(\sigma\phi^{-1})\beta = \beta^{n/p}$.

Let L be the kernel of the group homomorphism $\mathbf{Z}^2 \to \langle\sigma\rangle$ that maps (i, j) to $(\sigma\phi^{-1})^i\phi^j$. Since the image $\langle\sigma\rangle$ of the group homomorphism has order d, the group L is a lattice of determinant d (see [13, Chapter I]). Consider the closed convex symmetric subset

$$K = \{(x, y) \in \mathbf{R}^2 : \max\{|x \log(n/p)|, |y \log p|, |x \log(n/p) + y \log p|\} \leq t\}$$

of $\mathbf{R}^2$, where $t \in \mathbf{R}_{>0}$ is chosen such that the area $3 \cdot t^2/\big(\log(n/p) \cdot \log p\big)$ of K equals $4d$. (Note that K is the hexagonal region with vertices at $\pm(t/\log(n/p), 0)$, $\pm(0, t/\log p)$, and $\pm(1/\log(n/p), -t/\log p)$.) By the inequality of the means we have

$$t = 2\sqrt{d/3} \cdot \big(\log(n/p) \cdot \log p\big)^{1/2} \leq \sqrt{d/3} \cdot \log n.$$

According to Minkowski's lattice point theorem (see [13, Chapter III, Theorem II]), the set K contains a non-zero element (i, j) of L. Multiplying (i, j) by ± 1, we can achieve that either $i > 0$ and $j \geq 0$, or $i \geq 0$ and $j < 0$. Let it first be assumed that $i > 0$ and $j \geq 0$. From $(\sigma\phi^{-1})^i\phi^j = \mathrm{id}_R$ we see that for all $\beta \in G$ one has $\beta^{(n/p)^i p^j} = \beta$. By (6.3), we can choose β to be a generator of G. Then on the one hand the order of β equals $\#G$, which exceeds $n^{\sqrt{d/3}} - 1$. On the other hand, the order of β divides $(n/p)^i p^j - 1$, where

$$\log\bigl((n/p)^i p^j\bigr) = i\log(n/p) + j\log p \leq t \leq \sqrt{d/3}\cdot\log n$$

so that $1 \leq (n/p)^i p^j \leq n^{\sqrt{d/3}}$. It follows that $(n/p)^i p^j = 1$, which by $i > 0$ implies that n is a power of p, as required. Next assume $i \geq 0$ and $j < 0$. Then from $(\sigma\phi^{-1})^i = \phi^{-j}$ one sees that for each $\beta \in G$ one has $\beta^{(n/p)^i} = \beta^{p^{-j}}$, so that the order of β divides the difference of the positive integers $(n/p)^i$ and p^{-j}; by

$$\max\{\log\bigl((n/p)^i\bigr), \log(p^{-j})\} \leq \max\{i\log(n/p), |j\log p|\} \leq t \leq \sqrt{d/3}\cdot\log n$$

we have $\max\{(n/p)^i, p^{-j}\} \leq n^{\sqrt{d/3}}$. Again taking β to be a generator of G, one deduces $(n/p)^i = p^{-j}$. Since p is prime and $j < 0$, it follows that n is a power of p. This proves 6.4.

Proof of Proposition 2.8. We let the notation and the assumptions be as in Proposition 2.8, and in addition we write $B = \lfloor (d/3)^{1/2}(\log n)/\log 2\rfloor$. Note that $d > (\log n)^2/\bigl(3\cdot(\log 2)^2\bigr)$ implies $d > B$.

We apply the theory just developed to a prime factor p of n that satisfies $p > B$. From $\sigma\alpha = \alpha^n$ we see that the element α of $R = A/pA$ belongs to the subgroup G of R^*. From $\sigma(\alpha + a) = \sigma\alpha + a = \alpha^n + a = (\alpha + a)^n$ for $a = 1, 2, \ldots, B$ and from 5.4(a), which implies each $\alpha + a \neq 0$, we see that $\alpha + 1, \alpha + 2, \ldots, \alpha + B$ also belong to G. For each proper subset S of $\{0, 1, \ldots, B\}$, the element $\prod_{a\in S}(\alpha + a)$ also belongs to G. There are $2^{B+1} - 1$ such sets S, and we claim that they give rise to $2^{B+1} - 1$ different elements of G. To see this, note that by $p > B$ the polynomials $X + a$, $a = 0, 1, \ldots, B$, are distinct in $\mathbf{F}_p[X]$, and that by unique factorization in $\mathbf{F}_p[X]$ the polynomials $\prod_{a\in S}(X + a)$, with S as above, are pairwise distinct. By $d > B$, all these polynomials have degrees smaller than d, so by 5.4(a) (applied to R) they give rise to $2^{B+1} - 1$ different elements $\prod_{a\in S}(\alpha + a)$ of G, as asserted.

It follows that we have

$$\#G \geq 2^{B+1} - 1 > 2^{(d/3)^{1/2}(\log n)/\log 2} - 1 = n^{\sqrt{d/3}} - 1.$$

Applying 6.4 we conclude that n is a power of p. This proves 2.8.

7. Tensor products

Tensor products (see [4, Chapter 2; 19, Chapter XVI]) can be used to construct "large" pseudofields out of "small" ones, in the following manner.

Proposition 7.1. *Let* (A_1, α_1) *and* (A_2, α_2) *be pseudofields with* $\operatorname{char} A_1 = \operatorname{char} A_2 = n$, *and suppose that the degrees* d_1, d_2 *of these pseudofields satisfy* $d_1 > 1$, $d_2 > 1$, *and* $\gcd(d_1, d_2) = 1$. *Then the tensor product* $(A_1 \otimes_{\mathbf{Z}/n\mathbf{Z}} A_2, \alpha_1 \otimes \alpha_2)$ *is a pseudofield of characteristic* n *and degree* $d_1 d_2$.

Proof. We check that $A = A_1 \otimes_{\mathbf{Z}/n\mathbf{Z}} A_2$, $\alpha = \alpha_1 \otimes \alpha_2$, n, $d = d_1 d_2$, and $\sigma = \sigma_1 \otimes \sigma_2$ satisfy (2.1)–(2.5). By 5.4(a), each A_i is a free $\mathbf{Z}/n\mathbf{Z}$-module with basis 1, α_i, …, $\alpha_i^{d_i - 1}$, so from [19, Chapter XVI, Corollary 2.4] one sees that A is a free $\mathbf{Z}/n\mathbf{Z}$-module with basis $(\alpha_1^i \otimes \alpha_2^j)_{0 \le i < d_1, 0 \le j < d_2}$. This implies both (2.1) and (2.2). One has $\sigma(\alpha) = \sigma_1(\alpha_1) \otimes \sigma_2(\alpha_2) = \alpha_1^n \otimes \alpha_2^n = \alpha^n$, which is (2.3). Each $\sigma_i^{d_i}$ is the identity on A_i, so σ^d is the identity on A, which implies (2.4). Finally, to prove (2.5), let l be a prime number dividing d. Then l divides exactly one of d_1 and d_2; by symmetry we may assume it divides d_2. Let k be a prime number dividing d_1. By $\sigma_1 \alpha_1 = \alpha_1^n$, the A_1-ideal $A_1 \alpha_1$ is mapped to itself by σ_1 and therefore contains $\sigma_1^{d_1/k} \alpha_1 - \alpha_1$; the latter element is a unit of A_1, so α_1 is a unit of A_1 as well. Since d/l is divisible by d_1, we have $\sigma_1^{d/l} \alpha_1 = \alpha_1 \in A_1^*$. Since d/l is not divisible by d_2, Proposition 5.1 implies $\sigma_2^{d/l} \alpha_2 - \alpha_2 \in A_2^*$. It follows that the element $\sigma^{d/l} \alpha - \alpha = (\sigma_1^{d/l} \alpha_1) \otimes (\sigma_2^{d/l} \alpha_2) - \alpha_1 \otimes \alpha_2 = \alpha_1 \otimes (\sigma_2^{d/l} \alpha_2 - \alpha_2)$ is a product of two units, and therefore belongs to A^*. This proves 7.1.

We next address the problem of designing an algorithm that, given two pseudofields (A_i, α_i) as in 7.1, computes their tensor product. Here it is assumed, as in Section 2, that a pseudofield is specified by its characteristic and its characteristic polynomial. For the general context of our algorithm one may consult [10].

Let R be a commutative ring, let $m \in \mathbf{Z}$, $m \ge 0$, and write S for the ring $R[t]/(t^{m+1})$, where t denotes a polynomial variable. The elements 1, t, …, t^m form a basis for S over R, in the sense that every element of S has a unique representation of the form $\sum_{i=0}^m a_i t^i$, with each $a_i \in R$. The elements $\sum_{i=0}^m a_i t^i$ with $a_0 = 0$ form the ideal tS of S, and the elements with $a_0 = 1$ form a subgroup of the group S^* of units of S; we write $1 + tS$ for this subgroup. We define the maps $D\colon S \to tS$ and $L\colon 1 + tS \to tS$ by

$$D\Big(\sum_{i=0}^m a_i t^i\Big) = \sum_{i=0}^m i a_i t^i \qquad (a_i \in R),$$

$$L(u) = D(u) \cdot u^{-1} \qquad (u \in 1 + tS).$$

(The notation reflects that, up to a factor t, the maps D and L are differentiation and logarithmic differentiation, respectively.) One readily verifies that for u, $v \in S$ one has $D(uv) = uD(v) + vD(u)$ and that, consequently, L is a group homomorphism from the multiplicative group $1 + tS$ to the additive group tS. For a monic polynomial $g = X^k + \sum_{i=1}^{k} b_i X^{k-i} \in R[X]$, we write $g^\flat$ for the image of $1 + \sum_{i=1}^{k} b_i t^i$ in S, which belongs to $1 + tS$. Evidently, we have $(gh)^\flat = g^\flat \cdot h^\flat$ for any two monic polynomials g, $h \in R[X]$.

The *Hadamard product* $*$ is the operation defined on S by

$$\Big(\sum_{i=0}^{m} a_i t^i\Big) * \Big(\sum_{i=0}^{m} b_i t^i\Big) = \sum_{i=0}^{m} a_i b_i t^i,$$

for a_i, $b_i \in R$. In the following result we use the definitions just given for the ring $R = \mathbf{Z}/n\mathbf{Z}$.

Proposition 7.2. *Let the hypotheses and notation be as in* 7.1. *Moreover, write* f_1, f_2, f *for the characteristic polynomials of the pseudofields* (A_1, α_1), (A_2, α_2), *and* $(A_1 \otimes_{\mathbf{Z}/n\mathbf{Z}} A_2, \alpha_1 \otimes \alpha_2)$, *respectively. Then for any non-negative integer* m *we have the identity*

$$L(f^\flat) = -L(f_1^\flat) * L(f_2^\flat)$$

in $t(\mathbf{Z}/n\mathbf{Z})[t]/(t^{m+1})$.

Proof. Let the notation A, α, d, σ_1, σ_2, σ be as in the proof of 7.1. We view A_1 and A_2 as subrings of A, identifying α_1 with $\alpha_1 \otimes 1$ and α_2 with $1 \otimes \alpha_2$, so that $\alpha = \alpha_1\alpha_2$. It suffices to prove the identity in $tA[t]/(t^{m+1})$. From $f = \prod_{i=0}^{d-1}(X - \sigma^i\alpha)$ we obtain $f^\flat = \prod_{i=0}^{d-1}\big(1 - (\sigma^i\alpha)t\big)$. From $L\big(1 - (\sigma^i\alpha)t\big) = -(\sigma^i\alpha)t/\big(1 - (\sigma^i\alpha)t\big) = -\sum_{j=1}^{m}(\sigma^i\alpha)^j t^j$ we thus obtain

$$L(f^\flat) = \sum_{i=0}^{d-1} L\big(1 - (\sigma^i\alpha)t\big) = -\sum_{j=1}^{m}\Big(\sum_{i=0}^{d-1}(\sigma^i\alpha)^j\Big)t^j.$$

Likewise, we have

$$L(f_1^\flat) = -\sum_{j=1}^{m}\Big(\sum_{i=0}^{d_1-1}(\sigma_1^i\alpha_1)^j\Big)t^j, \qquad L(f_2^\flat) = -\sum_{j=1}^{m}\Big(\sum_{i=0}^{d_2-1}(\sigma_2^i\alpha_2)^j\Big)t^j.$$

The identity to be proved now follows from

$$\sum_{i=0}^{d-1}(\sigma^i\alpha)^j = \Big(\sum_{i=0}^{d_1-1}(\sigma_1^i\alpha_1)^j\Big) \cdot \Big(\sum_{i=0}^{d_2-1}(\sigma_2^i\alpha_2)^j\Big)$$

for all $j \geq 1$, which is a consequence of $\sigma^i\alpha = (\sigma_1^i\alpha_1) \cdot (\sigma_2^i\alpha_2)$ and the fact that the orders d_1 and d_2 of σ_1 and σ_2 are coprime. This proves 7.2.

