Canad. J. Math. Vol. **70** (1), 2018 pp. 117–141
http://dx.doi.org/10.4153/CJM-2017-023-x
© Canadian Mathematical Society 2017

Smooth Polynomial Solutions to a Ternary Additive Equation

Junsoo Ha

Abstract. Let $\mathbf{F}_q[T]$ be the ring of polynomials over the finite field of q elements and Y a large integer. We say a polynomial in $\mathbf{F}_q[T]$ is Y-smooth if all of its irreducible factors are of degree at most Y. We show that a ternary additive equation $a+b=c$ over Y-smooth polynomials has many solutions. As an application, if S is the set of first s primes in $\mathbf{F}_q[T]$ and s is large, we prove that the S-unit equation $u+v=1$ has at least $\exp(s^{1/6-\epsilon}\log q)$ solutions.

1 Introduction and Statement of Results

The S-unit equation has been studied over the years. For the simplest case, let S be a finite set of (rational) primes. We consider an additive equation $a+b=c$ over integers, where all prime factors of abc lie in S and a, b, c are coprime. This equation is a special case of the binary S-unit equation (so-called because it is usually written as $u+v=1$ with $u=a/c$ and $v=b/c$). One of the classic theorems on S-unit equations is the famous theorem of Siegel: every binary S-unit equation has finitely many solutions.

After Siegel, several upper bounds on the number of solutions to general or special S-unit equations were established. Evertse [4] established a uniform upper bound that only depends on the cardinality of S (and on the extension degree of the number field in the number field case.) Let s be the cardinality of S. In the case of $a+b=c$, Evertse's bound showed that the equation has at most $3\cdot 7^{2s}$ solutions. On the other hand, S-unit equations sometimes have many solutions, particularly when S contains many small primes. Erdös, Stewart, and Tijdeman [3] showed that there exist arbitrarily large sets S for which the S-unit equation $u+v=1$ has at least $\exp((4-\epsilon)(s/\log s)^{1/2})$ solutions. Konyagin and Soundararajan [13] refined the argument to produce arbitrarily large sets S for which the equation $a+b=c$ has at least $\exp(s^{2-\sqrt{2}-\epsilon})$ solutions. The set S constructed by these authors starts with the set of the first few primes, and expands by a small number of primes accordingly as the argument proceeds, and may contain a large prime.

When S is the set of the first s primes, the problem is particularly interesting in connection with smooth numbers. We say an integer is *y-smooth* if all of its prime factors are less than or equal to y, and we write $\Psi(x,y)$ for the number of y-smooth integers up to x. For clarity, we say a solution to $a+b=c$ is *primitive* if $\gcd(a,b,c)=1$. Then

Received by the editors November 4, 2015; revised July 6, 2017.
Published electronically October 5, 2017.
AMS subject classification: **11T55**, 11D04, 11L07, 11T23.
Keywords: smooth number, polynomial over a finite field, circle method.

https://doi.org/10.4153/CJM-2017-023-x Published online by Cambridge University Press

when S is the set of primes up to y, each solution to the S-unit equation $u + v = 1$ corresponds to a primitive solution to an additive equation $a+b = c$ over y-smooth numbers. Lagarias and Soundararajan studied this additive equation [15,16]. Let $N^*(x, y)$ be the number of solutions to $a + b = c$, and where $\gcd(a, b, c) = 1$, a, b, c are all y-smooth and less than or equal to x. Then they investigated the equation by the circle method to show that, in a rough form,

$$N^*(x, (\log x)^\kappa) \sim C(\kappa)\frac{\Psi(x, (\log x)^\kappa)^3}{x}$$

under the Generalized Riemann Hypothesis (GRH) when $\kappa > 8$ and x is large and where $C(\kappa)$ is a constant depending on κ.

We let S be the set of primes up to y and $s = \pi(y) \sim y/\log y$. Then for given $\kappa > 1$, we have a smooth asymptote $\Psi(x, y) \sim x^{1-1/\kappa+o(1)}$ when $y = (\log x)^\kappa$. We denote by $S(x, y)$ the set of y smooth numbers up to x. Then if y is large and $\kappa > 8$, we have

$$\text{(1.1)} \quad \big\{(a, b, c) \in S(x, y)^3 : a + b = c,\ \gcd(a, b, c) = 1\big\} \gg x^{2-3/\kappa+o(1)} \geq \exp(s^{1/\kappa}).$$

Thus, as a corollary of their work, Lagarias and Soundararajan showed that under GRH the binary S-unit equation has at least $\exp(s^{1/8-\epsilon})$ solutions when S is the set of the first s primes and s is large.

We remark here that a few unconditional results have been made in recent years; Drappeau [1] proved unconditionally the number of solutions to $a + b = c$ in the range $\exp(c\sqrt{\log x}\log\log x) \leq y \leq x$ in terms of exponential sums as Lagarias and Soundararajan did, and some nontrivial estimates on major arcs in a wide range of x and y. Very recently, Harper [8] gave an unconditional proof in the range $(\log x)^\kappa \leq y \leq x$ for some κ, using the result of Bourgain on the restriction theory for the minor arcs.

We turn our attention to the polynomials over a finite field. Let q be a prime power and $\mathbf{F}_q$ be the field of q elements. Throughout this paper, we take T as an indeterminate and we simply say m is a polynomial if m is a polynomial in $\mathbf{F}_q[T]$. The analogy of the ring of integers and the ring of polynomials over a finite field has been studied by numerous authors, and one may ask if a comparable result in the ring of polynomials can be established. We say a solution to $a + b = c$ is *primitive* if $\gcd(a, b, c) = 1$, and *separable* if not all of a, b, c are p-th power, where p is the characteristic of $\mathbf{F}_q$. In this paper, we prove the following theorem.

Theorem 1.1 *Let Y be a large integer and S be the set of irreducible polynomials of degree at most Y. Then the S-unit equation $a + b = c$ with all irreducible factors of abc lying in S has at least $q^{|S|^{1/6-\epsilon}}$ separable primitive solutions.*

This result improves that of the previous paper [5], where we obtained $1/8 - \epsilon$ instead of $1/6 - \epsilon$.

We remark here that a few features of the xyz conjecture of Lagarias and Soundararajan [15] need a slight modification in the case of $\mathbf{F}_q[T]$. Over the integers, Lagarias and Soundararajan defined the smoothness exponent κ for a solutions to $a + b = c$ by

$$\text{(1.2)} \qquad \kappa^{\mathrm{LS}}(a, b, c) = \frac{\log \max_{p|abc} |p|}{\log\log \max(|a|, |b|, |c|)}.$$

Then the original xyz smoothness exponent is defined by

$$\kappa_0^{\mathrm{LS}} = \liminf_{\substack{\gcd(a,b,c)=1\\ \max(|a|,|b|,|c|)\to\infty}} \kappa^{\mathrm{LS}}(a,b,c).$$

Then they conjectured that κ_0 is finite and that $\kappa_0^{\mathrm{LS}} = 3/2$. As mentioned earlier, the finiteness part was proved by Harper.

In $\mathbf{F}_q[T]$, we define the norm of polynomial (with respect to prime $1/T$) by

$$|m| = q^{\deg m}, \tag{1.3}$$

which is used to compute the size of a solution. We also define the analogue for smooth numbers; we say a polynomial is Y-smooth if all of its prime factors have degree less than or equal to Y. Then we can consider the distribution function $\Psi(X,Y)$, the number of Y-smooth polynomials of degree X. The work of Manstavičius [18,19] can be used to prove $\Psi(X,Y) = q^{X(1-1/\kappa+o(1))}$ when $\kappa = \log_q X/Y$ under some range of X and Y (see (2.1) for the precise condition). This formula suggests that the natural logarithms in (1.2) should be replaced by the base q logarithms. Therefore in $\mathbf{F}_q[T]$, we can define

$$\kappa(a,b,c) = \frac{\log_q \max_{\varpi|abc}|\varpi|}{\log_q \log_q \max(|a|,|b|,|c|)} = \frac{\max_{\varpi|abc}\deg\varpi}{\log_q \max(\deg a, \deg b, \deg c,)} \tag{1.4}$$

where ϖ denotes an irreducible polynomial, and we want to define

$$\kappa_0^{\mathrm{naive}} = \liminf_{\substack{\gcd(a,b,c)=1\\ \max(|a|,|b|,|c|)\to\infty}} \kappa(a,b,c)$$

in a similar way. Then we expect that the heuristics on smooth numbers should be applied with a minor change. However, we need to consider the following two examples, which only arise in the function field case.

The first example comes from the separability. If (a,b,c) is a primitive solutions to $a+b=c$, so is (a^p, b^p, c^p), where p is the characteristic of $\mathbf{F}_q$. Then raising by the p-th power of the solution, (a,b,c) increases the maximum degree of the solution by p times, whereas the maximum degree of the prime divisors of abc remains the same. Then the xyz exponent

$$\kappa(a^{p^k}, b^{p^k}, c^{p^k})$$

clearly converges to 0 for any fixed solution a, b, c as $k \to \infty$. To remedy such a case, we call the solution *separable* when at least one of a, b, c is not the p-th power when p is the characteristic, and we require that the solutions should be separable.

The next example is more pathological and seems unique to $\mathbf{F}_q[T]$. For any n, consider the additive equation $a+b=c$ when $(a,b,c) = (T^{q^n-1}, -1, T^{q^n-1}-1)$. From the fact that $\mathbf{F}_{q^n}$ is a splitting field of $T^{q^n} - T$ over $\mathbf{F}_q$, we have the identity

$$T^{q^n} - T = \prod_{\substack{\deg f|n\\ f\text{:irreducible}}} f(T),$$

and thus we observe that all prime factors of abc have degree at most n. Then this family of solutions to $a+b=c$ satisfies $\max_{\varpi|abc}|\varpi| = q^n$ whereas $\max(|a|,|b|,|c|) = q^{q^n-1}$. As n grows, the corresponding smoothness exponent converges to 1, and thus

https://doi.org/10.4153/CJM-2017-023-x Published online by Cambridge University Press

$\kappa_0^{\text{naive}} \le 1$. On the contrary, one can show that $\kappa_0^{\text{naive}} \ge 1$, as a corollary of *abc* conjecture for function fields (see [21, Theorem 7.17]). Thus we easily deduce that $\kappa_0^{\text{naive}} = 1$. However, it is not only inconsistent with the heuristic conjecture that $\kappa_0^{\text{LS}} = 3/2$, but it also may not be applied to producing many S-unit equations when S is the set of the first few primes, which is one of the main implications of the work of Lagarias and Soundararajan. It may be noted that this particular family of solutions is insufficient to produce many solutions to an S-unit equation. Therefore, we may refine the conjecture of Lagarias and Soundararajan for the rational function field case in the following new form.

Conjecture 1.2 *Let (a, b, c) be a triple in the ring of polynomials over $\mathbf{F}_q$, and let H be a large positive integer. Let $B(H) = \{(a, b, c) : \max(\deg a, \deg b, \deg c) \le H\}$. We consider the set of all primitive and separable triples that are solutions to $a + b = c$, and denote those with bounded smoothness exponent $\widetilde{\kappa}$ by*

$$(1.5)\quad A(\widetilde{\kappa}) = \Big\{(a, b, c) : a + b = c,\ \gcd(a, b, c) = 1,\ (a, b, c) \textit{ is separable and } \kappa(a, b, c) < \widetilde{\kappa}\Big\}.$$

The new xyz exponent $\kappa_0^{\text{new}} > 0$ is defined by

$$(1.6)\qquad \kappa_0^{\text{new}} = \inf\left\{\kappa : \liminf_{H\to\infty} \frac{\log \#\left(A(\kappa) \cap B(H)\right)}{\log \# B(H)} > 0\right\}.$$

Then $\kappa_0^{\text{new}} = 3/2$.

