Math. Res. Lett.
Volume 30, Number 2, 413–461, 2023

Local and global densities for Weierstrass models of elliptic curves

John E. Cremona and Mohammad Sadek

In memory of John Tate, 1925–2019

We prove local results on the p-adic density of elliptic curves over $\mathbb{Q}_p$ with different reduction types, together with global results on densities of elliptic curves over $\mathbb{Q}$ with specified reduction types at one or more (including infinitely many) primes. These global results include: the density of integral Weierstrass equations which are minimal models of semistable elliptic curves over $\mathbb{Q}$ (that is, elliptic curves with square-free conductor) is $1/\zeta(2) \approx 60.79\%$, the same as the density of square-free integers; the density of semistable elliptic curves over $\mathbb{Q}$ is $\zeta(10)/\zeta(2) \approx 60.85\%$; the density of integral Weierstrass equations which have square-free discriminant is $\prod_p \left(1 - \frac{2}{p^2} + \frac{1}{p^3}\right) \approx 42.89\%$, which is the same (except for a different factor at the prime 2) as the density of monic integral cubic polynomials with square-free discriminant (and agrees with a 2013 result of Baier and Browning for short Weierstrass equations); and the density of elliptic curves over $\mathbb{Q}$ with square-free minimal discriminant is $\zeta(10) \prod_p \left(1 - \frac{2}{p^2} + \frac{1}{p^3}\right) \approx 42.93\%$.

The local results derive from a detailed analysis of Tate's Algorithm, while the global ones are obtained through the use of the Ekedahl Sieve, as developed by Poonen, Stoll, and Bhargava.

1. Introduction

In this paper we first study purely local results on the p-adic density of elliptic curves over $\mathbb{Q}_p$ with different reduction types, and then apply these, using a version of the Ekedahl Sieve, to establish several global results on densities of elliptic curves over $\mathbb{Q}$.

In the local setting, we use Tate's Algorithm [29] to determine in Section 2 the local density of Weierstrass equations having each possible reduction type. For example, the proportion of Weierstrass equations over $\mathbb{Z}_p$ which have good reduction (at p) is $1 - 1/p$, those with reduction of type I_m (respectively I_m^*) have density $(p-1)^2/p^{m+2}$ (respectively $(p-1)^2/p^{m+7}$), and the density of elliptic curves over $\mathbb{Q}_p$ which are semistable is $(1-p^{-2})/(1-p^{-10})$. See Propositions 2.2 and 2.5 for details. Here we distinguish between the set of local integral Weierstrass equations with some property, and the larger set of those which may not be minimal models but define elliptic curves whose minimal model has the property. For example, the density of integral Weierstrass equations defining elliptic curves with good reduction is $(1-p^{-1})/(1-p^{-10})$, which is greater than the density $1-p^{-1}$ of equations which are themselves minimal models of curves with good reduction, after allowing for non-minimal models, as the local density of non-minimal Weierstrass equations is p^{-10}.

We show that the local densities of minimal Weierstrass equations with prime conductor and prime discriminant are, respectively, $(p-1)/p^2$ and $(p-1)^2/p^3$.

The local results mentioned so far all generalise immediately to any p-adic field, replacing p in each formula with the cardinality of the residue field.

Further local results over $\mathbb{Q}_p$ are obtained in Section 5, again by studying Tate's Algorithm in great detail. In Theorems 5.3 and 5.6, we establish the densities of elliptic curves over $\mathbb{Q}_p$ with each possible conductor exponent and each possible Tamagawa number. (Note that Tate's Algorithm in [29] includes the determination of both these quantities.) For example (see Theorem 5.6), among elliptic curves over $\mathbb{Q}_3$ with additive reduction the densities of the possible conductor exponents $f_3 = 2, 3, 4, 5$ are in the ratio 189 : 366 : 122 : 61 or approximately 25.6% : 49.6% : 16.5% : 8.3%. Extending these results to general extensions of $\mathbb{Q}_p$ is not so straightforward, as the analysis depends on the precise valuations of certain integers (such as the coefficients of the discriminant of a long Weierstrass equation).

In order to pass from local results to global statements, we make use of a version of the Ekedahl Sieve from [16] as developed by Poonen and Stoll in [25] and further by Bhargava in [6], by Bhargava, Shankar and Wang in [9], and elsewhere. Provided that certain conditions are met, it is often the case that global densities may be expressed as a convergent infinite product (over all primes) of local densities. In order to be able to apply these methods with some flexibility, we develop them systematically in Section 3.

The global results, for elliptic curves over $\mathbb{Q}$, follow in Section 4. For a set S of Weierstrass equations with integer coefficients $\boldsymbol{a} = (a_1, a_2, a_3, a_4, a_6) \in \mathbb{Z}^5$, we define the *weighted density* of S to be

$$\rho^{\mathbf{k}}(S) = \lim_{X\to\infty} \frac{\#\{\boldsymbol{a} \in S \mid |a_i| \leq X^i\}}{\#\{\boldsymbol{a} \in \mathbb{Z}^5 \mid |a_i| \leq X^i\}}, \tag{1}$$

when this limit exists, where $\mathbf{k} = (1, 2, 3, 4, 6)$. (More general weighted densities will be defined in subsection 3.1 below.) An alternative way of expressing density results is to define the *height* of a Weierstrass equation with integer coefficients $\boldsymbol{a}$ to be

$$\mathrm{ht}(\boldsymbol{a}) = \max_i |a_i|^{1/i},$$

and then order such equations by height; then we may say that when integral Weierstrass equations are ordered by height, the proportion which lie in the set S is $\rho^{\mathbf{k}}(S)$, whose definition may now be written as

$$\rho^{\mathbf{k}}(S) = \lim_{X\to\infty} \frac{\#\{\boldsymbol{a} \in S \mid \mathrm{ht}(\boldsymbol{a}) \leq X\}}{\#\{\boldsymbol{a} \in \mathbb{Z}^5 \mid \mathrm{ht}(\boldsymbol{a}) \leq X\}}.$$

Each of our results will have two versions, depending on whether we restrict to Weierstrass equations which are globally minimal, or include all equations. (Those with zero discriminant, which define singular curves, may always be ignored as they form a set of measure zero.)

In general, the global density exists and equals the product of the corresponding local densities, provided that the local condition specified at all but finitely many primes is to have good or multiplicative reduction. We state here a summary of the results from Section 4, which allow more flexibility in specifying local conditions at any finite set of primes.

Theorem 1.1. *When ordered by height, the proportion of integral Weierstrass equations with each of the following properties is as given:*

- *globally minimal:* $1/\zeta(10) = 93555/\pi^{10} \approx 99.9\%$*;*
- *minimal models of semistable elliptic curves:* $1/\zeta(2) = 6/\pi^2 \approx 60.8\%$*;*
- *minimal models of semistable elliptic curves with good reduction at all the primes in the finite set* S*:* $\zeta(2)^{-1}\prod_{p\in S}\frac{p}{p+1}$*;*
- *minimal models of elliptic curves with square-free discriminant:*

$$\prod_p \left(1 - \frac{2}{p^2} + \frac{1}{p^3}\right) \approx 42.9\%.$$

In each case, the proportion of integral Weierstrass equations which are not necessarily minimal models of elliptic curves with the stated property may be obtained by multiplying by $\zeta(10) \approx 1.001$*.*

It would be interesting to extend the global results here to number fields other than $\mathbb{Q}$, which would entail several additional challenges.

1.1. Related work

Our result for the density of integral Weierstrass equations which have square-free discriminant is—apart from a different local factor at 2—the same as the density of monic integral cubic polynomials with square-free discriminant: see the 2016 paper [9] of Bhargava *et al.*, and also Theorem 6.8 in the 2007 paper [1] by Ash, Brakenhoff, and Zarrabi. We note that this is also in agreement with a result of Baier and Browning in their 2013 paper [3] (see also Baier's 2016 paper [2]) for short Weierstrass equations $Y^2 = X^3 + AX + B$ with squarefree discriminant, established using quite different methods.

In their famous 1990 paper [10], Brumer and McGuinness give heuristics for the number of elliptic curves whose minimal discriminant is less than X, separating the cases of positive and negative discriminant. In each case the number is conjectured to be a constant multiple of $X^{5/6}$ with a constant which is the value of an elliptic integral divided (in each case) by $\zeta(10)$, the latter to allow for non-minimal discriminants. This was revisited by Watkins in 2008 in [30], who re-derives the same heuristic estimate, and also discusses the factor $\zeta(10)$. Watkins also gives applications to the distribution of curves by conductor instead of discriminant, and also to the distribution of odd and even analytic ranks.

