

# Beating the Gilbert–Varshamov Bound at Alphabet Size 42

Gil Cohen\*      Tomer Manket†

October 24, 2026

## Abstract

The question of whether the Gilbert–Varshamov bound (1952, 1957) gives the optimal tradeoff between rate and distance is one of the most fundamental in coding theory. It is a central instance of the broader question of how far algebraic structure can surpass the benchmark provided by the probabilistic method.

In a celebrated breakthrough, Tsfasman, Vlăduț, and Zink (1982) combined Goppa’s construction of algebraic-geometric codes, introduced a year earlier, with deep results from algebraic geometry and number theory to beat the asymptotic Gilbert–Varshamov bound over sufficiently large fields, starting at size 49. Building on their work, Zinoviev and Litsyn (1985) showed that nonlinear codes beat this bound over every alphabet of size at least 46.

We prove the existence of a family of codes beating the asymptotic Gilbert–Varshamov bound over an alphabet of size 42, improving a threshold that has stood for over forty years. Like Zinoviev and Litsyn, we restrict the allowed symbols in each coordinate of an algebraic-geometric code. We then improve the distance by removing relatively few codewords. This is made possible by averaging the number of surviving close pairs over divisor classes and using the zeta function of the underlying function field to bound this average.

---

\*Tel Aviv University. [gil@tauex.tau.ac.il](mailto:gil@tauex.tau.ac.il). Supported by the European Research Council (ERC) Consolidator Grant No. 101231313, and by the Israel Science Foundation grant 2989/24.

†Tel Aviv University. [tomermanket@mail.tau.ac.il](mailto:tomermanket@mail.tau.ac.il). Supported by the European Research Council (ERC) Consolidator Grant No. 101231313.

# Contents

|          |                                                                |           |
|----------|----------------------------------------------------------------|-----------|
| <b>1</b> | <b>Introduction</b>                                            | <b>1</b>  |
| 1.1      | Our result . . . . .                                           | 3         |
| 1.2      | Related work . . . . .                                         | 5         |
| <b>2</b> | <b>Proof overview</b>                                          | <b>6</b>  |
| <b>3</b> | <b>Preliminaries</b>                                           | <b>10</b> |
| <b>4</b> | <b>Proof of the main theorem</b>                               | <b>11</b> |
| 4.1      | Alphabet restriction and expurgation . . . . .                 | 12        |
| 4.2      | Average weight distributions via the zeta function . . . . .   | 13        |
| 4.3      | Zeta functions and class numbers of optimal families . . . . . | 17        |
| 4.4      | A general asymptotic rate bound . . . . .                      | 20        |
| 4.5      | Choosing the parameters . . . . .                              | 23        |
| <b>A</b> | <b>A friendly exposition of algebraic-geometric codes</b>      | <b>32</b> |
| A.1      | An introduction to algebraic function fields . . . . .         | 33        |
| A.2      | Goppa codes and optimal towers . . . . .                       | 46        |
| A.3      | Further tools from function-field theory . . . . .             | 50        |

# 1 Introduction

The modern theory of error-correcting codes grew out of Shannon’s work on reliable communication over noisy channels [Sha48] and Hamming’s work on detecting and correcting errors [Ham50]. A code  $C \subseteq \Sigma^n$  is a collection of words of length  $n$  over a finite alphabet  $\Sigma$ , used to represent messages. Its *rate*,  $R = \log_q |C|/n$ , where  $q = |\Sigma|$ , measures the amount of information carried per transmitted alphabet symbol. Its *relative minimum distance*,  $\delta$ , is the smallest fraction of coordinates on which two distinct codewords differ. A large rate means little redundancy, while a large distance allows the correction of many errors.

The tradeoff between distance and rate is perhaps the most fundamental question in coding theory, and certainly one of its oldest. Since the 1950s, the Gilbert–Varshamov (GV) bound [Gil52, Var57] has been a central benchmark for this tradeoff. For every integer  $q \geq 2$  and every  $0 < \delta < 1 - 1/q$ , there are arbitrarily long  $q$ -ary codes of relative distance at least  $\delta$  and rate at least

$$\text{GV}_q(\delta) - o(1) \quad \text{where} \quad \text{GV}_q(\delta) = 1 - H_q(\delta), \quad (1.1)$$

and

$$H_q(\delta) = \delta \log_q(q-1) - \delta \log_q \delta - (1-\delta) \log_q(1-\delta)$$

is the  $q$ -ary entropy function.

Gilbert [Gil52] established the bound for general codes through a greedy packing argument, while Varshamov [Var57] obtained a refinement for linear codes. These bounds also follow from elementary probabilistic arguments: a random linear code attains the bound by a direct union-bound argument, while for arbitrary alphabets one can use random sampling followed by expurgation (removing a small fraction of the sampled words to eliminate pairs that are too close). These arguments place the GV bound within the early tradition of probabilistic existence proofs pioneered by Erdős in combinatorics [Erd47] and Shannon in information theory [Sha48].

Despite the simplicity of these arguments, the GV bound resisted asymptotic improvement for roughly three decades. As Tsfasman, Vlăduț, and Nogin recall in their preface [TVN07, p. ix],<sup>1</sup> the repeated failure to improve the bound had led to a widespread,

---

<sup>1</sup>Their account includes a charming anecdote: in early 1981, Goppa presented his connection between algebraic geometry and coding theory at the algebra seminar of Moscow State University. Manin attended the talk and subsequently explained the ideas in his own seminar, where Tsfasman encountered coding theory for the first time. Only a few months later, Tsfasman, Vlăduț, and Zink obtained their improvement over the GV bound. Initially unsure whether the result even merited publication, they were surprised by the enthusiastic response; only then did they learn of the long history of unsuccessful attempts to improve the bound.

though not unanimous, belief in its tightness. In our notation, this belief was that the largest achievable asymptotic rate  $R_q(\delta)$  might equal  $1 - H_q(\delta)$ . Thus, the question was whether the elementary existence arguments already captured the ultimate limit of error-correcting codes.

A breakthrough came in 1982, when Tsfasman, Vlăduț, and Zink [TVZ82], building on Goppa’s construction of codes from algebraic curves [Gop81], showed that the GV bound is not always tight. Their proof brought exceptionally deep algebraic geometry and number theory into coding theory. Goppa’s construction uses the Riemann–Roch theorem to turn curves with many rational points relative to their genus into good codes. The requisite families of curves with many rational points were obtained using the arithmetic theory of modular and Shimura curves [TVZ82, Iha81]. Through this connection, deep results about algebraic curves yield codes surpassing the benchmark provided by the probabilistic method. For square prime powers  $q$ , the resulting TVZ bound is

$$R_q(\delta) \geq \text{TVZ}_q(\delta) := 1 - \delta - \frac{1}{\sqrt{q} - 1}. \quad (1.2)$$

This line lies strictly above the GV curve on a nonempty interval of distances whenever  $q \geq 49$  (see Figure 1). Moreover, the TVZ bound can be attained by polynomial-time constructible, and hence efficiently encodable, linear codes; such constructions were established by Vlăduț, Katsman, and Tsfasman [VKT84].

Later, Garcia and Stichtenoth [GS95, GS96] gave a substantially simpler realization of the same bound, constructing asymptotically optimal towers of function fields defined by explicit recursive equations. Although their construction and analysis still require substantial machinery from the theory of algebraic function fields, they avoid the additional machinery of modular and Shimura curves, making the algebraic-geometric route to beating GV more accessible.

The alphabet size is crucial to this phenomenon. At  $q = 49$ , the TVZ line lies only slightly above the GV curve on an interval of distances; for smaller fields, it does not exceed it anywhere. In 1985, Zinoviev and Litsyn [ZL85] showed that the TVZ bound, combined with an elementary alphabet-reduction argument, implies the existence of nonlinear codes exceeding GV for every integer  $q \geq 46$ . The idea is to start with a TVZ code over a larger alphabet and randomly restrict the allowed symbols in each coordinate. An averaging argument guarantees a choice of restrictions that retains sufficiently many codewords while preserving their pairwise distances. The resulting loss in rate is small enough that the improvement supplied by TVZ survives at alphabet size 46. This averaging argument establishes existence but does not provide an efficient encoding algorithm for the resulting codes.

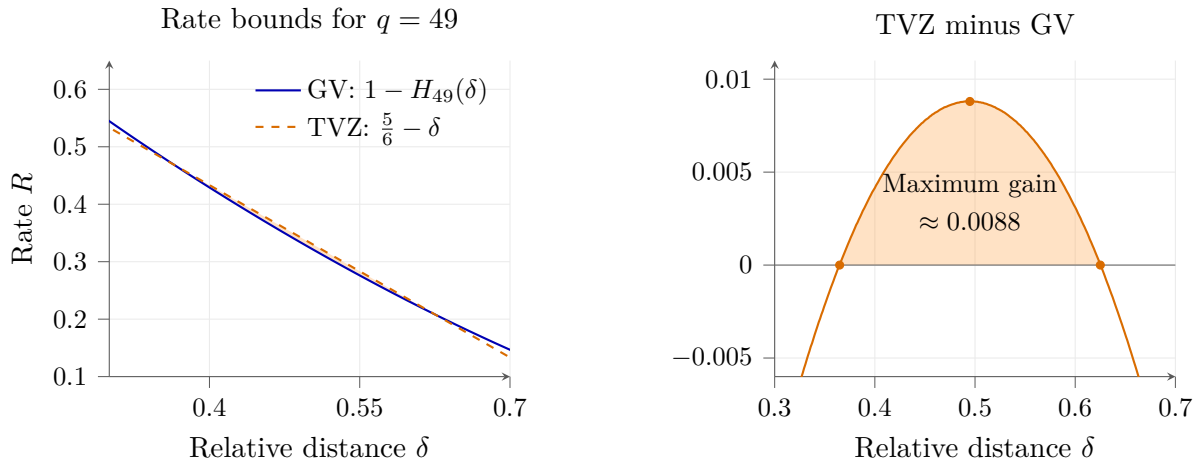


Figure 1: The Gilbert–Varshamov and Tsfasman–Vlăduț–Zink bounds for alphabet size  $q = 49$ . The right panel shows their difference,  $\Delta R = H_{49}(\delta) - \delta - 1/6$ ; the shaded region is where TVZ exceeds GV, approximately  $0.365 < \delta < 0.625$ . The maximum gain is approximately 0.0088, attained at  $\delta = 48/97$ .

The TVZ breakthrough, followed by the explicit towers of Garcia and Stichtenoth, helped shape a rich line of research on algebraic-geometric codes. Alongside these developments, a central question has remained: how small can the alphabet be while still admitting an asymptotic improvement over the GV bound? This is an instance of a broader question about the power of algebraic structure to surpass the benchmark provided by the probabilistic method: over which alphabets can carefully arranged codewords be packed asymptotically more densely than this benchmark guarantees?

Despite these advances and many further contributions (see [Section 1.2](#)), alphabet size 46 has, to the best of our knowledge, remained the frontier for beating the asymptotic GV bound. The question of how far this phenomenon extends toward smaller alphabets remains unresolved, with the binary case being the most prominent instance.

## 1.1 Our result

In this work, we lower the alphabet-size threshold for beating the asymptotic GV bound from 46 to 42, improving the result of Zinoviev and Litsyn after more than four decades.

**Theorem 1.1.** *There exists a universal constant  $\beta > 0$  such that, for every  $\delta \in [0.39, 0.48]$ , there exists a family of codes over an alphabet of size 42 with asymptotic relative distance at least  $\delta$  and asymptotic rate at least*

$$1 - H_{42}(\delta) + \beta.$$

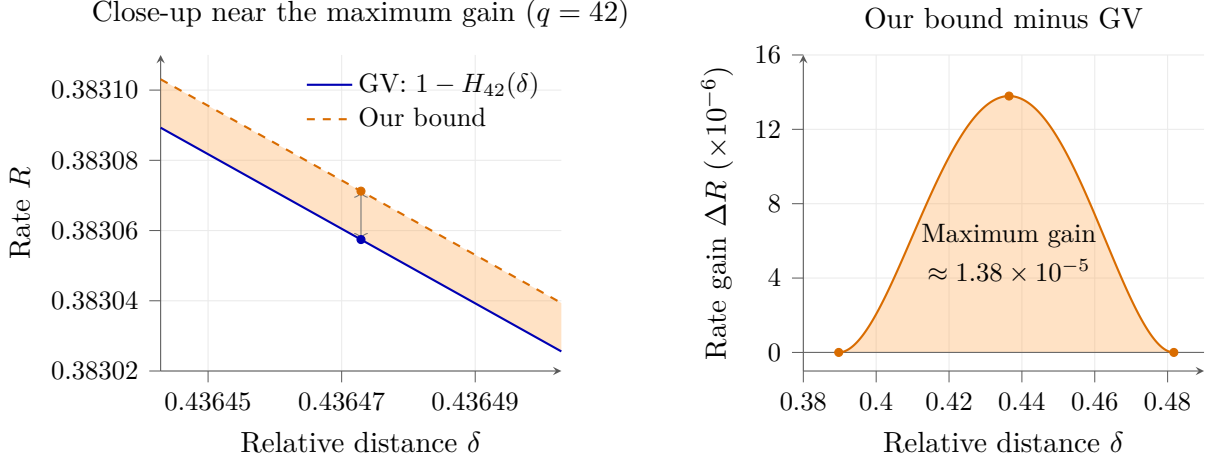


Figure 2: The Gilbert–Varshamov bound and our optimized asymptotic rate bound for alphabet size  $q = 42$ . The left panel zooms in near  $\delta = 0.436473$ , where the rate gain is largest; the vertical segment joins the two rates at this distance. The right panel shows the gain over the full improvement interval, approximately  $0.3897184 < \delta < 0.4817103$ , with vertical values in units of  $10^{-6}$ . The maximum gain is approximately  $1.37905 \times 10^{-5}$ . The two panels use different horizontal scales.

Like Zinoviev and Litsyn [ZL85], we begin with algebraic-geometric codes over a larger alphabet and restrict the allowed symbols in each coordinate. Their argument retains the distance guarantee of the starting code. Our approach additionally controls the number of close pairs that survive alphabet restriction, allowing us to increase the guaranteed distance by removing relatively few codewords. We obtain this control by averaging weight distributions over divisor classes and using the zeta function of the underlying function field.

**Theorem 1.1** is an existence result: it improves our understanding of the combinatorial limits of the rate–distance tradeoff, but does not provide an explicit construction. Matching this guarantee explicitly is a separate algorithmic challenge. Indeed, even matching the GV bound by explicit codes remains open for the alphabet considered in **Theorem 1.1**. The corresponding binary problem remains a major open problem, with Ta-Shma’s breakthrough construction [TS17] and subsequent advances [CC26, HMZ26] providing explicit codes that approach the bound in the high-distance regime. Ta-Shma’s approach was also extended to nonbinary finite fields by Jalan and Moshkovitz [JM21].

Our result surpasses the GV bound for every alphabet size  $q \geq 42$ . For concreteness, we focus on  $q = 42$ , the smallest alphabet size covered by our analysis. For  $q = 46$ , our bound improves upon that of Zinoviev and Litsyn [ZL85]: it exceeds GV over a larger interval of relative distances and achieves a strictly higher rate throughout the interval on

which their bound exceeds GV. We expect our technique to yield further improvements over some of the existing bounds for larger alphabets. Determining the full extent of these improvements requires alphabet-specific comparisons with the extensive literature on algebraic-geometric codes and improvements to asymptotic rate–distance bounds (see [Section 1.2](#)); we leave such comparisons to the interested reader.

Naturally, the holy grail remains to resolve the binary case: either to surpass the GV bound asymptotically or to prove that it is tight. On the converse side, the celebrated linear-programming bounds of McEliece, Rodemich, Rumsey, and Welch [[MRRW77](#)] remained the strongest general asymptotic bounds for almost half a century, until the recent breakthroughs of Alrabiah and Guruswami [[AG26](#)] and OpenAI [[Ope26](#)], followed by further advances of Gay, Jeronimo, and Liu [[GJL26](#)] and Salmon [[Sal26](#)]. Despite this progress, a substantial gap between the known achievable rates and the converse bounds remains.

Our improvement for an alphabet of size 42 naturally raises the question of how much further the alphabet size can be reduced, and whether the remaining gap to the binary case can ultimately be closed. Although our result beats the GV bound for every alphabet size  $q \geq 42$ , one should be cautious about viewing alphabet size as the sole relevant parameter: the algebraic structure of the fields used to obtain such codes may be essential to achieving these dense packings. More concretely, we ask whether the possibility of beating the GV bound is monotone in the alphabet size: does a strict asymptotic improvement over the  $q$ -ary GV bound at some relative distance imply a strict asymptotic improvement over the  $(q+1)$ -ary GV bound, possibly with a smaller gap and at a different relative distance?

## 1.2 Related work

We briefly review several lines of work related to our results. This overview is intended to place our work in context rather than provide a comprehensive survey of the literature.

**Finite-length improvements.** Jiang and Vardy [[JV04](#)] improved the finite-length GV bound by an  $\Omega(n)$  factor in code size for binary codes, and Vu and Wu [[VW05](#)] extended this result to general alphabets. For linear codes, Gaborit and Zémor [[GZ08](#)] obtained an  $\Omega(n)$ -factor improvement for certain binary families of rate  $1/2$ , while Hao, Huang, Livshyts, and Tikhomirov [[HHLT22](#)] obtained an  $\Omega(\sqrt{n})$ -factor improvement over general prime-power alphabets. Yuan and Zhu [[YZ26](#)] subsequently gave an inclusion–exclusion proof and an extension to quantum codes. These polynomial factors vanish in the asymptotic rate. Our proof shares the classical viewpoint of codes as independent sets in a graph of close pairs, but obtains an asymptotic improvement through arithmetic control of the

pairs surviving alphabet restriction.

**Refinements of AG codes and their weight distributions.** Litsyn and Tsfasman [LT86] obtained asymptotic bounds for nonlinear codes improving, for certain alphabets, on both GV and AG-code bounds. Further refinements include work of Xing [Xin01, Xin03], Elkies [Elk01, Elk03], Niederreiter and Özbudak [NÖ07], Yang and Qi [YQ09], and Liu, Wu, and Xing [LWX23], whose use of effective-divisor counts is particularly related to our approach. Another close connection is Duursma’s [Duu99] expression of AG-code weight distributions in terms of  $L$ -series through effective divisors in divisor classes. We use such counting information to control the close pairs surviving alphabet restriction and the resulting cost of expurgation.

**Towers over nonprime fields.** Bassa, Beelen, Garcia, and Stichtenoth [BBGS15, BBGS14] constructed recursive towers over nonprime finite fields and showed that linear codes beat the asymptotic GV bound over every such field of size at least 49, with the possible exception of 125. See Beelen’s survey [Bee22] for an overview of recursive towers and Ihara’s constant.

