J. reine angew. Math. 438 (1993), 143—161

Journal für die reine und angewandte Mathematik

Estimates for Soto-Andrade sums

By *Nicholas M. Katz* at Princeton

Introduction and statement of results

Let k be a finite field of characteristic p, with q elements, K/k a finite extension field of degree $n \geqq 2$. Denote by

$$\mathbb{N}: K \to k, \quad \mathrm{Tr}\colon K \to k,$$

the norm and trace maps for K/k. Denote by U the multiplicative subgroup of $K^\times$ consisting of those elements α in $K^\times$ with $\mathbb{N}\alpha = 1$. Given any character of U,

$$\omega: U \to \mathbb{C}^\times,$$

any character of $k^\times$,

$$\varepsilon\colon k^\times \to \mathbb{C}^\times, \text{ extended to } k \text{ by putting } \varepsilon(0) := 0,$$

and any element t in k, consider the "Soto-Andrade sum" (cf. [S-A])

$$\sum_{x \text{ in } U} \varepsilon(\mathrm{Tr}(x) + t)\,\omega(x).$$

Theorem 1. *Suppose that $n = 2$. Then we have the estimate*

$$\Big|\sum_{x \text{ in } U} \varepsilon(\mathrm{Tr}(x) + t)\,\omega(x)\Big| \leqq 2(q)^{1/2}$$

unless we are in one of the following two exceptional cases:

exceptional case 1) *both ε and ω are trivial,*

exceptional case 2) *both ε and ω have exact order* 2, *and $t = \pm 2$.*

For higher n, it is more convenient to consider a slightly more general sum, depending also on a parameter α in $K^\times$:

$$\sum_{x \text{ in } U} \varepsilon\big(t - \mathrm{Tr}(\alpha x)\big)\,\omega(x)\,.$$

(For $\alpha = -1$, we recover the Soto-Andrade sum as defined above.)

Theorem 2. *Suppose that $n \geqq 2$, α in $K^\times$, and ε nontrivial. Then we have*

$$\Big|\sum_{x \text{ in } U} \varepsilon\big(t - \mathrm{Tr}(\alpha x)\big)\,\omega(x)\Big| \leqq n(q)^{(n-1)/2}$$

unless we are in the following exceptional case:

n is prime to p, n is even, ε has exact order 2, ω has exact order n, the n characters $\omega^{((q^i-1)/(q-1))}$, $0 \leqq i \leqq n-1$, are all distinct, and $t^n = n^n\, \mathbb{N}(\alpha)$.

In this exceptional case, we have the estimate

$$\left|\,\Big|\sum_{x \text{ in } U} \varepsilon\big(t - \mathrm{Tr}(\alpha x)\big)\,\omega(x)\Big| - (q)^{n/2}\right| \leqq n(q)^{(n-1)/2}\,.$$

Theorem 3. *Suppose that $n \geqq 2$, α in $K^\times$, ε is trivial, and ω is nontrivial. Then we have*

$$\Big|\sum_{x \text{ in } U} \varepsilon\big(t - \mathit{Tr}(\alpha x)\big)\,\omega(x)\Big| \leqq n(q)^{(n-2)/2}$$

unless we are in the following exceptional case:

n is prime to p, n is odd, ω has exact order n, the n characters $\omega^{((q^i-1)/(q-1))}$, $0 \leqq i \leqq n-1$, are all distinct, and $t^n = n^n\, \mathbb{N}(\alpha)$.

In this exceptional case, we have the estimate

$$\left|\,\Big|\sum_{x \text{ in } U} \varepsilon\big(t - \mathrm{Tr}(\alpha x)\big)\,\omega(x)\Big| - (q)^{(n-1)/2}\right| \leqq n(q)^{(n-2)/2}\,.$$

Theorem 4. *Suppose that $n \geqq 2$, α in $K^\times$, b in $k^\times$, t in k. Denote by*

$$f(t) := \mathrm{Card}\{x \text{ in } K^\times \text{ with } \mathbb{N}x = b \text{ and } \mathrm{Tr}(\alpha x) = \mathrm{t}\}\,.$$

We have the estimate

$$|f(t) - (q^n - 1)/q(q-1)| \leqq n(q)^{(n-2)/2}\,.$$

Theorem 1 is, of course, a formal consequence of Theorems 2 and 3: take $n = 2$ and $\alpha = -1$. It is isolated here both because of its interest to graph theorists (it proves that Audrey Terras' "upper half plane H_q" is a "Ramanujan graph", cf. [Ter], pp. 89–91), and because, unlike Theorems 2 and 3, it admits an elementary reduction to the Riemann Hypothesis for curves over finite fields, proven by Weil a half-century ago. For these reasons, we will devote a few pages to a separate proof of Theorem 1, which we hope will be at least slightly intelligible to those who are not experts in ℓ-adic cohomology. It is a pleasure to thank Ron Evans for having told me both about the problem of estimating Soto-Andrade sums with $n = 2$, and about its relevance to proving that H_q is a Ramanujan graph.

Discussion and proof of Theorem 1

We will see that, because $n = 2$, Theorem 1 is an immediate consequence of the truth of the "Riemann Hypothesis" for curves over finite fields ([We]), together with Grothendieck's Euler-Poincaré formula ([Ray]).

Let us first recall how, for general n, to use Weil's "restriction of the ground field" functor to reduce the study of such sums to the study of more "usual" character sums in finite fields. The "trick" is to view the group U as the group of k-valued points of a certain algebraic torus $\boldsymbol{U}$ over k which over $\bar{k}$ (and indeed already over K) becomes isomorphic (for general n) to the $n-1$ fold self-product of the multiplicative group $\mathbb{G}_m$. The algebraic group $\boldsymbol{U}$ is defined as follows. For any commutative, associative, unitary k-algebra A, $K \otimes_k A$ is an A-algebra which is free of rank n as an A-module. Therefore we may speak of the relative norm and trace

$$\mathbb{N}_A : K \otimes_k A \to A, \quad \mathrm{Tr}_A : K \otimes_k A \to A.$$

Then $\boldsymbol{U}$ as a functor on k-algebras is defined by

$$\boldsymbol{U}(A) := \{\alpha \text{ in } K \otimes_k A \text{ with } \mathbb{N}_A(\alpha) = 1\}.$$

Concretely, if we pick a k-basis $e_1, \ldots, e_n$ of K as a k-vector space, then for indeterminates $X_1, \ldots, X_n$, the polynomial in $K[X_i\text{'s}]$

$$\mathbb{N}(\sum_i X_i e_i) := \prod_{\sigma \text{ in } \mathrm{Gal}(K/k)} (\sum_i X_i(\sigma e_i))$$

actually has coefficients in k, and $\boldsymbol{U}$ is the hypersurface in affine n-space over k defined by the equation $\mathbb{N}(\sum_i X_i e_i) = 1$. In these coordinates, the trace map Tr_A becomes the restriction to A-valued points of the morphism of algebraic varieties over k

$$\mathbf{Tr} : \boldsymbol{U} \to \mathbb{A}^1, \quad (X_i) \mapsto \sum_i \mathrm{Tr}(e_i) X_i.$$

Notice that if α in $K \otimes_k A$ has $\mathbb{N}_A(\alpha) = 1$, then α is invertible in $K \otimes_k A$ (by the Cayley-Haimilton theorem). So $\boldsymbol{U}$ is naturally a subgroup-scheme of the algebraic group $\boldsymbol{K}^\times$ over k defined by

$$\boldsymbol{K}^\times(A) := (K \otimes_k A)^\times.$$

Concretely, $\boldsymbol{K}^\times$ is defined by the n quadratic equations in $k[X_i\text{'s}, Y_i\text{'s}]$ obtained from the equation

$$(\sum_i X_i e_i)(\sum_i Y_i e_i) = 1 \text{ in } K[X_i\text{'s}, Y_i\text{'s}]$$

by equating coefficients of the basis elements e_i of K/k. In these coordinates, the norm map $\mathbb{N}_A$ becomes the restriction to A-valued points of the homomorphism of algebraic groups over k

$$\mathbb{N}: \boldsymbol{K}^\times \to \mathbb{G}_m, \quad (X_i, Y_i) \to \mathbb{N}(\sum_i X_i e_i),$$

and U becomes the kernel of this norm homomorphism.

