Electronic Colloquium on Computational Complexity, Revision 1 of Report No. 69 (2024)

Trace Hermitian codes have vanishing bias

Swastik Kopparty[*] Amnon Ta-Shma[†] Kedem Yakirevitch[‡]

November 10, 2025

Abstract

In this work we give the first proof that Trace Hermitian codes have vanishing bias. This, and following work, bring to the front the question of the true behavior of Trace AG codes, and the fascinating possibility that they might match, perhaps even surpass, the GV bound.

1 Introduction

Error–correcting codes (ECC) are mathematical tools used to detect and correct errors in data transmission and storage. These codes add redundancy to the original message so that even when part of the data is corrupted (because of channel noise, hardware faults, etc.), the original message can still be recovered. The design of efficient ECC lies at the heart of information theory and has profound applications in communication systems, data storage, cryptography, and theoretical computer science.

A central trade-off in the study of ECC is the rate–distance trade-off. The *rate* of an $[n, k]_q$ code is $r = k/n$, the fraction of symbols that carry information, while the *relative distance* δ equals the minimum Hamming distance between distinct codewords divided by n and governs the code's error resilience. Increasing the rate typically decreases the distance and vice versa, and much of coding theory is concerned with the best possible trade-offs between r and δ.

[*] Department of Mathematics and Department of Computer Science, University of Toronto. Email: swastik.kopparty@utoronto.ca.

[†] Department of Computer Science, Tel Aviv University. The research leading to these results has received funding from the Israel Science Foundation (grant number 443/22). Email: amnon@tauex.tau.ac.il.

[‡] Department of Computer Science, Tel Aviv University. The research leading to these results has received funding from the Israel Science Foundation (grant number 443/22) and the Blavatnik fund. Email: kedem12@gmail.com

ISSN 1433-8092

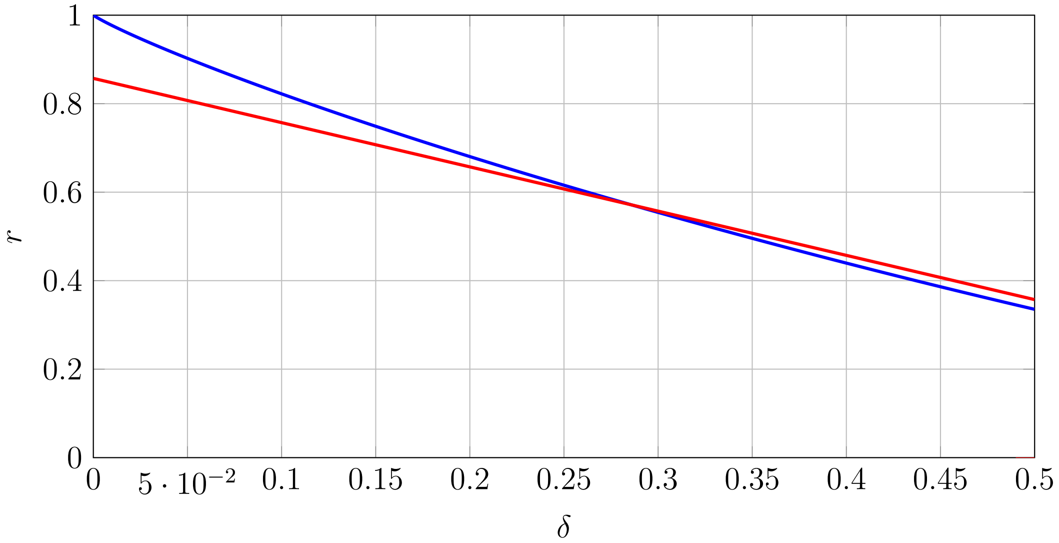

Figure 1: Rate vs distance bounds for $q = 64$. The blue curve is the Gilbert–Varshamov lower bound $r \geq 1 - H_{64}(\delta)$. The red line is the Algebraic–Geometry (AG) lower bound $r \geq 1 - \delta - \frac{1}{\sqrt{q}-1}$ and it is better than the GV bound for δ large enough.

The Gilbert–Varshamov (GV) bound [Gil52, Var57] shows non-constructively that for alphabet size q and relative distance δ there exist codes with

$$r \;\geq\; 1 - H_q(\delta),$$

where H_q is the q-ary entropy function. The GV bound guarantees the existence of asymptotically good codes but is non-constructive. A major breakthrough came in the 1980s with Algebraic–Geometry (AG) codes (Tsfasman, Vlăduţ, and Zink [TVZ82], building on Goppa), which are explicit constructions from algebraic curves over finite fields. AG codes were shown to satisfy

$$r \;\geq\; 1 - \delta - \frac{1}{\sqrt{q}-1},$$

and — remarkably — they beat the GV bound for square alphabets of size $q \geq 49$ by exploiting curves with many rational points relative to their genus. For illustration, the GV bound and the AG codes bound for $q = 64$ are shown in Figure 1.

Upper bounds such as the McEliece–Rodemich–Rumsey–Welch (MRRW) bounds place limits on achievable rates; the MRRW bound is the strongest known asymptotic upper bound. The binary case ($q = 2$) is especially important. The best known lower bound is again the nonconstructive GV bound, while the best upper bound is the MRRW bound. Although AG codes beat GV for sufficiently large alphabets, it remains open whether the GV bound is tight for smaller alphabets and, in particular, for binary codes. Several explicit constructions exist for binary alphabets, but they are generally far from the GV bound; for

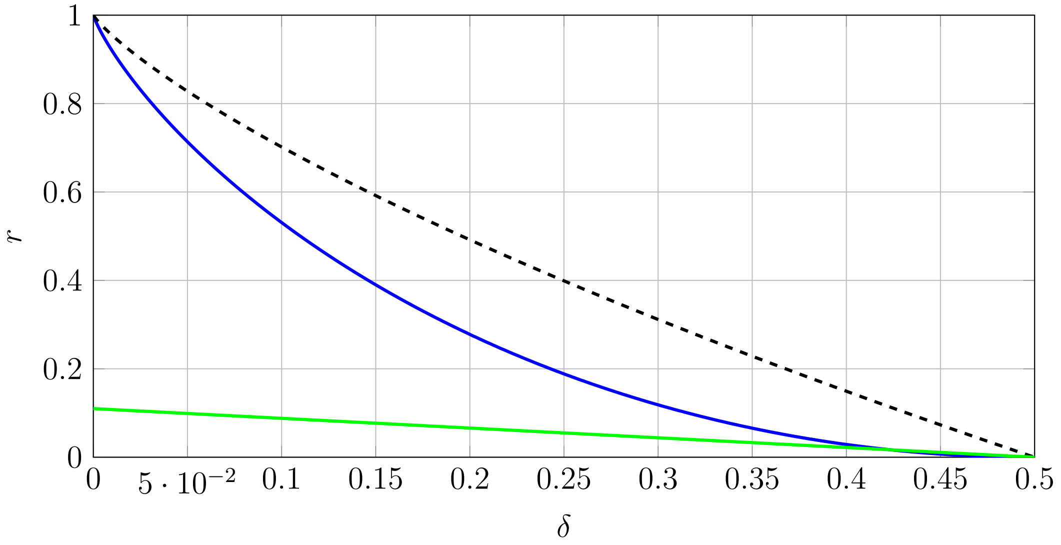

Figure 2: Binary bounds: The green line marks an approximation to the *explicit* $[n, \frac{R}{2}n, (1-R-\varepsilon)H_2^{-1}(\frac{1}{2}-\varepsilon)n]_2$ Justensen code. The blue line is the GV lower bound $r \geq 1 - H_2(\delta)$, where non-explicit codes are known, the dashed line is the Elias/Bassalygo upper bound $r <= 1 - H_2(\frac{1-\sqrt{1-2\delta}}{2})$ where it is known that no codes exist above it. The MRRW bound is a better upper bound, but it is more complicated to draw and we omit it.

example, the Justesen code (with RS as the outer code, the Wonzencraft ensemble as the family of inner codes) is shown in Figure 2. [1]

A particularly interesting regime is binary codes with very large relative distance,

$$\delta = \frac{1-\varepsilon}{2} \quad \text{for small } \varepsilon > 0.$$

The GV bound guarantees binary codes with such distance and rate $r = \Omega(\varepsilon^2)$, while the MRRW bound shows that any binary code with $\delta = (1-\varepsilon)/2$ must satisfy

$$r = O\Big(\varepsilon^2 \log \frac{1}{\varepsilon}\Big).$$

Many explicit binary constructions start from a good code over a larger alphabet and reduce to binary by concatenation. Two common approaches are:

- Start with an AG code over $\mathbb{F}_q$ (with $q > 2$) and reduce the alphabet by concatenation; a natural inner code is the Hadamard code. Using AG codes obtained from curves above the genus threshold yields
$$r = \widetilde{\Omega}(\varepsilon^3),$$

[1] For every $\varepsilon > 0$, the Justensen code is $[n, \frac{R}{2}n, (1-R-\varepsilon)H^{-1}(\frac{1}{2}-\varepsilon)n]_2$ and the distance is about $0.11(1-R)$ for small ε.

where the $\widetilde{\Omega}$ hides polylogarithmic factors in $1/\varepsilon$. AG codes below the genus lead to other trade-offs but, as shown in [BATS13], but this approach cannot reach rates near the GV bound [BATS13].

- Alon et al. [ABN+92] proposed amplifying the distance of an asymptotically good binary code by a random walk on an expander (or any sampler), which increases the alphabet size; the alphabet is then reduced by concatenation with the Hadamard code. That construction achieves $r = \widetilde{\Omega}(\varepsilon^3)$ [ABN+92].

Concatenation appears to introduce an extra factor of ε in the rate in known analyses. A natural idea is to replace concatenation by a well-chosen deterministic binary function. This idea was explored in two directions. For the expander/amplification approach (not the focus of this paper), Alon and independently Rozenman and Wigderson (see [Bog12]) studied replacing concatenation with parity; a straightforward implementation gives the weaker bound $r = \Omega(\varepsilon^4)$, but combining parity with a wide zig-zag product leads to the best explicit construction so far of rate $r = \Omega(\varepsilon^{2+o(1)})$ [TS17].

This paper focuses on the first approach: replacing concatenation with a deterministic binary-valued function applied symbolwise to an outer AG evaluation code. An early related construction appears in [AGHP92]: it produces a binary code with distance $\delta = (1-\varepsilon)/2$, dimension k, and length $n = O\big((k/\varepsilon)^2\big)$. We review a variant of that construction because it motivates our work.

Assume $q = 2^\ell$ and let the outer code be the evaluation code on the vector space spanned by all *odd* monomials x^i with $i \le d$. To reduce the alphabet size we apply the trace map $\mathsf{Tr}_{\mathbb{F}_q/\mathbb{F}_2} : \mathbb{F}_q \to \mathbb{F}_2$ coordinate-wise to each outer symbol. The key technical tool we use is the Weil bound for additive character sums:

Theorem 1.1 (Weil bound, additive characters)**.** *Let $q = 2^\ell$, and let $f \in \mathbb{F}_q[x]$ be a polynomial of degree d that is not of the form $g(x)^2 - g(x)$. Then the additive character sum satisfies*

$$\Big|\sum_{x\in\mathbb{F}_q}(-1)^{\mathsf{Tr}(f(x))}\Big| \le (d-1)\sqrt{q},$$

so in particular the normalized bias satisfies

$$\Big|\frac{1}{q}\sum_{x\in\mathbb{F}_q}(-1)^{\mathsf{Tr}(f(x))}\Big| \le \frac{d-1}{\sqrt{q}}.$$

Applying this bound with $d = \deg(f)$ and

$$\varepsilon \approx \frac{d}{\sqrt{q}}$$

gives a binary $[q, k = \Omega(d),\ \delta = \frac{1-\epsilon}{2}q]_2$ code, and the block length satisfies $q = \Omega((\frac{k}{\varepsilon})^2)$. We remark that one of the constructions appearing in AGHP uses the quadratic multiplicative character (quadratic residue) together with an extra linearization trick to obtain linear codes; here we use the additive character (trace) to avoid the linearization trick.

The Weil bound is not known to be tight in all regimes — a potentially stronger bound of the form $O\big(\sqrt{\frac{d}{q}},\big)$ would immediately give $\varepsilon = \sqrt{\frac{d}{q}}$, i.e. $q = d/\varepsilon^2$, and hence rate

$$r = \Omega\Big(\frac{d}{q}\Big) = \Omega(\varepsilon^2),$$

matching the GV bound. This raises natural questions:

- What is the true behavior of this trace-based construction? Can it match the GV bound? How close can it get to MRRW?
- What happens when the outer code is an AG code which already beats GV on a larger alphabet — does taking the trace preserve that advantage when reducing to a smaller alphabet?

In this work we give the first proof that trace AG codes have vanishing bias. For concreteness we focus on Trace of the Hermitian code, and later on we survey what we know about other AG codes, and following work on the problem. Also, for simplicity, the following discussion concerns a multiplicative character of order 2 (namely, the quadratic residue function χ) rather than the trace map - but the results we obtain usually equally apply on additive and multiplicative characters.

The Hermitian code is an evaluation code with low-degree bi-variate polynomials as the vector space of evaluation functions and the Hermitian curve $H = \{(x,y) \in \mathbb{F}_q^p \mid \mathsf{Tr}(y) = N(x)\}$ where $q = p^2$, p is prime, Tr and N are the trace and norm functions from $\mathbb{F}_q$ to $\mathbb{F}_p$, namely, $\mathsf{Tr}(y) = y^p + y$ and $N(x) = x^{p+1}$. The Hermitian curve has a maximal number of rational points for its genus, and was extensively studied. The question we study is: given a non-square polynomial $p(x,y)$, how large $bias(p) = \frac{1}{|H|}\sum_{(a,b)\in H}\chi(p(a,b))$ can be?

