hexagon:2611.00285v1

WATKINS'S CONJECTURE FOR ELLIPTIC CURVES WITH RATIONAL TWO-TORSION

QIYUE TANG

Abstract. We prove that the modular degree of every elliptic curve $E/\mathbb{Q}$ with $E(\mathbb{Q})[2] \neq 0$ is divisible by $2^{\operatorname{rank} E(\mathbb{Q})}$.

1. Introduction

For an elliptic curve $E/\mathbb{Q}$ of conductor N, the modularity theorem gives a nonconstant morphism $X_0(N) \to E$ defined over $\mathbb{Q}$ [8]. Its least possible degree, denoted by m_E, is the modular degree of E. This invariant is related to congruences between modular forms [1]. In 2002, Watkins developed a method for computing it in large families using the Rankin–Selberg formula relating the modular degree to a special value of the symmetric-square L-function. He attributes this formula to Shimura's work, in the form given by Flach [33, Section 1]. The analogy with class-number formulas suggested studying prime divisors of the modular degree. His computations revealed a further dependence of its divisibility by powers of two on the Mordell–Weil rank. These observations led to his conjecture [33, Conjecture 4.1] that

$$2^{\operatorname{rank} E(\mathbb{Q})} \mid m_E.$$

We prove this divisibility for elliptic curves with nontrivial rational two-torsion.

Theorem 1.1. *Let $E/\mathbb{Q}$ be an elliptic curve, let N be its conductor, and put $r = \operatorname{rank} E(\mathbb{Q})$. If $E(\mathbb{Q})[2] \neq 0$, then*

$$2^r \mid \deg(\phi)$$

for every nonconstant $\mathbb{Q}$-morphism $\phi : X_0(N) \to E$. In particular, $2^r \mid m_E$.

The parity of the modular degree was an early focus. Stein and Watkins determined it for the Neumann–Setzer curves in 2004 [32]. Calegari and Emerton subsequently obtained restrictions on the conductor and the residual representation, and proved that the analytic rank is even [10, Theorem 1.1]. Their argument includes a multiplicity-one theorem at two. Yazdani studied the corresponding congruence number using the rational cuspidal subgroup [35]. Kazalicki and Kohen proved that odd congruence number forces Mordell–Weil rank zero, with the corrected statement given in their 2019 corrigendum [23, 24]. Since congruence number and modular degree have the same parity when $4 \nmid N$, this settles the odd-modular-degree case under that hypothesis. More recently, Kazalicki and Slijepčević obtained divisibility by four for positive-rank curves of prime conductor and root number $+1$, proving Watkins's conjecture for the rank-two curves in this class [25].

A deformation-theoretic approach was introduced by Dummigan in 2006 [17]. Dummigan and Krishnamoorthy corrected the original Selmer-group comparison and extended the approach to modular abelian varieties [18, Proposition 8.2]. The original comparison lost one factor of two;

Date: November 10, 2026.

2020 *Mathematics Subject Classification.* 11G05, 11G18, 11G40, 14G35.

Key words and phrases. Elliptic curve, modular degree, rational two-torsion, quadratic twist, Watkins's conjecture.

see the discussion in [2]. The method compares a two-Selmer group with a cotangent space of a deformation ring, and then relates this space to a congruence ideal through a suitable complete-intersection 2-adic $R = T$ theorem. Allen carries out this approach for curves with $E(\mathbb{Q})[2] = 0$, $2 \parallel N$, and odd valuation of the minimal discriminant at every prime of multiplicative reduction [2]. His lower bound incorporates both the two-Selmer rank and the number of prime divisors of the conductor.

For curves with rational two-torsion, quadratic twists provide another approach. In 2021, Esparza-Lozano and Pasten proved that, for each fixed E with $E(\mathbb{Q})[2] \neq 0$, Watkins's conjecture holds for twists by fundamental discriminants with sufficiently many prime divisors [19]. Their effective threshold gives the conjecture for a density-one subset of each such twist family. Caro subsequently proved it for every quadratic twist of a curve with rational two-torsion and prime-power conductor [11]; these families include the congruent-number curves.

Caro and Pasten proved the conjecture for semistable curves with exactly one rational point of order two when the number of primes of non-split multiplicative reduction is odd, or when all bad primes have non-split multiplicative reduction [12]. Bhakta, Krishnamoorthy, and Pasupulati obtained further semistable cases, including curves with full rational two-torsion, under additional hypotheses involving reduction, the two-torsion in the Tate–Shafarevich group, or rank parity [6, Theorem 3.6]. Bhakta and Krishnamoorthy obtained quantitative results for families with rational torsion, including a positive proportion of curves with rational two-torsion ordered by height [5]. Hatley and Kundu studied curves with rational two-torsion and conductor $2^m p$, and established the conjecture for specified prime twists of these curves [21]. Their work treats additive reduction at two together with multiplicative reduction at an odd prime.

Our proof compares two-isogeny descent with quadratic twists having a common Atkin–Lehner character. At odd additive primes, the local descent conditions contain a matrix of Legendre symbols. Its rank decreases the rank bound, while its kernel parametrizes the twist family. Character sums of the corresponding oldforms produce differentials divisible by powers of two. These differentials lie in the cotangent lattice of the same quotient through which the original modular parametrization factors. Pullback under the dual homomorphism supplies a divisor of the remaining map degree. The semistable equality case is completed by specialization at two and a parity calculation with CM fixed points on the quotient.

For $v_2(N) \leq 1$, integrality follows from the good or semistable model of the modular curve. For additive reduction at two, we use the dyadic normalizer and divide morphisms over number fields unramified at two. The integral lifting argument follows the method of Chi, Liu, and Nie [14]. Section 2 proves the number-field lifting and component statements used here, from the unbounded-denominators theorem of Calegari, Dimitrov, and Tang [9], the connectedness of Igusa curves [22], and the local Manin-constant theorem of Česnavičius [13]. It also proves the required oddness of the optimal Manin constant. When every odd additive prime is removable by twisting, an exact Petersson-norm formula completes the comparison. The final step passes from the optimal curve to every target in its rational isogeny class.

1.1. **Notation and normalization.** Write v_p for the valuation with $v_p(p) = 1$, and $\omega(n)$ for the number of distinct prime divisors of n. The identity point of an elliptic curve is denoted by O, and $L(E, s)$ is its Hasse–Weil L-function. For a number field K, write $\mathcal{O}_K$ for its ring of integers. Let $\mathfrak{H} = \{z \in \mathbb{C} : \operatorname{Im} z > 0\}$. Under reduction $\mathrm{SL}_2(\mathbb{Z}) \to \mathrm{SL}_2(\mathbb{Z}/N\mathbb{Z})$, the inverse images of the upper-triangular, upper-unipotent, and trivial subgroups are denoted by $\Gamma_0(N)$, $\Gamma_1(N)$, and $\Gamma(N)$, respectively. We write $\mathcal{E}^0$ for the identity component of a Néron model and

$E^0(K)$ for its points over a local field K. For an odd prime p, put

$$p^* = (-1)^{(p-1)/2}p, \qquad [u/p] = \begin{cases} 0 & (u/p) = 1, \\ 1 & (u/p) = -1. \end{cases}$$

The bracket is used only when $p \nmid u$. The quadratic character attached to a discriminant D is χ_D, and $E^{(D)}$ is the quadratic twist. For an abelian variety $A/\mathbb{Q}$, let $\mathcal{A}$ be its Néron model and e its identity section. Write $\Omega(A) = e^*\Omega^1_{\mathcal{A}/\mathbb{Z}}$ for its lattice of invariant differentials. For a curve Y, we identify $\Omega(J_Y) \otimes \mathbb{Q}$ with $H^0(Y, \Omega^1_Y)$, where $J_Y = \mathrm{Jac}(Y)$. Over a number field the analogous notation is used locally at each finite place. For an elliptic curve $C/\mathbb{Q}$, let Δ_C denote its minimal discriminant and let $c_p(C) = [C(\mathbb{Q}_p) : C^0(\mathbb{Q}_p)]$ be its Tamagawa number.

Put $q = e^{2\pi iz}$. For a cusp form g, write $a_n(g)$ for its coefficient of q^n. Let $f_E = \sum_{n\geq 1} a_n q^n$, $a_1 = 1$, be the newform of E, where $a_n = a_n(f_E)$. The Petersson pairing at weight two is

$$\langle g_1, g_2 \rangle_N = \int_{\Gamma_0(N)\backslash\mathfrak{H}} g_1(z)\overline{g_2(z)}\, dx\, dy, \qquad z = x + iy.$$

We identify a weight-two cusp form f with the differential $f(q)\, dq/q$. A minimal modular parametrization is normalized by $\phi(\infty) = O$, and its Manin constant is defined, up to sign, by

$$\phi^*\omega_E = c_E f_E(q)\frac{dq}{q},$$

where ω_E is a global minimal Néron differential. Changing the sign of ω_E lets us take $c_E > 0$. For a general morphism, its pullback differential determines its multiplier.

Atkin–Lehner involutions are indexed by exact divisors $d \mid N$, that is, $(d, N/d) = 1$. Their group is denoted by W. Its character on f_E is $\varepsilon : W \to \{\pm 1\}$. Write $J_0(N) = \mathrm{Jac}(X_0(N))$. The arithmetic form $X_\mu(N)$ of $X_1(N)$ parametrizes a subgroup with a specified embedding of μ_N; this form has the rational Tate cusp used in the lifting argument.

Lemma 1.2. *It suffices to prove Theorem 1.1 for a positive-rank optimal elliptic quotient of $J_0(N)$, with its optimal parametrization.*

Proof. Rank zero is immediate. A map to E, after translation at ∞, induces a homomorphism $J_0(N) \to E$. The newform quotient with connected kernel is an elliptic curve E_0; the homomorphism factors through an isogeny $E_0 \to E$. The factorization follows either from the connected-kernel quotient or from the multiplicity-one decomposition of $J_0(N)$ at the newform of exact level N. Thus degrees multiply. This also applies to any larger-degree map to E_0.

The existence of rational two-torsion is invariant under rational isogeny. Indeed, a minimum-degree isogeny between two fixed curves has cyclic kernel: a noncyclic finite subgroup of an elliptic curve contains the full $E[\ell]$ for some prime ℓ, and factoring through $[\ell]$ would reduce the degree. An odd isogeny preserves the rational two-torsion. If a cyclic isogeny has even degree, both its kernel and its dual kernel have a unique order-two subgroup, which is rational. In particular, the optimal curve has a nonzero rational point of order two. Ranks are invariant under isogeny. These facts prove the reduction. □

1.2. **Translations and degree on a quotient.** Let w be an automorphism with $w^* f_E = \varepsilon(w) f_E$, where $\varepsilon(w) \in \{1, -1\}$. Equality of differentials gives

$$\phi w = \varepsilon(w)\phi + t_w, \qquad t_w = \phi(w\infty). \tag{1.1}$$

Write $t(w) = t_w$, and $t_d = t_{w_d}$ for an exact divisor d. The restriction of t to $H = \ker \varepsilon$ is a homomorphism. For an involution in H, its image is killed by two. Atkin–Lehner translations

are rational cuspidal torsion by the Manin–Drinfeld theorem [28, 16]. We shall also use the vanishing

$$r > 0 \quad \Longrightarrow \quad L(E,1) = 0 \quad \Longrightarrow \quad t_{w_N} = \phi(0) = O. \tag{1.2}$$

The first implication is the contrapositive of the analytic-rank-zero theorem [26]; the second follows by integrating the newform from $i\infty$ to 0. The vanishing integral gives $\phi(0) = O$.

We record a polarization calculation used in both integral arguments. Suppose that $\phi = \psi\pi$, where $\pi : X_0(N) \to Y$ is a quotient, and let $i : E \to J_Y$ be the dual of $\psi_* : J_Y \to E$. Then

$$\psi_* i = [\deg \psi]. \tag{1.3}$$

If f_Y denotes the descent of $f_E(q)dq/q$, then $\psi^*\omega_E = c_E f_Y$, and hence

$$i^*\left(\frac{f_Y + g_Y}{2^b}\right) = \frac{\deg \psi}{c_E 2^b}\,\omega_E \tag{1.4}$$

for $b \in \mathbb{Z}$ and a holomorphic differential g_Y orthogonal to f_Y for the Petersson, or equivalently the polarization, pairing. The vanishing on g_Y is the adjointness of pullback and trace. Pullback from Y multiplies this pairing by $\deg \pi$, so orthogonality proved on $X_0(N)$ descends to Y. If the differential on the left belongs to the Néron cotangent lattice of J_Y, then the coefficient on the right is in $\mathbb{Z}_2$, because homomorphisms extend to Néron models. When c_E is odd this implies $v_2(\deg \psi) \geq b$.

2. Integral lifts and identity components

We use the lifting method of Chi–Liu–Nie [14, Sections 3–5] for morphisms from $X_0(N)$ over number fields. The passage from integral expansions to a congruence subgroup uses the unbounded-denominators theorem of Calegari–Dimitrov–Tang [9].

For an elliptic curve A over a nonarchimedean local field L, write $A^0(L)$ for the points reducing into the identity component of its Néron model, and $A^1(L)$ for the kernel of reduction. A minimal Weierstrass equation has parameter $t = -x/y$ at the origin. Its normalized invariant differential has formal logarithm $L_A(t) = t + O(t^2)$.

Lemma 2.1. *Let R be the ring of integers of a finite unramified extension of $\mathbb{Q}_p$, and let $A/\operatorname{Frac}(R)$ have additive reduction. Then*

$$L_A(t),\ L_A^{-1}(t) \in tR[[t]].$$

In particular $A^1(\operatorname{Frac}(R))$ is torsion-free. If $A/\mathbb{Q}_p$ has additive reduction, then

$$\#A^0(\mathbb{Q}_p)[p^\infty] \leq p.$$

For an arbitrary finite extension $L/\mathbb{Q}_p$, additive reduction implies that $A^0(L)$ has no nonzero torsion of order prime to p.

Proof. The identity component of the special fibre is the additive group. Hence multiplication by p on its formal group is zero. Since p is a uniformizer of R, the formal multiplication law has the form

$$[p](t) = pg(t), \qquad g(t) = t + \sum_{n\geq 2} g_n t^n \in tR[[t]].$$

Write $L_A(t) = t + \sum_{n\geq 2} b_n t^n$. The identity $L_A([p](t)) = pL_A(t)$, divided by p and compared in degree $n \geq 2$, gives

$$(1 - p^{n-1})b_n = g_n + \sum_{k=2}^{n-1} p^{k-1} b_k [t^n] g(t)^k.$$

The coefficient on the left is a unit. Induction proves $b_n \in R$. The inverse of a series in $tR[[t]]$ with linear coefficient one again belongs to $tR[[t]]$. Both series converge on pR and give mutually

inverse group isomorphisms between $A^1(\mathrm{Frac}(R))$ and the additive group pR. This proves torsion-freeness.

For $L = \mathbb{Q}_p$, smooth reduction gives

$$0 \longrightarrow A^1(\mathbb{Q}_p) \longrightarrow A^0(\mathbb{Q}_p) \longrightarrow (\mathbb{F}_p, +) \longrightarrow 0.$$

The torsion subgroup therefore injects into $(\mathbb{F}_p, +)$. For a general finite extension $L/\mathbb{Q}_p$, multiplication by an integer prime to p is an automorphism on the formal group, whereas the smooth reduction quotient is the additive group of the residue field. Neither group has nonzero prime-to-p torsion, which proves the last assertion. □

2.1. **The arithmetic curve and a congruence lift.** Let $X_\mu(N)$ be the smooth projective curve over $\mathbb{Q}$ compactifying the coarse moduli curve $Y_\mu(N)$ of pairs

$$(A, \iota), \qquad \iota : \mu_N \hookrightarrow A[N].$$

Its complex uniformization is by $\Gamma_1^{\pm}(N) = \langle \Gamma_1(N), -I \rangle$. The cusp ∞ is given by the Tate curve together with its canonical inclusion of μ_N; it is rational and has parameter q. For $\delta \in (\mathbb{Z}/N\mathbb{Z})^\times$, the diamond operator is

$$\langle \delta \rangle (A, \iota) = (A, \iota \circ [\delta]).$$

It is defined over $\mathbb{Q}$, and $\langle -1 \rangle$ acts trivially. Forgetting the embedding and retaining its image defines $\varpi_N : X_\mu(N) \to X_0(N)$ and identifies

$$X_\mu(N)/((\mathbb{Z}/N\mathbb{Z})^\times/\langle -1 \rangle) \simeq X_0(N).$$

These conventions, including the arithmetic choice of cusp, are those of the inclusion moduli problem in [22, Sections 4.9 and 8.6].