Proposition 7.3. *For positive integers n, m, let $S_{n,m}$ denote the ring $(\mathbf{Z}/n\mathbf{Z})[t]/(t^{m+1})$.*

(a) *Let n and m be positive integers such that each prime factor of n exceeds m. Then the map $L\colon 1+tS_{n,m} \to tS_{n,m}$ is a group isomorphism.*

(b) *There is an algorithm that, given positive integers n and m, and an element $u \in 1+tS_{n,m}$, computes the element $L(u)$ of $tS_{n,m}$ in time $\tilde{O}(m \log n)$.*

(c) *There is an algorithm that, given positive integers n and m, and an element $s \in tS_{n,m}$, either computes a prime factor of n that is at most m or correctly decides that no such prime factor exists, and in the latter case computes the element $L^{-1}(s)$ of $1+tS_{n,m}$, all in time $\tilde{O}(m \log n)$.*

Proof. (a) Since each prime factor of n exceeds m, we have $i \in (\mathbf{Z}/n\mathbf{Z})^*$ for $i = 1, \ldots, m$, so D restricts to a group automorphism of $tS_{n,m}$. For the same reason, there are well-defined maps $\log\colon 1+tS_{n,m} \to tS_{n,m}$ and $\exp\colon tS_{n,m} \to 1+tS_{n,m}$ with

$$\log(1-x) = -\sum_{i=1}^{m} x^i/i, \qquad \exp(x) = \sum_{i=0}^{m} x^i/i!$$

for $x \in tS_{n,m}$. It is well known that log and exp are inverse group isomorphisms. An easy computation shows $L = D \circ \log$. It follows that L is an isomorphism, with inverse $\exp \circ D^{-1}$.

(b) In [6, Section 8] one finds an algorithm that computes $L(u)$ by means of $\tilde{O}(m)$ ring operations in $\mathbf{Z}/n\mathbf{Z}$; this particular algorithm does not depend on the condition, in [6, Section 8], that the field $\mathbf{Q}$ of rational numbers be contained in the coefficient ring. By [26, Sections 8.3 and 9.1], each ring operation in $\mathbf{Z}/n\mathbf{Z}$ can be done in time $\tilde{O}(\log n)$.

(c) We describe an algorithm with the stated properties. Using the extended Euclidean algorithm, see [26, Corollary 11.10], one attempts to compute $i^{-1} \in \mathbf{Z}/n\mathbf{Z}$ for $i = 1, 2, \ldots, m$; this can only fail if among those i a prime factor of n is found, in which case the algorithm halts. Suppose it does not fail. Then one computes $D^{-1}(s)$ directly from the definition of D by means of m multiplications in $\mathbf{Z}/n\mathbf{Z}$, and next one uses the algorithm from [6, Section 9] to compute $L^{-1}(s) = \exp\big(D^{-1}(s)\big)$ using $\tilde{O}(m)$ ring operations in $\mathbf{Z}/n\mathbf{Z}$; inspection of this algorithm shows that the condition from [6, Section 9] that $\mathbf{Q}$ be contained in the coefficient ring may be replaced by the weaker condition that multiplicative inverses of each of $i = 1, 2, \ldots, m$ be available; this condition is satisfied in the present case.

This proves 7.3.

Proposition 7.4. *There is an algorithm with the following property. Given an integer n and two pseudofields of characteristic n and of coprime degrees d_1, d_2 greater than 1, it* *either finds a prime factor of n that is at most $d_1 d_2$ or it constructs the tensor product of the two given pseudofields, and it does so in time $\tilde{O}(d_1 d_2 \log n)$.*

Proof. The following algorithm has the stated properties. Let f_1, f_2 be the characteristic polynomials of the two given pseudofields. Put $m = d_1 d_2$ and $S = (\mathbf{Z}/n\mathbf{Z})[t]/(t^{m+1})$, and compute $f_1^\flat$, $f_2^\flat \in 1 + tS$ from the definition of $\flat$. Next compute $L(f_1^\flat)$ and $L(f_2^\flat)$ by means of the algorithm of 7.3(b), and compute $L(f_1^\flat) * L(f_2^\flat)$ by $d_1 d_2$ multiplications in $\mathbf{Z}/n\mathbf{Z}$. Finally, apply the algorithm of 7.3(c) to $s = -L(f_1^\flat) * L(f_2^\flat)$; this either yields a prime factor of n that is at most $m = d_1 d_2$, or it finds $L^{-1}(s) \in 1 + tS$; in the latter case, the characteristic polynomial of the tensor product is the unique monic polynomial $f \in (\mathbf{Z}/n\mathbf{Z})[X]$ of degree $d_1 d_2$ that satisfies $f^\flat = L^{-1}(s)$. This completes the description of the algorithm. It is correct by 7.2, and 7.3 readily implies that it runs in time $\tilde{O}(d_1 d_2 \log n)$. This proves 7.4.

8. Gaussian periods

In this section we let n be an integer with $n > 1$. Let r be a prime number not dividing n, and define $\Phi_r = \sum_{i=0}^{r-1} X^i \in (\mathbf{Z}/n\mathbf{Z})[X]$. The element $(X \bmod \Phi_r)$ of the ring $(\mathbf{Z}/n\mathbf{Z})[X]/(\Phi_r)$ is denoted by ζ_r, and that ring itself by $(\mathbf{Z}/n\mathbf{Z})[\zeta_r]$. We have $\zeta_r^r = 1 \neq \zeta_r$, so ζ_r is an element of $(\mathbf{Z}/n\mathbf{Z})[\zeta_r]^*$ of order r. From $\deg \Phi_r = r-1$ and $1+\zeta_r+\ldots+\zeta_r^{r-1} = 0$ one sees that the elements ζ_r^i, $1 \leq i \leq r-1$, form a basis for $(\mathbf{Z}/n\mathbf{Z})[\zeta_r]$ over $\mathbf{Z}/n\mathbf{Z}$.

For each $a \in \mathbf{Z}$, $a \notin r\mathbf{Z}$, the ring $(\mathbf{Z}/n\mathbf{Z})[\zeta_r]$ has a unique automorphism mapping ζ_r to ζ_r^a; we write σ_a for this automorphism. The set Δ of all automorphisms of the form σ_a is a group under composition, and the map $\sigma_a \mapsto (a \bmod r)$ is a group isomorphism $\Delta \cong \mathbf{F}_r^*$. One concludes that Δ is cyclic of order $r-1$, and that the elements $\tau\zeta_r$, $\tau \in \Delta$, form a basis for $(\mathbf{Z}/n\mathbf{Z})[\zeta_r]$ over $\mathbf{Z}/n\mathbf{Z}$.

Next let q be a positive integer dividing $r-1$. Then $\Delta^q = \{\tau^q : \tau \in \Delta\}$ is a subgroup of index q of Δ. The subset $(\mathbf{Z}/n\mathbf{Z})[\zeta_r]^{\Delta^q} = \{\beta \in (\mathbf{Z}/n\mathbf{Z})[\zeta_r] : \rho\beta = \beta \text{ for all } \rho \in \Delta^q\}$ is a subring of $(\mathbf{Z}/n\mathbf{Z})[\zeta_r]$. An element $\sum_{\tau\in\Delta} a_\tau \cdot \tau\zeta_r$, with each $a_\tau \in \mathbf{Z}/n\mathbf{Z}$, belongs to this subring if and only if $a_\tau = a_{\tau\rho}$ for all $\tau \in \Delta$, $\rho \in \Delta^q$. Hence, if we put $\eta_{r,q} = \sum_{\rho\in\Delta^q} \rho\zeta_r$, then the elements $\tau\eta_{r,q}$, with τ ranging over a set of coset representatives for Δ modulo Δ^q, form a basis for $(\mathbf{Z}/n\mathbf{Z})[\zeta_r]^{\Delta^q}$ over $\mathbf{Z}/n\mathbf{Z}$; in particular, one has $\#(\mathbf{Z}/n\mathbf{Z})[\zeta_r]^{\Delta^q} = n^q$. The elements $\tau\eta_{r,q}$ are called *Gaussian periods* of degree q and conductor r. For example, we have $\eta_{r,r-1} = \zeta_r$ and $\eta_{r,1} = -1$. We write

$$f_{r,q} = \prod_{\tau\Delta^q \in \Delta/\Delta^q} (Y - \tau\eta_{r,q}).$$

This is a monic polynomial in Y of degree q with $f_{r,q}(\eta_{r,q}) = 0$. Its coefficients, which belong to $(\mathbf{Z}/n\mathbf{Z})[\zeta_r]$, are readily checked to be invariant under all $\rho \in \Delta$, so they belong to $(\mathbf{Z}/n\mathbf{Z})[\zeta_r]^{\Delta^1} = (\mathbf{Z}/n\mathbf{Z}) \cdot \eta_{r,1} = \mathbf{Z}/n\mathbf{Z}$. Thus, one has $f_{r,q} \in (\mathbf{Z}/n\mathbf{Z})[Y]$.

Proposition 8.1. *Let $n \in \mathbf{Z}$, $n > 1$, let r be a prime number not dividing n, and let q be a divisor of $r-1$ with the property that the element $(n^{(r-1)/q} \bmod r)$ of $\mathbf{F}_r^*$ has order q. Let the notation ζ_r, σ_a, Δ, $\eta_{r,q}$, $f_{r,q}$ be as just defined. Then we have:*

(a) *if n is prime, then in the ring $(\mathbf{Z}/n\mathbf{Z})[\zeta_r]$ one has $\eta_{r,q}^n = \sigma_n \eta_{r,q}$;*

(b) *if in the ring $(\mathbf{Z}/n\mathbf{Z})[\zeta_r]$ one has $\eta_{r,q}^n = \sigma_n \eta_{r,q}$, then $\left((\mathbf{Z}/n\mathbf{Z})[\zeta_r]^{\Delta^q}, \eta_{r,q}\right)$ is a pseudo-field of characteristic n and degree q, with characteristic polynomial $f_{r,q}$.*

Proof. To prove (a), suppose that n is prime. Then the map from $(\mathbf{Z}/n\mathbf{Z})[\zeta_r]$ sending each β to β^n is a ring homomorphism, and since it agrees with σ_n on ζ_r it coincides with σ_n on all of $(\mathbf{Z}/n\mathbf{Z})[\zeta_r]$. This implies (a).

To prove (b), we first observe that the kernel of the group homomorphism $\mathbf{F}_r^* \to \mathbf{F}_r^*$ sending each x to $x^{(r-1)/q}$ equals the subgroup $\mathbf{F}_r^{*q}$ of index q of $\mathbf{F}_r^*$. Hence the condition that $(n^{(r-1)/q} \bmod r)$ have order q implies that the coset $(n \bmod r)\mathbf{F}_r^{*q}$ generates the group $\mathbf{F}_r^*/\mathbf{F}_r^{*q}$ and, consequently, that the coset $\sigma_n \Delta^q$ generates Δ/Δ^q.

For brevity, write $A = (\mathbf{Z}/n\mathbf{Z})[\zeta_r]^{\Delta^q}$. Define the ring homomorphism $\phi\colon (\mathbf{Z}/n\mathbf{Z})[Y] \to A$ by $\phi(g) = g(\eta_{r,q})$. Its image is the subring of A generated by $\eta_{r,q}$. From $\sigma_n \eta_{r,q} = \eta_{r,q}^n$ it follows that that subring is mapped to itself by σ_n. Since all elements of Δ^q act as the identity on A, and $\sigma_n \Delta^q$ generates Δ/Δ^q, the subring is mapped to itself by *all* $\tau \in \Delta$. Hence, in addition to $\eta_{r,q}$ it contains all $\tau\eta_{r,q}$, so that it is equal to all of A; in other words, ϕ is surjective. The kernel of ϕ contains the $(\mathbf{Z}/n\mathbf{Z})[Y]$-ideal generated by $f_{r,q}$, and since both of these ideals have index n^q in $(\mathbf{Z}/n\mathbf{Z})[Y]$, we must have equality. Thus, ϕ induces a ring isomorphism $(\mathbf{Z}/n\mathbf{Z})[Y]/(f_{r,q}) \cong A$.