The natural approach for attacking the problem is looking at the function field

$$\mathbb{F}_q(x,y,z) \bmod \{\mathsf{Tr}(y) - N(x), z^2 - p(x,y)\}$$

and bounding its number of rational points using the Hasse-Weil bound. However, this approach meets an immediate obtacle: the error term in the Hasse-Weil bound is in the order of $g\sqrt{q}$ where g is the genus of the curve, and as a result the Hasse-Weil bound is vacuous when the genus is larger than $\sqrt{q}$ - as is the case with the Hermitian curve.

In [Vla96], Vladut uses properties of zeta functions of curves and extremal properties of the Hermitian curve to prove that $bias(p)$ is smaller than some constant smaller than 1.[2] This was the first non-trivial result on the problem, and the best prior to our work.

[2] The material in Section 1.4 is related to this approach.

In this work we use the Stepanov method (which we explain below) to prove the bias is vanishing (i.e., it goes down to 0 together with p). In Sections 1.1 and 1.2 we explain our approach, in Section 1.3 we present the bounds we get on the bias, and finally in Section 1.5 we conclude with some remarks, open problems and summary of following work.

1.1 The Stepanov method

In 1924 Artin conjectured the "Riemann hypothesis for finite fields". In a landmark result Weil proved in 1948 Artin's conjecture [Wei48]. A corollary of this is the Weil bound for character sums, of which a special (representing) case is:

Theorem 1.2. *Let $\mathbb{F}_q$ be a finite field of odd characteristic and $h \in \mathbb{F}_q[X]$ a polynomial of degree d such that h cannot be written as cg^2 with $c \in \mathbb{F}_q$ and $g \in \mathbb{F}_q[X]$. Then*

$$| \sum_{x \in \mathbb{F}_q} \chi(h(x)) \mid \leq (d-1)\sqrt{q}.$$

Thus, the evaluations of h over $\mathbb{F}_q$ have almost the same number of quadratic residues and non-residues with a bias of at most $O(d\sqrt{q})$.

Weil's proof of the Riemann hypothesis for curves over finite fields uses algebraic-geometric arguments. Stepanov, in a series of papers around 1970, gave alternative "elementary" proofs for special cases of Weil's result. Stepanov's proof uses what we nowadays call the "polynomial method" [Ste69, Ste70, Ste72b, Ste72a, Ste73]. The method was refined (independently) by Bombieri [Bom74] and Schmidt [Sch73, Sch74] who were able to use it to prove the full "Riemann hypothesis for finite fields". Further work used the method to prove bounds even in cases where algebraic geometry seems unable to provide any nontrivial ones [Mit92, HB96].

Very roughly, using the terminology of Theorem 1.2, our objective is the following. Let A denote the set of points $x \in \mathbb{F}_q$ such that $h(x)$ is a quadratic residue. Our goal is to show A is small (and then, proving a corresponding bound for non-residues the theorem follows). The polynomial method approach to this would be to show that there exists a low-degree, non-zero polynomial R vanishing on A with high multiplicity. If we are able to show such an R exists then the number of elements in A is bounded by the degree of R divided by the multiplicity. However, the naive approach of finding R by solving a linear system, where the intermediates are R's coefficients and there is one equation per each element of A is doomed to fail.

The crux of Stepanov's argument is the observation that A is a *variety*, i.e., it is the zero locus of polynomial equations. Specifically, A is the set of ponits x s.t. $x^q = x$ (i.e., $x \in \mathbb{F}_q$) and $(h(x))^{\frac{q-1}{2}} = 1$ (i.e., $h(x)$ is a quadratic residue). Stepanov chooses R with many degrees of freedom, such that when R is restricted to the variety it is simplified to the zero polynomial. Thus, the vanishing of R on A is obtained not by a set of local constraints (one constraint per each element of A) but rather globally, by forcing R to be the zero polynomial when taken modulo the equations that define the variety.

The same approach also works to count the number of $\mathbb{F}_q$-points on more general algebraic-geometric curves (generalizing the result above, which is about the curve $y^2 = h(x)$), and, as we mentioned before, Bombieri [Bom74] and Schmidt [Sch73, Sch74] used Stepanov's method to prove the Riemann hypothesis for finite fields, which shows that the number of rational points on any small-genus curve over $\mathbb{F}_q$ is close to q: the error term is $O(g\sqrt{q})$.

These same results also automatically handle character sums of polynomials over small genus curves. Let C be an algebraic-geometric curve over $\mathbb{F}_q$ given by the equation $R(x,y) = 0$, and let g be its genus. Let $h(x,y) \in \mathbb{F}_q[x,y]$ be a low degree polynomial, viewed as an algebraic function on C. Let χ be a multiplicative character of $\mathbb{F}_q^*$. If we want to bound:

$$|\sum_{(x,y)\in C(\mathbb{F}_q)} \chi(h(x,y))|,$$

which measures how frequently h evaluates to a quadratic residue on C, it suffices to show that the number of points on the curve C_h embedded in 3 dimensional space given by the equations:

$$R(x,y) = 0,$$
$$z^2 = h(x,y),$$

is close to q. Thus we again are trying to estimate the number of $\mathbb{F}_q$-points on a curve C_h. The Riemann hypothesis over finite fields again applies; the key fact that is needed is the Riemann-Hurwitz formula, which shows that C_h also has small genus ($\approx 2g$).

For C of high genus, the above method does not work, and the Hasse-Weil bound is usually meaningless. Indeed, Vladut [Vla96] notes the lack of results for high-genus curves. He then studied these kinds of character sums over the high genus Hermitian curve, and obtained very mild results, using general properties of the zeta functions of curves, and the fact that the Hermitian curve has the maximum possible number of points for its genus.

Instead, we develop a version of the Stepanov method which works directly with the Hermitian curve function field, generalizing the original Stepanov method that works with the rational function field $\mathbb{F}_q(X)$. The main difficulty that we face, which stems from the fact that the Hermitian curve has high genus, is that derivatives (that are needed for vanishing with high multiplicity) carry a penalty in high-genus curves: the derivative of a low degree function on a curve can be of significantly higher degree.

To handle this, we develop the *universal derivative-fix lemma* that we explain next.

1.2 The universal derivative-fix lemma

Our proofs use the Stepanov method, but instead of applying it over the polynomial ring, we apply it over algebraic curves of high genus. For that we need to understand high-order derivatives over algebraic curves, which we discuss first.[3]

[3] We give some necessary mathematical background in Appendices A.1 and A.2.

We start by considering two examples:

the polynomial ring: In the polynomial ring $F = \mathbb{F}_q[X]$, the derivative of a non-constant polynomial is a polynomial of a strictly smaller degree, and the more times we derive, the smaller the degree gets until we reach the zero polynomial. This gives the impression that derivatives are simpler, i.e. have less poles than the original functions.

When transitioning to rational functions, this is no longer the case. For example, when m is smaller then the characteristic of F, $D_x^m(\frac{1}{x}) = \frac{c_m}{x^{m+1}}$ for some non-zero constant c_m.[4] Now, the more we derive the more poles we get, and each derivation increases the pole order by one. Similarly, if we look at $D_x^m(\frac{f}{g})$ we get some polynomial in the derivatives of f and g divided by g^{m+1}, meaning the poles at the zeroes of g increase many-fold as we derive. Thus, both in the case of $\frac{1}{x}$ and in the more general case of $\frac{f}{g}$, the poles "stay where they were", but the pole order increases.

Now consider derivatives of the form $D_g^1(f)$ where $f, g \in F$. From the chain rule $D_g^1(f) = \frac{df}{dg} = \frac{df}{dx}\frac{dx}{dg} = \frac{f'}{g'}$ we see that $D_g^1(f)$ may have poles also where g' has zeroes, and the more zeroes g has, the more new poles we introduce when deriving. Thus, it greatly matters with respect to which function g we choose to derive.

The Hermitian function field: Now we look at $H = \mathbb{F}_{p^2}(x, y) \mod y^p + y - x^{p+1}$. The elements x and y are regular, i.e., they only have poles at a single degree one place, which we denote P_∞. It holds that $v_{P_\infty}(x) = -p$ and $v_{P_\infty}(y) = -(p+1)$. Now, $x^p = D_x(x^{p+1}) = D_x(y^p + y) = D_x(y^p) + D_x(y) = D_x(y)$, and so, $D_x(y) = x^p$ has p^2 poles at P_∞ while y has only $p+1$ poles at P_∞, an increase of $p^2 - p - 1 = 2 \cdot \text{genus}(F) - 1$.

We now give a theorem that applies generally to all function fields. We call it *the universal derivative-fix lemma*, because it fixes all new poles created by the derivation operator, and does so by multiplying by a universal element, that is independent of the function that we derive.

For a divisor D we let $(D)_0$ denote the zero-divisor of D, and $(D)_\infty$ the pole divisor of D, so that $D = (D)_0 - (D)_\infty$. We let $\text{DegSupp}((x)_\infty)$ be the degree of the support of the pole divisor of x, i.e., the degree of the pole divisor of x when all positive coefficients are reduced to one. In Section 2 we prove:

Theorem 1.3. *Let F/K be a function field of genus g and characteristic p. Let $x \in F$ be a separating element of F/K and P_∞ a degree one place of F. Let*

$$G = 3g - 2 + \deg(x) + \text{DegSupp}((x)_\infty),$$

and denote

$$W = G - \max\{v_\infty(dx), 0\},$$
$$\Delta = G + \min\{v_\infty(dx), 0\}.$$

[4] D_x denotes derivation by x. D_x^m is iterating D_x m times. H_x^m is the m'th Hasse derivative. For background on iterated and Hasse derivatives, and for the notation we use, see Appendix A.1.

Then there exists an element

$$0 \neq \omega = \omega(x, P_\infty) \in \mathcal{L}(G \cdot P_\infty - (dx)_0) \subseteq \mathcal{L}(WP_\infty)$$

such that for every $A \geq 0$ and every $f \in \mathcal{L}(AP_\infty)$ it holds that

$$\omega \cdot H_x(f) \in \mathcal{L}((A + \Delta + 1) \cdot P_\infty).$$

We stress that ω does not depend on A or f, and depends only on x (and the place P_∞).

Theorem 1.3 can be generalized to any derivation order m. Let $H_x^m(f)$ denote the m-th Hasse derivative of f with respect to x. Then,

Theorem 1.4. *Keeping the notation as in Theorem 1.3, there exists an element $0 \neq \omega = w(x, P_\infty) \in \mathcal{L}(G \cdot P_\infty - (dx)_0) \subseteq \mathcal{L}(WP_\infty)$ such that for every positive integer $m < p$ (or any positive integer m, if $p = 0$) and every $f \in \mathcal{L}(A \cdot P_\infty)$,*

$$\omega^{2m-1} \cdot H_x^m(f) \in \mathcal{L}(A_m \cdot P_\infty).$$

where $A_m = A - W + m \cdot (\Delta + W + 1)$.

As an example, consider the rational function field $F = K(x)$. Then $g = 0$,$(dx) = -2P_\infty$ and therefore $G = W = 0$ and $\Delta = -2$. Then $w \in \mathcal{L}(0)$ and for any $f \in \mathcal{L}(AP_\infty)$, $H_x(f) \in \mathcal{L}((A-1)P_\infty)$ which is indeed tight.

Next, we examine the Hermitian function field.

- First consider the Hermitian function field when we derive by x. $D_x(y) = x^p$. In fact, $(dx) = (2g-2)P_\infty$, i.e., it has no poles, and all its zeroes are at P_∞. Then, we can take $w = 1$. Furthermore, for every $f \in \mathcal{L}(AP_\infty)$, $-v_\infty(H_x(f)) = v_\infty(dx) - v_\infty(df) \leq 2g - 2 + A + 1$. For a general m, $A_m \leq A + m(2g-1)$.
- Next consider the Hermitian function field when we derive by y. Then, $D_y(x) = \frac{1}{x^p}$ and $(dy) = (p+2)P_\infty - p(x)_0$. Nevertheless, since all functions in $\mathcal{L}(P_\infty)$ are polynomials in x and y, we get that if we are deriving with respect to y we can choose w to be x^p to cancel out the $\frac{1}{x^p}$ which is the derivative of x with respect to y. With this choice of ω we again get that if $f \in \mathcal{L}(AP_\infty)$ then $\omega^m H^m(f) \in \mathcal{L}((A + m(2g-1))P_\infty)$.

In both cases the bounds that we get are worse, and there are several reasons for that:

- We paid an additive g to guarantee a certain Riemann-Roch space is nonempty, by forcing the degree of its divisor to be at least g. While there are divisors of degree $g-1$ which have empty Riemann-Roch spaces, there are divisors of degree 0 which have non-empty Riemann-Roch spaces. It is conceivably possible that the $3g-1$ we have is not mandatory and can be replaced with $2g-1$ as we have in the Hermitian curve. Perhaps, using the Riemann-Roch theorem with canonical divisors would do the trick.

- Additionally, the $2m$ factor is a side effect of the inductive argument which requires us to apply the induction hypothesis twice - once for $H^{m-1}(f)$ and once for $H^1\omega$. If, however, $H^1\omega$ is regular, we can apply the induction hypothesis once and so ω^m would be sufficient. Alternatively, if the poles of $D^m(f)$ which exceed those of $D^{m-1}(f)$ behave like "dividing by a function again and again", similarly to what we saw with $D_x^m(\frac{f}{g})$ in $K(x)$ or to $D_y(f)$ for regular f in the Hermitian function field, we would again get that ω^m is sufficient.