Some similar results, including local conditions, are given in the 2001 paper [31] of Wong, who defines the height of an elliptic curve over $\mathbb{Q}$ to be

$$\mathrm{ht}_c(E) = \max\{|c_4(E)|^{1/4}, |c_6(E)|^{1/6}\},$$

where $c_4(E), c_6(E)$ are the invariants of a minimal model for E. This is comparable with our height: in one direction, standard formulae for c_4, c_6 imply that, for E defined by a minimal Weierstrass equation with coefficients $\boldsymbol{a}$, we have $\mathrm{ht}_c(E) \ll \mathrm{ht}(\boldsymbol{a})$. In the other direction, given a pair (c_4, c_6) which satisfy Kraus's conditions from [21], using the formulas in the first author's book [12, p. 61] to recover Weierstrass coefficients $\boldsymbol{a}$ from these, one obtains $\mathrm{ht}(\boldsymbol{a}) \ll \mathrm{ht}_c(E)$. In Theorem 1 of [31], Wong gives asymptotic expansions of the number of curves of height up to X together with the number which are semistable, and the number which are semistable and have good reduction at both 2 and 3. In each case, the leading coefficient gives the value of the density in our sense. To compare these with our results, we first need to take into account the density of (c_4, c_6) pairs which satisfy Kraus's conditions, which may easily be seen to be $2^{-7}3^{-3} = 1/3456$, and also the minimality condition which leads to a factor of $1/\zeta(10) = 9355/\pi^{10} = 3^5 \cdot 5 \cdot 7 \cdot 11/\pi^{10}$ as in our theorem above. The density given in [31] is a rational multiple of $1/\pi^{10}$, but with a different rational factor. We would also expect, from our theorem above, that the density for semistable curves should be multiplied by $1/\zeta(2) = 6/\pi^2$, and that if in addition we impose the condition of having good reduction at 2 and 3, the density should be multiplied by $(2/3)(3/4) = 1/2$, rather than $7/9$ as in [31]. These discrepancies lead to Wong's statement that the proportion of semistable curves is 17.9%, compared with our value of 60.85%. We should emphasize that the majority of Wong's results in [31] do not depend on precise values of any densities, only that they exist and are positive.

Over general number fields, not all elliptic curves have global minimal Weierstrass equations when the class group is non-trivial. In her 2004 paper [4], Bekyel determined the density of elliptic curves defined over any number field K which have global minimal models to be $\zeta_K(\mathcal{C}_0, 10)/\zeta_K(10)$, where $\zeta_K(s)$ is the Dedekind zeta function of K and $\zeta_K(\mathcal{C}_0, s)$ is the partial zeta function associated to the trivial ideal class. Of course this equals 1 when the class group is trivial. Note that once again the factor of $\zeta_K(10)$ appears.

Earlier work of Papadopoulos in [23] uses a close analysis of Tate's Algorithm similar to our approach in Sections 2 and 5, working over a general local field, but not including the quantification of the local densities which we require.

This paper grew out of independent work of each of the authors: unpublished notes on purely local densities (at arbitrary primes) by Cremona, and a 2017 preprint [28] on global densities (excluding conditions at the primes $p = 2$ and $p = 3$) by Sadek. After the first version of the current paper appeared online, we noticed a new preprint [11] by Cho and Jeong, whose subject matter has some overlap with the current paper, but with several differences: conditions at the primes 2 and 3 are excluded in [11], and only conditions at finitely many primes are considered, through the use of short Weierstrass equations. On the other hand, they consider additional local conditions we do not, including the condition of having a fixed trace of Frobenius a_p at a prime p of good reduction, and their paper also contains applications to the distribution of analytic ranks.

2. Local densities I

2.1. Weierstrass equations and coordinate transformations

For any integral domain R denote by

$$\mathcal{W}(R) = R^5 = \{\boldsymbol{a} = (a_1, a_2, a_3, a_4, a_6) | a_i \in R\}$$

the set of all 5-tuples of coefficients in R of plane cubic curves $E_{\boldsymbol{a}}$ in long Weierstrass form over R:

$$E_{\boldsymbol{a}}: \; Y^2 + a_1 XY + a_3 Y = X^3 + a_2 X^2 + a_4 X + a_6.$$

Denote by $\Delta(\boldsymbol{a})$ the discriminant of $E_{\boldsymbol{a}}$; when $\Delta(\boldsymbol{a})$ is non-zero, $E_{\boldsymbol{a}}$ is a model for an elliptic curve defined over the fraction field of R; otherwise, we say that $\boldsymbol{a}$ is *singular*. Below we will also refer to the standard associated quantities b_2, b_4, b_6, b_8, c_4 and c_6; together with Δ these may all be viewed as elements of $\mathbb{Z}[a_1, a_2, a_3, a_4, a_6]$.

The *translation group* $\mathcal{T}(R) = \{\tau(r, s, t) \mid r, s, t \in R\}$ acts on $\mathcal{W}(R)$ in the standard way, with $\tau(r, s, t)$ induced by the coordinate substitutions $(X, Y) \mapsto (X + r, Y + sX + t)$; we call elements of $\mathcal{T}(R)$ *translations*.

In the case $R = \mathbb{Z}_p$, we make further definitions of certain subsets of $\mathcal{W}(\mathbb{Z}_p)$ and subgroups of $\mathcal{T}(\mathbb{Z}_p)$. We denote by v the normalised p-adic valuation.

Given non-negative integers v_i for $i = 1, 2, 3, 4, 6$, define

$$\mathcal{W}(v_1, v_2, v_3, v_4, v_6) = \{\boldsymbol{a} \in \mathcal{W}(\mathbb{Z}_p) \mid v(a_i) \geq v_i \text{ for } i = 1, 2, 3, 4, 6\}.$$

To specify further that $v(a_i) = v_i$ exactly, we indicate this by writing "$=v_i$": for example, $\mathcal{W}(1, 1, 1, 1, =1)$. Below we will also need notation for subsets of these satisfying an additional condition, for example $\mathcal{W}(1, 1, 1, 2, 2 \mid v(b_2) = 2)$ and $\mathcal{W}(1, 1, 1, 2, 2 \mid v(\Delta) = 6)$, whose meaning should be clear.

For $e, f, g \geq 0$ we define

$$\mathcal{T}_{e,f,g} = \{\tau(r, s, t) \in \mathcal{T}(\mathbb{Z}_p) : p^e \mid r,\ p^f \mid s,\ p^g \mid t\},$$

which is a subgroup of $\mathcal{T}(\mathbb{Z}_p)$ provided $e + f \geq g$, of index p^{e+f+g}.

2.2. Local densities and Tate's Algorithm

For each non-singular $\boldsymbol{a} \in \mathcal{W}(\mathbb{Z}_p)$, the equation $E_{\boldsymbol{a}}$ defines an elliptic curve over $\mathbb{Q}_p$. With the usual p-adic measure μ on $\mathbb{Z}_p$ such that $\mu(\mathbb{Z}_p) = 1$, we have $\mu(\mathcal{W}(\mathbb{Z}_p)) = 1$, and for any measurable subset $S \subseteq \mathcal{W}(\mathbb{Z}_p)$ we refer to $\mu(S)$ as the *density* (or p-adic density) of the associated set of equations $E_{\boldsymbol{a}}$, and also think of $\mu(S)$ as the probability that a random Weierstrass equation lies in S. Note that the subset of singular $\boldsymbol{a}$ has measure zero, and may be tacitly ignored.

For example,

$$\mu(\mathcal{W}(v_1, v_2, v_3, v_4, v_6)) = 1/p^{v_1+v_2+v_3+v_4+v_6}, \tag{2}$$

while if any of the v_i is replaced by $=v_i$, then the measure should be multiplied by $(1 - 1/p)$; so for $i = 6$,

$$\mu(\mathcal{W}(v_1, v_2, v_3, v_4, =v_6)) = (p-1)/p^{v_1+v_2+v_3+v_4+v_6+1}, \tag{3}$$

and similarly for $i = 1, 2, 3, 4$.

We write $\mathcal{T} = \mathcal{T}(\mathbb{Z}_p)$ for the rest of this section. The action of $\mathcal{T}$ on $\mathcal{W}(\mathbb{Z}_p)$ is measure-preserving and also leaves the discriminant, and c_4 and c_6, invariant. Translations induce isomorphisms of elliptic curves (when $\Delta \neq 0$).

Given some property or *type* T of isomorphism classes of elliptic curves over $\mathbb{Q}_p$, we associate a subset $\mathcal{W}_T(\mathbb{Z}_p) \subseteq \mathcal{W}(\mathbb{Z}_p)$:

$$\mathcal{W}_T(\mathbb{Z}_p) = \{\boldsymbol{a} \in \mathcal{W}(\mathbb{Z}_p) \mid E_{\boldsymbol{a}} \text{ is smooth and has type } T\},$$

and define the density of curves with property T as the p-adic measure of this set.

Definition 1. *The local density $\rho_T(p)$ of elliptic curves over $\mathbb{Q}_p$ with type T is the p-adic measure $\mu(\mathcal{W}_T(\mathbb{Z}_p))$ of the associated subset $\mathcal{W}_T(\mathbb{Z}_p) \subseteq \mathcal{W}(\mathbb{Z}_p)$:*

$$\rho_T(p) = \mu(\mathcal{W}_T(\mathbb{Z}_p)).$$

In this section p is fixed and we abbreviate: $\mathcal{W}_T = \mathcal{W}_T(\mathbb{Z}_p)$ and $\rho_T = \rho_T(p)$.

The types of interest to us are the following Kodaira types of reduction of elliptic curves over $\mathbb{Q}_p$:

- I_0 (good reduction);
- $\mathrm{I}_{\geq 1}$ (bad multiplicative reduction, of type I_m for some $m \geq 1$);
- bad additive reduction; with subtypes II, III, IV, II^*, III^*, IV^*, I_0^*, $\mathrm{I}_{\geq 1}^*$, the latter meaning type I_m^* for some $m \geq 1$.