We prove that A, $\alpha = \eta_{r,q}$, n, $d = q$, and σ equal to the restriction of σ_n to A, satisfy (2.1)–(2.5). Conditions (2.1)–(2.3) are clearly satisfied, and (2.4) follows from $\sigma_n^q \in \Delta^q$. We prove (2.5). Since $\sigma_n \Delta^q$ generates the group Δ/Δ^q of order q, we may rewrite the definition of $f_{r,q}$ as

$$f_{r,q} = \prod_{i=0}^{q-1} (Y - \sigma^i \eta_{r,q}).$$

It follows that the derivative $f'_{r,q} = \mathrm{d}f_{r,q}/\mathrm{d}Y$ satisfies $f'_{r,q}(\eta_{r,q}) = \prod_{i=1}^{q-1}(\eta_{r,q} - \sigma^i \eta_{r,q})$, so that to prove (2.5) it will suffice to prove $f'_{r,q}(\eta_{r,q}) \in A^*$.

Let p be a prime number dividing n. Taking the isomorphism $(\mathbf{Z}/n\mathbf{Z})[Y]/(f_{r,q}) \cong A$ modulo p, we see that the ring $\mathbf{F}_p[Y]/(f)$, where $f = (f_{r,q} \bmod p) \in \mathbf{F}_p[X]$, is isomorphic to a subring of $\mathbf{F}_p[X]/(g)$, where $g = \sum_{i=0}^{r-1} X^i$. Since g divides $X^r - 1$, where r is a prime number different from p, one has $\gcd(g, \mathrm{d}g/\mathrm{d}X) = 1$ in the ring $\mathbf{F}_p[X]$. From Lemma 8.2, stated and proved below, it follows that one has $\gcd(f, \mathrm{d}f/\mathrm{d}Y) = 1$ in the ring $\mathbf{F}_p[Y]$. Thus, there are u, $v \in \mathbf{F}_p[Y]$ with $uf + v\mathrm{d}f/\mathrm{d}Y = 1$. Lifting u, v to $(\mathbf{Z}/n\mathbf{Z})[Y]$, one obtains u_p, v_p, $w_p \in (\mathbf{Z}/n\mathbf{Z})[Y]$ such that $u_p f_{r,q} + v_p f'_{r,q} = 1 + p w_p$. Applying ϕ one gets, for each prime number p dividing n, an identity in A of the form $v_p(\eta_{r,q}) \cdot f'_{r,q}(\eta_{r,q}) - p \cdot w_p(\eta_{r,q}) = 1$. Take the product over p, repeating the pth identity just as many times as p occurs in n. On the right, we get 1. On the left, the only term that does not have a factor $f'_{r,q}(\eta_{r,q})$ is divisible by n and is therefore 0. Hence, 1 is divisible by $f'_{r,q}(\eta_{r,q})$ in A, so that the latter element is a unit, as required. The formula we gave for $f_{r,q}$ shows that it is indeed the characteristic polynomial for the pseudofield.

Lemma 8.2. *Let p be a prime number, and let f, $g \in \mathbf{F}_p[X]$ be non-zero polynomials for which the ring $\mathbf{F}_p[X]/(f)$ is isomorphic to a subring of $\mathbf{F}_p[X]/(g)$. Suppose also* $\gcd(g, \mathrm{d}g/\mathrm{d}X) = 1$. *Then one has* $\gcd(f, \mathrm{d}f/\mathrm{d}X) = 1$.

Proof. A non-zero polynomial $h \in \mathbf{F}_p[X]$ satisfies $\gcd(h, \mathrm{d}h/\mathrm{d}X) = 1$ if and only if h is squarefree in the ring $\mathbf{F}_p[X]$, and if and only if there is no non-zero nilpotent element in the ring $\mathbf{F}_p[X]/(h)$. Thus, the lemma follows from the trivial observation that if a ring has no non-zero nilpotent element, then the same is true for a subring. This proves 8.2 and completes the proof of 8.1.

We next describe an algorithm that will prove Propositions 2.13 and 2.14.

Algorithm 8.3. Given an integer $n > 1$, which may or may not be known to be prime, and a period system $\mathcal{P}$ for n satisfying $n > \prod_{(r,q)\in\mathcal{P}} q$, this algorithm attempts to construct a pseudofield of characteristic n and degree $\prod_{(r,q)\in\mathcal{P}} q$.

Step 1. For all $(r, q) \in \mathcal{P}$ in succession, do the following. Compute $\eta_{r,q} \in (\mathbf{Z}/n\mathbf{Z})[\zeta_r]$ as well as all of its conjugates $\tau\eta_{r,q}$, and form the product of the q polynomials $Y - \tau\eta_{r,q}$ in the polynomial ring $(\mathbf{Z}/n\mathbf{Z})[\zeta_r][Y]$; the result is $f_{r,q}$, which has coefficients in the subring $\mathbf{Z}/n\mathbf{Z}$ of $(\mathbf{Z}/n\mathbf{Z})[\zeta_r]$. If n is not known to be prime, compute by an nth powering in the ring $(\mathbf{Z}/n\mathbf{Z})[Y]/(f_{r,q})$ the unique polynomial $g_{r,q} \in (\mathbf{Z}/n\mathbf{Z})[Y]$ satisfying $Y^n \equiv g_{r,q} \bmod f_{r,q}$ and $\deg g_{r,q} < q$, and test whether in the ring $(\mathbf{Z}/n\mathbf{Z})[\zeta_r]$ one has $g_{r,q}(\eta_{r,q}) = \sigma_n\eta_{r,q}$; if this test fails, declare n composite and halt.

Step 2. [If the algorithm arrives at this point then, as we shall prove below, for each $(r,q)\in\mathcal{P}$ the pair $(n, f_{r,q})$ specifies a pseudofield.] Applying the algorithm of 7.4 at most $\#\mathcal{P}-1$ times, either find a prime factor of n that is at most $\prod_{(r,q)\in\mathcal{P}} q$, or construct the repeated tensor product of the $\#\mathcal{P}$ pseudofields specified by the pairs $(n, f_{r,q})$ for $(r,q)\in\mathcal{P}$. In the former case, declare n composite and halt, and in the latter case return the tensor product computed by the algorithm and halt.

This completes the description of Algorithm 8.3.

Proposition 8.4. *Algorithm* 8.3*, on input* n, $\mathcal{P}$ *satisfying* $n > \prod_{(r,q)\in\mathcal{P}} q$*, runs in time*

$$\tilde{O}\Big(\Big(\prod_{(r,q)\in\mathcal{P}} q + \sum_{(r,q)\in\mathcal{P}} qr\Big)\log n\Big) \quad \text{or} \quad \tilde{O}\Big(\Big(\prod_{(r,q)\in\mathcal{P}} q + \sum_{(r,q)\in\mathcal{P}} q(r+\log n)\Big)\log n\Big)$$

according as n *is or is not known to be prime, and either correctly declares* n *composite or constructs a pseudofield of characteristic* n *and degree* $\prod_{(r,q)\in\mathcal{P}} q$.

Proof. We first prove the correctness of the algorithm. By $f_{r,q}(\eta_{r,q}) = 0$, the congruence $Y^n \equiv g_{r,q} \bmod f_{r,q}$ in Step 1 implies $g_{r,q}(\eta_{r,q}) = \eta_{r,q}^n$. Thus, by 8.1(a), the condition $g_{r,q}(\eta_{r,q}) = \sigma_n\eta_{r,q}$ is necessary for n to be prime, and the algorithm is correct if it halts in Step 1. If it passes Step 1, then by 8.1(b) there is for each $(r,q)\in\mathcal{P}$ a pseudofield of characteristic n with characteristic polynomial $f_{r,q}$. Hence by 7.4 the algorithm either constructs the desired tensor product, or it finds a prime factor of n that is at most $\prod_{(r,q)\in\mathcal{P}} q$; in the latter case, n is composite because $n > \prod_{(r,q)\in\mathcal{P}} q$. This proves the correctness of the algorithm.

The run time of Step 1 is dominated by the computation of the polynomials $f_{r,q}$ and, if n is not known to be prime, the polynomials $g_{r,q}$ and their values at $\eta_{r,q}$. The computation of $f_{r,q}$, if done by means of Algorithm 10.3 from [26], runs in time $\tilde{O}(qr\log n)$. The computation of $g_{r,q}$ involves $O(\log n)$ multiplications in the ring $(\mathbf{Z}/n\mathbf{Z})[Y]/(f_{r,q})$ and can therefore be performed in time $\tilde{O}\big(q\cdot(\log n)^2\big)$. The computation of $g_{r,q}(\eta_{r,q})$ runs in time $\tilde{O}(qr\log n)$. By 7.4, Step 2 runs in time $\tilde{O}(\log n\cdot\prod_{(r,q)\in\mathcal{P}} q)$.

This proves 8.4.

Proposition 2.13 is an immediate corollary of 8.4. In addition, if n is prime, then it is not declared composite, so that the algorithm returns a pseudofield; whence by 2.7, its characteristic polynomial is irreducible in $\mathbf{F}_n[X]$. So Proposition 2.14 follows from 8.4 as well.

9. The continuous Frobenius problem

The famous Frobenius postage problem asks for the largest number which is not in the additive semigroup generated by a set of coprime positive integers. In this section we present a new result of Bleichenbacher [9] that might be considered a continuous version of this problem. A similar result was also recently obtained by Lev [20].

Proposition 9.1. (Daniel Bleichenbacher) *Suppose S is an open subset of the positive reals that is closed under addition, and such that $1 \notin S$. Then for any number $t \in (0,1]$, the dx/x measure of $S \cap (0,t)$ is at most t.*

Proof. Let M be a positive differentiable measure on the positive reals, with continuous derivative m. Thus, if $\mathcal{S}$ is any measurable subset of the positive reals with characteristic function $\chi_{\mathcal{S}}$, we have

$$M(\mathcal{S}) \;=\; \int_0^\infty \chi_{\mathcal{S}}(x) m(x)\,\mathrm{d}x.$$

Let S be as in the hypothesis of the theorem, and first suppose that $S_t := S \cap (0,t)$ is a finite union of open intervals; that is, for some positive integer n,

$$S_t \;=\; \bigcup_{i=1}^{n} (a_i, b_i),$$

where

$$t \ge b_1 \ge a_1 \ge \cdots \ge b_n \ge a_n \ge 0. \tag{9.1}$$

Let $\mathbf{a} = (a_1, \ldots, a_n), \mathbf{b} = (b_1, \ldots, b_n)$. The condition that 1 is not in the additive semigroup generated by S_t is equivalent to the assertion that for each vector $\mathbf{h} \in (\mathbf{N}_{\ge 0})^n$,

$$\text{either} \;\; \mathbf{h}\cdot\mathbf{a} \ge 1 \;\; \text{or} \;\; \mathbf{h}\cdot\mathbf{b} \le 1. \tag{9.2}$$

That is, it is not the case that $\mathbf{h}\cdot\mathbf{a} < 1 < \mathbf{h}\cdot\mathbf{b}$.

Suppose now that we fix the vector $\mathbf{b}$ and assume that

$$t \ge b_1 > b_2 > \cdots > b_n > 0. \tag{9.3}$$

If j is an integer with $j > 1/b_n$, then (9.2) implies that $a_n \ge 1/j$. In particular, $a_n \ge b_n/2$. Hence, the set of vectors $\mathbf{a}$ which, with the fixed vector $\mathbf{b}$, satisfy (9.1) and (9.2) forms a compact subset of $(\mathbf{R}_{>0})^n$. Thus there is a choice of the vector $\mathbf{a}$ which maximizes $M(S_t)$

for the given vector $\mathbf{b}$. Call this maximum value $M_{\mathbf{b}}$ and assume that $\mathbf{a}$ is fixed at a choice which produces this maximum.