In other words, κ_0^{new} is the least exponent so that for any $\widetilde{\kappa} > \kappa_0^{\text{new}}$, there is $\delta > 0$ such that for all sufficiently large H, the equation $a + b = c$ with the degrees bounded by H and the smoothness exponent at most $\widetilde{\kappa}$ has at least $q^{\delta H}$ separable and primitive solutions. We can show that for any $\widetilde{\kappa} > \kappa_0^{\text{new}}$, there are at least $\exp(s^{1/\widetilde{\kappa}})$ (separable) solutions to the S-unit equation when S is the set of first s primes and s is large. The results of Lagarias and Soundararajan and of Harper can be migrated to this setting, showing that over the integers, the infimum of the exponent κ for which the equation $a + b = c$ has at least x^δ solutions for sufficiently large y and $y = (\log x)^\kappa$ is at most 8 under GRH, and is unconditionally finite. In connection with this new conjecture, our result can be stated as follows.

Theorem 1.3 *In $\mathbf{F}_q[T]$, the new xyz exponent $\kappa_0^{\text{new}} \le 6$.*

In the next section, we shall state a more technical version of this theorem. Meanwhile, we would like to survey some technical aspects of the theorem.

We note two features of the above theorem compared with the work of Lagarias and Soundararajan. The first is that it is an unconditional theorem and this property is rather clear, because we have a direct substitute for the GRH in the case of function fields, namely the Weil bound on curves over finite fields. On the other hand, the obvious decrease in the exponent from 8 to 6 is the new part of this theorem. The key element here is the use of well-factorizability of the smooth polynomials, which stems from the work of Harper [6] and an unpublished work of Soundararajan, who used a similar technique to produce many solutions to certain S-unit equation.

https://doi.org/10.4153/CJM-2017-023-x Published online by Cambridge University Press

2 Main Technical Result and the Idea of the Proof

In this section, we state our main theorem more precisely, and briefly survey the history of some technical elements of the proof that appear in the remaining sections.

The proof uses several major ingredients from previous authors. The basic setup of this paper is based on the work of Lagarias and Soundararajan, who used the circle method in the additive problem over smooth numbers, and their techniques such as character decomposition, and the major and minor arc estimates work fairly parallel in our main theorem.

Another underlying setup is the circle method in the polynomial ring over a finite field. One of the first studies is due to Hayes [10] who proved that a polynomial in $\mathbf{F}_q[T]$ can be written as the sum of three irreducible polynomials, an analogue of the ternary Goldbach problem. Another nice application of the circle method can be found in the more recent work of Liu and Wooley [17] who studied Waring's problem over $\mathbf{F}_q[T]$.

The central advantage of this paper relies on the observation of Harper [6] in the study of the Diophantine equation $a + 1 = c$ where all prime factors of ac lie in a set S. Harper observed that the average behavior of smooth numbers in an arithmetic progression can be managed in terms of bilinear form due to well-factorizability of the set of smooth numbers, *i.e.*, the ability to factor into any two sets of desirable sizes. In this way, he managed to prove that the character sums on average over certain sets of smooth numbers can be well estimated. Soundararajan gave a reformulation of Harper's idea in his unpublished note, and his setup is closely related to our new setting.

2.1 Main Technical Result: Counting Certain Solutions to $a + b = c$

Let $\mathbf{F}_q$ be the finite field of q elements, and $\mathbf{F}_q[T]$ be the ring of polynomials. We write $\mathbf{M}$ for monic polynomials, and $\mathbf{M}_0$ for squarefree monic polynomials. Throughout this paper, ϖ represents a monic irreducible polynomial, and m represents a monic polynomial. We usually use a capital letter to denote a positive integer, and lowercase for polynomials. To emphasize the parallel argument between $\mathbf{Z}$ and $\mathbf{F}_q[T]$, we adopt some handy notation from the work of Liu and Wooley. We write $\widehat{Z} = q^Z$ for any real number $Z > 1$. Similarly, we write $\mathcal{L}(Z) = \max(\log_q Z, 1)$. Obviously, $\mathcal{L}(\widehat{Z})$ reduces to Z, but we sometimes keep the former form when we have a comparable form of equations over the ring of integers.

Let X and Y be large and we write $\psi(X, Y)$ for the number of monic Y-smooth polynomials of degree X. We are mainly interested in the range when $\widehat{Y} = \mathcal{L}(\widehat{X})^\kappa$ for some $\kappa > 1$. The work of Manstavičius [18, 19] can be applied in our range so that if $\widehat{Y} = \mathcal{L}(\widehat{X})^\kappa$ as a particular case,

$$\psi(X, Y) = \widehat{X}^{1-1/\kappa+o(1)}, \tag{2.1}$$

when Y, X/Y, and $\widehat{Y}/\mathcal{L}(\widehat{X})$ are large. This result is comparable to the counting function of y-smooth integers up to x, which is $\Psi(x, y) = x^{1-1/\kappa+o(1)}$ when $y = (\log x)^\kappa$ and y is large and $\kappa > 1$ is fixed.

https://doi.org/10.4153/CJM-2017-023-x Published online by Cambridge University Press

Instead of $a + b = c$, we will use $a + b = 2c$, to maintain that all variables be monic. Heuristically, the number of solutions to $a + b = 2c$ can be expected to be of size $\asymp \psi(X, Y)^3/\widehat{X}$, because the choice of each variable is considered to be $\psi(X, Y)$, and the chance of two randomly chosen monic polynomials of the same leading coefficient being equal can be thought of as $1/\widehat{X}$. In this point of view, we may conjecture that the number of solutions is $\asymp \psi(X, Y)^3/\widehat{X} = \widehat{X}^{2-3/\kappa+o(1)}$, and we have at least $\gg \widehat{X}^{\delta}$ solutions if $\kappa > 3/2$. For a lower bound, suppose Y divides X and Y is large. Instead of using the whole set of Y-smooth polynomials of degree X, we consider the subset $\mathcal{S}_1$ defined by

$$\mathcal{S}_1 = \Big\{ m \in \mathbf{M}_0 : \deg m = X, \varpi \mid m \Rightarrow \deg \varpi = Y \Big\}.$$

Finding the solutions over $\mathcal{S}_1$ is useful, because we want to establish the lower bound for the number of solutions, and the cardinality of the set $\mathcal{S}_1$ is still fairly large. Indeed, we let $K = X/Y$ be the number of prime factors for elements in $\mathcal{S}_1$. Then for $\widehat{Y} = \mathcal{L}(\widehat{X})^{\kappa}$ and for large X, we have

$$|\mathcal{S}_1| = \binom{\pi(Y)}{K} = \frac{\pi(Y)^{K(1+o(1))}}{K!} = (\widehat{Y}/K)^{K(1+o(1))} = \widehat{X}^{1-1/\kappa+o(1)},$$

where $\pi(Y)$ is the number of irreducible polynomials of degree Y, and it is known that $\pi(Y) \sim \widehat{Y}/\mathcal{L}(\widehat{Y})$ (see Lemma 3.1). Therefore we solve the equation $a + b = 2c$ over $a, b, c \in \mathcal{S}$ and we still expect that the equation has many solutions. Indeed, we have the following theorem.

Theorem 2.1 *Let Y be a large integer, $\kappa > 6$, and let X be an integer such that X is divisible by Y and $\mathcal{L}(\widehat{X})^{\kappa} \le \widehat{Y}$. Then the number of solutions to $a + b = 2c$ is*

$$N(\mathcal{S}_1) = \mathfrak{S}\frac{|\mathcal{S}_1|^3}{\widehat{X}}\Big(1 + O\Big(\frac{1}{\widehat{Y}^{1/2}}\Big)\Big),$$

where the singular series $\mathfrak{S}$ is defined by

$$\mathfrak{S} = \prod_{\varpi}\Big(1 - \frac{1}{(|\varpi| - 1)^2}\Big),$$

where the product is taken over all monic irreducible polynomials of any degree.

We may impose the coprimality condition without much loss.

Theorem 2.2 *Retain the assumption of Theorem 2.1. Then the number of primitive solutions to $a + b = 2c$ is*

$$N^*(\mathcal{S}_1) = N(\mathcal{S}_1)\Big(1 + O\Big(\frac{1}{\widehat{Y}^{1/2}}\Big)\Big).$$

We remark that from the construction, the separability is already attained.

2.2 A Remark on Applications to the Number Field Case

The reader may ask if the original problem of finding many solutions to S-unit equations over the integers can be attacked by the methods used in this paper. For instance,

https://doi.org/10.4153/CJM-2017-023-x Published online by Cambridge University Press

if we assume GRH, we choose a parameter y large and $\kappa > 6$ so that one may write $a + b = c$ with a, b, c being of the form $p_1 \cdots p_k$, where p_i are primes in the interval $[y/2, y]$ and k is in the range comparable to $y^{1/\kappa}$. Indeed, one obtains a similar estimate on the minor arcs in such a case. On the other hand, technical difficulties arise on the major arcs, because we no longer have a sharp estimate for the exponential sums over prime like (5.1) in Proposition 5.1. This is partly because, in the function field case, all polynomials that appear in the exponential sum are of constant size, whereas for the case of integers, the size varies to the scale of 2^z, and we cannot use a combinatorial argument without producing large errors. This problem does not arise in the case of Lagarias and Soundararajan, because the counting function for smooth numbers is smooth, whereas the variant accumulative function that we used as the product of k primes in a short interval is not.

The reason we took $\mathcal{S}_1$ instead of all smooth polynomials is that we need the set $\mathcal{S}$ to be decomposed into the multiplication of two sets of the desirable size in order to use the well-factorizability of the smooth polynomials. On the other hand, Harper [7] recently proved a Bombieri–Vinogradov type estimate for smooth numbers, that was later improved by Drappeau [2]. In their proofs, they showed how we can use the well-factorizability of smooth numbers using Fourier analysis when the set itself cannot be decomposed directly into two sets, but each element can be decomposed into the desired size [7, §4.2].

Briefly speaking, when we decompose a smooth integer n into $n_1 n_2$, where n_1 is the factor of the desired size, the range of n_2 depends on the variable n_1 and cannot be summed separately. However, when we use the Mellin transform on the indicator function of the interval of the desired size, the factors n_1^{-s} and n_2^{-s} are separated, and we can use the well-factorizability directly on the set of y-smooth numbers in each dyadic interval. Combined with their new technique, we can obtain (1.1) for $\kappa > 6$. We leave it for a future paper.

3 Number Theory in a Polynomial Ring

The parallel theory between the ring of integers and the ring of polynomial rings over finite fields was well known. In this section, we cite a few relevant results in $\mathbf{F}_q[T]$ whose counterparts in $\mathbf{Z}$ are familiar to number theorists. We start from the prime number theorem for polynomials.

Lemma 3.1 *Let N be an integer, and write $\pi(N)$ for the number of monic irreducible polynomials of degree N. Then*

$$\pi(N) = \sum_{\deg \varpi = N} 1 = \frac{\widehat{N}}{\mathcal{L}(\widehat{N})}\big(1 + O(\widehat{N}^{-1/2})\big).$$

Moreover, $\pi(N) \le \widehat{N}/\mathcal{L}(\widehat{N})$.

Proof See [21, Theorem 2.2]. The last assertion is from

$$\mathcal{L}(\widehat{N})\pi(N) \le \sum_{D|N} \mathcal{L}(\widehat{D})\pi(D) = N. \qquad \blacksquare$$

https://doi.org/10.4153/CJM-2017-023-x Published online by Cambridge University Press

Next, we introduce the congruence characters [9]. Let g be a monic polynomial and l an integer. We define the dual of a polynomial $f \in \mathbf{F}_q[T]$ by

$$f^\star = T^{-\deg f} f(T) \in \mathbf{F}_q\Big[\frac{1}{T}\Big]. \tag{3.1}$$

Then we define the equivalence relation $\mathcal{R}_{g,l}$ on the set of monic polynomials as follows.