We call these types *finite*, since, as we will see below (see Proposition 2.3), we only need know the coefficients $\boldsymbol{a}$ to finite p-adic precision in order to determine whether the curve $E_{\boldsymbol{a}}$ has each of these reduction types, provided that $E_{\boldsymbol{a}}$ is a minimal model. Moreover, the condition that $E_{\boldsymbol{a}}$ is minimal also only depends on $\boldsymbol{a}$ to finite precision (modulo p^6 suffices). Note that $\mathrm{I}_{\geq 1}$ and $\mathrm{I}_{\geq 1}^*$, the unions of types I_m and I_m^* for all $m \geq 1$ respectively, are finite in this sense. However, while for each fixed m it is true that finite p-adic precision suffices to detect the individual types I_m and I_m^*, this precision depends on m. For this reason we do not regard these types as finite, and for some of our results it will be necessary to consider these together, rather than individually.

Note that while each model $E_{\boldsymbol{a}}$ with $\Delta(\boldsymbol{a}) \neq 0$ defines an elliptic curve over $\mathbb{Q}_p$ whose type is well-defined, the set of $\boldsymbol{a} \in \mathcal{W}_T(\mathbb{Z}_p)$ for which $E_{\boldsymbol{a}}$ is itself a minimal model is a strictly smaller subset, with a smaller density, since scaling (replacing each a_i by $p^{ni}a_i$ for some $n \geq 1$) does not change the isomorphism class of $E_{\boldsymbol{a}}$. We will relate these two densities.

Define

$$\mathcal{W}_M = \{\boldsymbol{a} \in \mathcal{W}(\mathbb{Z}_p) \mid E_{\boldsymbol{a}} \text{ is a minimal model}\},$$

call $\boldsymbol{a} \in \mathcal{W}_M$ *minimal*, and set $\mathcal{W}_N$ to be the complement $\mathcal{W}(\mathbb{Z}_p) \setminus \mathcal{W}_M$. This complement contains the set $\mathcal{W}(1,2,3,4,6)$ of all "trivially non-minimal" $\boldsymbol{a}$, satisfying $p^i \mid a_i$ for $i = 1,2,3,4,6$, which has measure p^{-16}. It is clear that the action of $\mathcal{T}$ preserves minimality, so $\mathcal{T}$ maps both $\mathcal{W}_M$ and $\mathcal{W}_N$ to themselves.

Proposition 2.1.

1) *The subgroup of* $\mathcal{T}$ *preserving* $\mathcal{W}(1,2,3,4,6)$ *is* $\mathcal{T}_{2,1,3}$.
2) *Each orbit of* $\mathcal{T}$ *on* $\mathcal{W}_N$ *contains an element of* $\mathcal{W}(1,2,3,4,6)$.
3) $\mu(\mathcal{W}_M) = 1 - p^{-10}$.

Proof. (1) follows from the standard formulas linking the coefficients $\boldsymbol{a}$ to the transformed coefficients $\boldsymbol{a}'$ after translation by $\tau(r,s,t) \in \mathcal{T}$; in case $p \geq 5$ this is almost trivial, and it is straightforward to check for $p = 2$ and $p = 3$. See the proof of Theorem 5.3 for details.

(2) follows directly from Tate's algorithm [29], in which, given any non-minimal $\boldsymbol{a}$, one constructs a sequence of translations taking $\boldsymbol{a}$ to some $\boldsymbol{a}' \in \mathcal{W}(1,2,3,4,6)$.

(3): from (2), since $\mathcal{T}_{2,1,3}$ has index p^6 in $\mathcal{T}$, it follows that $\mathcal{W}_N$ is partitioned into p^6 disjoint subsets, each a translation of $\mathcal{W}(1,2,3,4,6)$ by an element of one coset of $\mathcal{T}_{2,1,3}$. Since $\mu(\mathcal{W}(1,2,3,4,6)) = 1/p^{16}$, it follows that $\mu(\mathcal{W}_N) = p^6/p^{16} = 1/p^{10}$ and hence $\mu(\mathcal{W}_M) = 1 - 1/p^{10}$. □

For each type T, we set $\mathcal{W}_T^M = \mathcal{W}_T \cap \mathcal{W}_M$, the set of $\boldsymbol{a} \in \mathcal{W}(\mathbb{Z}_p)$ for which $E_{\boldsymbol{a}}$ is a minimal model of an elliptic curve of type T, and make the following definition:

Definition 2. *The local density* $\rho_T^M = \rho_T^M(p)$ *of minimal Weierstrass equations defining elliptic curves over* $\mathbb{Q}_p$ *of type* T *is the* p*-adic measure of* $\mathcal{W}_T^M$*:*

$$\rho_T^M = \mu(\mathcal{W}_T^M) = \mu(\mathcal{W}_T \cap \mathcal{W}_M).$$

Although the properties we consider are invariants of elliptic curves up to isomorphism over $\mathbb{Q}_p$, and not properties of specific models or equations, we can still determine local densities by studying Weierstrass models, by relating ρ_T and ρ_T^M. For example, the model $E_{\boldsymbol{a}}$ will have bad reduction modulo p when $\Delta(\boldsymbol{a}) \equiv 0 \pmod{p}$, but the curve over $\mathbb{Q}_p$ which this model defines may still have good reduction if the model is non-minimal.

Just as all non-minimal $\boldsymbol{a}$ can be translated into the set $\mathcal{W}(1,2,3,4,6)$, which is defined by simple valuation conditions on the coefficients, Tate's algorithm implies that, for each type T, there is a "base set" $\mathcal{B}_T$ also defined by valuation conditions, such that

$$\boldsymbol{a} \text{ is minimal and of type } T \quad \Longleftrightarrow \quad \boldsymbol{a} \text{ has a translate in } \mathcal{B}_T.$$

In the following proposition and table, we define such a set $\mathcal{B}_T \subseteq \mathcal{W}(\mathbb{Z}_p)$ for each finite type T, and give its measure and the subgroup $\mathcal{T}_T \subseteq \mathcal{T}$ which stabilises it. For example, in the first line of the table for $T = \mathrm{I}_0$ (good reduction), we have $\mathcal{B}_T = \mathcal{W}(0,0,0,0,0 \mid v(\Delta) = 0)$, since the only condition required for good reduction apart from integrality (all coefficients have valuation ≥ 0) is that the discriminant has valuation zero. This condition is invariant under all translations, so $\mathcal{T}_{\mathrm{I}_0} = \mathcal{T}_{0,0,0} = \mathcal{T}$.

Proposition 2.2. *For each minimal $\boldsymbol{a} \in \mathcal{W}(\mathbb{Z}_p)$ there exists $\tau \in \mathcal{T}$ such that $\tau(\boldsymbol{a}) \in \mathcal{B}_T$ for exactly one of the base sets $\mathcal{B}_T$ in the following table. The table also shows the measure $\mu(\mathcal{B}_T)$, the stabiliser $\mathcal{T}_T$ and its index, and the measure $\rho_T^M = \mu(\mathcal{W}_T^M)$. The last row refers to the set of non-minimal $\boldsymbol{a}$, which has density $1/p^{10}$, with base set the set of trivially non-minimal $\boldsymbol{a}$. The discriminant of the cubic $x^3 + a_2x^2 + a_4x + a_6$ is denoted[1] $\tilde{\Delta}$.*

T	$\mathcal{B}_T$	$\mu(\mathcal{B}_T)$	$\mathcal{T}_T$	$[\mathcal{T} : \mathcal{T}_T]$	ρ_T^M
I_0	$\mathcal{W}(0,0,0,0,0 \mid v(\Delta) = 0)$	$(p-1)/p$	$\mathcal{T}_{0,0,0}$	1	$(p-1)/p$
$\mathrm{I}_{\geq 1}$	$\mathcal{W}(0,0,1,1,1 \mid v(b_2) = 0)$	$(p-1)/p^4$	$\mathcal{T}_{1,0,1}$	p^2	$(p-1)/p^2$
II	$\mathcal{W}(1,1,1,1,{=}1)$	$(p-1)/p^6$	$\mathcal{T}_{1,1,1}$	p^3	$(p-1)/p^3$
III	$\mathcal{W}(1,1,1,{=}1,2)$	$(p-1)/p^7$	$\mathcal{T}_{1,1,1}$	p^3	$(p-1)/p^4$
IV	$\mathcal{W}(1,1,1,2,2 \mid v(b_6) = 2)$	$(p-1)/p^8$	$\mathcal{T}_{1,1,1}$	p^3	$(p-1)/p^5$
I_0^*	$\mathcal{W}(1,1,2,2,3 \mid v(\tilde{\Delta}) = 6)$	$(p-1)/p^{10}$	$\mathcal{T}_{1,1,2}$	p^4	$(p-1)/p^6$
$\mathrm{I}_{\geq 1}^*$	$\mathcal{W}(1,{=}1,2,3,4)$	$(p-1)/p^{12}$	$\mathcal{T}_{2,1,2}$	p^5	$(p-1)/p^7$
IV^*	$\mathcal{W}(1,2,2,3,4 \mid v(b_6) = 4)$	$(p-1)/p^{13}$	$\mathcal{T}_{2,1,2}$	p^5	$(p-1)/p^8$
III^*	$\mathcal{W}(1,2,3,{=}3,5)$	$(p-1)/p^{15}$	$\mathcal{T}_{2,1,3}$	p^6	$(p-1)/p^9$
II^*	$\mathcal{W}(1,2,3,4,{=}5)$	$(p-1)/p^{16}$	$\mathcal{T}_{2,1,3}$	p^6	$(p-1)/p^{10}$
	$\mathcal{W}(1,2,3,4,6)$	$1/p^{16}$	$\mathcal{T}_{2,1,3}$	p^6	

Proof. The conditions defining each basic set $\mathcal{B}_T$ in the table are equivalent to the exit conditions in Tate's algorithm. The density of $\mathcal{B}_T$ is given by (2) or (3) when there is no extra condition (such as $v(b_6) = 2$ for type IV); the extra condition always has the effect of multiplying the density by $1 - 1/p$. For type I_0 this is Proposition 5.2, while for types $\mathrm{I}_{\geq 1}$, IV, and IV^* see (9), (10), and (11) in Section 5 respectively.