Since we allow empty intervals, that is, we allow $a_i = b_i$, it is clear that if some coordinates of $\mathbf{b}$ are deleted to form a shorter vector $\mathbf{b}'$ then $M_{\mathbf{b}'} \leq M_{\mathbf{b}}$. Thus, by possibly replacing $\mathbf{b}$ with a shorter vector, we may assume that each $a_i < b_i$. We now show that we may assume that each $a_{i-1} > b_i$ for $2 \leq i \leq n$. For suppose some $a_{i-1} = b_i$. We may then consolidate the two intervals $(a_i, b_i), (a_{i-1}, b_{i-1})$ into one interval (a_i, b_{i-1}). Indeed, if not, then now 1 is representable by a sum of members of $S_t \cup b_i$, so that b_i must be involved in the sum, say with positive integral coefficient c. If $c = 1$, then replace b_i in the sum with $b_i + \epsilon$, for a suitably small $\epsilon > 0$, and then replace another member $x \in S_t$ of the sum with $x - \epsilon$. (There must be another number in the sum since $b_i < 1$.) If ϵ is small enough, both $b_i + \epsilon$ and $x - \epsilon$ are in S_t, and we have represented 1 as a sum of members of S_t. And if $c \geq 2$, then since $b_i + \epsilon/(c-1)$ and $b_i - \epsilon$ are both in S_t for ϵ small enough, we can replace the c copies of b_i in the sum with $c - 1$ copies of $b_i + \epsilon/(c-1)$ and one copy of $b_i - \epsilon$, and so represent 1 as a sum of members of S_t. Either way, we reach a contradiction, and so the consolidation of the two abutting intervals continues to enjoy the property that 1 is not in the additive semigroup generated by the intervals. Hence, we may assume that $a_{i-1} > b_i$ for $2 \leq i \leq n$. Thus, we may assume that the vector $\mathbf{a}$ satisfies

$$t \geq b_1 > a_1 > \cdots > b_n > a_n > 0. \tag{9.4}$$

Now let

$$\begin{aligned} H_0 &= \{\mathbf{h} \in (\mathbf{N}_{\geq 0})^n \ : \ \mathbf{h} \cdot \mathbf{a} < 1\}, \\ H_1 &= \{\mathbf{h} \in (\mathbf{N}_{\geq 0})^n \ : \ \mathbf{h} \cdot \mathbf{a} = 1\}, \\ H_2 &= \{\mathbf{h} \in (\mathbf{N}_{\geq 0})^n \ : \ \mathbf{h} \cdot \mathbf{a} > 1\}. \end{aligned}$$

Since each $a_i > 0$, it follows that H_0, H_1 are finite sets. We now show that H_1 is nonempty. Suppose not. Let $\mathbf{u} = (1, 1, \ldots, 1) \in (\mathbf{N}_{\geq 0})^n$. We claim that if $\epsilon > 0$ is small enough, then the pair $\mathbf{a} - \epsilon\mathbf{u}, \mathbf{b}$ still satisfies (9.2) and (9.4). This would create a choice for S_t with strictly larger $M(S_t)$, a contradiction, thus showing that H_1 is nonempty. It is clear that we may choose $\epsilon > 0$ small enough so as to preserve the condition (9.4). For $\mathbf{h} \in H_0$ we have $\mathbf{h} \cdot \mathbf{b} \leq 1$, so that the vectors in H_0 do not pose a problem for condition (9.2), and since H_1 is assumed empty, H_1 also does not pose a problem. There are only finitely many $\mathbf{h} \in H_2$ with $\mathbf{h} \cdot \mathbf{a} \leq 2$. We may choose $\epsilon > 0$ small enough so that $\mathbf{h} \cdot (\mathbf{a} - \epsilon\mathbf{u}) \geq 1$ for all such $\mathbf{h}$. But if we choose $\epsilon < a_n/2$, then if $\mathbf{h} \cdot \mathbf{a} > 2$, then $\mathbf{h} \cdot (\mathbf{a} - \epsilon\mathbf{u}) > \frac{1}{2}\mathbf{h} \cdot \mathbf{a} > 1$.

Hence, as claimed, if $\epsilon > 0$ is small enough, $\mathbf{a} - \epsilon\mathbf{u}, \mathbf{b}$ still satisfy (9.2) and (9.4), providing a contradiction which shows that H_1 is nonempty.

Let $\mathbf{h} \in H_1$. For notational convenience, let $a_{n+1} = b_{n+1} = 0$. And let $\mathbf{e}_k$ be the k-th standard basis vector in $\mathbf{R}^n$. For $k = 1, \ldots, n$, since $\mathbf{h} \cdot \mathbf{a} = 1$ and $a_k > a_{k+1}$, we have

$$\mathbf{h} \cdot \mathbf{a} - a_k + a_{k+1} < 1.$$

Suppose that $h_k > 0$. Let $\mathbf{h}' = \mathbf{h} - \mathbf{e}_k + \mathbf{e}_{k+1}$ in the case that $k < n$, and let $\mathbf{h}' = \mathbf{h} - \mathbf{e}_k$ in the case that $k = n$. Then $\mathbf{h}' \in H_0$. Hence, from (9.2), we have that $\mathbf{h}' \cdot \mathbf{b} \leq 1$. That is,

$$\mathbf{h} \cdot \mathbf{b} - b_k + b_{k+1} \leq 1.$$

Using that $\mathbf{h} \in H_1$ we get that

$$\mathbf{h} \cdot (\mathbf{b} - \mathbf{a}) \;=\; \mathbf{h} \cdot \mathbf{b} - 1 \;\leq\; b_k - b_{k+1}.$$

Thus, we have

$$h_k \mathbf{h} \cdot (\mathbf{b} - \mathbf{a}) \;\leq\; h_k(b_k - b_{k+1}), \tag{9.5}$$

an inequality that clearly continues to hold even if $h_k = 0$.

Let $\mathbf{v} \in \mathbf{R}^n$ and let

$$f_{\mathbf{v}}(x) \;=\; M\left(\bigcup_{i=1}^{n}(a_i + xv_i, b_i)\right).$$

Note that

$$f'_{\mathbf{v}}(0) \;=\; -\mathbf{v} \cdot m(\mathbf{a}),$$

where $m(\mathbf{a}) = (m(a_1), \ldots, m(a_n))$. Note too that by the maximality of $\mathbf{a}$, if the vector $\mathbf{a} + x\mathbf{v}$ satisfies (9.2) and (9.4) for all x in some interval $[0, \epsilon)$ with $\epsilon > 0$, then $f'_{\mathbf{v}}(0) \leq 0$, that is, $\mathbf{v} \cdot m(\mathbf{a}) \geq 0$. In fact, this event occurs whenever $\mathbf{h} \cdot \mathbf{v} \geq 0$ for all $\mathbf{h} \in H_1$. Indeed, suppose so, and suppose that $\mathbf{h}' \cdot (\mathbf{a} + x\mathbf{v}) < 1 < \mathbf{h}' \cdot \mathbf{b}$ for some $\mathbf{h}' \in (\mathbf{N}_{\geq 0})^n$. Since $\mathbf{h} \cdot \mathbf{b} \leq 1$ for all $\mathbf{h} \in H_0$, we have $\mathbf{h}' \notin H_0$. If $\mathbf{h} \in H_1$, then $\mathbf{h} \cdot (\mathbf{a} + x\mathbf{v}) = 1 + x\mathbf{h} \cdot \mathbf{v} \geq 1$ for all $x \geq 0$, so that $\mathbf{h}' \notin H_1$. For any given $\epsilon > 0$, there are only finitely many $\mathbf{h} \in H_2$ with $\mathbf{h} \cdot (\mathbf{a} + \epsilon\mathbf{v}) < 1 < \mathbf{h} \cdot \mathbf{a}$. Reducing the size of ϵ to a small enough positive quantity makes this set of $\mathbf{h}$ empty, and so $\mathbf{h}' \notin H_2$. It follows that for $\epsilon > 0$ small enough, if $\mathbf{h} \cdot \mathbf{v} \geq 0$ for all $\mathbf{h} \in H_1$, then $\mathbf{a} + x\mathbf{v}$ satisfies (9.2) and (9.4) for $0 \leq x < \epsilon$, and so $\mathbf{v} \cdot m(\mathbf{a}) \geq 0$.

We now apply a theorem of Farkas [18].

Lemma. (J. Farkas) *Suppose A is an $n \times k$ real matrix and $\mathbf{m} \in \mathbf{R}^n$. Then the inequalities $A\mathbf{v} \geq \mathbf{0}$, $\mathbf{m} \cdot \mathbf{v} < 0$ are unsolvable for a vector $\mathbf{v} \in \mathbf{R}^n$ if and only if there is a vector $\mathbf{p} \in \mathbf{R}^k$ with $\mathbf{p} \geq \mathbf{0}$ and $\mathbf{p}^T A = \mathbf{m}$.*

(Note that we say a vector is $\geq \mathbf{0}$ when each entry of the vector is ≥ 0.) We apply this lemma to the matrix A whose rows are the u vectors in H_1 and to the vector $\mathbf{m} = m(\mathbf{a})$. We have shown that $A\mathbf{v} \geq \mathbf{0}$ implies that $\mathbf{m} \cdot \mathbf{v} \geq 0$. Thus the lemma implies there is a vector $\mathbf{p} \in \mathbf{R}^u$ with $\mathbf{p} \geq \mathbf{0}$ and $\mathbf{p}^T A = \mathbf{m}$. Say $\mathbf{p} = (p_1, \ldots, p_u)$, $H_1 = \{\mathbf{h}_1, \ldots, \mathbf{h}_u\}$, and let each $\mathbf{h}_j = (h_{j1}, \ldots, h_{jn})$. We have

$$\sum_{j=1}^{u} p_j h_{ji} \;=\; m(a_i) \text{ for } 1 \leq i \leq n.$$

Take (9.5) applied to $\mathbf{h}_j$, multiply it by p_j, and sum over j. For $k = 1, \ldots, n$, we have,

$$\sum_{j=1}^{u} p_j h_{jk} \sum_{i=1}^{n} h_{ji}(b_i - a_i) \;\leq\; \sum_{j=1}^{u} p_j h_{jk}(b_k - b_{k+1}) \;=\; m(a_k)(b_k - b_{k+1}).$$

Multiplying corresponding inequalities by a_k and summing over k, we get

$$\sum_{k=1}^{n} a_k \sum_{j=1}^{u} p_j h_{jk} \sum_{i=1}^{n} h_{ji}(b_i - a_i) \;\leq\; \sum_{k=1}^{n} a_k m(a_k)(b_k - b_{k+1}). \tag{9.6}$$

The left side of (9.6) is

$$\begin{aligned} \sum_{j=1}^{u} p_j \sum_{k=1}^{n} a_k h_{jk} \sum_{i=1}^{n} h_{ji}(b_i - a_i) \;&=\; \sum_{j=1}^{u} p_j \sum_{i=1}^{n} h_{ji}(b_i - a_i) \\ &=\; \sum_{i=1}^{n} (b_i - a_i) \sum_{j=1}^{u} p_j h_{ji} \\ &=\; \sum_{i=1}^{n} (b_i - a_i) m(a_i). \end{aligned}$$

Thus,

$$\sum_{i=1}^{n} m(a_i)(b_i - a_i) \;\leq\; \sum_{k=1}^{n} a_k m(a_k)(b_k - b_{k+1}). \tag{9.7}$$

We now apply (9.7) with the measure M being dx/x. Then each $m(a_i) = 1/a_i$, so that

$$\sum_{i=1}^{n}(b_i/a_i - 1) \;\le\; \sum_{k=1}^{n}(b_k - b_{k+1}) \;=\; b_1 \;\le\; t. \tag{9.8}$$

However, $M((a_i, b_i)) = \log(b_i/a_i) < b_i/a_i - 1$. Hence, by (9.8),

$$M_{\mathbf{b}} \;=\; \sum_{i=1}^{n} \log(b_i/a_i) \;<\; t.$$

Since $M_{\mathbf{b}} < t$ for each choice of $\mathbf{b}$ satisfying (9.3), it remains to handle the case of S_t being the union of infinitely many disjoint open intervals. If $S_t(n)$ is the union of n of these disjoint open intervals with $S_t(n) \subset S_t(n+1)$ and $\bigcup S_t(n) = S_t$, we have $M(S_t(n)) < t$ for each n, and $M(S_t) = \lim_{n\to\infty} M(S_t(n)) \le t$. This concludes the proof of the theorem.

Remarks. We have seen that if S_t is a finite union of intervals, then $M(S_t) < t$. That is, we have a strict inequality. Moreover, this inequality for a finite union of intervals is best possible. Indeed, suppose S^n is the additive semigroup generated by $(1/(n+1), 1/n)$, where n is a positive integer. Then 1 is not in S^n. Further, we have

$$M(S_t^n) \ge \sum_{j=1}^{\lfloor tn \rfloor} \log(1 + 1/n) \;=\; \lfloor tn \rfloor (1/n + O(1/n^2)) \;\sim\; t \;\text{ as }\; n \to \infty.$$

We finally remark that it is not difficult to obtain inequalities for $M_\alpha(S_t)$, where M_α is the dx/x^α measure and $0 \le \alpha < 1$. This may be done as a corollary of the result for the dx/x measure, or as a consequence of (9.7).

10. The distribution of primes in residue classes

For a natural number q, an integer a coprime to q, and a real number x, let $\pi(x, q, a)$ denote the number of primes $p \le x$ with $p \equiv a \bmod q$. Also, let

$$\psi(x, q, a) \;=\; \sum_{\substack{n \le x \\ n \equiv a \bmod q}} \Lambda(n), \qquad \theta(x, q, a) \;=\; \sum_{\substack{p \le x,\ p \text{ prime} \\ p \equiv a \bmod q}} \log p,$$

where Λ is von Mangoldt's function.