We shall also use the integral Tate neighbourhood of ∞ on $X_0(N)$. Over the integer ring R of any local field its completion is $R[[q]]$, and the curve is smooth along this cusp. Here the Tate parameter is as in [22, Theorem 8.11.10], and smoothness of the cusp section on the integral model is recorded in [13, Remark 5.3]. Consequently, if $F : X_0(N)_K \to B$ is a morphism over a number field and $F(\infty) = O$, the Néron mapping property gives

$$t_B(F(q)) \in q\mathcal{O}_{K_v}[[q]] \tag{2.1}$$

at every finite place v, for a parameter from a minimal local equation of B, by the smooth Tate neighbourhood of the cusp.

Lemma 2.2. *Let K be a number field, let $\nu : A \to B$ be a K-isogeny, and let $F : X_0(N)_K \to B$ satisfy $F(\infty) = O$. Fix an embedding $K \hookrightarrow \mathbb{C}$, and let $\psi : \mathfrak{H} \to A(\mathbb{C})$ be the lift of F through ν which tends to O at ∞. Suppose that, for every finite place v of K, the formal expansion of this lift in a minimal parameter of A belongs to $q\mathcal{O}_{K_v}[[q]]$. Then ψ induces a K-morphism*

$$\psi_\mu : X_\mu(N)_K \longrightarrow A, \qquad \psi_\mu(\infty) = O, \qquad \nu \circ \psi_\mu = F \circ \varpi_N.$$

For every diamond operator there is a point $\tau_\delta \in \ker \nu(K)$ such that

$$\psi_\mu \circ \langle \delta \rangle = \psi_\mu + \tau_\delta. \tag{2.2}$$

Proof. The result is immediate for a constant F, so suppose that F is nonconstant. The isogeny is an analytic covering and $\mathfrak{H}$ is simply connected. Its lift is determined by the branch taking the value O over the cusp. Equivalently, it is obtained by inverting the formal isogeny at O in characteristic zero. This also shows that its formal coefficients belong to K. The two lifts $\psi(z+1)$ and $\psi(z)$ have the same limit at the cusp, so they are equal.

Pass to the Hilbert class field H of K. The principal ideal theorem makes every ideal class of K principal in H [29]. In particular it principalizes the Weierstrass ideal class of A. Minimal equations remain minimal under unramified extensions, and the criterion for a global minimal

equation in [31, Chapter VIII, Proposition 8.2] therefore gives a minimal Weierstrass equation for A_H over $\mathcal{O}_H$. The hypothesis on formal expansions persists under this unramified extension and under integral changes between minimal equations.

On this equation put

$$t = -\frac{x}{y}, \qquad w = -\frac{1}{y}.$$

Both functions are regular at O. If the coefficients of the equation are $a_i \in \mathcal{O}_H$, they satisfy

$$w = t^3 + a_1 tw + a_2 t^2 w + a_3 w^2 + a_4 t w^2 + a_6 w^3.$$

The formal implicit-function recursion yields $w \in t^3 \mathcal{O}_H[[t]]$. Thus both $t(\psi(q))$ and $w(\psi(q))$ lie in $\mathcal{O}_H[[q]]$. The functions t, w generate the function field of A, because $x = t/w$ and $y = -1/w$.

For $\gamma \in \Gamma_0(N)$, the difference

$$\rho(\gamma) = \psi(\gamma z) - \psi(z)$$

is independent of z and belongs to $\ker \nu(\mathbb{C})$. It defines a homomorphism, and $G = \ker \rho$ has finite index in $\Gamma_0(N)$. The compact curve associated to G is the component, selected by (∞, O), of the finite étale pullback

$$X_0(N) \times_{B,\nu} A.$$

In particular $t \circ \psi$ and $w \circ \psi$ are meromorphic modular functions for G, algebraic over the function field of $X_0(N)$.

For either function r, its noncuspidal poles have algebraic j-values. Choose a nonzero polynomial $P_r(j) \in \mathcal{O}_H[j]$ whose zeros cancel all these poles, taking their conjugates and increasing multiplicities if necessary. A sufficiently high power of the discriminant modular form Δ then cancels the remaining poles at all cusps. Hence

$$\Delta^{m_r} P_r(j) r$$

is a holomorphic modular form for G. Its Fourier coefficients have bounded denominators: r has integral coefficients and j, Δ have integral Laurent expansions. By [9, Theorem 1.0.1 and Remark 6.3.1], this modular form is congruence. Dividing by its level-one factors shows that each of $t \circ \psi, w \circ \psi$ is a congruence modular function. Their joint stabilizer inside $\Gamma_0(N)$ is exactly G, since t, w separate the generic points of A. Thus G contains a principal congruence subgroup.

The compact pullback is étale, so the monodromy around each cusp and each elliptic point is trivial. Consequently the cusp widths for G are those of their images on $X_0(N)$, and their least common multiple is N. Moreover $-I \in G$. Wohlfahrt's theorem [34] gives $\Gamma(N) \subseteq G$. The width-one translation $T = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ also belongs to G. Since the image of $\Gamma_1(N)$ modulo N is generated by the image of T, we obtain $\Gamma_1^{\pm}(N) \subseteq G$. This proves the asserted descent over $\mathbb{C}$, including extension across the cusps.

The resulting graph is the unique geometric component through (∞, O) of

$$X_\mu(N)_K \times_{F \circ \varpi_N, B, \nu} A,$$

and its projection to $X_\mu(N)_K$ has degree one. The pullback and this point are defined over K. Galois therefore fixes the component, so its degree-one projection and the resulting morphism descend to K. Finally, applying ν to the difference in (2.2) gives zero; connectedness makes the difference a constant in $\ker \nu$. Evaluation at ∞ shows that it is K-rational, since all diamond translates of this cusp are rational. □

2.2. **A connected source at every prime of the level.** The next lemma extends the component argument of [14, Section 5] to finite extensions of $\mathbb{Q}_p$.

Lemma 2.3. *Let $L/\mathbb{Q}_p$ be finite, let $p \mid N$, and let A/L be an elliptic curve. Suppose that a morphism $\psi : X_\mu(N)_L \to A$ satisfies*

$$\psi \circ \langle\delta\rangle = \psi + \tau_\delta, \qquad \tau_\delta \in A(L),$$

for every $\delta \in (\mathbb{Z}/N\mathbb{Z})^\times$. Then $\tau_\delta \in A^0(L)$ for every δ.

Proof. Write $N = p^a M$ with $p \nmid M$. Choose $n \geq 3$ prime to p and divisible by M, and choose a primitive nth root of unity ζ. We first construct a smooth auxiliary source over $R_0 = \mathbb{Z}_p[\zeta]$. Let B classify triples (C, P, Q) in which C is an elliptic curve and (P, Q) is a basis of $C[n]$ satisfying $e_n(P, Q) = \zeta$. Full level $n \geq 3$ is a fine moduli problem, and fixing this Weil pairing selects an open and closed component of the smooth full-level curve [22, Corollary 4.7.2; 9.1.5, 9.1.6 and 9.4.3.1]. Thus B/R_0 is smooth and carries a universal elliptic curve. Remove its supersingular special-fibre points, obtaining an open U with entire generic fibre and ordinary special fibre [22, Lemmas 12.5.2 and 12.5.4].

Over U consider the functor of quadruples

$$(C, P, Q, \iota_{p^a}), \qquad \iota_{p^a} : \mu_{p^a} \hookrightarrow C[p^a].$$

By the inclusion representability theorem [22, Section 4.9], this functor is represented by an affine étale U-scheme S. It is of finite presentation, so S/R_0 is smooth of finite type. On its generic fibre the first basis vector defines

$$\iota_M : \mu_M \hookrightarrow C[M], \qquad \iota_M(\zeta^{n/M}) = [n/M]P.$$

When $M = 1$ this is the trivial structure. Multiplication identifies $\mu_M \times \mu_{p^a}$ with μ_N. Under this identification $(x, y) \mapsto \iota_M(x) + \iota_{p^a}(y)$ is an inclusion $\iota_N : \mu_N \hookrightarrow C[N]$, since the two image groups have coprime orders. It defines a generic-fibre morphism

$$j : S_{\mathrm{Frac}(R_0)} \longrightarrow Y_\mu(N).$$

Each diamond has an individual lift to S. Given δ modulo N, choose $v \in (\mathbb{Z}/n\mathbb{Z})^\times$ reducing to δ modulo M; this is possible because reduction on unit groups is surjective. Set

$$\widetilde{\delta}(C, P, Q, \iota_{p^a}) = (C, [v]P, [v^{-1}]Q, \iota_{p^a} \circ [\delta]). \tag{2.3}$$

The pairing is unchanged, and this formula defines an R_0-automorphism with $j\widetilde{\delta} = \langle\delta\rangle j$.

We verify that the geometric special fibre of S is nonempty and connected. Let k be an algebraic closure of the residue field of R_0. Write $F^a : C \to C^{(p^a)}$ for the iterated relative Frobenius and $V^a : C^{(p^a)} \to C$ for its dual Verschiebung. On the ordinary universal elliptic curve over U_k there is an exact sequence

$$0 \longrightarrow \ker F^a \longrightarrow C[p^a] \xrightarrow{F^a} \ker V^a \longrightarrow 0,$$

where $\ker V^a$ is étale and locally cyclic of order p^a [22, Lemmas 12.2.1, 12.2.3, 12.2.5 and Proposition 12.3.6]. Every homomorphism from μ_{p^a} to $\ker V^a$ is zero: it is zero on geometric fibres, and its vanishing locus is open and closed [22, Lemma 4.10.0(3)]. Therefore an inclusion ι_{p^a} factors through $\ker F^a$ and is an isomorphism onto it. Cartier duality and the canonical principal polarizations identify $(\ker F^a)^D$ with $\ker V^a$ [22, (2.8.2.1)]. The map $\iota_{p^a} \mapsto (\iota_{p^a}^D)^{-1}$ consequently identifies our inclusions with isomorphisms

$$(\mathbb{Z}/p^a\mathbb{Z}) \xrightarrow{\sim} \ker V^a.$$

This is precisely an Igusa generator on the ordinary locus.

The Igusa problem over B_k assigns a section $Z \in C^{(p^a)}$ such that $V^a Z = 0$ and

$$\sum_{i=1}^{p^a} [iZ] = \ker V^a$$

as effective Cartier divisors. This definition makes sense at supersingular points as well [22, Definition 12.3.1]. Let C_a be the scheme classifying the simultaneous full-level and Igusa structures. The product construction of moduli problems and relative representability give a finite flat morphism

$$C_a \longrightarrow B_k \quad \text{of degree } p^{a-1}(p-1)$$

[22, 4.3.4 and Theorem 12.6.1(1)]. The preceding Cartier-dual description, being functorial in the base, proves

$$S_k \simeq C_a \times_{B_k} U_k.$$

The fixed-pairing full-level problem is representable, finite étale, and surjective over the moduli problem of elliptic curves over k: an n-torsion basis exists étale locally, and rescaling one basis vector gives the prescribed Weil pairing. The fixed-pairing full-level curve B_k is connected by connectedness of its smooth compactification [22, Corollary 10.13.12]. Katz–Mazur's Igusa theorem gives connectedness of C_a and its smoothness [22, Corollary 12.6.2(2) and (1)]. A connected smooth curve over k is irreducible. Since the supersingular j-values form a finite set and $C_a \to B_k$ is finite flat of positive degree, its ordinary open is nonempty. It is therefore connected, proving the claim about S_k.

Now put $L' = L(\zeta)$. This is unramified over L. Base changing the construction to $\mathcal{O}_{L'}$ gives a smooth scheme with nonempty geometrically connected special fibre. Let $\mathcal{A}'$ be the Néron model of $A_{L'}$ over $\mathcal{O}_{L'}$. The morphism ψj extends uniquely to $h : S_{\mathcal{O}_{L'}} \to \mathcal{A}'$, and τ_δ extends to a section $\widetilde{\tau}_\delta$. The Néron mapping property and (2.3) give

$$h\widetilde{\delta} = h + \widetilde{\tau}_\delta.$$

The image of the connected geometric special fibre of S lies in a single component of $\mathcal{A}'_k$. For any geometric special-fibre point s, the points $h(s)$ and $h(\widetilde{\delta}s)$ thus have the same component class. Their difference is $\widetilde{\tau}_\delta$, whose component class is zero. Hence $\tau_\delta \in A^0(L')$. Néron models commute with unramified base change [7, Section 1.2, Proposition 2(c)], and their identity components descend under this extension. Thus $\tau_\delta \in A^0(L)$. □

We shall also use the following consequence of the ordinary modular charts.

Lemma 2.4. *Let $p \mid N$, let $L/\mathbb{Q}_p$ be finite, and let $P : X_0(N)_L \to A$ be a morphism to an elliptic curve, with $P(\infty) = O$. If $d \mid N$ is an exact divisor and $p \nmid d$, then*

$$P(w_d \infty) \in A^0(L).$$

Proof. Write $N = p^a M$, with $p \nmid M$. Choose $n \geq 3$ such that $M \mid n$ and $p \nmid n$, and put $L' = L(\zeta_n)$, $R = \mathcal{O}_{L'}$. The extension L'/L is unramified. Let B/R be the fine full-level-n modular curve with Weil pairing fixed to ζ_n. Its compactification is smooth with geometrically connected fibers [22, Corollary 10.13.12]. Let X/R be the flat Katz–Mazur compactification obtained by adding a cyclic subgroup of order p^a. On the generic fiber, retain this subgroup together with the order-M subgroup generated by (n/M) times the first auxiliary basis vector. Forgetting the rest of the basis gives a morphism

$$j : X_{L'} \longrightarrow X_0(N)_{L'}.$$

Over an algebraic closure k of the residue field, the noncuspidal block of X_k indexed by $(a, 0)$ parametrizes the subgroup $\ker(F_C^a)$ of the universal elliptic curve C. Its map to B_k is finite flat of degree one [22, Proposition 13.4.4], and hence is an isomorphism. This block is reduced,

has multiplicity one, and meets the other noncuspidal blocks only at supersingular points [22, Theorem 13.4.7]. Thus X is relatively smooth along its ordinary part: the geometric fiber there is smooth and X is flat over R.

The closure of this ordinary component also contains every cusp whose p-primary subgroup is μ_{p^a}. Such a cusp has a full-level Tate chart with elliptic curve $\mathrm{Tate}(u^n)$ and completed local ring $R[[u]]$; adding the canonical subgroup μ_{p^a} leaves this ring unchanged [22, Section 8.11, especially Theorem 8.11.10]. These cusp sections therefore lie in the relative smooth locus. There is one, denoted by c_∞, above ∞. The exact-divisor cusp $w_d\infty$ has denominator N/d. Since $p^a \mid N/d$, the matrix

$$\begin{pmatrix} 1 & 0 \\ N/d & 1 \end{pmatrix} \in \Gamma_0(p^a)$$

sends ∞ to this cusp class and preserves the connected p-primary branch. Explicitly, in multiplicative coordinates on $\mathrm{Tate}(u^n)$, the two full-level bases are

$$(\zeta_n, u), \qquad (\zeta_n u^{N/d}, u).$$

They have the same Weil pairing. The second basis retains an order-M subgroup generated by $\zeta_n^{n/M} u^{nN/(Md)}$, together with the canonical subgroup μ_{p^a}. Both charts have ring $R[[u]]$. The second therefore gives a section c_d above $w_d\infty$ on the same component.

Remove the other special-fiber components and the nonsmooth points of this component. The resulting open $U \subset X$ is smooth over R, contains c_∞ and c_d, and has geometrically connected special fiber, since this fiber is a nonempty open of the irreducible $(a, 0)$ component. The morphism $P \circ j$ extends from $U_{L'}$ to the Néron model of $A_{L'}$. Its special-fiber image lies in one connected component, which is the identity component because c_∞ maps to O. Hence c_d maps into that component, and $P(w_d\infty) \in A^0(L')$. Néron models and their identity components commute with unramified base change, so this membership descends to L. □

2.3. **Optimal quotients and the Manin constant.**

Proposition 2.5. *Let $\pi : J_0(N) \to E$ be an optimal elliptic quotient over $\mathbb{Q}$. Suppose that $\nu : A \to E$ is a $\mathbb{Q}$-isogeny satisfying*

$$\nu^* \omega_E = \omega_A$$

for minimal global invariant differentials. Then

$$\ker \nu(\overline{\mathbb{Q}}) \subseteq A(\mathbb{Q}) \cap \bigcap_{p \mid N} A^0(\mathbb{Q}_p).$$

In particular, if E has an odd additive prime, there is no such isogeny of even degree.