**Trace codes.** A different approach to alphabet reduction takes coordinatewise field traces of AG codes, with the goal of retaining their advantage over GV over smaller alphabets. Kopparty, Ta-Shma, and Yakirevitch [KTSY26] refine Stepanov’s method to prove vanishing bias for families of trace Hermitian codes. Cohen, Doron, Goldgraber, and Manket [CDGM26] develop Hasse–Weil-type bounds adapted to trace codes over more general function fields. These results provide tools toward meeting or surpassing GV over small alphabets, which remains open for these trace constructions.

## 2 Proof overview

This section describes the main ideas that go into the proof and assumes some familiarity with algebraic function fields. To help bridge this gap for readers unfamiliar with the subject, we give a friendly exposition of the required background in [Section A](#). Written independently of the results obtained in this paper, it introduces the relevant notions through a running example and explains the results we use, with the aim of providing enough background to follow the ideas below.

The relevance of this theory extends well beyond improvements to the GV bound. Algebraic-geometric codes have played an important role in the development of list-

decoding algorithms [SW99, GS99] and in obtaining efficiently list-decodable codes with near-optimal rate over bounded alphabets [GX12, GX13, GX14]. Related algebraic methods have also led to explicit constructions of subspace designs [GK13], with algebraic function fields enabling constructions over small fields and further applications to dimension expanders and rank condensers [GXY17, CDG26]. These examples represent only part of a broader range of applications, which also includes probabilistically checkable proofs [BSKK<sup>+</sup>13], fine-grained hardness of approximation [Rub18], locally correctable codes [BSGK<sup>+</sup>13], and secret sharing and secure multiparty computation [CC06]. We hope that the exposition will serve both as background for our proof and as an invitation to explore these methods and their broader applications in theoretical computer science.

For a more comprehensive treatment, we warmly recommend Stichtenoth’s book [Sti09], the two volumes by Tsfasman, Vlăduț, and Nogin, *Basic Notions* and *Advanced Chapters* [TVN07, TVN19], and Niederreiter and Xing’s *Algebraic Geometry in Coding Theory and Cryptography* [NX09]. The survey by Couvreur and Randriambololona [CR20] provides a shorter overview of algebraic-geometric codes and their applications, while Beelen’s survey [Bee22] focuses on recursive towers of function fields and Ihara’s constant.

**Starting with a larger alphabet.** For concreteness, we begin with an asymptotically optimal tower of function fields over  $\mathbb{F}_{49}$ . At each level, let  $g$  be the genus and let  $P_1, \dots, P_n$  be all the rational places of the corresponding function field. Optimality gives  $g/n \rightarrow 1/6$ . Write  $D = P_1 + \dots + P_n$ , fix a parameter  $1/3 < a < 1$  to be specified later, and set  $m = \lfloor an \rfloor$ . Consider the Goppa code  $C_L(D, G)$  associated with a degree- $m$  divisor  $G$  whose support avoids the evaluation places. At sufficiently high levels of the tower, we have  $2g - 2 < m < n$ , so the Riemann–Roch theorem gives dimension exactly  $m - g + 1$ , while the usual zero-counting argument gives minimum distance at least  $n - m$ . These are the familiar ingredients of Goppa codes, instantiated with an optimal tower. We refer to this code as the “mother code.”

**Reducing the alphabet and improving the distance.** Our aim is to reduce the alphabet from 49 to 42 symbols and, at the same time, increase the guaranteed relative distance from approximately  $1 - a$  to a target value  $\delta > 1 - a$ . Both operations require discarding codewords, so we must begin with a sufficiently high-dimension mother code to absorb this loss. The question is whether we can exploit this additional dimension while removing relatively few codewords to enforce the smaller alphabet and the larger distance, leaving a code whose rate exceeds the 42-ary GV bound. To answer this, we need to control how many pairs of codewords lie below the target distance, rather than merely know the minimum distance of the mother code. Since the mother code is linear,

this amounts to understanding its weight distribution.

**Random alphabet restriction and expurgation.** For a linear mother code  $C \subseteq \mathbb{F}_{49}^n$ , independently choose a uniformly random 42-element subset  $S_j \subseteq \mathbb{F}_{49}$  at each coordinate and retain  $C' = C \cap (S_1 \times \cdots \times S_n)$ . A simple calculation gives the relevant survival probabilities: every word survives with probability  $p^n$ , where  $p = 42/49$ , and if two words are at distance  $w$ , then, conditional on the first surviving, the second survives with probability  $b^w$ , where  $b = 41/48$ . By linearity, every word of  $C$  has exactly  $A_w(C)$  other words at distance  $w \geq 1$ , where  $A_w(C)$  denotes the number of weight- $w$  words in  $C$ . Setting  $d = \lceil \delta n \rceil$ , define

$$W(C, d) = \sum_{1 \leq w < d} A_w(C) b^w.$$

This weighted sum is proportional to the expected number of surviving pairs at distance less than  $d$ , and hence bounds, up to the same normalization factor, the expected number of codewords we need to remove to eliminate all such pairs.

To make this precise, form a graph  $G = (V, E)$  on the surviving words, joining pairs at distance less than  $d$ . Linearity of expectation gives

$$\mathbb{E}|V| = \mu := |C|p^n, \quad \mathbb{E}|E| = \frac{\mu}{2} W(C, d).$$

Deleting one endpoint of each edge leaves at least  $|V| - |E|$  words with the desired distance. Consequently, if  $W(C, d) \leq 1$ , then  $\mathbb{E}[|V| - |E|] \geq \mu/2$ , so an averaging argument guarantees a choice of the coordinate subsets leaving at least  $\mu/2$  words after expurgation. Relabeling each  $S_j$  by a common 42-element alphabet preserves all distances.

For our Goppa mother code,  $|C| = 49^{m-g+1}$ , so the code  $B$  obtained after restriction and expurgation satisfies

$$|B| \geq \frac{\mu}{2} = \frac{1}{2} 49^{m-g+1} \left( \frac{42}{49} \right)^n.$$

Taking logarithms to base 42, dividing by the block length, and using  $m/n \rightarrow a$  and  $g/n \rightarrow 1/6$ , we obtain an asymptotic rate of at least

$$R(a) = 1 - \left( \frac{7}{6} - a \right) \frac{\ln 49}{\ln 42}.$$

Thus, it remains to find a degree- $m$  divisor  $G$  for which the mother code  $C = C_L(D, G)$  satisfies  $W(C, d) \leq 1$ , while choosing  $a$  large enough that  $R(a) > 1 - H_{42}(\delta)$ . The usual minimum-distance bound only tells us that  $A_w(C) = 0$  for  $w < n - m$ ; it does not control the contributions from  $n - m \leq w < d$ .

**Averaging over divisor classes.** To control the number of codewords with weights between the designed distance  $n - m$  and the target distance  $d$ , we introduce an additional layer of randomness into the construction: rather than fix the mother code, we choose its defining divisor class uniformly among all classes of degree  $m$ , keeping the function field and evaluation places fixed. The resulting codes all have the same dimension and designed minimum distance, but their weight distributions may differ. The key insight is that, although the weight distribution of an individual code may be difficult to control, its average over divisor classes admits an exact expression in terms of the zeta function of the underlying function field.

It is known that the number of degree- $m$  divisor classes is exactly the class number  $h_F$ , a fundamental invariant of the function field  $F$ . For each such class  $\gamma$ , fix a representative  $G_\gamma$  whose support avoids the evaluation places; the approximation theorem guarantees that such a representative exists. We now choose  $\gamma$  uniformly at random and take  $C_\gamma = C_L(D, G_\gamma)$  as our mother code.

All these codes have the same dimension. Changing the representative within a fixed class only rescales the coordinates by nonzero field elements and therefore preserves the weight distribution. The weight distribution may nevertheless vary between classes, but its average admits the following exact expression:

$$\mathbb{E}_\gamma A_w(C_\gamma) = \frac{48}{h_F} \binom{n}{w} [T^{m-n+w}] (Z_F(T)(1-T)^w), \quad 1 \leq w \leq n. \quad (2.1)$$

Here  $[T^r]$  denotes the coefficient of  $T^r$ , and  $Z_F(T)$  is the zeta function of  $F$ , whose coefficient of  $T^r$  counts the effective divisors of degree  $r$ . This identity can be extracted from Duursma's generating-function formula [Duu99]. It connects the average weight distribution of interest to a classical object in the arithmetic of function fields, allowing us to bring the extensive theory of zeta functions to bear on our counting problem.

The reason for this identity is worth spelling out. Since  $m < n$ , evaluation on  $D$  is injective, so counting functions in  $\mathcal{L}(G_\gamma)$  is equivalent to counting their evaluation vectors. Fix a set  $S \subseteq \{1, \dots, n\}$  of size  $w$ . We wish to count functions  $f \in \mathcal{L}(G_\gamma)$  whose evaluation vector  $(f(P_1), \dots, f(P_n))$  is nonzero precisely at the coordinates in  $S$ . Each such function determines an effective divisor  $G_\gamma + (f)$ , which contains every  $P_i$  with  $i \notin S$  with multiplicity at least one and contains none of the  $P_i$  with  $i \in S$ . Subtracting one copy of each of the  $n - w$  places outside  $S$  leaves an effective divisor  $E$  of degree  $m - n + w$  whose support avoids every  $P_i$  with  $i \in S$ . Conversely, given such an  $E$ , the divisor  $E + \sum_{i \notin S} P_i$  determines the degree- $m$  class  $\gamma$ . For our fixed representative  $G_\gamma$ , the equation  $G_\gamma + (f) = E + \sum_{i \notin S} P_i$  then has exactly 48 solutions  $f$ , differing by multiplication by elements of  $\mathbb{F}_{49}^\times$ .

To count the possible residual divisors  $E$ , recall that each rational place contributes a factor  $1 + T + T^2 + \cdots = (1 - T)^{-1}$  to the Euler product for  $Z_F(T)$ , recording its possible multiplicities in an effective divisor. Requiring multiplicity zero at each of the  $w$  places indexed by  $S$  removes these factors, leaving the generating function  $Z_F(T)(1 - T)^w$ . Taking its coefficient of degree  $m - n + w$ , multiplying by  $48 \binom{n}{w}$ , and dividing by the number  $h_F$  of divisor classes gives [Equation \(2.1\)](#).

Although  $(1 - T)^w$  has alternating coefficients, the product  $Z_F(T)(1 - T)^w$  has non-negative coefficients, since it counts effective divisors avoiding the specified places. For any  $0 < z < 1/49$ , each term of this convergent generating function is therefore bounded by its sum, giving

$$[T^r](Z_F(T)(1 - T)^w) \leq z^{-r} Z_F(z)(1 - z)^w, \quad r \geq 0.$$

**From divisor counts to a gain over GV.** To evaluate the coefficient bound, we use two asymptotic consequences of optimality:

$$\frac{\ln Z_F(z)}{n} \longrightarrow -\ln(1 - z), \quad \frac{\ln h_F}{n} \longrightarrow \frac{\ln 49}{6} + \ln \frac{49}{48},$$

where the first limit holds for every fixed  $0 < z < 1/49$ . These follow from the generalized Drinfeld–Vlăduț inequality and Weil’s theorem.

Substituting these estimates into the coefficient bound and summing over weights below the target distance reduces the proof to choosing the degree fraction  $a$  so that  $\mathbb{E}_\gamma W(C_\gamma, d)$  decays exponentially while the resulting rate  $R(a)$  exceeds  $1 - H_{42}(\delta)$ . The remaining argument consists of optimizing the bound and verifying that both requirements can be met for every  $\delta \in [0.39, 0.48]$ . Thus, at every sufficiently high level of the tower, some mother code satisfies  $W(C_\gamma, d) \leq 1$ , and alphabet restriction followed by expurgation gives the desired codes.

### 3 Preliminaries

We use the standard language of algebraic function fields, following Stichtenoth [\[Sti09\]](#), and briefly recall the notation needed below. A more detailed exposition appears in [Section A](#).

Throughout,  $Q$  is a prime power, and  $F/\mathbb{F}_Q$  is an algebraic function field of one variable with full constant field  $\mathbb{F}_Q$ . We write  $g = g(F)$  for its genus and  $\mathbb{P}_F$  for its set of places. For a place  $P$ , let  $\nu_P$  denote the normalized discrete valuation and let  $\deg P = [F_P : \mathbb{F}_Q]$ , where  $F_P$  is the residue field. A place is *rational* if it has degree one. We denote by  $B_r(F)$

the number of places of degree  $r$ , and write  $N(F) = B_1(F)$ . If  $f \in F$  is regular at a rational place  $P$ , its residue  $f(P)$  is an element of  $\mathbb{F}_Q$ .

The divisor group  $\text{Div}(F)$  is the free abelian group on  $\mathbb{P}_F$ . For a divisor  $G = \sum_P n_P P$ , its degree and support are  $\deg G = \sum_P n_P \deg P$  and  $\text{supp}(G) = \{P : n_P \neq 0\}$ . We call  $G$  *effective*, and write  $G \geq 0$ , if all  $n_P \geq 0$ . For  $f \in F^\times$ , its principal divisor is  $(f) = \sum_P \nu_P(f) P$  and has degree zero. Two divisors are linearly equivalent if their difference is principal. We write  $\mathcal{C}_F$  for the divisor class group,  $[G]$  for the class of  $G$ , and  $\mathcal{C}_F^m$  for the set of classes of degree  $m$ . The class number is  $h_F = |\mathcal{C}_F^0|$ ; whenever  $F$  has a rational place,  $|\mathcal{C}_F^m| = h_F$  for every integer  $m$ .

The Riemann–Roch space associated with a divisor  $G$  is

$$\mathcal{L}(G) = \{0\} \cup \{f \in F^\times : (f) + G \geq 0\}, \quad \ell(G) = \dim_{\mathbb{F}_Q} \mathcal{L}(G).$$

The Riemann–Roch theorem gives  $\ell(G) \geq \deg G + 1 - g$ , with equality when  $\deg G > 2g - 2$ . Moreover,  $\ell(G) = 0$  when  $\deg G < 0$ .

For distinct rational places  $P_1, \dots, P_N$ , write  $D = P_1 + \dots + P_N$ . If  $\text{supp}(G) \cap \text{supp}(D) = \emptyset$ , the associated Goppa code is

$$C_L(D, G) = \{(f(P_1), \dots, f(P_N)) : f \in \mathcal{L}(G)\} \subseteq \mathbb{F}_Q^N.$$

When  $2g - 2 < \deg G < N$ , this code has dimension  $\deg G + 1 - g$  and minimum distance at least  $N - \deg G$ .

Finally, the zeta function of  $F$  is

$$Z_F(T) = \sum_{r=0}^{\infty} |\{E \in \text{Div}(F) : E \geq 0, \deg E = r\}| T^r = \prod_{P \in \mathbb{P}_F} (1 - T^{\deg P})^{-1}.$$

We write  $[T^r]A(T)$  for the coefficient of  $T^r$  in a power series  $A(T)$ , interpreted as zero when  $r < 0$ . All logarithms denoted by  $\ln$  are natural.

## 4 Proof of the main theorem

In this section we prove [Theorem 1.1](#). Throughout the section,  $Q$  is a prime power and  $2 \leq q \leq Q$  is an integer (not necessarily a prime power). For a linear code  $C \subseteq \mathbb{F}_Q^N$  and  $w \in [N]$ , let

$$A_w(C) = |\{c \in C : \text{wt}(c) = w\}|. \quad (4.1)$$

For each  $d \in [N + 1]$ , define the *truncated weight enumerator*

$$W_q(C, d) := \sum_{w=1}^{d-1} A_w(C) \left( \frac{q-1}{Q-1} \right)^w. \quad (4.2)$$

## 4.1 Alphabet restriction and expurgation

The purpose of this section is to establish the following lemma.

**Lemma 4.1.** *Let  $C_1 \subseteq \mathbb{F}_Q^N$  be a linear code and let  $d \in [N+1]$ . If  $W_q(C_1, d) \leq 1$ , then there exists a  $q$ -ary code  $C_2$  of length  $N$  and minimum distance  $d(C_2) \geq d$  satisfying*

$$|C_2| \geq \frac{1}{2} \cdot |C_1| \cdot \left(\frac{q}{Q}\right)^N. \quad (4.3)$$

*Proof.* The proof is by the probabilistic method. Choose  $N$  subsets  $S_1, \dots, S_N \subseteq \mathbb{F}_Q$  independently and uniformly at random from the collection of all  $q$ -element subsets of  $\mathbb{F}_Q$ . Let  $C' := C_1 \cap (S_1 \times \dots \times S_N)$ . For every codeword  $c = (c_1, \dots, c_N) \in C_1$  we have

$$\Pr[c \in C'] = \prod_{i=1}^N \Pr[c_i \in S_i] = \left(\frac{q}{Q}\right)^N,$$

where the probability is taken over the random choice of  $S_1, \dots, S_N$ . Moreover, for  $x, y \in C_1$  with Hamming distance  $\Delta(x, y) = w$  we have

$$\begin{aligned} \Pr[x, y \in C'] &= \Pr[x \in C'] \cdot \Pr[y \in C' \mid x \in C'] \\ &= \left(\frac{q}{Q}\right)^N \cdot \prod_{i: x_i \neq y_i} \Pr[y_i \in S_i \mid x_i \in S_i] \\ &= \left(\frac{q}{Q}\right)^N \cdot \left(\frac{q-1}{Q-1}\right)^w. \end{aligned} \quad (4.4)$$

For each  $w \in [N]$ , let  $B_w = \{(x, y) \in C_1 \times C_1 \mid \Delta(x, y) = w\}$ . Since the code  $C_1$  is linear,

$$|B_w| = |C_1| \cdot A_w(C_1). \quad (4.5)$$

Now, consider the undirected random graph  $G = (V, E)$  whose vertices are the codewords in  $C'$ , with an edge between any two vertices  $x, y \in C'$  such that  $0 < \Delta(x, y) < d$ . Then

$$\mathbb{E}[|V|] = \mathbb{E}[|C'|] = |C_1| \cdot \left(\frac{q}{Q}\right)^N$$

and, by Equation (4.4) and Equation (4.5),

$$\begin{aligned} \mathbb{E}[|E|] &= \frac{1}{2} \sum_{w=1}^{d-1} \sum_{(x,y) \in B_w} \Pr[x, y \in C'] \\ &= \frac{1}{2} \sum_{w=1}^{d-1} |B_w| \cdot \left(\frac{q}{Q}\right)^N \cdot \left(\frac{q-1}{Q-1}\right)^w \\ &= \frac{1}{2} |C_1| \left(\frac{q}{Q}\right)^N \cdot \sum_{w=1}^{d-1} A_w(C_1) \cdot \left(\frac{q-1}{Q-1}\right)^w \\ &= \frac{1}{2} |C_1| \left(\frac{q}{Q}\right)^N \cdot W_q(C_1, d). \end{aligned}$$

It follows that if  $W_q(C_1, d) \leq 1$ , then

$$\mathbb{E}[|V| - |E|] \geq \frac{1}{2} \cdot |C_1| \cdot \left(\frac{q}{Q}\right)^N.$$

In particular, by an averaging argument, there exist subsets  $S_1, \dots, S_N$  for which the obtained graph  $G$  satisfies the same lower bound.