- The requirement $m < p$ is also a side effect of the induction, but when looking at the p-th Hasse derivative of y^p we get from claim A.4 $H_x^p(y^p) = H_x^1(y)^p = x^{p^2}$ which is of pole order $p^3 = p(p+1) + (2g-1)p$, an increase of exactly $2g-1$ times the order of the derivative.

To summarize this, we believe that the following version of Theorem 1.4 could hold:

Conjecture 1.5. *Let F/K be a function field of genus g. Let $x \in F$ be a separating element of F/k. Let P_∞ be a degree one place of F. There exists an element $0 \neq \omega = w(x, P_\infty) \in F$ such that for every $m \in \mathbb{N}$ and every $f \in \mathcal{L}(P_\infty)$, $\omega^m \cdot H_x^m(f) \in \mathcal{L}(P_\infty)$. Furthermore, $\omega \in \mathcal{L}((2g-1+\deg(x)+\mathrm{DegSupp}((x)_\infty)) \cdot P_\infty)$ and so if $f \in \mathcal{L}(A \cdot P_\infty)$ then $\omega^m \cdot H_x^m(f) \in \mathcal{L}(A_m \cdot P_\infty)$ for $A_m = A + m(2g-1+\deg(x)+\mathrm{DegSupp}((x)_\infty) + \min\{v_\infty(dx), 0\})$.*

1.3 Our results

We start with multiplicative characters. In Section 3.1 we give, as a warm-up, a version of Stepanov's proof of the Weil bound. In this proof we work in the function field

$$\widehat{F} = F(z) \pmod{z^2 - h(x)}, \tag{1}$$

and we use the universal derivative-fix lemma (for the function field $\widehat{F}$). We prove:

Theorem 1.6. *Let $\mathbb{F}_q$ be a finite field of odd characteristic and $h \in \mathbb{F}_q[X]$ a square-free polynomial of odd degree d. Then*

$$|\sum_{x \in \mathbb{F}_q} \chi(h(x))\ | \leq O(d\sqrt{q}).$$

This theorem is weaker then the Weil bound in two aspects. First, the error is bounded by $O(d\sqrt{q})$ instead of the tight bound of $(d-1)\sqrt{q}$ in Weil's bound. Also, we prove the claim only for square free polynomials h of odd degree, while Weil's bound holds for any h that is not a square over the algebraic closure of $\mathbb{F}_q$. On the bright side, the proof is quite generic.

After this warm-up exercise we give in Section 3.2 a corresponding proof for the Hermitian function field, i.e., p is a prime power, $q = p^2$,

$$F = \mathbb{F}_q(x) \tag{2}$$

is the rational function field, and,

$$H = \mathbb{F}_q(x,y) \mod \varphi(x,y), \tag{3}$$

where $\varphi(x,y) = y^p + y - x^{p+1}$, is the Hermitian function field. Let $\mathcal{C}$ denote all the $\mathbb{F}_q$-rational points on the curve φ. We prove

Theorem 1.7. *Suppose p is prime and $h = \sum_{k,\ell} c_{k,\ell} x^k y^\ell \in \mathbb{F}_q[x,y]$ is a polynomial with total degree $d < \sqrt{p}$, and $(p, p+1)$-weighted degree w. The weighted degree w is obtained by a unique monomial $x^k y^\ell$. Suppose $d = k + \ell$ and $w = pk + (p+1)\ell$ and further assume $k + \ell$ is odd. Then*

$$\mid \sum_{(x,y)\in\mathcal{C}} \chi(h(x,y)) \mid \leq O(\sqrt{d} \cdot p^{2.5}).$$

As before, we use the universal derivative-fix lemma (this time for the function field $\widehat{H} = H(z) \pmod{z^2 - h(x))}$, and as before, a drawback of the proof is that it requires additional assumptions on the polynomial h beyond assuming h is not a square polynomial. On the bright side the proof is quite generic, e.g., it also applies for the first levels of the Hermitian tower (we omit the specific claims because the parameters become more involved).

Next, we look for a character sum bound that holds for any non-square polynomial, rather than more specific polynomials. Schmidt in [Sch06, Chapter I, section 5] showed how to use the Stepanov method to prove the Weil bound for any non-square polynomial. Schmidt works over the rational function field $\mathbb{F}_q(X)/\mathbb{F}_q$. We extend the argument to work over the Hermitian function field. The proof we give is specific for the Hermitian function field and delicate. In essence, we still follow the generic proof, which uses the Riemann-Roch theorem and the derivative-fix Lemma as black-box theorems, but we replace the black-box components with specific functions, tailored for the Hermitian function field. Consequently, in Section 4 we prove

Theorem 1.8. *Suppose p is a prime power, $q = p^2$, $h \in \mathbb{F}_q[x,y]$, $\deg(h) = d$. If $Z^2 - h$ is absolutely irreducible then*

$$\mid \sum_{(x,y)\in\mathcal{C}} \chi(h(x,y)) \mid = O(d \cdot p^{2.5}).$$

Thus the relative bias (i.e., bias divided by the number of evaluation points) is $O(\frac{d}{\sqrt{p}})$. This is worse than the bias $O(\sqrt{\frac{d}{p}})$ in Theorem 1.7, but it holds for any non-square polynomial.

Finally, in Section 5 we give a corresponding bound for additive characters. Here, the Hasse derivatives have a very explicit structure, and we can explicitly point out the derivative fix element w guaranteed by Theorem 1.3. We prove:

Theorem 1.9. *Suppose $h = \sum_{k,\ell} c_{k,\ell} x^k y^\ell \in \mathbb{F}_q[x, y]$ is a polynomial with total degree $d < \sqrt{p}$, and $(p, p + 1)$-weighted degree w. Suppose w is odd. Then*

$$|\sum_{(x,y)\in\mathcal{C}} (-1)^{\mathsf{Tr}(h(x,y))}| \leq O(d \cdot p^{2.5}).$$

1.4 A surprising remark about the roots of the zeta function

Finally, we make a comment about what our results mean from the point of view of the zeta function of the relevant curves. For a curve C of genus g_C over a finite field $\mathbb{F}_q$, there is a zeta function $Z_C(T) \in \mathbb{C}(T)$ of the form $\frac{P_C(T)}{(1-T)(1-qT)}$. $P_C(T)$ is always of the form

$$P_C(T) = \prod_{i=1}^{2g_C} (1 - \omega_i T),$$

where each $\omega_i \in \mathbb{C}$ has $|\omega_i| = \sqrt{q}$. (These are called the reciprocal roots of $P_C(T)$). These ω_i have the property that the number of $\mathbb{F}_q$ points on C equals

$$q + 1 - \sum_i \omega_i.$$

Our result on the quadratic residue character sums on the Hermitian curve H are essentially an estimate on the number of points on some curve H' which is a degree 2 cover of H. By general properties of zeta functions, we have that $P_H(T)$ divides $P_{H'}(T)$.

Let $b = g_{H'} - g_H$. Let $\omega_1, \ldots, \omega_{2g_H}$ be the reciprocal roots of $P_H(T)$, and let $\alpha_1, \ldots, \alpha_{2b}$ be the remaining reciprocal roots of $P_{H'}(T)$ (other than the ω_i). By the Riemann-Hurwitz formula, $g_{H'} \approx 2g_H$, and thus $b \approx g_H \approx q/2$.

The number of $\mathbb{F}_q$ points on H equals $q^{3/2} + 1$, and also equals

$$q + 1 - \sum_{i=1}^{2g_H} \omega_i$$

By our results, the number of $\mathbb{F}_q$ points on H' is $q^{3/2} + 1 \pm o(q^{3/2})$, and also equals:

$$q + 1 - \sum_{i=1}^{2g_H} \omega_i - \sum_{i=1}^{2b} \alpha_i = q^{3/2} + 1 - \sum_{i=1}^{2b} \alpha_i.$$

Thus,

$$|\sum_{i=1}^{2b} \alpha_i| = o(q^{3/2}).$$

On the other hand, each α_i has magnitude $\sqrt{q}$, and b has magnitude $\Theta(q)$. This means that there is significant cancellation in the sum $\sum_{i=1}^{2b} \alpha_i$. We find this to be quite special. In general one cannot guarantee much cancellation in the sum of the reciprocal roots of a zeta function. Our result thus finds a special instance where there is cancellation.

1.5 Conclusions and following work

Understanding the rate vs. distance problem for *binary* codes is the holy grail of error correcting codes. The question is two-fold:

- What is the correct combinatorial bound, and,
- How close can we get to it explicitly.

Currently, the best lower bound (showing non-constructive existence) for binary codes is the GV bound, the best upper bound (showing impossibility results) is the MRRW bound, and there is a significant gap between the GV and the MRRW bounds. Two major questions are open:

1. What is the best combinatorial problem, and where (between the GV bound and the MRRW bound) does it lie?
2. The best explicit codes are far from the GV bound. The situation is the same for the important regime where the distance is close to half, $\delta = \frac{1-\varepsilon}{2}$. Can one get better explicit codes?

Surprisingly, for alphabet size $q = 49$ and above the GV bound is not tight, and there are better AG codes. Furthermore, these AG codes are explicit. In fact, as the GV bound captures the behavior of random codes, it is not surprising that codes that surpass the GV bound are explicit (but it is surprising that such codes exist at all). Can a similar phenomenon occur for binary codes?

A natural approach for getting explicit codes close to the GV bound (or better than it!) is to start with such a code over a constant alphabet, and, somehow reduce the alphabet size. Concatenation is a natural candidate, but it seems it is not strong enough to approach the GV bound. Trace codes are also natural candidates, but the results obtained so far are poor. A natural question is whether these results can be improved. Two questions of this type are:

- How good the code described before Theorem 1.1 is? Is it close to the GV bound? If not, is there a natural subclass of polynomials that is good? We believe these questions deserve further study.
- If we work over the Hermitian function field, and our set of evaluation function also allows polynomials in x or y alone, then the relative bias of the Trace Hermitian code cannot be better then the relative bias of the Trace RS code. To see that suppose f is a non-square polynomial in $\mathbb{F}_q[x]$ of degree d that has bias e over $\mathbb{F}_q$. As every $x \in \mathbb{F}_q$ has p rational solutions in y that do not affect f, we get a bias of ep over the

Hermitian curve, which has the same relative bias. However, it is possible that the bias of the Trace Hermitian code is much worse. A natural open question is understanding whether this happens or not. Currently, the bias we know for the Trace RS code is $O(d\sqrt{q})$. Lifting this to the Hermitian curve (by taking a polynomial that only depends on one variable) we get an absolute bias of $O(dq)$. For some minor restriction on the evaluation polynomial we can bound the bias of Trace Hermitian curve by $O(\sqrt{dpq})$, which is worse (as $d < \sqrt{p}$). Is it possible to improve the bound we get for the Trace Hermitian curve to $O(dq)$?

We believe the techniques we develop are of independent interest. The results on Trace AG codes prior to our work, in the interesting case of high genus curves, were very poor, and the only non-trivial result we are aware of is [Vla96] who managed to bound the bias by some constant. In this work we use the Stepanov bound, together with the general derivative-fix-lemma to show vanishing bounds on the bias. We believe there is much more to learn and understand about the surprising behavior of derivatives in function fields.

We conclude with an intriguing question where further progress has been made following our work: What one can say about general Trace AG codes? Our techniques (and the derivative-fix-lemma in particular) allow us to study not only the trace of the Hermitian curve, but also the trace of the codes in the first few levels of the Hermitian tower. We do not state these results because technically they get more involved. Recently, Gil Cohen, Dean Doron, Noam Goldgraber and Tomer Manket [CDGM25] informed us they managed to generalize the techniques to general AG curves. At a very high level, the techniques of this paper and of [CDGM25] are related in the same way as Stepanov's method and Bombieri's version of Stepanov's method are related:

1. Their proof works in a more general setting of an arbitrary algebraic curve, and directly works with Riemann-Roch spaces of algebraic functions on the curve, while our proof works concretely with the equation of the Hermitian curve (and closely related curves), and represents algebraic functions on the curve more explicitly as polynomials.

2. Unlike our proof, the proof in [CDGM25] does not use multiplicities, and instead relies on Bombieri's "raising to the power $\sqrt{q}$" trick.

We are hopeful that these developments would trigger further work on this fascinating question.

2 The universal derivative-fix lemma

2.1 First order derivatives

We first prove Theorem 1.3 about first order derivatives. Let us denote $f' := H_x^1(f) = D_x^1(f)$. As $f' = \frac{df}{dx}$ we have $(f') = (df) - (dx)$. It follows that the poles of f' can come either from poles of df, or, from zeroes of dx. Since $f \in \mathcal{L}(AP_\infty)$, Claim A.8 tells us all the poles of f and df are at P_∞. Claim A.8 also tells us that $v_\infty(df) \geq v_\infty(f) - 1 \geq -(A+1)$, and so df has at most $A+1$ poles, all of which must be at P_∞. We wish to find $\omega \in F$ s.t. $\omega \cdot f' \in \mathcal{L}(P_\infty)$ so we need to choose $\omega \in F$ that cancels the poles of f' at all places other than P_∞. These poles can arise only from zeroes of dx *outside* P_∞. Indeed, we will find $w \in \mathcal{L}(G \cdot P_\infty - (dx)_0)$ for some large enough G.

While we are interested in the zeroes of dx, we first consider the *poles* of dx. By Claim A.8:

- The poles of dx are at the same places as the poles of x, i.e., $v_P(dx) < 0$ implies $v_P(x) < 0$, and,
- At any place P where dx and x have a pole, dx may have at most one more pole than x, i.e., $v_P(dx) \geq v_P(x) - 1$.

It therefore follows that $\deg((dx)_\infty) \leq \deg(x) + \mathrm{DegSupp}((x)_\infty)$.

