THE CHAIN AND THE SQUARE ROOT

Prerequisites. Finite fields and Euler's criterion; polynomials over a field, their roots and multiplicities; a homogeneous linear system with more unknowns than equations has a nonzero solution. Euclidean lattices and their covolumes. The slope inequality of Bost, which is stated in Section 8.4 and proved in Chapter 6. The von Mangoldt function Λ and Mertens' estimate $\sum_{n\le x}\Lambda(n)/n = \log x + O(1)$. Sections 8.1–8.3 use no earlier chapter. Section 8.5 uses the function G of (1) and Theorem 1.2.3 of Chapter 1.

8.1. The auxiliary polynomial

How can a polynomial show that a set of about $q/2$ elements of $\mathbb{F}_q$ has at most $q/2+O(\sqrt{q})$ elements? Stepanov's answer [Ste69] is to make it vanish to high order on the set. We follow his method for the curve $y^2 = f(x)$.

Let q be a power of an odd prime p, and let $f \in \mathbb{F}_q[X]$ have odd degree $d = 2g+1$ with $3 \le d < q$. Let

$$c := \frac{q-1}{2}, \qquad D := \frac{q-d}{2}.$$

For $\eta \in \{1,-1\}$ let $S_\eta := \{x \in \mathbb{F}_q : f(x)^c = \eta\}$, let $S_0 := \{x \in \mathbb{F}_q : f(x) = 0\}$, and let N^{aff} be the number of solutions $(x,y) \in \mathbb{F}_q^2$ of $y^2 = f(x)$.

Problem 8.1.1. $\mathbb{F}_q$ is the disjoint union of S_1, S_{-1} and S_0, and $N^{\mathrm{aff}} = 2|S_1| + |S_0|$. For $q = 7$ and $f = X^3 - X$, $S_1 = \{4,5\}$, $S_{-1} = \{2,3\}$ and $N^{\mathrm{aff}} = 7$.

Hence $N^{\mathrm{aff}} - q = |S_1| - |S_{-1}|$, and a bound $|S_\eta| \le q/2 + B$ for both signs gives $|N^{\mathrm{aff}} - q| \le 2B + d$. For squarefree f, Weil's theorem [Wei48] gives $|N^{\mathrm{aff}} - q| \le 2g\sqrt{q}$.

The principle is the count of zeros against the degree: if $F \in \mathbb{F}_q[X]$ is nonzero and vanishes to order at least ℓ at every point of S_η, then

$$\ell\,|S_\eta| \le \deg F. \tag{1}$$

The polynomials $f^c - \eta$ and $X^q - X$ vanish on S_η, and give $|S_\eta| \le dc$ and $|S_\eta| \le q$. To reach $q/2$ one needs $\deg F/\ell$ close to $q/2$. Stepanov combines the two polynomials. Let $J \ge 1$, and let $r_j, s_j \in \mathbb{F}_q[X]$, for $0 \le j < J$, have degree less than D and be not all zero. The *auxiliary polynomial* is

$$F := \sum_{j<J} \bigl(r_j + s_j f^c\bigr) X^{jq}. \tag{2}$$

Problem 8.1.2. $\deg f^c = gq + D$, and $2D < q$.

Problem 8.1.3. The nonzero terms $r_j X^{jq}$ and $s_j f^c X^{jq}$ of (2) have pairwise distinct degrees. Hence $F \ne 0$ and $\deg F \le (J+g)q - d - 1$.

For $x \in S_\eta$ we have $x^{jq} = x^j$ and $f(x)^c = \eta$, so $F(x) = C_0(x)$, where

$$C_0 := \sum_{j<J} (r_j + \eta s_j) X^j.$$

Date: November 25, 2026.

If $C_0 = 0$, then F vanishes on all of S_η.

Problem 8.1.4. If $C_0 = 0$, then $\deg F \geq q$. If $D \geq 2$ and $J \geq 2$, the tuple $r_0 = -X$, $r_1 = 1$, all others zero, gives $F = X^q - X$.

So the gain must come from the order of vanishing.