Dirichlet proved in 1837 that if q is a positive integer coprime to the integer a, then $\pi(x, q, a)$ is unbounded, in fact, he showed that the sum of the reciprocals of the primes

$p \equiv a \bmod q$ diverges. In 1896, de la Vallée Poussin proved the prime number theorem for arithmetic progressions. This result asserts that for q, a as in Dirichlet's theorem, we have $\pi(x, q, a) \sim \pi(x)/\varphi(q)$ as $x \to \infty$ (where as usual, $\pi(x)$ is the number of primes in $[1, x]$). In the last 100+ years people have been trying to improve on this result, by allowing $q \to \infty$ as well. Clearly q cannot be as large as x, since then the assertion loses meaning. We know that if the ERH is assumed then we can take q up to nearly $x^{1/2}$. But rigorously, we only have asymptotics for each individual $\pi(x, q, a)$, with effective error estimates, for $q < (\log x)^{2-\epsilon}$, see [14], page 123. Allowing the ineffective theorem of Siegel allows us to extend this range to $q < (\log x)^A$ for any fixed A, giving us the Page-Siegel-Walfisz theorem. However, since we wish to use only effective tools, we will bypass this result.

Other ways that the prime number theorem for arithmetic progressions has been extended is to allow for a few exceptional moduli, and then to prove results about the remaining unexceptional moduli. One such theorem is found in [3]. Another type of theorem is to show that the exceptional moduli *in toto* do not contribute too much to the error when averaging over many moduli. An example of such a result is the Bombieri–Vinogradov theorem, which we discuss below. As it stands, this result uses Siegel's theorem to show that the contribution from exceptional moduli is small. We give a result that instead just ignores the exceptional moduli, if there are any.

Finally, barring asymptotics, or asymptotics on average, we have inequalities. In particular, the Brun-Titchmarsh inequality gives useful upper bounds for $\pi(x, q, a)$. However, this inequality degrades as q grows larger, so people have tried to get results that do not degrade so rapidly or are at least better on average. A culmination of these efforts is found in the series of papers of Bombieri–Friedlander–Iwaniec. However, these papers and many others, use Siegel's theorem. Further, unlike with the Bombieri–Vinogradov theorem, it does not seem so simple to disentangle Siegel's theorem from the result. As it turns out, we do not need a great improvement on the Brun-Titchmarsh inequality, just a small improvement. And a result of Deshouillers–Iwaniec from 1981 fills the bill: it is effective, and strong enough for our needs.

In this section we collect the main results we shall use on $\pi(x, q, a)$, including a proof-sketch of a version of the Bombieri–Vinogradov theorem that is effective.

Lemma 10.1. [Brun-Titchmarsh inequality] *If $x > q$ we have*

$$\pi(x, q, a) \le \frac{2x}{\varphi(q) \log(x/q)}.$$

The lemma in this form is due to Montgomery and Vaughan [22]. Note that the inequality gives an upper bound for $\pi(x, q, a)$ that is of the expected order of magnitude, namely $x/(\varphi(q)\log x)$, if $q < x^{1-\epsilon}$. When q is of order of magnitude x^α, the upper bound provided by the lemma is presumably too large by a factor $2/(1-\alpha)$.

A result similar to the following lemma can be found in Timofeev [25], Theorem 2.

Lemma 10.2. [effective Bombieri–Vinogradov inequality] *There are absolute, effectively computable positive numbers* c_6, c_7 *such that for all numbers* $x \geq 3$, *there is an integer set* $\mathcal{S}(x) \subset [(\log x)^{1/2}, \exp\left((\log x)^{1/2}\right)]$ *of cardinality* 0 *or* 1, *such that for each number* $Q \in [x^{1/3}\log x,\, x^{1/2}]$,

$$\sum_{q\leq Q}{}' \max_{2\leq y\leq x} \max_{\gcd(a,q)=1} \left|\psi(y,q,a) - \frac{y}{\varphi(q)}\right| \;\leq\; c_6 x^{1/2} Q(\log x)^5 + c_6 x \exp\left(-c_7(\log x)^{1/2}\right),$$

where the dash indicates that if $\mathcal{S}(x) = \{s_1\}$, *then no* q *in the sum is divisible by* s_1.

Proof. We follow Vaughan's proof of Bombieri's theorem, see Davenport [14, Chapter 28]. There is an effectively computable positive number c_8 such that for any number $X > 2$, there is at most one natural number $s_1 \leq X$ for which there is a primitive (real) character χ_1 with modulus s_1, and for which the L-function $L(s,\chi_1)$ has a real zero $\beta_1 > 1 - c_8/\log X$. Further, if s_1 exists, it exceeds $\log X$. Let $\mathcal{S}(x)$ be the set of such integers s_1 for $X = \exp\left((\log x)^{1/2}\right)$. Thus $\mathcal{S}(x)$ is either $\{s_1\}$ or the empty set.

For a Dirichlet character χ to the modulus q, let

$$\psi(y,\chi) \;=\; \sum_{n\leq y} \Lambda(n)\chi(n).$$

Also, let $\delta(\chi) = 1$ if χ is the principal character, and otherwise let $\delta(\chi) = 0$. We consider $|\psi(y,\chi) - \delta(\chi)y|$ for $q \leq \exp\left((\log x)^{1/2}\right)$, q not divisible by s_1 if s_1 exists, and $2 \leq y \leq x$. Any real zero of the L-function $L(z,\chi)$ must be at most $1 - c_8/(\log x)^{1/2}$. It then follows from the argument on pages 121 and 122 of [14], especially (6), that

$$|\psi(y,\chi) - \delta(\chi)y| \;=\; O\left(y^{1-c_8/(\log x)^{1/2}} + y^{1-c_9/(\log y)^{1/2}}\right),$$

where c_9 is positive and effectively computable. Thus, uniformly for $q \leq \exp\left((\log x)^{1/2}\right)$ with q not divisible by any member of $S(x)$, if χ has modulus q, then

$$\max_{2\leq y\leq x} |\psi(y,\chi) - \delta(\chi)y| \;=\; O\left(x\exp\left(-c_{10}(\log x)^{1/2}\right)\right), \tag{10.1}$$

where $c_{10} = \min\{c_8, c_9\}$.

Let

$$E(x,q) \;=\; \max_{2\le y\le x}\ \max_{\gcd(a,q)=1} \left|\psi(y;q,a) - \frac{y}{\varphi(q)}\right|.$$

We have from the argument on page 163 of [14] that

$$(10.2)\quad \sum_{q\le Q}{}' E(x,q) \;=\; O\left(Q(\log x)^2 + (\log x)\sum_{q\le Q}{}'\frac{1}{\varphi(q)}\sum_{\chi \bmod q}{}^{*}\max_{2\le y\le x}|\psi(y,\chi)-\delta(\chi)y|\right),$$

where $\sum^*$ indicates the summation is over primitive characters. Let $c_{11} = \min\{1, c_{10}/2\}$ and let $Q' = \exp\left(c_{11}(\log x)^{1/2}\right)$. Then by (10.1),

(10.3)

$$\sum_{q\le Q'}{}'\frac{1}{\varphi(q)}\sum_{\chi \bmod q}{}^{*}\max_{2\le y\le x}|\psi(y,\chi)-\delta(\chi)y| \;=\; O\left(xQ'\exp\left(-c_{10}(\log x)^{1/2}\right)\right) \;=\; O(x/Q').$$

From (2) on page 162 of [14] (Vaughan's inequality), we have for any number U with $1 \le U < x$,

$$\sum_{U<q\le 2U}\frac{1}{\varphi(q)}\sum_{\chi \bmod q}{}^{*}\max_{2\le y\le x}|\psi(y,\chi)| \;=\; O\left(\left(x/U + x^{5/6} + x^{1/2}Q\right)(\log x)^4\right).$$

Thus, as on page 164 of [14], we have

$$\sum_{Q'<q\le Q}\frac{1}{\varphi(q)}\sum_{\chi \bmod q}{}^{*}\max_{2\le y\le x}|\psi(y,\chi)| \;=\; O\left(\left(\frac{x}{Q'} + x^{5/6}\log x + x^{1/2}Q\right)(\log x)^4\right),$$

where there is no restriction on the divisibility of q by a member of $\mathcal{S}(x)$. Note that since $q > 1$ in the sum, any primitive $\chi \bmod q$ is nonprincipal, so that $\delta(\chi) = 0$. Putting this estimate together with (10.2) and (10.3), we have

$$\sum_{q\le Q}{}' E(x,q) \;=\; O\left(x^{1/2}Q(\log x)^5 + x\exp\left(-c_7(\log x)^{1/2}\right)\right)$$

for any choice of c_7 with $c_7 < c_{11}$. This completes the proof of 10.2.

Lemma 10.3. *With the same notation and hypotheses as in 10.2, we have*

$$\sum_{q\le Q}{}'\ \max_{\gcd(a,q)=1}\left|\pi(x,q,a) - \frac{\mathrm{li}(x)}{\varphi(q)}\right| \;\le\; c_{12}x^{1/2}Q(\log x)^5 + c_{12}x\exp\left(-c_7(\log x)^{1/2}\right),$$

where c_7 is as in 10.2, and c_{12} is an absolute, effectively computable number.

Proof. First note that one may replace the expressions $\psi(y,q,a)$ in 10.2 with $\theta(y,q,a)$, since

$$|\psi(y,q,a)-\theta(y,q,a)| \;\le\; \sum_{\substack{n\le y\\ n \text{ is a power}}} \log y \;=\; O\left(y^{1/2}\log y\right).$$

Thus, the result follows directly from 10.2 and the identity

$$\pi(x,q,a) \;=\; \frac{\theta(x,q,a)}{\log x}+\int_2^x \frac{\theta(y,q,a)}{y(\log y)^2}\,\mathrm{d}y.$$

In fact, one can save a factor of $\log x$ using this identity, but this is unimportant.

Lemma 10.4. [Deshouillers–Iwaniec] *For each integer m with $m\ge 3$ there is an effectively computable integer x_m and absolute and effectively computable positive numbers c_{13}, c_{14} with the following property. For arbitrary numbers x, Q with $x\ge x_m$, and $x^{1/2}\le Q\le x^{1-1/m}$, and for an arbitrary integer a with $0<|a|<x^{1/m}$, we have*

$$\pi(x,q,a) \;\le\; \frac{(4/3+c_{13}/m)x}{\varphi(q)\log(x/q)}$$

for almost all integers $q\in[Q,2Q]$ with $\gcd(q,a)=1$, the number of exceptions being less than $Qx^{-c_{14}/m}$.

This result was announced in [15], and a sketch of the proof was presented in [16]. No claim of effectivity for c_{13}, c_{14}, x_m was made by the authors, but their methods are effective.

11. Sieved primes

The goal of this section is to get a lower bound for the distribution of primes r with $r-1$ free of prime factors in some given set. Our proof closely follows an argument of Balog [5]. Before stating this result we first present an elementary lemma.

Lemma 11.1. *We have for any real number $t>1$ that*

$$\sum_{d<t}\frac{1}{\varphi(d)} \;=\; \frac{\zeta(2)\zeta(3)}{\zeta(6)}\log t+\nu+O\left(\frac{\log(2t)}{t}\right),$$

where ζ is the Riemann zeta-function and where ν is a constant identified below.

Proof. By writing

$$\frac{1}{\varphi(d)} \;=\; \frac{1}{d}\sum_{u|d}\frac{\mu^2(u)}{\varphi(u)},$$

with μ the Möbius function, we have (with γ the Euler-Mascheroni constant)

$$\begin{aligned}\sum_{d<t}\frac{1}{\varphi(d)} &= \sum_{u<t}\frac{\mu^2(u)}{\varphi(u)}\sum_{d<t,\ u\mid d}\frac{1}{d} = \sum_{u<t}\frac{\mu^2(u)}{\varphi(u)}\frac{1}{u}\left(\log\left(\frac{t}{u}\right)+\gamma+O\left(\frac{u}{t}\right)\right)\\ &= (\log t)\sum_{u<t}\frac{\mu^2(u)}{u\varphi(u)}+\sum_{u<t}\frac{\mu^2(u)(\gamma-\log u)}{u\varphi(u)}+O\left(\frac{1}{t}\sum_{u<t}\frac{\mu^2(u)}{\varphi(u)}\right)\\ &= (\log t)\prod_{p\text{ prime}}\left(1+\frac{1}{p(p-1)}\right)+\sum_{u=1}^{\infty}\frac{\mu^2(u)(\gamma-\log u)}{u\varphi(u)}+O\left(\frac{\log(2t)}{t}\right)\\ &= \frac{\zeta(2)\zeta(3)}{\zeta(6)}\log t+\nu+O\left(\frac{\log(2t)}{t}\right),\end{aligned}$$

where $\nu=\sum_u \mu^2(u)(\gamma-\log u)/(u\varphi(u))$.