Because K/k is galois, for any K-algebra B, the B-algebra $K \otimes_k B$ is isomorphic to the n-fold product (indexed by elements of $\mathrm{Gal}(K/k)$) of B with itself, by the B-linear map

$$K \otimes_k B \to \prod_{\sigma \text{ in } \mathrm{Gal}(K/k)} B$$

defined by

$$\sum \alpha_i \otimes b_i \mapsto \prod_{\sigma \text{ in } \mathrm{Gal}(K/k)} (\Sigma \sigma(\alpha_i) b_i).$$

Therefore, after we extend scalars from k to K, the algebraic group $\boldsymbol{K}^\times$ becomes isomorphic to the n-fold product of $\mathbb{G}_m$ over K:

$$\boldsymbol{K}^\times \otimes K \cong \prod_{\sigma \text{ in } \mathrm{Gal}(K/k)} \mathbb{G}_{m,K}.$$

Concretely, the new coordinates Z_σ on this product are defined by

$$Z_\sigma := \sum_i \sigma(e_i) X_i.$$

In these coordinates, the norm and trace maps are given by

$$\prod_\sigma Z_\sigma \quad \text{and} \quad \sum_\sigma Z_\sigma \quad \text{respectively}.$$

The subgroup $\boldsymbol{U} \otimes K$ of $\boldsymbol{K}^\times \otimes K$ is defined by the equation $\prod_\sigma Z_\sigma = 1$. If we identify $\mathrm{Gal}(K/k)$ as a set with $\{1, 2, \ldots, n\}$, then $\boldsymbol{U} \otimes K$ becomes the $n-1$ torus with coordinates $Z_1, \ldots, Z_{n-1}$, and the trace map restricted to $\boldsymbol{U} \otimes K$ is

$$\mathrm{Tr}: \boldsymbol{U} \otimes K \to \mathbb{A}^1_K,$$

$$(Z_1, \ldots, Z_{n-1}) \mapsto Z_1 + Z_2 + \cdots + Z_{n-1} + 1/(Z_1 Z_2 \ldots Z_{n-1}).$$

With these preliminaries out of the way, we now return to the more direct discussion of the Soto-Andrade sum

$$\sum_{x \text{ in } U} \varepsilon(\mathrm{Tr}(x) + t)\, \omega(x),$$

attached to a character of U,

$$\omega : U \to \mathbb{C}^\times ,$$

a character of $k^\times$,

$$\varepsilon : k^\times \to \mathbb{C}^\times, \text{ extended to } k \text{ by putting } \varepsilon(0) := 0 ,$$

and an element t in k.

Pick a prime $\ell \neq \mathrm{char}(k)$, and an embedding of $\mathbb{C}$ into $\bar{\mathbb{Q}}_\ell$. The "Lang torsor" construction (cf. [De-AFT]) attaches to the character ω of $U = \mathbb{U}(k)$ a lisse, rank on $\bar{\mathbb{Q}}_\ell$-sheaf $\mathscr{L}_\omega$ on $\mathbb{U}$, whose trace of Frobenius at a rational point α in $\mathbb{U}(k) = U$ is $\omega(\alpha)$. Similarly, it attaches to the character ε of $k^\times = \mathbb{G}_m(k)$ a lisse, rank one $\bar{\mathbb{Q}}_\ell$-sheaf $\mathscr{L}_\varepsilon$ on $\mathbb{G}_{m,k}$ whose trace of Frobenius at a rational point a in $\mathbb{G}_m(k) = k^\times$ is $\varepsilon(a)$. The sheaves $\mathscr{L}_\varepsilon$ and $\mathscr{L}_\omega$ on $\mathbb{G}_{m,k}$ and $\mathbb{U}$ respectively are of finite order, hence both are pure of weight zero.

For t in k, we have the function

$$\mathbb{U} \to \mathbb{A}^1_k \quad \text{defined by} \quad u \mapsto \mathrm{Tr}(u) + t .$$

We denote by $\mathbb{U}[1/(\mathrm{Tr} + t)]$ the open set of $\mathbb{U}$ where this function is invertible. On this open set, we can speak of the pullback

$$(\mathrm{Tr} + t)^*(\mathscr{L}_\varepsilon) := \mathscr{L}_{\varepsilon(\mathrm{Tr}+t)} .$$

By Grothendieck's Lefschetz trace formula ([Gr-FL]), the Soto-Andrade sum is the alternating sum of the traces of Frobenius F_k on the ℓ-adic etale cohomology groups with compact support:

$$\sum_{x \text{ in } U} \varepsilon(\mathrm{Tr}(x) + t)\, \omega(x)$$

$$= \sum_i (-1)^i \,\mathrm{Trace}\big(F_k, H^i_c(\mathbb{U}[1/(\mathrm{Tr} + t)] \otimes \bar{k}, \mathscr{L}_{\varepsilon(\mathrm{Tr}+t)} \otimes \mathscr{L}_\omega)\big) .$$

In the case $n = 2$ (when $\mathbb{U}[1/(\mathrm{Tr} + t)]$ is a curve, indeed it is $\mathbb{G}_m$ with at most two points removed) we know by [We], and in the case of general n (when $\mathbb{U}[1/(\mathrm{Tr} + t)]$ is a smooth geometrically connected affine variety of dimension $n - 1$) we know by [De-Weil II], that the above H^i_c is mixed of weight $\leqq i$, and that it vanishes unless

$$n - 1 \leqq i \leqq 2(n - 1) .$$

We henceforth restrict our attention to the case $n = 2$. Then the only possibly nonvanishing cohomology groups are H^1_c and H^2_c. We must show that, save in the two exceptional cases, we have

a) $$H^2_c(\mathbb{U}[1/(\mathrm{Tr} + t)] \otimes \bar{k}, \mathscr{L}_{\varepsilon(\mathrm{Tr}+t)} \otimes \mathscr{L}_\omega) = 0 ,$$

and

b) $$\dim H_c^1(U[1/(\mathrm{Tr}+t)]\otimes \bar{k}, \mathscr{L}_{\varepsilon(\mathrm{Tr}+t)}\otimes \mathscr{L}_\omega) \leqq 2\,.$$

These are both *geometric* statements. So we may identify $U \otimes \bar{k} \cong \mathbb{G}_{m,\bar{k}}$, with coordinate Z. Then $\mathscr{L}_\omega$ becomes the sheaf $\mathscr{L}_{\tilde{\omega}}$ on $\mathbb{G}_{m,\bar{k}}$ which is the pullback from $\mathbb{G}_{m,K}$ of the Lang torsor attached to the character $\tilde{\omega}$ of $K^\times$ defined as the composition of ω with the surjective K/k norm map from $K^\times = U(K)$ to $U = U(k)$. (In fact, this map is $x \mapsto x^{1-q}$.)

The variety $U[1/\mathrm{Tr}+t)]\otimes \bar{k}$ itself becomes

$$\mathbb{G}_{m,\bar{k}}[1/(Z+1/Z+t)] = \mathbb{G}_{m,\bar{k}}[1/(Z^2+tZ+1)],$$

the last equality because Z is an invertible function on $\mathbb{G}_m$.

The sheaf $\mathscr{L}_{\varepsilon(\mathrm{Tr}+t)}\otimes \mathscr{L}_\omega$ becomes

$$\mathscr{L}_{\varepsilon(Z+1/Z+t)}\otimes \mathscr{L}_{\tilde{\omega}(Z)} = \mathscr{L}_{\varepsilon(Z^2+tZ+1)}\otimes \mathscr{L}_{\bar{\varepsilon}\tilde{\omega}(Z)}\,.$$

This is a lisse sheaf of rank one, and finite order prime to p (because ε has order dividing $q-1$, and $\tilde{\omega}$ has order dividing $q+1$ (the order of U). So this sheaf is tamely ramified at all the "points at ∞". Therefore its Euler-Poincaré characteristic is the same as for the constant sheaf:

$$\begin{aligned}
&\chi_c(U[1/(\mathrm{Tr}+t)]\otimes \bar{k}, \mathscr{L}_{\varepsilon(\mathrm{Tr}+t)}\otimes \mathscr{L}_\omega)\\
&\quad = \chi_c(U[1/(\mathrm{Tr}+t)]\otimes \bar{k}, \bar{\mathbb{Q}}_\ell)\\
&\quad = \chi_c(\mathbb{G}_{m,\bar{k}}[1/(Z^2+tZ+1)], \bar{\mathbb{Q}}_\ell)\\
&\quad = \chi_c(\mathbb{G}_{m,\bar{k}}, \bar{\mathbb{Q}}_\ell) - \{\text{number of zeroes of } Z^2+tZ+1 \text{ in } \bar{k}\}\\
&\quad = -\{\text{number of zeroes of } Z^2+tZ+1 \text{ in } \bar{k}\}\\
&\quad = -1 \text{ or } -2.
\end{aligned}$$

Since the only possibly nonvanishing groups are H_c^1 and H_c^1, if H_c^2 vanishes, then $\dim H_c^1 \leqq 2$ is automatic.