Definition 3.2 We define the equivalence relation on $\mathbf{M}$ such that

$$m_1 \sim m_2 \Longleftrightarrow m_1 \equiv m_2 \pmod g \text{ and } m_1^\star = m_2^\star \pmod{T^{-l-1}}$$

for $m_1, m_2 \in \mathbf{M}$ and write $m_1 \equiv m_2 \pmod{\mathcal{R}_{g,l}}$ for this case. Equivalently, $m_1 \equiv m_2 \pmod{\mathcal{R}_{g,l}}$ if $m_1 \equiv m_2 \pmod g$ and the first $l+1$ coefficients of m_1 and m_2 coincide, including the leading coefficient 1.

Now we define the congruence characters, which extend the Dirichlet characters (mod g).

Definition 3.3 We say m is *invertible* $(\text{mod}\mathcal{R}_{g,l})$ if there is m^*such that $mm^* \equiv 1 \pmod{\mathcal{R}_{g,l}}$. Then we say $\xi : \mathbf{M} \to \mathbf{C}$ is a *congruence character* (mod $\mathcal{R}_{g,l}$) if ξ satisfies the following.

- $\xi(f_1 f_2) = \xi(f_1)\xi(f_2)$ for any $f_1, f_2 \in \mathbf{F}_q[T]$,
- $\xi(f_1) = \xi(f_2)$ if $f_1 \equiv f_2 \pmod{\mathcal{R}_{g,l}}$,
- $\xi(1) = 1$,

and with the convention that $\xi(f) = 0$ when f is not invertible (mod $\mathcal{R}_{g,l}$).

The invertible congruence classes form a multiplicative group and we have

$$(\mathbf{F}_q[T]/\mathcal{R}_{g,l})^\times \simeq (\mathbf{F}_q[T]/\langle g\rangle \times \mathbf{F}_q[T^{-1}]/\langle T^{-l-1}\rangle)^\times$$

with an isomorphism $f \mapsto (f, f^\star)$. Thus any congruence character can be written as $\xi = \chi\theta$ for a Dirichlet character χ (mod g) and a character θ (mod $\mathcal{R}_{1,l}$). For the latter, we simply say θ is a congruence character (mod $1/T^{l+1}$). Similar to Dirichlet characters, we say a congruence character θ (mod $1/T^{l+1}$) is primitive if it is not induced from a congruence character $\theta^\star$ (mod $1/T^{l_1+1}$) for some $l_1 < l$; nevertheless, θ and $\theta^\star$ define the same function on monic polynomials, and thus the character (mod $1/T^{l+1}$) being primitive only means that the conductor is chosen to be minimal. Likewise, we say ξ (mod $\mathcal{R}_{g,l}$) is primitive if $\xi = \chi\theta$ for some primitive characters χ (mod g) and θ (mod $1/T^{l+1}$). As a convention, we write $\theta = 1$ when θ is the trivial congruence character (mod $1/T$), *i.e.*, $\theta(m) = 1$ for all monic polynomial m.

Now we state an analogue of the Generalized Riemann Hypothesis in $\mathbf{F}_q[T]$. Hayes [11] explicitly constructed the maximal abelian extension of $\mathbf{K} = \mathbf{F}_q(T)$ using the torsion point of a Calitz module. We may find a Galois extension $\mathbf{K}_{g,l}$ of $\mathbf{K}$ whose Galois group is isomorphic to the group of invertible congruence classes mod $\mathcal{R}_{g,l}$ via an Artin map. Then we apply the work of Weil on algebraic curves over finite fields to bound the character sums. (See Hsu [12], for an explicit bound.)

https://doi.org/10.4153/CJM-2017-023-x Published online by Cambridge University Press

Theorem 3.4 *Suppose ξ is a nontrivial primitive character (mod $\mathcal{R}_{g,l}$). Then we have*

$$\pi(N,\xi) = \sum_{\deg \varpi = N} \xi(\varpi) \le (l + \deg g + 3)\frac{\widehat{N}^{1/2}}{\mathcal{L}(\widehat{N})}.$$

Proof Apart from the explicit constant 3 in the sum, this estimate can be deduced from the Weil bound [21, Theorems 9.16 A–B]. This explicit form of the upper bound is a direct consequence of Hsu [12, Corollary 2.5]. ■

Next we discuss the analogue of the exponential sum and a unit interval. Let $\mathbf{K}_\infty$ be the ring of formal power series

$$\mathbf{K}_\infty = \mathbf{F}_q((1/T)) = \Big\{ x = \sum_{-\infty < i \le n} x_i T^i,\ x_i \in \mathbf{F}_q \Big\}$$

equipped with the norm

$$|x| = q^{-\operatorname{ord} x}, \tag{3.2}$$

where $\operatorname{ord} x$ is the smallest (possibly negative) integer k for which the T^{-k} coefficient is nonzero. Then the definition of the norm (3.2) is consistent with the norm defined earlier in (1.3) when $x \in \mathbf{F}_q[T]$. We take $\mathbf{T} = \{x \in \mathbf{K}_\infty : |x| < 1\}$, and fix an additive Haar measure on $\mathbf{K}_\infty$ normalized so that $\mathbf{T}$ has measure 1.

Let p be the characteristic of $\mathbf{F}_q$. We define a nontrivial additive character on $\mathbf{K}_\infty$ by

$$e(x) = \exp\Big(\frac{2\pi i}{p} \operatorname{tr}_{\mathbf{F}_q/\mathbf{F}_p} x_{-1}\Big)$$

for $x \in \mathbf{K}_\infty$, where x_{-1} denotes the coefficient of T^{-1} in the Laurent series expansion of x.

Lemma 3.5 *For $\lambda \in \mathbf{K}_\infty$, we have*

$$\int_{\mathbf{T}} e(\lambda x)\, dx = \begin{cases} 1 & |\lambda| < 1, \\ 0 & |\lambda| \ge 1. \end{cases}$$

In particular, when $m \in \mathbf{F}_q[T]$ is a monic polynomial, the integral $\int_{\mathbf{T}} e(mx)$ is 1 *if and only if $m = 0$, and is* 0 *otherwise.*

Proof See [10, Corollary 2.5] or [14, Lemma 1]. ■

Later, we want to decompose $e(\lambda x)$ (as a function of x) into characters, when λ is near a/g for small denominator g. Let χ be a Dirichlet character (mod g) where g is chosen to be monic. Then we define the Gauss sum of character χ by

$$\tau(\chi) = \sum_{|b|<|g|} \chi(b) e(b/g),$$

where b runs over all polynomials of degree less than g (including non-monic ones.) We introduce the Euler totient function and the Möbius function for $\mathbf{F}_q[T]$ by

$$\varphi(m) = |m| \prod_{\varpi | m} \Big(1 - \frac{1}{|\varpi|}\Big)$$

and

$$\mu(m) = \begin{cases} (-1)^r & \text{if } m \text{ is squarefree and has } r \text{ prime factors,} \\ 0 & \text{otherwise.} \end{cases}$$

Then, using the orthogonality of characters, for $(b,g) = 1$,

$$e(b/g) = \frac{1}{\varphi(g)} \sum_{\chi \pmod g} \tau(\overline{\chi})\chi(b).$$

The size estimate for the Gauss sum remains to hold.

Lemma 3.6 (i) *If χ is a character (mod g) and is induced by primitive character $\chi^\star$ (mod $g^\star$), we have $\tau(\chi) = \mu(g/g^\star)\chi(g/g^\star)\tau(\chi^\star)$.*

(ii) *Moreover, if χ is primitive, $|\tau(\chi)| = \sqrt{|g|}$ and thus for any χ (primitive or imprimitive), then $|\tau(\chi)| \leq \sqrt{|g|}$.*

Proof The proof is similar to the number theoretic case [5, Lemma 3.7]. ■

Lagarias and Soundararajan introduced a joint Mellin–Fourier transform of a smooth weight function, a variant of the Gauss sum, to control the exponential sums on the major arcs. In our case, however, the cutoff function for a fixed degree is already a Schwartz–Bruhat (locally constant) function, and the analogue of a Mellin transform is a Fourier transform by a continuous character of $\mathbf{K}_\infty^\times$.

We [5] gave a few properties of Fourier analysis on $\mathbf{K}_\infty^\times$ for which one may find a parallel result in Tate's thesis; although Tate did not state the function field case explicitly, the same method can be applied. The role of quasi-characters $n \mapsto n^{it}$ ($t \in \mathbf{R}$) is played by the continuous characters on a locally compact group $U_1 = 1 + \mathbf{T} \subset \mathbf{K}_\infty^\times$. We use the following (unconventional) notation.

Definition 3.7 We say $\theta\colon U_1 \to \{z \in \mathbf{C} : |z| = 1\}$ is *a character at pole* if θ is totally multiplicative and $\theta(U_{r+1}) = 1$ for some integer r. We say *the conductor of* θ by $1/T^{r+1}$ if the choice of r is minimal.

When θ is a character at pole with the conductor $1/T^{l+1}$, θ induces a (primitive) congruence character θ^* (mod $1/T^{l+1}$) by $\theta^*(m) = \theta(m^\star)$ for any monic polynomial m where $m^\star$ is the dual of m defined in (3.1). One may easily check that this (non-canonical) correspondence is one-to-one. Now we are ready to define the Fourier–Mellin transform of the cutoff function on U_1.

Definition 3.8 Let Φ be a characteristic function of U_1. For $\lambda \in \mathbf{K}_\infty$ and a character at pole θ, we define $\check{\Phi}(\theta,\lambda)$ by $\check{\Phi}(\theta,\lambda) = \int_{U_1} e(\lambda x)\overline{\theta(x)}\,dx$, where dx is the (additive) Haar measure of $\mathbf{K}_\infty$.

Then we have the following lemma.

Lemma 3.9 *Let θ be a character at pole and $\lambda \in \mathbf{K}_\infty$. The following statements are true.*

https://doi.org/10.4153/CJM-2017-023-x Published online by Cambridge University Press

(i) *For $\theta = 1$, we have $\check{\Phi}(1,\lambda) = e(\lambda)\cdot 1_{|\lambda|<1}$, where $1_{|\lambda|<1}$ is the characteristic function of $|\lambda| < 1$.*

(ii) *For $\theta \neq 1$, let l be such that the conductor of θ is $1/T^{l+2}$. Then $\check{\Phi}(\theta,\lambda) = 0$ if $|\lambda| \neq q^{l}$.*

(iii) *For $\theta \neq 1$ and if $|\lambda| = q^{l}$, there are exactly q^{l} characters at pole (mod $1/T^{l+2}$) that satisfy $|\check{\Phi}(\theta,\lambda)| = q^{-l/2}$, and $\check{\Phi}(\theta,\lambda) = 0$ for the remaining characters. Indeed, the character θ satisfies $\check{\Phi}(\theta,\lambda) \neq 0$ if and only if $\theta \in \theta^{\star}\widehat{U_1/U_{l+1}}$ for some primitive character $\theta^{\star} \in \widehat{U_1/U_{l+2}}$.*

Proof (i) This assertion is immediate from Lemma 3.5.

(ii) (See [5, Lemma 3.14] and [20, Lemma 7.4].) Let $|\lambda| = q^{r}$. Fix a set of representatives in U_1/U_{r+2} and write $x = au$ with $aU_{r+2} \in U_1/U_{r+2}$ and $r \in U_{r+2}$. If $r < l$, we have $e(\lambda a r) = e(\lambda a)$ because $|\lambda a r - \lambda a| \leq q^{-2}$. Thus

$$\int_{U_1} e(\lambda x)\overline{\theta(x)} = \sum_{a\in U_1/U_{r+2}} \overline{\theta(a)}\, e(\lambda a) \int_{U_{r+1}} \theta(u)\, du.$$

However, $\theta|_{U_{r+1}}$ is nontrivial and the integral is 0 by the orthogonality of characters. If $r > l$, we write $x = a(1 + T^{-l-1}u)$ with $a \in U_1/U_{l+2}$, $u \in \mathbf{T}$. Then $\theta(1 + T^{-l-1}u) = 1$, and using the change of variable, $dx = du/q^{l+1}$. Thus

$$\sum_{a\in U_1/U_{l+2}} \overline{\theta(a)}\, e(\lambda a)\frac{1}{q^{l+1}} \int_{\mathbf{T}} e(\lambda a T^{-l-1}u) du.$$

By Lemma 3.5, the integral is 0.