Fix such  $G$  and choose one endpoint of each edge in  $E$  and delete the union of the chosen endpoints from  $V$ . The remaining vertices form an independent set of size at least  $|V| - |E|$  in  $G$ . Thus, they form a subcode  $C'' \subseteq C'$  with minimum distance  $d(C'') \geq d$  and

$$|C''| \geq |V| - |E| \geq \frac{1}{2} \cdot |C_1| \cdot \left(\frac{q}{Q}\right)^N.$$

The desired code  $C_2 \subseteq [q]^N$  is then obtained from  $C''$  by applying coordinate-wise bijections  $\phi_i: S_i \rightarrow [q]$ . That is,  $C_2 = \phi(C'')$  where

$$\phi(c_1, \dots, c_N) = (\phi_1(c_1), \dots, \phi_N(c_N)).$$

This operation clearly preserves the code distance. □

## 4.2 Average weight distributions via the zeta function

Let  $F/\mathbb{F}_Q$  be a function field of genus  $g$  with distinct rational places  $P_1, \dots, P_N$ , and let  $D = P_1 + \dots + P_N$ . For  $m \in \mathbb{N}$  we denote by  $\mathcal{C}_F^m$  the set of degree- $m$  divisor classes,

$$\mathcal{C}_F^m = \{[G] \in \mathcal{C}_F : \deg[G] = m\}.$$

We recall that this is a finite set of size  $|\mathcal{C}_F^m| = h_F$ , the class number of  $F/\mathbb{F}_Q$ .

Fix an integer  $2g - 2 < m < N$ . We record the following basic claim.

**Claim 4.2.** *Every divisor class  $\gamma \in \mathcal{C}_F^m$  has a representative  $G_\gamma \in \text{Div}(F)$  satisfying*

$$\text{supp}(G_\gamma) \cap \text{supp}(D) = \emptyset.$$

*Proof.* Let  $G \in \text{Div}(F)$  be an arbitrary representative of the class  $\gamma$ . By the Weak Approximation Theorem ([Sti09, Theorem 1.3.1]), there exists a function  $f \in F$  satisfying

$$\nu_{P_i}(f) = -\nu_{P_i}(G) \quad \text{for all } i \in [N].$$

The divisor  $G_\gamma = G + (f)$  satisfies  $[G_\gamma] = [G] = \gamma$  and  $\nu_{P_i}(G_\gamma) = 0$  for all  $i \in [N]$ , hence  $\text{supp}(G_\gamma) \cap \text{supp}(D) = \emptyset$  as desired. □

For each class  $\gamma \in \mathcal{C}_F^m$  fix a representative  $G_\gamma$  as in [Claim 4.2](#), and consider the family of AG codes  $\{C_\gamma\}$ , where

$$C_\gamma := C_L(D, G_\gamma) = \{(f(P_1), \dots, f(P_N)) : f \in \mathcal{L}(G_\gamma)\}.$$

Note that each  $C_\gamma$  is a linear code with

$$\dim(C_\gamma) = \ell(G_\gamma) = m + 1 - g \quad \text{and} \quad d(C_\gamma) \geq N - m. \quad (4.6)$$

The following identity is a coefficient form of Duursma's formula for the average weight distribution of algebraic-geometric codes [\[Duu99\]](#); see also [\[Duu08, Section 1.4.8\]](#). We include a direct proof for completeness.

**Lemma 4.3.** *For each  $w \in [N]$  we have*

$$\mathbb{E}_\gamma [A_w(C_\gamma)] = \frac{Q-1}{h_F} \binom{N}{w} \left[ T^{m-N+w} \right] (Z_F(T)(1-T)^w), \quad (4.7)$$

where the expectation is taken over  $\gamma \in \mathcal{C}_F^m$  chosen uniformly at random.

Here  $Z_F(T)$  is the zeta function of  $F/\mathbb{F}_Q$ , and  $[T^k]f(T)$  denotes the coefficient of  $T^k$  in the power series expansion of  $f(T)$  (interpreted as 0 if  $k < 0$ ).

*Proof.* Let  $w \in [N]$  be given. For a subset  $S \subseteq [N]$ , let  $W_S(C_\gamma)$  denote the set of codewords in  $C_\gamma$  whose nonzero symbols occur exactly at the indices in  $S$ . It is clear that

$$A_w(C_\gamma) = \sum_{\substack{S \subseteq [N] \\ |S|=w}} |W_S(C_\gamma)|.$$

Together with the fact that  $|\mathcal{C}_F^m| = h_F$ , we obtain

$$\begin{aligned} \mathbb{E}_\gamma [A_w(C_\gamma)] &= \frac{1}{h_F} \sum_{\gamma \in \mathcal{C}_F^m} A_w(C_\gamma) \\ &= \frac{1}{h_F} \sum_{\gamma \in \mathcal{C}_F^m} \sum_{\substack{S \subseteq [N] \\ |S|=w}} |W_S(C_\gamma)| \\ &= \frac{1}{h_F} \sum_{\substack{S \subseteq [N] \\ |S|=w}} \sum_{\gamma \in \mathcal{C}_F^m} |W_S(C_\gamma)|. \end{aligned}$$

Since the number of subsets  $S \subseteq [N]$  of size  $w$  is  $\binom{N}{w}$ , it suffices to show that for every  $S \subseteq [N]$  of size  $w$ ,

$$\sum_{\gamma \in \mathcal{C}_F^m} |W_S(C_\gamma)| = (Q-1) \left[ T^{m-N+w} \right] (Z_F(T)(1-T)^w). \quad (4.8)$$

Indeed, fix  $S \subseteq [N]$  with  $|S| = w$ . To prove [Equation \(4.8\)](#), we first define for each  $\gamma \in \mathcal{C}_F^m$ ,

$$\mathcal{F}_\gamma = \{f \in \mathcal{L}(G_\gamma) : (f(P_1), \dots, f(P_N)) \in W_S(C_\gamma)\}.$$

Note that the disjoint union

$$\mathcal{U} := \bigcup_{\gamma \in \mathcal{C}_F^m} (\mathcal{F}_\gamma \times \{\gamma\})$$

has size  $|\mathcal{U}| = \sum_{\gamma \in \mathcal{C}_F^m} |W_S(C_\gamma)|$ , which is exactly the LHS of [Equation \(4.8\)](#).

We also introduce the set  $\mathcal{H}$  of effective divisors of degree  $m - N + w$  whose supports are disjoint from  $\{P_i : i \in S\}$ , that is,

$$\mathcal{H} = \{H \in \text{Div}(F) : H \geq 0, \deg H = m - N + w \text{ and } \nu_{P_i}(H) = 0 \text{ for all } i \in S\}.$$

To proceed, we need the following claim.

**Claim 4.4.** *The map  $\Phi : \mathcal{U} \rightarrow \mathcal{H}$  given by*

$$(f, \gamma) \mapsto (f) + G_\gamma - \sum_{i \in [N] \setminus S} P_i$$

*is  $(Q - 1)$ -to-one.*

*Proof of Claim 4.4.* We first show that the map  $\Phi$  is well-defined. Let  $(f, \gamma) \in \mathcal{U}$  and

$$H := \Phi(f, \gamma) = (f) + G_\gamma - \sum_{i \in [N] \setminus S} P_i.$$

It is clear that

$$\deg H = \deg(G_\gamma) - \sum_{i \in [N] \setminus S} \deg P_i = m - (N - |S|) \cdot 1 = m - N + w.$$

In addition,  $f \in \mathcal{F}_\gamma$  and hence  $0 \neq f \in \mathcal{L}(G_\gamma)$ , with  $f(P_i) = 0$  precisely when  $i \notin S$ . In particular,  $(f) + G_\gamma \geq 0$  and  $\nu_{P_i}(f) \geq 1$  for  $i \notin S$ . Since  $\text{supp}(G_\gamma) \cap \text{supp}(D) = \emptyset$  we conclude that  $\nu_{P_i}(H) = \nu_{P_i}(f) - 1 \geq 0$  for all  $i \notin S$  and therefore  $H \geq 0$ . Finally, for  $i \in S$  we have  $f(P_i) \neq 0$  and hence  $\nu_{P_i}(H) = \nu_{P_i}(f) = 0$ . This shows that  $H \in \mathcal{H}$ , as desired.

Now, let  $H \in \mathcal{H}$  be arbitrary. We need to show that  $|\Phi^{-1}(H)| = Q - 1$ . Since  $\deg H = m - N + w$ , the divisor

$$H' := H + \sum_{i \in [N] \setminus S} P_i$$

has degree  $m$ . Observe that for any pair  $(f, \gamma) \in \mathcal{U}$ ,

$$\Phi(f, \gamma) = H \iff (f) + G_\gamma - \sum_{i \in [N] \setminus S} P_i = H \iff (f) + G_\gamma = H'. \quad (4.9)$$

Thus, any preimage of  $H$  must have  $\gamma = [H']$ . Conversely, set  $\gamma = [H']$ . Since  $G_\gamma$  represents this class, there exists  $h \in F^\times$  such that  $H' = (h) + G_\gamma$ . Since  $H' \geq 0$  we get that  $h \in \mathcal{L}(G_\gamma)$ , and since  $\nu_{P_i}(h) = \nu_{P_i}(H')$  we see that  $\nu_{P_i}(h) = 0$  for  $i \in S$  and  $\nu_{P_i}(h) > 0$  otherwise. It follows that  $\alpha h \in \mathcal{F}_\gamma$  and hence  $(\alpha h, \gamma) \in \mathcal{U}$  for all  $\alpha \in \mathbb{F}_Q^\times$ . Moreover, as  $(\alpha h) = (h)$  and  $(h) + G_\gamma = H'$ , we get (in view of Equation (4.9)) that  $\Phi(\alpha h, \gamma) = H$ . Thus,

$$\Phi^{-1}(H) \supseteq \{(\alpha h, \gamma) \mid \alpha \in \mathbb{F}_Q^\times\}.$$

For the reverse inclusion, note that if  $(f, \gamma) \in \Phi^{-1}(H)$  then  $\gamma = [H']$  and

$$(f) + G_\gamma = H' = (h) + G_\gamma \implies (f) = (h),$$

which implies that  $f = \alpha h$  for some  $\alpha \in \mathbb{F}_Q^\times$ .  $\square$

Note that Claim 4.4 implies that

$$\sum_{\gamma \in \mathcal{C}_F^m} |W_S(C_\gamma)| = |\mathcal{U}| = (Q - 1) \cdot |\mathcal{H}|.$$

Thus, in order to establish Equation (4.8) it remains to show that

$$|\mathcal{H}| = \left[ T^{m-N+w} \right] (Z_F(T)(1 - T)^w). \quad (4.10)$$

Let  $\mathbb{P}_S = \mathbb{P}_F \setminus \{P_i : i \in S\}$ . The classical Euler-product argument shows that  $|\mathcal{H}|$  is the coefficient of  $T^{m-N+w}$  in the generating function

$$\begin{aligned} A(T) &= \prod_{\mathfrak{p} \in \mathbb{P}_S} \sum_{n=0}^{\infty} T^{\deg(n\mathfrak{p})} = \prod_{\mathfrak{p} \in \mathbb{P}_S} \sum_{n=0}^{\infty} \left( T^{\deg \mathfrak{p}} \right)^n \\ &= \prod_{\mathfrak{p} \in \mathbb{P}_S} \left( 1 - T^{\deg \mathfrak{p}} \right)^{-1} = \frac{Z_F(T)}{\prod_{i \in S} (1 - T^{\deg P_i})^{-1}} = Z_F(T)(1 - T)^w. \end{aligned} \quad (4.11)$$

This completes the proof of Lemma 4.3.  $\square$

**Corollary 4.5.** *For every  $w \in [N]$  and  $0 < t < 1/Q$  it holds that*

$$\mathbb{E}_\gamma [A_w(C_\gamma)] \leq \frac{Q-1}{h_F} \binom{N}{w} \cdot Z_F(t) \cdot t^{-(m-N+w)} (1-t)^w.$$

*Proof.* Let  $w \in [N]$  and  $0 < t < 1/Q$ . By [Lemma 4.3](#) it suffices to show that

$$\left[ T^{m-N+w} \right] (Z_F(T)(1-T)^w) \leq Z_F(t) \cdot t^{-(m-N+w)}(1-t)^w.$$

The generating function  $A(T) = \sum_{n=0}^{\infty} A_n T^n$  defined in [Equation \(4.11\)](#) satisfies  $A_n \geq 0$  for all  $n$ . As a complex-valued function, it converges to the function  $Z_F(T)(1-T)^w$  in the disk  $|T| < 1/Q$ . In particular, for  $T = t$  and  $k \in \mathbb{N}$  we obtain

$$A_k t^k \leq A(t) = Z_F(t)(1-t)^w \implies \left[ T^k \right] A(T) \leq Z_F(t) \cdot t^{-k}(1-t)^w.$$

Note that the last inequality holds trivially for  $k < 0$  as well. Taking  $k = m - N + w$  yields the desired inequality.  $\square$

### 4.3 Zeta functions and class numbers of optimal families

Let  $(F_i/\mathbb{F}_Q)_{i=1}^{\infty}$  be a sequence of function fields with genera  $g_i \rightarrow \infty$ . For an integer  $r \geq 1$ , let  $B_r(F_i)$  denote the number of places of degree  $r$  in  $F_i$ . In particular,  $N_i = B_1(F_i)$ .

**Definition 4.6.** *The sequence  $(F_i/\mathbb{F}_Q)_{i=1}^{\infty}$  is asymptotically optimal if*

$$\lim_{i \rightarrow \infty} \frac{N_i}{g_i} = \sqrt{Q} - 1.$$

*The sequence is asymptotically exact if for all  $r \geq 1$ , the limit*

$$\beta_r := \lim_{i \rightarrow \infty} \frac{B_r(F_i)}{g_i}$$

*exists.*

The following generalization of the Drinfeld–Vlăduț bound goes back to Tsfasman [\[Tsf92\]](#); we use the formulation of Tsfasman and Vlăduț [\[TV02, Theorem 3.2\]](#).

**Theorem 4.7.** *Let  $(F_i/\mathbb{F}_Q)_{i=1}^{\infty}$  be an asymptotically exact sequence of function fields. Then*

$$\sum_{r=1}^{\infty} \frac{r\beta_r}{Q^{r/2} - 1} \leq 1. \tag{4.12}$$

**Lemma 4.8.** *Let  $(F_i/\mathbb{F}_Q)_{i=1}^{\infty}$  be an asymptotically optimal sequence of function fields. Then*

$$\lim_{i \rightarrow \infty} \frac{\ln h_{F_i}}{N_i} = \frac{\ln Q}{\sqrt{Q} - 1} - \ln(1 - 1/Q), \tag{4.13}$$

*and for every  $0 < t < 1/Q$ ,*

$$\lim_{i \rightarrow \infty} \frac{\ln Z_{F_i}(t)}{N_i} = -\ln(1 - t). \tag{4.14}$$

*Proof.* We begin by showing that the sequence  $(F_i/\mathbb{F}_Q)_{i=1}^\infty$  is asymptotically exact, with  $\beta_1 = \sqrt{Q} - 1$  and  $\beta_r = 0$  for all  $r > 1$ . To this end, it suffices to show that every subsequence of  $(F_i)_{i=1}^\infty$  has a further subsequence that is asymptotically exact with the prescribed limits.

Take an arbitrary subsequence, which we continue to denote by  $(F_i)_{i=1}^\infty$  for simplicity of notation. Let  $N_r(F_i)$  denote the number of rational places in the constant field extension  $F_i\mathbb{F}_{Q^r}/\mathbb{F}_{Q^r}$ . By the Hasse–Weil bound and the invariance of the genus under constant field extensions,

$$|N_r(F_i) - (Q^r + 1)| \leq 2g_i Q^{r/2}.$$

Since

$$N_r(F_i) = \sum_{d|r} dB_d(F_i), \quad (4.15)$$

we obtain

$$B_r(F_i) \leq \frac{1}{r} N_r(F_i) \leq \frac{Q^r + 1 + 2g_i Q^{r/2}}{r}.$$

Thus, for every  $i$  sufficiently large so that  $g_i > 0$  we have

$$0 \leq \frac{B_r(F_i)}{g_i} \leq \frac{Q^r + 1}{rg_i} + \frac{2Q^{r/2}}{r} \leq \frac{4 \cdot Q^r}{r}.$$

It follows that for each  $r$ , the sequence  $(B_r(F_i)/g_i)_{i=1}^\infty$  is bounded and hence has a convergent subsequence. By a standard diagonalization argument, there exists an asymptotically exact subsequence  $(F_{i_j})_{j=1}^\infty$  – along which all the limits

$$\beta'_r := \lim_{j \rightarrow \infty} \frac{B_r(F_{i_j})}{g_{i_j}} \quad (r \geq 1)$$

simultaneously exist. Since the original sequence is asymptotically optimal, we have  $\beta'_1 = \sqrt{Q} - 1$  (and clearly  $\beta'_r \geq 0$  for  $r > 1$ ). By the generalized Drinfeld–Vlăduț bound [Equation \(4.12\)](#) we conclude that  $\beta'_r = 0$  for all  $r > 1$ , as desired.