Proposition 11.2. *For each integer $m \ge 4$, there are effectively computable positive numbers X_m, δ_m, with X_m an integer, satisfying the following property. If $x \ge X_m$ and $\mathcal{Q}$ is a set of primes in the interval $(1, x^{1/2}]$ with*

$$\sum_{q\in\mathcal{Q}}\frac{1}{q-1} \le \frac{3}{11}-\frac{1}{m}, \tag{11.3}$$

then there are at least $\delta_m x/(\log x)^2$ primes $r \le x$ such that every prime factor q of $r-1$ satisfies $q \le x^{1/2}$ and $q \notin \mathcal{Q}$.

Proof. Let m be an integer with $m \ge 4$, let x be a positive real number, and suppose we have a set of primes $\mathcal{Q} \subset (1, x^{1/2}]$ satisfying (11.3). Let m' be an integer with $m' \ge 2m$ to be determined later and let $\beta = 1/m'$. Let

$$\mathcal{L}=(x^{1/2-2\beta},x^{1/2-\beta})\cap\mathbf{Z},\quad \mathcal{H}=(x^{1/2+\beta},x^{1/2+2\beta})\cap\mathbf{Z}.$$

For a prime $r \le x$, let $g(r)$ denote the number of factorizations of $r-1$ as lh, where

$$\begin{aligned}&l\in\mathcal{L},\quad h\in\mathcal{H},\\ &lh\ \text{ is not divisible by any member of }\ \mathcal{Q},\\ &l\ \text{ is not divisible by any member of }\ \mathcal{S}(x),\\ &h\ \text{ is not divisible by any prime larger than }\ x^{1/2},\end{aligned}$$

where $\mathcal{S}(x)$ is defined in 10.2. It is possible that $g(r)=0$; let N denote the number of primes $r \le x$ with $g(r) > 0$. Our goal is to get a good lower bound for N.

From Cauchy's inequality, we obtain

$$N \;\geq\; \left(\sum_{r\leq x} g(r)\right)^2 \left(\sum_{r\leq x} g(r)^2\right)^{-1}.$$

Our first task is to get an upper bound for $\sum_{r\leq x} g(r)^2$, and to do this we shall ignore the non-divisibility requirements in the definition of $g(r)$ and use only the relatively simple 10.1. We have, with $[a,b]$ denoting the least common multiple of a, b,

$$\sum_{\text{prime } r\leq x} g(r)^2 \;\leq\; \sum_{\text{prime } r\leq x} \sum_{\substack{l_1,l_2|r-1\\ l_1,l_2\in\mathcal{L}}} 1 \;=\; \sum_{l_1,l_2\in\mathcal{L}} \pi(x,[l_1,l_2],1).$$

By 10.1, we thus have

$$\begin{aligned}\sum_{\text{prime } r\leq x} g(r)^2 \;&\leq\; 2x \sum_{l_1,l_2\in\mathcal{L}} \frac{1}{\varphi([l_1,l_2])\log(x/[l_1,l_2])}\\ &<\; \frac{x}{\beta\log x}\sum_{l_1,l_2\in\mathcal{L}}\frac{1}{\varphi([l_1,l_2])}.\end{aligned}$$

We have

$$\begin{aligned}\sum_{l_1,l_2\in\mathcal{L}}\frac{1}{\varphi([l_1,l_2])} \;&=\; \sum_{d<x^{1/2-\beta}} \sum_{\substack{\gcd(l_1,l_2)=d\\ l_1,l_2\in\mathcal{L}}} \frac{1}{\varphi(l_1l_2/d)} \;\leq\; \sum_{d<x^{1/2-\beta}} \sum_{a,b<x^{1/2-\beta}/d} \frac{1}{\varphi(abd)}\\ &\leq\; \left(\sum_{d<x^{1/2}}\frac{1}{\varphi(d)}\right)^3 \leq (\log x)^3,\end{aligned}$$

the last inequality following from 11.1 for all x beyond an absolute bound. We conclude that

$$\sum_{\text{prime } r\leq x} g(r)^2 \;\leq\; \beta^{-1}x(\log x)^2.$$

We now turn our attention to the heart of the proof, which is to obtain a reasonable lower bound for $\sum_{r\leq x} g(r)$, and for this we shall use 10.3 and 10.4. Let $\mathcal{L}_1$ denote the set of integers $l\in\mathcal{L}$ with l not divisible by any member of $\mathcal{S}(x)$. To begin, we have

$$\begin{aligned}\sum_{\text{prime } r\leq x} g(r) \;\geq\; &\sum_{l\in\mathcal{L}_1}\pi(x,l,1) - \sum_{l\in\mathcal{L}_1}\pi(x^{1/2+\beta}l,l,1) - \sum_{\substack{l\in\mathcal{L}_1\\ q|l \text{ for some } q\in\mathcal{Q}}}\pi(x,l,1)\\ &- \sum_{\substack{h\in\mathcal{H}\\ q|h \text{ for some } q\in\mathcal{Q}}}\pi(x,h,1) - \sum_{\substack{h\in\mathcal{H}\\ q|h \text{ for some prime } q>x^{1/2}}}\pi(x,h,1)\\ =\;& S_1 - S_2 - S_3 - S_4 - S_5, \quad\text{say.}\end{aligned}$$

Indeed, S_1 counts the number of pairs l, h where $lh+1$ is a prime $r \le x$ and $l \in \mathcal{L}_1$, while S_2 removes from this count those pairs where $h \notin \mathcal{H}$, S_3 removes those pairs where l is divisible by some prime in $\mathcal{Q}$, etc.

For S_1 we use 10.3, getting

$$S_1 \;=\; \mathrm{li}(x) \sum_{l \in \mathcal{L}_1} \frac{1}{\varphi(l)} + O\left(\frac{x}{(\log x)^2}\right).$$

By 11.1 and using that $\mathcal{S}(x)$ is either empty or has a single member greater than $(\log x)^{1/2}$, it follows that with $\xi = \zeta(2)\zeta(3)/\zeta(6)$,

$$S_1 \;=\; \xi\beta x + O(x/(\log x)^{1/4}).$$

By 10.1, we have

$$S_2 \;=\; O\left(\frac{x^{1/2+\beta}}{\log x} \sum_{l \in \mathcal{L}_1} \frac{l}{\phi(l)}\right).$$

Using 11.1 and partial summation, $\sum_{l \in \mathcal{L}_1} l/\phi(l) \;=\; O(x^{1/2-\beta})$, so $S_2 = O(x/\log x)$.

For S_3 we use 10.3, getting

$$\begin{aligned} S_3 \;&\le\; \mathrm{li}(x) \sum_{q \in \mathcal{Q}} \sum_{l \in \mathcal{L}_1,\, q|l} \frac{1}{\varphi(l)} + O\left(\frac{x}{(\log x)^2}\right) \\ &\le\; \mathrm{li}(x) \sum_{q \in \mathcal{Q}} \frac{1}{q-1} \sum_{qm \in \mathcal{L}} \frac{1}{\varphi(m)} + O\left(\frac{x}{(\log x)^2}\right). \end{aligned}$$

By 11.1 we have uniformly for $q \in \mathcal{Q}$ that

$$\sum_{qm \in \mathcal{L}} \frac{1}{\varphi(m)} \quad \begin{cases} = \;\xi\beta \log x + O(q \log(2x) x^{2\beta-1/2}), & \text{for } q < x^{1/2-2\beta} \\ \le \;\xi\beta \log x + \nu + O(q \log(2x) x^{\beta-1/2}), & \text{for } x^{1/2-2\beta} \le q \le x^{1/2-\beta} \\ = \;0, & \text{for } q > x^{1/2-\beta}. \end{cases}$$

Thus,

$$S_3 \;\le\; \xi\beta x \sum_{q \in \mathcal{Q}} \frac{1}{q-1} + O\left(\frac{x}{\log x}\right).$$

We estimate S_4 by using 10.4 with "m" chosen as m' and with "Q" being various powers of 2 so that the intervals $[Q, 2Q]$ cover the interval $(x^{1/2+\beta}, x^{1/2+2\beta})$. If h is an exceptional modulus in 10.4, we use the trivial estimate $\pi(x,h,1) \le x/h$. Thus, for $x \ge x_{m'}$,

$$\begin{aligned}
S_4 &= \sum_{\substack{h\in\mathcal{H} \\ q|h \text{ for some } q\in\mathcal{Q}}} \pi(x,h,1) \\
&\le (4/3+O(\beta))x \sum_{\substack{h\in\mathcal{H} \\ q|h \text{ for some } q\in\mathcal{Q}}} \frac{1}{\varphi(h)\log(x/h)} + O\left(\frac{x}{\log x}\right) \\
&\le (8/3+O(\beta))\frac{x}{\log x} \sum_{\substack{h\in\mathcal{H} \\ q|h \text{ for some } q\in\mathcal{Q}}} \frac{1}{\varphi(h)} + O\left(\frac{x}{\log x}\right) \\
&\le (8/3+O(\beta))\frac{x}{\log x} \sum_{q\in\mathcal{Q}} \frac{1}{q-1} \sum_{qm\in\mathcal{H}} \frac{1}{\varphi(m)} + O\left(\frac{x}{\log x}\right) \\
&= (8/3+O(\beta))\xi\beta x \sum_{q\in\mathcal{Q}} \frac{1}{q-1} + O\left(\frac{x}{\log x}\right).
\end{aligned}$$

For S_5 it is sufficient to use 10.1. Note that

$$\sum_{\substack{h\in\mathcal{H} \\ q|h \text{ for some prime } q>x^{1/2}}} \frac{1}{\varphi(h)} \le \sum_{\substack{x^{1/2}<q\le x^{1/2+2\beta} \\ q \text{ prime}}} \frac{1}{q-1} \sum_{t\le x^{2\beta}} \frac{1}{\varphi(t)}.$$

By Mertens' theorem, the first sum on the right is $O(\beta)$, and by 11.1, the second sum is $O(\beta \log x)$. Thus, the sum $\sum 1/\varphi(h)$ is $O(\beta^2 \log x)$, so that we have

$$\begin{aligned}
S_5 &\le 2x \sum_{\substack{h\in\mathcal{H} \\ q|h \text{ for some prime } q>x^{1/2}}} \frac{1}{\varphi(h)\log(x/h)} \\
&= O\left(\frac{x}{\log x} \sum_{\substack{h\in\mathcal{H} \\ q|h \text{ for some prime } q>x^{1/2}}} \frac{1}{\varphi(h)}\right) = O(\beta^2 x).
\end{aligned}$$

Putting together our estimates for S_1, S_2, S_3, S_4, S_5 we have that for $x \ge x_{m'}$,

$$\begin{aligned}
\sum_{\text{prime } r\le x} g(r) &\ge S_1 - S_2 - S_3 - S_4 - S_5 \\
&\ge \xi\beta x\left(1 - (11/3+O(\beta))\sum_{q\in\mathcal{Q}} \frac{1}{q-1}\right) + O(\beta^2 x) + O(x/(\log x)^{1/4}) \\
&\ge \xi\beta x\big(1 - (11/3+O(\beta))(3/11 - 1/m)\big) + O(x/(\log x)^{1/4}) \\
&= \xi\beta x(11/(3m) + O(\beta)) + O(x/(\log x)^{1/4}).
\end{aligned}$$

Thus, there is some absolute, computable, positive integer c_{15} such that if $m' = c_{15}m$ and $\beta = 1/m'$, we have

$$\sum_{\text{prime } r \le x} g(r) \ \ge\ \xi x/(mm') \ =\ \xi x/(c_{15}m^2)$$

for $x \ge X_m$, where $X_m \ge x_{m'}$ and X_m is also larger than an absolute constant. Using this with our upper bound for $\sum_{r \le x} g(r)^2$, we get the desired estimate for N, where we may choose $\delta_m = \xi^2/(c_{15}^2 m^5)$. This completes the proof of 11.2.

Remarks. By using the results of Bombieri–Friedlander–Iwaniec instead of 10.4 one can do better. In fact, by the method of Friedlander [18] we can not only replace "3/11" with "1/2" in 11.2, but the number of primes r satisfying the condition is of order of magnitude $\pi(x)$. However, the results of Bombieri–Friedlander–Iwaniec involve constants that are not effectively computable. If one is not concerned with effective constants, this stronger form of 11.2 would support the conclusion of 2.15 for $D > (\log n)^{1+\epsilon}$.