Next, we use the fact that (dx) is a canonical divisor (being a divisor associated with a differential), and therefore has degree $2g-2$. Thus, the number of zeroes of dx is exactly $2g-2$ more than the number of poles of dx, and in total we get

$$\deg((dx)_0) \leq \deg(x) + \mathrm{DegSupp}((x)_\infty) + 2g - 2 = G - g,$$

and for $D = G \cdot P_\infty - (dx)_0$ we have

$$\deg(D) = G - \deg((dx)_0) \geq g.$$

By the Riemann-Roch Theorem there exists some $0 \neq \omega \in \mathcal{L}(D)$. Fix any such ω. Set $\Delta = G + \min\{v_\infty(dx), 0\}$. Then,

Claim 2.1. $\omega f' = \omega \cdot \frac{df}{dx} \in \mathcal{L}((A + \Delta + 1)P_\infty)$.

Proof. For any $P \neq P_\infty$, $v_P(\omega) \geq -v_P(D) = v_P((dx)_0)$, and

$$v_P(\omega f') = v_P(\omega) + v_P(df) - v_P(dx) \geq v_P(\omega) + v_P(df) - v_P((dx)_0) \geq v_P(df) \geq 0,$$

where we have used that $v_P((dx)_0) = \max\{0, v_P(dx)\} \geq v_P(dx)$.

Next we compute the pole order of wf' at P_∞. We have $w \in \mathcal{L}(D) \subseteq \mathcal{L}((G - \max\{v_\infty(dx), 0\})P_\infty)$. Thus,

$$\begin{aligned} -v_\infty(\omega f') &= v_\infty(dx) - v_\infty(\omega) - v_\infty(df) \\ &= v_\infty(dx) + G - \max\{v_\infty(dx), 0\} - v_\infty(df) \\ &\leq A + 1 + G + v_\infty(dx) - \max\{v_\infty(dx), 0\}, \end{aligned}$$

because $v_\infty(df) \geq v_\infty(f) - 1 \geq -A - 1 = -(A+1)$. However,

$$v_\infty(dx) - \max\{v_\infty(dx), 0\} = \min\{0, v_\infty(dx)\},$$

and the proof is complete. □

2.2 High order derivatives

Proof. We use the same w as before. We prove by induction. We already saw the $m = 1$ case. Assume for m and let us prove for $m+1$. The $m+1$-th Hasse derivative is the same as the $m+1$-th iterated derivative D_x^{m+1} up to multiplication by a non-zero scalar (and using $m+1 < p$). Now,

$$\begin{aligned}\omega^2 D_x(\omega^{2m-1} D_x^m f) &= \omega^2 \ [\ D_x(\omega^{2m-1}) D_x^m f + \omega^{2m-1} D_x(D_x^m f)\ ] \\ &= (2m-1)\omega^{2m} \cdot D_x(\omega) \cdot D_x^m f + \omega^{2m+1} D_x^{m+1} f\end{aligned}$$

Thus,

$$\omega^{2m+1} D_x^{m+1} f = \omega^2 D_x(\omega^{2m-1} D_x^m f) - (2m-1)\omega^{2m} D_x(\omega) \cdot D_x^m f.$$

By the induction hypothesis and the $m = 1$ case:

$$\begin{aligned}\omega^{2m-1} \cdot D_x^m f &\in \mathcal{L}(A_m \cdot P_\infty), \\ \omega D_x(\omega^{2m-1} D_x^m f) &\in \mathcal{L}((A_m + (\Delta+1)) \cdot P_\infty).\end{aligned}$$

Also $\omega \in \mathcal{L}(W P_\infty)$. By the $m = 1$ case,

$$\omega \cdot D_x(\omega) \in \mathcal{L}((W + (\Delta+1)) \cdot P_\infty)$$

The term $\omega^2 D_x(\omega^{2m-1} D_x^m f)$ is in $\mathcal{L}((A_m + W + \Delta + 1)P_\infty)$. The term $\omega D_x(\omega) \cdot \omega^{2m-1} D_x^m f$ is also in $\mathcal{L}((A_m + W + \Delta + 1)P_\infty)$. Altogether, $\omega^{2m+1} D_x^{m+1} f$ is in $\mathcal{L}(A_{m+1} P_\infty) = \mathcal{L}((A_m + W + \Delta + 1)P_\infty)$. □

Remark 2.2. *Note that if $D_x(\omega)$ is in $\mathcal{L}(P_\infty)$, we can multiply by a single ω per derivative, instead of multiplying by ω^2.*

3 Proving the Hermitian Curve Weil bound with the universal derivative-fix lemma, multiplicative characters

In this section we prove a Weil bound for multiplicative characters over the Hermitian curve. We bound the bias for a restricted class of polynomials - those spanned by products of $x^i y^j$ where $i + j$ is odd. All such polynomials are non-squares, but not all non-squares are of this form. For this class of polynomials we bound the bias in the order of $O(\sqrt{\frac{d}{p}})$. We remark that in Section 4 we bound the bias for all non-square polynomials, but we get the worse bound $O(\frac{d}{\sqrt{p}})$ on the bias.

We begin in Section 3.1, as a warm-up, by first proving the Weil bound over $\mathbb{F}_q[x]$ using the same technique, and then in Section 3.2, we prove the result for the Hermitian curve.

3.1 Proof over $\mathbb{F}_q[x]$

Proof. (Of Theorem 1.6) $F = \mathbb{F}_q(x)$ is the rational function field and $h \in \mathbb{F}_q[x]$ a square-free polynomial of odd degree. Let

$$\mathcal{C} = \{(x, z) \in \overline{\mathbb{F}}_q^2 \mid z^2 - h(x) = 0\}. \tag{4}$$

Let P_∞ be the unique pole of x and v_∞ its corresponding valuation function. We will create a nonzero algebraic function $R \in \mathcal{L}(D \cdot P_\infty)$ such that for each point $P = (\alpha, \beta)$ of $\mathcal{C}$ with $h(\alpha) \neq 0$, R vanishes on P with multiplicity at least M (i.e., $v_P(R) \geq M$). This gives an upper bound of $\frac{D}{M}$ on the number of such points. It follows that the number of elements $\alpha \in \mathbb{F}_q$ for which $h(x)$ is a square is at most $d + \frac{D}{2M}$ (because there are at most d points with $h(\alpha) = 0$ and at most $\frac{D}{2M}$ points with non-zero square evaluation).

Let $A = q - 1$ and $B < Md$ an integer to be determined later. We choose R to be of the form:

$$R = \sum_{i=0}^{B-1} u_i x^{iq} + \sum_{i=0}^{B-1} v_i x^{iq} z^q,$$

where each u_i and v_i comes from $\mathcal{L}(A \cdot P_\infty)$. Observe that all the x^{iq} and $x^{iq}z^q$ all have v_∞ valuations differing by at least q, because $v_\infty(x) = -2$ is even, and $v_\infty(z) = -d$ (see Appendix A.4) is by assumption odd. Since $A = q - 1$ we get that R is nonzero provided some u_i or some v_i is nonzero. Now we calculate:

- **Degrees of freedom:** The total number of degrees of freedom is at least $2B(A-g+1)$, because there are $2B$ functions from $\mathcal{L}(A \cdot P_\infty)$ that we have to choose.

- **Number of constraints:** For a separating element t, we let $H_t^{(m)}$ denote the m'th order Hasse derivative with respect to t.[5] By Claim A.1 it is enough to show that for every rational place $P = (\alpha, \beta)$ on $\mathcal{C}$ with $h(\alpha) \neq 0$, there exists a separating element t such that $v_P(t) = 1$ and for each $m \in \{0, 1, \ldots, M-1\}$, $H_t^{(m)}R(P) = 0$. In fact, if $P = (\alpha, \beta)$ is as above, then $t = x - \alpha$ is a separating element with $v_P(t) = 1$ (see Claim A.13). Thus, by Fact A.2 it suffices to show that for every such point P, and every $m \in \{0, 1, \ldots, M-1\}$, $H_x^{(m)}R(P) = 0$.

  For $h \in F$ and $m \in \mathbb{N}$, $m < q$, let $h^{(m)}$ denote $H_x^{(m)}h$. Then, $R^{(m)}$ vanishes on all $\mathbb{F}_q$ points of $\mathcal{C}$ follows from the condition:

  $$\omega_m R^{(m)} = \sum_{i=0}^{B-1} \omega_m u_i^{(m)} x^i + \sum_{i=0}^{B-1} \omega_m v_i^{(m)} x^i z = 0,$$

  where we have used Corollary A.6 and the fact that $x^q = x$ and $z^q = z$ over all the points of $\mathcal{C}$ (as they are $\mathbb{F}_q$ rational).

[5] See Appendix A.1 for more details about Hasse derivatives.

By Theorem 1.4, this asks for a certain member of

$$\mathcal{L}\left((A+O(gm)+B\deg(X)+\deg(Z))\cdot P_\infty\right)$$

to equal 0. Thus the total number of constraints is:

$$\sum_{m=0}^{M-1}(A+2B+O(md)+O(d))\leq MA+2BM+O(M^2d).$$

We need the number of degrees of freedom to be larger than the number of constraints. Taking $2B=M+E$ we want

$$(M+E)(A-g)>MA+2BM+O(M^2d),$$

i.e.,

$$E(A-g)>Mg+2BM+O(M^2d)=O(M^2d)$$

because $g=O(d)$ and $B<Md$. As $A=q-1\geq 2d\geq 2(g-1)$ (see Appendix A.3) it is enough that $EA=\Omega(M^2d)$. Remembering that $A=q-1$ it is enough that $E=\Theta(\frac{M^2d}{q})$.

The number of points $\alpha\in\mathbb{F}_q$ for which $h(\alpha)$ is a square is at most $\frac{\deg(R)}{2M}+d$, and

$$|\frac{\deg(R)}{2M}+d-\frac{q}{2}|=|\frac{2Bq+A+dq}{2M}+d-\frac{q}{2}|=\frac{Eq}{2M}+\frac{A+dq}{2M}+d=O(Md+\frac{dq}{M}).$$

Equating the two error terms we get $M^2d=dq$ or $M=\sqrt{q}$.

It follows that the number of points $\alpha\in F_q$ for which $h(\alpha)$ is a square is at most $\frac{q}{2}+O(d\sqrt{q})$. We can get the same bound for the number of points $\alpha\in F_q$ for which $h(\alpha)$ is a non-square (e.g., by counting the number of squares ch obtains, when c is a non-square in $\mathbb{F}_q$). We therefore conclude that $|\sum_{x\in\mathbb{F}_q}\chi(h(x))|$, which is the difference between the number of squares and non-squares, is at most $O(d\sqrt{q})$, completing the proof. □

3.2 Proof over H

Proof. (Of Theorem 1.7) Let $\widehat{H}=H(z)\pmod{z^2-h(x,y)}$. Let $\mathcal{C}_H$ be the set of all $\mathbb{F}_q$ rational points of the curve,

$$\mathcal{C}_H=\{(x,y,z)\in\overline{\mathbb{F}}_q^3\mid y^p+y=x^{p+1},z^2-h(x,y)=0\}.$$

Let P_∞ be the unique pole of x in $\mathcal{C}_H$ and v_∞ its corresponding valuation function. P_∞ is also the unique pole of y and z, and $v_\infty(x)=-2p$, $v_\infty(y)=-2(p+1)$ and $v_\infty(z)=-w$. The genus g of $\mathcal{C}_H$ is at most $g=p^2+w\leq p^2+d(p+1)\leq 2p^2$ (see Appendix A.5). We will create a nonzero algebraic function $R\in\mathcal{L}(D\cdot P_\infty)$ such that for each $\mathbb{F}_q$ point $P=(\alpha,\beta,\gamma)$

of $\mathcal{C}_H$ with $h(\alpha, \beta) \neq 0$, R vanishes on P with multiplicity at least M (i.e., $v_P(R) \geq M$). This will give an upper bound of $\frac{D}{M} + w$ on the number of $\mathbb{F}_q$ points of $\mathcal{C}_H$.

Let $A = pq - d$ and $B < p$ an integer to be determined later. We choose R to be of the form:

$$R = \sum_{i=0}^{B-1} u_i x^{iq} + \sum_{i=0}^{B-1} v_i x^{iq} z^q,$$

where each u_i and v_i comes from $\mathcal{L}(A \cdot P_\infty)$. Observe that $v_\infty(x^{iq})$ is an even multiple of pq. Also, $v_\infty(z) = -w = -pk - (p+1)\ell = -(k+\ell)p - \ell$ and so $v_\infty(z^q)$ is ℓq close to an odd multiple of pq. As $\ell < d$ and $A = pq - d$, all the $u_i x^{iq}$ and $v_i x^{iq} z^q$ have different v_∞ valuations, and so, R is nonzero provided some u_i or some v_i is nonzero.

Now we calculate:

- **Degrees of freedom:** There are $2B$ functions from $\mathcal{L}(A \cdot P_\infty)$ that we have to choose. So the total number of degrees of freedom is at least $2B(A - g + 1)$,

- **Number of constraints:** By Claim A.1 it is enough to show that for every rational place $P = (\alpha, \beta, \gamma)$ on $\mathcal{C}_H$ with $h(\alpha, \beta, \gamma) \neq 0$, there exists a separating element t such that $v_P(t) = 1$ and for each $m \in \{0, 1, \ldots, M-1\}$, $H_t^{(m)} R(P) = 0$. In fact, if $P = (\alpha, \beta, \gamma)$ is as above, then $t = x - \alpha$ is a separating element with $v_P(t) = 1$ (see Claim A.13). Thus, by Fact A.2 it suffices to show that for every such point P, and every $m \in \{0, 1, \ldots, M-1\}$, $H_x^{(m)} R(P) = 0$.