8.2. Vanishing to order ℓ

In characteristic p the derivative of order p vanishes identically, so one uses Hasse derivatives. For $G \in \mathbb{F}_q[X]$ define $H^k G \in \mathbb{F}_q[X]$ by

$$G(X+Y) = \sum_{k \geq 0} (H^k G)(X)\, Y^k \qquad \text{in } \mathbb{F}_q[X,Y].$$

Problem 8.2.1. $H^k(X^n) = \binom{n}{k} X^{n-k}$, and $H^k(G_1 G_2) = \sum_{i \leq k} H^i G_1 \cdot H^{k-i} G_2$. For $t \in \mathbb{F}_q$, $(X-t)^\ell$ divides G if and only if $(H^k G)(t) = 0$ for all $k < \ell$.

Problem 8.2.2. $(X+Y)^{jq} = (X^q + Y^q)^j$. Hence $H^k(GX^{jq}) = (H^k G) X^{jq}$ for $k < q$.

To differentiate f^c, write $f(X+Y)^c = f(X)^c \big(f(X+Y)/f(X)\big)^c$ and define $U_i \in \mathbb{F}_q(X)$ by

$$\Big(\frac{f(X+Y)}{f(X)}\Big)^c = \sum_{i \geq 0} \frac{U_i}{f^i}\, Y^i \qquad \text{in } \mathbb{F}_q(X)[[Y]].$$

Problem 8.2.3. Each U_i is a polynomial of degree at most $i(d-1)$. Moreover $U_0 = 1$ and $U_1 = cf'$.

For $k \geq 0$ let

$$C_k := \sum_{j<J} \Big(f^k H^k r_j + \eta \sum_{i \leq k} f^{k-i} U_i\, H^{k-i} s_j \Big) X^j. \tag{3}$$

For $k = 0$ this is the polynomial C_0 of Section 8.1.

Problem 8.2.4. Let $t \in S_\eta$ and $k < q$. Then $(H^k F)(t) = f(t)^{-k} C_k(t)$.

Problem 8.2.5. $\deg C_k \leq D + J - 2 + k(d-1)$.

By Problems 8.2.1 and 8.2.4, if $\ell \leq q$ and

$$C_0 = C_1 = \cdots = C_{\ell-1} = 0, \tag{4}$$

then F vanishes to order at least ℓ at every point of S_η. By Problem 8.2.5 the identities (4) cost at most

$$\sum_{k<\ell} \big(D + J - 1 + k(d-1)\big) = \ell(D+J-1) + (d-1)\frac{\ell(\ell-1)}{2}$$

linear conditions, whatever S_η is. The kth identity costs $k(d-1)$ more than the first, and this growth will limit ℓ.

8.3. The count and the balance

The unknowns are the $2JD$ coefficients of the r_j and s_j.

Problem 8.3.1. If $2JD > \ell(D+J-1)+(d-1)\ell(\ell-1)/2$, then there is an auxiliary polynomial (2) satisfying (4).

Problem 8.3.2. Let ℓ be even with $\ell < q-d$, and let $J = \ell/2 + t$. Then

$$2JD - \ell(D+J-1) - (d-1)\frac{\ell(\ell-1)}{2} = t\,(q-d-\ell) - \frac{d\ell^2 - (d+1)\ell}{2}.$$

Hence the hypothesis of Problem 8.3.1 holds if $t > d\ell^2/\big(2(q-d-\ell)\big)$.

Inserting $\deg F \le (J+g)q - d - 1$ from Problem 8.1.3 into (1), Problems 8.2.1–8.3.2 prove the following.

Theorem 8.3.1 (Stepanov)**.** *Let ℓ be even with $2 \le \ell < q-d$, and let $t := \lfloor d\ell^2/(2(q-d-\ell))\rfloor + 1$. Then, for $\eta = \pm 1$,*

$$|S_\eta| \le \frac{q}{2} + \frac{(g+t)\,q}{\ell}.$$

This is Stepanov's method [Ste69]; see [Sch76] for Hasse derivatives. For example, for $q = 1009$, $d = 3$ and $\ell = 24$ one has $t = 1$, and the bound is $588.58\ldots$; for $f = X^3 + X + 1$ the sets have $|S_1| = 516$ and $|S_{-1}| = 492$.