Proof. Normalize the modular parametrization $\phi = \pi \circ \mathrm{AJ}_\infty$, where $\mathrm{AJ}_\infty(P) = [P] - [\infty]$. At every prime the isogeny extends to Néron models, and its formal expansion in minimal parameters has integral coefficients and linear coefficient one. Its compositional inverse is therefore integral. Equation (2.1) shows that the normalized analytic lift of ϕ through ν satisfies the hypothesis of Lemma 2.2. It gives a rational morphism on $X_\mu(N)$ with rational diamond translations in $\ker \nu$. Lemma 2.3 puts all these translations in $A^0(\mathbb{Q}_p)$ for each $p \mid N$.

It remains to see that the translations exhaust the geometric kernel. The connected kernel of π gives a surjection on integral homology

$$H_1(J_0(N)(\mathbb{C}), \mathbb{Z}) \longrightarrow H_1(E(\mathbb{C}), \mathbb{Z}).$$

Indeed, in the exact sequence of complex tori the quotient lattice is the target lattice precisely when the kernel is connected. The Abel–Jacobi map identifies $H_1(X_0(N)(\mathbb{C}), \mathbb{Z})$ with the homology of its Jacobian, so ϕ_* is also surjective. Using the compatible differentials, write

$$A(\mathbb{C}) = \mathbb{C}/\Lambda_A, \qquad E(\mathbb{C}) = \mathbb{C}/\Lambda_E, \qquad \Lambda_A \subseteq \Lambda_E.$$

The monodromy of the lift is reduction of the periods of ϕ modulo Λ_A. Surjectivity on homology makes its image all of $\Lambda_E/\Lambda_A = \ker \nu(\mathbb{C})$. By Lemma 2.2, this monodromy factors through $\Gamma_0(N)/\Gamma_1^{\pm}(N)$, so each kernel point is one of the diamond translations. This proves the inclusion.

Isogenous elliptic curves have the same conductor and reduction category at each prime. If an odd prime p is additive for E, it is additive for A. An even-order geometric kernel has an element of order two, and the inclusion just proved would put it in $A^0(\mathbb{Q}_p)$. This contradicts Lemma 2.1. □

Proposition 2.6. *Let $E/\mathbb{Q}$ have conductor N, and let $\phi : X_0(N) \to E$ be an optimal modular parametrization normalized at ∞. Write*

$$\phi^*\omega_E = c_E f_E(q)\frac{dq}{q}$$

for a minimal global invariant differential. Then c_E is an odd integer.

Proof. The leading coefficient of the integral cusp parameter $t_E(\phi(q))$ is c_E. Thus $c_E \in \mathbb{Z}$; changing the sign of ω_E makes it positive. The local Manin-constant theorem [13, Theorem 2.11] gives

$$p^2 \nmid N \quad \Longrightarrow \quad p \nmid c_E.$$

Suppose that $2 \mid c_E$. Then $4 \mid N$, so E has additive reduction at 2, and the local newform relation gives $a_{2n} = 0$ for every n. In particular

$$\ell_E(q) = \sum_{n\geq 1} \frac{a_n}{n} q^n = \sum_{n \text{ odd}} \frac{a_n}{n} q^n \in q\mathbb{Z}_2[[q]].$$

Consider the normalized lift of ϕ through $[2] : E \to E$. At each odd prime the formal inverse of [2] has integral coefficients, so this lift is integral by (2.1). At 2, Lemma 2.1 and integration with zero constant term give

$$L_E(t_E(\phi(q))) = c_E\ell_E(q), \qquad t_E(\psi(q)) = L_E^{-1}((c_E/2)\ell_E(q)) \in q\mathbb{Z}_2[[q]].$$

Lemma 2.2 therefore applies. Surjectivity on homology, as in Proposition 2.5, says that the monodromy of this lift is all of $E[2](\mathbb{C})$. Consequently the four points of $E[2]$ are rational diamond translations. Lemma 2.3 at 2 places them all in $E^0(\mathbb{Q}_2)[2]$. Lemma 2.1 bounds this group by two elements, a contradiction. Hence $2 \nmid c_E$. □

2.4. **Division of an arbitrary morphism.**

Proposition 2.7. *Let K be a number field unramified at 2, and let A/K be an elliptic curve with additive reduction at every place above 2. Suppose that A also has additive reduction at some place w of odd residue characteristic p, where $p \mid N$. Let*

$$F : X_0(N)_K \longrightarrow A, \qquad F(\infty) = O,$$

be a morphism. For each $v \mid 2$, choose a minimal invariant differential ω_v and write

$$F^*\omega_v = \sum_{n\geq 1} b_{n,v} q^n \frac{dq}{q}.$$

If, for an integer $h \geq 0$,

$$b_{2n,v} = 0 \quad \text{and} \quad b_{n,v} \in 2^h\mathcal{O}_{K_v} \qquad (n \geq 1,\ v \mid 2),$$

then there is a unique K-morphism $F_0 : X_0(N)_K \to A$ with

$$F_0(\infty) = O, \qquad F = [2^h] \circ F_0.$$

Proof. The assertions are immediate for $h = 0$ or for constant F, so assume otherwise. Take the normalized analytic lift of F through $[2^h]$. At places of odd residue characteristic its formal expansion is integral, since the formal multiplication law has unit linear coefficient and (2.1) applies.

For $v \mid 2$, the unramified hypothesis and additive reduction allow Lemma 2.1 over $\mathcal{O}_{K_v}$. In the minimal parameter corresponding to ω_v, integration of the displayed differential gives

$$L_A(t_A(F(q))) = \sum_{n \text{ odd}} \frac{b_{n,v}}{n} q^n \in 2^h q \mathcal{O}_{K_v}[[q]].$$

There is no integration constant because $F(\infty) = O$. Thus

$$t_A(\psi(q)) = L_A^{-1}\left(2^{-h} \sum_{n \text{ odd}} \frac{b_{n,v}}{n} q^n\right) \in q\mathcal{O}_{K_v}[[q]].$$

Lemma 2.2 now yields a K-morphism $\psi_\mu : X_\mu(N)_K \to A$ and translations $\tau_\delta \in A[2^h](K)$. Applying Lemma 2.3 over K_w puts each translation in $A^0(K_w)$. This group has no nonzero 2-primary torsion, by the odd additive hypothesis and Lemma 2.1. Therefore every τ_δ vanishes. The morphism ψ_μ is invariant under all diamonds and descends over K to the quotient $X_0(N)$, giving F_0 with the required normalization and equality. Finally, the difference between any two such morphisms takes values in the finite group $A[2^h]$, hence is constant; its value at ∞ is zero. □

3. Two-isogeny descent and integral isogeny directions

The local conditions in two-isogeny descent will also index the twist construction. All vector spaces in this section are over $\mathbb{F}_2$.

3.1. **The two descent spaces.** Let T be a nonzero rational point of order two. Completing the square, and translating T to the origin, gives equations

$$E : y^2 = x(x^2 + Ax + B), \qquad E' : y^2 = x(x^2 - 2Ax + D), \qquad D = A^2 - 4B. \tag{3.1}$$

Here $\eta : E \longrightarrow E'$ has kernel $\langle T \rangle$, and $\widehat{\eta} : E' \longrightarrow E$ is its dual. We choose $A, B \in \mathbb{Z}$ so that the first equation is minimal at every odd prime. Such a choice is obtained from a global minimal equation by a change of variables whose only possible denominator is a power of two. The second equation is also minimal at odd primes: a further odd scaling would give an odd prime factor in the minimal differential multiplier of a degree-two isogeny, which is impossible because the multiplier and that of the dual have product 2.

The descent maps are

$$\alpha_B : E(\mathbb{Q}) \longrightarrow \mathbb{Q}^\times/\mathbb{Q}^{\times 2}, \qquad \alpha_D : E'(\mathbb{Q}) \longrightarrow \mathbb{Q}^\times/\mathbb{Q}^{\times 2},$$

where $\alpha_B(O) = 1$, $\alpha_B(T) = B$, and $\alpha_B(x, y) = x$ for $x \neq 0$, and likewise with D on E'. Their kernels are $\widehat{\eta}E'(\mathbb{Q})$ and $\eta E(\mathbb{Q})$, respectively. These assertions follow either from the explicit two-isogeny formulas, or from the homogeneous equation

$$du^4 + Au^2v^2 + (B/d)v^4 = w^2$$

for the class d on the E side; see also [31, Chapter X]. For each place v, let $L_{B,v}$ and $L_{D,v}$ be the images of the corresponding maps on $E(\mathbb{Q}_v)$ and $E'(\mathbb{Q}_v)$. Set

$$\mathcal{S}_B = \{s \in \mathbb{Q}^\times/\mathbb{Q}^{\times 2} : s_v \in L_{B,v} \text{ for every } v\},$$

and define $\mathcal{S}_D$ in the same way using $L_{D,v}$. Here s_v is the image of s in $\mathbb{Q}_v^\times/\mathbb{Q}_v^{\times 2}$, with $\mathbb{Q}_\infty = \mathbb{R}$. The global Kummer images lie in these spaces.

Lemma 3.1. *If $r = \operatorname{rank} E(\mathbb{Q})$, then*

$$\dim \alpha_B(E(\mathbb{Q})) + \dim \alpha_D(E'(\mathbb{Q})) = r + 2. \tag{3.2}$$

Consequently $r + 2 \le \dim \mathcal{S}_B + \dim \mathcal{S}_D$.

Proof. Put $t = \dim E(\mathbb{Q})[2]$. The exact sequence

$$0 \longrightarrow \frac{E'(\mathbb{Q})[\widehat{\eta}]}{\eta(E(\mathbb{Q})[2])} \longrightarrow \frac{E'(\mathbb{Q})}{\eta E(\mathbb{Q})} \xrightarrow{\widehat{\eta}} \frac{E(\mathbb{Q})}{2E(\mathbb{Q})} \longrightarrow \frac{E(\mathbb{Q})}{\widehat{\eta}E'(\mathbb{Q})} \longrightarrow 0$$

has first term of dimension $2 - t$ and middle term $E(\mathbb{Q})/2E(\mathbb{Q})$ of dimension $r + t$. Taking dimensions proves (3.2). □

Let a and m denote the numbers of odd additive and odd multiplicative primes, and put $f = v_2(N)$. Write S for the set of odd additive primes and Q for the set of odd multiplicative primes. The odd primes dividing both B and D are precisely S, and the odd primes dividing at least one are precisely $S \cup Q$. Indeed, at an odd additive prime the identity component has no two-torsion, so T is singular and $p \mid A, B$. Conversely, $p \mid B, D$ implies $p \mid A$, hence additive reduction. A bad nonadditive prime divides exactly one of B, D.

The ambient squareclass space on the B side is generated by -1 and the primes dividing B; the analogous assertion holds on the D side. For example, if $q \nmid B$ and $v_q(x) \neq 0$, comparison of valuations in (3.1) shows that $v_q(x)$ is even. The same argument at a local point proves the assertion for the Selmer space. In particular the total number of odd-prime coordinates in the two spaces is $2a + m$, in addition to their two separate sign coordinates.

We use the matrix and row

$$K = \big([q/p]\big)_{q \in Q,\ p \in S}, \qquad \zeta = \big([2/p]\big)_{p \in S}, \qquad k = \operatorname{rank} K. \tag{3.3}$$

The rows of K always refer to *odd* multiplicative primes. Define $K^* = K$ when f is even and $K^* = \left(\begin{smallmatrix} K \\ \zeta \end{smallmatrix}\right)$ when f is odd, and put $k^* = \operatorname{rank} K^*$. Thus $k \le k^* \le k + 1$. In particular, when $f = 1$, K^* is the matrix obtained by including the multiplicative prime 2.

3.2. **Independent local conditions.** At an odd additive prime p, the local squareclass group $\mathbb{Q}_p^\times/\mathbb{Q}_p^{\times 2}$ has dimension two, with valuation and unit Legendre coordinates. The local index formula gives

$$|L_{B,p}| = 2\frac{c_p(E)}{c_p(E')}, \qquad |L_{D,p}| = 2\frac{c_p(E')}{c_p(E)},$$

and therefore

$$\dim L_{B,p} + \dim L_{D,p} = 2. \tag{3.4}$$

To obtain this formula, compare the component groups, the smooth identity reductions, and the formal subgroups under the two isogenies. On the formal subgroups both maps are isomorphisms at odd p, their rational kernels have order two, and the smooth identity reductions of the two additive curves both have p points.

For a local squareclass s, choose a representative $p^e u$ with $e \in \{0, 1\}$ and $u \in \mathbb{Z}_p^\times$, and set $v_p(s) = e \in \mathbb{F}_2$ and $\operatorname{unit}_p(s) = [u/p]$. Thus valuations of squareclasses in the following linear equations are taken modulo two. Call p *regular* if both local images are ramified lines. Each such image is cut out, for a suitable $b_p \in \mathbb{F}_2$ depending on the image, by an equation

$$\operatorname{unit}_p(s) + b_p v_p(s) = 0. \tag{3.5}$$

Thus, in either global block, its coefficient at a generator $q \neq p$ is $[q/p]$, and its sign coefficient is $[-1/p]$. If p is not regular, one local image is contained in the unit line. Otherwise both would

contain a ramified vector, and (3.4) would make both ramified lines. Consequently, at every nonregular prime we have a pure valuation condition $v_p(s) = 0$ in at least one of the two blocks.

Suppose there are e nonregular primes, and let K_R be the restriction of K to the regular columns. Project the regular local rows onto the multiplicative-prime coordinates. Splitting these coordinates between the B and D blocks shows that their combined rank is at least $\operatorname{rank} K_R$. Choose this many rows with independent projections. The e pure valuation rows are independent of them and of each other: their projections are zero and their nonzero odd-prime coordinates are distinct. Since deleting e columns decreases rank by at most e, these rows give at least k independent conditions.

There is one further, independent, real condition. At least one of the real Kummer images consists of positive classes. Indeed, if $B < 0$, then $D > 0$ and $x^2 - 2Ax + D$ is everywhere positive; if $D < 0$, then $x^2 + Ax + B$ is everywhere positive. If $B, D > 0$, the real quadratic roots show that positivity holds on the D side when $A > 0$, and on the B side when $A < 0$. The resulting row is nonzero only on one sign coordinate. It is independent of the selected rows, since their independent multiplicative projections first force their coefficients to vanish, after which the pure valuation rows have zero sign coordinates. We have proved that the two Selmer spaces have total codimension at least $k + 1$.

If both blocks contain the dyadic coordinate, project onto it as well. Let ζ_R be the restriction of ζ to the regular columns, and let $K_{B,R}$ and $K_{D,R}$ be the row restrictions of K_R to multiplicative primes dividing B and D, respectively. The regular local rows have dyadic coefficient $[2/p]$ in each block. Their projections in the two blocks have disjoint coordinate sets and combined rank

$$\operatorname{rank}\begin{pmatrix} K_{B,R} \\ \zeta_R \end{pmatrix} + \operatorname{rank}\begin{pmatrix} K_{D,R} \\ \zeta_R \end{pmatrix} \geq \operatorname{rank}\begin{pmatrix} K_R \\ \zeta_R \end{pmatrix}.$$

Choose regular rows with independent projections. The e pure odd-valuation rows have zero projection and are independent of these rows and of each other. Deleting the e nonregular columns gives

$$e + \operatorname{rank}\begin{pmatrix} K_R \\ \zeta_R \end{pmatrix} \geq \operatorname{rank}\begin{pmatrix} K \\ \zeta \end{pmatrix}.$$

The real condition supplies one further independent row as above. Thus the total codimension is at least $1 + \operatorname{rank}\left(\begin{smallmatrix} K \\ \zeta \end{smallmatrix}\right)$.

3.3. **The imaginary self-twist condition.** Products of the discriminants p^*, $p \in S$, form an a-dimensional space of quadratic characters. Let $U = \ker K^*$, and let κ be 1 if the newform of E has a nontrivial quadratic self-twist in U, and 0 otherwise. The usual characterization of self-twists of elliptic newforms says that the nontrivial self-twist, if present, is unique and is the imaginary quadratic character of complex multiplication. The endomorphism-algebra proof of this fact is given in Lemma 5.3.

