Gate Complexity of the Algebraic Torus

Anonymous

Abstract

A (p,q)-*gate* is a function $\mathbb{F}_q^n \to \mathbb{F}_p$ of the form $g \circ \ell$, where $\ell\colon \mathbb{F}_q^n \to \mathbb{F}_q$ is affine and $g\colon \mathbb{F}_q \to \mathbb{F}_p$ is arbitrary; equivalently, it is a depth-2 subcircuit consisting of a single MOD-q gate followed by a MOD-p gate. We determine the *gate complexity* $t(p,q,n)$—the minimum number of (p,q)-gates whose $\mathbb{F}_p$-linear combination equals the indicator function of the algebraic torus $(\mathbb{F}_q^*)^n$—for all primes p and prime powers q with $\mathrm{char}(\mathbb{F}_q) \neq p$.

The answer exhibits a dichotomy governed by a single divisibility condition:

$$t(p,q,n) = \begin{cases} (q-1)^{n-1} & \text{if } p \mid (q-1), \\ \dfrac{q^n-1}{q-1} = |\mathbb{P}^{n-1}(\mathbb{F}_q)| & \text{if } p \nmid (q-1). \end{cases}$$

When $p \mid (q-1)$, the $\mathbb{F}_{p^k}$-Fourier transform of $\mathbf{1}_T$ is supported on the torus T, and the optimal construction uses $(q-1)^{n-1}$ gates indexed by $(\mathbb{F}_q^*)^{n-1}$. When $p \nmid (q-1)$, the Fourier transform has full support on $\mathbb{F}_q^n \setminus \{0\}$, and the optimal construction requires one gate per point of $\mathbb{P}^{n-1}(\mathbb{F}_q)$.

In both cases, the upper bound is a Fourier inversion identity and the lower bound is a Frobenius orbit counting argument. We also show that depth-3 circuits escape the exponential barrier: for $n < p$, the torus indicator can be computed with $O(n)$ gates, giving an exponential depth-2 vs. depth-3 separation. We discuss implications for $\mathsf{AC}^0[6]$: the same-vs-cross-characteristic dichotomy in gate span (cross-characteristic gates span all functions; same-characteristic gates do not) isolates the algebraic core of the difficulty, while the depth-3 escape shows that the Fourier-based lower bound technique is inherently limited to depth 2.

1 Introduction

A central open problem in circuit complexity is to prove super-polynomial lower bounds for $\mathsf{AC}^0[6]$, the class of constant-depth circuits with AND, OR, NOT, and MOD-6 gates. The Razborov–Smolensky method [6, 7] gives exponential lower bounds for $\mathsf{AC}^0[p]$ when p is prime, but breaks down for composite moduli like $6 = 2 \times 3$.

The key difficulty is the interaction between different characteristics. MOD-6 gates can simulate both MOD-2 and MOD-3, combining information from $\mathbb{F}_2$ and $\mathbb{F}_3$ in a way that resists standard polynomial methods. In this paper we isolate this cross-characteristic interaction in its simplest form and study it through the lens of algebraic coding theory.

The model. The gate complexity model is inherently depth-2: a single layer of affine maps $\ell_i : \mathbb{F}_q^n \to \mathbb{F}_q$ composed with arbitrary functions $g_i : \mathbb{F}_q \to \mathbb{F}_p$, followed by an $\mathbb{F}_p$-linear combination. This corresponds to a depth-2 circuit with one layer of MOD-q gates feeding into a single MOD-p output gate.

We consider the gate complexity $t(p,q,n)$: the minimum number of (p,q)-gates needed to represent the indicator function $\mathbf{1}_T$ of the algebraic torus $T = (\mathbb{F}_q^*)^n$ as an $\mathbb{F}_p$-linear combination.

Here a (p, q)-gate is a composition $g \circ \ell$ where $\ell : \mathbb{F}_q^n \to \mathbb{F}_q$ is affine and $g : \mathbb{F}_q \to \mathbb{F}_p$ is arbitrary. The function $\mathbf{1}_T$ is the canonical "hard function" for this model: it is nonzero precisely on the torus, the complement of the union of coordinate hyperplanes.

Example 1.1. *For $p = 2$, $q = 3$, $n = 2$, a typical $(2, 3)$-gate on $\mathbb{F}_3^2$ is*

$$g(x_1 + 2x_2), \quad \text{where } g\colon \mathbb{F}_3 \to \mathbb{F}_2 \text{ is defined by } g(0) = 1,\ g(1) = 0,\ g(2) = 0.$$

This gate outputs 1 if and only if $x_1 + 2x_2 \equiv 0 \pmod 3$. The affine map $\ell(x_1, x_2) = x_1 + 2x_2$ computes a MOD-3 linear form, and g applies an arbitrary Boolean post-processing. The gate complexity $t(2, 3, n)$ asks: how many such gates must be summed over $\mathbb{F}_2$ to produce the indicator of $(\mathbb{F}_3^*)^n$? Our main theorem gives $t(2, 3, n) = 2^{n-1}$.

Scope and limitations. Our exponential lower bound applies to this restricted depth-2 setting. A full $\mathsf{AC}^0[6]$ circuit has arbitrary constant depth, and the central open problem is precisely to understand how cross-characteristic interactions compose across multiple layers. We show in Section 9 that depth-3 already escapes the exponential barrier, reducing the gate count from exponential to linear.

1.1 Main Results

Our main result determines the gate complexity for all primes p and prime powers q with $\mathrm{char}(\mathbb{F}_q) \neq p$.

Theorem 1.2 (Main Theorem)**.** *Let p be a prime and q a prime power with $\mathrm{char}(\mathbb{F}_q) \neq p$. Then*

$$t(p, q, n) = \begin{cases} (q-1)^{n-1} & \text{if } p \mid (q-1), \\ \dfrac{q^n - 1}{q-1} = |\mathbb{P}^{n-1}(\mathbb{F}_q)| & \text{if } p \nmid (q-1). \end{cases}$$

The dichotomy is governed by a single divisibility condition. Note that for $p = 2$, the condition $2 \mid (q-1)$ holds for all odd q, so the formula simplifies to $t(2, q, n) = (q-1)^{n-1}$. For $q = 2$, we have $q - 1 = 1$, so $p \nmid 1$ for all primes $p \geq 3$, giving $t(p, 2, n) = 2^n - 1 = |\mathbb{P}^{n-1}(\mathbb{F}_2)|$.