[1]For $p \neq 2$ we have $v(\tilde{\Delta}) = v(\Delta)$ for $\boldsymbol{a} \in \mathcal{W}(1,1,2,2,3)$, but this is not the case when $p = 2$.

The last column is the product of the index $[\mathcal{T} : \mathcal{T}_T]$ and the measure of $\mathcal{B}_T$, since the subset of $\boldsymbol{a}$ of type T is the disjoint union of $[\mathcal{T} : \mathcal{T}_T]$ translates of $\mathcal{B}_T$.

The side conditions for types $I_{\geq 1}$, IV and IV* ensure that a certain quadratic has distinct roots modulo p, while that for I_0^* ensures that a certain cubic has distinct roots modulo p. In the algorithm, if the exit condition for types I_0, $\mathrm{I}_{\geq 1}$, IV, I_0^* and $\mathrm{I}_{\geq 1}^*$ fails, a translation is required before continuing, and hence the stabiliser becomes smaller, by index p except in the first step when the index is p^2.

Tate's Algorithm itself takes an arbitrary $\boldsymbol{a} \in \mathcal{W}(\mathbb{Z}_p)$ and applies to it a sequence of translations, each well-defined up to an element in the next stabiliser, until it has been transformed into one of the base sets $\mathcal{B}_T$, at which point one concludes that the reduction type is T, or that the equation was not minimal.

Additional detail will be given in the proof of Theorem 5.3 below. □

The next proposition implies that for each of the finite[2] types T, the condition that $\boldsymbol{a} \in \mathcal{W}_T^M$ only depends on the class of $(a_i \pmod{p^6})$ in $(\mathbb{Z}_p/p^6\mathbb{Z}_p)^5$. We denote this product by $\mathcal{W}(\mathbb{Z}_p/p^6)$ and the image of $\boldsymbol{a}$ in $\mathcal{W}(\mathbb{Z}_p/p^6)$ by $\boldsymbol{a}(p^6)$; there are p^{30} classes in $\mathcal{W}(\mathbb{Z}_p/p^6)$, each of measure $1/p^{30}$. Similarly for $\mathcal{W}_N$.

Proposition 2.3. *Let $\boldsymbol{a}, \boldsymbol{a}' \in \mathcal{W}(\mathbb{Z}_p)$ be such that $\boldsymbol{a}(p^6) = \boldsymbol{a}'(p^6)$. Then*

$$\boldsymbol{a} \in \mathcal{W}_M \iff \boldsymbol{a}' \in \mathcal{W}_M,$$

and for each finite type T we have

$$\boldsymbol{a} \in \mathcal{W}_T^M \iff \boldsymbol{a}' \in \mathcal{W}_T^M.$$

Proof. This again follows from Tate's Algorithm. At each step the exit criterion is a test for membership of one of the basis sets $\mathcal{B}_T$, which only depends on $\boldsymbol{a}(p^6)$. Also, whenever a coordinate transformation $\tau(r,s,t)$ is required, in each case it is taken from the finite set of cosets of one of the subgroups $\mathcal{T}_{e,f,g}$. It is clear that the action of $\mathcal{T}$ is well-defined on $\mathcal{W}(\mathbb{Z}_p/p^6)$, in the sense that for each $\tau \in \mathcal{T}$, $\boldsymbol{a}(p^6) = \boldsymbol{a}'(p^6)$ implies $\tau(\boldsymbol{a})(p^6) = \tau(\boldsymbol{a}')(p^6)$.

It follows that the outcome of the algorithm (up to the point of determining that the initial $\boldsymbol{a}$ was non-minimal, and excluding the exact index m for types I_m and I_m^*) also only depends on the initial value of $\boldsymbol{a}(p^6)$. □

[2]Recall that we do not consider the individual types I_m and I_m^* as finite.

Corollary 2.4. *For each finite type T,*

$$\rho_T^M = N(T)/p^{10}$$

where $N(T) = p^k - p^{k-1}$ for some integer k with $1 \le k \le 10$, depending on the type T, such that

$$\#\{\boldsymbol{a} \in \mathcal{W}_T^M \mid 0 \le a_i < p^6 \text{ for } i = 1, 2, 3, 4, 6\} = p^{20} N(T).$$

Proof. This follows immediately from the table above. □

The precise index m for types I_m and I_m^* when $m \ge 1$ depends on the discriminant valuation which can be arbitrarily large, so no fixed p-adic precision will suffice to determine this value in all cases. However, for later reference we can determine the densities of these types for each m:

Proposition 2.5. *For each $m \ge 1$ we have $\rho_{I_m}^M = (p-1)^2/p^{m+2}$ and $\rho_{I_m^*}^M = (p-1)^2/p^{m+7}$.*

Proof. Consideration of Tate curves shows that $\rho_{I_m}^M = p \cdot \rho_{I_{m+1}}^M$. Explicitly, in [13, §2.2], the first author proved that if $p^m \mid \Delta$, then there is a translation of the form $\tau(r, 0, t)$ to a Weierstrass model such that p^m divides all of a_3, a_4, a_6, b_4, b_6, and b_8. For such a model we have $v(\Delta) = m \iff v(b_8) = m \iff v(a_6) = m$. Hence the relative density of models of type $\mathrm{I}_{\ge m+1}$ within those of type $\mathrm{I}_{\ge m}$ is $1/p$. Since $\sum_{m \ge 1} \rho_{I_m}^M = \rho_{I_{\ge 1}}^M = (p-1)/p^2$ (see Proposition 2.2), the first result follows.

For the second result, a careful analysis of Tate's algorithm (see the proof of Theorem 5.3 below) again shows that the density is reduced by a factor of p when m increases by 1, since the criterion for increasing m is that a certain monic quadratic has a repeated root modulo p, which has probability $1/p$. □

The preceding proof shows that

$$\mathcal{B}_{\mathrm{I}_m} = \mathcal{W}(0, 0, m, m, {=}\, m \mid v(b_2) = 0),$$

with measure $(p-1)^2/p^{3m+2}$ and stabiliser of index p^{2m}. In Section 5 we will show that

$$\mathcal{B}_{\mathrm{I}_m^*} = \begin{cases} \mathcal{W}(1, {=}1, k+1, k+2, 2k+2 \mid v(b_6) = 2k+2) & \text{if } m = 2k-1; \\ \mathcal{W}(1, {=}1, k+2, k+2, 2k+3 \mid v(b_8) = 2k+4) & \text{if } m = 2k; \end{cases}$$

with measure $(p-1)^2/p^{2m+11}$ and stabiliser of index p^{m+4}.

Hence we have an explicit upper bound on the p-adic precision to which we must know $\boldsymbol{a} \in \mathcal{W}(\mathbb{Z}_p)$ in order to determine the type of $E_{\boldsymbol{a}}$, provided that $\boldsymbol{a}$ is minimal, for all finite types. This is false without the minimality condition—that is, we cannot replace the subsets $\mathcal{W}_T^M$ by $\mathcal{W}_T$ in Proposition 2.3—since scaling (replacing each a_i by $p^{ni}a_i$ for some $n \geq 1$) does not change the isomorphism class of $E_{\boldsymbol{a}}$. Later we will consider $\boldsymbol{a}$ to higher p-adic precision in order to handle non-minimal models. On the other hand, for most finite types, lower p-adic precision than $\boldsymbol{a}(p^6)$ is required: for example, to distinguish between good reduction, multiplicative reduction and additive reduction of a minimal model only requires knowledge of $\boldsymbol{a} \pmod{p}$. However the individual finite types of additive reduction require successively higher precision, as does the condition of minimality itself, and to treat all finite types uniformly it is more convenient to work modulo p^6.

The two densities ρ_T and ρ_T^M are related as follows.

Proposition 2.6. *For each finite type T,*

$$\rho_T = \frac{p^{10}}{p^{10}-1}\rho_T^M.$$

Proof. Recall from Proposition 2.1 that the set $\mathcal{W}(1,2,3,4,6)$ of trivially non-minimal $\boldsymbol{a}$ has measure p^{-16}, and the set $\mathcal{W}_N$ of all non-minimal $\boldsymbol{a}$ is the union of p^6 translates of $\mathcal{W}(1,2,3,4,6)$ under a set of translations τ which are coset representatives for $\mathcal{T}_{2,1,3}$ in $\mathcal{T}$.