Lemma 3.2. *If $\kappa = 1$, the codimension estimates above improve by one. This also holds for the augmented projection whenever that projection is used and the self-twist lies in its kernel.*

Proof. At an odd prime, rational two-torsion excludes inertia of order divisible by three. At $p = 3$ this includes wild inertia: its pro-3 image in $\mathrm{GL}_2(\mathbb{Z}_2)$ reduces into the stabilizer of a nonzero vector in $\mathbb{F}_2^2$, a group of order two, while the reduction kernel is pro-2. Hence that wild image is trivial. The usual tame reduction calculation then gives, at odd additive primes, conductor exponent two and types III, III*, I_0^*, or I_n^* with $n > 0$. Ramified quadratic twisting preserves additive reduction in the first two cases and removes it in the last two. A self-twist cannot be supported at a removable prime, since its conductor would change there.

At types III and III*, the two Tamagawa numbers are both two and the distinguished classes B, D both have odd valuation. This can also be read directly from the minimal equations: their valuations are respectively $(1,1)$ or $(3,3)$, up to exchanging the two isogenous curves. Thus both local images are ramified lines, by (3.4); these primes are regular. Let $v = (v_p)$ be the self-twist vector. Since its discriminant is negative,

$$\sum_{p \in S} v_p[-1/p] = 1.$$

The v-weighted sums of the regular rows in each block vanish on the multiplicative coordinates because $Kv = 0$, and have sign coefficient one. If the dyadic projection is used, they also vanish there because $\zeta v = 0$. Choose the weighted row in the block opposite to the real positivity row. Its sign coordinate proves its independence from that row and the pure valuation rows; its zero projection proves independence from the selected projected row basis. This is the extra condition. □

3.4. **The dyadic coordinates.** We next specify the ambient dimensions at two.

Assume first that E is optimal, $f = 0$, and $a > 0$. Good reduction at two together with rational two-torsion implies ordinary reduction. A connected rational order-two subgroup would give a degree-two quotient whose dual has unit minimal differential multiplier at every prime. Proposition 2.5 would place its nonzero kernel point in the identity component at an odd additive prime, an impossibility. Hence T is étale at two, and full rational two-torsion is impossible, since it would contain the connected order-two subgroup.

More generally, whenever E has good reduction at two and T is étale, translation of a minimal integral equation puts it in the form

$$y^2 + a_1xy = x^3 + a_2x^2 + a_4x,$$

with a_1, a_4 odd. Completing the square with $X = 4x$, $Y = 8y + 4a_1x$ gives (3.1) with

$$A = a_1^2 + 4a_2 \equiv 1 \pmod 4, \qquad B = 16a_4, \qquad v_2(B) = 4, \quad D \text{ odd}. \tag{3.6}$$

The integral kernel of η is the constant group $\mathbb{Z}/2$, and that of its dual is μ_2. Finite flat Kummer theory over $\mathbb{Z}_2$, and the vanishing of the first cohomology of the smooth elliptic special fiber over $\mathbb{F}_2$, give

$$L_{D,2} = H^1(\mathbb{Z}_2, \mathbb{Z}/2) = \{1, 5\}, \qquad L_{B,2} = H^1(\mathbb{Z}_2, \mu_2) = \mathbb{Z}_2^\times/\mathbb{Z}_2^{\times 2}. \tag{3.7}$$

Here H^1 denotes fppf cohomology. In particular neither block has a dyadic valuation coordinate. This also explains ordinary reduction in the case used above: a supersingular elliptic curve over $\mathbb{Z}_2$ has no rational order-two point, since its height-two formal group has nonzero two-torsion parameter of valuation $1/3$, whereas an integral point of order two cannot reduce to a nonzero point on its supersingular special fiber.

If $f = 1$ and E is optimal, then

$$E(\mathbb{Q})[2] \cap E^0(\mathbb{Q}_2) = 0. \tag{3.8}$$

Indeed, a nonzero point in this intersection is -1 in a Tate uniformization over an unramified splitting extension. The quotient and dual maps are

$$\overline{\mathbb{Q}}_2^\times/q^{\mathbb{Z}} \xrightarrow{z \mapsto z^2} \overline{\mathbb{Q}}_2^\times/q^{2\mathbb{Z}} \xrightarrow{z \mapsto z} \overline{\mathbb{Q}}_2^\times/q^{\mathbb{Z}}.$$

The dual has globally unit differential multiplier, but its nonzero kernel point, represented by q, is outside the identity component. This contradicts Proposition 2.5. The component group at a multiplicative prime is cyclic, so (3.8) also excludes full rational two-torsion. The point T is integral and singular; the same normalization as above now has a_1 odd and a_4 even. Thus A, D

are odd and exactly the B block contains the coordinate 2. Writing $\mu = m + 1$ for the number of all multiplicative primes, the total ambient dimension is $2a + \mu + 2$. The local row argument applies with the extra multiplicative row $q = 2$, hence with k^*.

Suppose finally that $f \geq 2$. A reduced minimal equation at two has a_1, a_3 even, and completing the square is integral. The integral additive formal logarithm of Lemma 2.1 makes $E^1(\mathbb{Q}_2)$ torsion-free. Thus T is integral and can be translated to zero without losing minimality. In (3.1) put

$$d = \mathbf{1}_{2|B} + \mathbf{1}_{2|D} \leq 2, \tag{3.9}$$

where $\mathbf{1}$ denotes the indicator of a condition. The total ambient dimension is $2a + m + 2 + d$.

There is a sharper statement when $f = 2$. The prime-to-two Tate module has invariant codimension two and Swan conductor zero. Additive potentially multiplicative reduction would require a ramified quadratic character at two and hence positive Swan conductor. The reduction is therefore potentially good, with nontrivial tame inertia of odd order. The odd part of the automorphism group of a geometric elliptic curve in characteristic two has order at most three. Consequently, over a strict unramified base, good reduction is attained in a tamely ramified cubic extension. Restriction followed by norm is multiplication by three. More precisely, Weil restriction of the good elliptic model from this totally ramified cubic extension has connected special fiber. Its reduction to the elliptic special fiber has a filtration whose two successive kernels are additive groups, obtained from the two nilpotent layers of the ramified cubic special fiber. Its generic norm map extends to the original Néron model by smoothness and the Néron mapping property, so its image lies in the identity component. Thus the geometric component group is killed by three. In particular its order is odd. The same conclusions hold for E' because its conductor is the same.

Both distinguished rational order-two points are consequently nonsingular. In their respective minimal translated short equations, the linear coefficient is odd. For such an equation, if $v_2(x) < 0$ then x^3 dominates, while if $v_2(x) > 0$ then $x^2 + Ax + B$ is a unit. In either case the square equation forces $v_2(x)$ to be even. The Kummer value at the order-two point is the unit B itself. Changing between two translated models preserving this point changes x by a rational square. The two local Kummer images therefore have even valuation at two, even if the displayed equation for E' is not minimal there. Both dyadic coordinates disappear. Moreover, full rational two-torsion is impossible: the odd component group would place all of it in $E^0(\mathbb{Q}_2)$, whereas Lemma 2.1 gives $|E^0(\mathbb{Q}_2)[2^\infty]| \leq 2$.

Combining these facts with Lemmas 3.1 and 3.2 proves the estimates we shall use.

Proposition 3.3. *With the notation above, the following bounds hold.*

(i) *If $f = 0$, $a > 0$, and E is optimal, then*

$$r \leq 2a + m - 1 - k - \kappa.$$

The bound also holds for any E whose chosen order-two point is étale at two.

(ii) *If $f = 1$ and E is optimal, with $\mu = m + 1$, then*

$$r \leq 2a + \mu - 1 - k^*.$$

There is no CM self-twist in this case, because a CM elliptic curve cannot have multiplicative reduction. The inequality also holds for any E whose chosen point lies outside $E^0(\mathbb{Q}_2)$.

(iii) *If $f \geq 2$, then*

$$r \leq 2a + m + d - 1 - k - \kappa. \tag{3.10}$$

If $d = 2$, one also has

$$r \leq 2a + m + 1 - \operatorname{rank} \binom{K}{\zeta} - \kappa, \tag{3.11}$$

provided, for the subtraction of κ, that the self-twist is in the kernel of the augmented matrix. This is the case when f is odd.

(iv) *If $f = 2$, the stronger bound is*

$$r \le 2a + m - 1 - k - \kappa. \tag{3.12}$$

3.5. **The semistable extreme case.**

Lemma 3.4. *If E is semistable and has rational two-torsion, then $r \le \omega(N) - 1$.*

Proof. At good reduction at two, choose an étale rational order-two kernel. If the given kernel is connected, pass to its quotient; the rational dual kernel is then étale. At multiplicative reduction at two, choose a rational order-two point outside the identity component. If the given point is inside it, the Tate calculation above shows that the dual kernel on its quotient is outside it. Isogeny preserves rank and conductor. We may thus apply the first or second ambient calculation with $a = 0$, K empty, and the single real positivity condition. Lemma 3.1 then gives $r \le \omega(N) - 1$. □

Lemma 3.5. *Suppose N is odd and squarefree, $r = \omega(N) - 1 > 0$, and a nonzero point*

$$T \in E(\mathbb{Q})[2] \cap \bigcap_{p \mid N} E^0(\mathbb{Q}_p)$$

is étale at two. In the normalization (3.6) one has

$$B = \pm 16, \qquad D > 0, \qquad q \mid D \Longrightarrow q \equiv 1 \pmod 4.$$

If $B = 16$, then $A > 8$. In these circumstances $E(\mathbb{Q})[2]$ is cyclic of order two.

Proof. An odd prime dividing B would make T singular at a multiplicative prime. Hence (3.6) gives $B = \pm 16$. Put $n = \omega(N)$. The two ambient spaces have total dimension $n + 2$, whereas their Mordell–Weil images have total dimension $r + 2 = n + 1$. The D-side condition in (3.7) contains the row with coefficient one on -1 and on exactly those odd primes $q \equiv 3 \pmod 4$. The real positivity row must coincide with it: if the two rows were independent, the total dimension would be at most n. Thus positivity occurs on the D side, all primes dividing D are 1 modulo four, and $D > 0$. For $B = 16$ the real sign calculation requires $A > 0$; then $D = A^2 - 64 > 0$ gives $A > 8$.

Full rational two-torsion would make $D = C^2$ for an odd integer C. If $B = -16$, the equation $C^2 - A^2 = 64$, together with $A \equiv 1 \pmod 4$, gives $A = -15$, $C = 17$. The odd conductor is then 17, contradicting $r > 0$ by Lemma 3.4. If $B = 16$, $A > 8$ and $A^2 - C^2 = 64$ give $A = 17$, $C = 15$. Then $3 \mid D$, contrary to the established congruence condition. □

3.6. **The possible loss on a nonoptimal target.** For an isogeny $\nu : C \longrightarrow C'$ between elliptic curves over $\mathbb{Q}$, write $\nu^* \omega_{C'} = c_\nu \omega_C$, using global minimal differentials and choosing signs so that $c_\nu > 0$. The Néron mapping property makes c_ν an integer, and the product of the multipliers of a degree-two isogeny and its dual is 2.

Lemma 3.6. *If $C/\mathbb{Q}_2$ has additive reduction, at most one rational cyclic two-isogeny from C has nonunit minimal differential multiplier.*

Proof. There is nothing to prove if the rational two-torsion is cyclic. Otherwise complete the square in a minimal equation and write $y^2 = \prod_{i=1}^{3}(x - r_i)$. The additive formal subgroup is torsion-free, so $r_i \in \mathbb{Z}_2$. The three difference valuations are n, n, m with $m > n$: after the common smallest power of two is removed, three residues cannot be distinct in $\mathbb{F}_2$. Also $n \le 1$. If all roots were congruent modulo four, translation by one root followed by $x = 4X$, $y = 8Y$ would lower the discriminant by 2^{12} and still give an integral equation.

Translate either root of the close pair to zero. In (3.1) this gives

$$v_2(A) = n, \qquad v_2(B) = n + m, \qquad v_2(D) = 2n.$$

The displayed quotient has $c_4 = 16(A^2 + 12B)$ and differential multiplier one before minimization. A nonunit multiplier would therefore require a scaling by two to a minimal target. If $n = 0$, its minimal c_4 would be a unit, making the target semistable, contrary to invariance of additive reduction under isogeny. If $n = 1$, put the minimal target in integral complete-square form. The change of coordinates has $x_{\text{naive}} = 4x_{\min} + r$. The dual order-two point is integral on the minimal additive target, so $r \in 4\mathbb{Z}_2$. The transformed linear coefficient would be

$$\frac{D - 4Ar + 3r^2}{16}.$$

Its numerator has valuation exactly two, since $v_2(D) = 2$ and the other terms are divisible by sixteen. This contradicts integrality. Thus neither close-pair kernel is a nonunit direction; only the third kernel can be one. □

Proposition 3.7. *Let F_0 be the optimal curve in an isogeny class with rational two-torsion, and let F be any curve in that class. If the class has an odd additive prime, then the Manin constant of a minimal parametrization to F is odd and*

$$v_2(m_F/c_F^2) \geq v_2(m_{F_0}). \tag{3.13}$$

If the reduction at two is additive, then

$$v_2(m_F/c_F^2) \geq v_2(m_{F_0}) - 1. \tag{3.14}$$

Proof. Choose a rational isogeny $\nu : F_0 \longrightarrow F$ of least degree. Its kernel is cyclic: otherwise it contains $F_0[\ell]$ for some prime ℓ, and factoring out $[\ell]$ lowers the degree. Take the two-primary part first and write it as a nonbacktracking chain of k two-isogenies, with successive curves $F_0, F_1, \ldots, F_k$. In good ordinary reduction, the nonunit direction is the unique connected kernel; in multiplicative reduction the same assertion follows from Tate uniformization. At additive reduction it is Lemma 3.6. Consequently, after a unit step the dual is the only possible nonunit next step. Cyclicity excludes this backtracking. All nonunit steps therefore form an initial run, say of length s.

The dual of that initial composite is an incoming isogeny $F_s \longrightarrow F_0$ whose minimal differential multiplier is a unit at every prime. Proposition 2.5 makes its entire cyclic kernel rational and places it in $F_s^0(\mathbb{Q}_p)$ at every bad prime. An odd additive identity component has no nonzero two-primary torsion, so $s = 0$ if an odd additive prime exists. At additive reduction at two, Lemma 2.1 gives $|F_s^0(\mathbb{Q}_2)[2^\infty]| \leq 2$, hence $s \leq 1$.

Odd-degree isogeny factors have unit multiplier at two. By Proposition 2.6, $v_2(c_{F_0}) = 0$, and hence

$$v_2(c_F) = s, \qquad v_2(m_F/c_F^2) = v_2(m_{F_0}) + k - 2s.$$

Here minimal modular degrees multiply by $\deg \nu$, by the optimal factorization. If $s = 0$ this proves the first assertion. If $s \leq 1$, the inequality $k \geq s$ gives the second. □

Lemma 3.8. *Suppose $C/\mathbb{Q}$ has additive reduction at two and a rational two-primary torsion subgroup of order at least eight. Then it has a rational nonzero point of order two in $C^0(\mathbb{Q}_2)$. For this point the minimal translated short equation has B odd. The same is true for its image on any quadratic twist unramified at two; in particular the dyadic support number d in (3.9) is at most one on that twist.*

Proof. The component group of an additive elliptic curve has order at most four, by the minimal regular reduction calculation. The torsion subgroup therefore has a nontrivial kernel in the

component group. Since the two-primary torsion of $C^0(\mathbb{Q}_2)$ has order at most two, this kernel supplies the asserted point. It is integral and nonsingular. On a minimal complete-square equation translated by it, nonsingularity of $(0,0)$ in characteristic two is exactly the condition that the linear coefficient B is odd. An unramified quadratic extension preserves the identity component, and an order-two point is unchanged by the sign defining a quadratic twist. The assertion thus persists after such a twist and unramified descent. □

4. The semistable case

For squarefree conductor, the equality case in the rank bound is settled by a parity argument on a quotient of the modular curve.

Lemma 4.1. *Suppose that $E/\mathbb{Q}$ is semistable, has positive rank, and has conductor N. Normalize $\phi(\infty) = O$, and put $H = \ker(\varepsilon : W \to \{\pm 1\})$. Then*

$$\operatorname{im}(t|_H) \;\subset\; L := E(\mathbb{Q})[2] \cap \bigcap_{p \mid N} E^0(\mathbb{Q}_p), \qquad \dim_{\mathbb{F}_2} L \le 1.$$

Consequently $v_2(\deg \phi) \ge \omega(N) - 2$.

Proof. At a bad prime p, Lemma 2.4 shows that $t_d \in E^0(\mathbb{Q}_p)$ whenever $p \nmid d$. If $p \mid d$, use the complementary divisor N/d and (1.2); its translation is $\pm t_d$. Thus the assertion holds for every d in H. On H the translation is killed by two.