Next, in order to derive the claimed limits, we write

$$Z_{F_i}(T) = \frac{L_i(T)}{(1-T)(1-QT)}$$

where, by Weil’s Theorem,

$$L_i(T) = \prod_{j=1}^{2g_i} (1 - \alpha_{i,j}T) \quad \text{where } |\alpha_{i,j}| = \sqrt{Q}.$$

Let  $0 < t < \frac{1}{\sqrt{Q}}$  be arbitrary. Then  $|\alpha_{i,j}t| = \sqrt{Q}t < 1$  and

$$\frac{\ln L_i(t)}{N_i} = \frac{1}{N_i} \sum_{j=1}^{2g_i} \ln(1 - \alpha_{i,j}t) = -\frac{1}{N_i} \sum_{j=1}^{2g_i} \sum_{r=1}^{\infty} \frac{(\alpha_{i,j}t)^r}{r} = -\sum_{r=1}^{\infty} \left( \frac{1}{N_i} \sum_{j=1}^{2g_i} \alpha_{i,j}^r \right) \frac{t^r}{r}. \quad (4.16)$$

For all  $i$  sufficiently large and all  $r \geq 1$ , the triangle inequality gives

$$\left| \left( \frac{1}{N_i} \sum_{j=1}^{2g_i} \alpha_{i,j}^r \right) \frac{t^r}{r} \right| \leq \frac{2g_i}{N_i} \cdot \frac{(\sqrt{Q}t)^r}{r}.$$

Since  $g_i/N_i \rightarrow 1/(\sqrt{Q}-1)$ , the RHS is at most  $M_r = C \cdot \frac{(\sqrt{Q}t)^r}{r}$  for some constant  $C > 0$ . Since  $\sum_{r=1}^{\infty} M_r < \infty$ , letting  $i \rightarrow \infty$  in [Equation \(4.16\)](#) and applying the dominated convergence theorem gives

$$\lim_{i \rightarrow \infty} \frac{\ln L_i(t)}{N_i} = - \sum_{r=1}^{\infty} \left( \lim_{i \rightarrow \infty} \frac{1}{N_i} \sum_{j=1}^{2g_i} \alpha_{i,j}^r \right) \frac{t^r}{r},$$

provided that the limit in the RHS exists. Showing that this limit is  $-1$  yields the equality

$$\lim_{i \rightarrow \infty} \frac{\ln L_i(t)}{N_i} = \sum_{r=1}^{\infty} \frac{t^r}{r} = -\ln(1-t). \quad (4.17)$$

[Equation \(4.14\)](#) then easily follows, as for every fixed  $0 < t < 1/Q$  the expressions  $\ln Z_{F_i}(t)$  and  $\ln L_i(t)$  differ by the additive constant  $\ln(1-t) + \ln(1-Qt)$ , and

$$N_i = (\sqrt{Q} - 1 + o(1)) \cdot g_i \rightarrow \infty.$$

To establish [Equation \(4.13\)](#), note that by the class number formula and the functional equation ([[Sti09](#)], Theorem 5.1.15),

$$h_{F_i} = L_i(1) = Q^{g_i} L_i(1/Q).$$

Combining this with [Equation \(4.17\)](#) evaluated at  $t = \frac{1}{Q} < \frac{1}{\sqrt{Q}}$  yields

$$\lim_{i \rightarrow \infty} \frac{\ln h_{F_i}}{N_i} = \lim_{i \rightarrow \infty} \frac{g_i \ln Q + \ln L_i(1/Q)}{N_i} = \frac{\ln Q}{\sqrt{Q}-1} - \ln(1-1/Q).$$

At this point, it only remains to show that for all  $r \geq 1$ ,

$$\lim_{i \rightarrow \infty} \frac{1}{N_i} \sum_{j=1}^{2g_i} \alpha_{i,j}^r = -1.$$

Using [Equation \(4.15\)](#) and the equality ([[Sti09](#)], Corollary 5.1.16)

$$N_r(F_i) = Q^r + 1 - \sum_{j=1}^{2g_i} \alpha_{i,j}^r,$$

we obtain

$$\frac{1}{N_i} \sum_{j=1}^{2g_i} \alpha_{i,j}^r = \frac{Q^r + 1 - N_r(F_i)}{N_i} = \frac{Q^r + 1}{N_i} - \sum_{d|r} \frac{dB_d(F_i)}{N_i}.$$

To conclude, observe that  $\lim_{i \rightarrow \infty} \frac{Q^r+1}{N_i} = 0$  and

$$\lim_{i \rightarrow \infty} \frac{dB_d(F_i)}{N_i} = d \cdot \lim_{i \rightarrow \infty} \frac{g_i}{N_i} \frac{B_d(F_i)}{g_i} = \frac{d}{\sqrt{Q}-1} \cdot \beta_d = \begin{cases} 1 & d = 1 \\ 0 & d > 1 \end{cases}.$$

□

## 4.4 A general asymptotic rate bound

We now combine the preceding results to obtain a sufficient condition for the existence of codes with a prescribed asymptotic rate and relative distance.

Throughout this section, let  $Q$  be a square prime power, let  $q$  be an integer satisfying  $2 \leq q \leq Q$ , and let  $a \in (\frac{2}{\sqrt{Q}-1}, 1)$  be a parameter to be specified later. We set

$$\delta_0 = 1 - a \quad , \quad \kappa = \frac{\ln Q}{\sqrt{Q}-1} - \ln(1 - 1/Q)$$

and define  $\Psi_{a,q}: (0, 1)^2 \rightarrow \mathbb{R}$  by

$$\Psi_{a,q}(t, x) = h(x) - (a + x - 1) \ln t - (1 - x) \ln(1 - t) - \kappa + x \ln \frac{q - 1}{Q - 1},$$

where

$$h(x) = -x \ln x - (1 - x) \ln(1 - x).$$

We are now ready to state the following proposition.

**Proposition 4.9.** *Suppose there exist  $\delta \in (\delta_0, 1 - 1/q)$  and  $t \in (0, 1/Q)$  such that  $\Psi_{a,q}(t, \delta) < 0$ . Then there exists a family  $\{C_i\}_{i=1}^\infty$  of  $q$ -ary codes of lengths  $n_i \rightarrow \infty$ , with asymptotic relative distance  $\delta$  and asymptotic rate at least*

$$R(a, q) := 1 - \left(1 + \frac{1}{\sqrt{Q}-1} - a\right) \frac{\ln Q}{\ln q}.$$

*Proof.* Take an asymptotically optimal sequence  $(F_i/\mathbb{F}_Q)_{i=1}^\infty$  of function fields over  $\mathbb{F}_Q$ , such as the one obtained from the Garcia-Stichtenoth tower [GS95]. We denote the genus of  $F_i$  by  $g_i$  and the number of its rational places by  $N_i$ .

For  $i \geq 1$ , let  $m_i = \lfloor aN_i \rfloor$ . Since  $a > \frac{2}{\sqrt{Q}-1}$  and  $N_i/g_i \rightarrow \sqrt{Q}-1$ , there exists  $i_0$  such that  $aN_i > 2g_i$  for every  $i \geq i_0$ . Consequently,  $2g_i - 2 < m_i < N_i$  for every  $i \geq i_0$ . Fix  $i \geq i_0$ . We can now specialize the setting of Section 4.2 to the case where the function field is  $F_i/\mathbb{F}_Q$  and  $D = P_1 + \dots + P_{N_i}$  is the sum of its rational places. This gives rise to a family of codes  $\{C_\gamma^{(i)}\}_{\gamma \in \mathcal{C}_{F_i}^{m_i}}$  with

$$\dim(C_\gamma^{(i)}) = m_i + 1 - g_i \quad \text{and} \quad d(C_\gamma^{(i)}) \geq N_i - m_i. \quad (4.18)$$

Applying Corollary 4.5, we conclude that for the given  $t \in (0, 1/Q)$  and any  $w \in [N_i]$ ,

$$\mathbb{E}_\gamma [A_w(C_\gamma^{(i)})] \leq \frac{Q-1}{h_{F_i}} \binom{N_i}{w} \cdot Z_{F_i}(t) \cdot t^{-(m_i - N_i + w)} (1-t)^w \quad (4.19)$$

where the expectation is taken over  $\gamma \in \mathcal{C}_{F_i}^{m_i}$  chosen uniformly at random.

Next, let  $d_i := \lceil \delta N_i \rceil$  and consider the truncated weight enumerator  $W_q(C_\gamma^{(i)}, d_i)$  (as defined at the beginning of the section)

$$W_q(C_\gamma^{(i)}, d_i) = \sum_{w=1}^{d_i-1} A_w(C_\gamma^{(i)}) \left( \frac{q-1}{Q-1} \right)^w.$$

To proceed, we need the following claim.

**Claim 4.10.** *For all  $i$  sufficiently large, we have*

$$\mathbb{E}_\gamma [W_q(C_\gamma^{(i)}, d_i)] \leq 1,$$

where the expectation is taken over  $\gamma \in \mathcal{C}_{F_i}^{m_i}$  chosen uniformly at random.

Once this claim is established, an averaging argument shows that there exists  $\gamma \in \mathcal{C}_{F_i}^{m_i}$  for which the code  $C_\gamma^{(i)}$  satisfies  $W_q(C_\gamma^{(i)}, d_i) \leq 1$ . We can therefore apply [Lemma 4.1](#) with the code  $C_1 := C_\gamma^{(i)} \subseteq \mathbb{F}_Q^{N_i}$  and  $d := d_i$ . We conclude that there exists a  $q$ -ary code  $C_2^{(i)} \subseteq [q]^{N_i}$  with minimum distance

$$d(C_2^{(i)}) \geq d_i = \lceil \delta N_i \rceil \geq \delta N_i \quad (4.20)$$

and size

$$|C_2^{(i)}| \geq \frac{1}{2} \cdot |C_\gamma^{(i)}| \cdot \left( \frac{q}{Q} \right)^{N_i} = \frac{1}{2} \cdot Q^{m_i+1-g_i} \cdot \left( \frac{q}{Q} \right)^{N_i}. \quad (4.21)$$

The desired family of codes is then the family  $\{C_2^{(i)}\}_{i>i_0}$ . Indeed, this is a family of  $q$ -ary codes of lengths  $n_i := N_i \rightarrow \infty$ . By [Equation \(4.20\)](#), its relative distance is at least  $\delta$ . As for its rate, by [Equation \(4.21\)](#) we have

$$\frac{\log_q |C_2^{(i)}|}{N_i} \geq \frac{\log_q \frac{1}{2} + (m_i + 1 - g_i) \log_q Q + N_i(1 - \log_q Q)}{N_i}.$$

This, together with

$$\frac{g_i}{N_i} \rightarrow \frac{1}{\sqrt{Q}-1} \quad \text{and} \quad \frac{m_i}{N_i} = \frac{\lfloor aN_i \rfloor}{N_i} \rightarrow a,$$

yields the desired lower bound on the rate

$$\left( a - \frac{1}{\sqrt{Q}-1} \right) \log_q Q + (1 - \log_q Q) = 1 - \left( 1 + \frac{1}{\sqrt{Q}-1} - a \right) \frac{\ln Q}{\ln q}.$$

To complete the proof, it remains to prove [Claim 4.10](#).

*Proof of Claim 4.10.* First, note that by Equation (4.18), for all  $i$  sufficiently large and all  $1 \leq w < N_i - m_i$  we have  $A_w(C_\gamma^{(i)}) = 0$ . Thus, by linearity of expectation,

$$\mathbb{E}_\gamma [W_q(C_\gamma^{(i)}, d_i)] = \sum_{w=N_i-m_i}^{d_i-1} \mathbb{E}_\gamma [A_w(C_\gamma^{(i)})] \left( \frac{q-1}{Q-1} \right)^w.$$

Using Equation (4.19) and the basic inequality  $\binom{N_i}{w} \leq \exp(N_i \cdot h(w/N_i))$ , we obtain

$$\mathbb{E}_\gamma [W_q(C_\gamma^{(i)}, d_i)] \leq \frac{(Q-1) \cdot Z_{F_i}(t)}{h_{F_i}} \cdot t^{N_i-m_i} \sum_{w=N_i-m_i}^{d_i-1} \exp\left(N_i \cdot h\left(\frac{w}{N_i}\right)\right) \cdot \left(\frac{1-t}{t}\right)^w \left(\frac{q-1}{Q-1}\right)^w.$$

Observe that as  $i \rightarrow \infty$ , Lemma 4.8 gives

$$\ln \frac{Z_{F_i}(t)}{h_{F_i}} = N_i \left( \frac{\ln Z_{F_i}(t)}{N_i} - \frac{\ln h_{F_i}}{N_i} \right) = N_i \cdot (o(1) - \ln(1-t) - \kappa).$$

Similarly, since  $m_i/N_i \rightarrow a$ ,

$$\ln(t^{N_i-m_i}) = N_i \left(1 - \frac{m_i}{N_i}\right) \ln t = N_i \cdot (1-a+o(1)) \ln t.$$

Combining, we obtain

$$\begin{aligned} \mathbb{E}_\gamma [W_q(C_\gamma^{(i)}, d_i)] &\leq \exp(N_i \cdot [o(1) - \ln(1-t) - \kappa + (1-a) \ln t]) \\ &\quad \cdot \sum_{w=N_i-m_i}^{d_i-1} \exp\left(N_i \cdot \left[h\left(\frac{w}{N_i}\right) + \frac{w}{N_i} \ln \frac{1-t}{t} + \frac{w}{N_i} \ln \frac{q-1}{Q-1}\right]\right) \end{aligned}$$

which, after rearranging, yields

$$\mathbb{E}_\gamma [W_q(C_\gamma^{(i)}, d_i)] \leq \sum_{w=N_i-m_i}^{d_i-1} \exp\left(N_i \left[o(1) + \Psi_{a,q}\left(t, \frac{w}{N_i}\right)\right]\right). \quad (4.22)$$

We proceed by showing that for our fixed given  $t$ , the function  $\Phi_{a,q}(x) := \Psi_{a,q}(t, x)$  is strictly increasing on the interval  $[1-a, \delta]$ . Differentiating, we obtain

$$\begin{aligned} \Phi'_{a,q}(x) &= h'(x) - \ln t + \ln(1-t) + \ln \frac{q-1}{Q-1} \\ &= -\ln x + \ln(1-x) + \ln \frac{1-t}{t} + \ln \frac{q-1}{Q-1} \\ &= \ln \left( \frac{1-x}{x} \cdot \frac{1-t}{t} \cdot \frac{q-1}{Q-1} \right). \end{aligned}$$

Observe that if  $x \in [1-a, \delta]$  then  $x \leq \delta < 1-1/q$ , hence  $\frac{1-x}{x} > \frac{1}{q-1}$ . Similarly, since  $0 < t < 1/Q$  it holds that  $\frac{1-t}{t} > Q-1$ . Together, we obtain that for all  $x \in [1-a, \delta]$ ,

$$\frac{1-x}{x} \cdot \frac{1-t}{t} \cdot \frac{q-1}{Q-1} > 1.$$

Thus,  $\Phi'_{a,q}(x) > 0$  on this interval, implying that  $\Phi_{a,q}$  is strictly increasing there.

Returning to [Equation \(4.22\)](#) and recalling that  $m_i = \lfloor aN_i \rfloor$  and  $d_i = \lceil \delta N_i \rceil$ , we see that for all  $w$  satisfying  $N_i - m_i \leq w \leq d_i - 1$ , we have that  $\frac{w}{N_i} \in [1 - a, \delta)$ . By the monotonicity of  $\Phi_{a,q}$ , it follows that

$$\Psi_{a,q}\left(t, \frac{w}{N_i}\right) = \Phi_{a,q}\left(\frac{w}{N_i}\right) \leq \Phi_{a,q}(\delta) = \Psi_{a,q}(t, \delta).$$

Since the RHS of [Equation \(4.22\)](#) contains at most  $N_i$  summands and the  $o(1)$ -term is independent of  $w$ , we conclude that

$$\mathbb{E}_\gamma \left[ W_q \left( C_\gamma^{(i)}, d_i \right) \right] \leq N_i \cdot \exp \left( N_i \cdot [o(1) + \Psi_{a,q}(t, \delta)] \right).$$

Recall that by the assumption,  $c := \Psi_{a,q}(t, \delta) < 0$ . For all  $i$  sufficiently large, the  $o(1)$ -term is at most  $-\frac{c}{2}$ , so that

$$\mathbb{E}_\gamma \left[ W_q \left( C_\gamma^{(i)}, d_i \right) \right] \leq N_i \exp(cN_i/2) \longrightarrow 0 \quad \text{as } i \rightarrow \infty.$$

Thus, for all sufficiently large  $i$  we have  $\mathbb{E}_\gamma \left[ W_q \left( C_\gamma^{(i)}, d_i \right) \right] \leq 1$ , establishing [Claim 4.10](#).  $\square$

This completes the proof of [Proposition 4.9](#).  $\square$

## 4.5 Choosing the parameters

We now apply [Proposition 4.9](#) with  $Q = 49$  and  $q = 42$  to prove [Theorem 1.1](#). Write

$$R(a) := R(a, 42) = 1 - \left( \frac{7}{6} - a \right) \frac{\ln 49}{\ln 42}, \quad I := [0.39, 0.48].$$

Also, denote

$$\Psi_a(t, x) := \Psi_{a,42}(t, x).$$

For  $\delta \in I$ , define

$$a_{\text{GV}}(\delta) := \frac{7}{6} - \frac{h(\delta) + \delta \ln 41}{\ln 49},$$

so that

$$R(a_{\text{GV}}(\delta)) = 1 - H_{42}(\delta).$$

We will show that, for this choice of  $a$  and a suitable  $t < 1/49$ , it holds that  $\Psi_a(t, \delta) < 0$ . We can then slightly increase  $a$  to obtain a rate strictly above GV.

Although [Proposition 4.9](#) requires  $t < 1/49$ , the expression defining  $\Psi_a(t, \delta)$  is also well-defined at  $t = 1/49$ . Substituting this value gives

$$\Psi_a(1/49, \delta) = h(\delta) + \delta \ln 41 - \left( \frac{7}{6} - a \right) \ln 49.$$

In particular,

$$\Psi_{a_{\text{GV}}(\delta)}(1/49, \delta) = 0.$$

Differentiating with respect to  $t$ , we obtain

$$\begin{aligned} \left. \frac{\partial}{\partial t} \Psi_{a_{\text{GV}}(\delta)}(t, \delta) \right|_{t=1/49} &= 49 \left( \frac{49}{48}(1 - \delta) - a_{\text{GV}}(\delta) \right) \\ &= \frac{49}{\ln 49} F(\delta), \end{aligned}$$

where

$$F(\delta) := h(\delta) + \delta \ln 41 - \frac{49\delta + 7}{48} \ln 49.$$

Thus, positivity of  $F(\delta)$  means that decreasing  $t$  slightly from  $1/49$  makes  $\Psi_{a_{\text{GV}}(\delta)}(t, \delta)$  negative.

The function  $F$  is strictly concave, since

$$F''(\delta) = -\frac{1}{\delta(1 - \delta)} < 0.$$

A direct calculation gives

$$F(0.39) \approx 5.31098734 \cdot 10^{-5} > 0, \quad F(0.48) \approx 3.12486226 \cdot 10^{-4} > 0.$$

Consequently,  $F(\delta) > 0$  throughout  $I$ .

We also check that the parameter  $a_{\text{GV}}(\delta)$  is in the required range  $(1/3, 1)$ . To this end, note that the function  $h(x) + x \ln(41/49)$  attains its maximum at  $x = 41/90$ , where its value is  $\ln(90/49)$ . Since  $90^6 < 49^7$ , we obtain

$$h(\delta) + \delta \ln \frac{41}{49} \leq \ln \frac{90}{49} < \frac{\ln 49}{6}.$$

Therefore,  $a_{\text{GV}}(\delta) > 1 - \delta$ . In particular,  $a_{\text{GV}}(\delta) \geq 0.52 > 1/3$  for every  $\delta \in I$ .