Additional results.

(1) **Coding-theoretic framework (Section 2).** We reduce gate complexity to a minimum coset weight problem in a linear code over $\mathbb{F}_p$, with quotient dimension $\dim(C/C_0) = (q-1)^n$ in the cross-characteristic case.

(2) **Gate span completeness (Theorem 3.1).** Cross-characteristic gates span all functions $\mathbb{F}_q^n \to \mathbb{F}_p$. This fails in same characteristic, explaining the algebraic core of the $\mathsf{AC}^0[6]$ difficulty.

(3) **Fourier support dichotomy (Theorem 4.2).** Over $\mathbb{F}_{p^k}$, the Fourier transform $\widehat{\mathbf{1}_T}$ is supported on T when $p \mid (q-1)$ and on $\mathbb{F}_q^n \setminus \{0\}$ when $p \nmid (q-1)$.

(4) **Depth-3 escape (Section 9).** For $n < p$, the torus indicator can be computed with $n + 1$ gates at depth-3—an exponential improvement over depth-2.

(5) **Vandermonde obstruction (Section 8).** For $q = 3$, an alternative Fourier support induction gives the lower bound. This method fails for $q \geq 5$, motivating the orbit counting argument and illustrating the subtlety of the cross-characteristic setting.

1.2 Techniques

Upper bound. The construction is a Fourier inversion identity decomposed over projective lines. For each projective point $[a] \in \mathbb{P}^{n-1}(\mathbb{F}_q)$, we define a gate $g_{[a]} \circ \ell_a$ where $\ell_a(x) = a \cdot x$ and $g_{[a]}(v) = c_{[a]} \cdot \mathbf{1}_{v=0}$ with explicit coefficients $c_{[a]}$. The key observation is that the coefficients $c_{[a]}$ vanish in $\mathbb{F}_p$ precisely when $p \mid (q-1)$ and $a \notin T$, reducing the gate count from $|\mathbb{P}^{n-1}(\mathbb{F}_q)|$ to $(q-1)^{n-1}$ in this case.

Lower bound. The lower bound proceeds by a Frobenius orbit counting argument. The $\mathbb{F}_{p^k}$-Fourier transform (where $k = \mathrm{ord}_r(p)$ and $r = \mathrm{char}(\mathbb{F}_q)$) has the property that Fourier support is closed under the Frobenius action $\alpha \mapsto p\alpha$. We show:

- Each gate's Fourier support lies on a single $\mathbb{F}_q$-line through the origin.
- Each such line contains at most $(q-1)/k$ Frobenius orbits in its torus part.
- The Fourier support of $\mathbf{1}_T$ consists of all torus orbits (when $p \mid (q-1)$) or all nonzero orbits (when $p \nmid (q-1)$).
- Covering all required orbits forces $w \geq (q-1)^{n-1}$ or $w \geq (q^n-1)/(q-1)$ gates.

The factors of k cancel perfectly, so the final answer depends only on p, q, and n—not on the multiplicative order of p in $\mathbb{F}_r^*$.

Discussion. The conceptual message is a dichotomy: cross-characteristic gates always span the full function space, but doing so efficiently requires overcoming a Fourier-theoretic obstruction that grows exponentially in n. The formula reveals that the growth rate is controlled by either the torus dimension $|(\mathbb{F}_q^*)|^{n-1} = (q-1)^{n-1}$ or the projective space dimension $|\mathbb{P}^{n-1}(\mathbb{F}_q)| = (q^n-1)/(q-1)$, with the divisibility $p \mid (q-1)$ determining which regime applies.

1.3 Related Work

The polynomial method of Razborov [6] and Smolensky [7] gives exponential lower bounds for $\mathsf{AC}^0[p]$ for prime p, but fails for composite moduli. Barrington, Straubing, and Thérien [3] studied the algebraic structure of ACC^0 and showed connections to group theory. Viola [9] surveyed the state of small-depth computation and highlighted the $\mathsf{AC}^0[6]$ problem as a central challenge. Williams [10] proved nonuniform ACC^0 lower bounds via a different route (satisfiability algorithms), but the uniform case remains open.

Our model isolates a single cross-characteristic interaction layer, which is the building block that the Razborov–Smolensky method cannot handle. The exponential lower bound shows that even one such layer already requires exponentially many gates, while the depth-3 escape shows that the difficulty is not in the interaction itself but in composing interactions across layers—precisely the obstacle for $\mathsf{AC}^0[6]$.

In the purely linear setting, recent work of Alman and Li [1] studies depth-2 linear circuits for Kronecker power matrices $M^{\otimes k}$, using Strassen's asymptotic spectrum theory to give optimal constructions. A key finding in this line of work [2, 5, 1] is that Kronecker products of small matrices are generally *not* Valiant-rigid, enabling sub-$N^{1.5}$ circuits. Our setting differs in two respects: the gates include nonlinear post-processing $g\colon \mathbb{F}_q \to \mathbb{F}_p$, and the target function $\mathbf{1}_T$ has cross-characteristic structure that is not captured by Kronecker powers. The exponential lower

bound in our model, contrasted with the improved upper bounds for Kronecker powers, highlights the role of the nonlinear cross-characteristic interaction as the source of hardness.

The connection between gate complexity and coding theory parallels work on toric codes [4, 8], where code parameters are controlled by lattice geometry.

1.4 Organization

Section 2 establishes the coding-theoretic framework. Section 3 proves gate span completeness. Section 4 develops the $\mathbb{F}_{p^k}$-Fourier transform and proves the support dichotomy. Section 5 proves the lower bound via orbit counting. Section 6 proves the upper bound via Fourier inversion. Section 7 analyzes the special case $q = 3$ in detail. Section 8 gives the alternative Vandermonde induction proof for $q = 3$ and shows why it fails for $q \geq 5$. Section 9 proves that depth-3 circuits escape the exponential barrier. Section 10 discusses connections to AC^0[6] and future directions.

2 The Coding-Theoretic Framework

2.1 Setup and Notation

Throughout, p is a prime, q is a prime power with $\mathrm{char}(\mathbb{F}_q) = r \neq p$, and $n \geq 1$. Write $T = (\mathbb{F}_q^*)^n$ for the algebraic torus and $Z = \mathbb{F}_q^n \setminus T$ for the boundary.