Over an unramified splitting field at a multiplicative prime, Tate uniformization identifies the two-torsion of the identity component with a subgroup of μ_2. Hence $\#L \le 2$. The actual kernel of $t|_H$ has order at least $|H|/2$, and $|H| \ge 2^{\omega(N)-1}$. Factoring ϕ through this kernel gives the last assertion. □

Lemma 4.2. *Assume that N is odd, $r > 0$, and all the translations t_h, $h \in H$, specialize to O at two. Then $|H| \mid \deg \phi$.*

Proof. Both $X_0(N)$ and E have smooth proper models over $\mathbb{Z}_2$, and ϕ extends to them. The degree of the pullback of $\mathcal{O}_E(3O)$ is constant on the fibers. Its generic degree is $3 \deg \phi > 0$, so the special map is nonconstant and has degree $\deg \phi$.

The genus of $X_0(N)$ is at least two in the present situation. Genus zero is impossible for a nonconstant map to an elliptic curve. If the genus were one, the Abel–Jacobi map $\iota : X_0(N) \to J_0(N)$ based at ∞ would be an isomorphism. Its invariant differential pulls back to a nonzero multiple of $f_E(q)dq/q$. Since $L(E,1) = 0$, integration from ∞ to 0 would give $\iota(0) = \iota(\infty)$, contrary to the distinctness of these cusps. Automorphisms of a smooth curve of genus at least two specialize faithfully: its relative automorphism scheme is unramified, so two sections agreeing on the special fiber agree over the DVR. Thus the special action of H is faithful. The special map is H-invariant because its translations vanish. Artin's theorem on fixed fields gives the factor $|H|$ in its degree, including when the remaining extension is inseparable. □

Lemma 4.3. *Let N be odd and squarefree, with every prime divisor congruent to $1 \pmod 4$. Let $d > 1$ divide N, let x be a geometric fixed point of w_d on $X_0(N)$, and let y be its image on $X_0(N)/J$, where J is any subgroup of W. Then*

$$\mathbb{Q}(\sqrt{d}) \subseteq \mathbb{Q}(y).$$

In particular, y has even residue degree over $\mathbb{Q}$, and its base changes have even residue degree over $\mathbb{Q}(i)$.

Proof. An Atkin–Lehner involution $w_d \neq 1$ fixes no cusp at squarefree level. Choose z in the upper half-plane above x. A representative of the Atkin–Lehner coset fixing z has the form

$$g = \begin{pmatrix} da & b \\ Nc & de \end{pmatrix}, \qquad \det g = d.$$

Its trace is divisible by d, and $|\operatorname{Tr} g| < 2\sqrt{d}$. Since $d \geq 5$, its trace is zero. On $\Lambda = \mathbb{Z}z + \mathbb{Z}$, multiplication by $\lambda = Ncz + de$ sends 1 to $Ncz + de$ and z to $daz + b$. It is an endomorphism of trace zero and norm d. Consequently $\lambda^2 = -d$, and the endomorphism order is $\mathcal{O}_K = \mathbb{Z}[\sqrt{-d}]$, where $K = \mathbb{Q}(\sqrt{-d})$ has discriminant $-4d$. The order is maximal, because d is squarefree and $d \equiv 1 \pmod 4$.

The full level subgroup is stable under this order: on its generator $1/N$,

$$\lambda/N = cz + de/N.$$

Each cyclic subgroup of order $q \mid N$ is therefore the kernel of an $\mathcal{O}_K$-ideal of norm q. In particular, all the Atkin–Lehner isogenies from this point are horizontal, including those at the level primes outside d.

Let H_K be the Hilbert class field of K. The biquadratic field

$$K(i) = \mathbb{Q}(i, \sqrt{d})$$

is unramified over K: its three quadratic discriminants are $-4, d, -4d$, whose product is $16d^2 = \operatorname{disc}(K)^2$. It is thus contained in H_K. For the real principal singular modulus $j_0 = j(i\sqrt{d})$, the main theorem of complex multiplication gives $H_K = K(j_0)$. Since $[\mathbb{Q}(j_0) : \mathbb{Q}] = [H_K : K]$, $\mathbb{Q}(j_0)$ is the fixed field of complex conjugation. It contains $\sqrt{d}$. Every singular modulus for $\mathcal{O}_K$ is a conjugate of j_0, and $\mathbb{Q}(\sqrt{d})/\mathbb{Q}$ is Galois. Therefore

$$\mathbb{Q}(\sqrt{d}) \subseteq \mathbb{Q}(j(x)).$$

The ideal-class description and reciprocity law for maximal CM orders are given in [15, 20].

If $\mathfrak{q}$ is one of the level ideals of norm $q \mid N$, its Artin automorphism in $K(i)/K$ is trivial, since $q \equiv 1 \pmod 4$. This remains true for a prime ramified in $K/\mathbb{Q}$, because $K(i)/K$ is unramified and its Frobenius acts on i by the q-power map. Hence the horizontal ideal action associated with any $w \in W$ fixes $\sqrt{d}$.

If $\sigma \in \operatorname{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$ fixes y, then $\sigma x = wx$ for some $w \in J$. On the underlying j-invariant this equals the action of an ideal-class automorphism $\sigma_{\mathfrak{b}}$ fixing $\sqrt{d}$. The automorphism $\sigma_{\mathfrak{b}}^{-1}\sigma$ fixes $j(x)$, and hence $\sqrt{d}$. Thus σ fixes $\sqrt{d}$, proving the asserted containment in the residue field of the quotient point. Finally, the real quadratic field $\mathbb{Q}(\sqrt{d})$ remains quadratic after adjoining i. □

Lemma 4.4. *Let $u : C \to D$ be a finite map of smooth projective curves over a characteristic-zero field k, equivariant for involutions j on C and a on D. Suppose that a has a k-rational fixed point and that every geometric fixed point of j has even residue degree over k. Then* $\deg u$ *is even.*

Proof. Pull back a rational fixed point to a j-invariant divisor. Two-element orbits of closed points contribute even degree. A stable closed point on which j acts nontrivially on the residue field also has even degree, by the degree-two fixed-field extension. If the residue-field action is trivial, its geometric points are fixed and the hypothesis applies. These conclusions include the ramification multiplicities in the pullback. □

Lemma 4.5. *Suppose that*

$$E: \quad y^2 = x(x^2 + Ax + B), \qquad B = \pm 16, \quad A \equiv 1 \pmod 4,$$

has odd squarefree conductor N, unique nonzero rational two-torsion point $T = (0, 0)$, and every prime divisor of N congruent to 1 (mod 4)*. In the case $B = 16$, suppose also $A > 8$. If $\varepsilon \neq 1$ and $\operatorname{im}(t|_H) \subseteq \langle T \rangle$, then $|H| \mid \deg \phi$.*

Proof. Let $\eta : E \to E' = E/\langle T \rangle$. The map $\eta\phi$ descends to $\psi : Y = X_0(N)/H \to E'$, and

$$|H| \deg \psi = 2 \deg \phi. \tag{4.1}$$

A negative Atkin–Lehner element induces an involution j on Y, and $\psi j = a\psi$, where

$$a(R) = -R + \eta(P)$$

for a rational cuspidal torsion point $P \in E(\mathbb{Q})$. Every geometric j-fixed point lifts to a point fixed by some negative w_d: if $wx = hx$ with $h \in H$, use wh^{-1}. Here $d > 1$. Lemma 4.3 gives even residue degree over both $\mathbb{Q}$ and $\mathbb{Q}(i)$.

It remains to find a fixed point of a over one of these fields. The integral good model at two is

$$Y^2 + XY = X^3 + \frac{A-1}{4}X^2 + \frac{B}{16}X.$$

It has two or four points over $\mathbb{F}_2$, and T reduces to the nonidentity point $(0, 0)$. Rational torsion injects under reduction: its kernel is pro-two, and a nonzero torsion point in the kernel would have a multiple of order two there, contradicting uniqueness of T. The rational torsion is consequently cyclic of order two or four. The duplication formula

$$x(2P) = \frac{(x(P)^2 - B)^2}{4x(P)(x(P)^2 + Ax(P) + B)}$$

shows that a point satisfying $2P = T$ has $x(P)^2 = B$. For $B = -16$ this is impossible, so $\eta(P) = O$ and a fixes O.

For $B = 16$, write $u = A - 8$, $v = A + 8$. The standard two-isogenous equation is

$$E' : \quad y^2 = x(x - u)(x - v), \qquad v - u = 16.$$

If P has order at most two, again $\eta(P) = O$. Otherwise $x(P) = 4$ or -4. In the first case $v = t^2$ for some $t > 0$ in $\mathbb{Q}$, $\eta(P) = (v, 0)$, and

$$R = (t^2 + 4t, -4t(t + 4)) \in E'(\mathbb{Q}) \quad \text{satisfies} \quad 2R = (v, 0).$$

In the second case $u = s^2$, $s > 0$ rational, $\eta(P) = (u, 0)$, and

$$R = (s^2 + 4is,\ 16s - 4is^2) \in E'(\mathbb{Q}(i)), \qquad 2R = (u, 0).$$

For the first point, the three cubic factors are $t(t + 4), 4(t + 4), 4t$; for the second they are $s(s+4i), 4is, 4i(s+4i)$. Their products equal the respective squares of the displayed y-coordinates. The tangent and duplication formulas on E' are

$$\lambda = \frac{3x^2 - 2(u + v)x + uv}{2y}, \qquad x(2R) = \lambda^2 + u + v - 2x(R),$$

$$y(2R) = -y(R) + \lambda(x(R) - x(2R)).$$

The slopes are $-t - 4$ and $-s - 4i$, respectively. Substitution gives $2R = (v, 0)$ and $2R = (u, 0)$. The denominators are nonzero because $t > 4$ and $s > 0$. Thus a has a fixed point over $\mathbb{Q}(i)$. Lemma 4.4 now makes $\deg \psi$ even. Equation (4.1) proves the assertion. □

Proposition 4.6. *Theorem 1.1 holds when N is squarefree.*

Proof. By Lemma 1.2, take E optimal and $r > 0$, and put $n = \omega(N)$. Lemma 3.4 gives $r \le n-1$. If $2 \mid N$, equation (3.8) and Lemma 4.1 make all translations on H zero. Hence $v_2(\deg\phi) \ge n-1$.

Assume N odd. Lemma 4.1 settles $r \le n-2$. If $\varepsilon = 1$, its translation kernel already has order at least 2^{n-1}. If $t|_H = 0$, the same conclusion follows directly. The only remaining case is

$$r = n-1 > 0, \qquad \varepsilon \ne 1, \qquad \operatorname{im}(t|_H) = \langle T \rangle \ne 0.$$

If T specializes to O at two, apply Lemma 4.2. Otherwise T is étale at two. The normalization and equality analysis in Lemma 3.5 give

$$B = \pm 16, \quad D = A^2 - 4B > 0, \quad q \mid D \Longrightarrow q \equiv 1 \pmod 4,$$

and $A > 8$ for $B = 16$.

The same equality analysis excludes full rational two-torsion, so T is the unique nonzero rational two-torsion point. All primes of N divide D, because $B = \pm 16$. Lemma 4.5 supplies the last factor of two, proving $v_2(\deg\phi) \ge n-1 = r$. □

5. Quadratic twists and integral Fourier sums

Let $E/\mathbb{Q}$ have nonzero rational two-torsion, and write $N = N_E$. Throughout this section, S is the set of odd additive primes of E, $a = |S|$, and $f = v_2(N)$. Put

$$D_v = \prod_{p \in S} (p^*)^{v_p} \quad (v \in \mathbb{F}_2^S).$$

Every D_v is an odd fundamental discriminant, with $D_0 = 1$. Write χ_v for its quadratic character and g_v for the normalized newform of $E^{(D_v)}$. The extension generated by all $\sqrt{p^*}$ is unramified at 2. Its ramification index at any odd prime is at most two: over an odd local field there is only one nontrivial quadratic character on inertia.

5.1. **Local conductors and signs.** The image of the mod-2 representation fixes a nonzero vector and therefore has order at most two. The kernel of $\mathrm{GL}_2(\mathbb{Z}_2) \to \mathrm{GL}_2(\mathbb{F}_2)$ is pro-2. Consequently wild inertia acts trivially at every odd prime, including 3. At an odd additive prime the conductor exponent is therefore two. The tame potentially good inertia has order two or four; the potentially multiplicative case is a ramified quadratic twist of multiplicative reduction. These are precisely the following possibilities:

reduction of E	reduction after a ramified quadratic twist	new conductor exponent
I_0^*	good	0
$\mathrm{I}_n^*,\ n > 0$	multiplicative	1
III or III*	III* or III	2.

Here and below an unramified quadratic twist is allowed in addition to the indicated ramified twist. Such a twist does not change these conductor categories. One can obtain the table either from the tame inertia representation or directly from the minimal discriminant valuations in Tate's algorithm. In the order-four case, tensoring the inertia characters by the quadratic character interchanges the characters of order four; it cannot make the representation unramified.

Call an odd additive prime *removable* in the first two rows of the table. The remaining primes will be called primes of type III/III*. Thus $N_{E^{(D_v)}} \mid N$.

The Atkin–Lehner sign at a prime of type III/III* is unchanged by a ramified quadratic twist. Normalize the Weil representation to have determinant one. If $p \equiv 1 \pmod 4$, it is a sum $\xi \oplus \xi^{-1}$ of tame characters. The product formula for local epsilon factors gives root number $\xi(-1)$. A ramified quadratic twist multiplies this by $\chi(-1) = (-1/p) = 1$. If $p \equiv -1 \pmod 4$, the representation is induced from a character ξ of the unramified quadratic extension. Its restriction to units has order four and its value on a uniformizer is -1, by the determinant-one

condition. The restriction of a ramified quadratic character is its square on units and is one on a uniformizer. It therefore changes ξ to ξ^{-1}, giving an isomorphic induced representation. The local epsilon-factor identities used here are the usual one-dimensional product and induction identities; see [30, §1].

More generally, an unramified quadratic twist changes the sign at a prime q by $\chi(q)^{v_q(N)}$ [30, Proposition 8(ii)]. Thus the signs at odd additive primes do not change under an unramified twist, the sign at an odd multiplicative prime q changes by $\chi_v(q)$, and the sign at 2, when $2 \mid N$, changes by $\chi_v(2)^f$.

5.2. **Oldforms with prescribed signs.** The level can drop at a removable prime. We now raise the resulting newform to level N, keeping both integrality and the original Atkin–Lehner signs. Write

$$V_p g(z) = g(pz).$$

This convention is for forms: the corresponding pullback of differentials under the degeneracy map $z \mapsto pz$ is $pV_p g(q)\, dq/q$.

Lemma 5.1. *Suppose p is an odd prime at which twisting has lowered the conductor exponent from two. Let g be the resulting normalized newform, let $a_p = a_p(g)$, and let ϵ_p be the Atkin–Lehner sign of the original newform at p^2. There is a normalized integral oldform of sign ϵ_p given by*

$$\begin{cases} g - \epsilon_p a_p p V_p g, & \text{if } g \text{ has conductor exponent one at } p, \\ g - a_p V_p g + p^2 V_p^2 g, & \text{if } g \text{ is unramified at } p \text{ and } \epsilon_p = 1, \\ g - p^2 V_p^2 g, & \text{if } g \text{ is unramified at } p \text{ and } \epsilon_p = -1. \end{cases} \tag{5.1}$$

Every coefficient at p^j, $j \geq 1$, is even. Under a further unramified quadratic twist χ at p, the coefficient at p^j is multiplied by $\chi(p)^j$.

Proof. With the weight-two slash normalization, the Atkin–Lehner matrices on the ordered old bases $(g, V_p g)$ and $(g, V_p g, V_p^2 g)$, respectively, are

$$\begin{pmatrix} 0 & -a_p/p \\ -a_p p & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 0 & p^{-2} \\ 0 & 1 & 0 \\ p^2 & 0 & 0 \end{pmatrix}. \tag{5.2}$$

For the first matrix, use the lower-level identity $w_p g = -a_p g$, where $a_p = \pm 1$, followed by the degree-p degeneracy map. For the second, the two extreme degeneracy maps are interchanged and the middle one is fixed. These identities also follow by multiplying representatives for the Atkin–Lehner operators with the matrices $\left(\begin{smallmatrix} p^j & 0 \\ 0 & 1 \end{smallmatrix}\right)$; the determinant factors in weight two give p and p^2 in (5.2). This is the oldform normalization of [3, 27]. Multiplication of the displayed matrices by the coefficient vectors in (5.1) proves the required signs. Their first Fourier coefficient is one.