12. The existence of period systems

In this section we prove 2.15. We first show that there are many period pairs for n.

Proposition 12.1. *Let n be an integer, $n > 1$, and let w, y be real numbers. Each prime number r satisfies at least one of the following conditions:*

(i) *the element $(n \bmod r)$ of $\mathbf{F}_r$ is either zero or has multiplicative order at most w;*

(ii) *there is an integer m composed of primes at most y with $m \mid r-1$ and $m > w$;*

(iii) *there is an integer q with $q > y$ and $q^2 \mid r-1$;*

(iv) *there is a prime q such that $q > y$ and (r, q) is a period pair for n.*

Proof. If $(n \bmod r)$ does not belong to $\mathbf{F}_r^*$ then (i) holds. Suppose $(n \bmod r) \in \mathbf{F}_r^*$, and let m be the order of $(n \bmod r)$ in $\mathbf{F}_r^*$. Then m divides $r-1$, so if $m \le w$, then (i) holds. Suppose $m > w$. If m has no prime factor exceeding y, then (ii) holds. Suppose therefore that q is a prime factor of m with $q > y$; then q equals the order of $(n^{m/q} \bmod r)$. If q divides $(r-1)/m$, then (iii) holds. If q does not divide $(r-1)/m$, then the element $(n^{(r-1)/q} \bmod r) = (n^{m/q} \bmod r)^{(r-1)/m}$ has order q, and (iv) holds. This proves 12.1.

Let $\rho : \mathbf{R}_{\ge 0} \to \mathbf{R}_{>0}$ denote the Dickman–de Bruijn function, which satisfies (see [12])

$$\log \rho(u) = -u \cdot \log(u \log u) + O(u) \quad \text{for } u \ge 2. \tag{12.2}$$

Lemma 12.3. *Let x, u, v be real numbers with $x \geq 20$, $1 \leq v \leq u \leq \sqrt{(\log x) \log\log x}$, and put $y = x^{1/u}$, $w = y^v$. The number of prime numbers $r \leq x$ satisfying* 12.1(ii) *is at most*

$$O\left(u\pi(x)\Big(\frac{\rho(v)}{\log(2v)} + \rho(u)\Big)\right).$$

Proof. This is Theorem 2 from [23].

Proposition 12.4. *For all sufficiently large integers n, if x is a real number such that $x \geq (\log n)^{1+1/1800}$, then the number of prime numbers $r \leq x$ for which there does not exist a period pair (r, q) for n satisfying*

$$q \text{ is prime}, \qquad q > x^{1/(\log\log x)^2}$$

is at most $x/(\log x)^3$.

Proof. By 12.1, it suffices to show that when n is a sufficiently large integer and x is a real number with $x \geq (\log n)^{1+1/1800}$, the number of primes $r \leq x$ satisfying one of 12.1(i)–(iii), with $w = x^{1/\log\log x}$ and $y = x^{1/(\log\log x)^2}$, is at most $x/(\log x)^3$. We prove this by showing that the number of such primes r is $o\big(x/(\log x)^3\big)$ as $n \to \infty$.

If the prime r satisfies 12.1(i), then either $r \mid n$ or $r \mid n^m - 1$ for some integer m in $[1, w]$. Since the number of distinct prime divisors of an integer $k > 2$ is evidently smaller than $(\log k)/\log 2$, the number of primes r satisfying 12.1(i) is smaller than

$$\frac{\log n}{\log 2} + \sum_{m \leq w} m \cdot \frac{\log n}{\log 2} \;\leq\; w^2 \cdot \frac{\log n}{\log 2} \;\leq\; x^{1800/1801+o(1)} \;=\; o\big(x/(\log x)^3\big)$$

as $n \to \infty$.

To estimate the number of primes $r \leq x$ satisfying 12.1(ii) we apply 12.3 with $v = \log\log x$ and $u = v^2$; one finds that this number is at most

$$x/(\log x)^{(1+o(1))\log\log\log x} = o\big(x/(\log x)^3\big)$$

as $n \to \infty$.

The number of integers r with $1 < r \leq x$ satisfying 12.1(iii) is clearly at most $\sum_{q>y} x/q^2 < x/(y-1) = o\big(x/(\log x)^3\big)$ as $n \to \infty$.

This proves 12.4.

Let $\epsilon = 1/150$, let n be an integer with $n \geq 20$, and let x, u be real numbers with

$$x \;\geq\; (\log n)^{1+\epsilon/12} \;=\; (\log n)^{1+1/1800}, \quad u \;=\; (\log\log x)^2.$$

For a prime r, let $Q(r)$ denote the set of prime divisors q of $r-1$ with

$$x^{1/u} < q \le x^{1/2} \quad \text{and} \quad (r, q) \quad \text{is a period pair for } n.$$

Further, let $\mathcal{Q}$ denote the union of the sets $Q(r)$ over all primes $r \le x$. The interest of $\mathcal{Q}$ for us is that each subset $\mathcal{S}$ of $\mathcal{Q}$ corresponds to at least one period system for n with degree $\prod_{q\in\mathcal{S}} q$.

Proposition 12.5. *For all sufficiently large integers n and for all real numbers $x \ge (\log n)^{1+\epsilon/12}$, we have*

$$\sum_{q\in\mathcal{Q}} \frac{1}{q} \;>\; \frac{3-\epsilon}{11}.$$

Proof. Let

$\mathcal{A} = \{\text{prime } r \le x \,:\, \text{prime } q \mid r-1 \text{ implies } q \le x^{1/2} \text{ and } q \notin \mathcal{Q}\}$,

$\mathcal{B} = \{\text{prime } r \le x \,:\, \text{prime } q \mid r-1 \text{ implies } q \le x^{1/u} \text{ or } (r,q) \text{ is not a period pair for } n\}$.

Clearly $\mathcal{A} \subset \mathcal{B}$. We use 11.2, with "$m$" of that result being the current $11/\epsilon = 1650$; let $\delta = \delta_{1650}$. Suppose n is so large that 11.2 and 12.4 hold for all $x \ge (\log n)^{1+\epsilon/12}$. If $\sum_{q\in\mathcal{Q}} 1/q \le (3-\epsilon)/11$, then 11.2 implies that $\#\mathcal{A} \ge \delta x/(\log x)^2$. And so $\#\mathcal{B} \ge \delta x/(\log x)^2$. But 12.4 implies that $\#\mathcal{B} \le x/(\log x)^3$. These two inequalities for $\#\mathcal{B}$ are incompatible for large n. This contradiction completes the proof of 12.5.

With n, x, u as above, let N be an integer for which

$$6u\log x \;\le\; N \;\le\; \exp\big(2(\log x)^{3/5}(\log\log x)^{-3/2}\big). \tag{12.6}$$

For a bounded interval I, let $|I|$ denote the length of I.

Proposition 12.7. *For an integer N satisfying (12.6) and for $i = 1, 2, \ldots, N$, let*

$$I_i \;=\; [x^{(i-1)/N}, x^{i/N}), \quad M_i \;=\; x^{i/N}/i^2,$$

and

$$k_i \;=\; \begin{cases} 0, & \text{if } \#(I_i \cap \mathcal{Q}) < M_i \\ \min\{\#(I_i\cap\mathcal{Q}), \lfloor |I_i|/\log(x^{i/N})\rfloor\}, & \text{otherwise.} \end{cases}$$

For $i \le N/u$, $\#(I_i \cap \mathcal{Q}) = 0$, and for each $i = 1, 2, \ldots, N$, $k_i = 0$ or $k_i \ge M_i$.

Proof. Note that all primes $q \in \mathcal{Q}$ have $q > x^{1/u}$, so it follows that $\#(I_i \cap \mathcal{Q}) = 0$ for $i \le N/u$. Note too that for $i > 2$ we have

$$\frac{|I_i|}{\log(x^{i/N})} \;=\; \frac{x^{i/N}(1-x^{-1/N})}{(i/N)\log x} \;>\; \frac{x^{i/N}(\log x)/(2N)}{(i/N)\log x} \;>\; M_i,$$

where we use $x^{-1/N} < 1-(\log x)/(2N)$, which holds from (12.6) when $n \ge 20$. This proves 12.7.

Proposition 12.8. *For an integer N satisfying (12.6) and intervals I_i and integers k_i defined in 12.7, let $\mathcal{Q}_i$ denote the set of the least k_i primes in $I_i \cap \mathcal{Q}$. If n is sufficiently large we have*

$$\sum_{i=1}^{N} \sum_{q \in \mathcal{Q}_i} \frac{1}{q} > \frac{3}{11} - \frac{\epsilon}{10}.$$

Proof. The double sum here may be smaller than the sum $\sum_{q\in\mathcal{Q}} \frac{1}{q}$ in 12.5, the possible difference between them coming from two sources: intervals I_i with $0 < \#(I_i \cap \mathcal{Q}) < M_i$ and intervals I_i with $\#(I_i \cap \mathcal{Q}) > \lfloor |I_i| / \log(x^{i/N}) \rfloor$. By 12.7 we need only consider indices $i > N/u$. The sum of $1/q$ for primes q in intervals I_i with $\#(I_i \cap \mathcal{Q}) < M_i$ is at most

$$\sum_{i>N/u} \frac{M_i}{x^{(i-1)/N}} = \sum_{i>N/u} \frac{x^{1/N}}{i^2} < \frac{2u}{N} x^{1/N} \leq \frac{1}{3\log x} \mathrm{e}^{1/(6u)} < \frac{1}{\log x},$$

by the first inequality in (12.6). Thus, this contribution is $o(1)$ as $n \to \infty$, so is negligible.

The sum of $1/q$ for the largest $\#(I_i \cap \mathcal{Q}) - \lfloor |I_i| / \log(x^{i/N}) \rfloor$ primes q in an interval I_i with $\#(I_i \cap \mathcal{Q}) > \lfloor |I_i| / \log(x^{i/N}) \rfloor$ is estimated as follows. By the prime number theorem (see [24]), the total number of primes in I_i is

$$L_i + O\big(E(x^{i/N})\big),$$

where

$$L_i = \int_{x^{(i-1)/N}}^{x^{i/N}} \frac{\mathrm{d}t}{\log t} \quad \text{and} \quad E(z) = z / \exp\left(c_{16} (\log z)^{3/5} (\log\log z)^{-1/5}\right),$$

with c_{16} an effective, positive constant. As before, we may assume $i > N/u$. Note that

$$0 \leq L_i - \frac{|I_i|}{\log(x^{i/N})} \leq \frac{|I_i|}{\log(x^{(i-1)/N})} - \frac{|I_i|}{\log(x^{i/N})} = |I_i| \frac{N}{i(i-1)\log x}.$$

Further,

$$|I_i| = x^{(i-1)/N} \left(x^{1/N} - 1\right) < 2x^{(i-1)/N} \frac{\log x}{N}, \tag{12.9}$$

so that

$$L_i - \frac{|I_i|}{\log(x^{i/N})} < \frac{2x^{(i-1)/N}}{i(i-1)} = O(M_i).$$

Further, for $N/u < i \le N$,

$$E(x^{i/N}) \;\le\; \frac{x^{i/N}}{\exp\left(c_{16}u^{-3/5}(\log x)^{3/5}(\log\log x)^{-1/5}\right)} \;=\; \frac{x^{i/N}}{\exp\left(c_{16}(\log x)^{3/5}(\log\log x)^{-7/5}\right)},$$

so that from the upper bound for N in (12.6), $E(x^{i/N}) = O(M_i)$. Thus, the contribution in 12.5 from primes in I_i with $\#(I_i \cap \mathcal{Q}) > \lfloor |I_i|/\log(x^{i/N})\rfloor$ is

$$O\left(\sum_{N/u<i\le N} \frac{M_i}{x^{(i-1)/N}}\right),$$

a sum we have seen to be negligible. Thus, we have 12.8.