Definition 2.1. *A (p, q)-gate on $\mathbb{F}_q^n$ is a function $g \circ \ell : \mathbb{F}_q^n \to \mathbb{F}_p$, where $\ell(u) = a \cdot u + b$ is affine ($a \in \mathbb{F}_q^n$, $b \in \mathbb{F}_q$) and $g : \mathbb{F}_q \to \mathbb{F}_p$ is arbitrary.*

Let G denote the set of all distinct gate evaluation vectors, with $|G| = G$, and form the gate evaluation matrix $M \in \mathbb{F}_p^{q^n \times G}$.

Definition 2.2. *The gate complexity is*

$$t(p, q, n) = \min\{\mathrm{wt}(c) : c \in \mathbb{F}_p^G, M_Z c = 0, M_T c = \mathbf{1}_T\}.$$

2.2 The Code and Its Quotient

Define linear codes over $\mathbb{F}_p$:

$$C = \ker(M_Z) = \{c \in \mathbb{F}_p^G : M_Z c = 0\},$$
$$C_0 = \ker(M) = \{c \in \mathbb{F}_p^G : Mc = 0\}.$$

The quotient C/C_0 maps isomorphically onto $\mathbb{F}_p^T$: every function $T \to \mathbb{F}_p$ is realizable. The target $\mathbf{1}_T$ determines a coset $c_0 + C_0$ inside C, and $t(p, q, n) = \min_{c \in c_0 + C_0} \mathrm{wt}(c)$.

3 Gate Span Completeness

Theorem 3.1. *Let p be a prime and q a prime power with $\mathrm{char}(\mathbb{F}_q) \neq p$. Then $\mathrm{span}_{\mathbb{F}_p}(G) = \mathbb{F}_p^{\mathbb{F}_q^n}$, and consequently $\dim(C/C_0) = (q-1)^n$.*

Proof. We prove the contrapositive: any $\lambda : \mathbb{F}_q^n \to \mathbb{F}_p$ annihilating every gate must be zero.

Step 1. If $\sum_u \lambda(u)(g \circ \ell)(u) = 0$ for all gates, then choosing $g = \delta_v$ shows that each fiber sum $\sum_{\ell(u)=v} \lambda(u) = 0$ for all nonconstant ℓ and all v.

Step 2. Since $\mathrm{char}(\mathbb{F}_q) \neq p$, fix a nontrivial additive character $\psi : (\mathbb{F}_q, +) \to \mathbb{F}_p[\zeta]^*$. Multiplying fiber sums by $\psi(v)$ and summing gives $\widehat{\lambda}(\psi_a) = 0$ for all nonzero a.

Step 3. Since q^n is coprime to p, the DFT is invertible in $\mathbb{F}_p[\zeta]$. All Fourier coefficients vanishing implies $\lambda \equiv 0$.

The dimension formula follows: $\mathrm{rank}(M) = q^n$, $\mathrm{rank}(M_Z) = q^n - (q-1)^n$, so $\dim(C/C_0) = (q-1)^n$. □

Remark 3.2. *When $p = \mathrm{char}(\mathbb{F}_q)$, the DFT is not invertible and nontrivial annihilators exist. The quotient dimension collapses: for $p = q = 3$, $n = 2$, one has $\dim(C/C_0) = 1$ versus $(q-1)^n = 4$ in the cross-characteristic case. This dichotomy is the algebraic core of the difficulty of AC^0[6].*

4 The $\mathbb{F}_{p^k}$-Fourier Transform

4.1 Setup

Let $r = \mathrm{char}(\mathbb{F}_q)$ and $k = \mathrm{ord}_r(p)$, the multiplicative order of p in $\mathbb{F}_r^*$. Since $r \mid p^k - 1$, the field $\mathbb{F}_{p^k}$ contains a primitive rth root of unity ζ.

Fix the nontrivial additive character $\chi : \mathbb{F}_q \to \mathbb{F}_{p^k}^*$ defined by $\chi(x) = \zeta^{\mathrm{Tr}(x)}$, where $\mathrm{Tr} : \mathbb{F}_q \to \mathbb{F}_r$ is the field trace. (For q prime, this reduces to $\chi(x) = \zeta^x$.) The $\mathbb{F}_{p^k}$-Fourier transform of $f : \mathbb{F}_q^n \to \mathbb{F}_{p^k}$ is

$$\widehat{f}(\alpha) = \sum_{x \in \mathbb{F}_q^n} f(x)\chi(-\alpha \cdot x), \quad \alpha \in \mathbb{F}_q^n.$$

Since $\mathbb{F}_p \subset \mathbb{F}_{p^k}$, any function $f : \mathbb{F}_q^n \to \mathbb{F}_p$ has a well-defined $\mathbb{F}_{p^k}$-Fourier transform.

The Frobenius $\sigma : x \mapsto x^p$ acts on $\mathbb{F}_{p^k}$ with order k. Since Tr is $\mathbb{F}_r$-linear and $p \in \mathbb{F}_r$, we have $\sigma(\chi(v)) = \chi(v)^p = \zeta^{p\mathrm{Tr}(v)} = \zeta^{\mathrm{Tr}(pv)} = \chi(pv)$, so σ acts on $\mathbb{F}_q^n$ as $\alpha \mapsto p\alpha$ (scalar multiplication by $p \in \mathbb{F}_q$). For f taking values in $\mathbb{F}_p = \mathbb{F}_{p^k}^{\sigma}$:

$$\widehat{f}(p\alpha) = \widehat{f}(\alpha)^p, \tag{1}$$

so the Fourier support is a union of Frobenius orbits.