To prove a), we argue by contradiction, and suppose that H_c^2 is nonzero. Since our sheaf is lisse of rank one, then

$$\mathscr{L}_{\varepsilon(Z^2+tZ+1)}\otimes \mathscr{L}_{\bar{\varepsilon}\tilde{\omega}(Z)} \text{ is constant on } \mathbb{G}_{m,\bar{k}} - \{\text{zeroes of } Z^2+tZ+1\}\,.$$

Suppose first that $t \neq \pm 2$. Then Z^2+tZ+1 has 2 distinct roots in $\bar{k}$, and both roots are nonzero. If ε is nontrivial, the sheaf

$$\mathscr{L}_{\varepsilon(Z^2+tZ+1)}\otimes \mathscr{L}_{\bar{\varepsilon}\tilde{\omega}(Z)}$$

is visibly ramified at each of these two points, hence cannot be geometrically constant. If ε is trivial, this sheaf if visibly ramified at the origin $Z = 0$, provided only that ω (and hence $\tilde{\omega}$) is nontrivial. So only exceptional case 1) arises here.

Suppose now that $t = \pm 2$. Then

$$Z^2 + tZ + 1 = (Z \pm 1)^2 ,$$

$$\mathscr{L}_{\varepsilon(Z^2+tZ+1)} \otimes \mathscr{L}_{\bar{\varepsilon}\tilde{\omega}(Z)} \cong \mathscr{L}_{\varepsilon^2(Z\pm 1)} \otimes \mathscr{L}_{\bar{\varepsilon}\tilde{\omega}(Z)} .$$

If this sheaf is constant, then it is unramified at both 0 and ± 1, so we must have both ε^2 and $\tilde{\omega}/\varepsilon$ trivial. If ε is trivial, so is $\tilde{\omega}$, and hence ω itself is trivial. This is exceptional case 1). If ε has exact order 2, then $\tilde{\omega}$ must have exact order 2, and hence ω itself has exact order 2. This is exceptional case 2). QED for Theorem 1.

Proofs of Theorems 2 and 3, and a Corollary

Let us recall the situation. We are given K/k an extension of degree $n \geqq 2$ of finite fields, $p := \mathrm{char}(k)$, $q := \mathrm{Card}(k)$, ε a $\mathbb{C}$-valued character of $k^\times$, and ω a $\mathbb{C}$-valued character of the subgroup U of elements of $K^\times$ of norm one.

We pick a $\mathbb{C}$-valued character χ of $K^\times$ which agrees with ω on the subgroup U. Concretely, the group $K^\times$ is cyclic of order $q^n - 1$, U is its unique subgroup of index $q - 1$, and so if ζ is a generator of $K^\times$, then ζ^{q-1} is a generator of U. Thus $\chi^{q-1}(\zeta) = \omega(\zeta^{q-1})$, and the value $\chi(\zeta)$ may be chosen as any $q-1$'st root of $\omega(\zeta^{q-1})$. In particular, the order of ω as character of U is equal to the order of χ^{q-1} as character of $K^\times$. Once we have one choice of χ, any other is of the form $\chi\varrho$, where ϱ is a character of $K^\times$ with ϱ^{q-1} trivial.

In terms of χ, we have the following reformulations of Theorems 2 and 3.

Theorem 2bis. *Suppose that $n \geqq 2$, α in $K^\times$, and ε nontrivial. Then we have*

$$\Big| \sum_{x \text{ in } U} \varepsilon\big(t - \mathrm{Tr}(\alpha x)\big)\chi(x) \Big| \leqq n(q)^{(n-1)/2}$$

unless we are in the following exceptional case:

n is prime to p, n is even, ε has exact order 2, χ^{q-1} has (exact) order n, the characters χ^{q^i-1} for $0 \leqq i \leqq n-1$ are all distinct, and $t^n = n^n \mathbb{N}(\alpha)$.

In this exceptional case, we have the estimate

$$\left| \Big| \sum_{x \text{ in } U} \varepsilon\big(t - \mathrm{Tr}(\alpha x)\big)\chi(x) \Big| - (q)^{n/2} \right| \leqq n(q)^{(n-1)/2} .$$

Theorem 3bis. *Suppose that $n \geqq 2$, α in $K^\times$, and ε is trivial. Then we have*

$$\Big| \sum_{x \text{ in } U} \varepsilon\big(t - \mathrm{Tr}(\alpha x)\big)\chi(x) \Big| \leqq n(q)^{(n-2)/2}$$

unless we are in the one of the two follwing exceptional cases:

exceptional case 1) χ^{q-1} *is also trivial,*

exceptional case 2) *n is prime to p, n is odd, χ^{q-1} has exact order n, the characters χ^{q^i-1} for $0 \leqq i \leqq n-1$ are all distinct, and $t^n = n^n \mathbb{N}(\alpha)$.*

In exceptional case 2), *we have the estimate*

$$\left| \left| \sum_{x \text{ in } U} \varepsilon(t - \operatorname{Tr}(\alpha x))\chi(x) \right| - (q)^{(n-1)/2} \right| \leqq n(q)^{(n-2)/2}.$$

Corollary of Theorem 3 bis. *Let K/k be a degree n extension of finite fields of characteristic p, with $q := \operatorname{Card}(k)$. Fix elements t in k, and a in $k^\times$, and denote by*

$$X(t, a) := \{x \text{ in } K^\times \text{ with } \operatorname{Tr}(x) = t \text{ and } \mathbb{N}(x) = a\}.$$

For any $\mathbb{C}$-valued character χ of $K^\times$, consider the sum

$$\sum_{x \text{ in } X(t,a)} \chi(x).$$

We have the estimate

$$\left| \sum_{x \text{ in } X(t,a)} \chi(x) \right| \leqq n(q)^{(n-2)/2}$$

unless we are in the one of the two following exceptional cases:

exceptional case 1) χ^{q-1} *is trivial,*

exceptional case 2) *n is prime to p, n is odd, χ^{q-1} has exact order n, the characters χ^{q^i-1} for $0 \leqq i \leqq n-1$ are all distinct, and $t^n = n^n a$.*

In exceptional case 2), *we have the estimate*

$$\left| \left| \sum_{x \text{ in } X(t,a)} \chi(x) \right| - (q)^{(n-1)/2} \right| \leqq n(q)^{(n-2)/2}.$$

Proof of Corollary. Since the norm map $\mathbb{N}: K^\times \to k^\times$ is surjective, we may choose an element α in $K^\times$ with $\mathbb{N}(\alpha) = a$. Then

$$x \text{ in } X(t, a) \Leftrightarrow x = \alpha y \text{ with } y \text{ in } U \text{ and } \operatorname{Tr}(\alpha y) = t.$$

So we may rewrite the sum in question as

$$\sum_{x \text{ in } X(t,a)} \chi(x) = \sum_{y \text{ in } U, \operatorname{Tr}(\alpha y) = t} \chi(\alpha y) = \chi(\alpha) \sum_{y \text{ in } U, \operatorname{Tr}(\alpha y) = t} \chi(y).$$

Since $|\chi(\alpha)| = 1$, it suffices to prove the corollary for the sum

$$\sum_{y \text{ in } U, \operatorname{Tr}(\alpha y)=t} \chi(y).$$

If χ^{q-1} is nontrivial, then $\chi|U$ is nontrivial, $\sum_{y \text{ in } U} \chi(y) = 0$, and so

$$\sum_{y \text{ in } U, \operatorname{Tr}(\alpha y)=t} \chi(y) = -\sum_{y \text{ in } U, \operatorname{Tr}(\alpha y)\neq t} \chi(y).$$

In terms of the trivial character $\mathbb{1}$ of $k^\times$, extended to k by $\mathbb{1}(0) := 0$, this last sum is

$$-\sum_{y \text{ in } U, \operatorname{Tr}(\alpha y)\neq t} \chi(y) = -\sum_{y \text{ in } U} \mathbb{1}(t - \operatorname{Tr}(\alpha y))\chi(y),$$

to which we apply Theorem 3bis. QED for Corollary.