The scaling map $(a_1,a_2,a_3,a_4,a_6) \mapsto (pa_1,p^2a_2,p^3a_3,p^4a_4,p^6a_6)$ is a bijection from $\mathcal{W}_T$ to $\mathcal{W}_T \cap \mathcal{W}(1,2,3,4,6)$, so $\mu(\mathcal{W}_T \cap \mathcal{W}(1,2,3,4,6)) = p^{-16}\mu(\mathcal{W}_T)$. Hence

$$\mu(\mathcal{W}_T \cap \mathcal{W}_N) = p^6\mu(\mathcal{W}_T \cap \mathcal{W}(1,2,3,4,6)) = p^{-10}\mu(\mathcal{W}_T),$$

so $\mu(\mathcal{W}_T \cap \mathcal{W}_M) = (1-p^{-10})\mu(\mathcal{W}_T)$ and hence $\rho_T^M = (1-p^{-10})\rho_T$. $\square$

Writing this relation as $\rho_T = \rho_T^M \sum_{k=0}^{\infty} p^{-10k}$, we now give an interpretation of each term of the series in terms of the "level of non-minimality" for $\boldsymbol{a} \in \mathcal{W}(\mathbb{Z}_p)$, which we now define.

Definition 3. *Let $\boldsymbol{a} \in \mathcal{W}(\mathbb{Z}_p)$ with $\Delta(\boldsymbol{a}) \neq 0$. The* level $\lambda(\boldsymbol{a})$ *of $\boldsymbol{a}$ is defined by*

$$\lambda(\boldsymbol{a}) = \frac{1}{12}\left(v(\Delta(\boldsymbol{a})) - v(\Delta_{\min}(E_{\boldsymbol{a}}))\right),$$

where $\Delta_{\min}(E_{\boldsymbol{a}})$ is the discriminant of a minimal model for the elliptic curve $E_{\boldsymbol{a}}$.

With this definition, $\boldsymbol{a}$ is minimal if and only if $\lambda(\boldsymbol{a}) = 0$, and $\mathcal{W}(\mathbb{Z}_p)$ is the disjoint union of "level sets" $\mathcal{W}_k = \{\boldsymbol{a} \in \mathcal{W}(\mathbb{Z}_p) \mid \lambda(\boldsymbol{a}) = k\}$, together with the set of singular $\boldsymbol{a}$. Let $\mathcal{W}_{T,k} = \mathcal{W}_T \cap \mathcal{W}_k$.

Proposition 2.7. *For each $k \geq 0$,*

$$\mu(\mathcal{W}_k) = (1 - p^{-10})/p^{10k}$$

and

$$\mu(\mathcal{W}_{T,k}) = \rho_T^M/p^{10k}.$$

Proof. For $k = 0$ the first statement follows from the table and the second is by definition, using $\mathcal{W}_T^M = \mathcal{W}_{T,0}$ and the definition of ρ_T^M. Proceeding by induction, scaling by p maps $\mathcal{W}_k$ to $\mathcal{W}_{k+1} \cap \mathcal{W}(1, 2, 3, 4, 6)$ whose measure is $\mu(\mathcal{W}_{k+1})/p^6$. Hence $\mu(\mathcal{W}_k)/p^{16} = \mu(\mathcal{W}_{k+1})/p^6$, so $\mu(\mathcal{W}_{k+1}) = \mu(\mathcal{W}_k)/p^{10}$. Similarly when we restrict to any fixed finite type T, we obtain $\mu(\mathcal{W}_{T,k+1}) = \mu(\mathcal{W}_{T,_k})/p^{10}$. □

This proof implies the following generalisation of the statements above that minimality of $\boldsymbol{a}$, and the type of $E_{\boldsymbol{a}}$ when minimal, only depend on $\boldsymbol{a}$ $\pmod{p^6}$.

Corollary 2.8. *Let $k \geq 0$.*

1) *The class of $\boldsymbol{a}$ $\pmod{p^{6(k+1)}}$ determines $\lambda(\boldsymbol{a})$ exactly if $\lambda(\boldsymbol{a}) \leq k$.*
2) *When $\lambda(\boldsymbol{a}) \leq k$, the type T of $E_{\boldsymbol{a}}$ depends only on $\boldsymbol{a}$ $\pmod{p^{6(k+1)}}$, and each $\mathcal{W}_{T,k}$ is the union of $p^{20}N(T)$ classes modulo $p^{6(k+1)}$.*

For example, when $k = 2$, knowing $\boldsymbol{a}$ $\pmod{p^{18}}$ we can distinguish between the cases $\lambda(\boldsymbol{a}) = 0$, $\lambda(\boldsymbol{a}) = 1$, $\lambda(\boldsymbol{a}) = 2$ or $\lambda(\boldsymbol{a}) \geq 3$, and in all but the last case can also determine the type of $E_{\boldsymbol{a}}$ from $\boldsymbol{a}$ $\pmod{p^{18}}$; but to distinguish between $\lambda(\boldsymbol{a}) = 3$ and $\lambda(\boldsymbol{a}) \geq 4$ we would need to know $\boldsymbol{a}$ $\pmod{p^{24}}$.

3. General results relating p-adic densities and global densities

Our aim is to use the local density results of the previous section to obtain global density results for integral Weierstrass equations. This is straightforward if we only impose conditions at finitely many primes, the conclusion being in general that the global density is given, as one would expect, by the finite product of the local densities. This remains true when the local conditions are genuinely p-adic, and not only given by congruences to finite powers of each prime. However, when we impose local conditions at all primes, the passage from local to global densities is considerably more subtle. Some general methods in this direction have been developed, notably the "Ekedahl Sieve" introduced by Ekedahl in [16], and the approach of Poonen and Stoll in their paper [25] on the Cassels-Tate pairing on Abelian Varieties (see also the shorter note [26] by the same authors just on this issue). For applications to the existence of rational points on hypersurfaces, the results of Poonen and Voloch in [27] are often applicable, as for example in the case of plane cubic curves in the paper [7] of the first author with Bhargava and Fisher.

In the prior work mentioned so far, only uniform densities were used; in the case of quadrics in n variables, treated by the first author with Bhargava, Fisher, Jones, and Keating in [8], a different probability distribution was required at the real place, requiring additional analysis there. Further refinements to the methods may be found in the work of Bhargava, for example in [6]. Furthermore, some specific cases not covered by these have been handled individually, for example in the work of Bhargava such as his results with Shankar and Wang on square-free discriminants in [9].

The results and approaches of the papers cited cannot easily be applied directly in our situation, without additional discussion: for example, we need the flexibility to adjust local conditions at finitely many primes, and to introduce weights. For this reason, while our account in the rest of this section is firmly based on this prior work, it is almost self-contained, the main exception being the proof of the codimension 2 criterion of Proposition 3.5.

3.1. Global densities I: finitely many p-adic conditions

The standard definition of the *uniform density* of a subset $Z \subseteq \mathbb{Z}^d$ is as follows: we define the density of Z to be

$$\begin{aligned}\rho(Z) &= \lim_{X\to\infty} \frac{\#\{\boldsymbol{a} \in Z \mid |a_i| \le X \ \forall i\}}{\#\{\boldsymbol{a} \in \mathbb{Z}^d \mid |a_i| \le X \ \forall i\}} \\ &= \lim_{X\to\infty} (2X)^{-d} \#\{\boldsymbol{a} \in Z \mid |a_i| \le X \ \forall i\},\end{aligned} \tag{4}$$

if the limit exists. Similarly, we define the *upper density* $\overline{\rho}(Z)$ and *lower density* $\underline{\rho}(Z)$, replacing the limit by lim sup or lim inf respectively.

More generally given any vector of positive real *weights* $\mathbf{k} = (k_1, k_2, \dots, k_d)$ with sum $k = \sum_{i=1}^d k_i$, we can define a *weighted density*

$$(5)\qquad \begin{aligned} \rho^{\mathbf{k}}(Z) &= \lim_{X\to\infty} \frac{\#\{\boldsymbol{a} \in Z \mid |a_i| \le X^{k_i}\ \forall i\}}{\#\{\boldsymbol{a} \in \mathbb{Z}^d \mid |a_i| \le X^{k_i}\ \forall i\}} \\ &= \lim_{X\to\infty} 2^{-d} X^{-k} \#\{\boldsymbol{a} \in Z \mid |a_i| \le X^{k_i}\ \forall i\}. \end{aligned}$$

Note that neither the existence nor the value of this limit is affected if we scale the weight vector $\mathbf{k}$ by any positive real factor. When all the weights are equal we recover the uniform density as a special case.

We first determine the density of any subset $Z \subseteq \mathbb{Z}^d$ defined by congruence conditions at a finite set of primes, where it is given by a simple counting formula not depending on the weights. Let $M \ge 1$ and let $\Sigma \subseteq (\mathbb{Z}/M\mathbb{Z})^d$ be an arbitrary subset. One way to define such a set is locally, by choosing a finite set of primes p, a power p^e of each, and a subset $\Sigma_p \subseteq (\mathbb{Z}/p^e\mathbb{Z})^d$. Then set $M = \prod_p p^e$ and $\Sigma = \prod_{p|M} \Sigma_p$, where we identify $\mathbb{Z}/M\mathbb{Z}$ with $\prod_p \mathbb{Z}/p^e\mathbb{Z}$ by the Chinese Remainder Theorem.

Given Σ, define $Z(M,\Sigma) = \{\boldsymbol{a} \in \mathbb{Z}^d \mid \boldsymbol{a} \pmod M \in \Sigma\}$, and denote its weighted density by $\rho^{\mathbf{k}}(M,\Sigma) = \rho^{\mathbf{k}}(Z(M,\Sigma))$ (or simply $\rho(M,\Sigma) = \rho(Z(M,\Sigma))$ in the case of uniform density).