4.2 Fourier Support Dichotomy

Proposition 4.1. *Over $\mathbb{F}_{p^k}$, the Fourier transform of $\mathbf{1}_T$ is:*

$$\widehat{\mathbf{1}_T}(\alpha) = \prod_{j=1}^{n} S(\alpha_j), \quad S(a) = \sum_{c \in \mathbb{F}_q^*} \chi(-ac).$$

The per-coordinate factor satisfies:

$$S(a) = \begin{cases} q-1 & \text{if } a = 0, \\ -1 & \text{if } a \neq 0. \end{cases}$$

Proof. The torus indicator factorizes as $\mathbf{1}_T(x) = \prod_j \mathbf{1}_{x_j \neq 0}$, so the Fourier transform factorizes. For the sum $S(a) = \sum_{c \in \mathbb{F}_q^*} \chi(-ac)$: if $a = 0$, every term is 1 and $S(0) = q - 1$. If $a \neq 0$, the map $c \mapsto -ac$ is a bijection on $\mathbb{F}_q^*$, so $S(a) = \sum_{t \in \mathbb{F}_q^*} \chi(t) = \sum_{t \in \mathbb{F}_q} \chi(t) - 1 = 0 - 1 = -1$. □

Theorem 4.2 (Fourier Support Dichotomy)**.** *Let $m(\alpha) = |\{j : \alpha_j = 0\}|$ for $\alpha \in \mathbb{F}_q^n$. Then in $\mathbb{F}_{p^k}$:*

$$\widehat{\mathbf{1}_T}(\alpha) = (-1)^{n-m(\alpha)}(q-1)^{m(\alpha)}.$$

Consequently:

(i) If $p \mid (q-1)$: $\widehat{\mathbf{1}_T}(\alpha) \neq 0 \iff \alpha \in T$. In particular, $\widehat{\mathbf{1}_T}(\alpha) = (-1)^n = \mathbf{1}_T(\alpha)$ for $p = 2$, recovering self-duality.

(ii) If $p \nmid (q-1)$: $\widehat{\mathbf{1}_T}(\alpha) \neq 0 \iff \alpha \neq 0$. The Fourier transform has full support on $\mathbb{F}_q^n \setminus \{0\}$.

Proof. By Proposition 4.1, $\widehat{\mathbf{1}_T}(\alpha) = \prod_j S(\alpha_j) = (-1)^{n-m(\alpha)}(q-1)^{m(\alpha)}$. This vanishes in $\mathbb{F}_{p^k}$ if and only if $m(\alpha) \geq 1$ and $q - 1 \equiv 0 \pmod{p}$. □

5 Lower Bound

Lemma 5.1 (Gate Fourier support)**.** *If $g \circ \ell$ is a gate with $\ell(x) = a \cdot x + b$, then* $\operatorname{supp}(\widehat{g \circ \ell}) \subseteq \mathbb{F}_q \cdot a$.

Proof. The Fourier transform of $g \circ \ell$ at α involves a sum over the affine hyperplane $\{x : a \cdot x + b = v\}$. This sum vanishes unless $\alpha \in (\ker a)^\perp = \mathbb{F}_q \cdot a$. □

Lemma 5.2 (Frobenius orbits)**.** *Let $k = \operatorname{ord}_r(p)$. The Frobenius $\alpha \mapsto p\alpha$ acts on $\mathbb{F}_q^n \setminus \{0\}$ with orbits of size exactly k. Each line $\mathbb{F}_q \cdot a$ through a nonzero a contains:*

(a) $(q-1)/k$ Frobenius orbits lying in $\mathbb{F}_q^* \cdot a$ (the torus part of the line), and

(b) one additional orbit $\{0\}$ (which has size 1).

For $a \in T$, the line $\mathbb{F}_q \cdot a$ meets T in exactly $(q-1)/k$ Frobenius orbits.

Proof. The orbits of $\mathbb{F}_q^*$ under multiplication by p have size exactly $k = \operatorname{ord}_r(p)$ (since $p^j\alpha = \alpha$ with $\alpha \neq 0$ implies $p^j = 1$ in $\mathbb{F}_q$), giving $(q-1)/k$ orbits. The line $\mathbb{F}_q \cdot a$ intersected with $\mathbb{F}_q^n \setminus \{0\}$ is $\mathbb{F}_q^* \cdot a$, which inherits the orbit decomposition. □

Theorem 5.3 (Lower bound)**.** *For all primes p and prime powers q with* $\operatorname{char}(\mathbb{F}_q) \neq p$*:*

$$t(p,q,n) \geq \begin{cases} (q-1)^{n-1} & \text{if } p \mid (q-1), \\ 1 + q + \cdots + q^{n-1} & \text{if } p \nmid (q-1). \end{cases}$$

Proof. Suppose $\mathbf{1}_T = \sum_{i=1}^w c_i(g_i \circ \ell_i)$ with $c_i \in \mathbb{F}_p^*$. Taking $\mathbb{F}_{p^k}$-Fourier transforms:

$$\widehat{\mathbf{1}_T} = \sum_{i=1}^{w} c_i \widehat{g_i \circ \ell_i}.$$

For any α with $\widehat{\mathbf{1}_T}(\alpha) \neq 0$, at least one gate must satisfy $\widehat{g_i \circ \ell_i}(\alpha) \neq 0$, placing α on the line $\mathbb{F}_q \cdot a_i$ by Lemma 5.1. Since the Fourier support is a union of Frobenius orbits by (1), each such orbit must be covered by some gate.

Case $p \mid (q-1)$**:** By Theorem 4.2(i), the Fourier support is T. The torus has $(q-1)^n/k$ Frobenius orbits, and each gate line covers at most $(q-1)/k$:

$$w \cdot \frac{q-1}{k} \geq \frac{(q-1)^n}{k} \implies w \geq (q-1)^{n-1}.$$

Case $p \nmid (q-1)$: By Theorem 4.2(ii), the Fourier support is $\mathbb{F}_q^n \setminus \{0\}$, which has $(q^n-1)/k$ Frobenius orbits. Each gate line covers at most $(q-1)/k$ orbits in $\mathbb{F}_q^n \setminus \{0\}$ (namely the orbits in $\mathbb{F}_q^* \cdot a_i$):

$$w \cdot \frac{q-1}{k} \geq \frac{q^n-1}{k} \implies w \geq \frac{q^n-1}{q-1} = |\mathbb{P}^{n-1}(\mathbb{F}_q)|.$$

□

Remark 5.4. *The factors of k cancel perfectly in the lower bound. This means the gate complexity depends only on q and n, not on the multiplicative order of p. The extension field $\mathbb{F}_{p^k}$ serves as an auxiliary tool but leaves no trace in the final answer.*

6 Upper Bound

Theorem 6.1 (Upper bound)**.** *For all primes p and prime powers q with $\mathrm{char}(\mathbb{F}_q) \neq p$ and $n \geq 1$:*

$$t(p,q,n) \leq \begin{cases} (q-1)^{n-1} & \text{if } p \mid (q-1), \\ 1+q+\cdots+q^{n-1} & \text{if } p \nmid (q-1). \end{cases}$$

Proof. For each nonzero direction $a \in \mathbb{F}_q^n \setminus \{0\}$, define the homogeneous linear form $\ell_a(x) = a \cdot x$ and the gate function $g_a : \mathbb{F}_q \to \mathbb{F}_p$ by

$$g_a(v) = c_{[a]} \cdot \mathbf{1}_{v=0},$$

where $[a]$ denotes the projective class of a and

$$c_{[a]} = \frac{(-1)^{n-m(a)} \cdot (q-1)^{m(a)}}{q^{n-1}} \in \mathbb{F}_p, \tag{2}$$

with $m(a) = |\{j : a_j = 0\}|$ as before, and q^{n-1} is inverted in $\mathbb{F}_p$ (possible since $\mathrm{char}(\mathbb{F}_q) \neq p$). The coefficient $c_{[a]}$ depends only on the projective class $[a]$ since $m(ta) = m(a)$ for $t \in \mathbb{F}_q^*$.