  For $h \in F$ and $m \in \mathbb{N}$, let $h^{(m)}$ denote $H_x^{(m)} h$. Then, $R^{(m)}$ vanishes on all $\mathbb{F}_q$ points of $\mathcal{C}$ follows from the condition:

  $$\omega_m R^{(m)} = \sum_{i=0}^{B-1} \omega_m u_i^{(m)} x^i + \sum_{i=0}^{B-1} \omega_m v_i^{(m)} x^i z = 0,$$

  where we have used Corollary A.6 and the fact that $x^q = x$ and $z^q = z$ over all the points of $\mathcal{C}$ (as they are $\mathbb{F}_q$ rational).

  By Theorem 1.4, this asks for a certain member of

  $$\mathcal{L}\left((A + O(gm) + B\deg(X) + \deg(Z)) \cdot P_\infty\right)$$

  to equal 0. Thus the total number of constraints is:

  $$\sum_{m=0}^{M-1} (A + O(mp^2) + 2pB + w) \leq MA + 2pBM + O(M^2 p^2),$$

  because $w \leq d(p+1) = O(p^2)$.

We need the number of degrees of freedom to be larger than the number of constraints. Taking $2B = M + E$ we want

$$(M+E)(A-g) > MA + 2pBM + O(M^2p^2),$$

i.e.,

$$E(A-g) > Mg + 2pBM + O(M^2p^2) = O(M^2p^2 + pBM) = O(M^2p^2)$$

because $g = O(p^2)$ and $B < Mp$. As $A = pq - d \geq 4p^2 \geq 2g$, it is enough that $EA = \Omega(M^2p^2)$. Remembering that $A = pq - d = \Theta(pq)$ it is enough that $E = \Theta(\frac{M^2}{p})$.

The number of points $(\alpha, \beta) \in H \cap \mathbb{F}_q^2$ for which $h(\alpha, \beta)$ is a square is at most $\frac{\deg(R)}{2M} + w$, and

$$|\frac{\deg(R)}{2M} + w - \frac{pq}{2}| = |\frac{2Bpq + A + wq}{2M} + w - \frac{pq}{2}| = \frac{Epq}{2M} + \frac{A + wq}{2M} + w = O(Mq + \frac{dpq}{M}).$$

Equating the two error terms we get $M^2q = dpq$ or $M = \sqrt{dp}$.

It follows that the number of points $(\alpha, \beta) \in H \cap \mathbb{F}_q^2$ for which $h(\alpha, \beta)$ is a square is at most $\frac{pq}{2} + O(\sqrt{d}p^{2.5})$. We can get the same bound for the number of points $(\alpha, \beta) \in H \cap \mathbb{F}_q^2$ for which $h(\alpha, \beta)$ is a non-square (e.g., by counting the number of squares ch obtains, when c is a non-square in $\mathbb{F}_q$). We therefore conclude that $|\sum_{x \in \mathbb{F}_q} \chi(h(x))|$, which is the difference between the number of squares and non-squares, is at most $O(\sqrt{d}p^{2.5})$, completing the proof. □

4 Proving the Hermitian Curve Weil bound for all non-square polynomials

In this section we bound the bias of multiplicative characters over the Hermitian curve, and the bound works for all non-square polynomials. However, now we get again bias in the order of $O(\frac{d}{\sqrt{p}})$ (instead of the better $O(\sqrt{\frac{d}{p}})$ of Section 3).

We define:

Definition 4.1. *We say $h \in H$ is* special *if it can be represented as*

$$h(x,y) = (c(x))^{q^i} \cdot \frac{g_1^2(x,y)}{g_2^2(x,y)}$$

where $i \in \mathbb{N}$, $g_1, g_2 \in \mathbb{F}_q[x,y]$ and $c \in \mathbb{F}_q[x]$ with $\deg(c) \leq \deg(h)$. *We say h is* non-special *otherwise.*

By the original Weil bound for $\mathbb{F}_q[x]$:

Claim 4.2. *If h is as in Theorem 1.8 and in addition h is special then*

$$\mid \sum_{(x,y)\in\mathcal{C}} \chi(h(x,y)) \mid \leq (d-1)p\sqrt{q}.$$

Proof.

$$\mid \sum_{(x,y)\in\mathcal{C}} \chi(h(x,y)) \mid = \mid \sum_{(x,y)\in\mathcal{C}} \chi(c(x)^{q^i}) \mid = \mid \sum_{(x,y)\in\mathcal{C}} \chi(c(x)) \mid,$$

where the first equality is because χ is multiplicative and one on squares, and the second is because we only evaluate over $x \in \mathbb{F}_q$. Now if c is a square in $\overline{\mathbb{F}_q}[x]$, say, $c = z^2$, then c^{q^i} is also a square, $c^{q^i} = (z^{q^i})^2$, and therefore h is also a square. A contradiction. Hence, c is not a square, and therefore by the original Weil bound $\mid \sum_{(x,y)\in\mathcal{C}} \chi(c(x)) \mid \leq (d-1)p\sqrt{q}$ (using the fact that every $x \in \mathbb{F}_q$ appears in exactly p tuples $(x,y) \in H$). □

It therefore follows that to prove Theorem 1.8 it is enough to prove the following theorem:

Theorem 4.3. *Suppose p is a prime power, $q = p^2$, $h \in \mathbb{F}_q[x,y]$, $\deg(h) = d$, such that $Z^2 - h$ is absolutely reducible. Further assume h is not special. Then*

$$\mid \sum_{(x,y)\in\mathcal{C}} \chi(h(x,y)) \mid = O(d \cdot p^{2.5}).$$

Proof. (of Theorem 4.3) Let S be the set of $\mathbb{F}_q$ rational points of H such that $h(x,y)$ is a non-zero square in $\mathbb{F}_q$. Let P_∞ be the unique pole of x in H and v_∞ its corresponding valuation function. P_∞ is also the unique pole of y, $v_\infty(x) = -p$, $v_\infty(y) = -(p+1)$. The genus g of H is $\frac{p(p-1)}{2}$ (see Appendix A.5).

We create a nonzero algebraic function $R \in \mathcal{L}(D \cdot P_\infty)$ such that for each $P = (\alpha, \beta) \in S$, R vanishes on P with multiplicity at least M (i.e., $v_P(R) \geq M$). This gives an upper bound of $\frac{D}{M}$ on the number of $\mathbb{F}_q$ points of S.

Let $B_x = \lfloor \frac{p}{4} - d \rfloor$ and $B_y = \frac{p}{2} - (B_x + d) - 1 = \frac{p}{4} - 1$. We choose R to be of the form:

$$R = u + vh(x,y)^{\frac{q-1}{2}},$$

where each u and v come from $V_{B_x q, B_y}$ where $V_{c_1,c_2} = \mathsf{Span}\{x^i y^j \mid i \leq c_1, j \leq c_2\}$.

A crucial step in the proof is proving R is nonzero provided u or v is nonzero. We defer the proof to Section 4.1 and we continue assuming this claim.

Now we calculate:

- **Degrees of freedom:** There are $2qB_xB_y$ degrees of freedom,

- **Number of constraints:** By Claim A.1 it is enough to show that for every rational place $P=(\alpha,\beta)$ from S there exists a separating element t such that $v_P(t)=1$ and for each $m \in \{0,1,\ldots,M-1\}$, $H_t^{(m)}R(P)=0$. In fact, if $P=(\alpha,\beta)$ is as above, then $t=x-\alpha$ is a separating element with $v_P(t)=1$ (see Claim A.13). Thus, by Fact A.2 it suffices to show that for every such point P, and every $m \in \{0,1,\ldots,M-1\}$, $H_x^{(m)}R(P)=0$.

  For $f \in F$ and $m \in \mathbb{N}$, let $f^{(m)}$ denote $H_x^{(m)}f$. Then, $R^{(m)}$ vanishes on all points of S follows from the condition $h^m R^{(m)}=0$ on S.

  Claim 4.4. *Let $I=\langle x^q-x, h^{(q-1)/2}-1\rangle$. Then $h^m R^{(m)} \bmod I \in V_{1,B_y+md}$.*

  Proof. First, we determine the effect of taking Hasse derivatives with respect to x:

  - If $u \in V_{c_1,c_2}$, say, $u=x^i y^j$ with $i \le c_1$ and $j \le c_2$. Then for every $\ell < p$, $H_x^{(\ell)}(u) \in V_{c_1+p\ell,c_2}$, using $H_x^{(\ell)}(y^{i_2}) = c \cdot y^{i_2-\ell}x^{p\ell}$ for some constant c. Also,
  - By Fact A.3,

    $$h^\ell H_x^{(\ell)}(h^{\frac{q-1}{2}}) = \sum_{j=1}^{\ell} \binom{\frac{q-1}{2}}{j} h^{\frac{q-1}{2}+\ell-j} \sum_{\substack{i_1,\ldots,i_\ell \ge 0\\ i_1+\ldots+i_\ell=j\\ i_1+2i_2+\ldots+\ell i_\ell=\ell}} \binom{j}{i_1,\ldots,i_\ell} \cdot (H_x^{(1)}h)^{i_1} \cdot \ldots \cdot (H_x^{(\ell)}h)^{i_\ell}.$$

    As $h \in V_{d,d}$, for $k<p$ we have

    $$H_x^{(k)}(h) \in V_{d+pk,d}, \text{ and,}$$
    $$(H_x^{(1)}h)^{i_1} \cdot \ldots \cdot (H_x^{(\ell)}h)^{i_\ell} \in V_{d+p(i_1+2i_2+\ldots+\ell i_\ell),d(i_1+\ldots+i_\ell)}.$$

    Also, $h^{\frac{q-1}{2}+\ell-j} \bmod I = h^{\ell-j} \in V_{(\ell-j)d,(\ell-j)d}$.
    Altogether,

    $$h^\ell H_x^{(\ell)}(h^{\frac{q-1}{2}}) \bmod I \in V_{\max_j\{(\ell-j)d+d+\ell p\},(\ell-j)d+dj} = V_{\ell(p+d),\ell d}.$$

  By linearity, Corollary A.6 and the above items we see that

  $$h^m H_x^{(m)}R \bmod I \in V_{q,B_y+md},$$

  and notice that we work now in a smaller subspace because we may replace x^q with x and $h^{\frac{q-1}{2}}$ with 1. □

  We want to force $h^m H_x^{(m)}R \bmod I \in V_{q,B_y+md}$ to equal 0. The number of constraints this imposes is at most

  $$\sum_{m=0}^{M-1} q(B_y+md) = MqB_y + O(M^2qd).$$

We need the number of degrees of freedom to be larger than the number of constraints. Taking $2B_x = M + E$ we want

$$(M+E)qB_y > MqB_y + O(M^2qd),$$

i.e., $EqB_y = \Omega(M^2qd)$, or $E = \Omega(\frac{M^2d}{B_y}) = \Omega(\frac{M^2d}{p})$.

The number of points $(\alpha,\beta) \in H \cap \mathbb{F}_q^2$ for which $h(\alpha,\beta)$ is a square is at most $\frac{\deg(R)}{M} + d(p+1)$, and

$$\begin{aligned}|\frac{\deg(R)}{M} + d(p+1) - \frac{pq}{2}| &= |\frac{B_xpq + B_y(p+1) + d(p+1)\frac{q-1}{2}}{M} + d(p+1) - \frac{pq}{2}| \\ &= \frac{Epq}{2M} + \frac{dpq}{M} = O(Mdq + \frac{dpq}{M}),\end{aligned}$$

using $M, B_y < p$. Equating the two error terms we get $M = \Theta(\sqrt{p})$.

It follows that the number of points $(\alpha,\beta) \in H \cap \mathbb{F}_q^2$ for which $h(\alpha,\beta)$ is a square is at most $\frac{pq}{2} + O(dp^{2.5})$. We can get the same bound for the number of points $(\alpha,\beta) \in H \cap \mathbb{F}_q^2$ for which $h(\alpha,\beta)$ is a non-square (e.g., by counting the number of squares ch obtains, when c is a non-square in $\mathbb{F}_q$). We therefore conclude that $|\sum_{x\in\mathbb{F}_q} \chi(h(x))|$, which is the difference between the number of squares and non-squares, is at most $O(dp^{2.5})$, completing the proof. □

4.1 Independence

Theorem 4.5. *Suppose $h \in H$ is as in Theorem 4.3. Let $B_y < \frac{p}{2} - (B_x + d)$. Then the elements*

$$\{x^iy^jh(x,y)^{k(q-1)/2} \mid 0 \le i \le B_xq, j \le B_y, k \in \{0,1\}\}$$

are independent.

Proof. Suppose $\sum_{i\le B_xq, j\le B_y, k} c_{i,j,k}x^iy^jh(x,y)^{k(q-1)/2} = 0$. Denote

$$m_k(x,y) = \sum_{i\le B_xq, j\le B_y} c_{i,j,k}x^iy^j.$$

Then $m_0 + m_1h^{(q-1)/2} = 0$ and $m_0^2h = m_1^2h^q = m_1^2h(x^q,y^q)$.