Problem 8.3.3. If $q \ge 16d^2$ and ℓ is an even integer with $|\ell - \sqrt{q}| \le 1$, then $|S_\eta| \le q/2 + 2d\sqrt{q}$ for $\eta = \pm 1$, and $|N^{\mathrm{aff}} - q| \le 4d\sqrt{q} + d$.

Where does the square root come from? Write $(J+g)q - d - 1 = \frac{\ell}{2}q + \big((g+1)q - d - 1\big) + (t-1)q$. After division by ℓ the first term is the main term $q/2$. The second is paid once, however high the order. The third comes from the growing cost of the kth identity: $t-1$ is about $d\ell^2/(2q)$. So the error is about $(g+1)q/\ell + d\ell/2$, of the shape $q/m + m$, and the two terms balance at $\ell \approx \sqrt{q}$.

Problem 8.3.4. For every ℓ as in Theorem 8.3.1, $(g+t)q/\ell > \sqrt{2gdq}$.

So the method never gives $o(\sqrt{q})$. We expect the same for Bombieri's form [Bom74], but do not prove it here.

8.4. The chain inequality

Over $\mathbb{Z}$ there is no derivative in the direction of a prime. In Alpöge's argument [Alp] (see also [Zh], [Dup, §§1.4–1.6]), the order of vanishing is replaced by a chain of m linear maps between lattices, each divisible by many primes. What does a chain give that one map does not?

A *Euclidean lattice* L is a free $\mathbb{Z}$-module of finite rank with a Euclidean norm on $L_{\mathbb{R}} := L \otimes \mathbb{R}$. Let

$$\mu(L) := -\frac{\log \operatorname{covol}(L)}{\operatorname{rk} L},$$

and let $\mu_{\max}(L)$ be the largest slope $\mu(L')$ of a nonzero subgroup $L' \subseteq L$, with the restricted norm. L is *semistable* if $\mu_{\max}(L) = \mu(L)$. For Euclidean lattices L, M and a nonzero linear map $\varphi\colon L_{\mathbb{Q}} \to M_{\mathbb{Q}}$ let

$$h_{\mathrm{fin}}(\varphi) := -\log \lambda, \qquad h_\infty(\varphi) := \log\|\varphi_{\mathbb{R}}\|,$$

where λ is the largest positive rational number with $\varphi(L) \subseteq \lambda M$, and $\|\varphi_{\mathbb{R}}\|$ is the operator norm. Thus $-h_{\mathrm{fin}}(\varphi)$ measures divisibility and $h_\infty(\varphi)$ measures size. The following is due to Bost [Bo1], and is proved in Chapter 6.

Theorem 8.4.1 (Slope inequality)**.** *If φ is injective, then $-h_{\mathrm{fin}}(\varphi) \le \mu_{\max}(M) - \mu(L) + h_\infty(\varphi)$.*

For $L = M = \mathbb{Z}$ and φ the multiplication by a nonzero integer n, both sides are $\log|n|$. Now let $L_0, \dots, L_m$ be Euclidean lattices, let $\varphi_j \colon L_{j-1,\mathbb{Q}} \to L_{j,\mathbb{Q}}$ be nonzero linear maps, and let $\Phi := \varphi_m \circ \cdots \circ \varphi_1$.

Problem 8.4.1. If $\Phi \neq 0$, then $h_{\mathrm{fin}}(\Phi) \le \sum_j h_{\mathrm{fin}}(\varphi_j)$ and $h_\infty(\Phi) \le \sum_j h_\infty(\varphi_j)$. In particular, if $\varphi_j(L_{j-1}) \subseteq p^{e_j} L_j$ for each j, then $\Phi(L_0) \subseteq p^{e_1+\cdots+e_m} L_m$.

Problem 8.4.1 and Theorem 8.4.1 for Φ give the following.

Theorem 8.4.2 (Chain inequality)**.** *If Φ is injective, then*

$$\sum_{j=1}^{m} \bigl(-h_{\mathrm{fin}}(\varphi_j)\bigr) \le \mu_{\max}(L_m) - \mu(L_0) + \sum_{j=1}^{m} h_\infty(\varphi_j).$$

The gains add, the sizes of the steps add, the slopes of the two ends are paid once, and the intermediate lattices $L_1, \dots, L_{m-1}$ do not occur.