Proposition 3.1. *For all positive weights $\mathbf{k} = (k_1, \dots, k_d)$ and all subsets $\Sigma \subseteq (\mathbb{Z}/M\mathbb{Z})^d$ we have*

$$\rho^{\mathbf{k}}(M,\Sigma) = \frac{\#\Sigma}{M^d}.$$

In particular, this is independent of $\mathbf{k}$.

Proof. For $X > 0$, set $Z(M,\Sigma;X) = Z(M,\Sigma) \cap \prod_{i=1}^d [-X^{k_i}, X^{k_i}]$. For $1 \le i \le d$, set $b_i = \lfloor X^{k_i}/M \rfloor$, where, $\lfloor x \rfloor$ denotes the integer n such that $n \le x < n+1$. Then

$$M b_i \le X^{k_i} < M(b_i + 1),$$

and the interval $[-X^{k_i}, +X^{k_i}]$ contains between $2b_i$ and $2(b_i+1)$ complete sets of residue classes modulo M. Hence the box $\mathbb{Z}^d \cap \prod_{i=1}^d [-X^{k_i}, X^{k_i}]$ contains between $2^d \prod b_i$ and $2^d \prod (b_i + 1)$ complete sets of residue classes, each of which contains $\#\Sigma$ elements of $Z(M,\Sigma)$. So $\#Z(M,\Sigma;X)$ satisfies

$$2^d \#\Sigma \prod b_i \le \#Z(M,\Sigma;X) \le 2^d \#\Sigma \prod (b_i + 1).$$

Now $\prod b_i$ is approximately equal to $\prod X^{k_i}/M = X^k/M^d$ (where $k = \sum_{i=1}^d k_i$), so we approximate $\#Z(M,\Sigma;X)$ by $2^d \#\Sigma X^k / M^d$, and bound the error by noting that $\prod(b_i+1)-\prod b_i$ is a sum of 2^d-1 terms each bounded above by $X^{k-\min k_i}/M^n$ for some $n\le d-1$. This gives

$$\left|\frac{\#Z(M,\Sigma;X)}{2^dX^k}-\frac{\#\Sigma}{M^d}\right|=O(X^{-\min k_i}),$$

with implied constant depending on the weights, M and Σ but not X. The result follows on letting $X\to\infty$. □

In the uniform case ($d=k$ and all $k_i=1$), we see from this proof that

$$\#Z(M,\Sigma;X)=\frac{\#\Sigma}{M^d}\cdot(2X)^d+O(X^{d-1}).$$

The sets $Z(M,\Sigma)$ considered so far are cut out by congruence conditions at a *finite* set of primes (those which divide M), each congruence being modulo some *finite* power p^e. Our next step is to consider sets determined by a finite number of local p-adic conditions.

Let S be any set of primes (possibly including all primes). We may impose local p-adic conditions at all $p\in S$ by specifying a measurable subset $\mathcal{U}\subseteq\prod_{p\in S}\mathbb{Z}_p^d$. Embedding $\mathbb{Z}$ diagonally into $\prod_{p\in S}\mathbb{Z}_p$, this gives a subset $\mathcal{U}\cap\mathbb{Z}^d$ of $\mathbb{Z}^d$, which we denote $Z(\mathcal{U})$. When S contains all primes we denote $\prod_{p\in S}\mathbb{Z}_p$ by $\hat{\mathbb{Z}}$, and the local conditions are determined by a measurable subset of $\hat{\mathbb{Z}}^d$.

For example, we may take $\mathcal{U}=\prod_{p\in S}U_p$, where for each $p\in S$ we have a measurable subset $U_p\subseteq\mathbb{Z}_p^d$; we now have

$$Z(\mathcal{U})=\cap_{p\in S}(U_p\cap\mathbb{Z}^d)=\{\boldsymbol{a}\in\mathbb{Z}^d\mid \boldsymbol{a}\in U_p\ \forall p\in S\}.$$

Set

$$\rho^{\mathbf{k}}(\mathcal{U}):=\rho^{\mathbf{k}}(Z(\mathcal{U})),\tag{6}$$

and similarly define the upper and lower densities $\overline{\rho}^{\mathbf{k}}(\mathcal{U})$ and $\underline{\rho}^{\mathbf{k}}(\mathcal{U})$. The results which follow relate these densities with the measure $\mu(\mathcal{U})$, which in the special case equals $\prod_{p\in S}\mu(U_p)$, and we consider whether the equality $\rho^{\mathbf{k}}(\mathcal{U})=\mu(\mathcal{U})$ holds. It is easy to show that the inequality

$$\overline{\rho}^{\mathbf{k}}(\mathcal{U})\le\mu(\overline{\mathcal{U}})\tag{7}$$

always holds, where $\overline{\mathcal{U}}$ denotes the closure of $\mathcal{U}$. See Lemma 1.1 in [16], where[3] Ekedahl defines densities by counting the intersection with $[1, X]^d$ rather than $[-X, X]^d$, but the result is the same.

From now on, we always take $\mathcal{U}$ to be a subset of the form $\prod_{p\in S} U_p$ with $U_p \subseteq \mathbb{Z}_p^d$ measurable, and boundary of measure zero: $\mu(\partial(U_p)) = 0$. When S is finite, the density of $Z(\mathcal{U}) = \mathcal{U} \cap \mathbb{Z}^d$ always exists and equals the measure $\mu(U)$.

Proposition 3.2. *Let S be a finite set of primes, for each $p \in S$ let $U_p \subseteq \mathbb{Z}_p^d$ with $\mu(\partial(U_p)) = 0$, and set $\mathcal{U} = \prod_{p\in S} U_p$. Then for an arbitrary weight vector $\mathbf{k}$,*

$$\rho^{\mathbf{k}}(\mathcal{U}) = \prod_{p\in S} \mu(U_p).$$

Remark. Note that this is essentially contained in the proof by Poonen and Stoll in Lemma 20 of [25], but there they include a condition at the infinite place and do not have weights. When the infinite place is included we expect the density to depend on the weights. By restricting our attention to sets defined by conditions only at the finite places we obtain a simplification.

Proof. Set $M = \prod_{p\in S} p$. For $\lambda \geq 1$, define

$$Y_\lambda = \{\boldsymbol{a} \in \mathbb{Z}^d \mid (\forall p \in S)(\exists \boldsymbol{a}_p \in U_p) : \boldsymbol{a} \equiv \boldsymbol{a}_p \pmod{p^\lambda}\}.$$

Then $\rho^{\mathbf{k}}(Y_\lambda) = \#\Sigma_\lambda / M^{d\lambda}$ by Proposition 3.1, where Σ_λ is the reduction modulo M^λ of Y_λ, noting that Y_λ is a union of complete residue classes modulo M^λ.

The sets Y_λ are nested ($Y_{\lambda+1} \subseteq Y_\lambda$), their intersection is the closure of $Z(\mathcal{U})$, which has the same measure as $Z(\mathcal{U})$ by our assumption on the boundary measures. Hence $\rho^{\mathbf{k}}(\mathcal{U}) = \lim_{\lambda\to\infty} \rho^{\mathbf{k}}(Y_\lambda) = \lim_{\lambda\to\infty} \#\Sigma_\lambda/M^{d\lambda} = \prod_p \mu(U_p)$, where the last equality follows by the Chinese Remainder Theorem and the definition of the p-adic measure. □

3.2. Alternative choices of global density

The results of the previous subsection show that, provided that we only consider subsets $Z \subseteq \mathbb{Z}^d$ defined by local conditions at finitely many primes,

[3] As noted in Remark 4.13(1) of [15], Ekedahl omits the closure, without which the inequality fails, for example with $\mathcal{U} = \mathbb{Z}^d$.

weighted density does not depend on the weight vector, so is the same as uniform density.

A different independence on the choice of a real probability distribution of the global form of p-adic density was also noted in [8] by the first author with Bhargava *et al.*. Let D be a sufficiently well-behaved[4] probability distribution on $\mathbb{R}^d$, so $\int_{\mathbb{R}^d} D(\boldsymbol{x})d\boldsymbol{x} = 1$. Then for $Z \subseteq \mathbb{Z}^d$, define

$$\rho^D(Z) = \lim_{X\to\infty} \frac{\sum_{\boldsymbol{a}\in Z} D(\boldsymbol{a}/X)}{\sum_{\boldsymbol{a}\in\mathbb{Z}^d} D(\boldsymbol{a}/X)}.$$

To recover our original (unweighted) definition take $D = U$, the uniform distribution on the box $[-1,1]^d$.

It follows from [8, §2] that $\rho^D(Z)$ is independent of the distribution D. A similar result (with a similar proof) would hold for an analogous weighted definition of $\rho^D(Z)$:

$$\rho^{D,\mathbf{k}}(Z) = \lim_{X\to\infty} \frac{\sum_{\boldsymbol{a}\in Z} D(\dots, a_i/X^{k_i}, \dots)}{\sum_{\boldsymbol{a}\in\mathbb{Z}^d} D(\dots, a_i/X^{k_i}, \dots)}.$$

For the applications in this paper, we are not concerned with constraints at the infinite place, so will not need this generality, but it might be useful in other applications. For example, we could compute the density of elliptic curves over $\mathbb{R}$ with positive and negative discriminant, and hence include a fixed sign of the discriminant in density results for elliptic curves over $\mathbb{Q}$. This will depend on the distribution.