$v_{(0,0)}(y) = p+1$, and, in fact, as $y^p + y = x^{p+1}$ we have:

$$y = \sum_{i=0}^{\infty}(-1)^i x^{p^i(p+1)}. \tag{5}$$

We observe that

$$\begin{aligned} y &= x^{p+1} - x^{p(p+1)} \pmod{x^{q(p+1)}} \\ y^p &= x^{p(p+1)} \pmod{x^{q(p+1)}} \\ y^q &= 0 \pmod{x^{q(p+1)}}. \end{aligned}$$

We have

$$m_0^2 h(x,y) = m_1^2 h(x^q,0) \pmod{x^{q(p+1)}}.$$

Notice that

$$\Delta = m_1^2 h(x^q,0) - m_0^2 h(x,y) \in \mathsf{Span}\{x^i y^j \mid i < (2B_x+d)q, j \leq 2B_y + d < p - (2B_x+d)\},$$

and so by Claim 4.6 that we will shortly prove we have:

$$m_0^2 h(x,y) = m_1^2 h(x^q,0).$$

Now,

- If $m_0 \neq 0$ then
$$h(x,y) = h(x,0)^q \cdot (\frac{m_1(x,y)}{m_0(x,y)})^2,$$
and h is special.
- If $m_0 = 0$. Then $m_1 h^{(q-1)/2} = 0$ and therefore $m_1 = 0$. Therefore, all the coefficients $c_{i,j,k}$ are zero as desired.

□

Claim 4.6. *Let $c \in \mathbb{N}$. Suppose $\Delta \in \mathsf{Span}\{x^i y^j \mid i < cq, j < p - c\}$. Then $\Delta = 0 \pmod{x^{q(p+1)}}$ iff $\Delta = 0$.*

Proof. The elements in $\mathsf{Span}\{x^i y^j \mid i \in \mathbb{N}, j < p\}$ have different v_∞ valuations and are independent in H, so there is a unique way to express $\Delta = \sum_{i<cq,j<p-c} \Delta_{i,j} x^i y^j$. We have

$$\begin{aligned} \Delta &= \sum_{i<cq,j<p-c} \Delta_{i,j} x^i (x^{p+1} - x^{p(p+1)})^j \pmod{x^{q(p+1)}} \\ &= \sum_{i<cq,j<p-c} \Delta_{i,j} x^i (x^{p+1} - x^{p(p+1)})^j \end{aligned}$$

where in the first line we have used Equation (5), and in the second we notice that the degree of Δ in x is smaller than $q(p+1)$, because $i < cq$ and $j < p - c$.

Now, if $\Delta = 0$ then clearly $\Delta (\mathrm{mod} x^{q(p+1)}) = 0$. If $\Delta \neq 0$, order the tuples (i,j) first by j and then by i, and let (i_0, j_0) be the maximal tuple in that order such that $\Delta_{i_0,j_0} \neq 0$. Then, the monomial $x^{j_0 p(p+1)+i_0}$ is obtained in a unique way, and therefore has coefficient $\Delta_{i_0,j_0} \neq 0$. Hence $\Delta (\mathrm{mod} x^{q(p+1)}) \neq 0$. □

5 Proving the Hermitian Curve Weil bound for additive characters

Theorem 5.1. *Assume $p = 2^n$, $q = p^2$, $H = \mathbb{F}_q(x, y) \bmod \varphi$. Let $f \in \mathbb{F}_q[x, y]$ be a degree $d \leq \sqrt{p}$ polynomial such that $d' = -v_\infty(f)$ is odd. Let $\theta \in \{0, 1\}$ and denote*

$$V_\theta = \{(a, b) \in \mathbb{F}_q \times \mathbb{F}_q \mid \varphi(a, b) = 0\ ,\ \mathsf{Tr}_{\mathbb{F}_q/\mathbb{F}_2}(f(a, b)) = \theta\}.$$

Then, $|V_\theta| \leq \frac{p^3}{2}(1 + O(\frac{d}{\sqrt{p}}))$.

Proof. We create a nonzero algebraic function $R \in \mathcal{L}(D \cdot P_\infty)$ such that for each point $P = (\alpha, \beta) \in V_\theta$ for which $\beta \neq 0$, R vanishes on P with multiplicity at least M (i.e., $v_P(R) \geq M$). This gives an upper bound of $\frac{D}{M} + 1$ on the cardinality of V_θ. Specifically,

- We look for $R \in \mathbb{F}_q[x, y]$ in the $\mathbb{F}_q$-vector space

  $$\mathsf{Span}\{x^{i_1} \cdot y^{i_2+i_3p+i_4q} \cdot \mathsf{Tr}(f(x, y))^{i_5} \mid i_1 + i_2 \in [A_1], i_3 \in [A_3], i_4 \in [A_4], i_5 \in \{0, 1\}\}, \tag{6}$$

  where A_1, A_3, A_4 will be specified later. Then, $R \in \mathcal{L}(D \cdot P_\infty)$ with $D = (p + 1)(A_1 + pA_3 + qA_4) + d(p + 1)\frac{q}{2}$.

- Define the ideal $I = \langle y^q - y, \mathsf{Tr}_{\mathbb{F}_q/\mathbb{F}_2}(f(x, y)) - \theta\rangle$ of $\mathbb{F}_q[x, y]$. In Section 5.2 we show that $H_y^{(m)}R \in \mathbb{F}_q[x, y]$. We look for a non-zero polynomial R such that for every $m < M$, $y^m H_y^{(m)}R \in I$ or, equivalently, $y^m H_y^{(m)}R \bmod I = 0$. This in particular guarantees that for every $P = (\alpha, \beta) \in V_\theta$ with $\beta \neq 0$, $H_y^{(m)}R(P) = 0$ which implies that $v_P(R) \geq M$.[6]

Now we calculate:

- **Degrees of freedom:** In Section 5.1 we prove all the elements $x^{i_1} \cdot y^{i_2+i_3p+i_4q} \cdot \mathsf{Tr}(f(x, y))^{i_5}$ appearing in Equation (6) have different valuations at P_∞, and therefore are independent over $\mathbb{F}_q$. Hence, the number of degrees of freedom is $2\binom{A_1+2}{2}(A_3 + 1)(A_4 + 1)$. It is straight forward to see that all the elements $x^{i_1} \cdot y^{i_2+i_3p+i_4q}$ are independent. The crucial thing to notice is that because we may also include $\mathsf{Tr}(f(x, y))$, the dimension of the span doubles.

- **Number of constraints:** For every $0 < m < M$, we would like $y^m H_y^{(m)}R$ to be in the ideal I. In Section 5.2 we prove this can be guaranteed by imposing at most $M(A_3 + M)\frac{(A_1+A_4+(M-1)d+2)^2}{2}$ linear constraints on the coefficients of R.

[6] To see that formally, first note that y is S-useful, for the set S of rational points of H (see Definition A.12). Thus, by Fact A.2, for any point $P = (\alpha, \beta) \in V_\theta$, $H_y^{(m)}(R) = H_{y-\beta}^{(m)}(R)$ and Claim A.1 implies $v_P(R) \geq M$.

We now choose parameters. The number of variables is at least $2(A_4+1)\frac{A_3A_1^2}{2}$. We choose

$$A_4+1 = \lceil (1+4\alpha)\frac{M}{2} \rceil,$$

for some constant parameter $0 < \alpha < \frac{1}{4}$ that we fix later. Thus, the number of variables is at least $(1+4\alpha)\frac{MA_3A_1^2}{2}$. We choose parameters such that $A_3 + M \approx A_3$, $A_1 + A_4 + Md \approx A_1$, say, such that the number of constraints is at most $(1+2\alpha)\frac{MA_3A_1^2}{2}$, and so there are more variables than constraints.

Specifically, we set

$$M = \lfloor \frac{\alpha A_1}{d+1} \rfloor = \lfloor \alpha A_3 \rfloor,$$

The number of linear constraints is upper bounded by $M \cdot (1+\alpha)A_3 \cdot \frac{(1+\alpha)^2A_1^2}{2}$.[7] The claim then follows because $1+4\alpha \geq (1+\alpha)^3$ (for $\alpha \leq \frac{1}{4}$).

We therefore see that $|V_\theta| \leq \frac{D}{M} + 1 = \frac{(p+1)(A_1+pA_3+qA_4)+d(p+1)\frac{q}{2}}{M} + 1$. For the second term, $\frac{d(p+1)q}{2M} \leq \frac{p^3}{2}(\frac{d}{M} + \frac{d}{pM})$. The first term contributes $\frac{p^3}{2}(\frac{2A_4}{M} + \frac{2A_3}{pM} + \frac{2A_1}{p^2M} + \frac{2A_4}{pM} + \frac{2A_3}{p^2M} + \frac{2A_1}{p^3M})$. Notice that $\frac{2A_4}{M} \leq 1+4\alpha$. Thus, basically, $|\ |V_\theta| - \frac{p^3}{2}| \leq p^3(2\alpha + \frac{d}{2M})$ plus lower order terms, and so we set $\alpha = \frac{d}{4M}$. We also set A_1 as large as we can and we choose $A_1 = \Theta(p)$. This means $M^2 = \Theta(A_1) = \Theta(p)$. Thus, $M = \Theta(\sqrt{p})$, $A_3 = \Theta(\frac{M^2}{d}) = \Theta(\frac{p}{d})$, $A_1 = \Theta(p)$. Altogether,

$$d < A_4 \approx \frac{M}{2} < M < A_3 < A_1.$$

This puts the error term at $O(\frac{1}{p} + \frac{d}{M}) = O(\frac{d}{\sqrt{p}})$ as stated. □

5.1 Independence

Lemma 5.2. *Assume $A_1 < \frac{p}{4}$, $A_3, A_4 < p$. The elements*

$$\{x^{i_1} \cdot y^{i_2+i_3p+i_4q} \cdot \mathsf{Tr}(f(x,y))^{i_5}\}_{i_1+i_2 \in [A_1], i_3 \in [A_3], i_4 \in [A_4], i_5 \in \{0,1\}}$$

have different valuations at P_∞ and hence are independent over $\mathbb{F}_q$.

Proof. Let $v = v_\infty$ be the valuation at P_∞. $v(x) = -(p+1)$,$v(y) = -p$. Before we start we first calculate $v(\mathsf{Tr}_{\mathbb{F}_q/\mathbb{F}_2}f(x,y))$. We have: $\mathsf{Tr}_{\mathbb{F}_q/\mathbb{F}_2}f(x,y) = \sum_{j=0}^{2n-1} f(x,y)^{2^j}$. Also, $v(f(x,y)^{2^j}) = 2^j v(f(x,y) = 2^j(-d')$. Hence,

$$v(\mathsf{Tr}_{\mathbb{F}_q/\mathbb{F}_2}f(x,y)) = \min\{2^j \cdot (-d')\}_{j=0}^{2n-1} = -2^{2n-1}d' = -\frac{q}{2}d'.$$

[7] As $M \leq \frac{\alpha}{d+1}A_1$, $A_4 < \frac{1+4\alpha}{2}M < M$ and $M(d+1) \leq \alpha A_1$.

Assume $x^{i_1} \cdot y^{i_2+i_3p+i_4q} \cdot \mathsf{Tr}(f(x,y))^{i_5} = x^{j_1} \cdot y^{j_2+j_3p+j_4q} \cdot \mathsf{Tr}(f(x,y))^{j_5}$, and w.l.o.g. $j_5 \geq i_5$. Set $\Delta_\ell = i_\ell - j_\ell$. Then

$$(\Delta_1, \Delta_2, \Delta_3, \Delta_4, \Delta_5) \cdot (p+1, p, p^2, p^3, \frac{q}{2}d') = 0,$$

and $\Delta_5 \in \{0,1\}$. In particular,

$$d'\Delta_5 \frac{q}{2} \quad = \quad (p+1)\Delta_1 + p\Delta_2 + p^2\Delta_3 + p^3\Delta_4. \tag{7}$$

Assume $\Delta_5 \neq 0$. Then $\Delta_5 = 1$. Taking the equation modulo q (and remembering that d' is odd) we get:

$$\frac{q}{2} \quad = \quad p\left[\Delta_1 + \Delta_2\right] + \Delta_1 (\mathrm{mod}\ q),$$

Now $|\Delta_1 + \Delta_2| \leq 2A_1 < \frac{p}{2}$, and so the RHS of the equation is in the range $(-q/2, q/2)$. A contradiction. Hence $\Delta_5 = 0$. With $\Delta_5 = 0$, Equation (7) becomes $\Delta_1 + (\Delta_1 + \Delta_2)p + \Delta_3 p^2 + \Delta_4 p^3 = 0$. As $|\Delta_1| + |\Delta_2|, |\Delta_3|, |\Delta_4| < p$ this implies $\Delta_1 = \Delta_3 = \Delta_4 = 0$ and $\Delta_1 + \Delta_2 = 0$, hence $(i_1, i_2, i_3, i_4, i_5) = (j_1, j_2, j_3, j_4, j_5)$. □

5.2 The derivatives $H_y^{(m)}$ and the constraints

Let us denote $\mathbf{I}(A_1, A_3, A_4, A_5) = \mathsf{Span}\{x^{i_1} y^{i_2+i_3p+i_4p^2} \mathsf{Tr}(f(x,y))^{i_5} \mid i_1 + i_2 \leq A_1, i_3 \leq A_3, i_4 \leq A_4, i_5 \leq A_5\}$.

Lemma 5.3. *For every $i \in \mathbb{N}$, $m < p$,*

1. $y^m H_y^{(m)}(y^i) \in \mathsf{Span}\{y^i\}$,
2. $y^m H_y^{(m)}(x^i) \in \mathbf{I}(i, \min\{m, i\}, 0, 0)$,
3. $y^m H_y^{(m)} f \in \mathbf{I}(d, \min\{m, d\}, 0, 0)$,
4. *If* $m > 0$, $y^m H_y^{(m)}(\mathsf{Tr}_{\mathbb{F}_q/\mathbb{F}_2} f(x,y)) \in \mathbf{I}(dm, m, 0, 0)$.