We now turn to the simultaneous proofs of Theorems 2bis and 3bis. Thus fix characters ε of $k^\times$ and χ of $K^\times$, and assume that either ε or χ^{q-1} is nontrivial. For each t in k, we have the sum

$$f(t) := \sum_{x \text{ in } U} \varepsilon(t - \operatorname{Tr}(\alpha x))\chi(x).$$

The idea is to view $t \mapsto f(t)$ as a complex valued function on k, and then to exploit the fact that its Fourier Transform has a simple expression in terms of Kloosterman sums.

Fix a nontrivial $\mathbb{C}$-valued additive character ψ of k. The Fourier Transform $\operatorname{FT}(f)$ of f is the $\mathbb{C}$-valued function on k given by

$$\begin{aligned} s \mapsto \sum_{t \text{ in } k} \psi(st)f(t) &= \sum_{t \text{ in } k}\sum_{x \text{ in } U} \psi(st)\varepsilon(t - \operatorname{Tr}(\alpha x))\chi(x) \\ &= \sum_{x \text{ in } U} \chi(x) \sum_{t \text{ in } k} \psi(st)\varepsilon(t - \operatorname{Tr}(\alpha x)). \end{aligned}$$

Making the change of variable $t \mapsto t + \operatorname{Tr}(\alpha x)$ in the inner sum, we get

$$\begin{aligned} \operatorname{FT}(f)(s) &= \sum_{x \text{ in } U} \chi(x) \sum_{t \text{ in } k} \psi(s(t + \operatorname{Tr}(\alpha x)))\varepsilon(t) \\ &= \sum_{x \text{ in } U} \chi(x)\psi(s\operatorname{Tr}(\alpha x)) \sum_{t \text{ in } k} \psi(st)\varepsilon(t). \end{aligned}$$

If $s = 0$, we have

$$\operatorname{FT}(f)(0) = \sum_{x \text{ in } U} \chi(x) \sum_{t \text{ in } k} \varepsilon(t) = 0,$$

because either ε or χ^{q-1} is nontrivial. (This is the *only* place we use the hypothesis that either ε or χ^{q-1} is nontrivial.)

For $s \neq 0$, we have

$$\sum_{t \text{ in } k} \psi(st)\varepsilon(t) = \bar{\varepsilon}(s) \sum_{t \text{ in } k} \psi(t)\varepsilon(t) := \bar{\varepsilon}(s) g(\psi, \varepsilon),$$

so

$$\begin{aligned}
\mathrm{FT}(f)(s) &= \bar{\varepsilon}(s) g(\psi, \varepsilon) \sum_{x \text{ in } U} \chi(x)\psi(s\mathrm{Tr}(\alpha x)) \\
&= \bar{\varepsilon}(s) g(\psi, \varepsilon) \sum_{x \text{ in } U} \chi(x)\psi(\mathrm{Tr}(s\alpha x)) \\
&= \bar{\varepsilon}(s) g(\psi, \varepsilon) \sum_{y \text{ in } K^\times,\, Ny = s^n N\alpha} \chi(y/s\alpha)\psi(\mathrm{Tr}(y)) \\
&= \bar{\varepsilon}(s) g(\psi, \varepsilon)\bar{\chi}(s\alpha) \sum_{y \text{ in } K^\times,\, Ny = s^n N\alpha} \chi(y)\psi(\mathrm{Tr}(y)).
\end{aligned}$$

In order to go further, we must recall a bit of the theory of the "exotic" Kloosterman sheaves [Ka-GKM], 8.8.4–7 attached to a finite etale k-algebra B of rank n (in our case K), a nondegenerate $\mathbb{C}$-valued additive character $\tilde{\psi}$ of B (in our case $\psi \circ \mathrm{Tr}$), and a $\mathbb{C}$-valued multiplicative character χ of B (in our case χ). Once we pick a prime $\ell \neq p$, and an embedding of $\mathbb{C}$ into $\bar{\mathbb{Q}}_\ell$, there exists a lisse $\bar{\mathbb{Q}}_\ell$-sheaf $\mathrm{Kl}(B, \tilde{\psi}, \chi)$ on $\mathbb{G}_m \otimes k$ with the property that for any finite extension E of k, and any point s in $E^\times = \mathbb{G}_m(E)$, the trace of Frobenius $F_{s,E}$ on $\mathrm{Kl}(B, \tilde{\psi}, \chi)$ is given by

$$\mathrm{Trace}(F_{s,F} | \mathrm{Kl}(B, \tilde{\psi}, \chi)) = (-1)^{n-1} \sum_{y \text{ in } B \otimes_k E;\, N_{B \otimes_k E/E}(y) = s} \tilde{\psi}(\mathrm{Tr}_{B \otimes_k E/B}(y))\chi(N_{B \otimes_k E/B}(y)).$$

In our case ($B = K, \tilde{\psi} = \psi \circ \mathrm{Tr}, \chi$), this sheaf, once we pull it back to $\mathbb{G}_m \otimes K$, becomes isomorphic to a usual Kloosterman sheaf [Ka-GKM], 8.8.7:

$$\mathrm{Kl}(K, \psi \circ \mathrm{Tr}, \chi) | \mathbb{G}_m \otimes K \cong \mathrm{Kl}(\psi \circ \mathrm{Tr}, \chi, \chi^q, \chi^{q^2}, \ldots, \chi^{q^{n-1}}).$$

In terms of $\mathrm{Kl}(K, \psi \circ \mathrm{Tr}, \chi)$, we may rewrite the formula for $\mathrm{FT}(f)(s)$, $s \neq 0$, as

$$\mathrm{FT}(f)(s) = (-1)^{n-1}\bar{\varepsilon}(s) g(\psi, \varepsilon)\bar{\chi}(s\alpha)\,\mathrm{Trace}(F_{s^n N\alpha, k} | \mathrm{Kl}(K, \psi \circ \mathrm{Tr}, \chi)).$$

(Recall that $\mathrm{FT}(f)(0) = 0$.) So by Fourier inversion, we have, for any t in k,

$$\begin{aligned}
(-1)^{n-1} f(t) &= (-1)^{n-1}(1/q) \sum_{s \text{ in } k} \bar{\psi}(st)\,\mathrm{FT}(f)(s) \\
&= (\bar{\chi}(\alpha) g(\psi, \varepsilon)/q) \sum_{s \text{ in } k^\times} \bar{\varepsilon}(s)\bar{\chi}(s)\bar{\psi}(st)\,\mathrm{Trace}(F_{s^n N\alpha, k} | \mathrm{Kl}(K, \psi \circ \mathrm{Tr}, \chi)).
\end{aligned}$$

To go further, consider the pullback sheaf

$$\mathscr{F} := [s \mapsto s^n N\alpha]^* \mathrm{Kl}(K, \psi \circ \mathrm{Tr}, \chi)$$

on $\mathbb{G}_m \otimes k$. Its trace at $s \neq 0$ in k is tautologically given by

$$\mathrm{Trace}(F_{s,k} | \mathscr{F}) = \mathrm{Trace}(F_{s^n N\alpha, k} | \mathrm{Kl}(K, \psi \circ \mathrm{Tr}, \chi)).$$

Let us temporarily denote by ψ_t the additive character of k defined by

$$\psi_t(s) := \psi(st),$$

and by ϱ the multiplicative character of $k^\times$ which is the restriction to $k^\times$ of χ:

$$\varrho(s) := \chi(s).$$

Notice that as characters of $K^\times$, we have

$$\varrho \circ \mathbb{N} = \chi^{1+q+q^2} + \cdots + q^{n-1},$$

and this uniquely characterizes ϱ as a character of $k^\times$. In terms of the Artin-Schreier sheaf $\mathscr{L}_{\psi_t}$ on $\mathbb{A}_k^1$, restricted to $\mathbb{G}_m \otimes k$, the Kummer sheaf $\mathscr{L}_{\varrho\varepsilon}$ on $\mathbb{G}_m \otimes k$, and the above sheaf $\mathscr{F}$, we may rewrite the above formula for $f(t)$ as follows:

$$(-1)^{n-1} q f(t) / (\bar{\chi}(\alpha) g(\psi, \varepsilon)) = \sum_{s \text{ in } k^\times} \mathrm{Trace}(F_{s,k} | \mathscr{F} \otimes \mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\psi_t}).$$