Proposition 12.10. *For an integer N satisfying (12.6) and integers k_i defined in 12.7, let S_i be the image of the interval*

$$(x^{(i-1)/N},\, x^{(i-1)/N} + k_i \log(x^{i/N}))$$

under the natural logarithm map. If n is sufficiently large, then

$$\sum_{i=1}^{N} \int_{S_i} \frac{\mathrm{d}t}{t} \;>\; \frac{3}{11} - \frac{\epsilon}{9}.$$

Proof. Since $\sum_{q\in\mathcal{Q}_i} 1/q \le k_i/x^{(i-1)/N}$, it follows from 12.8 that for n sufficiently large,

$$\sum_{i=1}^{N} \frac{k_i}{x^{(i-1)/N}} \;>\; \frac{3}{11} - \frac{\epsilon}{10}. \tag{12.11}$$

Further, if $S_i \ne \emptyset$, that is, if $k_i > 0$, then

$$\int_{S_i} \frac{\mathrm{d}t}{t} \;=\; \log\left(\frac{\log\left(x^{(i-1)/N} + k_i\log(x^{i/N})\right)}{\log\left(x^{(i-1)/N}\right)}\right).$$

Now, $\log(a+b) > \log(a) + b/a - (b/a)^2$ when $a, b > 0$, so that if $a > \mathrm{e}$ and $0 < b < a$,

$$\log\left(\frac{\log(a+b)}{\log a}\right) \;>\; \frac{b}{a\log a} - \frac{2}{\log a}\left(\frac{b}{a}\right)^2 \;=\; \frac{b}{a\log a}\left(1 - \frac{2b}{a}\right).$$

Hence,

$$\int_{S_i} \frac{\mathrm{d}t}{t} \;>\; \frac{k_i \log(x^{i/N})}{x^{(i-1)/N}\log(x^{(i-1)/N})}\left(1 - \frac{2k_i \log(x^{i/N})}{x^{(i-1)/N}}\right)$$
$$\;>\; \frac{k_i}{x^{(i-1)/N}}\left(1 - \frac{2k_i \log(x^{i/N})}{x^{(i-1)/N}}\right).$$

Note that, using the definition of k_i, (12.9), and (12.6),

$$\frac{2k_i \log(x^{i/N})}{x^{(i-1)/N}} \;\le\; \frac{2|I_i|}{x^{(i-1)/N}} \;<\; \frac{4\log x}{N} \;<\; \frac{1}{u},$$

so that

$$\int_{S_i} \frac{\mathrm{d}t}{t} \;>\; \frac{k_i}{x^{(i-1)/N}}(1-u^{-1}).$$

Thus, 12.10 follows from (12.11) for sufficiently large n.

Proof of Proposition 2.15. Let $\epsilon = 1/150$, let n be an integer so large that 12.10 holds, and let D be an integer satisfying

$$D \;>\; (\log n)^{11/6+\epsilon} \;=\; (\log n)^{46/25}.$$

Let $x = D^{6/11-\epsilon/4}$ so that $x > (\log n)^{1+1/1800}$, let $u = (\log\log x)^2$, and let integer N satisfy (12.6). Let $D' = D\exp(2u(\log x)/N)$ and let S be the additive semigroup generated by

$$\bigcup_{i=1}^{N} \frac{1}{\log D'} S_i,$$

where S_i is as in 12.10. Note that if $S_i \neq \emptyset$ we have $x^{(i-1)/N} \le x^{1/2}$, so that

$$\frac{\log(x^{i/N})}{\log D'} \;\le\; \left(\frac{1}{2}+\frac{1}{N}\right)\frac{\log x}{\log D} \;=\; \left(\frac{1}{2}+\frac{1}{N}\right)\left(\frac{6}{11}-\frac{\epsilon}{4}\right) \;<\; \frac{3}{11}-\frac{\epsilon}{9}$$

for sufficiently large n; that is, $S_i/\log D' \subset (0, 3/11-\epsilon/9)$. We suppose that n is so large. Thus, from 12.10 and the fact that the intervals S_i are disjoint,

$$\int_0^{3/11-\epsilon/9} \frac{\chi_S(t)}{t}\,\mathrm{d}t \;\ge\; \sum_{i=1}^{N}\int_{S_i/\log D'} \frac{\mathrm{d}t}{t} \;=\; \sum_i \int_{S_i}\frac{\mathrm{d}t}{t} \;>\; \frac{3}{11}-\frac{\epsilon}{9}.$$

It thus follows from 9.1 that $1 \in S$. Hence, there is a finite subset F of $\bigcup_i S_i$ and positive integers $\kappa(f)$ for each $f \in F$ such that

$$\sum_{f\in F} \kappa(f) f \;=\; \log D'.$$

Let $F_i = F \cap S_i$ for $i = 1, 2, \ldots, N$, and let

$$\kappa_i \; = \; \sum_{f \in F_i} \kappa(f).$$

Then, using $S_i = \emptyset$ for $i \le N/u$ from 12.7,

$$\begin{aligned} \sum_{i=1}^{N} \kappa_i \; &= \; \sum_i \sum_{f \in F_i} \kappa(f) \; \le \; \sum_i \frac{1}{\log(x^{(i-1)/N})} \sum_{f \in F_i} \kappa(f) f \\ &< \; \frac{1}{\log\left(x^{1/u - 1/N}\right)} \sum_{f \in F} \kappa(f) f \; = \; \frac{\log D'}{(1/u - 1/N) \log x} \; < \; 2u, \end{aligned} \tag{12.12}$$

the last inequality holding when n is sufficiently large. If $S_i \ne \emptyset$, then 12.7 implies that $k_i \ge M_i$, so that $k_i > x^{1/u}/N^2 > 2u > \kappa_i$. Thus, for each i with $\kappa_i > 0$ there are at least κ_i distinct primes in $\mathcal{Q}_i$. Label the least such primes $q_{1,i}, q_{2,i}, \ldots, q_{\kappa_i,i}$ and let

$$d = \prod_{i=1}^{N} \prod_{j=1}^{\kappa_i} q_{j,i}.$$

If $r_{j,i} \le x$ is a prime with $q_{j,i} \in Q(r_{j,i})$, then

$$\mathcal{P} = \{(r_{j,i}, q_{j,i}) \; : \; i = 1, \ldots, N, \, j = 1, \ldots, \kappa_i\}$$

is a period system for n with degree d. We have

(12.13)

$$\begin{aligned} |\log D' - \log d| \; &= \; \left| \sum_{f \in F} \kappa(f) f - \sum_{i=1}^{N} \sum_{j=1}^{\kappa_i} \log(q_{j,i}) \right| = \left| \sum_{i=1}^{N} \left(\sum_{f \in F_i} \kappa(f) f - \sum_{j=1}^{\kappa_i} \log(q_{j,i}) \right) \right| \\ &< \; \sum_i \kappa_i \left(\log(x^{i/N}) - \log(x^{(i-1)/N}) \right) \; = \; \frac{\log x}{N} \sum_i \kappa_i \; < \; \frac{2u \log x}{N}, \end{aligned}$$

using (12.12). Thus,

$$D \; = \; D' \exp(-2u(\log x)/N) \; < \; d \; < \; D' \exp(2u(\log x)/N) \; < \; D(1 + 6u(\log x)/N).$$

By choosing N near the upper end of the interval in (12.6), we have 2.15.

Remark. We may use the same argument to show that there are many integers in $[D, 2D)$ that are degrees of period systems for n, in fact more than $D/\exp(5(\log\log D)^3)$, once n

is sufficiently large and $D > (\log n)^{46/25}$. We choose N near the lower limit in (12.6) and we let $D' = \sqrt{2}D$. For each i with $\kappa_i > 0$ choose κ_i primes from $\mathcal{Q}_i$ and let d denote the product of all of these primes over all choices for i. Then, as in (12.13) and by our choice of N,

$$|\log D' - \log d| \;<\; \frac{2u\log x}{N} \;=\; \frac{1}{3}.$$

Since $e^{1/3} < \sqrt{2}$, it follows that $D < d < 2D$. It remains to count the number of choices for d in the argument. Since $\#\mathcal{Q}_i \geq M_i$ when $\kappa_i > 0$, the number of choices for d is at least

$$\prod_{\kappa_i>0}\binom{\lceil M_i\rceil}{\kappa_i} \;\geq\; \prod_{\kappa_i>0}\left(\frac{M_i}{\kappa_i}\right)^{\kappa_i} \;=\; \prod_{\kappa_i>0}\left(\frac{x^{i/N}}{i^2\kappa_i}\right)^{\kappa_i} \;>\; \frac{D}{\prod_{\kappa_i>0}(i^2\kappa_i)^{\kappa_i}}.$$

Now, by (12.12),

$$\prod_{\kappa_i>0} i^{2\kappa_i} \;<\; N^{2\sum_i \kappa_i} \;<\; N^{4u}$$

and

$$\prod_{\kappa_i>0} \kappa_i^{\kappa_i} \;<\; \left(\sum_i \kappa_i\right)^{\sum_i \kappa_i} \;<\; (2u)^{2u}.$$

Thus, the number of choices for d exceeds $D/(2uN^2)^{2u}$ and it remains to note that

$$(2uN^2)^{2u} \;<\; \exp\left(5(\log\log x)^3\right) \;<\; \exp\left(5(\log\log D)^3\right).$$

References

1 L. M. Adleman and H. W. Lenstra, Jr., *Finding irreducible polynomials over finite fields*, Proc. 18th Annual ACM Symp. on Theory of Computing (STOC), Berkeley, May 28–30, 1986, 350–355.

2. M. Agrawal, N. Kayal, and N. Saxena, *PRIMES is in P*, Ann. of Math. **160** (2004), 781–793.

3. W. R. Alford, A. Granville, and C. Pomerance, *There are infinitely many Carmichael numbers*, Ann. of Math. (2) **139** (1994), 703–722.

4. M. F. Atiyah and I. G. Macdonald, *Introduction to commutative algebra*, Addison-Wesley, Reading, Mass., 1969.

5. A. Balog, *$p + a$ without large prime factors*, in Seminar on number theory, 1983–1984 (Talence, 1983/1984), Exp. No. 31, 5 pp., Univ. Bordeaux I, Talence, 1984.

6. D. J. Bernstein, *Fast multiplication and its applications*, in J. P. Buhler, P. Stevenhagen (eds), *Algorithmic number theory*, 325–384, MSRI Publications **44**, Cambridge U. Press, New York, 2008.

7. D. J. Bernstein, *Proving primality in essentially quartic random time*, Math. Comp. **76** (2007), 389–403.

8. D. J. Bernstein, H. W. Lenstra, Jr., and J. Pila, *Detecting perfect powers by factoring into coprimes*, Math. Comp. **76** (2007), 385–388.

9. D. Bleichenbacher, *The continuous postage problem*, unpublished manuscript, 2003.

10. A. Bostan, P. Flajolet, B. Salvy, and É. Schost, *Fast computation of special resultants*, J. Symbolic Comput. **41** (2006), 1–29.

11. R. P. Brent, *Fast multiple-precision evaluation of elementary functions*, J. Assoc. Comput. Mach. **23** (1976), 242–251.

12. N. G. de Bruijn, *The asymptotic behaviour of a function occurring in the theory of primes*, J. Indian Math. Soc. (N.S.) **15** (1951), 25–32.

13. J. W. S. Cassels, *An introduction to the geometry of numbers*, Springer-Verlag, Berlin, second edition, 1971.

14. H. Davenport, *Multiplicative number theory*, Second edition (revised by H. L. Montgomery), Graduate Texts in Mathematics, 74. Springer-Verlag, New York–Berlin, 1980.

15. J.-M. Deshouillers and H. Iwaniec, *Kloosterman sums and Fourier coefficients of cusp forms*, Invent. Math. **70** (1982/83), 219–288.

16. J.-M. Deshouillers and H. Iwaniec, *On the Brun-Titchmarsh theorem on average*, in Topics in classical number theory (G. Halász, ed.), Vol. I, (Budapest, 1981), 319–333, Colloq. Math. Soc. János Bolyai, **34**, North-Holland, Amsterdam, 1984.

17. J. Farkas, *Theorie der einfachen Ungleichungen*, J. Reine Angew. Math. **124** (1902), 1–27.

18. J. B. Friedlander, *Shifted primes without large prime factors*, in Number theory and applications (R. A. Mollin, ed.), Kluwer Academic Publishers, Dordrecht, 1989, pp. 393–401.

19. S. Lang, *Algebra*, revised third edition, Springer-Verlag, New York, 2002.

20. V. F. Lev, *The continuous postage stamp problem*, J. London Math. Soc. **73** (2006), 625–638.

21. R. Lidl and H. Niederreiter, *Finite fields*, Addison-Wesley Publishing company, Read-

ing (Mass.), 1983.

22. H. L. Montgomery and R. C. Vaughan, *The large sieve*, Mathematika **20** (1973), 119–134.

23. C. Pomerance and I. E. Shparlinski, *Smooth orders and cryptographic applications*, Algorithmic Number Theory (Sydney, 2002), 338–348, Lecture Notes in Comput. Sci. **2369**, Springer, Berlin, 2002.

24. G. Tenenbaum, *Introduction to analytic and probabilistic number theory*, Cambridge University Press, Cambridge, UK, 1995.

25. N. M. Timofeev, *The Vinogradov–Bombieri theorem* (in Russian), Mat. Zametki **38** (1985), 801–809, 956.

26. J. von zur Gathen and J. Gerhard, *Modern computer algebra*, Cambridge University Press, Cambridge, 1999.