In the multiplicative case, $a_{p^j}(g) = a_p^j$, so the coefficient in the first line of (5.1) is

$$a_p^j(1 - \epsilon_p p) \qquad (j \geq 1).$$

It is even because p is odd. Replacing a_p by $\chi(p)a_p$ proves the asserted covariance in this case.

In the good case put $A_j = a_{p^j}(g)$, with $A_{-1} = A_{-2} = 0$. Reduction of rational two-torsion at the odd good prime gives $a_p \equiv 0 \pmod 2$. The recurrence is

$$A_0 = 1, \qquad A_1 = a_p, \qquad A_j = a_p A_{j-1} - p A_{j-2} \quad (j \geq 2).$$

For the plus vector the coefficient at p is zero, and for $j \geq 2$ it is

$$A_j - a_p A_{j-1} + p^2 A_{j-2} = p(p-1)A_{j-2}.$$

For the minus vector it is $A_j - p^2 A_{j-2}$. Modulo two the recurrence gives $A_j = 0$ for odd j and $A_j = 1$ for even j; hence the latter coefficient is also even, including $j = 1$. Finally the same recurrence shows that replacing a_p by $\chi(p)a_p$ replaces A_j by $\chi(p)^j A_j$. Since $\chi(p)^{j-2} = \chi(p)^j$, both oldvectors have the claimed covariance. □

Apply Lemma 5.1 at every removable prime where the conductor of g_v drops. Operators at distinct primes commute. Denote the resulting level-N form by F_v. Then

$$F_0 = f_E, \qquad a_1(F_v) = 1, \qquad F_v \in \mathbb{Z}[[q]].$$

The coefficients remain multiplicative, with the modified local factors in (5.1). At an additive prime that has not been removed, all positive prime-power coefficients vanish. If $f \geq 2$, every F_v has zero even-index coefficients; in particular the half-translation $z \mapsto z + 1/2$ acts on every F_v by -1.

Let K be the matrix of Section 3, with rows indexed by the odd multiplicative primes q, columns indexed by $p \in S$, and

$$K_{q,p} = [q/p] \in \mathbb{F}_2.$$

Put

$$\zeta_p = [2/p], \qquad K^* = \begin{cases} K, & f \text{ even}, \\ \binom{K}{\zeta}, & f \text{ odd}, \end{cases} \qquad U = \ker K^*. \tag{5.3}$$

Write $k^* = \operatorname{rank}_{\mathbb{F}_2} K^*$, so $\dim U = a - k^*$. For $f = 1$, the additional row is precisely the row belonging to the multiplicative prime 2. Quadratic reciprocity gives $\chi_{p^*}(q) = (q/p)$, also for $q = 2$. The preceding sign calculations show that all F_v, $v \in U$, have the same full Atkin–Lehner character as f_E.

5.3. **Divisibility at every Fourier index.** We require a statement for affine subspaces, since those appear when the twist field is nonsplit at 2. An affine sign character on an affine $\mathbb{F}_2$-space A is a function $\rho : A \to \{\pm 1\}$ whose restriction after choosing any origin is a constant sign times a group character.

Lemma 5.2. *Let $A \subseteq \mathbb{F}_2^S$ be any nonempty affine subspace of dimension b, and let ρ be an affine sign character. Then*

$$G_{A,\rho} := \sum_{v \in A} \rho(v) F_v \in 2^b \mathbb{Z}[[q]]. \tag{5.4}$$

Proof. Fix an integer $n \geq 1$. If a prime of type III/III* divides n, every $a_n(F_v)$ is zero. Otherwise let I be the set of removable primes dividing n. The coefficient can be nonzero only on the affine slice

$$A_I = \{v \in A : v_p = 1 \text{ for all } p \in I\},$$

since an unremoved additive prime has zero positive prime-power coefficients. If the slice is empty there is nothing to prove. Otherwise choose $v_0 \in A_I$, and let L be the translation space of A. The translation space of A_I is

$$L_I = \ker(L \longrightarrow \mathbb{F}_2^I), \qquad \dim L_I = b - \operatorname{rank}(L \longrightarrow \mathbb{F}_2^I).$$

Lemma 5.1 and multiplicativity give

$$2^{|I|} \mid a_n(F_{v_0}).$$

For $t \in L_I$, the character χ_t is unramified at every prime dividing n. At primes in I use the exact prime-power covariance in Lemma 5.1; at the other primes use the ordinary unramified twist identity. Together they give the equality

$$a_n(F_{v_0+t}) = \chi_t(n)\, a_n(F_{v_0}).$$

Thus the signed sum over A_I is either zero or, up to a sign, $2^{\dim L_I} a_n(F_{v_0})$. Its valuation is at least

$$|I| + b - \operatorname{rank}(L \longrightarrow \mathbb{F}_2^I) \geq b.$$

This also covers $n = 1$, for which $I = \varnothing$. If $2 \mid n$ and $f \geq 2$, the coefficients vanish. If $f \leq 1$, the twist at 2 is unramified and the same covariance applies. □

Lemma 5.3. *The subgroup of quadratic characters χ for which $f_E \otimes \chi = f_E$ has order at most two. If it is nontrivial, its nontrivial element is the character of the imaginary quadratic field of complex multiplication of E. If its vector belongs to $\mathbb{F}_2^S$, it is supported only at primes of type* III/III*.

Let $\kappa \in \{0, 1\}$ be the dimension of the self-twist subgroup contained in U, and put $h = a - k^* - \kappa$. There is a subspace $U' \subseteq U$ of dimension h such that the F_v, $v \in U'$, have pairwise distinct prime-to-N Hecke eigensystems. For any character $\rho : U' \to \{\pm 1\}$,

$$G_{U',\rho} = f_E + g_\rho, \qquad \langle f_E, g_\rho \rangle_N = 0. \tag{5.5}$$

Proof. A rational isogeny from E to a quadratic twist $E^{(D)}$, transported back to E over $\overline{\mathbb{Q}}$, gives a nonzero element of $\operatorname{End}_{\overline{\mathbb{Q}}}(E) \otimes \mathbb{Q}$ on which Galois acts through χ_D. Without complex multiplication this algebra is $\mathbb{Q}$, so only the trivial character occurs. With complex multiplication it is an imaginary quadratic field, whose Galois representation is the sum of the trivial line and its trace-zero line. The character on the latter is exactly the character of that quadratic field. The equivalence between equality of the newform eigensystems and rational isogeny gives the first assertion.

A self-twist cannot involve a removable prime: twisting at that prime would lower the conductor exponent from two. Consequently a self-twist supported on S is supported on the last row of the local table. Its fundamental discriminant is negative.

Choose U' complementary to the self-twist subgroup of U. Equality of two eigensystems in U' would give a nonzero self-twist in that complement, which is impossible. Distinct prime-to-N eigensystems are orthogonal for the Petersson inner product, since the operators T_ℓ, $\ell \nmid N$, are self-adjoint and distinguish them. The term $v = 0$ is f_E with coefficient $\rho(0) = 1$, which proves (5.5). □

Lemma 5.2 therefore supplies

$$2^{-h}(f_E + g_\rho) \in \mathbb{Z}[[q]].$$

The passage from this Fourier statement to a lattice on an actual quotient of $X_0(N)$ will be made in the following sections.

5.4. **The exact degree change at removable primes.** The following identity compares a curve with its quadratic twist.

Proposition 5.4. *Let $F/\mathbb{Q}$ have nonzero rational two-torsion, let D be an odd fundamental discriminant supported at odd primes where F is semistable, and put $E = F^{(D)}$. For a curve C, define*

$$V(C) = v_2(m_C / c_C^2),$$

using a minimal-degree parametrization to that particular curve and its Manin constant. Then

$$\begin{aligned} V(E) - V(F) = &\sum_{\substack{p \mid D \\ F \text{ good at } p}} v_2\big((p-1)((p+1)^2 - a_p(F)^2)\big) \\ &+ \sum_{\substack{p \mid D \\ F \text{ multiplicative at } p}} v_2(p^2 - 1). \end{aligned} \tag{5.6}$$

Every summand is at least three. If $F_0/\mathbb{Q}$ is isogenous to F and $F_0(\mathbb{Q})$ contains a subgroup of order 2^ℓ, then at a good prime the summand is at least $3, 5, 7$ for $\ell = 1, 2, 3$, respectively.

Proof. For a global minimal invariant differential ω_C, set

$$A(C) = \frac{i}{2}\int_{C(\mathbb{C})} \omega_C \wedge \overline{\omega_C}, \qquad \|f_C\|_{N_C}^2 = \langle f_C, f_C\rangle_{N_C}.$$

Integrating the identity $\phi_C^*\omega_C = 2\pi i c_C f_C(z)\,dz$ gives

$$\frac{m_C}{c_C^2} = \frac{4\pi^2\|f_C\|_{N_C}^2}{A(C)}. \tag{5.7}$$

Compare [19, (2.4)]. Curves E and F have the same complex j-invariant. The weight-twelve scaling of the discriminant and the weight-two scaling of the area give

$$\frac{A(F)}{A(E)} = \left|\frac{\Delta_E}{\Delta_F}\right|^{1/6}.$$

A quadratic-twist isomorphism scales minimal differentials by a rational multiple of $\sqrt{D}$, so this area ratio is rational. Since the quadratic twist is unramified at 2, the curves become isomorphic over an unramified extension of $\mathbb{Q}_2$; minimal discriminant valuations are unchanged under such an extension. Thus $v_2(\Delta_E) = v_2(\Delta_F)$ and the area ratio has two-adic valuation zero.

It remains to compare Petersson norms. Put $I_M = [\mathrm{SL}_2(\mathbb{Z}) : \Gamma_0(M)]$. Rankin–Selberg unfolding gives

$$\|g\|_M^2 = \frac{I_M}{48\pi}\operatorname*{Res}_{s=2}\sum_{n\geq 1}\frac{|a_n(g)|^2}{n^s}$$

for a normalized weight-two newform with trivial character. Indeed the Eisenstein series at the width-one cusp has residue $3/(\pi I_M)$, whereas unfolding contributes $\Gamma(s)/(4\pi)^s$ to the coefficient Dirichlet series. Only the local squared-coefficient factors at primes dividing D change; all other coefficients are multiplied locally by unramified signs.

At a good prime $p \mid D$, the recurrence for $A_j = a_{p^j}(F)$ gives

$$R_p(X) := \sum_{j\geq 0} A_j^2 X^j = \frac{1+pX}{(1-pX)((1+pX)^2 - a_p(F)^2X)}. \tag{5.8}$$

To obtain this identity, write $A_j = (\alpha^{j+1} - \beta^{j+1})/(\alpha - \beta)$, where $\alpha + \beta = a_p(F)$ and $\alpha\beta = p$, and sum three geometric series. The roots are distinct since $a_p(F)^2 \neq 4p$. After the twist the local squared-coefficient series is one. The level-index ratio is $p(p+1)$. Consequently the local norm ratio is

$$\frac{p(p+1)}{R_p(p^{-2})} = \frac{(p-1)((p+1)^2 - a_p(F)^2)}{p}.$$

At a multiplicative prime, $a_{p^j}(F)^2 = 1$, so $R_p(X) = (1-X)^{-1}$; the level-index ratio is p, giving

$$\frac{p}{R_p(p^{-2})} = \frac{p^2-1}{p}.$$

Multiplying these local ratios and using (5.7) proves (5.6), since all the denominators p are odd.

At an odd good prime the rational point of order two reduces injectively, so $a_p(F)$ is even. Each of $p-1, p+1-a_p(F), p+1+a_p(F)$ is then even. At a multiplicative prime p^2-1 is divisible by eight. This proves the lower bound three.

For the refinement, isogenous curves have the same a_p. The rational 2-primary subgroup of F_0 injects under reduction at the odd good prime, so

$$2^\ell \mid A := p + 1 - a_p(F), \qquad B := p + 1 + a_p(F) = 2(p+1) - A.$$

For $\ell = 2$, if $p \equiv 3 \pmod 4$, both A and B are divisible by four and $v_2(p-1) = 1$, giving at least five. If $p \equiv 1 \pmod 4$, both are again divisible by four and $v_2(p-1) \geq 2$, which is more than required. For $\ell = 3$, if $p \equiv 3 \pmod 4$, both A, B are divisible by eight, giving at least seven. If $p \equiv 1 \pmod 4$, then $v_2(2(p+1)) = 2$, so $v_2(B) = 2$; together with $v_2(A) \geq 3$ and $v_2(p-1) \geq 2$ this again gives seven. □

6. Integral differentials on quotients at two

We now treat curves with an odd additive prime and $v_2(N) \leq 1$. Throughout, the target E is optimal and has positive rank. Let a be the number of odd additive primes, let μ be the total number of multiplicative primes. Use the matrix K^* of Section 3, and put $k^* = \operatorname{rank} K^*$. For odd N this is K; when $2 \parallel N$, the extra row is the multiplicative prime two.

Lemma 6.1. *For any conductor N, let $E/\mathbb{Q}$ have positive rank and an odd additive prime, and let ϕ be a normalized modular parametrization. The positive Atkin–Lehner group $H = \ker \varepsilon$ acts on ϕ without translations. Thus*

$$\phi = \psi\pi, \qquad \pi : X_0(N) \longrightarrow Y = X_0(N)/H, \qquad \deg \phi = |H| \deg \psi.$$

In particular, for $v_2(N) \leq 1$, $v_2(|H|) \geq a + \mu - 1$.

Proof. Fix an odd additive prime p. Lemma 2.4 puts t_d in $E^0(\mathbb{Q}_p)$ whenever $p \nmid d$. If d belongs to H and $p \mid d$, use the complementary exact divisor and (1.2) to reach the same conclusion. On H, the translation is two-torsion. At an odd additive prime, the formal subgroup is pro-p and the smooth reduction is additive, so $E^0(\mathbb{Q}_p)[2] = 0$. Every translation is consequently zero. The quotient factorization and degree formula follow. There are $a + \mu$ bad primes when $v_2(N) \leq 1$, and ε has image of order at most two. □

Lemma 6.2. *Let R be a complete DVR of characteristic zero with algebraically closed residue field of characteristic two. Let a finite group act over R on a semistable relative curve with smooth generic fiber. The quotient is smooth at the image of a smooth point. At a node its completed local ring is either a node or a smooth ring.*

Proof. Let π be a uniformizer of R, and put $k = R/(\pi)$. It is enough to consider the stabilizer of a geometric special point, replacing it by its effective image on the local ring. At a smooth point its complete local ring is $B = R[[t]]$. For a stabilizer G of order n, set $u = \prod_{\sigma \in G} \sigma(t)$, subtracting its constant term. Modulo π, this has order n in t. Weierstrass preparation makes B finite free of rank n over $A = R[[u]]$. Artin's fixed-field theorem gives $[\operatorname{Frac} B : (\operatorname{Frac} B)^G] = n$. Thus the fixed field is $\operatorname{Frac} A$. Since A is normal and B is integral over it, $B^G = A$.

At a node write

$$B = R[[x, y]]/(xy - \pi^e).$$

First let G, of order n, preserve both branches. The equality of the Weil divisors of x and $\sigma(x)$ gives $\sigma(x) = x u_\sigma$ for a unit u_σ; similarly $\sigma(y) = y u_\sigma^{-1}$. For the norms $X = \prod_\sigma \sigma(x)$ and $Y = \prod_\sigma \sigma(y)$ one has $XY = \pi^{en}$. The ideal (π, X, Y) is primary to the maximal ideal of B, so B is finite over

$$A = R[[X, Y]]/(XY - \pi^{en}).$$

At the vertical prime (π, X) there is just one prime of B above it. The ramification index is one because π is a uniformizer in both localized DVRs. The residue extension $k((y))/k((Y))$ has degree n: the reduction of Y has order n in y, and Weierstrass preparation applies. Since B is

normal and finite over A, it is the integral closure of A in $\operatorname{Frac} B$. The finite extension of the localized DVRs has rank equal to the product of the ramification index and residue degree. The field degree is therefore n. As above, Artin's theorem and normality give $B^G = A$.