By Grothendieck's Lefschetz trace formula, we have

$$\sum_{s \text{ in } k^\times} \mathrm{Trace}(F_{s,k} | \mathscr{F} \otimes \mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\psi_t}) = \sum_i (-1)^i \mathrm{Trace}(F_k | H_c^i(\mathbb{G}_m \otimes \bar{k}, \mathscr{F} \otimes \mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\psi_t})).$$

Thus we obtain the cohomological formula for $f(t)$:

$$(-1)^{n-1} q f(t) / (\bar{\chi}(\alpha) g(\psi, \varepsilon)) = \sum_i (-1)^i \mathrm{Trace}(F_k | H_c^i(\mathbb{G}_m \otimes \bar{k}, \mathscr{F} \otimes \mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\psi_t})).$$

We now turn to a detailed discussion of the weights and dimensions of the cohomology groups $H_c^i(\mathbb{G}_m \otimes \bar{k}, \mathscr{F} \otimes \mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\psi_t})$. Because the sheaf $\mathscr{F} \otimes \mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\psi_t}$ on $\mathbb{G}_m \otimes \bar{k}$ is lisse, only the terms with $i = 1$ or $i = 2$ are possibly nonzero. The sheaf $\mathscr{F}$ is lisse of rank n, pure of weight $n-1$, and tame at 0 [Ka-GKM], 8.8.7 and 4.1.1. If we write $n = n_0 p^e$ with n_0 prime to p, then all the ∞-slopes of $\mathscr{F}$ are n_0/n [Ka-GKM], 4.1.1 (3), 1.11 (1) and 1.13.1. The sheaves $\mathscr{L}_{\varrho\varepsilon}$ and $\mathscr{L}_{\psi_t}$ are lisse, and pure of weight zero.

By Deligne's fundamental result [De-Weil II], 3.3.1,

$H_c^2(\mathbb{G}_m \otimes \bar{k}, \mathscr{F} \otimes \mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\psi_t})$ is pure of weight $(n-1)+2 = n+1$,

$H_c^1(\mathbb{G}_m \otimes \bar{k}, \mathscr{F} \otimes \mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\psi_t})$ is mixed of weight $\leqq (n-1)+1 = n$.

In discussing the dimensions of these groups, it is convenient to consider first the case when the characteristic p divides n, i.e., when $n_0 < n$. In this case, all the ∞-slopes of $\mathscr{F}$ are $n_0/n < 1$. Because $\mathscr{L}_{\psi_t}$ has its ∞-slope $=1$ for $t \neq 0$, $=0$ for $t = 0$, the sheaf

$$\mathscr{F} \otimes \mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\psi_t}$$

has all its ∞-slopes equal to 1 for $t \neq 0$, and equal to n_0/n for $t = 0$. So for any t in $\bar{k}$, $\mathscr{F} \otimes \mathscr{L}_{\varrho z} \otimes \mathscr{L}_{\psi_t}$ is totally wild at ∞, and hence its H_c^2 vanishes (cf. [Ka-GKM], 2.1.1). Since $\mathscr{F} \otimes \mathscr{L}_{\varrho z} \otimes \mathscr{L}_{\psi_t}$ is lisse on $G_m \otimes \bar{k}$, and tame at zero, the Euler-Poincaré formula gives

$$\begin{aligned}\chi_c(G_m \otimes \bar{k}, \mathscr{F} \otimes \mathscr{L}_{\varrho z} \otimes \mathscr{L}_{\psi_t}) &= -n_0 \text{ if } t = 0 \text{ and } p \mid n, \\ &= -n \text{ if } t \neq 0 \text{ and } p \mid n.\end{aligned}$$

So in the case when p divides n, we have $H_c^2 = 0$, and $\dim H_c^1 \leqq n$. From the cohomological formula for $f(t)$, we get

$$|(-1)^{n-1} q f(t)/(\chi(\alpha) g(\psi, \varepsilon))| \leqq n(q)^{n/2}, \quad \text{if } p \mid n.$$

Since $|g(\psi, \varepsilon)|$ is 1 if ε is trivial, and $q^{1/2}$ if not, this gives Theorems 2bis and 3bis in the case when p divides n. It also gives the slight sharpening for $f(0)$:

$$|(-1)^{n-1} q f(0)/(\bar{\chi}(\alpha) g(\psi, \varepsilon))| \leqq n_0(q)^{n/2}, \quad \text{if } p \mid n.$$

We now turn to the case when p does not divide n.

Lemma 1. *Suppose that p does not divide n. Denote by Y the set*

$$Y := \{y \text{ in } \bar{k} \text{ such that } y^n = n^n N\alpha\}.$$

As P_∞-representation, $\mathscr{F}$ is isomorphic to the direct sum

$$\mathscr{F} | P_\infty \cong \bigoplus_{y \text{ in } Y} \mathscr{L}_{\psi_y} | P_\infty.$$

Proof. The question is geometric, so we may replace $\mathscr{F}$ by the sheaf

$$[s \mapsto s^n N\alpha]^* \mathrm{Kl}(\psi \circ \mathrm{Tr}, \chi, \chi^q, \chi^{q^2}, \ldots, \chi^{q^{n-1}}),$$

to which it is geometrically isomorphic. Moreover, the P_∞-representation of

$$\mathrm{Kl}(\psi \circ \mathrm{Tr}, \chi, \chi^q, \chi^{q^2}, \ldots, \chi^{q^{n-1}})$$

does not depend on the particular characters which occur [Ka-GKM], 10.1, so it is the same as that of $\mathrm{Kl}(\psi \circ \mathrm{Tr}, \mathbb{1}, \mathbb{1}, \mathbb{1}, \ldots, \mathbb{1})$, the sheaf on $G_m \otimes \bar{k}$ denoted $\mathrm{Kl}_n(\psi)$ in [Ka-GKM], 11.0.1. Thus

$$\mathscr{F} | P_\infty \cong [s \mapsto s^n N\alpha]^* \mathrm{Kl}_n(\psi).$$

Now pick an element z in $\bar{k}^\times$ with $z^n = N\alpha$. Then $s^n N\alpha = (sz)^n$, so

$$[s \mapsto s^n N\alpha]^* \mathrm{Kl}_n(\psi) = [s \mapsto zs]^* [s \mapsto s^n]^* \mathrm{Kl}_n(\psi).$$

It is known [Ka-GKM], 10.4.5, that as P_∞-representations, we have

$$[s \mapsto s^n]^* \mathrm{Kl}_n(\psi) | P_\infty \cong \bigoplus_{\zeta^n = 1} \mathscr{L}_{\psi_{\zeta/n}} | P_\infty.$$

Pulling back by $s \mapsto zs$, we find

$$\mathscr{F} | P_\infty \cong \bigoplus_{\zeta^n = 1} \mathscr{L}_{\psi_{\zeta z/n}} | P_\infty . \quad \text{QED}$$

Lemma 2. *Suppose that p does not divide n. Then the ∞-breaks of the sheaf $\mathscr{F} \otimes \mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\varphi_t}$ are all equal to 1, unless $t^n = n^n N\alpha$, in which case $n-1$ of the ∞-breaks are equal to 1, and the remaining one is equal to 0.*

Proof. This is immediate from Lemma 1. QED

Corollary. *Suppose that p does not divide n. If $t^n \neq n^n N\alpha$, then*

$$H^2_c(\mathbb{G}_m \otimes \bar{k}, \mathscr{F} \otimes \mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\varphi_t}) = 0 ,$$

$$\dim H^1_c(\mathbb{G}_m \otimes \bar{k}, \mathscr{F} \otimes \mathscr{L}_{\varrho\varepsilon} \otimes_{\varphi_t}) = n .$$

If $t^n = n^n N\alpha$, then

$$\chi_c(\mathbb{G}_m \otimes \bar{k}, \mathscr{F} \otimes \mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\varphi_t}) = 1 - n ,$$

$$\dim H^2_c(\mathbb{G}_m \otimes \bar{k}, \mathscr{F} \otimes \mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\varphi_t}) = 0 \text{ or } 1 ,$$

$$\dim H^1_c(\mathbb{G}_m \otimes \bar{k}, \mathscr{F} \otimes \mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\varphi_t}) = n - 1 \text{ or } n .$$

Proof. The sheaf $\mathscr{F} \otimes \mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\varphi_t}$ is lisse on $\mathbb{G}_m \otimes \bar{k}$ and tame at zero, so by the Euler-Poincaré formula, its Euler characteristic is minus the sum of the ∞-slopes. If all ∞-slopes are > 0, the H^2_c vanishes. In general, the dimension of H^2_c is bounded by the number of ∞-slopes which are zero (i.e., the dimension of the $\pi_1(\mathbb{G}_m \otimes \bar{k})$-coinvariants is at most the dimension of the P_∞-coinvariants). QED

To conclude the proof of Theorems 2 bis and 3 bis, it remains to analyse more closely the question of when H^2_c is nonzero. We know this can only happen if both p does not divide n, and $t^n = n^n N\alpha$. We also know that if it does happen, then the H^2_c is one-dimensional.