(iii) If $r = l$, we have

$$\begin{aligned} |\check{\Phi}(\theta,\lambda)|^{2} &= \iint_{U_1\times U_1} e(\lambda(x-y))\theta(yx^{-1})\, dx dy \\ &= \int_{U_1} \theta(z) \int_{U_1} e(\lambda(1-z)x)\, dx dz. \end{aligned}$$

Then the inner integral is $e(\lambda(1-z))$ if $|\lambda(1-z)| < 1$, and 0 otherwise. Therefore we have

$$\begin{aligned} |\check{\Phi}(\theta,\lambda)|^{2} &= \int_{U_{r+1}} \theta(z) e(\lambda(1-z))\, dz \\ &= \frac{1}{q^{l+1}} \sum_{a\in\mathbf{F}_q} \theta\Big(1 + \frac{a}{T^{l+1}}\Big) e\Big(a \cdot \frac{\lambda}{T^{l+1}}\Big). \end{aligned}$$

For given θ and λ, we write $\psi_\theta(a) = \theta(1 + a/T^{l+1})$ and $\chi_\lambda(a) = e(a\lambda T^{-l-1})$. Viewed as a function of a, both ψ_θ and χ_λ are additive characters on $\mathbf{F}_q$; thus the sum is either 0 or q, and the size of $|\check{\Phi}(\theta,\lambda)|$, if not zero, is $q^{-l/2}$.

Now we define a surjective group homomorphism $\mathcal{T}$ from $\widehat{U_1/U_{l+2}}$ to the group of additive characters on $\mathbf{F}_q$ by $\mathcal{T}(\theta) = \psi_\theta$ whose kernel is $\widehat{U_1/U_{l+1}}$. Therefore, $\mathcal{T}^{-1}(\overline{\chi_\lambda})$, which is the set of characters θ for which $\check{\Phi}(\theta,\lambda)$ is nonzero, is a coset written as $\theta^{\star}\widehat{U_1/U_{l+1}}$ for some character $\theta^{\star}$ (mod $1/T^{l+2}$). Because χ_λ is nontrivial ($\lambda_l \neq 0$), we also have $\theta^{\star} \notin \widehat{U_1/U_{l+1}}$ and thus $\theta^{\star}$ is primitive (mod $1/T^{l+2}$). Since $|\widehat{U_1/U_{l+1}}| = q^{l}$, the number of characters θ for which $|\check{\Phi}(\theta,\lambda)|$ is nonzero, is exactly q^{l}. ∎

https://doi.org/10.4153/CJM-2017-023-x Published online by Cambridge University Press

Combining all the above results, we can deduce the multiplicative character decomposition of $e(\lambda x)$.

Corollary 3.10 *Let $\lambda \in \mathbf{K}_\infty$. Then for any monic polynomial m,*

$$e(\lambda m) = \sum_\theta \check{\Phi}(\theta, \lambda T^{\deg m})\theta^*(m),$$

where the sum is taken over all characters at pole.

Proof We write

$$e(\lambda m) = e\Big(\lambda T^{\deg m} \cdot \frac{m}{T^{\deg m}}\Big).$$

Then $m/T^{\deg m} \in U_1$. Using the Fourier inversion formula,

$$e(\lambda m) = \sum_\theta \check{\Phi}(\theta, \lambda T^{\deg m})\theta\Big(\frac{m}{T^{\deg m}}\Big).$$

From Lemma 3.9 $\check{\Phi}(\theta, \lambda T^{\deg m}) = 0$ for all except finitely many θ, and thus the sum is defined. The result is immediate from the definition of θ^*. ∎

4 Exponential Sums

Let X and Y be two large integers and suppose Y divides X. Recall that we denote by $\mathcal{S}_1$ the set of squarefree monic polynomials of degree X whose prime factors are of degree Y. We want to count the number of solutions to $a + b = 2c$ where $a, b, c \in \mathcal{S}_1$. For $\alpha \in \mathbf{T}$ and a set of polynomials $\mathcal{S}$, we write

$$E(\alpha) = E(\alpha; \mathcal{S}) = \sum_{m \in \mathcal{S}} e(m\alpha). \tag{4.1}$$

Then the number of solutions to the equation $a + b = 2c$ with $a, b, c \in \mathcal{S}$ is

$$N(\mathcal{S}) = \int_{\mathbf{T}} E(\alpha; \mathcal{S})^2 E(-2\alpha; \mathcal{S})\, d\alpha. \tag{4.2}$$

Theorem 2.1 can be proved by estimating $N(\mathcal{S}_1)$. To prove Theorem 2.2 however, we need to use the inclusion-exclusion principle on $\mathcal{S}_1$ as our final step, and thus we need to estimate $N(\mathcal{S})$ when $\mathcal{S}$ varies over a slightly more general type of set, that is, the set $\mathcal{S}_1$ sifted by a small number of primes. We write for squarefree monic r,

$$\mathcal{S}_r = \big\{ m \in \mathcal{S}_1 : (m, r) = 1 \big\}. \tag{4.3}$$

We shall treat all $\mathcal{S}$ simultaneously in the next few sections.

4.1 The Hardy–Littlewood Method: Major and Minor Arcs

The main idea of the Hardy–Littlewood method, or the circle method, is to estimate an exponential sum (over integers) near a rational number with a small denominator. A straightforward analogy on $\mathbf{F}_q[T]$ are the rational functions with the small degree denominator. Some successful applications of this method can be found in [10,14,17].

We start with the following lemma on Diophantine approximation.

Lemma 4.1 *Let $\alpha \in \mathbf{T}$ and Q be a fixed positive integer. Then there exist $a, g \in \mathbf{F}_q[T]$ such that* $\deg a, \deg g \le Q$ *and* $\left|\alpha - \frac{a}{g}\right| < \frac{1}{|g|\widehat{Q}}$.

Proof See [5, Lemma 3.5] or [10, Theorem 4.3]. ■

For $a, g \in \mathbf{F}_q[T]$, we write

$$\mathcal{F}\Big(\frac{a}{g}, \widehat{R}\Big) = \Big\{x \in \mathbf{T} : \Big|x - \frac{a}{g}\Big| < \frac{1}{|g|\widehat{R}}\Big\}.$$

Then the collection of arcs $\mathcal{F}(a/g, \widehat{R})$ for $(a, g) = 1$, $\deg g \le R$ are disjoint, because, when $|\alpha - a_1/g_1| < 1/|g_1|\widehat{R}$ and $|\alpha - a_2/g_2| < 1/|g_2|\widehat{R}$,

$$\frac{1}{|g_1 g_2|} \le \Big|\frac{a_1}{g_1} - \frac{a_2}{g_2}\Big| \le \max\Big(\Big|\alpha - \frac{a_1}{g_1}\Big|, \Big|\alpha - \frac{a_2}{g_2}\Big|\Big)$$

by the strong triangular inequality, which fails when $\deg g_1, \deg g_2 \le R$. Therefore, we dissect the range $\mathbf{T}$ by the disjoint unions of $\mathcal{F}(a/g, \widehat{R})$.

Now we define the set of major arcs

$$\mathfrak{M} = \bigcup_{\substack{a,g \\ |g| < \widehat{X}^{1/2}}} \mathfrak{M}\Big(\frac{a}{g}\Big),$$

where $\mathfrak{M}\big(\frac{a}{g}\big) = \big\{x \in \mathbf{T} : \big|x - \frac{a}{g}\big| < \frac{1}{\widehat{X}}\big\}$, and we define the set of minor arcs by $\mathbf{T} - \mathfrak{M}$. Thus $\alpha \in \mathfrak{m}$ if and only if $\alpha \in \mathcal{F}(a/g, \widehat{X}^{1/2})$ and $|\alpha - a/g| \ge 1/\widehat{X}$ for some a, g with $|g| \le \widehat{X}^{1/2}$.

4.2 Character Decomposition

In this section, $\mathcal{S}$ will be always $\mathcal{S}_1$ or $\mathcal{S}_r$ for some r where $\mathcal{S}_r$ is defined in (4.3). The goal of this section is to decompose $E(\alpha; \mathcal{S})$ into character sums. First of all, we decompose the additive character into the Dirichlet characters. Recall that for $(b, g) = 1$,

$$e\Big(\frac{b}{g}\Big) = \frac{1}{\varphi(g)} \sum_{\chi \pmod{g}} \tau(\overline{\chi})\chi(b).$$

Then combined with 3.10, we have

$$e\Big(\Big(\frac{a}{g} + \gamma\Big) m\Big) = \frac{1}{\varphi(g/d)} \sum_{\chi \pmod{g}} \tau(\overline{\chi})\chi(a) \sum_{\theta} \check{\Phi}(\theta, \gamma T^X)\chi(m/d)\theta^*(m),$$

where $d = \gcd(m, g)$. For brevity, we write $\mathcal{S}(d) = \{m' : m'd \in \mathcal{S}\}$. Then for $\alpha = a/g + \gamma$, we have

$$E(\alpha) = \sum_{d|g} \frac{1}{\varphi(g/d)} \sum_{\chi \pmod{g/d}} \tau(\overline{\chi})\chi(a) \sum_{\theta} \check{\Phi}(\theta, \gamma T^X)\theta^*(d) E(\mathcal{S}(d), \chi\theta^*), \tag{4.4}$$

where

$$E(\mathcal{S}(d), \chi\theta^*) = \sum_{m' \in \mathcal{S}(d)} \chi(m')\theta^*(m').$$

https://doi.org/10.4153/CJM-2017-023-x Published online by Cambridge University Press

We remark that if $|\gamma| \geq 1/\widehat{X}$, $\check{\Phi}(\theta, \gamma T^X) = 0$ unless θ has a conductor $1/T^{l+2}$ for some $l \geq 0$. Therefore all those θ^* that contribute to the sum (4.4) are nonprincipal characters if $|\gamma| \geq 1/\widehat{X}$.

4.3 Character Sum Estimate

The next goal is to estimate $E(\mathcal{S}_r(d), \xi)$ when ξ is a congruence character $(\bmod\ \mathcal{R}_{g,l})$. We write $\kappa = \kappa(X, Y) := (\mathcal{L}(\widehat{Y}))/(\mathcal{L}\mathcal{L}(\widehat{X}))$ *i.e.*, $\widehat{Y} = \mathcal{L}(\widehat{X})^\kappa$. This definition is consistent with the previous definition (1.4) as the solution $a + b = 2c$ for a, b, $c \in \mathcal{S}_1$ for given X, Y is $\kappa(X, Y)$.

The advantage over previous papers is the average estimate over a certain set of characters. Let $X(g, l)$ be the set of congruence characters $(\bmod\ \mathcal{R}_{g,l})$. We consider the set of characters Ξ in either of the forms

$$\Xi = \begin{cases} \theta X(g, l) \text{ for some primitive } \theta \pmod{1/T^{l+2}}; & \text{or} \\ X(g, 0) \end{cases} \tag{4.5}$$

which is a typical set of characters that appears in the inner sum (4.4). We write $Q = \mathcal{L}(|\Xi|) = \deg g + l$ for the first case and $Q = \deg g$ for the second case to track the size of $|\Xi|$. The average estimate for character sums for $\xi \in \Xi$ in this section depends on Q but does not depend on the choice of θ, g, and l when the Q-value is the same. Finally, we use a shorthand notation Ξ' for the nonprincipal characters in Ξ, *i.e.*, $\Xi' = \Xi$ for the first case and $\Xi - \{\chi_{0,g}\}$ for the second case, where $\chi_{0,g}$ is the principal character $(\bmod\ g)$.