If the stabilizer exchanges the branches, first divide by its branch-preserving subgroup. The remaining involution τ exchanges the branches of a node B_0. The invariant $t = X + \tau(X)$ has order one on each branch, so $B_0/(\pi, t)$ has length two. It follows that B_0 is finite over $R[[t]]$. It is Cohen–Macaulay, and hence is free over this regular local ring; the displayed length makes its rank two. Artin's theorem and normality now give $B_0^\tau = R[[t]]$. These arguments prove the assertions, including inseparable residue-field extensions in the branch-preserving case. □

Proposition 6.3. *Assume $v_2(N) \leq 1$. Let α be a holomorphic characteristic-zero differential on $X_0(N)$ with the full Atkin–Lehner character ε. If its q-expansion at ∞ is integral over $\mathbb{Z}_2$, its descent to Y belongs to the Néron cotangent lattice of J_Y.*

Proof. The local model assertions can be checked after base change to the completion of the maximal unramified extension of $\mathbb{Q}_2$. Its residue field is algebraically closed, and integrality descends faithfully to $\mathbb{Z}_2$. If N is odd, $X_0(N)$ is smooth and proper over $\mathbb{Z}_2$. The quotient is smooth by Lemma 6.2, and its special fiber is geometrically integral, being a finite image of the geometrically integral special fiber upstairs. The distinct Atkin–Lehner translates of ∞ have distinct exact-divisor cusp types modulo two. Thus H acts freely at this special cusp, and the quotient is étale there. The integral q-expansion is therefore an integral expansion on Y. A possible denominator of a holomorphic differential is a vertical pole; on a smooth proper model with one special component it is detected at any smooth point of that component. There is no such pole. Hence the descended differential is integral, and it belongs to the cotangent lattice of the good-reduction Jacobian.

Suppose $N = 2M$, M odd. The regular semistable modular model has two principal components $X_0(M)_{\mathbb{F}_2}$, containing ∞ and $w_2\infty$, joined at supersingular points by chains of rational components. Each exceptional component has at most two nodal branches. This description and the identification of relative dualizing differentials with the Néron cotangent lattice hold also at $p = 2$; see [1, §5.2 and the proofs of Lemmas 5.19–5.20]. The group action extends to the regular model. Its quotient is semistable by Lemma 6.2; resolving thick nodes inserts further rational chains.

On the resulting regular semistable model Z of Y, call the images of the original two principal components principal, allowing the images to coincide. All remaining components are rational and have at most two nodal branches. Indeed, the normalization of the image of a rational component is rational, and a quotient node lies under a source node by Lemma 6.2. Its incident branches are images of the at most two incident branches upstairs. The special fiber is connected. The quotient is étale at the two distinguished cusps: odd Atkin–Lehner elements change their prime-to-two cusp type, and w_2 interchanges the principal components. The full eigencharacter hypothesis implies that the expansion at $w_2\infty$ is the same integral series up to sign. Thus the descended differential β is integral at a smooth point of every principal component.

If β were not integral on Z, choose $b > 0$ minimal such that $\gamma = 2^b\beta$ is a global relative dualizing differential. Its reduction is nonzero, since the kernel of reduction of this lattice is twice the lattice. Its reduction vanishes on each principal component, by the integral expansion of β at the corresponding cusp. On a semistable fiber a dualizing differential has at most simple poles at the nodal branches, with opposite residues on the two branches of a node. A rational component adjacent to a zero component has a zero residue there. It has at most two branches, so the residue theorem makes all its residues zero. Its differential is then holomorphic on $\mathbf{P}^1$, hence zero. Connectivity propagates the vanishing along every rational chain and tail, contradicting

the nonzero reduction of γ. Therefore β is integral on Z. The Néron/dualizing identification for a semistable Jacobian finishes the proof. □

Proposition 6.4. *Theorem 1.1 holds when $v_2(N) \le 1$.*

Proof. The semistable case is Proposition 4.6. Otherwise $a \ge 1$. Let $\kappa \in \{0,1\}$ be the dimension of the self-twist subgroup in $U = \ker K^*$. When $2 \parallel N$, multiplicative reduction excludes complex multiplication and $\kappa = 0$. Choose a linear complement to this subgroup in U, of dimension $h = a - k^* - \kappa$. By Lemma 5.2, a signed sum of the corresponding forms has the shape

$$G = f_E + g_\perp, \qquad G/2^h \in \mathbb{Z}_2[[q]]\, dq/q,$$

where $g_\perp$ is orthogonal to f_E, and every summand has the full Atkin–Lehner character ε. This remains valid for $h = 0$, with $G = f_E$. Let G_Y be the descent of G to Y. Proposition 6.3 puts $G_Y/2^h$ in the Néron cotangent lattice of J_Y. Equation (1.4) and Proposition 2.6 give $v_2(\deg \psi) \ge h$. Combining with Lemma 6.1, we obtain

$$v_2(\deg \phi) \ge a + \mu - 1 + a - k^* - \kappa = 2a + \mu - 1 - k^* - \kappa.$$

This is the rank bound in Proposition 3.3 for good or multiplicative reduction at two. □

7. Additive reduction at two

Assume from now on that $f = v_2(N) \ge 2$. Let a and m count the odd additive and odd multiplicative primes, respectively. The argument in this section applies a family of twists directly to the original curve when at least one odd prime has type III or III*. The case where every odd additive prime is removable is treated separately in Section 8.

7.1. The normalizer quotient.

Lemma 7.1. *If $8 \mid N$, put $s_2 : z \mapsto z + \frac{1}{2}$ and $W_2 = w_{2^f}$. Together with the odd Atkin–Lehner involutions they generate a group*

$$G = D_8 \times (\mathbb{Z}/2\mathbb{Z})^{\omega(N)-1}.$$

The group D_8 is dihedral of order eight. The character of G on f_E is nontrivial, with $\varepsilon(s_2) = -1$, and its positive kernel has order $2^{\omega(N)+1}$.

Proof. The matrix s_2 normalizes $\Gamma_0(N)$, has square one on $X_0(N)$, and acts on q by $q \mapsto -q$. It is defined over $\mathbb{Q}$. It commutes, modulo $\Gamma_0(N)$, with every odd Atkin–Lehner operator. These statements also follow by choosing the usual integral representatives; see [4]. To check the dyadic relation, replace W_2 temporarily by w_N, differing by the commuting odd involution $w_{N/2^f}$. With

$$s_2 = \begin{pmatrix} 1 & 1/2 \\ 0 & 1 \end{pmatrix}, \qquad w_N = \begin{pmatrix} 0 & -1 \\ N & 0 \end{pmatrix},$$

the determinant-one representative of $(s_2 w_N)^2$ is

$$C = \begin{pmatrix} N/4 - 1 & -1/2 \\ N/2 & -1 \end{pmatrix}.$$

It is not in $\Gamma_0(N)$, whereas

$$C^2 = \begin{pmatrix} N^2/16 - 3N/4 + 1 & 1 - N/8 \\ N(N/8 - 1) & 1 - N/4 \end{pmatrix} \in \Gamma_0(N).$$

Thus the product of s_2 and W_2 has exact order four, and the two involutions generate D_8.

The intersection with the odd Atkin–Lehner group is trivial. The images of ∞ under the eight dihedral elements have two-primary cusp-denominator valuations $f, 0, 1, f-1$, each occurring

twice. These four valuations are distinct for $f \geq 3$. Only $1, s_2$ preserve the two-primary type of ∞; a nonidentity odd Atkin–Lehner operator moves its odd cusp type, while s_2 fixes ∞ itself. This proves the product decomposition. Finally $a_{2n} = 0$, since $4 \mid N$, so $f_E|s_2 = -f_E$. Its character is nontrivial, and the stated order of the positive kernel follows. □

Proposition 7.2. *Let $E/\mathbb{Q}$ have positive rank. If $8 \mid N$, a normalized modular parametrization satisfies*

$$v_2(\deg \phi) \geq \omega(N).$$

If $8 \mid N$ and an odd additive prime exists, the full positive kernel H of Lemma 7.1 acts without translations, and

$$\phi = \psi\pi, \quad Y = X_0(N)/H, \quad v_2(\deg \pi) = a + m + 2. \tag{7.1}$$

If $4 \parallel N$ and $a > 0$, the ordinary Atkin–Lehner positive kernel gives a quotient with

$$v_2(\deg \pi) \geq a + m. \tag{7.2}$$

Proof. First suppose $8 \mid N$. Since s_2 fixes ∞, one has $\phi s_2 = -\phi$. For an odd Atkin–Lehner operator w, comparison of $\phi s_2 w$ and $\phi w s_2$ gives $2t_w = 0$, regardless of the sign of w. Lemma 2.4 at two gives $t_w \in E^0(\mathbb{Q}_2)$. The integral additive logarithm in Lemma 2.1 shows $\#E^0(\mathbb{Q}_2)[2^\infty] \leq 2$. By (1.2), $t_{w_N} = 0$. Writing $w_N = W_2 w_{N/2^f}$ therefore puts t_{W_2} in the same group $E^0(\mathbb{Q}_2)[2]$. The translation image of the positive kernel has order at most two. Its kernel has order at least $2^{\omega(N)}$, proving the first bound.

Suppose that an odd additive prime p exists. The relations $W_2^2 = 1$ and $(s_2W_2)^4 = 1$ show that t_{W_2} is killed by four; if $\varepsilon(W_2) = 1$, it is killed by two. The exact divisor of W_2 is prime to p, so Lemma 2.4 puts t_{W_2} in $E^0(\mathbb{Q}_p)$. This group has no two-primary torsion, and $t_{W_2} = 0$. Hence all dyadic generators act with zero translation.

Choose $h_0 = w_N$ or $s_2 w_N$, whichever belongs to H. It has zero translation and its odd label contains all the odd bad primes. For $h \in H$, write $h = uw_d$, with $u \in D_8$ and d odd. The vanishing of dyadic translations gives $t_h = \varepsilon(u)t_d$. If $p \nmid d$, Lemma 2.4 puts t_h in $E^0(\mathbb{Q}_p)$, and it is zero. If the label contains p, apply the same argument to hh_0; its odd label is complementary and its translation is still t_h. Thus the whole of H acts without translations. Its order is 2^{a+m+2}, giving (7.1). For $f = 2$, Lemma 6.1 applies to the ordinary Atkin–Lehner group, whose positive kernel has order at least 2^{a+m}. □

7.2. **Realizing the forms by maps.** Suppose some odd additive prime p_0 has type III or III*. Let S be the set of odd additive primes and $V = \mathbb{F}_2^S$. Use the discriminants D_v defined in Section 5, and put

$$L = \mathbb{Q}(\sqrt{D_v} : v \in V).$$

Every completion of L at two is unramified. At p_0 the ramification index is at most two: all the other generators are unramified there. Tame inertia of type III or III* has order four. It has order two after this ramification. Thus E_L remains additive at p_0.

Let $U = \ker K^*$, with K^* and k^* as in (3.3). The forms F_v, $v \in U$, of Section 5 have the same full character as f_E. All their even-index coefficients vanish. When $8 \mid N$, they also have the same character for G, because s_2 acts by minus one.

Lemma 7.3. *For every $v \in V$, the normalized old/new form F_v can be realized by a map $\Phi_v : X_0(N)_L \to E_L$, normalized at ∞, with*

$$\Phi_v^* \omega_E = \lambda_v F_v, \qquad \lambda_v = r_v\sqrt{D_v}, \quad r_v \in \mathbb{Q}^\times, \quad v_2(r_v) = 0.$$

Proof. The curve $E^{(D_v)}$, and its optimal representative, retain an odd additive prime. Proposition 3.7 therefore rules out nonunit two-isogeny steps from its optimal representative. Proposition 2.6 then gives an odd Manin constant for its minimal parametrization. The oldvectors in

Lemma 5.1 are realized by integer linear combinations of degeneracy maps after clearing only odd denominators: the pullback under $z \mapsto dz$ of $g(q)dq/q$ is $d\,g(q^d)dq/q$, and every such d is odd. The sum of these maps to the elliptic curve is again a map. Translate its value at ∞ to zero. The twist isomorphism $E_L^{(D_v)} \to E_L$ changes minimal differentials by a rational multiple of $\sqrt{D_v}$, which is a unit at every place over two. This proves the assertion. □

Lemma 7.4. *Let $P : X_0(N)_L \to E_L$ be any normalized map whose differential has the common character above. For the quotient Y in Proposition 7.2, the differential of $[2]P$ descends to an integral Néron cotangent differential on J_Y at every place over two.*

Proof. If $f \geq 3$, the relation $Ps_2 = -P$ makes all odd Atkin–Lehner translations two-torsion. The dyadic relation makes t_{W_2} killed by four. At the surviving additive prime p_0, choose a place $w \mid p_0$ of L and apply Lemma 2.4 over L_w. It puts t_{W_2} in $E_L^0(L_w)$, where there is no two-primary torsion. Thus this translation is zero. All translations on H are consequently killed by two, and $[2]P$ is a map from Y_L to E_L.

For $f = 2$, we chose only the ordinary Atkin–Lehner positive kernel. Every one of its elements is an involution with positive character, so its translation is killed by two directly. Thus $[2]P$ again descends.

The induced homomorphism $J_Y \to E_L$ extends to Néron models. Pullback of a minimal differential is therefore integral, as asserted. □

7.3. **An affine induction on the quotient lattice.**

Proposition 7.5. *Let $A = v_0 + U_0$ be any affine subspace of the common-character twist space, of dimension h, and let $\rho : A \to \{\pm 1\}$ be an affine sign character. Put*

$$G_A = \sum_{v \in A} \rho(v) F_v.$$

The descended differential satisfies

$$\frac{G_A}{2^{h-1}} \in \Omega(J_Y) \otimes \mathbb{Z}_2. \tag{7.3}$$

For $h = 0$, the assertion means that $2F_{v_0}$ is integral.

Proof. Work first at a fixed embedding $L \hookrightarrow \overline{\mathbb{Q}}_2$ with unramified valuation ring R. Minimal differentials remain minimal under this base change. By Lemmas 7.3 and 7.4, $2\lambda_v F_v$ is integral on Y. Since λ_v is a unit, this proves the dimension-zero assertion.

Induct simultaneously on all affine spaces and characters. Normalize by the unit λ_{v_0}, and put $\beta_v = \lambda_v/\lambda_{v_0}$. If $v = v_0 + t$, its Galois conjugates satisfy

$$\sigma(\beta_v) = \chi_t(\sigma)\beta_v. \tag{7.4}$$

Here $\chi_t(\sigma)$ denotes the sign of the quadratic Galois character associated with the relative twist. For $\sigma \in \mathrm{Gal}(L/\mathbb{Q})$, the function

$$\rho_\sigma(v) = \rho(v)\chi_{v-v_0}(\sigma)$$

is again an affine sign character on A.

First suppose that all relative twists split at two. Then $\beta_v \in \mathbb{Z}_2^\times$ in the chosen embedding. Choose integers n_v such that $n_v\beta_v \equiv \rho(v) \pmod{2^h R}$. Put $e_v = (n_v\beta_v - \rho(v))/2^h \in R$ and $E_A = \sum_{v \in A} e_v F_v$. For the sum of maps $P = \sum_v [n_v]\Phi_v$,

$$\frac{P^*\omega_E}{\lambda_{v_0}} = G_A + 2^h E_A, \qquad E_A \in \sum_{v \in A} RF_v. \tag{7.5}$$

Lemma 5.2 makes this differential divisible coefficientwise by 2^h. For every $\sigma \in \mathrm{Gal}(L/\mathbb{Q})$, (7.4) gives the equality in the fixed embedding

$$\sigma\left(\frac{P^*\omega_E}{\lambda_{v_0}}\right) = \sum_{v\in A} \rho_\sigma(v)F_v + 2^h \sum_{v\in A} \chi_{v-v_0}(\sigma)e_v F_v.$$

The first sum is divisible by 2^h by Lemma 5.2; the second sum has coefficients in R. Every conjugate of λ_{v_0} is a unit. Thus $P^*\omega_E$ is divisible coefficientwise by 2^h at every place above two. Its even-index coefficients vanish. Proposition 2.7 supplies an actual normalized map Ψ with $[2^h]\Psi = P$. Its differential still has the common character, so the descent argument applies anew to this divided map. Lemma 7.4 makes

$$\frac{P^*\omega_E}{\lambda_{v_0} 2^{h-1}}$$

integral on Y. Each $2F_v$ is integral by the base case, so $2E_A$ is integral too. Subtracting it in (7.5) proves the claim.