Problem 8.4.2. Suppose that every φ_j is injective. The sum of the inequalities of Theorem 8.4.1 for $\varphi_1, \dots, \varphi_m$ is the inequality of Theorem 8.4.2 with

$$\sum_{0<j<m} \bigl(\mu_{\max}(L_j) - \mu(L_j)\bigr) \ge 0$$

added to its right side. This sum vanishes if and only if $L_1, \dots, L_{m-1}$ are semistable.

Problem 8.4.3. Let $N \ge 2$, $L_0 = L_2 = \mathbb{Z}^2$ and $L_1 = \mathbb{Z}e_1 + \mathbb{Z}Ne_2$, all with the standard norm of $\mathbb{R}^2$, and let $\varphi_1(x,y) = (x, Ny)$ and $\varphi_2(x,y) = (x, y/N)$. Then $\mu_{\max}(L_1) - \mu(L_1) = \frac{1}{2}\log N$. The right side of Theorem 8.4.2 is $\log N$, the right side of the summed inequality of Problem 8.4.2 is $\frac{3}{2}\log N$, and $h_\infty(\Phi) = 0$.

In the argument the divisibility $\varphi_j(L_{j-1}) \subseteq p^{e_j} L_j$ is known only on a full sublattice of L_{j-1}, spanned by the images of Hecke operators, and not on all of L_{j-1}. The gain is then recorded on that sublattice as in Theorem 6.5.1, and the logarithm of its index, divided by the rank, is subtracted from it: the gains that add are the gains on the sublattices, and the colength of each sublattice is a loss [Adr, §§10, 12]. Chapter 19 computes this colength at one vertex of a tree, and Chapter 25 follows it along the chain.

Theorem 8.4.2 needs only the slopes of the two ends and the size $h_\infty(\varphi_j)$ of each step. How precisely must these be known?

8.5. The budget $x/m + m$

How long should the chain be? Let $G(x) = \sum_{n \le x} \Lambda(n)(x/n - 1) - x\log x + (1+\gamma)x$ be the function of (1) in Chapter 1. By Mertens' estimate, $G(x) = O(x)$. By Theorem 1.2.3, if $G(x) \le x^\theta$ for some $0 < \theta < 1$ and all large x, then $\zeta(s) \neq 0$ for $\operatorname{Re} s > \theta$. So a one-sided bound with exponent $\frac{1}{2}$ gives the Riemann hypothesis.

Let $a, b > 0$ and $x := a/b$. As in Section 1.3, $(a/p^e - b)_+ = b(x/p^e - 1)_+$, so that

$$\sum_p \log p \sum_{e \ge 1} \Bigl(\frac{a}{p^e} - b\Bigr)_+ = b \sum_{n \le x} \Lambda(n) \Bigl(\frac{x}{n} - 1\Bigr),$$

where $u_+ := \max(u, 0)$; the inner sum is a smoothed Legendre formula for $v_p(a!)$. In the argument such divisibilities come from jets of theta functions [Dup, §1.4]. We consider a chain as in Section 8.4, with Φ injective, and real numbers E and $c_1, \dots, c_m$ such that for $1 \le j \le m$

$$-h_{\mathrm{fin}}(\varphi_j) \ge b \sum_{n \le x} \Lambda(n)\Big(\frac{x}{n} - 1\Big), \qquad h_\infty(\varphi_j) \le a(\log x - 1 - \gamma) + c_j, \tag{5}$$
$$\mu_{\max}(L_m) - \mu(L_0) \le E.$$

Here E is the cost at the ends, and c_j the cost of the jth step. This models the bookkeeping only; the chains of Alpöge's argument, and the status of their inputs in [Adr], are described in Chapter 58.

Problem 8.5.1. Under (5), $G(x) \le (E + c_1 + \cdots + c_m)/(mb)$.

Problem 8.5.2. If $E = a$ and $c_j = (2j-1)b$, then $G(x) \le x/m + m$. With $m = \lceil \sqrt{x} \rceil$ this gives $G(x) \le 2\sqrt{x} + 1$. For $x = 10^6$ the bound is 2000 at $m = 1000$ and 10100 at $m = 100$.