On the other hand, since

$$\frac{49}{48}(1 - \delta) - a_{\text{GV}}(\delta) = \frac{F(\delta)}{\ln 49} > 0,$$

we have

$$a_{\text{GV}}(\delta) < \frac{49}{48}(1 - \delta) \leq \frac{49}{48} \cdot 0.61 < 1.$$

It remains to make the improvement uniform in  $\delta$ . The derivative computed above is continuous and strictly positive on the compact interval  $I$ . Therefore, a single  $t_0 < 1/49$ ,

sufficiently close to  $1/49$ , satisfies

$$\Psi_{a_{\text{GV}}(\delta)}(t_0, \delta) < 0 \quad \text{for every } \delta \in I.^2$$

By continuity and compactness, there is then a constant  $\varepsilon > 0$  such that, for every  $\delta \in I$ , the choice

$$a(\delta) := a_{\text{GV}}(\delta) + \varepsilon$$

still satisfies  $a(\delta) \in (1/3, 1)$ ,  $1 - a(\delta) < \delta$ , and

$$\Psi_{a(\delta)}(t_0, \delta) < 0.$$

All the hypotheses of [Proposition 4.9](#) now hold. It yields a family of 42-ary codes with asymptotic relative distance at least  $\delta$  and asymptotic rate at least

$$R(a(\delta)) = 1 - H_{42}(\delta) + \varepsilon \cdot \frac{\ln 49}{\ln 42}.$$

Taking  $\beta := \varepsilon \ln 49 / \ln 42 > 0$ , independently of  $\delta$ , completes the proof of [Theorem 1.1](#).

## AI Usage

This disclosure is included as part of the TCS community’s initial efforts to adapt to AI use, and in particular in response to the recommendations of the September 2026 report *AI and TCS: The Next Six Months* [\[Sep26\]](#).

The authors used AI (GPT-6 Astra) at several stages of this project. Initially, it helped evaluate a range of ideas the authors had developed over the years and wrote code to test these ideas computationally on specific function fields. Several promising ideas were then explored through a sustained series of exchanges with the AI, interspersed with independent work by the authors to develop new ideas and guide further exploration. This process led to a successful combination that yielded the desired result, first for an alphabet of size 43 and then 42. Most of the key ideas were developed by the authors, but the AI made a substantive contribution by combining and refining them and by identifying results in the literature that had already developed some of the ideas the authors had in mind.

---

<sup>2</sup>Write  $f(\delta, t) := \Psi_{a_{\text{GV}}(\delta)}(t, \delta)$  and  $b := 1/49$ . We have  $f(\delta, b) = 0$  and  $\partial_t f(\delta, b) > 0$  for every  $\delta \in I$ . Since  $I$  is compact and  $\partial_t f(\delta, b)$  is continuous in  $\delta$ , its minimum  $c$  on  $I$  is strictly positive. Moreover,  $\partial_t f$  is jointly continuous in  $(\delta, t)$ , hence uniformly continuous on a compact rectangle  $I \times [b/2, b]$ . Thus, there exists  $0 < \eta < b/2$ , independent of  $\delta$ , such that  $\partial_t f(\delta, t) \geq c/2$  whenever  $\delta \in I$  and  $b - \eta \leq t \leq b$ . Setting  $t_0 = b - \eta$ , the mean value theorem gives  $f(\delta, t_0) \leq f(\delta, b) - (c/2)\eta = -c\eta/2 < 0$  for every  $\delta \in I$ .

AI also helped draft the exposition (see [Section A](#)) from lecture notes for a course taught at Tel Aviv University by the first author, with the second author serving as teaching assistant.<sup>3</sup> It also assisted with checking the literature and revising the language and presentation. The paper, and in particular its proofs, was written by the authors with the AI assistance described above. The authors take full responsibility for its contents and results.

## References

- [AG26] Omar Alrabiah and Venkatesan Guruswami. Binary code rate bounds via classical–quantum channels. *arXiv preprint [arXiv:2608.09347](#)*, 2026.
- [BBGS14] Alp Bassa, Peter Beelen, Arnaldo Garcia, and Henning Stichtenoth. An improvement of the Gilbert–Varshamov bound over nonprime fields. *IEEE Transactions on Information Theory*, 60(7):3859–3861, 2014.
- [BBGS15] Alp Bassa, Peter Beelen, Arnaldo Garcia, and Henning Stichtenoth. Towers of function fields over non-prime finite fields. *Moscow Mathematical Journal*, 15(1):1–29, 2015.
- [Bee22] Peter Beelen. A survey on recursive towers and Ihara’s constant, 2022.
- [BSGK<sup>+</sup>13] Eli Ben-Sasson, Ariel Gabizon, Yohay Kaplan, Swastik Kopparty, and Shubhangi Saraf. A new family of locally correctable codes based on degree-lifted algebraic geometry codes. In *Proceedings of the 45th Annual ACM Symposium on Theory of Computing (STOC)*, 2013.
- [BSKK<sup>+</sup>13] Eli Ben-Sasson, Yohay Kaplan, Swastik Kopparty, Or Meir, and Henning Stichtenoth. Constant rate PCPs for Circuit-SAT with sublinear query complexity. In *Proceedings of the 54th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 320–329, 2013.
- [CC06] Hao Chen and Ronald Cramer. Algebraic geometric secret sharing schemes and secure multi-party computations over small fields. In *Advances in Cryptology—CRYPTO 2006*, volume 4117 of *Lecture Notes in Computer Science*, pages 521–536. Springer, 2006.

---

<sup>3</sup>The course is aimed at students of theoretical computer science. Lecture notes and links to video lectures on YouTube are available at <https://www.gilcohen.org/2024-agc>.

- [CC26] Gil Cohen and Itay Cohen. Wide replacement products meet Gray codes: Toward optimal small-bias sets. In *41st Computational Complexity Conference (CCC 2026)*, volume 383 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 21:1–21:27. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2026.
- [CDG26] Gil Cohen, Dean Doron, and Noam Goldgraber. Algebraic-geometric Parvaresh–Vardy subspace designs and rank condensers. Technical Report TR26-214, Electronic Colloquium on Computational Complexity (ECCC), 2026.
- [CDGM26] Gil Cohen, Dean Doron, Noam Goldgraber, and Tomer Manket. Tracing AG codes: Toward meeting the Gilbert–Varshamov bound. In Sayan Bhat-tacharya, Danupon Nanongkai, Michael Benedikt, and Gabriele Puppis, editors, *53rd International Colloquium on Automata, Languages, and Programming (ICALP 2026)*, volume 374 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 68:1–68:23. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2026.
- [CR20] Alain Couvreur and Hugues Randriambololona. Algebraic geometry codes and some applications. *arXiv preprint [arXiv:2009.01281](https://arxiv.org/abs/2009.01281)*, 2020.
- [Duu99] Iwan M. Duursma. Weight distributions of geometric Goppa codes. *Transactions of the American Mathematical Society*, 351(9):3609–3639, 1999.
- [Duu08] Iwan M. Duursma. Algebraic geometry codes: General theory. In Edgar Martínez-Moro, Carlos Munuera, and Diego Ruano, editors, *Advances in Algebraic Geometry Codes*, pages 1–48. World Scientific, 2008.
- [Elk98] Noam D. Elkies. Explicit modular towers. In *Proceedings of the Thirty-Fifth Annual Allerton Conference on Communication, Control and Computing*, pages 23–32. University of Illinois at Urbana-Champaign, 1998.
- [Elk01] Noam D. Elkies. Excellent nonlinear codes from modular curves. In *Proceedings of the 33rd Annual ACM Symposium on Theory of Computing (STOC)*, pages 200–208. ACM, 2001.
- [Elk03] Noam D. Elkies. Still better nonlinear codes from modular curves. *arXiv:math/0308046*, 2003.

- [Erd47] Paul Erdős. Some remarks on the theory of graphs. *Bulletin of the American Mathematical Society*, 53(4):292–294, 1947.
- [Gil52] Edgar N. Gilbert. A comparison of signalling alphabets. *Bell System Technical Journal*, 31(3):504–522, 1952.
- [GJL26] William Gay, Fernando Granha Jeronimo, and Lenny Liu. The Honeycomb framework for code bounds. *arXiv preprint [arXiv:2608.20287](https://arxiv.org/abs/2608.20287)*, 2026.
- [GK13] Venkatesan Guruswami and Swastik Kopparty. Explicit subspace designs. In *Proceedings of the 54th Annual IEEE Symposium on Foundations of Computer Science (FOCS)*, pages 608–617, 2013.
- [Gop81] V. D. Goppa. Codes on algebraic curves. *Doklady Akademii Nauk SSSR*, 259(6):1289–1290, 1981. In Russian; English translation in *Soviet Mathematics Doklady*, 24 (1981), 170–172.
- [GS95] Arnaldo Garcia and Henning Stichtenoth. A tower of Artin–Schreier extensions of function fields attaining the Drinfeld–Vlăduţ bound. *Inventiones Mathematicae*, 121:211–222, 1995.
- [GS96] Arnaldo Garcia and Henning Stichtenoth. On the asymptotic behaviour of some towers of function fields over finite fields. *Journal of Number Theory*, 61(2):248–273, 1996.
- [GS99] Venkatesan Guruswami and Madhu Sudan. Improved decoding of Reed–Solomon and algebraic-geometry codes. *IEEE Transactions on Information Theory*, 45(6):1757–1767, 1999.
- [GSR03] Arnaldo Garcia, Henning Stichtenoth, and Hans-Georg Rück. On tame towers over finite fields. *Journal für die reine und angewandte Mathematik*, 557:53–80, 2003.
- [GX12] Venkatesan Guruswami and Chaoping Xing. Folded codes from function field towers and improved optimal rate list decoding. In *Proceedings of the 44th Annual ACM Symposium on Theory of Computing (STOC)*, 2012.
- [GX13] Venkatesan Guruswami and Chaoping Xing. List decoding Reed–Solomon, algebraic-geometric, and Gabidulin subcodes up to the Singleton bound. In *Proceedings of the 45th Annual ACM Symposium on Theory of Computing (STOC)*, pages 843–852, 2013.

- [GX14] Venkatesan Guruswami and Chaoping Xing. Optimal rate list decoding of folded algebraic-geometric codes over constant-sized alphabets. In *Proceedings of the 25th Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 1858–1866, 2014.
- [GXY17] Venkatesan Guruswami, Chaoping Xing, and Chen Yuan. Subspace designs based on algebraic function fields. In *44th International Colloquium on Automata, Languages, and Programming (ICALP)*, volume 80 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 86:1–86:10. Schloss Dagstuhl–Leibniz-Zentrum für Informatik, 2017.
- [GZ08] Philippe Gaborit and Gilles Zémor. Asymptotic improvement of the Gilbert–Varshamov bound for linear codes. *IEEE Transactions on Information Theory*, 54(9):3865–3872, 2008.
- [Ham50] Richard W. Hamming. Error detecting and error correcting codes. *Bell System Technical Journal*, 29(2):147–160, 1950.
- [HHLT22] Jing Hao, Han Huang, Galyna Livshyts, and Konstantin Tikhomirov. Distribution of the minimum distance of random linear codes. *IEEE Transactions on Information Theory*, 68(10):6388–6401, 2022.
- [HMZ26] Jun-Ting Hsieh, Sidhanth Mohanty, and Rachel Yun Zhang. Explicit almost-optimal  $\varepsilon$ -balanced codes via free expander walks. *arXiv preprint [arXiv:2601.12606](https://arxiv.org/abs/2601.12606)*, 2026.
- [Iha81] Yasutaka Ihara. Some remarks on the number of rational points of algebraic curves over finite fields. *Journal of the Faculty of Science, University of Tokyo, Section IA, Mathematics*, 28(3):721–724, 1981.
- [JM21] Akhil Jalan and Dana Moshkovitz. Near-optimal Cayley expanders for abelian groups. In *41st IARCS Annual Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS 2021)*, volume 213 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 24:1–24:23. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2021.
- [JV04] Tao Jiang and Alexander Vardy. Asymptotic improvement of the Gilbert–Varshamov bound on the size of binary codes. *IEEE Transactions on Information Theory*, 50(8):1655–1664, 2004.

- [KTSY26] Swastik Kopparty, Amnon Ta-Shma, and Kedem Yakirevitch. Trace Hermitian codes have vanishing bias. In Dana Moshkovitz, editor, *41st Computational Complexity Conference (CCC 2026)*, volume 383 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 35:1–35:28. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, 2026.
- [LT86] S. N. Litsyn and M. A. Tsfasman. A note on lower bounds. *IEEE Transactions on Information Theory*, 32(5):705–706, 1986.
- [LWX23] Shu Liu, Tingyi Wu, and Chaoping Xing. Nonlinear codes exceeding the Gilbert–Varshamov and Tsfasman–Vlăduț–Zink bounds. In *Proceedings of the 2023 Annual ACM–SIAM Symposium on Discrete Algorithms (SODA)*, pages 4104–4114. SIAM, 2023.
- [MRRW77] Robert J. McEliece, Eugene R. Rodemich, Howard C. Rumsey, Jr., and Lloyd R. Welch. New upper bounds on the rate of a code via the Delsarte–MacWilliams inequalities. *IEEE Transactions on Information Theory*, 23(2):157–166, 1977.
- [NÖ07] Harald Niederreiter and Ferruh Özbudak. Improved asymptotic bounds for codes using distinguished divisors of global function fields. *SIAM Journal on Discrete Mathematics*, 21(4):865–899, 2007.
- [NX09] Harald Niederreiter and Chaoping Xing. *Algebraic Geometry in Coding Theory and Cryptography*. Princeton University Press, 2009.
- [Ope26] OpenAI. Improved bounds for binary and spherical codes. In *Ten Advances in Mathematics and Theoretical Computer Science*, chapter 2. OpenAI, August 2026.
- [Rub18] Aviad Rubinfeld. Hardness of approximate nearest neighbor search. In *Proceedings of the 50th Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, pages 1260–1268, 2018.
- [Sal26] Andrew Salmon. The half-rate linear programming bound for binary codes is  $\frac{1}{2} - \frac{1}{\pi}$ . *arXiv preprint [arXiv:2609.03736](https://arxiv.org/abs/2609.03736)*, 2026.
- [Sep26] September 2026 Working Group. AI and TCS: The next six months. Working group report, Simons Institute for the Theory of Computing, September 2026.
- [Sha48] Claude E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27(3):379–423, 1948.

- [Sti09] Henning Stichtenoth. *Algebraic Function Fields and Codes*, volume 254 of *Graduate Texts in Mathematics*. Springer, 2 edition, 2009.
- [SW99] M. Amin Shokrollahi and Hal Wasserman. List decoding of algebraic-geometric codes. *IEEE Transactions on Information Theory*, 45(2):432–437, 1999.
- [TS17] Amnon Ta-Shma. Explicit, almost optimal, epsilon-balanced codes. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing (STOC)*, pages 238–251. ACM, 2017.
- [Tsf92] Michael A. Tsfasman. Some remarks on the asymptotic number of points. In Henning Stichtenoth and Michael A. Tsfasman, editors, *Coding Theory and Algebraic Geometry*, volume 1518 of *Lecture Notes in Mathematics*, pages 178–192. Springer, Berlin, 1992.
- [TV02] Michael A. Tsfasman and Serge G. Vlăduț. Infinite global fields and the generalized Brauer–Siegel theorem. *Moscow Mathematical Journal*, 2(2):329–402, 2002.
- [TVN07] Michael Tsfasman, Serge Vlăduț, and Dmitry Nogin. *Algebraic Geometric Codes: Basic Notions*, volume 139 of *Mathematical Surveys and Monographs*. American Mathematical Society, 2007.
- [TVN19] Michael Tsfasman, Serge Vlăduț, and Dmitry Nogin. *Algebraic Geometry Codes: Advanced Chapters*, volume 238 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2019.
- [TVZ82] M. A. Tsfasman, S. G. Vlăduț, and Thomas Zink. Modular curves, Shimura curves, and Goppa codes, better than Varshamov–Gilbert bound. *Mathematische Nachrichten*, 109(1):21–28, 1982.
- [Var57] R. R. Varshamov. The evaluation of signals in codes with correction of errors. *Doklady Akademii Nauk SSSR*, 117(5):739–741, 1957. In Russian.
- [VD83] S. G. Vlăduț and V. G. Drinfeld. Number of points of an algebraic curve. *Functional Analysis and Its Applications*, 17(1):53–54, 1983.
- [VKT84] S. G. Vlăduț, G. L. Katsman, and M. A. Tsfasman. Modular curves and codes with polynomial construction complexity. *Problems of Information Transmission*, 20(1):35–42, 1984.

- [VW05] Van H. Vu and Lei Wu. Improving the Gilbert–Varshamov bound for  $q$ -ary codes. *IEEE Transactions on Information Theory*, 51(9):3200–3208, 2005.
- [Xin01] Chaoping Xing. Algebraic-geometry codes with asymptotic parameters better than the Gilbert–Varshamov and the Tsfasman–Vlăduț–Zink bounds. *IEEE Transactions on Information Theory*, 47(1):347–352, 2001.
- [Xin03] Chaoping Xing. Nonlinear codes from algebraic curves improving the Tsfasman–Vlăduț–Zink bound. *IEEE Transactions on Information Theory*, 49(7):1653–1657, 2003.
- [YQ09] Siman Yang and Lulu Qi. On improved asymptotic bounds for codes from global function fields. *Designs, Codes and Cryptography*, 53(1):33–43, 2009.
- [YZ26] Chen Yuan and Ruiqi Zhu. Improvement of the Gilbert–Varshamov bound for linear codes and quantum codes, 2026. arXiv:2601.18590.
- [ZL85] V. A. Zinoviev and S. N. Litsyn. On codes beyond the Gilbert bound. *Problemy Peredachi Informatsii*, 21(1):109–111, 1985. In Russian.

## A A friendly exposition of algebraic-geometric codes

In this section we give an exposition of algebraic-geometric codes, covering the notions and results used in this paper. We have made every effort to keep the exposition accessible, emphasizing the ideas and using a running example to make them concrete. Our hope is that this appendix will serve as an invitation to the broader theory community to explore these ideas and push further toward beating the GV bound over smaller alphabets.

The construction of algebraic-geometric codes starts with a simple idea: identify each message with a function and encode it by listing the function’s values at a collection of evaluation locations. For the familiar Reed–Solomon codes, the functions are low-degree polynomials over a finite field, and the locations are distinct field elements. A polynomial of degree at most  $m$  is specified by  $m + 1$  coefficients. The elementary polynomial root bound holds over any field: a nonzero polynomial of degree at most  $m$  has at most  $m$  roots. Applying this to the difference of two such polynomials shows that they can agree at no more than  $m$  locations unless they are identical. These observations give the dimension and distance of the code. The limitation is the number of available evaluation locations: over  $\mathbb{F}_Q$ , there are only  $Q$  field elements to choose from.

To obtain longer codes over the same field, we need more evaluation locations while retaining control over how often two distinct functions can agree. One way to seek both is

to replace univariate polynomials by functions of several variables constrained by algebraic equations. The solutions to these equations provide the evaluation locations, and suitably chosen spaces of functions provide the messages. Algebraic curves give a setting in which this idea works. Through a running example, we will explain how their algebraic structure controls the dimension and distance of the resulting codes, introducing places, divisors, Riemann–Roch spaces, and genus along the way. The same example will later illustrate the zeta function and its connection to counting.

Algebraic-geometric codes can be described in two closely related languages: the geometric language of algebraic curves and the algebraic language of function fields. Curves provide a picture of points and functions evaluated at them, while function fields let us study these functions without committing to a particular system of equations. Different equations may describe different models of the same underlying curve, and the function-field viewpoint helps separate the properties that matter from the details of a chosen model. We will use this algebraic language throughout, drawing on geometric intuition along the way.