3.3. Global densities II: infinitely many p-adic conditions

We will closely follow the form of the Ekedahl Sieve used by Poonen and Stoll, referring to their paper [25, §9.3] as needed. We find it convenient to discuss their results in terms of the notion of an *admissible family* to encapsulate the critical condition in equation (10) of [25, Lemma 20] but not given a name there.

Let $d \geq 1$, and let $\mathcal{U} = \prod U_p \subseteq \hat{\mathbb{Z}}^d$ be a subset determined by a family of subsets $U_p \subseteq \mathbb{Z}_p^d$, one for each rational prime p. As before, we suppose that each U_p is measurable, and assume that the boundaries have measure zero.

[4]Piecewise smooth and rapidly decaying, in the sense that $D(x)$ and all its partial derivatives are $O(|x|^{-N})$ for all $N > 0$.

For each $M > 0$ define

$$\begin{aligned} Z_M(\mathcal{U}) &= \{\boldsymbol{a} \in \mathbb{Z}^d \mid \boldsymbol{a} \in U_p \text{ for some prime } p > M\} \\ &= \bigcup_{p>M} (U_p \cap \mathbb{Z}^d). \end{aligned}$$

For a positive weight vector $\mathbf{k}$, set $\rho_M^{\mathbf{k}}(\mathcal{U}) = \overline{\rho}^{\mathbf{k}}(Z_M(\mathcal{U}))$.

Definition 4. *The family $\mathcal{U}$ is* admissible with respect to $\mathbf{k}$, *or* $\mathbf{k}$-admissible, *if* $\lim_{M\to\infty} \rho_M^{\mathbf{k}}(\mathcal{U}) = 0$.

We will omit $\mathbf{k}$ from the notation when all the weights are equal.

Example 1. Let $U_p = p^2\mathbb{Z}_p$ for all p. The associated set $\mathcal{U} = \prod_p U_p$ is admissible.

Proof of admissibility in this example uses the fact that $\sum_p \mu(U_p)$ converges; however, this is not sufficient for $\prod U_p$ to be admissible. The next example, where $\mu(U_p) = 0$ for all p but still $\mathcal{U}$ is not admissible, was shown to us by Michael Stoll.

Example 2. For $n \geq 1$ let p_n be the nth prime, and define $U_{p_n} = \{n\}$, the singleton set. Then $\mu(U_p) = 0$ for all p, but $Z_M(\mathcal{U})$ contains all positive integers n except for the finitely many for which $p_n \leq M$, so its density for each M is the same as the density of the set of all positive integers, namely $1/2$. So the limit is not zero and $\mathcal{U} = \prod U_p$ is not admissible.

It will be useful to have simple sufficient criteria for a family to be admissible. First we note the following easy consequences of the definition.

Lemma 3.3. 1) *Let $\mathcal{U}' = \prod U'_p$ be a second family such that $U'_p = U_p$ for all but finitely many primes p. Then $\mathcal{U}$ is $\mathbf{k}$-admissible if and only if $\mathcal{U}'$ is $\mathbf{k}$-admissible, for any weight vector $\mathbf{k}$.*

2) *Let $\mathcal{U}' = \prod U'_p$ be a second family with $\mathcal{U}' \subseteq \mathcal{U}$ (that is, $U'_p \subseteq U_p$ for all p). Then $\mathbf{k}$-admissibility of $\mathcal{U}$ implies $\mathbf{k}$-admissibility of $\mathcal{U}'$.*

Proof. The first statement holds, since $Z_M(\mathcal{U}) = Z_M(\mathcal{U}')$ for all M greater than the largest prime p for which $U_p \neq U'_p$; the second is clear, since $Z_M(\mathcal{U}') \subseteq Z_M(\mathcal{U})$. □

Let S be any set of primes. Define

$$\rho^{\mathbf{k}}(\mathcal{U}, S) = \rho^{\mathbf{k}}(\{\boldsymbol{a} \in \mathbb{Z}^d \mid \boldsymbol{a} \in U_p \iff p \in S\}),$$

the density of the set of integer vectors which lie in the distinguished subset U_p precisely for the primes in S. Taking S to be the set of all primes, we have $\rho^{\mathbf{k}}(\mathcal{U}, S) = \rho^{\mathbf{k}}(\mathcal{U})$ as defined in (6). In what follows we will use the subsets U_p to encode conditions to be avoided, so that the density we are most interested in is $\rho^{\mathbf{k}}(\mathcal{U}, \emptyset)$, which we hope under certain conditions to equal $\prod_p (1 - s_p)$, where $s_p = \mu(U_p)$.

The result from [25] which we will use is the following: for admissible families, the density exists and equals the measure, so we have the desired product formula.

Proposition 3.4. *Let $\mathcal{U} = \prod U_p$ be an admissible family with respect to the weight vector $\mathbf{k}$, with $s_p = \mu(U_p)$ and $\mu(\partial U_p) = 0$. Then $\sum_p s_p$ converges, and for every finite set S of primes,*

$$\rho^{\mathbf{k}}(\mathcal{U}, S) = \prod_{p \in S} s_p \prod_{p \notin S} (1 - s_p). \tag{8}$$

In particular, the density of the set of $\boldsymbol{a} \in \mathbb{Z}^d$ which do not lie in U_p for any prime p is $\rho^{\mathbf{k}}(\mathcal{U}, \emptyset) = \prod_p (1 - s_p)$, and $\rho^{\mathbf{k}}(\mathcal{U}, S) = 0$ if S is infinite.

Proof. Replacing U_p by its complement in $\mathbb{Z}_p^d$ for $p \in S$ gives another admissible family by Lemma 3.3, and the general formula (8) follows from the same result for this latter family. Hence we may assume that $S = \emptyset$.

To ease notation we omit the superscript $\mathbf{k}$, writing ρ for $\rho^{\mathbf{k}}$.

Assume that $U_p = \emptyset$ for all $p > M$ for some M. Let U_p' be the complement of U_p in $\mathbb{Z}_p^d$. Now

$$\rho(\mathcal{U}, \emptyset) = \rho(\prod_{p \le M} U_p') = \prod_{p \le M} (1 - s_p),$$

by Proposition 3.2. This gives (8) since $s_p = 0$ for all $p > M$.

Hence, we have in general for each $M > 0$,

$$\rho(\prod_{p \le M} U_p') = \prod_{p \le M} (1 - s_p).$$

Now $Z(\prod_p U'_p) \subseteq Z(\prod_{p\le M} U'_p)$; the sets $Z(\prod_{p\le M} U'_p)$ form a decreasing nested sequence whose intersection as $M \to \infty$ is $Z(\prod_p U'_p)$. The complement is

$$\begin{aligned} &Z(\prod_{p\le M} U'_p) \setminus Z(\prod_p U'_p)\\ &\qquad = \{\boldsymbol{a} \in \mathbb{Z}^d \mid \boldsymbol{a} \notin U_p \text{ for all } p \le M;\ \boldsymbol{a} \in U_p \text{ for some } p > M\}\\ &\qquad \subseteq Z_M(\mathcal{U}), \end{aligned}$$

whose density tends to zero by the admissibility condition. Hence

$$\rho(\mathcal{U},\emptyset) = \rho(Z(\prod_p U'_p)) = \lim_{M\to\infty} \rho(\prod_{p\le M} U'_p) = \lim_{M\to\infty} \prod_{p\le M}(1-s_p) = \prod_p (1-s_p)$$

as required. □

Note that it follows from Proposition 3.4 that the density $\rho^{\mathbf{k}}(\mathcal{U}, S)$ is independent of the weight vector $\mathbf{k}$, being equal to a product which does not depend on $\mathbf{k}$, provided that $\mathcal{U}$ is $\mathbf{k}$-admissible.

Example (Example 1 continued). Proposition 3.4, together with the admissibility statement of Example 1, implies the well-known result that the density of the set of square-free integers is $1/\zeta(2)$. Since $\mathcal{U} = \prod p^2\mathbb{Z}_p$ defines an admissible family, with $s_p = 1/p^2$, the density of square-free integers is $\rho(\mathcal{U},\emptyset) = \prod_p (1 - 1/p^2) = 1/\zeta(2)$.

Closed subschemes of $\mathbb{Z}^d$ of codimension at least 2 determine admissible conditions. The following is the simplest example:

Example 3. The set of coprime pairs $(a, b) \in \mathbb{Z}^2$ has density $1/\zeta(2)$.

This is a special case (with $d = 2$, $f = X_1$ and $g = X_2$) of the following much more general result of Poonen and Stoll (see [25, Lemma 21]), which will be crucial for our applications in the next section. Note that the proof given in [25] simply states that it follows immediately from a result of Ekedahl (Theorem 1.2 of [16]) applied to the closed subscheme of the affine scheme $\mathbb{A}^d_{\mathbb{Z}}$ cut out by $f = g = 0$, making use of the fact that the subscheme has codimension 2. However, while it is clear that Ekedahl's theorem implies that the product formula holds in this situation, for our applications in the next section we need to know that $\prod U_p$ is admissible, so that we can adjust the p-adic condition at $p = 2$ and $p = 3$. It is hard to extract this precise statement from Ekedahl's proof, but the necessary details have been supplied by Bhargava in [6, Theorem 3.3] which we use instead.