Claim: The function

$$F(x) = \sum_{[a] \in \mathbb{P}^{n-1}(\mathbb{F}_q)} c_{[a]} \cdot \mathbf{1}_{a \cdot x = 0}$$

satisfies $F(x) = \mathbf{1}_T(x) + C$ for a constant $C \in \mathbb{F}_p$.

Proof of claim. Expand each indicator using the additive characters of $\mathbb{F}_q$:

$$\mathbf{1}_{a \cdot x = 0} = \frac{1}{q} \sum_{s \in \mathbb{F}_q} \chi(s \cdot a \cdot x) = \frac{1}{q} + \frac{1}{q} \sum_{s \in \mathbb{F}_q^*} \chi(s \cdot a \cdot x).$$

Substituting into F and using $\alpha = sa$ to parametrize $\mathbb{F}_q^n \setminus \{0\}$:

$$F(x) = C_0 + \frac{1}{q} \sum_{\alpha \in \mathbb{F}_q^n \setminus \{0\}} \frac{c_{[\alpha]}}{q-1} \chi(\alpha \cdot x),$$

where we used the fact that each $\alpha \neq 0$ is counted once for each $s \in \mathbb{F}_q^*$ in its projective class, and the factor $1/(q-1)$ compensates.

By Fourier inversion, $\mathbf{1}_T(x) = q^{-n} \sum_\alpha \widehat{\mathbf{1}_T}(\alpha) \chi(\alpha \cdot x)$. Matching coefficients shows $F(x) = \mathbf{1}_T(x) + C$ for some constant C.

Since a constant function can be absorbed into any single gate (by adjusting $g_a(v)$ for one gate), the number of gates equals the number of projective classes $[a]$ for which $c_{[a]} \neq 0$ in $\mathbb{F}_p$.

Counting nonzero gates. The coefficient $c_{[a]} = (-1)^{n-m(a)}(q-1)^{m(a)}/q^{n-1}$ vanishes in $\mathbb{F}_p$ if and only if $p \mid (q-1)$ and $m(a) \geq 1$ (since q^{n-1} is invertible and $(-1)^{n-m(a)}$ is a unit).

- If $p \mid (q-1)$: $c_{[a]} \neq 0$ only when $m(a) = 0$, i.e., $a \in T$. The number of such projective classes is $|T|/(q-1) = (q-1)^{n-1}$.
- If $p \nmid (q-1)$: $c_{[a]} \neq 0$ for all $[a] \in \mathbb{P}^{n-1}(\mathbb{F}_q)$, giving $(q^n-1)/(q-1)$ gates.

□

Proof of Theorem 1.2. Combine Theorem 5.3 and Theorem 6.1. □

7 The Special Case $q = 3$

For $q = 3$ and $p = 2$, the formula gives $t(2, 3, n) = 2^{n-1}$. This case admits a more detailed analysis.

7.1 Explicit Construction

The gates are indexed by $s \in (\mathbb{F}_3^*)^{n-1} = \{1, 2\}^{n-1}$. For each $s = (s_1, \ldots, s_{n-1})$, define

$$\ell_s(x) = x_1 + \sum_{j=2}^{n} s_{j-1} x_j, \quad g_s = \mathbf{1}_{\ell_s \neq 0}.$$

Then $\bigoplus_{s \in \{1,2\}^{n-1}} g_s(\ell_s(x)) = \mathbf{1}_T(x)$ in $\mathbb{F}_2$.

7.2 Solution Structure

Theorem 7.1. *For $q = 3$: every weight-2^{n-1} gate combination representing $\mathbf{1}_T$ uses the 2^{n-1} linear forms $\{\ell_s : s \in (\mathbb{F}_3^*)^{n-1}\}$ (up to a choice of distinguished coordinate). The only freedom is in the gate function: each form ℓ_s can be paired with either $\mathbf{1}_{\ell_s \neq 0}$ or $\mathbf{1}_{\ell_s = 0}$, subject to an even-parity constraint. This gives $2^{2^{n-1}-1}$ solutions.*

Proof. The tightness of the lower bound (Theorem 5.3) forces each of the 2^{n-1} gates to cover a distinct Frobenius orbit in T, so the 2^{n-1} linear forms used must be exactly one representative per projective class in T. Up to a choice of distinguished coordinate, these are $\{\ell_s : s \in (\mathbb{F}_3^*)^{n-1}\}$.

Given these forms, on the torus $T = (\mathbb{F}_3^*)^n$, the functions $\mathbf{1}_{\ell_s \neq 0}|_T$ and $\mathbf{1}_{\ell_s = 0}|_T$ are complementary: their XOR is the constant function 1 on T. Flipping the gate function for ℓ_s changes the contribution on T by $\mathbf{1}|_T$, while preserving the vanishing on Z. Flipping an even number of gate functions preserves the global XOR being $\mathbf{1}_T$, giving $2^{2^{n-1}-1}$ valid assignments. □

7.3 The ψ-Independence Theorem

The construction uses 2^{n-1} canonical gates $g_s = \mathbf{1}_{\ell_s \neq 0}$. The following theorem shows these are linearly independent, so the canonical construction is locally optimal.