In this section, we prove the following estimate.

Proposition 4.2 *Let d, r be polynomials of degree $\leq X/2$ and all prime factors of dr are of degree Y. Let Ξ be the set of characters of the form (4.5), and $Q = \deg g + l$ or $\deg g$ when g, l are as in the definition of Ξ. Suppose $\mathcal{L}(\widehat{X})^2 \leq \widehat{Y}$ and $Q \leq X/2$. We have*

$$\frac{1}{|\Xi|} \sum_{\xi \in \Xi'} \Big| \sum_{m \in \mathcal{S}_r(d)} \xi(m) \Big| \leq \widehat{Q}^{-1/\kappa} \widehat{X}^{1/2 + O\left(\mathcal{L}\mathcal{L}(\widehat{Y})/\mathcal{L}(\widehat{Y}) \right)}.$$

The condition on r and d in the proposition appears frequently. Thus we define

$$\mathcal{D} = \left\{ d \in \mathbf{M}_0 : \varpi \mid d \implies |\varpi| \leq \widehat{Y}, |d| \leq \widehat{X}^{1/2} \right\}, \tag{4.6}$$

and for the remainder of this section, we assume r, $d \in \mathcal{D}$. It follows that $\omega(d)$, $\omega(r) \leq X/2Y$, where $\omega(m)$ denotes the number of prime factors of m. Thus when $\widehat{Y} \geq \mathcal{L}(\widehat{X})^2$, we have

$$\omega(dr) = \frac{X}{Y} \leq \frac{\widehat{Y}^{1/2}}{\mathcal{L}(\widehat{Y})}.$$

Therefore, the primes not dividing dr cause little change to the character sums that we write as the following lemma.

https://doi.org/10.4153/CJM-2017-023-x Published online by Cambridge University Press

Lemma 4.3 *Let d be a monic polynomial and $\xi \in \Xi$. Suppose $\omega(d) \le \widehat{Y}^{1/2}/\mathcal{L}(\widehat{Y})$. Then*

$$\Big| \sum_{\substack{\deg \varpi = Y \\ \varpi \nmid d}} \xi(\varpi) \Big| \le (Q+5) \frac{\widehat{Y}^{1/2}}{\mathcal{L}(\widehat{Y})}.$$

Proof Immediate from Theorem 3.4. We increase the constant by 2, because the conductor for ξ is at most $Q+1$. ■

The estimate in Lemma 4.3 is very powerful, and perhaps the best we can hope for, when Q is not too large compared to Y. However, the interesting case in our problem is when $\widehat{Y}$ is comparable to some power of $\mathcal{L}(\widehat{X})$; then if $\widehat{Q}$ is as large as a small power of $\widehat{X}$, we still have a power saving when $\mathcal{L}(\widehat{Q}) \le Y^{1/2-\delta}$ for some $\delta > 0$, but not as powerful as the square-root cancellation. However, the orthogonality of character, or the large sieve in more general settings, always ensures the square-root cancellation when you average the (second) moment of character sums over all characters. Harper observed that when a set is well factorizable, one may incorporate this idea to produce extra saving over characters. The following lemma is the interpolation of this observation combined with the Weil bound estimate.

Lemma 4.4 *Let d be a polynomial with $\omega(d) \le \widehat{Y}^{1/2}/\mathcal{L}(\widehat{Y})$ and let k be an integer with $k \le X/Y$. Suppose $Q \le X/2$ and $\widehat{Y} \ge \mathcal{L}(\widehat{X})^2$. We write $s = \lfloor Q/Y \rfloor$,* i.e., *the largest integer satisfying $\widehat{Y}^s \le \widehat{Q}$. Then we have*

$$\frac{1}{|\Xi|} \sum_{\xi \in \Xi'} \Big| \sum_{\substack{\deg \varpi = Y \\ \varpi \nmid d}} \xi(\varpi) \Big|^k \le \begin{cases} \mathcal{L}(\widehat{X})^{k-s} \widehat{Y}^{k/2} & k \ge 2s, \\ O(\widehat{Y}^{k/2} \mathcal{L}(\widehat{X})^{\lceil k/2 \rceil}) & k < 2s. \end{cases}$$

Proof Suppose first that $k \ge 2s$. Then by Lemma 4.3,

$$\begin{aligned} \frac{1}{|\Xi|} \sum_{\xi \in \Xi'} \Big| \sum_{\varpi} \xi(\varpi) \Big|^k &= \frac{1}{|\Xi|} \sum_{\xi \in \Xi'} \Big| \sum_{\varpi} \xi(\varpi) \Big|^{2s} \Big| \sum_{\varpi} \xi(\varpi) \Big|^{k-2s} \\ &\le \frac{1}{|\Xi|} \sum_{\xi \in \Xi'} \Big| \sum_{m} a(m) \xi(m) \Big|^2 \Big((\mathcal{L}(\widehat{Q}) + 5) \frac{\widehat{Y}^{1/2}}{\mathcal{L}(\widehat{Y})} \Big)^{k-2s}, \end{aligned}$$

where $a(m)$ is the number of ways to write m as $\varpi_1 \cdots \varpi_s$ for primes ϖ_i (not necessarily distinct). Then $a(m) \le s!$ and $a(m)$ is supported on polynomials of degree sY, which is at most Q from the construction. Then if m_1 and m_2 are two monic polynomials of degree at most Q and $m_1 \equiv m_2 \pmod{\mathcal{R}_{g,l}}$, we have $m_1 = m_2$. Now we expand the square and use the orthogonality of characters (mod $\mathcal{R}_{g,l}$) to get

$$\begin{aligned} \frac{1}{|\Xi|} \sum_{\xi \in \Xi'} \Big| \sum_{m} a(m) \xi(m) \Big|^2 &\le \frac{1}{|\Xi|} \sum_{\chi \in \mathcal{R}_{g,l}} \Big| \sum_{m} a(m) \theta(m) \chi(m) \Big|^2 \\ &= \sum_{m} a(m)^2 \le s! \pi(Y)^s, \end{aligned}$$

where θ, g, l are defined in (4.5), with the convention that $\theta = 1$ when $\Xi = X(g, 0)$. Then we apply Lemma 3.1 to have $\pi(Y) \le \widehat{Y}/\mathcal{L}(\widehat{Y})$.

To summarize, we have

$$\frac{1}{|\Xi|}\sum_{\xi\in\Xi'}\Big|\sum_{\varpi}\chi(\varpi)\Big|^{k}\le s!\Big(\frac{\widehat{Y}}{\mathcal{L}(\widehat{Y})}\Big)^{s}\Big((\mathcal{L}(\widehat{Q})+5)\frac{\widehat{Y}^{1/2}}{\mathcal{L}(\widehat{Y})}\Big)^{k-2s}.$$

Then use a crude bound $s!\le s^s\le\mathcal{L}(\widehat{X})^s$, drop the powers of $\mathcal{L}(\widehat{Y})$ in the denominators, and $(\mathcal{L}(\widehat{Q})+5)^{k-2s}\ll\mathcal{L}(\widehat{X})^{k-2s}$ to obtain the desired result.

For $s>k/2$, if k is even, we split the k-th power into two $(k/2)$-th powers and use the orthogonality to have

$$\frac{1}{|\Xi|}\sum_{\xi\in\Xi}\Big|\sum_{\varpi}\xi(\varpi)\Big|^{k}\le(k/2)!\widehat{Y}^{k/2}\le\mathcal{L}(\widehat{X})^{k/2}\widehat{Y}^{k/2},$$

which is the desired result; when k is odd, we apply Lemma 4.3 for the single leftover. ■

If we expand the k-th power in Lemma 4.4, we have the (average) character sums over k-products of irreducibles, not necessarily distinct, of degree m, and each polynomial is counted with multiplicity up to $k!$. We expect that the polynomials that are the products of k distinct polynomials make up the majority, with multiplicity $k!$. To prove this claim rigorously, we need the following combinatorial lemma.

Lemma 4.5 *Let $\{x_i\}_{i=1}^n$ be any sequence of complex numbers. We define*

$$\pi_k=\sum_{1\le i_1<i_2<\cdots<i_k\le n}x_{i_1}\cdots x_{i_k}$$

to be the sum of products of k distinct x_i and $s_k=\sum_i x_i^k$ to be k-th power sum of x_i. Suppose the power sums satisfy $|s_1|\le a$ and $|s_k|\le b$ for $k\ge 2$. For any $k\le n$,

$$k!|\pi_k|\le\sum_r\binom{k}{r}a^r(bk)^{(k-r)/2}.$$

The proof will be given in Section 4.4. Suppose the lemma for the moment, and we shall prove Proposition 4.2 using Lemmas 4.4 and 4.5.

Proof of Proposition 4.2 Let $K=X/Y$ and write $s=\lfloor Q/Y\rfloor$. For each character $\xi\in\Xi$, we shall take our sequence to be $\{\xi(\varpi)\}_{\deg\varpi=Y,\varpi\nmid dr}$. Then $|s_1|=|\sum_{\deg\varpi=Y,\varpi\nmid dr}\xi(\varpi)|$ and $|s_k|\le\pi(Y)$ by the trivial estimate. Thus,

$$\begin{aligned}\Big|\sum_{m\in\mathcal{S}_r(d)}\xi(m)\Big|&\le\frac{1}{K!}\Big|\sum_{(\varpi_i):\text{distinct }K\text{-tuple}}\xi(\varpi_1\varpi_2\cdots\varpi_K)\Big|\\&\le\frac{1}{K!}\sum_{k\le K}\binom{K}{k}\Big|\sum_{\varpi}\xi(\varpi)\Big|^{k}\Big(\pi(Y)K\Big)^{(K-k)/2}.\end{aligned}$$

Now if we sum over all characters $\xi\in\Xi'$, we use Lemma 4.4 to have

$$\begin{aligned}\frac{1}{|\Xi|}\sum_{\xi\in\Xi'}\Big|\sum_{m\in\mathcal{S}_r(d)}\xi(m)\Big|\le\frac{1}{K!}\Big(&\sum_{2s\le k\le K}\binom{K}{k}\widehat{Y}^{k/2}\mathcal{L}(\widehat{X})^{k-s}(\pi(Y)K)^{(K-k)/2}\\&+\sum_{k\le 2s}\binom{K}{k}\widehat{Y}^{k/2}\mathcal{L}(\widehat{X})^{\lceil k/2\rceil}(\pi(Y)K)^{(K-k)/2}\Big).\end{aligned}$$

https://doi.org/10.4153/CJM-2017-023-x Published online by Cambridge University Press