We follow the notation and conventions of Stichtenoth [Sti09]. The general theorems will be stated without proof, but accompanied by explanations and calculations in our running example. Throughout,  $Q$  denotes the size of a finite field, whereas  $q$  denotes the size of the final alphabet and need not be a prime power. Our example is over  $\mathbb{F}_Q$  with  $Q = 49$ , the field used in the proof before reducing the alphabet to  $q = 42$  symbols.

## A.1 An introduction to algebraic function fields

### A.1.1 A function field to keep in mind

An *algebraic function field*  $F/\mathbb{F}_Q$  is a finite extension of the rational function field  $\mathbb{F}_Q(x)$ , where  $x$  is transcendental over  $\mathbb{F}_Q$ .<sup>4</sup> Thus  $x$  is an indeterminate, rather than an element of a finite field, and should be thought of as a function we can evaluate. The rational function field  $\mathbb{F}_Q(x)$  consists of ratios of polynomials in  $x$  with coefficients in  $\mathbb{F}_Q$ . This is already the setting for Reed–Solomon codes, though they use only polynomials of bounded degree. More generally, the elements of  $F$  are the functions we will eventually evaluate.

---

<sup>4</sup>We assume throughout that  $\mathbb{F}_Q$  is the *full field of constants*: every element of  $F$  algebraic over  $\mathbb{F}_Q$  already belongs to  $\mathbb{F}_Q$ . This rules out a hidden enlargement of the field of constants. For example,  $\mathbb{F}_{Q^2}(x)$  is a finite extension of  $\mathbb{F}_Q(x)$ , but its full field of constants is  $\mathbb{F}_{Q^2}$ . The reader may safely disregard this technicality on a first reading.

**Two steps of the modular construction.** Our running example is obtained by making two quadratic extensions:

$$\begin{aligned} F_0 &= \mathbb{F}_{49}(x), & F_1 &= F_0(y), & E &= F_1(z), \\ y^2 &= \frac{x^2 + 1}{2x}, & z^2 &= \frac{y^2 + 1}{2y}. \end{aligned} \tag{A.1}$$

These are the first two extensions in the recursive tower studied by Garcia and Stichtenoth, with an appendix by Rück [GSR03]. This tower is related to the explicit modular towers of Elkies [Elk98]. For the running example we stop here and study the fixed field  $E$ . Each step has degree two, so  $[E : \mathbb{F}_{49}(x)] = 4$ .

There is a particularly convenient way to describe the same field. Set

$$u = \frac{3xyz}{x+1}. \tag{A.2}$$

Using the two relations above, and doing the arithmetic in characteristic seven, gives <sup>5</sup>

$$E = \mathbb{F}_{49}(x, u), \quad u^4 = x^3 + x. \tag{A.3}$$

Conversely,

$$y = \frac{2u^2}{x}, \quad z = -\frac{x+1}{u},$$

so the two descriptions generate the same field. In fact, this is another model of the classical Fermat quartic function field. Indeed, in characteristic seven we have

$$(x+1)^4 - (x-1)^4 = x^3 + x = u^4,$$

so the change of variables  $s = (x-1)/(x+1)$  and  $t = u/(x+1)$  gives  $s^4 + t^4 = 1$ . The inverse formulas  $x = (1+s)/(1-s)$  and  $u = 2t/(1-s)$  show that the two equations describe the same function field. All of these are identities of functions; a denominator in one expression does not mean that the underlying location has been removed. When discussing evaluation, we will keep track of whether the function itself has a pole.

The polynomial  $U^4 - (x^3 + x)$  is irreducible over  $\mathbb{F}_{49}(x)$ . One can see this from Eisenstein's criterion at the factor  $x$  of  $\mathbb{F}_{49}[x]$ : its constant coefficient is divisible by  $x$  but not by  $x^2$ . Consequently, every element of  $E$  can be written uniquely as

$$a_0(x) + a_1(x)u + a_2(x)u^2 + a_3(x)u^3, \quad a_j(x) \in \mathbb{F}_{49}(x).$$

We will use the quadratic description to see the extension process and the quartic description to make calculations. In particular, the relation  $u^4 = x^3 + x$  lets us reduce polynomial expressions to  $u$ -degree at most three. What is still missing is an analogue of polynomial degree, together with a careful account of where these functions can be evaluated.

---

<sup>5</sup>The first relation gives  $y^2 + 1 = (x+1)^2/(2x)$ . Substituting this and the expression for  $z^2$  into the definition of  $u$  yields  $u^2 = 4xy$ , and hence  $u^4 = 16x^2y^2 = 8x(x^2 + 1) = x^3 + x$ , where the constants are reduced modulo seven.

### A.1.2 Places: where functions are evaluated

Geometrically, a place corresponds to a point on a (smooth projective model of the) curve, possibly defined over an extension of the base field.<sup>6</sup> To describe such a location algebraically, we look at how functions behave there. A function may have a pole, vanish, or take a nonzero finite value; in the first two cases, we also record the order of the pole or zero. Quite surprisingly, this information suffices to recover the point itself. This illustrates an elegant and recurring idea in mathematics: an object can be characterized by how functions behave on it, without choosing coordinates or a particular representation. A basic theorem makes this viewpoint precise: the normalized discrete valuations of the function field that are trivial on the constants correspond exactly to the points of its smooth projective model.

We therefore describe a place  $P$  through its associated *discrete valuation*

$$\nu_P : F^\times \longrightarrow \mathbb{Z}.$$

The value  $\nu_P(f) = r > 0$  means that  $f$  has a zero of order  $r$  at  $P$ , while  $\nu_P(f) = -r < 0$  means that it has a pole of order  $r$ . If  $\nu_P(f) = 0$ , then  $f$  can be evaluated at  $P$  and its value is nonzero. Thus, the sign tells us which of the three behaviors occurs, and the magnitude records the order of vanishing or the pole order. We set  $\nu_P(0) = +\infty$ , reflecting the fact that the zero function vanishes to every order at every place. The set of all places of a function field  $F$  is denoted  $\mathbb{P}_F$ .

Valuations satisfy  $\nu_P(fh) = \nu_P(f) + \nu_P(h)$  and  $\nu_P(f + h) \geq \min\{\nu_P(f), \nu_P(h)\}$ . The second inequality allows for cancellation: adding functions may increase their order of vanishing, but cannot decrease it below the smaller of their two orders. We normalize the valuation so that some function  $t$  satisfies  $\nu_P(t) = 1$ . Such a function is called a *local parameter*; it plays the role of  $x - \alpha$  at an ordinary evaluation point.

The valuation tells us whether a function can be evaluated at  $P$ , but we still need to explain what its value is. We begin with all the functions that have no pole at  $P$ . These form a ring,

$$\mathcal{O}_P = \{f \in F : \nu_P(f) \geq 0\},$$

called the *valuation ring* at  $P$ . Two such functions should have the same value at  $P$  precisely when their difference vanishes there. Thus, to define evaluation, we identify functions whose difference has positive valuation.

The functions vanishing at  $P$  form an ideal of  $\mathcal{O}_P$ , which, following Stichtenoth's convention, we also denote by  $P$ . This is the unique maximal ideal of  $\mathcal{O}_P$ , and the

---

<sup>6</sup>More precisely, points over an algebraic closure that are conjugate under Frobenius correspond to the same place. Rational places correspond to individual  $\mathbb{F}_Q$ -rational points.

quotient

$$F_P = \mathcal{O}_P/P$$

is therefore a field, called the *residue field* at  $P$ . Its elements are the possible values of functions at  $P$ . For  $f \in \mathcal{O}_P$ , we denote its residue class by  $f(P)$ . In particular,  $f(P) = 0$  exactly when  $\nu_P(f) > 0$ .

For the rational function field  $\mathbb{F}_Q(x)$ , this recovers ordinary evaluation. The place  $P_\alpha$  corresponding to  $x = \alpha \in \mathbb{F}_Q$  has residue field  $\mathbb{F}_Q$ , and if  $f = a(x)/b(x)$  with  $b(\alpha) \neq 0$ , then

$$f(P_\alpha) = \frac{a(\alpha)}{b(\alpha)}.$$

The abstract definition thus extends the familiar operation of substituting a field element for  $x$ .

In general, the values at a place need not lie in  $\mathbb{F}_Q$  itself. A basic theorem guarantees that  $F_P$  is a finite extension of  $\mathbb{F}_Q$ , and we call its extension degree the *degree of the place*:

$$\deg P = [F_P : \mathbb{F}_Q], \quad |F_P| = Q^{\deg P}.$$

A place of degree one is called *rational*. At such a place, evaluation takes values in  $\mathbb{F}_Q$ , so each rational place provides a potential coordinate for a code over this field. These are the places we will use for evaluation. Higher-degree places will also play a role, but only later, when we count divisors.

**The rational function field.** Let us make these notions concrete in  $\mathbb{F}_Q(x)$ . At an ordinary evaluation point  $x = \alpha$ , the order of a zero or pole is determined by how many factors of  $x - \alpha$  occur in the numerator and denominator. For example,  $(x - \alpha)^3/(x - \beta)$  has valuation 3 at  $P_\alpha$  and valuation  $-1$  at  $P_\beta$ , provided  $\alpha \neq \beta$ .

The same idea works with any monic irreducible polynomial  $p(x)$  in place of  $x - \alpha$ : count its multiplicity in the numerator and subtract its multiplicity in the denominator. This defines a place whose residue field is  $\mathbb{F}_Q[x]/(p)$  and whose degree is therefore  $\deg p$ . Thus, the degree of a place extends the familiar notion of the degree of an irreducible polynomial. One can think of evaluation at this place as substituting a root of  $p$ , which may require passing to an extension of  $\mathbb{F}_Q$ . The familiar evaluation points are exactly the cases  $p(x) = x - \alpha$ , where the root already lies in  $\mathbb{F}_Q$ .

Have we found all the rational places? We have already accounted for every point  $x = \alpha$  on the affine line over  $\mathbb{F}_Q$ , so it may seem that there is nowhere else to evaluate. The natural geometric object associated with  $\mathbb{F}_Q(x)$ , however, is the *projective line*, which includes one additional point at infinity. Algebraically, this point gives one more rational place, which we can describe by introducing the coordinate  $t = 1/x$ . Infinity in

the  $x$ -coordinate becomes zero in the  $t$ -coordinate, so we can study functions there by substituting  $x = 1/t$  and examining their behavior at  $t = 0$ . For a nonzero polynomial  $f$  of degree  $m$ , the lowest power of  $t$  appearing in  $f(1/t)$  is  $t^{-m}$ . Thus,  $f$  has a pole of order  $m$  at infinity when  $m > 0$ . Applying this observation to the numerator and denominator of a rational function gives

$$\nu_{\infty}\left(\frac{a(x)}{b(x)}\right) = \deg b - \deg a.$$

In particular, for every nonzero polynomial  $f$ ,

$$\deg f = -\nu_{\infty}(f).$$

The additional place therefore gives us more than another evaluation location: it reveals that polynomial degree is precisely a pole order. This is the connection we will use to formulate an analogue of polynomial degree in more general function fields.

The place at infinity is rational: its residue field is  $\mathbb{F}_Q$ , and  $t = 1/x$  is a local parameter there. A basic classification theorem says that the places arising from monic irreducible polynomials, together with this place at infinity, account for all places of  $\mathbb{F}_Q(x)$ . Thus, its rational places are the  $Q$  ordinary evaluation points  $x = \alpha \in \mathbb{F}_Q$  and the place at infinity, giving  $Q + 1$  in total.

**Splitting and ramification in our example.** What happens to an evaluation location when we pass from  $\mathbb{F}_{49}(x)$  to  $E = \mathbb{F}_{49}(x, u)$ ? Specifying  $x = a$  no longer determines all the values: we must also choose a value  $u = b$  satisfying  $b^4 = a^3 + a$ . Each pair  $(a, b) \in \mathbb{F}_{49}^2$  satisfying this equation gives a rational place  $P_{a,b}$  of  $E$ . We say that this place lies *above* the place  $x = a$  of the rational function field.

For a concrete example, take  $a = 1$ . The equation for  $u$  becomes  $u^4 = 2$ , which has four solutions in  $\mathbb{F}_{49}$ :  $2, -2, 2\iota, -2\iota$ , where  $\iota^2 = -1$ . Here we may write  $\mathbb{F}_{49} = \mathbb{F}_7(\iota)$ , since  $-1$  is not a square in  $\mathbb{F}_7$ . Thus, the single evaluation location  $x = 1$  gives four distinct rational places in  $E$ . We say that it *splits completely*. More generally, whenever  $a^3 + a$  is a nonzero fourth power in  $\mathbb{F}_{49}$ , it has four distinct fourth roots, giving four rational places above  $x = a$ . At each of them,  $x - a$  still has a simple zero and serves as a local parameter.

The behavior changes when  $a^3 + a = 0$ , which happens at  $a = 0, \iota, -\iota$ . There is now only one place above  $x = a$ , namely  $P_{a,0}$ . Moreover,  $x - a$  no longer has a simple zero there. To see why, factor

$$x^3 + x = (x - a)h_a(x), \quad h_a(a) \neq 0.$$

Geometric intuition suggests what is happening here. At  $a = 0$ , imagine the same equation  $u^4 = x^3 + x$  over the real numbers. Close to the origin, the term  $x^3$  is much smaller than  $x$ , so the curve looks like  $x = u^4$ : as  $u$  moves away from zero,  $x$  changes only to fourth order. Thus,  $u$  is a natural local coordinate, while  $x$  has a zero of order four; see Figure 3. Although this picture uses the real numbers, the underlying observation has a purely algebraic formulation that applies over our finite field. Since  $h_a(x)$  is nonzero at  $P_{a,0}$ , it has valuation zero, and the identity  $u^4 = (x - a)h_a(x)$  gives

$$4\nu_{P_{a,0}}(u) = \nu_{P_{a,0}}(x - a).$$

In fact,  $u$  is a local parameter at this place, so

$$\nu_{P_{a,0}}(u) = 1, \quad \nu_{P_{a,0}}(x - a) = 4.$$

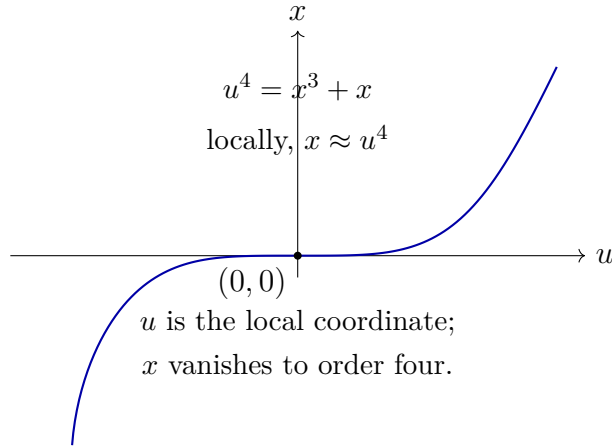


Figure 3: Real geometric intuition for ramification at  $P_{0,0}$ . With  $u$  horizontal and  $x$  vertical, the smooth curve has fourth-order contact with the line  $x = 0$ . The picture illustrates the local behavior; the argument over  $\mathbb{F}_{49}$  uses valuations.

This behavior is called *total ramification*: there is a single place above the base place, and the orders of zeros and poles of functions from the base field are multiplied by the extension degree, here four. Thus, an extension can produce more evaluation locations, as in complete splitting, or increase the orders of vanishing at a single location, as in total ramification.

We must also account for infinity. There is a unique place  $P_\infty$  of  $E$  above the place at infinity of  $\mathbb{F}_{49}(x)$ . It too is rational and totally ramified. Since  $x$  has a simple pole in the base field, it has a pole of order four in  $E$ . The defining equation then determines the pole order of  $u$ :

$$\nu_{P_\infty}(x) = -4, \quad \nu_{P_\infty}(u) = -3. \quad (\text{A.4})$$

Indeed,  $x^3 + x$  has a pole of order twelve, so  $u^4$  must have the same pole order. Neither  $x$  nor  $u$  has any other poles. **Figure 4** illustrates complete splitting and total ramification through the two quadratic extensions that define  $E$ .

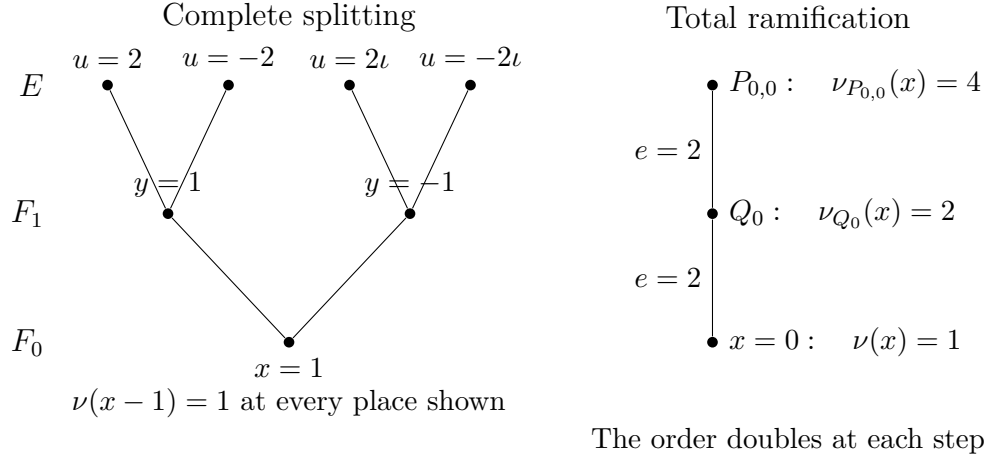


Figure 4: Splitting and ramification through the two quadratic extensions  $F_0 \subset F_1 \subset E$ . Each dot represents a rational place, and an edge connects a place to the one below it. On the left, the place  $x = 1$  splits completely at each step; the ramification index along every edge is  $e = 1$ . On the right,  $Q_0$  denotes the unique place of  $F_1$  above  $x = 0$ , and each step has ramification index  $e = 2$ . Thus the unique place  $P_{0,0}$  of  $E$  above  $x = 0$  has total ramification index four.

How many rational places does this give us altogether? We can answer by checking, for each  $a \in \mathbb{F}_{49}$ , how many values of  $b \in \mathbb{F}_{49}$  satisfy  $b^4 = a^3 + a$ . A direct calculation in  $\mathbb{F}_7(\iota)$  gives

| $a^3 + a$              | number of choices of $a$ | rational places above each $a$ |
|------------------------|--------------------------|--------------------------------|
| 0                      | 3                        | 1                              |
| a nonzero fourth power | 22                       | 4                              |
| not a fourth power     | 24                       | 0                              |

There are therefore  $3 + 22 \cdot 4 = 91$  finite rational places. Adding  $P_\infty$  gives 92 rational places in total, compared with the 50 we had in  $\mathbb{F}_{49}(x)$ .