Lemma 3. *Suppose p does not divide n, and $t^n = n^n N\alpha$. Then*

$$H^2_c(\mathbb{G}_m \otimes \bar{k}, \mathscr{F} \otimes \mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\varphi_t}) \neq 0$$

if and only if the following two conditions hold:

1) *χ^{q-1} has (exact) order n, and the characters $\chi^{q^i - 1}$ for $0 \leq i \leq n-1$ are all distinct.*

2) *Either n is even and ε has exact order 2, or n is odd and ε is trivial.*

Proof. The sheaf $\mathscr{F} \otimes \mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\varphi_t}$ is geometrically semisimple, by [Ka-GKM], 4.1.2. So $H^2_c(\mathbb{G}_m \otimes \bar{k}, \mathscr{F} \otimes \mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\varphi_t})$ is, up to a Tate twist, the space

$$\operatorname{Hom}_{\mathbb{G}_m \otimes \bar{k}}(\mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\psi_t}, \mathscr{F}) .$$

The question being geometric, we may replace $\mathscr{F}$ by

$$[s \mapsto s^n N\alpha]^* \mathrm{Kl}(\psi \circ \mathrm{Tr}, \chi, \chi^q, \chi^{q^2}, \ldots, \chi^{q^{n-1}}).$$

Now pick an element z in $\bar{k}^\times$ with $z^n = N\alpha$. Then $s^n N\alpha = (sz)^n$, so

$$\begin{aligned}&[s \mapsto s^n N\alpha]^* \mathrm{Kl}(\psi \circ \mathrm{Tr}, \chi, \chi^q, \chi^{q^2}, \ldots, \chi^{q^{n-1}})\\ &\cong [s \mapsto zs]^* [s \mapsto s^n]^* \mathrm{Kl}(\psi \circ \mathrm{Tr}, \chi, \chi^q, \chi^{q^2}, \ldots, \chi^{q^{n-1}}).\end{aligned}$$

Via the automorphism $s \to s/z$, we have

$$\begin{aligned}&\mathrm{Hom}_{G_m \otimes \bar{k}}(\mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\psi_t}, \mathscr{F})\\ &\cong \mathrm{Hom}_{G_m \otimes \bar{k}}([s \mapsto s/z]^*(\mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\psi_t}), [s \mapsto s/z]^* \mathscr{F})\\ &= \mathrm{Hom}_{G_m \otimes \bar{k}}(\mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\psi_{t/z}}, [s \mapsto s/z]^* \mathscr{F})\\ &= \mathrm{Hom}_{G_m \otimes \bar{k}}(\mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\psi_{t/z}}, [n]^* \mathrm{Kl}(\psi \circ \mathrm{Tr}, \chi, \chi^q, \chi^{q^2}, \ldots, \chi^{q^{n-1}})),\end{aligned}$$

where we have written $[n]$ for the map $s \mapsto s^n$.

By Frobenius reciprocity for the finite etale map $[n]$ of $G_m \otimes \bar{k}$, this last Hom group is

$$= \mathrm{Hom}_{G_m \otimes \bar{k}}([n]_*(\mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\psi_{t/z}}), \mathrm{Kl}(\psi \circ \mathrm{Tr}, \chi, \chi^q, \chi^{q^2}, \ldots, \chi^{q^{n-1}})).$$

But both of the sheaves inside this last Hom are Kloosterman sheaves [Ka-GKM], 5.6.2, so both are geometrically irreducible. Therefore this Hom group vanishes unless there exists an isomorphism of lisse sheaves on $G_m \otimes \bar{k}$

$$[n]_*(\mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\psi_{t/z}}) \cong \mathrm{Kl}(\psi \circ \mathrm{Tr}, \chi, \chi^q, \chi^{q^2}, \ldots, \chi^{q^{n-1}}).$$

If such an isomorphism exists, the Hom group is one-dimensional.

Over $G_m \otimes \bar{k}$ we know from [Ka-GKM], 5.6.2, that

$$[n]_*(\mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\psi_{t/z}}) \cong \mathrm{Kl}(\psi_{t/nz}, \text{all the } n\text{'th roots of } \varrho\varepsilon).$$

Since $t^n = n^n N\alpha$, and z was *any* element in $\bar{k}^\times$ with $z^n = N\alpha$, we may choose $z = t/n$. For this choice, we find

$$[n]_*(\mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\psi_{t/z}}) \cong \mathrm{Kl}(\psi, \text{all the } n\text{'th roots of } \varrho\varepsilon).$$

Thus our H_c^2 is nonzero if and only if, on $G_m \otimes \bar{k}$, we have an isomorphism of Kloosterman sheaves

$$\mathrm{Kl}(\psi, \text{all the } n\text{'th roots of } \varrho\varepsilon) \cong \mathrm{Kl}(\psi \circ \mathrm{Tr}, \chi, \chi^q, \chi^{q^2}, \ldots, \chi^{q^{n-1}}).$$

But two such Kloosterman sheaves are isomorphic if and only if they have the *same* sets of characters with multiplicities (all characters viewed as characters of some sufficiently large common finite extension of both K and k, by composition with the norm) [Ka-GKM], 4.1.1 (1)–(2) and 4.1.2 (2a) for "if", 7.4.1 for "only if". So we are reduced to asking when

$$\{\text{all the } n\text{'th roots of } \varrho\varepsilon\} = \{\chi, \chi^q, \chi^{q^2}, \ldots, \chi^{q^{n-1}}\}.$$

To analyze this question, notice first that in order for the set

$$\{\chi, \chi^q, \chi^{q^2}, \ldots, \chi^{q^{n-1}}\}$$

to be all the n'th roots of any character, it is necessary for the set of ratios to χ,

$$\{\mathbb{1}, \chi^{q-1}, \chi^{q^2-1}, \ldots, \chi^{q^{n-1}-1}\},$$

to consist of the n distinct characters of order dividing n. Since the characters χ^{q^i-1} are all powers of χ^{q-1}, this holds if and only if χ^{q-1} is a character of exact order n, and the n characters

$$\{\mathbb{1}, \chi^{q-1}, \chi^{q^2-1}, \ldots, \chi^{q^{n-1}-1}\}$$

are all distinct. If this condition holds, then $\{\chi, \chi^q, \chi^{q^2}, \ldots, \chi^{q^{n-1}}\}$ is the set of all n-th roots of the character χ^n. So in order to have

$$\mathrm{Kl}(\psi, \text{all the } n\text{'th roots of } \varrho\varepsilon) \cong \mathrm{Kl}(\psi \circ \mathrm{Tr}, \chi, \chi^q, \chi^{q^2}, \ldots, \chi^{q^{n-1}}),$$

we must have in addition

$$\chi^n = \varrho\varepsilon.$$

Since $\varrho \circ \mathbb{N} = \chi^{1+q+q^2+\cdots+q^{n-1}}$, this condition may be rewritten

$$\chi^n = (\varepsilon \circ \mathbb{N})\,\chi^{1+q+q^2+\cdots+q^{n-1}} \text{ as characters of } K^\times,$$

or equivalently, as

$$\begin{aligned}\bar{\varepsilon} \circ \mathbb{N} &= \mathbb{1}(\chi^{q-1})(\chi^{q^2-1})\ldots(\chi^{q^{n-1}-1})\\ &= \text{the product of all the characters of order dividing } n\\ &= \mathbb{1} \text{ if } n \text{ is odd, the quadratic character is } n \text{ is even.} \quad \text{QED}\end{aligned}$$

Combining this last lemma with the preceding corollary, we find that Theorems 2bis and 3bis are proven.