Definition 7.2. *For $m \geq 0$ and $s = (s_1, \ldots, s_m) \in \{1, 2\}^m$, define $\psi_s : \mathbb{F}_3^{m+1} \to \mathbb{F}_2$ by*

$$\psi_s(x_1, \ldots, x_{m+1}) = \mathbf{1}_{x_1 + \sum_{k=1}^{m} s_k x_{k+1} \equiv 0 \pmod 3}.$$

Theorem 7.3 (ψ-Independence)**.** *For all $m \geq 0$, the 2^m functions $\{\psi_s : s \in \{1,2\}^m\}$ satisfy:*

(a) They are $\mathbb{F}_2$-linearly independent on $\mathbb{F}_3^{m+1}$.

(b) The constant function 1 is not in their $\mathbb{F}_2$-span.

Proof. By strong induction on m, proving (a) and (b) simultaneously.

Base case ($m = 0$). The single function $\psi(x_1) = \mathbf{1}_{x_1=0}$ is nonzero, hence independent. And $\psi \neq 1$ since $\psi(1) = 0$.

Inductive step. Assume both statements hold for all $m' < m$. Suppose $\bigoplus_{s\in S} \psi_s = 0$ for some nonempty $S \subseteq \{1,2\}^m$.

Step 1: Restrict to $\{x_{m+1} = 0\}$. On this slice, $\psi_{(s',s_m)}$ reduces to $\psi_{s'}^{(m-1)}$, independently of s_m. Write $\varepsilon_j(s') = \mathbf{1}_{(s',j)\in S}$ for $j \in \{1,2\}$. The restricted equation becomes $\bigoplus_{s'}(\varepsilon_1(s') \oplus \varepsilon_2(s'))\psi_{s'}^{(m-1)} = 0$. By induction (a) for $m-1$, we conclude $\varepsilon_1(s') = \varepsilon_2(s')$ for all s'.

Define $S_0 = \{s' \in \{1,2\}^{m-1} : (s',1) \in S\} = \{s' : (s',2) \in S\}$.

Step 2: Restrict to $\{x_{m+1} = 1\}$. On this slice, $\psi_{(s',1)}|_{x_{m+1}=1} \oplus \psi_{(s',2)}|_{x_{m+1}=1} = \mathbf{1}_{\ell_{s'}\neq 0} = 1 \oplus \psi_{s'}^{(m-1)}$. Summing over $s' \in S_0$:

$$\bigoplus_{s'\in S_0} (1 \oplus \psi_{s'}^{(m-1)}) = 0, \quad \text{giving} \quad \bigoplus_{s'\in S_0} \psi_{s'}^{(m-1)} = |S_0| \mod 2.$$

If $|S_0|$ is even, induction (a) gives $S_0 = \emptyset$. If $|S_0|$ is odd, induction (b) is contradicted. Either way $S = \emptyset$, proving (a). Part (b) follows similarly by restricting the equation $\bigoplus_S \psi_s = 1$ to $\{x_{m+1} = 0\}$ and applying induction (b). □

Corollary 7.4. *The 2^{n-1} canonical gates $g_s = \mathbf{1}_{\ell_s\neq 0}$ for $s \in (\mathbb{F}_3^*)^{n-1}$ are $\mathbb{F}_2$-linearly independent as functions on $\mathbb{F}_3^n$.*

8 Vandermonde Induction for $q = 3$

For the special case $q = 3$, we give an alternative lower bound proof that establishes a stronger result: an $\mathbb{F}_4$-Fourier support theorem for all functions supported on T.

8.1 Coordinate Slicing

Write $f : \mathbb{F}_3^n \to \mathbb{F}_4$ and define $f_1(x') = f(1,x')$, $f_2(x') = f(2,x')$ for $x' \in \mathbb{F}_3^{n-1}$. Then

$$\widehat{f}(\alpha_1,\alpha') = \omega^{-\alpha_1}\widehat{f_1}(\alpha') + \omega^{\alpha_1}\widehat{f_2}(\alpha'),$$

since $-2\alpha_1 = \alpha_1$ in $\mathbb{F}_3$, where $\omega \in \mathbb{F}_4$ is a primitive cube root of unity (satisfying $\omega^2 + \omega + 1 = 0$).

For fixed α', the three values $\widehat{f}(0,\alpha'), \widehat{f}(1,\alpha'), \widehat{f}(2,\alpha')$ are the entries of

$$\begin{pmatrix} 1 & 1 \\ \omega^2 & \omega \\ \omega & \omega^2 \end{pmatrix} \begin{pmatrix} \widehat{f_1}(\alpha') \\ \widehat{f_2}(\alpha') \end{pmatrix}.$$

Since this 3×2 Vandermonde matrix over $\mathbb{F}_4$ has every 2×2 submatrix nonsingular:

Lemma 8.1 (Slicing Lemma)**.** *For each $\alpha' \in \mathbb{F}_3^{n-1}$:*

(a) If $\widehat{f_1}(\alpha') = \widehat{f_2}(\alpha') = 0$, then $\widehat{f}(\alpha_1,\alpha') = 0$ for all α_1.

(b) If exactly one is nonzero, then $\widehat{f}(\alpha_1, \alpha') \neq 0$ for all α_1.

(c) If both are nonzero, then $\widehat{f}(\alpha_1, \alpha') = 0$ for exactly one α_1.

Theorem 8.2 ($\mathbb{F}_4$-Support Theorem)**.** *Let $f : \mathbb{F}_3^n \to \mathbb{F}_2$ be nonzero with $\mathrm{supp}(f) \subseteq T$. Then $|\mathrm{supp}(\widehat{f})| \geq 2^n$.*

Proof. By induction on n. The base case $n = 1$ is verified directly. For the inductive step, let $K_i = \mathrm{supp}(\widehat{f_i})$ with $k_i = |K_i|$. By Lemma 8.1:

$$|\mathrm{supp}(\widehat{f})| = 3|K_1 \triangle K_2| + 2|K_1 \cap K_2| \geq 2\max(k_1, k_2).$$

Since each nonzero f_i satisfies $\mathrm{supp}(f_i) \subseteq T' = (\mathbb{F}_3^*)^{n-1}$, induction gives $k_i \geq 2^{n-1}$, yielding $|\mathrm{supp}(\widehat{f})| \geq 2 \cdot 2^{n-1} = 2^n$. □

Corollary 8.3. $t(2, 3, n) \geq 2^{n-1}$.