What about the remaining 24 values of  $a$ ? There are still places of  $E$  above them, but none are rational: evaluating  $u$  requires a larger residue field in which  $U^4 = a^3 + a$  has a root. These places do not give evaluation coordinates over  $\mathbb{F}_{49}$ , but they remain part of the function field's structure. The degree of a place records precisely this distinction.

### A.1.3 Divisors: recording zeros and poles

So far, we have examined the behavior of a function at one place at a time. To understand its zeros and poles together, we need a way to record orders at several places simultaneously. A *divisor* does exactly this: it is a formal sum

$$A = \sum_{P \in \mathbb{P}_F} a_P P, \quad a_P \in \mathbb{Z},$$

with only finitely many nonzero coefficients. The word “formal” means that we are not adding points geometrically; we are simply assigning an integer to each place. These integers can record the zeros and poles of a function, but we may also choose them independently to specify conditions that we want a function to satisfy.

The *support* of  $A$  is the set of places with nonzero coefficient. Its *degree* counts the places with their assigned multiplicities, weighting each place by its own degree:

$$\deg A = \sum_{P \in \mathbb{P}_F} a_P \deg P.$$

Thus, a rational place contributes its coefficient, while a place of degree  $r$  contributes  $r$  times its coefficient. A divisor is called *effective*, written  $A \geq 0$ , if all its coefficients are nonnegative. We also write  $A \geq B$  when  $A - B$  is effective, so this comparison is made separately at each place.

Every nonzero function  $f$  gives a divisor by recording all its orders of vanishing and pole orders:

$$(f) = \sum_{P \in \mathbb{P}_F} \nu_P(f) P.$$

This is called the *principal divisor* of  $f$ . A basic theorem guarantees that a nonzero function has zeros and poles at only finitely many places, so this sum is indeed a divisor. Positive coefficients record zeros, negative coefficients record poles, and places where  $f$  has a nonzero finite value do not appear.

The key fact is that these zeros and poles balance:

$$\deg(f) = 0. \tag{A.5}$$

In other words, the total degree of the zeros equals the total degree of the poles, counting multiplicities on both sides. We have already seen this in the rational function field. A polynomial of degree  $m$  has zeros of total degree  $m$ , as follows by factoring it into irreducible polynomials and counting each factor with its multiplicity and degree. Its only pole is at infinity, also of order  $m$ . The identity above extends this balance to every nonzero function in an algebraic function field.

This gives a useful way to control zeros: a bound on the total degree of the poles is automatically a bound on the total degree of the zeros. In particular, it bounds the number of rational places at which the function vanishes. This will take the place of the polynomial root bound in our construction of codes. Another basic fact is that  $(f) = 0$  precisely when  $f$  is a nonzero constant in  $\mathbb{F}_Q$ : a nonconstant function cannot avoid both zeros and poles.

**Principal divisors in our example.** Let us collect the local calculations from the previous section. Write  $P_0 = P_{0,0}$ ,  $P_+ = P_{\iota,0}$ , and  $P_- = P_{-\iota,0}$ . The function  $x$  has a zero of order four at  $P_0$  and a pole of order four at  $P_\infty$ , with no other zeros or poles. The function  $u$  has a simple zero at each of  $P_0, P_+, P_-$  and a pole of order three at  $P_\infty$ . In divisor notation, these statements become

$$(x) = 4P_0 - 4P_\infty, \quad (u) = P_0 + P_+ + P_- - 3P_\infty. \quad (\text{A.6})$$

All these places are rational, so we can check the balance simply by adding coefficients:  $4 - 4 = 0$  for  $x$ , and  $1 + 1 + 1 - 3 = 0$  for  $u$ .

There is a distinction here that will matter for coding. The zero divisor of  $x$  has degree four, but  $x$  vanishes at only one place. By contrast,  $u$  has zeros of total degree three spread over three distinct places. A zero coordinate in a codeword is counted only once, regardless of the order of vanishing. Thus, the total degree of the zeros gives an upper bound on the number of zero coordinates, rather than necessarily counting them exactly.

Complete splitting gives an example where this bound is attained. If  $a^3 + a$  is a nonzero fourth power in  $\mathbb{F}_{49}$ , then

$$(x - a) = \sum_{\substack{b \in \mathbb{F}_{49} \\ b^4 = a^3 + a}} P_{a,b} - 4P_\infty.$$

Here the same pole of order four is balanced by four distinct simple zeros.

#### A.1.4 Riemann–Roch spaces: putting a bound on the poles

We now have the ingredients for an analogue of low-degree polynomials. The balance between zeros and poles tells us how to bound the number of zeros of a function: restrict its poles. A divisor lets us specify these restrictions place by place. For example,  $A = 3P + 2Q$ , with  $P, Q$  distinct rational places, allows a pole of order at most three at  $P$  and at most two at  $Q$ , with no poles anywhere else. Every nonzero function satisfying these restrictions then has zeros of total degree at most five.

For a general divisor  $A = \sum_P a_P P$ , the corresponding requirement is  $\nu_P(f) \geq -a_P$  at every place  $P$ . Equivalently,  $(f) + A \geq 0$ . This leads to the *Riemann–Roch space*

$$\mathcal{L}(A) = \{0\} \cup \{f \in F^\times : (f) + A \geq 0\}, \quad \ell(A) = \dim_{\mathbb{F}_Q} \mathcal{L}(A). \quad (\text{A.7})$$

When  $A$  is effective, its coefficients specify the allowed pole orders. A negative coefficient imposes a different requirement: if  $a_P = -r < 0$ , then  $\nu_P(f) \geq r$ , so  $f$  must vanish at  $P$  to order at least  $r$ . Thus, a single divisor can specify both allowed poles and required zeros.

These conditions are preserved under addition and multiplication by constants. Indeed, adding two functions cannot create a pole of larger order than either already has, nor can it destroy a zero order shared by both. Consequently,  $\mathcal{L}(A)$  is a vector space over  $\mathbb{F}_Q$ . A basic theorem says that it is finite-dimensional. Its dimension  $\ell(A)$  measures how many independent functions remain after we impose the restrictions encoded by  $A$ .

Let us check that this recovers the familiar message spaces for Reed–Solomon codes. In  $\mathbb{F}_Q(x)$ , take  $A = mP_\infty$  with  $m \geq 0$ . A function in  $\mathcal{L}(mP_\infty)$  has no finite poles, so it must be a polynomial. Its pole order at infinity is its degree, which must therefore be at most  $m$ . Hence

$$\mathcal{L}(mP_\infty) = \text{span}_{\mathbb{F}_Q} \{1, x, \dots, x^m\}, \quad \ell(mP_\infty) = m + 1.$$

Bounding the poles has given us exactly the usual degree restriction.

**An explicit basis in the example.** What does the same restriction give in our running example? Recall that  $x$  and  $u$  have no poles except at  $P_\infty$ , where their pole orders are four and three, respectively. Their products therefore provide a supply of functions with controlled poles:

$$\nu_{P_\infty}(x^i u^j) = -(4i + 3j).$$

Thus, whenever  $4i + 3j \leq m$ , the monomial  $x^i u^j$  belongs to  $\mathcal{L}(mP_\infty)$ .

Do these monomials account for every function in the space? In this example, they do. The affine curve  $u^4 = x^3 + x$  is smooth, and its only missing place is  $P_\infty$ . A standard fact about smooth affine curves therefore tells us that every function with no finite poles is a polynomial expression in  $x, u$ . Using  $u^4 = x^3 + x$ , we can write such an expression uniquely with  $u$ -degree at most three.

There is one more issue to check: could monomials with large poles cancel, leaving a function with a smaller pole? Here this cannot happen, because the remaining monomials have distinct pole orders. Indeed,

$$4i + 3j = 4i' + 3j', \quad 0 \leq j, j' \leq 3$$

forces  $j = j'$  by reduction modulo four, and then  $i = i'$ . In any nonzero linear combination, the monomial with the largest pole order therefore determines the pole order of the sum. We conclude that

$$\mathcal{L}(mP_\infty) = \text{span}_{\mathbb{F}_{49}}\{x^i u^j : i \geq 0, 0 \leq j \leq 3, 4i + 3j \leq m\}, \quad m \geq 0, \quad (\text{A.8})$$

and the displayed monomials form a basis.

For the first few values of  $m$ , this gives

| $m$     | a basis of $\mathcal{L}(mP_\infty)$ | $\ell(mP_\infty)$ |
|---------|-------------------------------------|-------------------|
| 0, 1, 2 | 1                                   | 1                 |
| 3       | 1, $u$                              | 2                 |
| 4, 5    | 1, $u, x$                           | 3                 |
| 6       | 1, $u, x, u^2$                      | 4                 |
| 7       | 1, $u, x, u^2, xu$                  | 5                 |
| 8       | 1, $u, x, u^2, xu, x^2$             | 6                 |

The table reveals a difference from the rational function field. There, increasing the allowed pole order from  $m - 1$  to  $m$  always adds a new function, namely  $x^m$ . Here, some increases give us nothing new. Allowing a pole of order one or two still leaves only the constants; allowing order five gives no more functions than allowing order four. In fact, there is no function whose only pole is at  $P_\infty$  and whose pole order is one, two, or five. These missing orders will give us a concrete way to understand genus.

**Adding requirements on the zeros.** Negative coefficients in the divisor let us impose vanishing conditions as well. For example, consider

$$\mathcal{L}(4P_\infty - P_0).$$

The coefficient at infinity allows a pole of order at most four, while the coefficient  $-1$  at  $P_0$  requires a zero there. We therefore start with  $\mathcal{L}(4P_\infty) = \text{span}\{1, u, x\}$  and retain only the functions that vanish at  $P_0$ . Since  $u(P_0) = x(P_0) = 0$ , this simply requires the constant coefficient to be zero:

$$\mathcal{L}(4P_\infty - P_0) = \text{span}\{u, x\}.$$

What if we require a zero of order at least two? The function  $u$  has only a simple zero at  $P_0$ , whereas  $x$  has a zero of order four. These different orders cannot cancel, so a combination  $cu + dx$  with  $c \neq 0$  still has a simple zero. Consequently,

$$\mathcal{L}(4P_\infty - 2P_0) = \text{span}\{x\}.$$

Thus, changing a coefficient of the divisor translates directly into a condition on the functions in the space.

### A.1.5 Genus and the Riemann–Roch theorem

How many independent functions can we obtain by allowing poles of a given total degree? In the rational function field, we know the answer:  $\ell(mP_\infty) = m + 1$ . Our running example shows that the answer can be smaller. Some pole orders are missing, and each missing order means that increasing the allowance does not add a new function. The *genus* measures the deficit that can arise in a general function field.

One algebraic characterization of the genus is

$$g = g(F) = \max_A (\deg A + 1 - \ell(A)),$$

where the maximum ranges over all divisors of  $F$ . Thus, we compare the dimension  $\ell(A)$  with the polynomial benchmark  $\deg A + 1$  and take the largest deficit. It is a theorem that this maximum is finite. The genus is a nonnegative integer, and the rational function field has genus zero.

The *Riemann–Roch theorem* tells us much more: once the degree of  $A$  is large enough, this deficit is always exactly  $g$ , regardless of how the coefficients of  $A$  are distributed among the places. Its full statement involves a special divisor called a *canonical divisor*. There is such a divisor  $K$ , of degree  $2g - 2$ , for which

$$\ell(A) = \deg A + 1 - g + \ell(K - A) \tag{A.9}$$

holds for every divisor  $A$ . Canonical divisors arise from the theory of differentials, which lies beyond this exposition. For our purposes, it is enough to view  $\ell(K - A)$  as a nonnegative correction to the main term  $\deg A + 1 - g$ .

To see when this correction disappears, first observe that a divisor of negative degree cannot have any nonzero functions in its Riemann–Roch space. Intuitively, we cannot require more zeros than we allow poles, with both counted by degree and multiplicity. Indeed, if  $f \neq 0$  belonged to  $\mathcal{L}(A)$ , then  $(f) + A$  would be effective, but its degree would be  $\deg A < 0$ , a contradiction. Consequently, when  $\deg A > 2g - 2$ , the divisor  $K - A$  has negative degree and the correction term vanishes.

We will use the following three consequences:

$$\ell(A) = 0 \quad \text{if } \deg A < 0, \tag{A.10}$$

$$\ell(A) \geq \deg A + 1 - g, \tag{A.11}$$

$$\ell(A) = \deg A + 1 - g \quad \text{if } \deg A > 2g - 2. \tag{A.12}$$

The first follows from the degree argument above. The second follows because the correction term in Riemann–Roch is nonnegative, and the third follows because it vanishes in the stated range. Thus, for sufficiently large degree, genus costs exactly  $g$  dimensions compared with the polynomial case.

**Seeing the genus in the example.** The missing pole orders we observed earlier give a concrete way to recover the genus. At any rational place  $P$ , the *gap theorem* states that precisely  $g$  positive integers fail to occur as the order of a function whose only pole is at  $P$  [Sti09, Chapter 1]. The missing integers may depend on the place, but their number is always the genus.

In our example, Equation (A.8) shows that the possible pole orders at  $P_\infty$ , including zero for the constants, are

$$\{3j + 4i : i, j \geq 0\} = \{0, 3, 4, 6, 7, 8, 9, \dots\}.$$

The missing positive orders are exactly 1, 2, 5, so  $g(E) = 3$ . Once  $m \geq 5$ , all three gaps have been passed, and counting the basis monomials gives

$$\ell(mP_\infty) = (m + 1) - 3 = m - 2.$$

This agrees with Riemann–Roch, whose exact range here is  $m > 2g - 2 = 4$ . At  $m = 4$ , only two of the gaps have been passed, and the dimension is three rather than  $4 + 1 - 3 = 2$ . The correction term in Equation (A.9) accounts for this extra dimension.

**Genus through ramification.** There is another way to find the genus of  $E$ , using the extension  $E/\mathbb{F}_{49}(x)$ . Recall the two behaviors we saw when passing to a larger function field: a place may split into several places, or it may ramify, multiplying the orders of zeros and poles of functions from the smaller field. The *Riemann–Hurwitz formula* describes how this ramification affects the genus.

The formula is particularly simple when the ramification is *tame*, meaning that every ramification index is relatively prime to the characteristic. For a finite separable extension  $L/F$  with the same full field of constants and with tame ramification, it reads

$$2g(L) - 2 = [L : F](2g(F) - 2) + \sum_{Q \in \mathbb{P}_L} (e_Q - 1) \deg Q,$$

where  $e_Q$  is the ramification index of  $Q$  over the place below it. Unramified places have  $e_Q = 1$  and contribute nothing to the sum. Thus, the formula starts with the genus of the smaller field and the extension degree, then adds a contribution from each ramified place.

In  $E/\mathbb{F}_{49}(x)$ , the only ramification occurs above  $x = 0, \iota, -\iota, \infty$ . Above each there is a single rational place of ramification index four. Since the characteristic is seven, the ramification is tame. The base field has genus zero and the extension has degree four, so

$$2g(E) - 2 = 4(-2) + 4(4 - 1) = 4.$$

We again obtain  $g(E) = 3$ .

The same calculation applies to the first quadratic extension  $F_1/\mathbb{F}_{49}(x)$ . Its four ramified places lie above the same values of  $x$ , now with ramification index two, giving

$$2g(F_1) - 2 = 2(-2) + 4(2 - 1) = 0.$$

Thus, the two extensions take us from genus zero to genus one and then to genus three. We have gained evaluation locations, but we have also increased the genus. The code parameters will make this tradeoff explicit.

## A.2 Goppa codes and optimal towers

### A.2.1 Goppa codes

We can now return to the coding idea that started this exposition: identify messages with functions and encode them by evaluation. Rational places give us evaluation locations over  $\mathbb{F}_Q$ , while Riemann–Roch spaces give us finite-dimensional spaces of functions with a bound on their zeros. The construction combines these two ingredients.

Choose distinct rational places  $P_1, \dots, P_N$  of  $F$ , and collect them in the divisor

$$D = P_1 + \dots + P_N.$$

Choose also a divisor  $G$  whose support avoids these places. This ensures that every function in  $\mathcal{L}(G)$  is regular at every  $P_i$ , so the evaluation map

$$\text{ev}_D : \mathcal{L}(G) \longrightarrow \mathbb{F}_Q^N, \quad f \longmapsto (f(P_1), \dots, f(P_N))$$

is well-defined. Evaluation respects addition and multiplication by constants, so its image is a linear code. This is the *algebraic-geometric evaluation code*, or *geometric Goppa code*:

$$C_L(D, G) = \text{ev}_D(\mathcal{L}(G)). \tag{A.13}$$

We will use “Goppa code” to refer to this construction.<sup>7</sup>

Concretely, choose a basis  $f_1, \dots, f_k$  of  $\mathcal{L}(G)$ . A message  $(c_1, \dots, c_k)$  specifies the function  $f = \sum_j c_j f_j$ , which we evaluate at the chosen places. As with Reed–Solomon codes, we need to check that distinct messages give distinct codewords and that these codewords disagree at many coordinates.

---

<sup>7</sup>A different construction, based on subfield subcodes, is also commonly called a Goppa code. We will not use that construction here.

**Dimension and distance.** Write  $m = \deg G$ . Let us first ask whether two different functions can have the same evaluation vector. Their difference would have to vanish at all  $N$  evaluation places. The functions in  $\mathcal{L}(G)$  satisfying these additional zero requirements form exactly the space  $\mathcal{L}(G - D)$ . Thus, this space is the kernel of evaluation, and the dimension of the code is

$$k = \ell(G) - \ell(G - D).$$

If  $m < N$ , then  $\deg(G - D) = m - N < 0$ , so  $\mathcal{L}(G - D) = \{0\}$ . Evaluation is therefore injective: a nonzero function cannot meet all  $N$  zero requirements with a divisor of degree only  $m$ . In this range, no message dimensions are lost, and  $k = \ell(G)$ .

The same reasoning gives the distance bound. Since the code is linear, its minimum distance is the smallest weight of a nonzero codeword. Let  $f \in \mathcal{L}(G)$  be nonzero. By definition,  $G + (f)$  is an effective divisor, and its degree is  $m$  because  $(f)$  has degree zero. Each zero coordinate  $f(P_i) = 0$  contributes at least one to this degree: the coefficient of  $G$  at  $P_i$  is zero, while  $\nu_{P_i}(f) \geq 1$  and  $\deg P_i = 1$ . There can therefore be at most  $m$  zero coordinates. Every nonzero codeword has at least  $N - m$  nonzero coordinates, giving

$$k = \ell(G) - \ell(G - D), \quad d \geq N - m, \quad (\text{A.14})$$

where the distance statement applies when the code is nonzero and  $m < N$ .

To turn the dimension formula into an explicit parameter bound, we use Riemann–Roch. In the range

$$2g - 2 < m < N, \quad (\text{A.15})$$

the two inequalities serve different purposes:  $m < N$  makes evaluation injective, while  $m > 2g - 2$  gives  $\ell(G) = m - g + 1$ . Consequently,

$$k = m - g + 1, \quad |C_L(D, G)| = Q^{m-g+1}, \quad d \geq N - m. \quad (\text{A.16})$$

The bound  $N - m$  is called the *designed distance*. It is the distance guaranteed by the argument; the actual minimum distance may be larger. For the rational function field, taking  $G = mP_\infty$  and evaluating at finite rational places gives precisely Reed–Solomon codes.