The first sum in the parentheses is

$$\leq 2^K \cdot \widehat{Y}^{K/2} \mathcal{L}(\widehat{X})^{K-s} \sum_{t\geq 0} \Big(\frac{\pi(Y)K}{\widehat{Y}\mathcal{L}(\widehat{X})^2} \Big)^{t/2}$$

by the substitution of $t = K - s$; the second sum is

$$\leq 2^K \cdot \mathcal{L}(\widehat{X})^{1/2} \cdot 2s \cdot \widehat{Y}^{K/2} \mathcal{L}(\widehat{X})^{K/2}$$

using the trivial estimates $\pi(Y) \leq \widehat{Y}$ and $K \leq \mathcal{L}(\widehat{X})$. The leading factor $1/K! \leq e^{O(K)}/K^K = \mathcal{L}(\widehat{X})^{-K}\widehat{X}^{O(\mathcal{L}\mathcal{L}(\widehat{Y})/\mathcal{L}(\widehat{Y}))}$, and 2^K,$\mathcal{L}(\widehat{X})$, and $2s$ are absorbed in the error term. Thus, overall,

$$\frac{1}{|\Xi|} \sum_{\xi\in\Xi'} \Big| \sum_{m\in\mathcal{S}_r(d)} \xi(m) \Big| \leq \mathcal{L}(\widehat{X})^{-K} \widehat{X}^{O\left(\mathcal{L}\mathcal{L}(\widehat{Y})/\mathcal{L}(\widehat{Y})\right)} \left(\widehat{Y}^{K/2}\mathcal{L}(\widehat{X})^{K-s} + \widehat{Y}^{K/2}\mathcal{L}(\widehat{X})^{K/2}\right)$$

and from $Q \leq X/2$, $s \leq K/2$. Thus the second term is less than the first term. By substituting $\widehat{Y}^K = \widehat{X}$, we have

$$\frac{1}{|\Xi|} \sum_{\xi\in\Xi'} \Big| \sum_{m\in\mathcal{S}_r(d)} \xi(m) \Big| \leq \widehat{X}^{1/2+O\left(\mathcal{L}\mathcal{L}(\widehat{Y})/\mathcal{L}(\widehat{Y})\right)} \mathcal{L}(\widehat{X})^{-s},$$

as desired. ∎

4.4 Proof of Lemma 4.5

The proof of Lemma 4.5 is a direct consequence of the Newton–Girard formula in its matrix form. Let n be fixed, x_i be complex number, and let

$$\pi_k = \sum_{1\leq i_1 < i_2 < \cdots < i_k \leq n} x_{i_1} \cdots x_{i_k}$$

and $s_k = \sum_i x_i^k$. Then the Newton–Girard formula states that the symmetric polynomial can be evaluated by the determinant of power sums as follows:

$$k!\pi_k = \begin{vmatrix} s_1 & 1 & 0 & \cdots & 0 \\ s_2 & s_1 & 2 & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ s_{k-1} & s_{k-2} & \cdots & s_1 & k-1 \\ s_k & s_{k-1} & \cdots & s_2 & s_1 \end{vmatrix}.$$

In our application, $|x_i| \leq 1$ for all i and k is smaller than n, but may grow up to some small power of n. The main issue here is to find an upper bound of the determinant uniformly in k. We start with the following observation.

Lemma 4.6 *Suppose f is a permutation of $\{1, \dots, n\}$ satisfying $f(i) \leq i + 1$ for all i. Then f is a product of a disjoint cycle of consecutive numbers (possibly of length 1).*

Proof We proceed by induction on n. When $n = 1$, the result is trivial. Let k be the least $k > 0$ such that $f^k(1) = 1$. Then $\{1, \dots, f^{i-1}(1)\} \subseteq \{1, \dots, i\}$ for all $i \leq k$ and since all $f^j(1)$ are distinct for $j < k$, we have $f(i) = i + 1$ for $i \leq k - 1$ and $f(k) = f^k(1) = 1$. Now we apply the induction hypothesis for the remaining $n - k$ terms. Take g_k to be the substitution of the first k terms and the remaining terms, *i.e.*,

https://doi.org/10.4153/CJM-2017-023-x Published online by Cambridge University Press

$g_k(i) = i + (n-k)$ for $i \le k$ and $g_k(i) = i - k$ for $i > k$. Then applying the induction hypothesis on $g_k f g_k^{-1}|_{\{1,\dots,n-k\}}$ proves the lemma. ∎

Then Lemma 4.5 follows immediately from the next lemma.

Lemma 4.7 *Suppose $A = (x_{ij})$ is the n by n matrix satisfying $x_{ij} = 0$ for $j > i+1$. Let $|x_{ii}| \le a$, and suppose $|x_{ij}| \le b$ for $i < j$, $|x_{ij}| \le c$ for $i = j+1$ for some $b, c \ge 2$. Then if $b \ge 3c$, $\det A \ll \sum_{0\le s\le n} \binom{n}{s} a^s (bc)^{\frac{n-s}{2}}$, where the implied constants are absolute.*

Proof Since $x_{ij} = 0$ for $j > i+1$,

$$\det A = \sum_{\substack{\sigma\in S_n\\ \sigma(i)\le i+1}} \operatorname{sgn}(\sigma) x_{1\sigma(1)} \cdots x_{n\sigma(n)}.$$

Let $S_{r,s}$ be the set of permutation σ such that $\sigma(i) \le i+1$. The cycle decomposition of σ consists of r cycles, including trivial ones, and $\sigma(i) = i$ has s solutions. By Lemma 4.6, each cycle consists of consecutive numbers, and

$$|x_{i,i+1} \cdots x_{i+j-1,i+j}\, x_{i+j,i}| \le bc^j,$$

i.e., each nontrivial cycle of length j contributes bc^{j-1}. Therefore. if $\sigma \in S_{r,s}$,

$$|x_{1\sigma(1)} \cdots x_{k\sigma(k)}| \le a^s b^{r-s} c^{k-r}.$$

By Lemma 4.6, $|S_{ij}|$ is determined by the length of each cycle, say $a_1, a_2, \dots, a_r$, where exactly s of these variables are equal to 1. Therefore $|S_{r,s}|$ is equal to the number of positive integer solutions to the equation $a_1 + \cdots + a_r = k$ when exactly s of these are equal to 1. By simple combinatorics, when $s \le r$ and $2(r-s) + s \le k$, we have

$$\begin{aligned} |S_{r,s}| &= \binom{r}{s} \cdot \binom{(r-s) + (k-s-2(r-s)) - 1}{(k-s-2(r-s))} = \binom{r}{s}\binom{k-r-1}{k-2r+s} \\ &\le \binom{k}{s} \cdot 2^{k-r} \end{aligned}$$

Fixing s and summing over r, we have

$$\begin{aligned} \sum_r |S_{r,s}| a^s b^{r-s} c^{k-r} &\le \binom{k}{s} a^s b^{-s} (2c)^k \sum_{s\le r\le (k+s)/2} b^r (2c)^{-r} \\ &\ll \binom{k}{s} a^s b^{[\frac{k+s}{2}]-s} (2c)^{k-[\frac{k+s}{2}]} = \binom{k}{s} a^s (bc)^{\frac{k-s}{2}} \end{aligned}$$

if $b \ge 3c$. Therefore, summing over s, $|\det A| \ll \sum_{s=0}^k \binom{k}{s} a^s (bc)^{(k-s)/2}$, as desired. ∎

5 Estimate for $E(\alpha)$

Let $\mathcal{S}$ be $\mathcal{S}_1$ or $\mathcal{S}_r$ for some $r \in \mathcal{D}$, where $\mathcal{D}$ is defined in (4.6), and $E(\alpha) = E(\alpha; \mathcal{S})$ denotes the exponential sum defined in (4.1). In this section, we combine the character decomposition of $E(\alpha)$ in (4.4) and the estimate in Proposition 4.2 to produce an estimate for $E(\alpha)$ near rational polynomials.

Proposition 5.1 *Let $\alpha = a/g + \gamma$, where $|g| \le \widehat{X}^{1/2}$, $|\gamma| \le 1/(|g|\widehat{X}^{1/2})$, and $\epsilon > 0$. Then when Y is sufficiently large and $Y \ge \mathcal{L}(\widehat{X})^2$, $E(\alpha) = M(\alpha) + O(\widehat{X}^{3/4-1/(2\kappa)+\epsilon})$, where $M(\alpha)$ is nonzero only when $|\gamma| < 1/\widehat{X}$, and is equal to*

$$M(\alpha) = e(\gamma T^X)\sum_{d|g}\frac{\mu(g/d)}{\varphi(g/d)}|\mathcal{S}(d)| \quad \Big(|\gamma| < \frac{1}{\widehat{X}}\Big) \tag{5.1}$$

and $\kappa = \kappa(X,Y) = \mathcal{L}(\widehat{Y})/\mathcal{L}(\mathcal{L}(\widehat{X}))$.

The main improvement of this paper is the extra saving on the exponent by $1/(2\kappa)$; for instance, when $\kappa = 6+o(1)$ and for large Y (and thus for large X), the error estimate is $O(\widehat{X}^{2/3+o(1)})$. We separate the principal character terms in (4.4), let $M(\alpha)$ be those terms involving principal characters ($\chi = \chi_{0,g/d}$ for $d \mid g$ and $\theta = 1$), and let $R(\alpha)$ be the nonprincipal parts. Then $E(\alpha) = M(\alpha) + R(\alpha)$, where

$$M(\alpha) = \sum_{d|g}\frac{\tau(\chi_{0,g/d})}{\varphi(g/d)}\check{\Phi}(1,\gamma T^X)|\mathcal{S}(d)|$$

and $\chi_{0,g/d}$ is the principal character (mod g/d).

Proof of Proposition 5.1 From Lemma 3.6 we have $\tau(\chi_{0,g/d}) = \mu(g/d)$ and from Lemma 3.9, $\check{\Phi}(1,\gamma T^X) = e(\gamma T^X)$ if $|\gamma| < 1/\widehat{X}$, and 0 otherwise. Therefore, if $\alpha \in \mathfrak{M}$,

$$M(\alpha) = e(\gamma T^X)\sum_{d|g}\frac{\mu(g/d)}{\varphi(g/d)}|\mathcal{S}(d)|$$

as desired. The remainder $R(\alpha)$ can be considered in two cases. If $\alpha \in \mathfrak{M}$, then $\check{\Phi}(\theta,\gamma T^X) = 0$ unless θ is trivial, and $\tau(\chi) \le \sqrt{|g/d|}$ from Lemma 3.6. Thus we have

$$|R| \le \sum_{d|g}\frac{\sqrt{|g/d|}}{\varphi(g/d)}\sum_{\substack{\chi \pmod{g/d}\\ \chi\ne\chi_{0,g/d}}}\Big|\sum_{m\in\mathcal{S}(d)}\chi(m)\Big|.$$

If $\alpha \in \mathfrak{m}$, we write $|\gamma T^X| = q^l$ for some $l \ge 0$. From Lemma 3.9, there is $\theta_\gamma \in \widehat{U_1/U_{l+2}}$ such that $\check{\Phi}(\theta,\gamma T^X)$ is nonzero if and only if $\theta \in \theta_\gamma\widehat{U_1/U_{l+1}}$, and $\check{\Phi}(\theta,\gamma T^X) = |\gamma T^X|^{-1/2}$ when $\alpha \in \mathfrak{m}$ and $\theta \in \theta_\gamma\widehat{U_1/U_{l+1}}$. In such case, the products of characters $\chi\theta^*$ in the sum (4.4) run over $\theta_\gamma^* X(g/d,l)$ where θ_γ^* is a congruence character (mod $1/T^{l+2}$) induced by θ_γ. Thus we have

$$\begin{aligned}
|R| &\le \sum_{d|g}\frac{1}{\varphi(g/d)}\sum_{\substack{\chi \pmod{g/d}\\ \theta\in\theta_\gamma\widehat{U_1/U_{l+1}}}}|\tau(\overline{\chi})|\,|\check{\Phi}(\theta,\gamma T^X)|\Big|\sum_{m\in\mathcal{S}(d)}\chi(m)\theta^*(m)\Big|\\
&\le \frac{1}{\sqrt{|\gamma T^X|}}\sum_{d|g}\frac{\sqrt{|g/d|}}{\varphi(g/d)}\sum_{\xi\in\Xi}\Big|\sum_{m\in\mathcal{S}(d)}\xi(m)\Big|\\
&= |\gamma T^X|^{1/2}|g|^{1/2}\sum_{d|g}\frac{1}{|d|^{1/2}|\Xi|}\sum_{\xi\in\Xi}\Big|\sum_{m\in\mathcal{S}(d)}\xi(m)\Big|,
\end{aligned}$$

where $\Xi = \theta_\gamma^* X(g/d,l)$ (and we used $|\Xi| = \varphi(g/d)|\gamma T^X|$).

https://doi.org/10.4153/CJM-2017-023-x Published online by Cambridge University Press

We may apply Proposition 4.2 in either case, and when Y is large,

$$|R| \ll \max(1, |\gamma T^X|)^{1/2} |g|^{1/2} \sum_{d|g} |d|^{-1/2} \widehat{X}^{1/2+O(\mathcal{L}\mathcal{L}(\widehat{Y})/\mathcal{L}(\widehat{Y}))} \Big|\frac{g}{d}\Big|^{-1/\kappa} (1 + |\gamma T^X|)^{-1/\kappa}$$
$$\leq \max(1, |\gamma T^X|)^{1/2-1/\kappa} |g|^{1/2-1/\kappa} \Big(\sum_{d|g} |d|^{1/\kappa - 1/2}\Big) \widehat{X}^{1/2+\epsilon/2},$$

where $\kappa = \kappa(X, Y) = \mathcal{L}(\widehat{Y})/\mathcal{L}(\mathcal{L}(\widehat{X}))$. The inner sum on $d|g$ is bounded by $|g|^{\epsilon}$ and since $|g| \max(1, |\gamma T^X|) \leq \widehat{Q}$ and $\widehat{Q} \leq \widehat{X}^{1/2}$, $R \ll \widehat{X}^{3/4-1/(2\kappa)+\epsilon}$, which finishes the proof. ■

6 Proof of Theorem 2.1

Now we prove the main theorem. Indeed, we prove a stronger theorem, and Theorem 2.1 follows directly from the following proposition.