Proof. For $f \in C \setminus C_0$, Theorem 8.2 gives $|\mathrm{supp}(\widehat{f})| \geq 2^n$, hence $|\mathrm{supp}(\widehat{f}) \setminus \{0\}| \geq 2^n - 1$. Since each gate covers at most one Frobenius pair, $2w \geq 2^n - 1$, giving $w \geq 2^{n-1}$. □

8.2 Failure for $q \geq 5$

Remark 8.4 (Failure for $q \geq 5$)**.** *The $\mathbb{F}_{16}$-Fourier support theorem does not hold for $q = 5$. Exhaustive computation for $n = 2$ reveals:*

- *The minimum Fourier support for a nonzero $f : \mathbb{F}_5^2 \to \mathbb{F}_2$ with $\mathrm{supp}(f) \subseteq T$ is $|\mathrm{supp}(\widehat{f})| = 8$, not $4^2 = 16$.*
- *The 10 worst-case functions have Hamming weight 8 or 12 and their Fourier support covers exactly 2 of the 4 Frobenius orbits.*
- *Several of these functions are coset indicators of index-2 subgroups of $(\mathbb{F}_5^*)^2 \cong (\mathbb{Z}/4\mathbb{Z})^2$.*

The obstruction is the Vandermonde structure: the 5×4 Vandermonde matrix V over $\mathbb{F}_{16}$ with nodes at the 5th roots of unity has 4×4 submatrices that can be singular (a degree-3 polynomial over $\mathbb{F}_{16}$ can vanish at up to 3 of the 5 nodes). The coordinate slicing induction yields only $|\mathrm{supp}(\widehat{f})| \geq 2 \cdot 4^{n-1}$, a factor of 2 short of the needed 4^n.

This failure motivated the orbit counting argument of Section 5, which sidesteps the Fourier support theorem entirely.

Remark 8.5 (Cohomological interpretation)**.** *The gate complexity admits a geometric restatement. The Fourier support $S = \mathrm{supp}(\widehat{\mathbf{1}_T})$ is closed under the diagonal $\mathbb{F}_q^*$-action, and $t(p, q, n) = |S/\mathbb{F}_q^*|$. When $p \mid (q-1)$, the orbit space is $T/\mathbb{F}_q^* \cong \mathbb{G}_m^{n-1}$ with $(q-1)^{n-1}$ rational points; when $p \nmid (q-1)$, it is $\mathbb{P}^{n-1}$ with $(q^n - 1)/(q-1)$ rational points. In both cases, $t(p, q, n)$ equals the Frobenius trace $\mathrm{Tr}(\mathrm{Frob}_q \mid H^*(S/\mathbb{F}_q^*, \mathbb{F}_p))$ via the Grothendieck–Lefschetz trace formula. The dichotomy arises because the Frobenius eigenvalues $1, q, \ldots, q^{n-1}$ on $H^*(\mathbb{P}^{n-1})$ collapse to 1 in $\mathbb{F}_p$ when $q \equiv 1 \pmod{p}$, making the boundary cohomology invisible.*

9 Depth-3 Circuits: Escaping the Exponential Barrier

The exponential lower bound of Theorem 1.2 applies to depth-2 circuits. A natural question is whether increased depth can circumvent this barrier. We show that depth-3 suffices to reduce the gate complexity from exponential to linear.

9.1 The Depth-3 Construction

Theorem 9.1. *For $n < p$, the torus indicator $\mathbf{1}_T$ can be computed by a depth-3 circuit with $n + 1$ gates.*

Proof. We construct a two-layer circuit:

Layer 1 (n gates): For each coordinate $i \in [n]$, define the gate

$$b_i = g_i(\ell_i(x)) \in \mathbb{F}_p$$

where $\ell_i(x) = x_i$ (the i-th coordinate projection) and $g_i : \mathbb{F}_q \to \mathbb{F}_p$ is defined by

$$g_i(v) = \begin{cases} 1 & \text{if } v \neq 0 \\ 0 & \text{if } v = 0 \end{cases}$$

Thus $b_i = \mathbf{1}[x_i \neq 0] \in \{0, 1\} \subset \mathbb{F}_p$.

Layer 2 (1 gate): The intermediate values $(b_1, \ldots, b_n) \in \mathbb{F}_p^n$ are fed into a single gate

$$h\left(\sum_{i=1}^{n} b_i\right)$$

where $h : \mathbb{F}_p \to \mathbb{F}_p$ is defined by

$$h(s) = \begin{cases} 1 & \text{if } s = n \\ 0 & \text{otherwise} \end{cases}$$

Correctness: We have $x \in T$ if and only if all $b_i = 1$, which occurs if and only if $\sum_i b_i = n$. The condition $n < p$ ensures that the sum $\sum_i b_i$ equals the integer n in $\mathbb{F}_p$ (no wraparound), so $h(\sum_i b_i) = 1$ if and only if all coordinates are nonzero.

The total gate count is $n + 1$. □

Remark 9.2. *The constraint $n < p$ is necessary for the construction. When $n \geq p$, the sum $\sum_i b_i$ can equal $n \mod p$ without all $b_i = 1$, breaking correctness.*

9.2 The Depth Gap

Corollary 9.3 (Exponential Depth Gap)**.** *For fixed p and q with $p \mid (q - 1)$ and $n < p$:*

$$\frac{t_2(p, q, n)}{t_3(p, q, n)} = \frac{(q-1)^{n-1}}{n+1} = \Omega\left(\frac{(q-1)^{n-1}}{n}\right).$$

The depth-2 complexity is exponential while depth-3 is linear—an exponential separation.

9.3 Why the Fourier Method Fails at Depth 3

The lower bound of Section 5 relied on covering the Fourier support with lines. This argument is inherently depth-2: it exploits the fact that each depth-2 gate has Fourier support on a single line.

At depth 3, intermediate values can be combined nonlinearly before the final output. The Fourier support of a depth-3 circuit is no longer constrained to a union of lines—the intermediate layer "mixes" Fourier modes in a way that defeats the covering argument.

Open Problem 9.4. *Determine the depth-3 gate complexity $t_3(p, q, n)$ for $n \geq p$. Our construction requires $n < p$; when $n \geq p$, is $t_3(p, q, n)$ still $O(n)$, or does it grow faster?*

10 Discussion

10.1 The Same-vs-Cross-Characteristic Dichotomy

The gate span completeness theorem (Theorem 3.1) and its failure in same characteristic (Remark following the proof) isolate the algebraic core of why $\mathsf{AC}^0[6]$ is harder than $\mathsf{AC}^0[p]$. In same characteristic ($p = \mathrm{char}(\mathbb{F}_q)$), the DFT is not invertible mod p, nontrivial annihilators exist, and the quotient C/C_0 collapses—so gates cannot even represent all functions on the torus. In cross-characteristic ($p \neq \mathrm{char}(\mathbb{F}_q)$), gates span everything, but the cost is exponential at depth 2.