Proof. Item 1 follows from $H_y^{(m)}(y^i) = \binom{i}{m} y^{i-m}$. Item Item 2 is Example A.5. For the last two items:

- f has total degree d, hence it is enough to evaluate $y^m H^{(m)}(x^{j_1} y^{j_2})$ for $j_1 + j_2 \leq d$. Using the product rule (Appendix A.1) and the first two items we see that it is in $\mathbf{I}(d, \min\{m, d\}, 0, 0)$,

- We have $\mathsf{Tr}_{\mathbb{F}_q/\mathbb{F}_2} f(x,y) = \sum_{j=0}^{2n-1} f(x,y)^{2^j}$. The Hasse derivative is linear, and therefore we need to evaluate $y^m H_y^{(m)} f(x,y)^{2^j}$. Using Claim A.4 it is non-zero only when $2^j | m$, i.e., $m = 2^j \ell$ for some natural number ℓ, and then $y^m H_y^{(m)}(f(x,y)^{2^j}) = (y^\ell H_y^{(\ell)} f(x,y))^{2^j}$. By the previous items, $y^\ell H_y^{(\ell)}(f(x,y)) \in \mathbf{I}(d, \min\{\ell, d\}, 0, 0)$, and $y^m H_y^{(m)}(f(x,y)^{2^j}) \in \mathbf{I}(d2^j, \min\{\ell, d\}2^j, 0, 0)$. As $\ell 2^j = m$, we see that $y^m H_y^{(m)}(\mathsf{Tr}(f(x,y))) \in \mathbf{I}(dm, m, 0, 0)$.

□

Lemma 5.4. *Suppose $Q \in \mathbf{I}(A_1, A_3, A_4, A_5)$ and $m > 0$. Then $y^m H_y^{(m)} Q \in \mathbf{I}(A_1 + dm, A_3 + m, A_4, A_5)$. Furthermore, each coefficient of $y^m H_y^{(m)} Q$ in the basis $\mathbf{I}(A_1 + dm, A_3 + m, A_4, A_5)$ is a linear combination (over $K = \mathbb{F}_q$) of the coefficients of Q in the basis $\mathbf{I}(A_1, A_3, A_4, A_5)$.*

Proof. It is enough to consider a single monomial $x^{i_1} y^{i_2 + i_3 p + i_4 p^2} \mathsf{Tr}(f(x,y))^{i_5}$. By the product rule it suffices to analyze $y^{m_1} H_y^{(m_1)} x^{i_1}$, $y^{m_2} H_y^{(m_2)} y^{i_2}$, $y^{m_3} H_y^{(m_3)} y^{i_3 p}$, $y^{m_4} H_y^{(m_4)} y^{i_4 p^2}$, $y^{m_5} H_y^{(m_5)}(\mathsf{Tr}(f(x,y))^{i_5}$ where $m_1 + m_2 + m_3 + m_4 + m_5 = m$. By Lemma 5.3 we see that

$$y^m H_y^{(m)}(x^{i_1} y^{i_2 + i_3 p + i_4 p^2} \mathsf{Tr}(f(x,y))^{i_5}) \in \mathbf{I}(i_1 + i_2 + dm, i_3 + m, i_4, i_5).$$

This shows that $y^m H_y^{(m)} Q \in \mathbf{I}(A_1 + dm, A_3 + m, A_4, A_5)$. It is also clear from the computation we have done that each coefficient of $y^m H_y^{(m)} Q$ is a linear combination (over $K = \mathbb{F}_q$) of the coefficients of Q. □

Say $R \in \mathbf{I}(A_1, A_3, A_4, A_5)$. We saw $y^m H^{(m)} R \in \mathbf{I}(A_1 + dm, A_3 + m, A_4, A_5)$, and each coefficient of $y^m H^{(m)} R$ is some linear combination of the coefficients of R. As $y^q = y \bmod I$ and $\mathsf{Tr}_{\mathbb{F}_q/\mathbb{F}_2} f(x,y) = \theta \bmod I$,

$$y^m H^{(m)} R \bmod I \quad \in \quad \mathbf{I}(A_1 + A_4 + dm, A_3 + m, 0, 0)$$

Thus, to force $y^m H^{(m)} R \bmod I = 0$ it is enough to force $\dim(\mathbf{I}(A_1 + A_4 + dm, A_3 + m, 0, 0))$ linear equations to be zero, which amounts to $\binom{A_1 + A_4 + dm + 2}{2} \cdot (A_3 + m + 1)$ linear equations. Altogether, we have $\sum_{k=0}^{M-1} \binom{A_1 + A_4 + dm + 2}{2} \cdot (A_3 + m + 1)$ linear equations which we upper bound by $M(A_3 + M)\frac{(A_1 + A_4 + (M-1)d + 2)^2}{2}$ as desired.

References

[ABN+92] Noga Alon, Jehoshua Bruck, Joseph Naor, Moni Naor, and Ron M Roth. Construction of asymptotically good low-rate error-correcting codes through pseudo-random graphs. *Information Theory, IEEE Transactions on*, 38(2):509–516, 1992.

[AGHP92] Noga Alon, Oded Goldreich, Johan Håstad, and René Peralta. Simple constructions of almost k-wise independent random variables. *Random Structures & Algorithms*, 3(3):289–304, 1992.

[BATS13] Avraham Ben-Aroya and Amnon Ta-Shma. Constructing small-bias sets from algebraic-geometric codes. *Theory OF Computing*, 9(5):253–272, 2013.

[Bog12] A. Bogdanov. A different way to improve the bias via expanders. Topics in (and out) the theory of computing, Lecture 12, 2012.

[Bom74] Enrico Bombieri. Counting points on curves over finite fields: d'après sa stepanov. In *Séminaire Bourbaki vol. 1972/73 Exposés 418–435*, pages 234–241. Springer, 1974.

[CDGM25] Gil Cohen, Dean Doron, Noam Goldgraber, and Tomer Manket. Tracing AG codes: Toward meeting the Gilbert–Varshamov bound. Private communication, 2025.

[Gil52] Edgar N Gilbert. A comparison of signalling alphabets. *The Bell system technical journal*, 31(3):504–522, 1952.

[Gol03] David M. Goldschmidt. *Algebraic Functions and Projective Curves*, volume 215. Springer New York, NY, 2003.

[HB96] DR Heath-Brown. An estimate for heilbronn's exponential sum. *PROGRESS IN MATHEMATICS-BOSTON-*, 139:451–464, 1996.

[Jeo11] Sangtae Jeong. Calculus in positive characteristic p. *Journal of Number Theory*, 131(6):1089–1104, 2011.

[Mas84] R. C. Mason. *Diophantine Equations over Function Fields.* London Mathematical Society Lecture Note Series. Cambridge University Press, 1984.

[Mit92] Dmitrii Alekseevich Mit'kin. Stepanov method of the estimation of the number of roots of some equations. *Mathematical Notes*, 51(6):565–570, 1992.

[Sch73] Wolfgang Schmidt. Zur methode von stepanov. *Acta Arithmetica*, 24(4):347–367, 1973.

[Sch74] Wolfgang M Schmidt. A lower bound for the number of solutions of equations over finite fields. *Journal of Number Theory*, 6(6):448–480, 1974.

[Sch06] Wolfgang M Schmidt. *Equations over finite fields: an elementary approach*, volume 536. Springer, 2006.

[Ste69] Sergei Aleksandrovich Stepanov. On the number of points of a hyperelliptic curve over a finite prime field. *Mathematics of the USSR-Izvestiya*, 3(5):1103, 1969.

[Ste70] Sergei Aleksandrovich Stepanov. An elementary method in algebraic number theory. *Acta Arith.*, 17:231–247, 1970.

[Ste72a] Sergei Aleksandrovich Stepanov. Congruences in two unknowns. *Izvestiya: Mathematics*, 6(4):677–704, 1972.

[Ste72b] Sergei Aleksandrovich Stepanov. An elementary proof of the hasse-weil theorem for hyperelliptic curves. *Journal of number theory*, 4(2):118–143, 1972.

[Ste73] Sergei Aleksandrovich Stepanov. On estimating rational trigonometric sums with prime denominator. *Collection of Articles Dedicated to Academician IM Vinogradov on the Eightieth Anniversary of His Birth*, 1:358, 1973.

[Sti09] Henning Stichtenoth. *Algebraic function fields and codes*, volume 254. Springer Science & Business Media, 2009.

[Tor00] Fernando Torres. The approach of stöhr-voloch to the hasse-weil bound with applications to optimal curves and plane arcs, 2000.

[TS17] Amnon Ta-Shma. Explicit, almost optimal, epsilon-balanced codes. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 238–251. ACM, 2017.

[TVZ82] Michael A Tsfasman, SG Vlădutx, and Th Zink. Modular curves, shimura curves, and goppa codes, better than varshamov-gilbert bound. *Mathematische Nachrichten*, 109(1):21–28, 1982.

[Var57] RR Varshamov. Estimate of the number of signals in error correcting codes, dokl. acad. nauk. sssr 117. *Dokl. Acad. Nauk. SSSR 117, in Russian*, 117:739–741, 1957.

[Vla96] Serge G Vladut. Two applications of weil-serre's explicit formula. *Applicable Algebra in Engineering, Communication and Computing*, 7(4):279–288, 1996.

[Wei48] André Weil. *Sur les courbes algébriques et les variétés qui s' en déduisent.* FeniXX, 1948.

A Some mathematical background

A.1 Hasse derivatives

Let F/K be a function field. $z \in F$ is called *separating* if $F/K(z)$ is a finite separable extension. If $v_P(z) \neq 0$ for some place P then z is separating. Furthermore, if K has characteristic p, then z is separating iff z is not a p power of some element in F [Sti09, Proposition 3.10.2].

Let $z \in F$ be separating. The m-th Hasse derivative with respect to z, denoted by H_z^m, is defined on $K[z]$ by the K-linear extension of $H_z^m(z^n) \triangleq \binom{n}{m} z^{n-m}$ to all of $K[z]$. H_z^m can be uniquely extended to all of F/K so that: (1) $H_z^0 = id_F$, (2), H_z^m vanish on K for all $m > 0$, (3) $H_z^m(fg) = \sum_{j=0}^m H_z^j(f) H_z^{m-j}(g)$ (Product Rule) and $H_z^m \circ H_z^n = \binom{m+n}{m} H_z^{m+n}$ (Composition Rule). These uniquely determined extensions are called the *Hasse derivatives*.

The following claim, which is an immediate corollary of [Gol03, Corollary 2.5.14 (Taylor's Theorem)], shows Hasse derivatives capture multiplicity:

Claim A.1. *Let P be a place in F/K and let t be a separating element with $v_P(t) = 1$. Let $f \in F$ and $M \in \mathbb{N} \setminus \{0\}$, then $v_P(f) \geq M \iff \forall m < M \quad (H_t^m(f))(P) = 0$.*

A simple fact is

Fact A.2 (Simple change of variable)**.** *Let z be a separating element in F/K, let $f \in F$ and $\alpha \in K$, then for every $m > 0$, $H_z^m(f) = H_{z-\alpha}^m(f)$.*

The Hasse chain rule is:

Fact A.3. *Suppose $f(t) \in K((t))$. Suppose $t = t(y) \in K((y))$. Let $m \geq 0$. Then*

$$\begin{aligned} H_y^{(m)}(f(t)) &= H_y^{(m)}(f \circ t(y)) \\ &= \sum_{j=1}^{m} H_t^{(j)}(f(t)) \cdot \sum_{\substack{i_1,\ldots,i_m \geq 0 \\ i_1+\ldots+i_m=j \\ i_1+2i_2+3i_3+\ldots+mi_m=m}} \binom{j}{i_1,\ldots,i_m} \cdot (H_y^{(1)}t)^{i_1} \cdot \ldots \cdot (H_y^{(m)}t)^{i_m} \end{aligned}$$

The following lemma tells us how Hasse derivatives behave on p-th powers:

Claim A.4. *[Tor00, Remark 2.4 and Remark 2.5], [Jeo11, Theorem 3.1] Let $z \in F/K$ be a separating element of a function field of characteristic $p > 0$. Let $q = p^k$ be a power of p and $f \in F$. If q divides m then $H_z^m(f^q) = (H_z^{m/q}(f))^q$, and otherwise $H_z^m(f^q) = 0$.*

Furthermore, the following conditions are equivalent,

1. *There exists some $g \in F$ such that $f = g^q$,*
2. *$H_z^m(f) = 0$ for all $m = 1, ..., q-1$,*
3. *$H_z^1(f) = H_z^p(f) = H_z^{p^2}(f) = ... = H_z^{p^{k-1}}(f) = 0$.*

Example A.5. *We look at $y^m H_y^{(m)}(x^i)$ in the Hermitian function field.*

- *We start with the $i = 1$ case. On the one hand $H_y^{(\ell)} y^{p+1} = \binom{p+1}{\ell} y^{p+1-\ell}$ and on the other hand, $H_y^{(\ell)}(y^{p+1}) = H_y^{(\ell)}(x^p + x) = H_y^{(\ell)}(x)$ (as by Claim A.4, $H_y^{(\ell)} x^p = 0$). Hence $y^\ell H_y^{(\ell)}(x) = \binom{p+1}{\ell} y^{p+1}$.*

- *Next we use Fact A.3 (with $t = x$ and $f(t) = x^i$) to understand general i. For every $1 \leq j \leq \min\{m, i\}$, $H_x^{(j)}(x^i) \in \mathsf{Span}\{x^{i-j}\}$, $y^\ell H_y^{(\ell)} x \in \mathsf{Span}\{y^{p+1}\}$, and for every $i_1 + \ldots + i_m = j$ such that $i_1 + 2i_2 + \ldots mi_m = m$, we have $(yH_y^{(1)}x)^{i_1} \cdot (y^2 H_y^{(2)} x)^{i_2} \ldots \cdot (y^m H_y^{(m)} x)^{i_m} \in \mathsf{Span}\{y^{j(p+1)}\}$. Altogether $y^m H_y^{(m)}(x^i) \in \mathsf{Span}\{x^{\ell_1} y^{\ell_2} y^{p\ell_3} \mid \ell_1 + \ell_2 \leq i, \ell_3 \leq \min\{m, i\}\}$.*

The following corollary will be useful to us later on:

Corollary A.6. *Let $z \in F/K$ be a separating element of a function field of characteristic $p > 0$. Let q be a power of p, $m < q$ and $f, g \in F$, then $H_z^m(fg^q) = H_z^m(f)g^q$.*

Proof. Since $m < q$, all of the derivatives of g^q that will appear in the expansion of $H_z^m(fg^q)$ by means of the product rule are zero (from Claim A.4) except for the term $H_z^m(f)g^q$. □

A.2 Differentials

We follow the presentation at [Sti09, Chapter 4]. Let F/K be a function field. A *derivation* is a K-linear map D from F to some F-module that upholds the product rule of derivatives, i.e. $D(fg) = fD(g) + gD(f)$. For example $H_z^1 : F \to F$ is a derivation for every separating $z \in F$. One can define an F-module Δ_F, such that all derivations of F factor through Δ_F via a canonical mapping $d : F \to \Delta_F$, i.e., if $\delta : F \to M$ is a derivation of F into some F-module M, then there exists a unique F-linear map $\mu : \Delta_F \to M$ such that $\delta = \mu \circ d$. It turns out that d is itself a derivation. For $x \in F$, $d(x)$ is called the differential associated with x and is denoted dx. The set Δ_F of differentials of F contains all elements udx where $u \in F$ and x is separating, where this set is taken modulo the equivalence relation $udx = vdy$ if and only if $\frac{u}{v} = D_x^1(y)$. With this we get a notion of division of differentials via $\frac{udy}{vdx} = \frac{u}{v} D_x^1(y)$.