The cost at the ends, $a = xb$, is divided by m because it is paid once. The costs of the steps grow linearly with j, as the cost of the kth identity does in Section 8.2. The square root is the balance of the two.

Problem 8.5.3. Suppose instead that every φ_j is injective and that Theorem 8.4.1 is applied to each φ_j separately, with $\mu_{\max}(L_j) - \mu(L_{j-1}) \le a$ for every j and $c_j = (2j-1)b$. Then the jth step gives $G(x) \le x + 2j - 1$, so the best of these bounds is $G(x) \le x + 1$, the bound for $m = 1$.

Problem 8.5.4. $\sum_{j=1}^{m} (3j^2 - 3j + 1) = m^3$. So if $E = a$ and $c_j = (3j^2 - 3j + 1)b$, then $G(x) \le x/m + m^2$. For every $m > 0$, $x/m + m^2 \ge 3 \cdot 2^{-2/3} x^{2/3}$. For $x = 10^6$ the least value over integers m is $18899.2\ldots$, at $m = 79$.

Thus steps whose cost grows quadratically in their index would give only $\zeta(s) \ne 0$ for $\operatorname{Re} s > \frac{2}{3}$. The exponent is decided by how the cost of a step grows along the chain, and so by how precisely the size of each step is known.

8.6. Ultrametric norms

Theorem 8.3.1 lives over $A := \mathbb{F}_q[X]$, where the product formula is the count (1). Can one model the chain there?

A *normed lattice* over A is a free A-module L with a basis (e_i) and real numbers (α_i). For $v = \sum_i c_i e_i$ let

$$\|v\| := \max_i \big(\deg c_i - \alpha_i\big),$$

with $\deg 0 = -\infty$. This norm is additive: it plays the role of the logarithm of a norm. A linear map $M \colon L \to L'$ of normed lattices has *norm at most* s if $\|Mv\| \le \|v\| + s$ for all v.

Problem 8.6.1. $\|v + w\| \le \max(\|v\|, \|w\|)$, with equality if $\|v\| \ne \|w\|$. A map with matrix $(m_{i'i})$ has norm at most s if and only if $\deg m_{i'i} \le s + \alpha'_{i'} - \alpha_i$ for all i, i'.

Problem 8.6.2. Let $L_0, \dots, L_m$ be normed lattices, with weights $\alpha_i^{(k)}$, and let $M_k \colon L_{k-1} \to L_k$ have norm at most s_k. Let $\rho \in A$ have degree n, and suppose that ρ divides every entry of every M_k and that $M_m \cdots M_1 \neq 0$. Then there are indices $i_0, \dots, i_m$ such that the entry μ_k of M_k in position (i_k, i_{k-1}) is nonzero for every k, and

$$n \le \deg \mu_k \le s_k + \alpha_{i_k}^{(k)} - \alpha_{i_{k-1}}^{(k-1)} \qquad (1 \le k \le m).$$

Problem 8.6.3. In Problem 8.6.2, $mn \le \sum_k s_k + \alpha_{i_m}^{(m)} - \alpha_{i_0}^{(0)}$. Moreover some k satisfies

$$n \le \deg \mu_k \le s_k + \frac{1}{m}\big(\alpha_{i_m}^{(m)} - \alpha_{i_0}^{(0)}\big).$$

The first inequality is the chain inequality over A, and by Problem 8.6.3 some single entry μ_k, a polynomial divisible by ρ, already does as well as the chain. With $\rho = \prod_{t \in S_\eta}(X - t)$ and entries auxiliary polynomials with $C_0 = 0$, Problem 8.1.4 gives at best $|S_\eta| \le q$. Over $\mathbb{Z}$ a lattice need not have an orthogonal basis, and the norm of a composite can be much smaller than the product of the norms (Problem 8.4.3); whatever lets a chain of theta lifts pay its end costs once must lie in such archimedean inputs.

Is there a model of the chain over a function field in which the costs at the ends are paid once and the chain does better than its single steps? By Problem 8.6.3, normed lattices over $\mathbb{F}_q[X]$ cannot provide one. We do not know the answer.

Hints

§8.1.

8.1.1: Euler's criterion: $f(x)$ is a nonzero square in $\mathbb{F}_q$ exactly when $f(x)^c = 1$.