To be more precise: the Razborov–Smolensky method succeeds for $\mathsf{AC}^0[p]$ because same-characteristic gates have limited span—there exist functions that no linear combination of gates can represent, and the method exploits this algebraic obstruction. For $\mathsf{AC}^0[6]$, the cross-characteristic gates that arise from MOD-2/MOD-3 interactions have full span (Theorem 3.1), so no such obstruction exists. The difficulty shifts from *whether* functions can be represented to *how efficiently*: our results show the cost is exponential at depth 2 but drops to linear at depth 3.

10.2 The Depth-3 Escape and Its Limitations

The exponential depth-2 vs. linear depth-3 separation (Corollary 9.3) shows that the Fourier orbit covering technique is inherently a depth-2 method: it relies on each gate's Fourier support lying on a single $\mathbb{F}_q$-line, a property that breaks at depth 3.

However, for the $\mathsf{AC}^0[6]$ problem, the relevant regime is n growing with p fixed, and our depth-3 construction requires $n < p$. This is a real limitation: when $n \geq p$, the sum $\sum_i b_i$ can wrap around in $\mathbb{F}_p$, and the construction fails. Whether $t_3(p, q, n)$ remains $O(n)$ for $n \geq p$, or whether some intermediate growth rate emerges, is the most natural open question from our work.

More broadly, a full $\mathsf{AC}^0[6]$ circuit has *constant* depth but *unbounded* width, with cross-characteristic interactions composing across multiple layers. Our results show that a single such interaction layer already costs exponentially many gates, and that one additional layer of composition (depth 3) suffices to bypass this cost—at least when $n < p$. Understanding how the cost evolves with further layers of composition is the essential remaining challenge. We note that in the purely linear setting, depth-2 circuits for Kronecker powers already achieve sub-$N^{1.5}$ size [1], so the exponential depth-2 barrier is specific to the cross-characteristic nonlinear model.

10.3 Why the Vandermonde Approach Fails

The Vandermonde induction of Section 8 gives a clean proof for $q = 3$ but fails for $q \geq 5$. The failure is instructive: it stems from the $q \times (q-1)$ Vandermonde matrix over $\mathbb{F}_{(q-1)^2}$ having singular maximal minors when $q \geq 5$. This means the coordinate slicing induction loses a constant factor at each step, accumulating to a gap of 2^{n-1} between the achieved bound and the needed one.

The orbit counting argument of Section 5 sidesteps this entirely by working with Frobenius orbits rather than Fourier support sizes. The lesson for future lower bound arguments is that counting orbits (a group-theoretic quantity) can be more robust than bounding support sizes (an analytic quantity), because orbit counts are insensitive to cancellations in the Vandermonde structure.

10.4 Comparison Across q

	$q = 2$	$q = 3$	$q = 5$	general q
Formula (when $p \mid (q-1)$)	—	2^{n-1}	4^{n-1}	$(q-1)^{n-1}$
Formula (when $p \nmid (q-1)$)	$2^n - 1$	$(3^n-1)/2$	$(5^n-1)/4$	$(q^n-1)/(q-1)$
Growth base	2	2 or 3/2	4 or 5/4	$q-1$ or q
$\lvert T \rvert$	1	2^n	4^n	$(q-1)^n$

The ratio between the two regimes is $(q^n-1)/((q-1)^n) \sim (q/(q-1))^{n-1}$, so for small q the phase transition at $p \mid (q-1)$ is significant: for $q = 3$, the jump from $p = 2$ to $p = 5$ changes the gate complexity from 2^{n-1} to $(3^n-1)/2$, a factor of roughly $(3/2)^{n-1}$.

10.5 Open Problems

1. **Exact depth-3 complexity.** Determine $t_3(p, q, n)$ precisely for all n, including the regime $n \geq p$. When $n \geq p$, does $t_3(p, q, n)$ remain $O(n)$, or does it grow faster? Even a lower bound of $\omega(n)$ for $t_3(p, q, n)$ when $n \geq p$ would be interesting.

2. **Composing layers.** The central barrier for $\mathsf{AC}^0[6]$ is composing cross-characteristic interactions across multiple layers. Can the Fourier-theoretic framework of this paper be extended to track how Fourier support evolves across composed layers, even if the line-covering argument no longer applies?

References

[1] J. Alman and B. Li. Kronecker powers, orthogonal vectors, and the asymptotic spectrum. In *Proc. 66th IEEE FOCS*, 2025.

[2] J. Alman. Kronecker products, low-depth circuits, and matrix rigidity. In *Proc. 53rd ACM STOC*, pages 772–785, 2021.

[3] D. A. M. Barrington, H. Straubing, and D. Thérien. Non-uniform automata over groups. *Information and Computation*, 89(2):109–132, 1990.

[4] J. P. Hansen. Toric varieties, Hirzebruch surfaces and error-correcting codes. *Applicable Algebra in Engineering, Communication and Computing*, 13(4):289–300, 2002.

[5] B. Kivva. Improved upper bounds for the rigidity of Kronecker products. In *Proc. 46th MFCS*, LIPIcs vol. 202, 2021.

[6] A. A. Razborov. Lower bounds on the size of bounded depth circuits over a complete basis with logical addition. *Mathematical Notes*, 41(4):333–338, 1987.

[7] R. Smolensky. Algebraic methods in the theory of lower bounds for Boolean circuit complexity. In *Proc. 19th ACM STOC*, pages 77–82, 1987.

[8] I. Soprunov and J. Soprunova. Toric surface codes and Minkowski length of polygons. *SIAM Journal on Discrete Mathematics*, 23(1):384–400, 2009.

[9] E. Viola. On the power of small-depth computation. *Foundations and Trends in Theoretical Computer Science*, 5(1):1–72, 2009.

[10] R. Williams. Nonuniform ACC circuit lower bounds. *Journal of the ACM*, 61(1):1–32, 2014.

Acknowledgment. Generative AI (Claude, Anthropic) was used for coding assistance and language polishing.