We now define the notion of valuation of differentials. If $P \in \mathbb{P}_F$ is a place of F and $udx \in \Delta_F$ a differential of F, we define $v_P(udx)$ as follows. We pick a local parameter t of P (i.e., $v_P(t) = 1$) and we find $b \in F$ such that $udx = bdt$. Then $v_P(udx) = v_P(b)$. One can show that this definition is independent of the specific choice of local parameter t. As differentials have valuations, differentials also have zeroes and poles, i.e., places where the valuation is strictly positive or strictly negative, and this can be encoded in a divisor, denoted (udx) and called the divisor associated with the differential udx. It turns out any differential $udx \in \Delta_F$ has only finitely many zeroes or poles and so the associated divisor is indeed well defined. A divisor which is associated to some differential in Δ_F is called a *canonical divisor*. All canonical divisors have degree $2g - 2$ and their Riemann-Roch space has dimension g where $g = genus(F)$. The following claims will be useful:

Claim A.7. *Let $u \in F$, $x, y \in F$ separating, then:*

- $(udx) = (u) + (dx)$

- $(D^1_x(y)) = (\frac{dy}{dx}) = (dy) - (dx)$

The first bullet is a restatement of [Sti09, Proposition 1.5.13] and the second one is an immediate consequence of the first bullet and the equality $1dy = D^1_x(y)dx$.

There is a close relationship between the zeroes and poles of f and the zeroes and poles of df. The following claim is stated in [Mas84, Chapter I (6)] for the case where K is algebraically closed, but by considering a constant-field extension of F one can verify it holds exactly as stated for any perfect base field K and even when P is not a place of degree one (which is not a consideration when K is algebraically closed). We give the proof for completeness.

Claim A.8. *Let $f \in F/K$, df its associated differential, then:*

- *For every place P, $v_P(df) \geq v_P(f) - 1$. In particular, If f has zeroes at P, df can lose at most one zero at P. Also, df can have at most one more pole at P. Also,*
- *If $v_P(f) \geq 0$ then $v_P(df) \geq 0$, i.e., df can have poles only at places where f has poles.*

Proof of claim A.8. Let $f \in F/K$, $P \in \mathbb{P}_F$, $t \in F$ with $v_P(t) = 1$. We are interested in $(df)^F$ the divisor of df over F. We move to $\bar{F}/\bar{K} = \bar{K} \cdot F/K = \bar{K}F/\bar{K}$ which is the constant field extension of F/K with all of $\bar{K}$, the algebraic closure of K. Let $\bar{P} \in \mathbb{P}_{\bar{F}}$ be a place lying over P. Since $\bar{K}$ is algebraically closed we know $\deg(\bar{P}) = 1$. From [Sti09, Theorem 3.6.3] we learn that $v_{\bar{P}}(t) = v_P(t) = 1$ and we can write:

$$f = \sum_{n=n_0 \in \mathbb{Z}}^{\infty} c_n t^n \quad (c_n \in \bar{k}, c_{n_0} \neq 0)$$

$$D^1_t(f) = \sum_{n=n_0 \in \mathbb{Z}}^{\infty} n \cdot c_n t^{n-1} \quad (c_n \in \bar{k}, c_{n_0} \neq 0)$$

From the definition of valuation for differential we know that

$$v_{\bar{P}}(1df) = v_{\bar{P}}(D^1_t(f)dt) = v_{\bar{P}}(D^1_t(f)) \geq n_0 - 1$$

From [Sti09, Theorem 3.6.3] we know that $v_{\bar{P}}(1df) = v_P(1df)$ and $v_P(f) = v_{\bar{P}}(f)$ and so we get $v_P(1df) \geq v_P(f) - 1$. Furthermore, if f has no pole at P, then f has no pole at $\bar{P}$ and so $n_0 \geq 0$ and so $D^1_t(f)$ has no pole at $\bar{P}$ and therefore at P, meaning $v_P(1df) \geq 0$. □

We mentioned earlier that $\deg(df) = 2g-2$ for canonical divisors (df), and therefore when g is large df has many more zeroes than f. We also know that df has a zero wherever f has a zero of multiplicity at least 2. Finding the other zeroes of df is a bit more complicated. It turns out that:

Claim A.9. *[Sti09, Sections 3.4 and 3.5] The zeroes of df are either at places that are zeroes of f, or, at places of F/K that are ramified when F is viewed as an extension of $K(f)/K$.*

A.3 Kummer extensions

An algebraic function field F'/K' is a *Kummer extension* of F/K if K contains a primitive n-th root of unity,[8] and $F' = F(Z) \mod (Z^n - u)$ where $u \in F$ and $u \neq w^d$ for all $w \in F$ and $d|n$ such that $d > 1$. Kummer extensions are Galois. For $P \in \mathbb{P}_F$ we denote by $\mathcal{L}(P) \overset{\text{def}}{=} \cup_{m\in\mathbb{N}}\mathcal{L}(m \cdot P)$ the K-linear, infinite dimensional vector space of all function that have poles only at P. The following claim follows from [Sti09, Corollary 3.7.4 and Proposition 3.11.1] and the Hurwitz Genus Formula [Sti09, Theorem 3.4.13]:

Claim A.10. *Let P_∞ be a degree one place of a function field F/K of genus g, ℓ a prime number and $u \in \mathcal{L}(P_\infty)$ which is not an ℓ-th power in F. Denote $d = \deg(u) = -v_{P_\infty}(u)$ and assume d is co-prime to ℓ. Let $F' = F(Z)$ where $Z^\ell = u$ be the Kummer extension with respect to u. Then:*

- *F' is a degree ℓ extension of F*
- *P_∞ is totally ramified in F'.[9] Also K is the full constant field of F'.*
- *$Z \in \mathcal{L}(P'_\infty)$ and $\deg(Z) = d$*
- *$g' \overset{\text{def}}{=} \text{genus}(F')$ satisfies $\ell(g-1) \leq g' - 1 \leq \ell(g-1) + d$*

From now on we assume that K is a finite field, $K = \mathbb{F}_q$ for some prime power q. Let ℓ be a prime number that divides $q-1$. Let P_∞ be a degree one place, and S a set of degree one places of F that does not contain P_∞. Let $u \in \mathcal{L}(P_\infty) \subset F$ such that u is not an ℓ-th power in F. We are interested in the number of $P \in S$ such that $u|_P \overset{\text{def}}{=} \phi_P(u) \in K$ is an ℓ-th power as an element of K (where ϕ_P is the evaluation function at P).[10] The following claim is standard and we omit its proof:

Claim A.11. *Suppose $u|_P \neq 0$ for some degree one place $P \in S$. Then:*

- *If $u|_P$ is not an ℓ-th power in K, then there is a single place above P in F' and it is a place of degree ℓ (and ramification 1).*
- *If, however, $u|_P$ is a non-zero ℓ-th power in K, then the place P is totally split in F', i.e. there are ℓ distinct degree one places above P (that have all ramification 1).*

Definition A.12. *Let F be a function field with constant field K. Let S be a set of degree one places of F/K. We say $z \in F$ is* derivative-useful *for S, or, in short, S-useful, if for every $P \in S$ there exists $\alpha \in K$ such that $v_P(z - \alpha) = 1$.*

[8] When $K = \mathbb{F}_q$ this means $n|(q-1)$.

[9] We remind the reader that this means that P'_∞ is the only place of F' above P_∞, has degree one and its ramification index over P_∞ is ℓ.

[10] Notice that $u|_P$ is defined because u does not have a pole at any $P \in S$, and $u|_P \in K$ because any $P \in S$ is degree one.

The term "S-useful" is not standard and does not hold any deeper meaning then saying z "works" for every place of S in the sense discussed above. We have:

Claim A.13. *Let $F' = F(Z) \mod (Z^\ell - u)$ be a Kummer extension with ℓ prime and $u \in \mathcal{L}(P_\infty) \subset F$ such that u is not an ℓ'th power of an element in F. Further assume K is the full constant field of F'. Let S be a set of degree one places F/K and assume $P_\infty \notin S$. Suppose $S_\ell \subset S$ is such that $u|_P$ is a non-zero ℓ-th power for all $P \in S_\ell$, and let S'_ℓ be the set of all places of F' lying over S_ℓ.*

Then:

- *If $Q \in S'_\ell$ lying above P then $e(Q|P) = 1$,*
- *If $x \in F$ is S_ℓ-useful then x when considered as an element of F' is S'_ℓ-useful.*

Proof. Let $Q \in S'_\ell$ and denote $P \in S_\ell$ the place of F lying below Q. Since $u|_P$ is a non-zero ℓ-th power, P is totally split in F' (by claim A.11). This means there are ℓ places lying over P (Q among them), each of them with relative degree one and ramification 1 over P. Since $x \in F$ is S-useful and $P \in S_\ell \subset S$ there exists an $\alpha \in K$ such that $v_P(X - \alpha) = 1$. Since Q is lying over P and has ramification 1 we get $v_Q(X - \alpha) = e(Q|P)v_P(X - \alpha) = 1 \cdot 1 = 1$. So for any $Q \in S'_\ell$ we found $\alpha \in K$ with $v_Q(X - \alpha) = 1$ completing the proof. □

A.4 Hyper-elliptic curve

Let $F = \mathbb{F}_q[x]$ and $\widehat{F}$ be as in Equations (1) and (2), and further assume q is odd. Let $h \in \overline{\mathbb{F}}_q[x]$ be a square-free polynomial of odd degree d and

$$\mathcal{C} = \{(x, z) \in \mathbb{F}_q^2 \mid z^2 - h(x) = 0\}.$$

By [Sti09, Proposition 6.2.3(b)] the genus g of $\mathcal{C}$ is $\frac{d-1}{2}$. Let P_∞^F be the unique pole of x in F. By [Sti09, Proposition 6.2.3(c)] P_∞^F totally ramifies in $\mathcal{C}/F$. Let P_∞ be the unique place of $\mathcal{C}$ that lies above P_∞^F and v_∞ its corresponding valuation function. Then $v_\infty(x) = -2$ and $2v_\infty(z) = v_\infty(z^2) = v_\infty(h(x)) = dv_\infty(x) = -2d$, i.e., $v_\infty(z) = -d$.

A.5 Extending the Hermitian curve

Suppose p is a prime power and $q = p^2$. Let $F = \overline{\mathbb{F}}_q(x, y) \bmod \phi(x, y)$ where $\phi(x, y) = y^p + y - x^{p+1}$. F is called the Hermitian function field. There is a unique place P_∞ where x and y have poles. Let v_∞ denote the valuation function at P_∞. $v_\infty(x) = -p$, $v_\infty(y) = -(p+1)$ and genus$(F) = \frac{p(p-1)}{2}$ (see [Sti09, Lemma 6.4.4]).

Let $h \in \overline{\mathbb{F}}_q[x, y]$ be a polynomial with $v_\infty(h) = -w$ for some odd w. Let

$$\mathcal{C} = \{(x, z) \in \overline{\mathbb{F}}_q^2 \mid z^2 - h(x) = 0\}.$$

P_∞ totally ramifies in $\mathcal{C}/F$, because if $P'_\infty|P_\infty$ with valuation v'_∞, then $v'_\infty(z^2) = e(P'_\infty|P_\infty)w$, and as w is odd, $2|e(P'_\infty|P_\infty)$. Let P'_∞ denote the unique place in $\mathcal{C}$ over P_∞, and v'_∞ its associated valuation function. Then, $v'_\infty(x) = -2p$, $v'_\infty(y) = -2(p+1)$ and $v'_\infty(z) = -w$. Also, clearly, $\{1, z\}$ is a basis for $\mathcal{C}/F$ and therefore by [Sti09, Proposition 3.11.1] $\text{genus}(\mathcal{C}) \leq 1 + 2(\text{genus}(F) - 1) + \deg(z) \leq p^2 + w$.