Remarks on Theorems 2bis and 3bis. 1) If p divides n, we have already seen that for $t = 0$ the constant in the bound may be taken to be n_0, the "prime to p part of n", rather than n itself.

2) Suppose p does not divide n, and we are not in any of the exceptional cases of Theorems 2bis or 3bis. By Lemma 3, we have

$$H_c^2(G_m \otimes \bar{k}, \mathscr{F} \otimes \mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\varphi_t}) = 0.$$

If t satisfies $t^n = n^n N\alpha$, we have seen in Lemma 2 that

$$\chi_c(\mathbb{G}_m \otimes \bar{k}, \mathscr{F} \otimes \mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\psi_t}) = 1 - n\,,$$

whence

$$\dim H_c^1(\mathbb{G}_m \otimes \bar{k}, \mathscr{F} \otimes \mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\psi_t}) = n - 1$$

in this case. So if we are not in an exceptional case, for t satisfying $t^n = n^n N\alpha$, the constant in the bound may be taken to be $n-1$ rather than n.

3) Suppose again that p does not divide n. If $\chi^n \neq \varepsilon\varrho$, the sheaf $\mathscr{F} \otimes \mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\psi_t}$ has no nonzero invariants under the local monodromy group I_0. (More generally, the semi-simplification of this sheaf as I_0-representation is the direct sum of the n tame characters $\chi^{nq^i}/\varepsilon\varrho$, for $0 \leq i \leq n-1$. As both ε and ϱ are characters of $k^\times$, $(\varepsilon\varrho)^q = \varepsilon\varrho$, so these are the characters $(\chi/\varepsilon\varrho)^{q^i}$, which all have the same order, being K/k-conjugates.) Thus if both $\chi^n \neq \varepsilon\varrho$ and $t^n \neq n^n N\alpha$, then $\mathscr{F} \otimes \mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\psi_t}$ has no nonzero inertial invariants at either 0 or ∞ (it has all ∞-slopes 1), and so by [De-Weil II], 3.2.3, $H_c^1(\mathbb{G}_m \otimes \bar{k}, \mathscr{F} \otimes \mathscr{L}_{\varrho\varepsilon} \otimes \mathscr{L}_{\psi_t})$ is *pure* of weight n.

4) One might ask about the compact cohomology groups which tautologically give rise to the Soto-Andrade sums. In the notations of our earlier discussion of Weil restriction, these are the groups

$$H_c^i(U[1/(t - \mathrm{Tr}_\alpha)] \otimes \bar{k}, \mathscr{L}_{\varepsilon(t-\mathrm{Tr}_\alpha)} \otimes \mathscr{L}_\chi)\,.$$

In terms of the inclusion $j: \mathbb{G}_m \to \mathbb{A}^1$, we may rewrite these groups as

$$H_c^i(U \otimes \bar{k}, \mathscr{L}_\chi \otimes [x \mapsto t - \mathrm{Tr}(\alpha x)]^*(j_!\mathscr{L}_\varepsilon))\,.$$

They are the stalks at t of the chomology sheaves of the complex $R(\mathrm{pr}_2)_!\mathscr{G}$ for the second projection

$$\mathrm{pr}_2 : U \times \mathbb{A}^1 \to \mathbb{A}^1\,,$$

and the sheaf on the source

$$\mathscr{G} := \mathrm{pr}_1^* \mathscr{L}_\chi \otimes [(x, t) \mapsto t - \mathrm{Tr}(\alpha x)]^*(j_!\mathscr{L}_\varepsilon)\,.$$

One can compute, in the derived category, the object $\mathrm{FT}_\psi(R(\mathrm{pr}_2)_!\mathscr{G})$, cf. [Ill-DFT]. By definition, it is the object defined as $R(\mathrm{pr}_3)_!L$ for the third projection

$$\mathrm{pr}_3 : U \times \mathbb{A}^1 \times \mathbb{A}^1 \to \mathbb{A}^1\,,$$

and the sheaf on the source

$$L := \mathrm{pr}_1^* \mathscr{L}_\chi \otimes [(x, t) \mapsto t - \mathrm{Tr}(\alpha x)]^*(j_!\mathscr{L}_\varepsilon) \otimes [(t, s) \mapsto st]^* \mathscr{L}_\psi\,.$$

Applying the shearing automorphism

$$(x, t, s) \mapsto (x, t + \mathrm{Tr}(\alpha x), s)$$

of $U \times \mathbb{A}^1 \times \mathbb{A}^1$ as $\mathbb{A}^1$-scheme by pr_3, L becomes

$$\mathrm{pr}_1^* \mathscr{L}_\chi \otimes \mathrm{pr}_2^*(j_! \mathscr{L}_\varepsilon) \otimes \mathscr{L}_{\psi(st)} \otimes \mathscr{L}_{\psi(s\mathrm{Tr}(\alpha x))}\,.$$

At $s = 0$, proper base change gives

$$\begin{aligned}
(R(\mathrm{pr}_3)_! L)_0 &= R\Gamma_c((U \times \mathbb{A}^1) \otimes \bar{k}, \mathrm{pr}_1^* \mathscr{L}_\chi \otimes \mathrm{pr}_2^*(j_! \mathscr{L}_\varepsilon)) \\
\text{(Kunneth)} \qquad &= R\Gamma_c(U \otimes \bar{k}, \mathscr{L}_\chi) \otimes R\Gamma_c((\mathbb{A}^1) \otimes \bar{k}, j_! \mathscr{L}_\varepsilon) \\
&= R\Gamma_c(U \otimes \bar{k}, \mathscr{L}_\chi) \otimes R\Gamma_c(\mathbb{G}_m \otimes \bar{k}, \mathscr{L}_\varepsilon) \\
&= 0, \text{ if either } \chi^{q-1} \text{ or } \varepsilon \text{ is nontrivial}.
\end{aligned}$$

Over the open set where $s \neq 0$, we apply the further isomorphism $(x, t, s) \to (xs\alpha, t, s)$ of $U \times \mathbb{A}^1 \times \mathbb{G}_m$ with the subgroup of

$$K^\times \times \mathbb{A}^1 \times \mathbb{G}_m \text{ of } (x, t, s) \text{ where } \mathbb{N}x = s^n \mathbb{N}\alpha\,.$$

For any β in $\bar{\mathbb{Q}}_\ell$, we denote by $\beta^{\deg}$ the geometrically constant lisse $\bar{\mathbb{Q}}_\ell$-sheaf of rank one on $\mathbb{G}_{m,k}$ on which, for any finite extension E of k and for any point s in $E^\times = \mathbb{G}_{m,k}(E)$, $F_{s,E}$ acts as $\beta^{\deg(E/k)}$: equivalenty, $\beta^{\deg}$ is the pullback to $\mathbb{G}_{m,k}$ of the lisse rank one $\bar{\mathbb{Q}}_\ell$-sheaf on $\mathrm{Spec}(k)$ on which F_k acts as β. By the relative Kunneth formula for pr_3, and [Ka-GKM], 8.8.5, we find

$$R(\mathrm{pr}_3)_! L \,|\, \mathbb{G}_m \cong \mathscr{F}[1-n] \otimes \mathscr{L}_\varrho \otimes \mathscr{L}_\varepsilon[-1] \otimes g(\psi, \varepsilon)^{\deg}\,.$$

Thus all in all we find

$$\mathrm{FT}_\psi(R(\mathrm{pr}_2)_! \mathscr{G}) = j_!(\mathscr{F}[1-n] \otimes \mathscr{L}_\varrho \otimes \mathscr{L}_\varepsilon[-1]) \otimes g(\psi, \varepsilon)^{\deg}$$

is a single lisse sheaf on $\mathbb{G}_m$, placed in degree n, extended by zero to $\mathbb{A}^1$.