8.1.2: $dc = \frac{d-1}{2}q + \frac{q-d}{2}$.

8.1.3: Look at the quotient and the remainder of each degree on division by q.

8.1.4: If $r_j = s_j = 0$ for all $j \ge 1$, then $s_0 \neq 0$. Otherwise use Problem 8.1.3.

§8.2.

8.2.1: For the last statement, expand $G(t + Y)$ and substitute $Y = X - t$.

8.2.2: The map $u \mapsto u^q$ is a ring homomorphism in characteristic p.

8.2.3: Write $f(X + Y) = f(X) + \Delta$ with $\Delta = \sum_{k \ge 1}(H^k f)Y^k$, and expand $(1 + \Delta/f)^c$ by the binomial theorem. Note $\deg H^k f \le d - k$.

8.2.4: Expand $F(t+Y)$ modulo Y^{k+1}. Use Problem 8.2.2, $t^q = t$, and $f(t)^c = \eta$.

8.2.5: Bound each of the two kinds of terms in (3) separately, using Problem 8.2.3.

§8.3.

8.3.1: Count unknowns and linear conditions, using Problem 8.2.5.

8.3.2: Use $2D = q - d$.

8.3.3: Check that $\sqrt{q} \ge 12$, $q - d - \ell \ge 8q/9$ and $11\sqrt{q}/12 \le \ell \le 13\sqrt{q}/12$. Then bound $t - 1$ and g by multiples of d.

8.3.4: $t > d\ell^2/(2q)$. Then use the inequality between arithmetic and geometric means.

§8.4.

8.4.1: If $\varphi_j(L_{j-1}) \subseteq \lambda_j L_j$ for each j, then $\Phi(L_0) \subseteq \lambda_1 \cdots \lambda_m L_m$. Operator norms are submultiplicative.

8.4.2: Write out the m inequalities and add them.

8.4.3: Every nonzero vector of L_1 has length at least 1.

§8.5.

8.5.1: Apply Theorem 8.4.2, and divide by mb.

8.5.2: $\sum_{j \le m}(2j-1) = m^2$.

8.5.3: Each step gives a bound for the same quantity.

8.5.4: Minimize over real $m > 0$ first.

§8.6.

8.6.1: $\deg(c + c') \le \max(\deg c, \deg c')$, with equality if the degrees differ. For the second statement test on the basis vectors.

8.6.2: An entry of the product is a sum over paths of products of entries. A nonzero multiple of ρ has degree at least n.

8.6.3: The weights telescope. The least of m numbers is at most their mean.

References

[Adr] J. Adriazola, *The Alpöge proof of the Riemann hypothesis*, manuscript, revised 24 November 2026, `https://github.com/jimjamzola/RH-Alpoge-Reconstruction`.

[Alp] L. Alpöge, untitled note announcing a proof of the Riemann hypothesis, 20 November 2026, `https://alpo.ge/remark.pdf`.

[Bom74] E. Bombieri, *Counting points on curves over finite fields (d'après S. A. Stepanov)*, Séminaire Bourbaki, Exp. 430, Lecture Notes in Math. **383**, Springer, 1974, 234–241.

[Bo1] J.-B. Bost, *Périodes et isogénies des variétés abéliennes sur les corps de nombres (d'après D. Masser et G. Wüstholz)*, Séminaire Bourbaki, Exp. 795, Astérisque **237** (1996), 115–161.

[Dup] *Notes on Alpöge's proof of the Riemann hypothesis*, printed by T. Dupuy, AI generated, 21 November 2026, `https://tdupu.github.io/riemann-public/riemann-book.pdf`.

[Sch76] W. M. Schmidt, *Equations over Finite Fields: An Elementary Approach*, Lecture Notes in Math. **536**, Springer, 1976.

[Ste69] S. A. Stepanov, *On the number of points of a hyperelliptic curve over a finite prime field*, Izv. Akad. Nauk SSSR Ser. Mat. **33** (1969), 1171–1181.

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

[Zh] R. Zhang, *A preliminary note on Alpöge's proposed proof of the generalized Riemann hypothesis*, preprint, 2026, hexagon:2611.00809.