Proposition 6.1 *Let Y be a sufficiently large integer, $\kappa > 6$, and let X be an integer that is a multiple of Y and satisfying $\mathcal{L}(\widehat{X})^{\kappa} \leq \widehat{Y}$. Let r be a squarefree polynomial of degree at most $X/2$ whose prime factors are of degree Y. Then the number of solutions to $a + b = 2c$ with a, b, $c \in \mathcal{S}_r$ is*

$$N(\mathcal{S}_r) = \mathfrak{S} \frac{|\mathcal{S}_r|^3}{\widehat{X}} \Big(1 + O\Big(\frac{1}{\widehat{Y}^{1/2}}\Big)\Big),$$

where the singular series $\mathfrak{S}$ is defined by

$$\mathfrak{S} = \prod_{\varpi} \Big(1 - \frac{1}{(|\varpi| - 1)^2}\Big)$$

and the product is taken over all monic irreducibles.

We first need a rough estimate on $|\mathcal{S}_r(d)|$.

Lemma 6.2 *Let d, r be monic squarefree polynomials of degree at most $X/2$ and all prime factors are of degree Y. Suppose $\widehat{Y} \geq \mathcal{L}(X)^2$ and let $\kappa = \kappa(X, Y)$ be such that $\widehat{Y} = \mathcal{L}(\widehat{X})^{\kappa}$. Then $|\mathcal{S}_r(d)| \ll |\mathcal{S}_r| |d|^{-1+1/\kappa}$.*

Proof Let $K = X/Y$. If $(r, d) > 1$, then $\mathcal{S}_r(d)$ is empty and the result is clear. Suppose $(r, d) = 1$. Then

$$|\mathcal{S}_r(d)| = \binom{\pi(Y) - \omega(r) - \omega(d)}{K - \omega(d)} = |\mathcal{S}_r| \cdot \binom{\pi(Y) - \omega(r) - \omega(d)}{K - \omega(d)} \Big/ \binom{\pi(Y) - \omega(r)}{K}$$
$$= |\mathcal{S}_r| \prod_{0 \leq j < \omega(d)} \Big(\frac{K - j}{\pi(Y) - \omega(r) - j}\Big).$$

Since $|d| \leq \widehat{X}^{1/2}$, we have $\omega(d) \leq X/2Y \leq K/2$ and $\omega(r) \leq K/2$. Since $K^2 = O(\pi(Y))$ and $\widehat{Y} \geq \mathcal{L}(\widehat{X})^2$, the denominator is

$$\prod_{0 \leq j < \omega(d)} (\pi(Y) - \omega(r) - j) = \pi(Y)^{\omega(d)} e^{O(\omega(d)\omega(rd)/\pi(Y))} \gg \pi(Y)^{\omega(d)}.$$

The numerator is less than $K^{\omega(d)}$, and thus $|\mathcal{S}_r(d)| \ll (\frac{K}{\pi(Y)})^{\omega(d)}|\mathcal{S}_r|$. We have

$$\frac{K}{\pi(Y)} = \frac{\mathcal{L}(\widehat{X})}{\mathcal{L}(\widehat{Y})\pi(Y)} = \widehat{Y}^{-1+1/\kappa}\big(1 + O(\widehat{Y}^{-1/2})\big)$$

by 3.1. Thus by taking $\omega(d)$-th power,

$$\Big(\frac{K}{\pi(Y)}\Big)^{\omega(d)} = |d|^{-1+1/\kappa} e^{O(\omega(d)/\widehat{Y}^{1/2})} \ll |d|^{-1+1/\kappa}$$

because $\omega(d) \le K/2 \le \widehat{Y}^{1/2}$ when $\kappa \ge 2$. Therefore, $|\mathcal{S}_r(d)| \ll |\mathcal{S}_r|\,|d|^{-1+1/\kappa}$ as desired. ■

Throughout this section, let $\epsilon > 0$ and suppose Y is sufficiently large according to ϵ, and $\mathcal{L}(\widehat{X})^2 \le \widehat{Y}$. The minor arc contribution of the integral (4.2) is

$$\int_{\mathfrak{m}} |E(\alpha)|^2|E(-2\alpha)|\,d\alpha \le \widehat{X}^{3/4-1/(2\kappa)+\epsilon} \int_{\mathbf{T}} |E(\alpha)|^2\,d\alpha \ll |\mathcal{S}|\widehat{X}^{3/4-1/(2\kappa)+\epsilon}.$$

Now we handle the main term M and find the contribution of

$$\int_{\mathfrak{M}} M(\alpha)^3 M(-2\alpha)\,d\alpha.$$

When $\alpha \in M(a/g)$, let

$$M_0(g) = M(a/g) = M(\alpha)e(-\gamma T^X) = \sum_{d|g} \frac{\mu(g/d)}{\varphi(g/d)}|\mathcal{S}(d)|,$$

which depends only on g, but not on a and γ. We have

$$\int_{\mathfrak{M}} M(\alpha)^2 M(-2\alpha)\,d\alpha = \sum_{|g|<\widehat{Q}} \sum_{(a,g)=1} \int_{\mathfrak{M}(a/g)} M_0(g)^3\,d\alpha = \frac{1}{\widehat{X}} \sum_{|g|<\widehat{Q}} \varphi(g)M_0(g)^3. \tag{6.1}$$

Thus the major arc contribution is understood through $M_0(g)$.

For simplicity, we write $P_Y = \prod_{\deg \varpi = Y} \varpi$. We observe that $|\mathcal{S}(d)|$ is nonzero only when d divides P_Y. Thus if we write $g = g_1g_2$ so that $(g_1, P_Y) = 1$ and all prime factors of g_2 divide P_Y, we have

$$M_0(g) = \frac{\mu(g_1)}{\varphi(g_1)} \sum_{d|g_2} \frac{\mu(g_2/d)}{\varphi(g_2/d)}|\mathcal{S}(d)|.$$

If g_2 has a cube factor, either $\mu(g_2/d)$ or $|\mathcal{S}(d)|$ vanishes, and $M_0 = 0$. Otherwise, we can write $g_2 = g_3^2g_4$ with g_3, g_4 both squarefree; then we require d to be a multiple of g_3. Thus

$$M_0 = \frac{\mu(g_1)}{\varphi(g_1)} \frac{\mu(g_3)}{\varphi(g_3)} \sum_{d|g_4} \frac{\mu(g_4/d)}{\varphi(g_4/d)}|\mathcal{S}(g_3d)|.$$

We use a few estimates for classical functions in our case. When $m|P_Y$,

$$\frac{\varphi(m)}{|m|} \ge \prod_{|\varpi|=\widehat{Y}} \Big(1 - \frac{1}{|\varpi|}\Big) = \Big(1 - \frac{1}{\widehat{Y}}\Big)^{\pi(Y)} \gg 1$$

https://doi.org/10.4153/CJM-2017-023-x Published online by Cambridge University Press

and thus $\varphi(m) \asymp m$. Also, when $m|P_Y$ with $\omega(m) \leq X/2Y = \widehat{Y}^{1/\kappa}/2\mathcal{L}(\widehat{Y})$, and for any $\nu \geq 1/\kappa$,

$$\sum_{d|m} \frac{1}{|d|^\nu} = \Big(1 + \frac{1}{|\varpi|^\nu}\Big)^{\omega(m)} \leq \exp\big(\widehat{Y}^{1/\kappa-\nu}/\mathcal{L}(\widehat{Y})\big) \ll 1,$$

and thus $\sum_{d|m} |d|^\nu = |m|^\nu \sum_{d|m} |d|^{-\nu} \asymp |m|^\nu$.

Combined with Lemma 6.2, we have

$$\begin{aligned} M_0/|\mathcal{S}| &\ll \frac{\mu^2(g_1)}{\varphi(g_1)} \cdot \frac{1}{|g_3|} \sum_{d|g_4} \frac{1}{|g_4/d||g_3 d|^{1-1/\kappa}} = \frac{\mu^2(g_1)}{\varphi(g_1)} \frac{1}{|g_3|^{2-1/\kappa}|g_4|} \sum_{d|g_4} |d|^{1/\kappa} \\ &\ll \frac{\mu^2(g_1)}{\varphi(g_1)} \frac{1}{|g_3|^{2-1/\kappa}|g_4|^{1-1/\kappa}} \leq \frac{\mu^2(g_1)}{\varphi(g_1)} \frac{1}{|g_2|^{1-1/\kappa}}, \end{aligned}$$

because $g_3 g_4 | P_Y$.

Then by writing $g = g_1 g_2$ and splitting $g_2 = 1$ and $|g_2| \geq \widehat{Y}$, (6.1) becomes

$$\begin{aligned} \int_{\mathfrak{M}} M(\alpha)^2 M(-2\alpha)\, d\alpha = \frac{1}{\widehat{X}} & \sum_{\substack{|g_1|<\widehat{Q} \\ (g_1,P_Y)=1}} \frac{\mu(g_1)}{\varphi(g_1)^2} |\mathcal{S}|^3 \\ & + O\Big(\frac{|\mathcal{S}|^3}{\widehat{X}} \sum_{\substack{|g_1|<\widehat{Q} \\ (g_1,P_Y)=1}} \frac{\mu^2(g_1)}{\varphi(g_1)^2} \sum_{\substack{\widehat{Y}\leq|g_2|<\widehat{Q}/|g_1| \\ \varpi|g_2 \Rightarrow \varpi|P_Y}} \frac{\varphi(g_2)}{|g_2|^{3-3/\kappa}} \Big) \end{aligned}$$

The inner sum over g_2 can be simplified to $O(\widehat{Y}^{-1+3/\kappa})$ when $\kappa > 3$, and thus the major arc contribution becomes $O(\frac{|\mathcal{S}|^3}{\widehat{X}} \cdot \widehat{Y}^{-1+3/\kappa})$. Also,

$$\sum_{\substack{|g_1|<\widehat{Q} \\ (g_1,P_Y)=1}} \frac{\mu(g_1)}{\varphi(g_1)^2} = \mathfrak{S} + O\Big(\frac{1}{\widehat{Y}}\Big),$$

where the singular series $\mathfrak{S}$ is defined by $\mathfrak{S} = \prod_\varpi \big(1 - \frac{1}{(|\varpi|-1)^2}\big)$, where the product is taken over all primes.