By Fourier inversion, we get

$$\begin{aligned}
R(\mathrm{pr}_2)_! \mathscr{G} &= \mathrm{FT}_{\bar\psi} \mathrm{FT}_\psi R(\mathrm{pr}_2)_! \mathscr{G}[2](1) \\
&= \mathrm{FT}_{\bar\psi}(j_!(\mathscr{F} \otimes \mathscr{L}_\varrho \otimes \mathscr{L}_\varepsilon))[2-n] \otimes (g(\psi, \varepsilon)/q)^{\deg}\,.
\end{aligned}$$

Thus we find that for t in $\bar{k}$, we have

$$\begin{aligned}
&H_c^i(U[1/(t - \mathrm{Tr}_\alpha)] \otimes \bar{k}, \mathscr{L}_{\varepsilon(t - \mathrm{Tr}_\alpha)} \otimes \mathscr{L}_\chi) \\
&= 0 \quad \text{if } i \neq n-1 \text{ or } n, \\
&= H_c^1(\mathbb{G}_m \otimes \bar{k}, \mathscr{F} \otimes \mathscr{L}_\varrho \otimes \mathscr{L}_\varepsilon) \otimes (g(\psi, \varepsilon)/q)^{\deg} \quad \text{for } i = n-1, \\
&= H_c^2(\mathbb{G}_m \otimes \bar{k}, \mathscr{F} \otimes \mathscr{L}_\varrho \otimes \mathscr{L}_\varepsilon) \otimes (g(\psi, \varepsilon)/q)^{\deg} \quad \text{for } i = n.
\end{aligned}$$

So we have been "speaking prose" all along.

Proof of Theorem 4

Recall that α in $K^\times$, b in $k^\times$, t in k, and

$$f(t) := \mathrm{Card}\{x \text{ in } K^\times \text{ with } \mathbb{N}x = b \text{ and } \mathrm{Tr}(\alpha x) = t\}.$$

We must show that

$$|qf(t) - (q^n - 1)/(q-1)| \leqq n(q)^{n/2}.$$

The idea is once again to compute the Fourier transform $\mathrm{FT}(f)$ of the function f. We compute, for s in k,

$$\mathrm{FT}(f)(s) = \sum_{t \text{ in } k} \psi(st) f(t) = \sum_{x \text{ in } K^\times,\ \mathbb{N}x = b} \psi(s\mathrm{Tr}(\alpha x)).$$

Since the norm map $\mathbb{N}: K^\times \to k^\times$ is surjective, for $s = 0$ we find

$$\mathrm{FT}(f)(0) = \mathrm{Card}\{x \text{ in } K^\times, \text{ with } \mathbb{N}x = b\} = (q^n - 1)/(q-1).$$

For $s \neq 0$, we make the substitution $y := s\alpha x$ to rewrite

$$\mathrm{FT}(f)(s) = \sum_{x \text{ in } K^\times,\ \mathbb{N}x = bs^n \mathbb{N}\alpha} \psi(\mathrm{Tr}(x)), \quad \text{for } s \neq 0.$$

In terms of the exotic Kloosterman sheaf $\mathrm{Kl}(K, \psi \circ \mathrm{Tr}, \mathbb{1})$ on $\mathbb{G}_m \otimes k$ and its pullback

$$\mathscr{F} = [s \mapsto s^n b \mathbb{N}\alpha]^* \mathrm{Kl}(K, \psi \circ \mathrm{Tr}, \mathbb{1}),$$

we have

$$\mathrm{FT}(f)(s) = (-1)^{n-1} \mathrm{Trace}(F_{s,k} | \mathscr{F}) \quad \text{for } s \neq 0.$$

So by Fourier inversion, we have

$$\begin{aligned} qf(t) &= \sum_{s \text{ in } k} \bar{\psi}(st) \mathrm{FT}(f)(s) \\ &= \mathrm{FT}(f)(0) + (-1)^{n-1} \sum_{s \text{ in } k^\times} \mathrm{Trace}(F_{s,k} | \mathscr{F} \otimes \mathscr{L}_{\bar{\psi}_t}). \end{aligned}$$

Thus we find

$$|qf(t) - (q^n - 1)/(q-1)| = \Big| \sum_{s \text{ in } k^\times} \mathrm{Trace}(F_{s,k} | \mathscr{F} \otimes \mathscr{L}_{\bar{\psi}_t}) \Big|.$$

By the Lefschetz trace formula,

$$\begin{aligned} &\sum_{s \text{ in } k^\times} \mathrm{Trace}(F_{s,k} | \mathscr{F} \otimes \mathscr{L}_{\bar{\psi}_t}) \\ &= \sum_{i=1,2} (-1)^i \mathrm{Trace}(F_k | H^i_c(\mathbb{G}_m \otimes \bar{k}, \mathscr{F} \otimes \mathscr{L}_{\bar{\psi}_t})). \end{aligned}$$

Geometrically, the sheaf $\mathrm{Kl}(K, \psi \circ \mathrm{Tr}, \mathbb{1})$ is isomorphic to the standard Kloosterman sheaf $\mathrm{Kl}_n(\psi) := \mathrm{Kl}(\psi; \mathbb{1}, \ldots, \mathbb{1})$. This sheaf is Lie-irreducible (by [Ka-GKM], 11.1) of rank n. Therefore its pullback by a finite map, $\mathscr{F}$, is geometrically irreducible, and hence $\mathscr{F} \otimes \mathscr{L}_{\psi_t}$ is geometrically irreducible, of rank $n \geqq 2$. Therefore its H_c^2 vanishes, and we find

$$|qf(t) - (q^n - 1)/(q - 1)| = |\mathrm{Trace}(F_k | H_c^1(\mathbb{G}_m \otimes \bar{k}, \mathscr{F} \otimes \mathscr{L}_{\psi_t}))|.$$

Now $\mathscr{F}$ is lisse on $\mathbb{G}_m$, pure of weight $n-1$, tame at zero, with all ∞-slopes $\leqq 1$. So the H_c^1 is mixed of weight $\leqq n$. As H_c^2 vanishes,

$$\dim H_c^1 = -\chi_c = \mathrm{Swan}_\infty(\mathscr{F} \otimes \mathscr{L}_{\psi_t}) \leqq n.$$

Thus we find

$$|qf(t) - (q^n - 1)/(q - 1)| \leqq n(q)^{n/2}, \quad \text{as required.} \quad \text{QED}$$

References

[De-AFT] *P. Deligne*, Application de la formule des traces aux sommes trigonométriques, in: Cohomologie Etale (SGA 4 1/2), Lect. Notes Math. **569** (1977), 168–232.

[De-RFT] *P. Deligne*, Rapport sur la formule des traces, in: Cohomologie Etale (SGA 4 1/2), Lect. Notes Math. **569** (1977), 76–109.

[De-Weil II] *P. Deligne*, La conjecture de Weil II, Publ. Math. I.H.E.S. **52** (1981), 313–428.

[Gr-FL] *A. Grothendieck*, Formule de Lefschetz et rationalité des fonctions *L*, Sém. Bourbaki 1964–65, Exposé 279, reprinted in: Dix Exposés sur la cohomologie des schémas, Amsterdam 1968.

[SGA] *A. Grothendieck* et al., Séminaire de Géométrie Algébrique du Bois-Marie, SGA 1, SGA 4 Parts I, II, and III, SGA 4 1/2, SGA 5, SGA 7 Parts I and II, Lect. Notes Math. **224** (1971) to (1977), 269–270–305, 569, 589, 288–340.

[Ill-DFT] *L. Illusie*, Deligne's ℓ-adic fourier transform, in: Algebraic Geometry: Bowdoin 1985, A.M.S., Providence 1987.

[Ka-GKM] *N. Katz*, Gauss sums, Kloosterman sums, and monodromy groups, Ann. Math. Study **116**, Princeton 1988.

[Ray] *M. Raynaud*, Caractéristique d'Euler-Poincaré d'un faisceau et cohomologie des variétés abéliennes, Sém. Bourbaki 1964/65, Exposé 286, New York 1966.

[S-A] *J. Soto-Andrade*, Geometrical Gelfand models, tensor quotients, and Weil representations, Proc. Symp. Pure Math. **47** (1987), 305–316.

[Ter] *A. Terras*, Fourier Analysis on Finite Groups and Applications, Chapter 2, The finite affine and Heisenberg groups and other non-abelian finite groups, preprint dated 4/6/92.

[We] *A. Weil*, On the Riemann Hypothesis in function-fields, Proc. Nat. Acad. Sci. U.S.A. **27** (1941), 345–347.

Department of Mathematics, Fine Hall, Princeton University, Princeton NJ 08544, USA

Eingegangen 17. Juni 1992