We can now see the tradeoff in choosing  $G$ . Increasing its degree allows more functions and hence more message dimensions, but also permits more zeros and weakens the distance guarantee. The genus measures the loss in dimension compared with the polynomial case. In terms of the rate  $R = k/N$  and relative distance  $\delta = d/N$ , our bounds give

$$R + \delta \geq 1 - \frac{g-1}{N}.$$

This explains why we seek function fields with many rational places relative to their genus. We want the number of evaluation locations to grow while keeping the genus small in comparison. Algebraic function fields make this possible even while the field  $\mathbb{F}_Q$  remains fixed.

**A code we can write down.** Let us carry out the construction in our running example  $E$ . Use all 91 finite rational places for evaluation, let  $D$  be their sum, and take  $G = 20P_\infty$ . The allowed functions have no finite poles and a pole of order at most 20 at infinity. By Equation (A.8), they have a basis consisting of 18 monomials.

A message therefore consists of 18 coefficients  $c_{ij} \in \mathbb{F}_{49}$ . To encode it, form

$$f(x, u) = \sum_{\substack{i \geq 0, \ 0 \leq j \leq 3 \\ 4i + 3j \leq 20}} c_{ij} x^i u^j$$

and evaluate at the 91 pairs  $(a, b) \in \mathbb{F}_{49}^2$  satisfying  $b^4 = a^3 + a$ . The resulting code has

$$N = 91, \quad k = 18, \quad d \geq 71.$$

The dimension agrees with Riemann–Roch, since  $20 - 3 + 1 = 18$ . The distance guarantee follows directly from the pole restriction: a nonzero function used here has zeros of total degree at most 20, so at most 20 of its 91 evaluation coordinates can be zero.

For an even smaller example, take  $\mathcal{L}(4P_\infty) = \text{span}\{1, u, x\}$ . A message  $(c_0, c_1, c_2)$  is encoded by evaluating  $c_0 + c_1 u + c_2 x$  at the same 91 places. This code has dimension three and designed distance  $91 - 4 = 87$ . Here we can check that the distance bound is exact:  $x - 1$  vanishes at the four rational places above  $x = 1$  and nowhere else among the evaluation places, so its codeword has weight 87.

Compare this with the codeword obtained from  $x$ . Its only zero coordinate is at  $P_0$ , so its weight is 90. Although  $x$  has a zero of order four there, this still gives only one zero coordinate. The distinction is exactly the one we saw with divisors: multiplicities enter the algebraic bound, while each evaluation location contributes only once to the Hamming weight.

### A.2.2 Optimal towers: obtaining arbitrarily long codes

Our running example gives a code of length 91 over  $\mathbb{F}_{49}$ . To obtain arbitrarily long codes over the same field, we need to continue the construction: find function fields with more and more rational places, while keeping their genera small enough that the dimension loss remains affordable.

A natural way to organize such a construction is a *tower* of function fields,

$$F_0 \subsetneq F_1 \subsetneq F_2 \subsetneq \cdots,$$

all with the same full field of constants  $\mathbb{F}_Q$ , and with genera tending to infinity. Our running example already gives the first three fields of such a tower. Continuing the recursion

$$F_0 = \mathbb{F}_{49}(x_0), \quad F_{i+1} = F_i(x_{i+1}), \quad x_{i+1}^2 = \frac{x_i^2 + 1}{2x_i}$$

produces an infinite tower over  $\mathbb{F}_{49}$  [GSR03]. Its first two extensions are exactly the fields  $F_1$  and  $E$  studied above. The constructions of Garcia and Stichtenoth [GS96] provide another important source of towers with many rational places.

How many rational places should we hope for? Write  $B_1(F)$  for the number of rational places of  $F$ . The code parameter bound

$$R + \delta \geq 1 - \frac{g - 1}{N}$$

suggests that the relevant quantity is the number of rational places relative to the genus. We want  $B_1(F)/g(F)$  to remain large as the genus grows. The Drinfeld–Vlăduț bound [VD83] places a fundamental limit on this ratio: for any sequence  $F_i/\mathbb{F}_Q$  with  $g_i = g(F_i) \rightarrow \infty$ ,

$$\limsup_{i \rightarrow \infty} \frac{B_1(F_i)}{g_i} \leq \sqrt{Q} - 1. \quad (\text{A.17})$$

A family attaining this bound, namely satisfying

$$\frac{B_1(F_i)}{g_i} \rightarrow \sqrt{Q} - 1,$$

is called *optimal*. For square  $Q$ , such families were obtained from modular curves by Ihara and, independently, Tsfasman, Vlăduț, and Zink [Iha81, TVZ82]. The modular tower above over  $\mathbb{F}_{49}$  and suitable Garcia–Stichtenoth towers are explicit examples of optimal families.

Combining an optimal family with Goppa’s construction gives the Tsfasman–Vlăduț–Zink bound [TVZ82],

$$R_Q(\delta) \geq 1 - \delta - \frac{1}{\sqrt{Q} - 1},$$

where  $R_Q(\delta)$  denotes the largest asymptotically achievable rate at relative distance at least  $\delta$ . To see the connection, use all but one of the rational places for evaluation and allow poles at the remaining place. Choose the divisor degree to be approximately  $(1 - \delta)N$ . The designed relative distance is then at least  $\delta$ , while the rate approaches at least  $1 - \delta - 1/(\sqrt{Q} - 1)$ . The bound is useful where its right-hand side is positive.

For  $Q = 49$ , optimality means that the number of rational places is asymptotically six times the genus. Our example has the much larger ratio  $92/3$ . There is no contradiction: the Drinfeld–Vlăduț bound concerns what can persist as the genus tends to infinity. A particularly favorable small example need not retain the same ratio throughout a tower.

This completes the basic picture behind asymptotically good algebraic-geometric codes. The tower supplies increasingly many evaluation locations over a fixed field, and Riemann–Roch controls the dimension available at each level. For our improvement over GV, we will need more information about the resulting codes than their dimension and designed distance. We now introduce the tools that let us study their weight distributions by averaging.

## A.3 Further tools from function-field theory

### A.3.1 Divisor classes and the class number

Fix a function field and a set of evaluation places. We can obtain different codes by changing the divisor  $G$ , even while keeping its degree fixed. Before averaging over these choices, it helps to ask when two divisors impose essentially the same conditions on functions.

Suppose  $B = A + (h)$  for some nonzero function  $h$ . Then

$$(f) + B = (fh) + A,$$

so  $f \in \mathcal{L}(B)$  precisely when  $fh \in \mathcal{L}(A)$ . Multiplication by  $h$  therefore gives an isomorphism between the two Riemann–Roch spaces:

$$\mathcal{L}(B) = h^{-1}\mathcal{L}(A).$$

We call  $A$  and  $B$  *linearly equivalent*, and write  $A \sim B$ , when their difference is a principal divisor. Equivalent divisors have the same degree, because principal divisors have degree zero, and the same Riemann–Roch dimension. Their zero and pole requirements may look different, but multiplication by a fixed function transforms one set of requirements into the other.

**An example of equivalence.** In our running example, the identity  $(x) = 4P_0 - 4P_\infty$  gives  $4P_0 \sim 4P_\infty$ . However,  $P_0$  and  $P_\infty$  themselves are not equivalent. If they were, there would be a function with divisor  $P_0 - P_\infty$ : a simple zero at  $P_0$ , a simple pole at  $P_\infty$ , and no other zeros or poles. We already know that no such function exists, since  $\mathcal{L}(P_\infty)$  contains only constants. Thus, having the same degree does not guarantee linear equivalence, and an equivalence between multiples cannot simply be divided by the multiplier.

**Counting the classes.** Divisors form an abelian group  $\mathcal{D}_F$  under addition, and principal divisors form a subgroup  $\mathcal{P}_F$ . Identifying linearly equivalent divisors gives the *divisor class group*

$$\mathcal{C}_F = \mathcal{D}_F / \mathcal{P}_F.$$

Since equivalent divisors have the same degree, we can speak of the degree of a class. Write  $\mathcal{C}_F^m$  for the set of classes of degree  $m$ .

Although there are infinitely many divisors of degree zero, a fundamental theorem says that they fall into only finitely many classes. Their number,

$$h_F = |\mathcal{C}_F^0|,$$

is called the *class number*. In our example,  $[P_0 - P_\infty]$  is a nonzero degree-zero class, while  $4[P_0 - P_\infty] = 0$ .

The class number also counts the classes of every other degree, provided the field has a rational place. Indeed, fix such a place  $P_*$  and consider

$$\mathcal{C}_F^0 \longrightarrow \mathcal{C}_F^m, \quad [A] \longmapsto [A + mP_*]. \quad (\text{A.18})$$

Subtracting  $mP_*$  gives the inverse map, so this is a bijection. Thus, there are exactly  $h_F$  classes of each degree. The full class group is infinite because the degree can vary, but fixing the degree leaves a finite set over which we can average.

**Choosing representatives for evaluation.** A divisor class contains many representatives, and some may involve the places where we want to evaluate. Fortunately, we can always choose a representative whose support avoids any prescribed finite set of places. This follows from the approximation theorem for function fields [Sti09, Chapter 1]. For a divisor  $A = \sum_P a_P P$  and evaluation places  $P_1, \dots, P_N$ , the theorem supplies a nonzero function  $h$  with

$$\nu_{P_i}(h) = -a_{P_i} \quad \text{for every } i.$$

The equivalent divisor  $A + (h)$  then has coefficient zero at every evaluation place. Its coefficients elsewhere may be negative, which is why it is useful to allow arbitrary divisors in the definition of a Riemann–Roch space.

Does the choice of representative affect the code? Suppose  $G' = G + (h)$  and both divisors avoid the evaluation places. Then  $\nu_{P_i}(h) = 0$ , so  $h(P_i)$  is a nonzero field element. The relation between the Riemann–Roch spaces becomes

$$C_L(D, G') = \left\{ \left( h(P_1)^{-1}c_1, \dots, h(P_N)^{-1}c_N \right) : c \in C_L(D, G) \right\}.$$

Thus, changing the representative only rescales each coordinate by a fixed nonzero value. It preserves which coordinates are zero, and hence preserves the weight distribution.

We may therefore choose one representative  $G_\gamma$  for each degree- $m$  class  $\gamma$ , avoiding the evaluation places, and study the resulting codes. Their weight distributions do not depend on these choices of representatives. When  $2g - 2 < m < N$ , they also all have the same dimension  $m - g + 1$ . The variation between classes is what our averaging argument exploits.

**Counting effective divisors within a class.** There is a useful connection between divisor classes and the functions in a Riemann–Roch space. An effective divisor equivalent to  $A$  has the form  $A + (f)$  for some nonzero  $f \in \mathcal{L}(A)$ . When do two functions give the same divisor? Exactly when

$$A + (f) = A + (h),$$

or equivalently  $(f/h) = 0$ . This means that  $f/h$  is a nonzero constant.

Consequently, the  $Q^{\ell(A)} - 1$  nonzero functions in  $\mathcal{L}(A)$  occur in groups of  $Q - 1$  scalar multiples, with each group giving one effective divisor. We obtain

$$|\{B \geq 0 : B \sim A\}| = \frac{Q^{\ell(A)} - 1}{Q - 1}. \quad (\text{A.19})$$

This is the bridge between counting functions and counting effective divisors. In particular, it explains the factors of  $Q - 1$  that appear when we translate divisor counts into codeword counts.

### A.3.2 The zeta function and the L-polynomial

We have just counted effective divisors within one class. What if we want to count all effective divisors of a given degree? Rather than study each degree separately, we collect the counts in a generating function.

Let  $a_r(F)$  be the number of effective divisors of degree  $r$ . There is exactly one effective divisor of degree zero, the zero divisor, so  $a_0(F) = 1$ . The *zeta function* of  $F$  is

$$Z_F(T) = \sum_{r \geq 0} a_r(F) T^r = \prod_{P \in \mathbb{P}_F} (1 - T^{\deg P})^{-1}. \quad (\text{A.20})$$

The product follows directly from the way an effective divisor is assembled. A place of degree  $r$  may appear with multiplicity  $0, 1, 2, \dots$ , contributing the factor

$$1 + T^r + T^{2r} + \dots = (1 - T^r)^{-1}.$$

Multiplying over the places chooses all their multiplicities, and the exponent of  $T$  records the resulting total degree. The identity holds as a formal power series; it also converges absolutely for  $|T| < 1/Q$ .

This interpretation lets us impose restrictions on the divisors we count. For example, to count effective divisors whose support avoids a specified rational place, we remove its factor  $(1 - T)^{-1}$  from the product. Avoiding  $w$  specified rational places therefore gives the generating function

$$Z_F(T)(1 - T)^w.$$

Although the factor  $(1 - T)^w$  has alternating coefficients, the product has nonnegative coefficients: each coefficient counts effective divisors satisfying the stated restriction. This observation will be useful in our weight-distribution argument.

**An infinite series controlled by a finite polynomial.** A fundamental theorem says that the zeta function is rational:

$$Z_F(T) = \frac{L_F(T)}{(1 - T)(1 - QT)}, \quad (\text{A.21})$$

where  $L_F(T) \in \mathbb{Z}[T]$  has degree  $2g$  and constant term one. This is the *L-polynomial* of  $F$ , also called its *L-function* in the function-field literature. Thus, the counts of effective divisors of every degree are determined by a polynomial of degree only  $2g$ .

For the rational function field, the genus is zero and the numerator is simply one. We can therefore view the denominator as the familiar genus-zero case, and the numerator as recording how the divisor counts change for a more general function field. The numerator also satisfies the symmetry

$$L_F(T) = Q^g T^{2g} L_F\left(\frac{1}{QT}\right), \quad (\text{A.22})$$

called its *functional equation*. Both rationality and this functional equation follow from Riemann–Roch and the counting of effective divisors within each class.

**Weil’s theorem: the Riemann hypothesis for curves.** The next theorem is much deeper. Factor the *L-polynomial* over the complex numbers as

$$L_F(T) = \prod_{j=1}^{2g} (1 - \alpha_j T). \quad (\text{A.23})$$

Weil’s theorem states that

$$|\alpha_j| = \sqrt{Q} \quad \text{for every } j.$$

The numbers  $\alpha_j$  are the *reciprocal roots*; the roots of  $L_F$  itself are  $\alpha_j^{-1}$ . Thus, every root lies on the circle of radius  $Q^{-1/2}$ .

This is the celebrated *Riemann hypothesis for curves over finite fields*. The connection with the usual Riemann hypothesis becomes visible under the substitution  $T = Q^{-s}$ : the circle  $|T| = Q^{-1/2}$  corresponds to the line  $\operatorname{Re}(s) = 1/2$ . We use this theorem without proof; a standard function-field treatment appears in [Sti09, Chapter 5].

**Bounding the number of rational places.** Weil's theorem connects the zeta function back to our supply of evaluation locations. An effective divisor of degree one is just a rational place, so the coefficient of  $T$  in  $Z_F(T)$  is  $B_1(F)$ . Comparing coefficients in Equation (A.21) gives

$$B_1(F) = Q + 1 - \sum_{j=1}^{2g} \alpha_j. \quad (\text{A.24})$$

There are  $2g$  summands, each of absolute value  $\sqrt{Q}$ . The triangle inequality therefore gives the *Hasse–Weil bound*:

$$\left| B_1(F) - (Q + 1) \right| \leq 2g\sqrt{Q}. \quad (\text{A.25})$$

The rational function field has exactly  $Q + 1$  rational places. This bound tells us how far that number can change at genus  $g$ . Unlike the Drinfeld–Vlăduț bound, it applies to each individual function field, without taking a limit.

**The L-polynomial of our example.** Our field  $E$  has genus three and 92 rational places. Substituting into Equation (A.24) gives

$$\sum_{j=1}^6 \alpha_j = 49 + 1 - 92 = -42.$$

Each summand has absolute value seven. Their sum can be  $-42$  only if every summand equals  $-7$ . Thus,

$$L_E(T) = (1 + 7T)^6, \quad Z_E(T) = \frac{(1 + 7T)^6}{(1 - T)(1 - 49T)}. \quad (\text{A.26})$$

Here the number of rational places determines the entire  $L$ -polynomial because our example attains the Hasse–Weil upper bound. A function field attaining this bound is called *maximal*. This is a special situation: in general, knowing the number of rational places does not determine the remaining coefficients.

Expanding the generating function gives

$$Z_E(T) = 1 + 92T + 5286T^2 + \cdots.$$

The coefficient 5286 counts effective divisors of degree two. Some are sums of two rational places, possibly the same place twice; the others are individual places of degree two. This illustrates why higher-degree places enter divisor counting even when our code coordinates use only rational places.

**Recovering the class number.** The same polynomial also counts divisor classes:

$$h_F = L_F(1) = Q^g L_F(1/Q). \quad (\text{A.27})$$

To understand why a connection should exist, suppose that  $F$  has a rational place and take  $r > 2g-2$ . Every degree- $r$  class has Riemann–Roch dimension  $r-g+1$ , so [Equation \(A.19\)](#) gives the same number of effective divisors in each class. There are  $h_F$  classes, and hence

$$a_r(F) = h_F \cdot \frac{Q^{r-g+1} - 1}{Q - 1}.$$

Thus, the class number determines all sufficiently late coefficients of the zeta function. The identity  $h_F = L_F(1)$  expresses this relationship through the numerator of the generating function, while the second equality in [Equation \(A.27\)](#) follows from the functional equation.

For our example,

$$h_E = L_E(1) = 8^6 = 262144.$$

There are therefore 262144 divisor classes of each fixed degree. Earlier we exhibited one nonzero degree-zero class,  $[P_0 - P_\infty]$ . The  $L$ -polynomial counts all the classes without requiring us to find their representatives.

**Returning to optimal towers.** We can now state the additional asymptotic information needed in our proof. Let  $F_i/\mathbb{F}_{49}$  be an optimal family, and put  $N_i = B_1(F_i)$ . Besides  $g_i/N_i \rightarrow 1/6$ , optimality implies

$$\frac{\ln Z_{F_i}(z)}{N_i} \longrightarrow -\ln(1-z) \quad \text{for every fixed } 0 < z < 1/49, \quad (\text{A.28})$$

$$\frac{\ln h_{F_i}}{N_i} \longrightarrow \frac{\ln 49}{6} + \ln \frac{49}{48}. \quad (\text{A.29})$$

These limits describe two quantities that enter our averaging argument: the generating function that counts effective divisors, and the number of divisor classes over which we average.

We state these consequences without proof here. Their derivation uses additional asymptotic results on places of higher degree, together with Weil’s theorem. The important point for our purposes is that they depend only on optimality. The special formula

$L_E(T) = (1 + 7T)^6$  need not persist as we move up the tower; the limits above provide the information we use in its place.