Thus if $\kappa > 3$,

$$\int_{\mathfrak{M}} M^3\, d\alpha = \frac{\mathfrak{S}|\mathcal{S}|^3}{\widehat{X}} \Big(1 + O\Big(\frac{1}{\widehat{Y}} + \frac{1}{\widehat{Y}^{1-3/\kappa}}\Big)\Big).$$

Thus

$$\int E(\alpha)^2 E(-2\alpha)\, d\alpha = \mathfrak{S}\frac{|\mathcal{S}|^3}{\widehat{X}} \Big(1 + O\Big(\frac{1}{\widehat{Y}^{1-3/\kappa}}\Big)\Big) + O(\widehat{X}^{3/4-1/2\kappa+\epsilon}|\mathcal{S}|).$$

Then the main term dominates the error term when $|\mathcal{S}| = \widehat{X}^{1-1/\kappa+o(1)} > \widehat{X}^{7/8-1/4\kappa+\epsilon/2}$. Thus when $\kappa \geq 6 + \delta$ and Y is large, we have

$$N(\mathcal{S}) = \mathfrak{S}\frac{|\mathcal{S}|^3}{\widehat{X}} \Big(1 + O\Big(\frac{1}{\widehat{Y}^{1/2}}\Big)\Big)$$

as desired.

7 Proof of Theorem 2.2

To impose the coprimality condition on $a + b = 2c$, we use the sieve argument on solutions. Let ϵ be given and Y be sufficiently large. In previous sections, the variable X does not vary from line to line and we omitted the implied dependency on $\mathcal{S}_r$; in this section, we clarify the dependency of X on $\mathcal{S}_r$. We write for any integer Z,

$$\mathcal{S}_1^{(Z)} = \{ m \in \mathbf{M}_0 : \deg m = Z, \varpi | m \Longrightarrow \deg \varpi = Y \},$$

where $\mathbf{M}_0$ is the set of squarefree monic polynomials, and similar to the previous notation, $\mathcal{S}_r^{(Z)}$ denotes the $m \in \mathcal{S}_1^{(Z)}$ such that $(m, r) = 1$. It can be shown from the definition that for given $r \in \mathcal{S}_1$, the set of elements in $\mathcal{S}_1$ divisible by r would be

$$\mathcal{S}_1^{(X)}(r) = r \cdot \mathcal{S}_r^{(X-\omega(r)Y)}. \tag{7.1}$$

As they appear frequently, we write $X_r = X - \omega(r)Y$ for simplicity. The following lemma is immediate.

Lemma 7.1 *Let r be a monic squarefree polynomial of degree at most $X/2$ and all prime factors of r are of degree Y. Then, for $\widehat{Y} \geq \mathcal{L}(\widehat{X})^2$, $|\mathcal{S}_r^{(X_r)}| \ll |\mathcal{S}_1^{(X)}||r|^{-1+1/\kappa}$.*

Proof Immediate from Lemma 6.2 and (7.1). ∎

Let r be a given squarefree polynomial all of whose prime factors are of degree Y and we count the number of solutions to $a + b = 2c$ with a, b, $c \in \mathcal{S}_1^{(X)}$, and $r \mid \gcd(a, b, c)$. Then by writing $a = ra'$, $b = rb'$, and $c = rc'$, the number of solutions to $a + b = 2c$ is equal to that of $a' + b' = 2c'$, where all variables are now from $\mathcal{S}_r^{(X_r)}$. When $\omega(r) \leq K/2$, the condition $\mathcal{L}(\widehat{X_r}) \leq \widehat{Y}^{1/\kappa}$ is satisfied when $\mathcal{L}(\widehat{X}) \leq 2\widehat{Y}^{1/\kappa}$. Thus Proposition 4.2 remains to hold for $\kappa > 6$ and all sufficiently large Y. Thus, the number of solutions is

$$N(\mathcal{S}_r^{(X_r)}) = \mathfrak{S} \frac{|\mathcal{S}_r^{(X_r)}|^3}{\widehat{X_r}} \Big(1 + O\Big(\frac{1}{\widehat{Y}^{1/2}}\Big)\Big).$$

By Lemma 7.1, we have $N(\mathcal{S}_r^{(X_r)}) \ll |r|^{-2+3/\kappa} \cdot |\mathcal{S}_1^{(X)}|^3/\widehat{X}$.

Now when $N^*(\mathcal{S})$ denotes the number of primitive solutions, we have the estimate

$$N^*(\mathcal{S}_1^{(X)}) = \sum_{r|P_Y} \mu(r) N(\mathcal{S}_r^{(X_r)}) \geq N(\mathcal{S}_1^{(X)}) - \sum_{|\varpi|=\widehat{Y}} N(\mathcal{S}_\varpi^{(X_\varpi)}).$$

Therefore, $N(\mathcal{S}_1^{(X)}) - N^*(\mathcal{S}_1^{(X)}) \ll \pi(Y)\widehat{Y}^{-2+3/\kappa} N(\mathcal{S}_1^{(X)}) \ll \widehat{Y}^{-1/2} N(\mathcal{S}_1^{(X)})$, which finishes the proof.

8 Proof of Theorem 1.1 and Theorem 1.3

We finally show that Theorem 2.2 implies Theorem 1.1 and Theorem 1.3. Let Y be a large integer, and $\kappa > 6$ be given. Take X to be the smallest integer that is multiple of Y and $\widehat{Y} \leq \mathcal{L}(\widehat{X})^\kappa$ and $H = X$. From the definition of $A(\kappa)$ in (1.5), we have $A(\kappa) \cap B(H) \supseteq N^*(\mathcal{S}_1^{(X)})$. Then for any $\epsilon > 0$ and $\kappa > 6$, we have $\# A(\kappa) \cap B(H) \geq$

https://doi.org/10.4153/CJM-2017-023-x Published online by Cambridge University Press

$\widehat{H}^{2-3/\kappa-\epsilon}$ and $\#B(H) \le \widehat{H}^3$, when Y is sufficiently large. Thus, the set in the definition of (1.6) contains every $\kappa > 6$, and it follows that $\kappa_0^{\text{new}} \le 6$, which proves Theorem 1.3.

When S is the set of irreducible polynomials of degree up to Y, we have

$$|S| = \sum_{Z \le Y} \pi(Z) \le \sum_{Z \le Y} \frac{q^Z}{Z} \le \sum_{Y/2 < Z \le Y} \frac{q^Z}{Z} + O\big(\widehat{Y}^{1/2}\big) \le \frac{2}{Y}\frac{q}{q-1}\widehat{Y} + O\big(\widehat{Y}^{1/2}\big),$$

and thus when Y is large, $|S| \le \widehat{Y}$. Let $\kappa > 6$ and $0 < \epsilon \le (\kappa - 6)/2$ be small, and let Y be sufficiently large, which depends on κ and ϵ. Let X be the least integer that is a multiple of Y and $\mathcal{L}(\widehat{X})^{\kappa-\epsilon} \ge \widehat{Y}$. Then the number of solutions to the S-unit equation is at least $N^*(\mathcal{S}_1)$. Then $N^*(\mathcal{S}_1) \ge q^{(2-3/(\kappa-\epsilon)-\epsilon)\widehat{Y}^{1/(\kappa-\epsilon)}} \ge q^{\widehat{Y}^{1/\kappa}} \ge q^{|S|^{1/\kappa}}$ for all sufficiently large Y.

Acknowledgments The author would like to thank Kannan Soundararajan for his advice and encouragement. A major portion of this work was done at Stanford University, when the author was granted a Samsung Scholarship. The author thanks both the Samsung Scholarship committee and the department of Mathematics at Stanford University for their support. This manuscript has been significantly improved from the earlier draft thanks to the anonymous referee(s) and the editor(s). The author would like to thank them for their efforts.

References

[1] Sary Drappeau, *Sur les solutions friables de l'équation $a + b = c$*. Math. Proc. Cambridge Philos. Soc. **154**(2013), no. 3, 439–463. http://dx.doi.org/10.1017/S0305004112000643

[2] ———, *Théorèmes de type Fouvry-Iwaniec pour les entiers friables*. Compos. Math. **151**(2015), no. 5, 828–862. http://dx.doi.org/10.1112/S0010437X14007933

[3] P. Erdös, C. L. Stewart, and R. Tijdeman, *Some Diophantine equations with many solutions*. Compositio Math. **66**(1988), no. 1, 37–56.

[4] J.-H. Evertse, *On equations in S-units and the Thue-Mahler equation*. Invent. Math. **75**(1984), no. 3, 561–584. http://dx.doi.org/10.1007/BF01388644

[5] Junsoo Ha, *Some problems in multiplicative number theory*. Ph.D. thesis, Stanford University, 2014.

[6] A. J. Harper, *On finding many solutions to S-unit equations by solving linear equations on average*. arxiv:1108.3819

[7] ———, *Bombieri–Vinogradov and Barban–Davenport–Halberstam type theorems for smooth numbers*. arxiv:1208.5992

[8] ———, *Minor arcs, mean values, and restriction theory for exponential sums over smooth numbers*. Compos. Math. **152**(2016), no. 6, 1121–1158. http://dx.doi.org/10.1112/S0010437X15007782

[9] David R. Hayes, *The distribution of irreducibles in* GF$[q, x]$. Trans. Amer. Math. Soc. **117**(1965), 101–127.

[10] ———, *The expression of a polynomial as a sum of three irreducibles*. Acta Arith. **11**(1966), 461–488.

[11] ———, *Explicit class field theory for rational function fields*. Trans. Amer. Math. Soc. **189**(1974), 77–91. http://dx.doi.org/10.1090/S0002-9947-1974-0330106-6

[12] Chih-Nung Hsu, *The distribution of irreducible polynomials in* $\mathbf{F}_q[t]$. J. Number Theory **61**(1996), no. 1, 85–96. http://dx.doi.org/10.1006/jnth.1996.0139

[13] S. Konyagin and Kannan Soundararajan, *Two S-unit equations with many solutions*. J. Number Theory **124**(2007), no. 1, 193–199 http://dx.doi.org/10.1016/j.jnt.2006.07.017

[14] R. M. Kubota, *Waring's problem for* $\mathbf{F}_q[x]$. Dissertationes Math. (Rozprawy Mat.) **117**(1974), 60.

[15] Jeffrey C. Lagarias and Kannan Soundararajan, *Smooth solutions to the abc equation: the xyz conjecture*. J. Théor. Nombres Bordeaux **23**(2011), no. 1, 209–234. http://dx.doi.org/10.5802/jtnb.757

[16] ———, *Counting smooth solutions to the equation $A + B = C$*. Proc. Lond. Math. Soc. **104**(2012), no. 4, 770–798. http://dx.doi.org/10.1112/plms/pdr037

[17] Yu-Ru Liu and Trevor D. Wooley, *Waring's problem in function fields*. J. Reine Angew. Math. **638**(2010), 1–67. http://dx.doi.org/10.1515/crelle.2010.001

[18] E. Manstavičius, *Remarks on elements of semigroups that are free of large prime factors*. Liet. Mat. Rink. **32**(1992), no. 4, 512–525.

[19] ———, *Semigroup elements free of large prime factors.* In: New trends in probability and statistics. Vol. 2 (Palanga, 1991), pages 135–153. VSP, Utrecht, 1992.

[20] Dinakar Ramakrishnan and Robert J. Valenza, *Fourier analysis on number fields*, Graduate Texts in Mathematics, 186. Springer-Verlag, New York, 1999.

[21] Michael Rosen, *Number theory in function fields*. Graduate Texts in Mathematics, 210. Springer-Verlag, New York, 2002.

Korea Institute for Advanced Study, 85 Hoegiro, Dongdaemungu, Seoul, Republic of Korea
e-mail: junsooha@kias.re.kr

https://doi.org/10.4153/CJM-2017-023-x Published online by Cambridge University Press