If the relative character at two is nontrivial, put $A_0 = \{v_0 + t : t \in U_0,\ D_t \in \mathbb{Q}_2^{\times 2}\}$ and $A_1 = A \setminus A_0$. These are parallel affine hyperplanes of A, with $v_0 \in A_0$. Write

$$G_i = \sum_{v\in A_i} \rho(v)F_v, \qquad G_A = G_0 + G_1.$$

Choose $t_1 \in U_0$ such that D_{t_1} is nonsquare in $\mathbb{Q}_2$, and put $\delta = D_{t_1}$, $s = \sqrt{D_{t_1}} \in L$. Since every p^* is congruent to one modulo four, $\delta \in 1 + 4\mathbb{Z}_2$. Then $u = (s-1)/2$ satisfies

$$u^2 + u + \frac{1-\delta}{4} = 0.$$

This monic equation has coefficients in $\mathbb{Z}_2$, so $u \in R$ and $s = 1 + 2u$. For $v \in A_i$, $\beta_v/s^i \in \mathbb{Z}_2^\times$. Choose integers n_v with $n_v\beta_v/s^i \equiv \rho(v) \pmod{2^h R}$, and put

$$e_v = \frac{n_v\beta_v - \rho(v)s^i}{2^h} \in R \quad (v \in A_i), \qquad E_A = \sum_{v\in A} e_v F_v.$$

For $P = \sum_v [n_v]\Phi_v$,

$$B := \frac{P^*\omega_E}{\lambda_{v_0}} = G_0 + sG_1 + 2^h E_A = G_A + 2uG_1 + 2^h E_A. \tag{7.6}$$

The full affine Fourier sum is divisible by 2^h, and the half sum G_1 by 2^{h-1}. Thus B is divisible by 2^h. The same holds at every place over two. Indeed, for each $\sigma \in \mathrm{Gal}(L/\mathbb{Q})$, put

$$E_\sigma = \sum_{v\in A} \chi_{v-v_0}(\sigma)e_v F_v.$$

Its coefficients lie in R, and (7.4) gives

$$\begin{aligned}\sigma(B) &= \sum_{v\in A_0} \rho_\sigma(v)F_v + s\sum_{v\in A_1} \rho_\sigma(v)F_v + 2^h E_\sigma \\ &= \sum_{v\in A} \rho_\sigma(v)F_v + 2u\sum_{v\in A_1} \rho_\sigma(v)F_v + 2^h E_\sigma.\end{aligned}$$

The full and half sums are divisible by 2^h and 2^{h-1}, respectively. Since $u \in R$ and $\sigma(\lambda_{v_0}) \in R^\times$, the required divisibility holds for every conjugate of $P^*\omega_E$.

Apply map division and Lemma 7.4 again. The divided map is normalized and has the same differential character. Then $B/2^{h-1}$ is integral on Y. The induction hypothesis on A_1 gives

$G_1/2^{h-2}$ integral on Y, including $h = 1$, when this is precisely the base case $2G_1$. Equation (7.6) now yields

$$\frac{G_A}{2^{h-1}} = \frac{B}{2^{h-1}} - u\,\frac{G_1}{2^{h-2}} - 2E_A,$$

which is integral. This proves the induction. It proves the assertion at every unramified completion above two. Since the differential in (7.3) is rational, integrality descends faithfully to the lattice over $\mathbb{Z}_2$. □

Proposition 7.6. *Theorem 1.1 holds if E is additive at two and has an odd prime of type* III *or* III*.

Proof. Use the optimal positive-rank reduction and the quotient in Proposition 7.2. The common-character space has dimension $a - k^*$. If its unique possible imaginary self-twist is present, remove that line by choosing a complement, and write $\kappa = 1$; otherwise put $\kappa = 0$. The remaining dimension is $h = a - k^* - \kappa$. By Lemma 5.3, the forms now have pairwise distinct eigensystems, and a character sum is $f_E + g_\perp$, with $g_\perp$ orthogonal to f_E. Proposition 7.5, (1.4), and the oddness of c_E give

$$v_2(\deg\psi) \geq \max(0, a - k^* - \kappa - 1).$$

For $f \geq 3$, the resulting valuation bound is

$$\begin{aligned} v_2(\deg\phi) &\geq a + m + 2 + \max(0, a - k^* - \kappa - 1) \\ &\geq 2a + m + 1 - k^* - \kappa. \end{aligned}$$

For even f, $k^* = k$, and the rank bound $2a + m + d - 1 - k - \kappa$, $d \leq 2$, is no larger. For odd f and $d = 2$, use the augmented rank bound $2a + m + 1 - k^* - \kappa$. For odd f and $d \leq 1$, use

$$r \leq 2a + m - k - \kappa \leq 2a + m + 1 - k^* - \kappa,$$

since $k^* \leq k + 1$. For $f = 2$, the ordinary quotient gives instead

$$v_2(\deg\phi) \geq a + m + \max(0, a - k - \kappa - 1) \geq 2a + m - 1 - k - \kappa.$$

The last expression is the rank bound after both dyadic valuation coordinates have been removed. □

8. Removal of the remaining additive primes and completion

It remains to treat additive reduction at two with no odd prime of type III or III*. We first record a degree bound for the untwisted curve.

Lemma 8.1. *Let $F_0/\mathbb{Q}$ be optimal, have nonzero rational two-torsion, be additive at two, and be semistable at every odd prime. Let b be the number of its odd multiplicative primes. If $v_2(N_{F_0}) \geq 3$, there is an integer $0 \leq \ell \leq 3$ such that*

$$2^\ell \mid \#F_0(\mathbb{Q})_{\mathrm{tors}}, \qquad v_2(m_{F_0}) \geq b + 2 - \ell.$$

If $v_2(N_{F_0}) = 2$, then

$$v_2(m_{F_0}) \geq b - 1.$$

Proof. For $v_2(N_{F_0}) \geq 3$, the normalizer positive kernel has order 2^{b+2}, by Lemma 7.1. All odd Atkin–Lehner translations are in the common group $F_0^0(\mathbb{Q}_2)[2]$, of order at most two: Lemma 2.4 applies at two, and their commutation with s_2 kills the translations by two. The relation $(s_2W_2)^4 = 1$ kills the dyadic translation by four. Thus the translation image on the positive kernel is a subgroup of $F_0(\mathbb{Q})$ of order 2^ℓ with $\ell \leq 3$. Factoring through its actual kernel gives $v_2(m_{F_0}) \geq b + 2 - \ell$.

For $v_2(N_{F_0}) = 2$, the ordinary Atkin–Lehner positive kernel has order at least 2^b. Its translation image is contained in $F_0(\mathbb{Q})[2]$. The conductor-two-exponent argument of Section 3 gives an odd component group at two; together with Lemma 2.1, this excludes full rational two-torsion. Its translation image therefore has order at most two, which gives the stated bound. □

Proposition 8.2. *Theorem 1.1 holds if E is additive at two and all its odd additive primes are removable by quadratic twisting.*

Proof. Take E optimal and of positive rank. Let m count its odd multiplicative primes and put $f = v_2(N)$. First assume that there are no odd additive primes. If $f \geq 3$, Proposition 7.2 gives

$$v_2(m_E) \geq m + 1, \qquad r \leq m + d - 1 \leq m + 1$$

by Proposition 3.3. If $f = 2$, the same rank proposition gives $r \leq m - 1$, and the ordinary positive Atkin–Lehner kernel loses at most one translation factor, as in Lemma 8.1. This also suffices.

Now suppose that at least one odd additive prime occurs. Twist by the product D of p^* over all these primes, and call the resulting particular curve $F = E^{(D)}$. Let g count those primes where F is good, and t those where it is multiplicative; thus $g + t \geq 1$. Let F_0 be the optimal curve in the isogeny class of F. It has $m + t$ odd multiplicative primes and no odd additive primes. The twist is unramified at two, so f is unchanged. We compare the degree of the original optimal target with the rank of $F_0^{(D)}$; the latter has the same rank as E.

Put $V(C) = v_2(m_C/c_C^2)$. The nonoptimal comparison of Proposition 3.7 and the exact twist formula of Proposition 5.4 give

$$v_2(m_E) = V(E) \geq v_2(m_{F_0}) - 1 + \sum_{\text{good } p} v_2((p-1)((p+1)^2 - a_p(F)^2)) + \sum_{\text{mult } p} v_2(p^2 - 1). \quad (8.1)$$

The sums are over $p \mid D$, grouped by the reduction of F. Each local gain is at least three. At a good prime, $a_p(F) = a_p(F_0)$ by isogeny invariance; thus rational torsion on F_0 supplies the refined gains in this same formula. The term -1 is the loss in passing from F_0 to F.

First take $f \geq 3$. Lemma 8.1, with $b = m + t$, gives

$$v_2(m_{F_0}) \geq m + t + 2 - \ell, \qquad 0 \leq \ell \leq 3.$$

The elementary descent bound on $F_0^{(D)}$, discarding the nonnegative matrix-rank defect, is

$$r \leq m + 2g + 2t + d - 1, \qquad d \leq 2. \quad (8.2)$$

If $\ell \leq 1$, the difference between the degree bound in (8.1) and the right side of (8.2) is at least

$$g + 2t + 2 - \ell - d \geq 0.$$

If $\ell = 2$, rational torsion of order at least four improves each good-prime gain from three to five. The difference is then at least

$$3g + 2t - d \geq 0.$$

If $\ell = 3$, there are at least eight rational two-primary torsion points on F_0. Lemma 3.8 permits the descent point to be chosen in the identity component at two, and gives $d \leq 1$ on the unramified twist $F_0^{(D)}$. Each good-prime gain is at least seven, so the difference is at least

$$5g + 2t - 2 \geq 0.$$

All three inequalities hold for every $g, t \geq 0$ with $g + t \geq 1$.

Finally, let $f = 2$. The untwist bound is $v_2(m_{F_0}) \geq m + t - 1$. Allowing the loss one and the gains three gives

$$v_2(m_E) \geq m + 3g + 4t - 2.$$

Both dyadic valuation coordinates are absent from descent, so

$$r \le m + 2g + 2t - 1.$$

Their difference is $g + 2t - 1 \ge 0$. □

Proof of Theorem 1.1. Lemma 1.2 reduces the theorem to an optimal positive-rank target. If $v_2(N) \le 1$, apply Proposition 6.4. If $v_2(N) \ge 2$, every odd additive prime has one of the types classified in Section 5. If a type III or III* prime occurs, Proposition 7.6 applies. Otherwise all odd additive primes are removable, including the possibility that none occurs, and Proposition 8.2 applies. These alternatives exhaust the reduction types and give $r \le v_2(m_E)$ for the optimal curve. Lemma 1.2 gives the assertion for every rational isogeny target and every parametrization. □

References

[1] A. Agashe, K. A. Ribet, and W. A. Stein, *The modular degree, congruence primes, and multiplicity one*, in *Number Theory, Analysis and Geometry: In Memory of Serge Lang*, Springer, New York, 2012, 19–49. doi:10.1007/978-1-4614-1260-1_2.

[2] P. B. Allen, *Powers of* 2 *in the modular degree via* 2*-adic* $R = T$, preprint, `https://patrick-allen.github.io/research/watkins.pdf`, accessed November 10, 2026.

[3] A. O. L. Atkin and W.-C. W. Li, *Twists of newforms and pseudo-eigenvalues of* W*-operators*, Invent. Math. **48** (1978), no. 3, 221–243.

[4] F. Bars, *The group structure of the normalizer of* $\Gamma_0(N)$ *after Atkin–Lehner*, preprint, 2007, arXiv:math/0701636.

[5] S. Bhakta and S. Krishnamoorthy, *Watkins's conjecture for elliptic curves with a rational torsion*, preprint, 2024, arXiv:2407.17680.

[6] S. Bhakta, S. Krishnamoorthy, and S. K. Pasupulati, *Modular degree and a conjecture of Watkins*, preprint, 2023, revised 2025, arXiv:2308.13708v3.

[7] S. Bosch, W. Lütkebohmert, and M. Raynaud, *Néron Models*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3), vol. 21, Springer-Verlag, Berlin, 1990.

[8] C. Breuil, B. Conrad, F. Diamond, and R. Taylor, *On the modularity of elliptic curves over* $\mathbb{Q}$*: wild* 3*-adic exercises*, J. Amer. Math. Soc. **14** (2001), no. 4, 843–939.

[9] F. Calegari, V. Dimitrov, and Y. Tang, *The unbounded denominators conjecture*, J. Amer. Math. Soc. **38** (2025), no. 3, 627–702. doi:10.1090/jams/1053.

[10] F. Calegari and M. Emerton, *Elliptic curves of odd modular degree*, Israel J. Math. **169** (2009), 417–444.

[11] J. Caro, *Watkins's conjecture for quadratic twists of elliptic curves with prime power conductor*, J. Théor. Nombres Bordeaux **37** (2025), no. 2, 647–663. doi:10.5802/jtnb.1335.

[12] J. Caro and H. Pasten, *Watkins's conjecture for elliptic curves with non-split multiplicative reduction*, Proc. Amer. Math. Soc. **150** (2022), no. 8, 3245–3251. doi:10.1090/proc/15942.

[13] K. Česnavičius, *The Manin constant in the semistable case*, Compos. Math. **154** (2018), no. 9, 1889–1920.

[14] Z. Chi, S. Liu, and P. Nie, *On the Manin constant*, preprint, 2026, arXiv:2610.25672v1.

[15] D. A. Cox, *Primes of the Form* $x^2 + ny^2$*: Fermat, Class Field Theory, and Complex Multiplication*, 2nd ed., John Wiley & Sons, Hoboken, NJ, 2013.

[16] V. G. Drinfeld, *Two theorems on modular curves*, Funct. Anal. Appl. **7** (1973), 155–156.

[17] N. Dummigan, *On a conjecture of Watkins*, J. Théor. Nombres Bordeaux **18** (2006), no. 2, 345–355.

[18] N. Dummigan and S. Krishnamoorthy, *Powers of* 2 *in modular degrees of modular abelian varieties*, J. Number Theory **133** (2013), no. 2, 501–522.

[19] J. A. Esparza-Lozano and H. Pasten, *A conjecture of Watkins for quadratic twists*, Proc. Amer. Math. Soc. **149** (2021), no. 6, 2381–2385. doi:10.1090/proc/15376.

[20] S. D. Galbraith, *Rational points on* $X_0^+(N)$ *and quadratic* $\mathbb{Q}$*-curves*, J. Théor. Nombres Bordeaux **14** (2002), no. 1, 205–219; expanded author version, February 1, 2011, `https://www.math.auckland.ac.nz/~sgal018/x0n.pdf`.

[21] J. Hatley and D. Kundu, *Elliptic curves of conductor* $2^m p$*, quadratic twists, and Watkins' conjecture*, Ann. Math. Qué. **50** (2026), 143–165. doi:10.1007/s40316-025-00249-8.

[22] N. M. Katz and B. Mazur, *Arithmetic Moduli of Elliptic Curves*, Annals of Mathematics Studies, vol. 108, Princeton University Press, Princeton, NJ, 1985.

[23] M. Kazalicki and D. Kohen, *On a special case of Watkins' conjecture*, Proc. Amer. Math. Soc. **146** (2018), no. 2, 541–545.

[24] M. Kazalicki and D. Kohen, *Corrigendum to "On a special case of Watkins' conjecture"*, Proc. Amer. Math. Soc. **147** (2019), no. 10, 4563.
[25] M. Kazalicki and S. Slijepčević, *Gross vectors modulo* 2 *and elliptic curves of prime conductor*, preprint, 2026, arXiv:2608.13371.
[26] V. A. Kolyvagin, *Finiteness of* $E(\mathbb{Q})$ *and* $\mathrm{Sha}(E,\mathbb{Q})$ *for a subclass of Weil curves*, Math. USSR-Izv. **32** (1989), no. 3, 523–541.
[27] W.-C. W. Li, *Newforms and functional equations*, Math. Ann. **212** (1975), no. 4, 285–315.
[28] Yu. I. Manin, *Parabolic points and zeta-functions of modular curves*, Math. USSR-Izv. **6** (1972), 19–64.
[29] J. S. Milne, *Class Field Theory*, version 4.03, August 6, 2020, https://www.jmilne.org/math/CourseNotes/CFT.pdf.
[30] D. E. Rohrlich, *Galois theory, elliptic curves, and root numbers*, Compos. Math. **100** (1996), no. 3, 311–349.
[31] J. H. Silverman, *The Arithmetic of Elliptic Curves*, 2nd ed., Graduate Texts in Mathematics, vol. 106, Springer, Dordrecht, 2009.
[32] W. A. Stein and M. Watkins, *Modular parametrizations of Neumann–Setzer elliptic curves*, Int. Math. Res. Not. **2004** (2004), no. 27, 1395–1405. doi:10.1155/S1073792804133916.
[33] M. Watkins, *Computing the modular degree of an elliptic curve*, Experiment. Math. **11** (2002), no. 4, 487–502.
[34] K. Wohlfahrt, *An extension of F. Klein's level concept*, Illinois J. Math. **8** (1964), 529–535.
[35] S. Yazdani, *Modular abelian varieties of odd modular degree*, Algebra Number Theory **5** (2011), no. 1, 37–62.

Zhili College, Tsinghua University, Beijing 100084, China
Email address: tangqy24@mails.tsinghua.edu.cn