Proposition 3.5. *Let $f, g \in \mathbb{Z}[X_1, \dots, X_d]$ be coprime polynomials. Let $\mathcal{U} = \prod U_p$ where*

$$U_p = \{\boldsymbol{a} \in \mathbb{Z}_p^d \mid f(\boldsymbol{a}) \equiv g(\boldsymbol{a}) \equiv 0 \pmod{p}\}.$$

Then $\mathcal{U}$ is $\mathbf{k}$-admissible, for all weight vectors $\mathbf{k}$.

Proof. The $\mathbb{Z}$-scheme Y cut out by $f = g = 0$ has codimension 2. In the case of uniform weights, we may apply Bhargava's estimate [6, Theorem 3.3] (with $n = d,\ k = 2,\ B = [-1, 1]^d,\ r = X$) to see that the cardinality of $Z_M(\mathcal{U})$ is $O(X^d/(M \log M) + X^{d-1})$, and hence $\rho_M(\mathcal{U}) = O(1/(M \log M))$ which tends to 0 as $M \to \infty$.

For the general case, we note that (as remarked by Bhargava *et al.* in [9, p. 4]), his result [6, Theorem 3.3] also holds in the weighted case. □

Remark. It has been observed by Bhargava (see the remarks on page 4 of [9] by Bhargava *et al.* for a similar observation) that among families $\mathcal{U} = \prod U_p$ with $\mu(U_p) = O(1/p^2)$, it is necessary to distinguish between those where U_p is defined by two independent "mod p" conditions, as in Proposition 3.5, and those defined by a single "mod p^2" condition. An example of the latter is to take a single square-free polynomial $f \in \mathbb{Z}[X_1, \dots, X_d]$ and define U_p to be the subset of $\boldsymbol{a} \in \mathbb{Z}_p^d$ where $f(\boldsymbol{a}) \equiv 0 \pmod{p^2}$, in order to determine the density of the set of $\boldsymbol{a} \in \mathbb{Z}^d$ such that $f(\boldsymbol{a})$ is square-free. In the former case, the family is $\mathbf{k}$-admissible for any weights $\mathbf{k}$, but in the latter case additional work is needed in order to establish $\mathbf{k}$-admissibility for suitable $\mathbf{k}$ and f, to conclude that the global weighted density is the product of local densities. The example treated in [9], of monic integral polynomials with square-free discriminant, is of the latter type.

Moreover, as discussed by Bhargava in [6, §1.3], a general result that the global density exists for all square-free f and is given by the product formula, is closely related to the *abc*-conjecture. In [18], Granville proved that the *abc*-conjecture implies the result for polynomials in one variable and arbitrary degree (the case of quadratics is easier, and for cubics was established by Hooley in [19], these cases being unconditional). In [24], Poonen proves this also for multivariable polynomials, using an unconditional reduction to the univariate case.

The simplest example where we require[5] non-uniform weights to establish $\mathbf{k}$-admissibility is the following. This example is closely related to the density of monic cubics in $\mathbb{Z}[X]$ with square-free discriminant, and we give details in the following example as a similar technique will be required in the next section when we consider the density of integral Weierstrass equations with square-free discriminant.

Example 4. Let $S = \{(a,b) \in \mathbb{Z}^2 \mid a^3 - b^2 \text{ is square-free}\}$, cut out by the local conditions $\mathcal{U} = \prod U_p$ where $U_p = \{(a,b) \in \mathbb{Z}_p^2 \mid a^3 \equiv b^2 \pmod{p^2}\}$. We show that $\mathcal{U}$ is $\mathbf{k}$-admissible for the weights $\mathbf{k} = (2,3)$, and hence that S has density given by the product formula

$$\rho^{\mathbf{k}}(S) = \prod_p \left(1 - 2/p^2 + 1/p^3\right).$$

Write U_p as the disjoint union $U'_p \cup U''_p$, where $U'_p = p\mathbb{Z}_p^2$ and $U''_p = \{(a,b) \in \mathbb{Z}_p^2 \mid p \nmid ab, p^2 \mid a^3 - b^2\}$. Set $\mathcal{U}' = \prod U'_p$ and $\mathcal{U}'' = \prod U''_p$.

Lemma 3.6. $\mu(U_p) = 2/p^2 - 1/p^3$.

Proof. Clearly $\mu(U'_p) = 1/p^2$. To compute $\mu(U''_p)$ it suffices to consider a, b modulo p^2 and note that there is a bijection between $\{(a,b) \in ((\mathbb{Z}/p^2\mathbb{Z})^*)^2 \mid a^3 = b^2\}$ and $(\mathbb{Z}/p^2\mathbb{Z})^*$ given by $(a,b) \mapsto b/a$ with inverse $t \mapsto (t^2, t^3)$. Hence $\mu(U''_p) = \varphi(p^2)/p^4 = 1/p^2 - 1/p^3$. $\square$

In the language of [9] by Bhargava *et al.*, $a^3 - b^2$ is "strongly divisible" by p^2 for $(a,b) \in U'_p$ but only "weakly divisible" for $(a,b) \in U''_p$. The proof of $\mathbf{k}$-admissibility for $\mathcal{U}' = \prod_p U_p$ is easier, and holds for arbitrary weights, while that for $\mathcal{U}'' = \prod_p U''_p$ is more subtle, and only works when $3k_1 \le 2k_2$. The choice $(2,3)$ for the weights is natural, considering that the discriminant of the cubic $X^3 - 3aX + 2b$ is $108(a^3 - b^2)$. Hence, apart from the conditions at 2 and 3 requiring adjustment, the density of S would give the density of monic cubics with square-free discriminant, with weights matching the natural ones for the coefficients of a monic univariate polynomial. The main result in [9] gives the density of monic polynomials in $\mathbb{Z}[X]$ with square-free discriminant as the product of local densities, in arbitrary degree, the result for degree 3 being that the density is $\frac{1}{2}\prod_{p\ge 3}(1 - 2/p^2 - 1/p^3)$, agreeing with our formula $\rho(S) = \prod_p (1 - 2/p^2 - 1/p^3)$ except for the local density at 2. In [9], the weights used for monic cubics $X^3 + a_1X^2 + a_2X + a_3$ are $\mathbf{k} = (1,2,3)$, consistent with our choice of weights $\mathbf{k} = (2,3)$ for S.

---

[5] See, however, the remark at the end of this section.

In showing that $\mathcal{U}'$ and $\mathcal{U}''$ are admissible, we may ignore the set Z_0 of pairs $(a, b) \in \mathbb{Z}^2$ with $a^3 = b^2$ (that is, pairs of the form (t^2, t^3) for some $t \in \mathbb{Z}$), as well as those for which $ab = 0$, since these form a subset of density zero.

We first show that $\mathcal{U}'$ is admissible, with arbitrary positive weights k_1, k_2. (This also follows from Proposition 3.5.) For this we must estimate the cardinality of the set

$$\bigcup_{p>M} \{(a, b) \in p\mathbb{Z}^2 \setminus Z_0 : |a| \le X^{k_1}, |b| \le X^{k_2}\},$$

divide by $4X^{k_1+k_2}$ and let $X \to \infty$ to obtain an estimate for the tail density $\rho_M(\mathcal{U}')$. The pth set in the union has cardinality $O((X^{k_1}/p)(X^{k_2}/p)) = O(X^{k_1+k_2}/p^2)$, and is empty for $p > X^{\min(k_1,k_2)}$, so the union has cardinality

$$O(\sum_{M<p\le X^{\min(k_1,k_2)}} X^{k_1+k_2}/p^2).$$

Dividing by $4X^{k_1+k_2}$ and letting $X \to \infty$, this is bounded above by $\sum_{p>M} 1/p^2$ and hence tends to 0 as $M \to \infty$.

Now we show that $\mathcal{U}''$ is $\mathbf{k}$-admissible for $\mathbf{k} = (2, 3)$; the same argument is valid whenever $3k_1 \le 2k_2$, but not for equal weights. We estimate the cardinality of the set

$$\bigcup_{p>M} \{(a, b) \in \mathbb{Z}^2 \setminus Z_0 : |a| \le X^2, |b| \le X^3, p \nmid ab, p^2 \mid a^3 - b^2\},$$

and show that, after dividing by $4X^5$ and letting $X \to \infty$, the resulting tail density $\rho_M(\mathcal{U}'')$ tends to 0 as $M \to \infty$. The pth set in this union is empty for $p > \sqrt{2}X^3$, since $p^2 \le |a^3 - b^2| \le 2X^6$. Let p be a prime with $M < p \le \sqrt{2}X^3$. For each integer a with $p \nmid a$, the number of solutions b to the congruence $b^2 \equiv a^3 \pmod{p^2}$ is either 2 or 0, according to whether a is a quadratic residue or not modulo p, as it follows from Hensel's Lemma that (since $p \nmid b$) each solution modulo p lifts uniquely to a solution modulo p^2. Since each residue class modulo p^2 has $2X^3/p^2 + O(1)$ representatives b in the interval $[-X^3, X^3]$, the number of pairs (a, b) to be counted (for each a) is $4X^3/p^2 + O(1)$, or zero. Hence the cardinality of the set above is at most

$$\sum_{M<p\le\sqrt{2}X^3} (2X^2)(4X^3/p^2 + O(1)).$$

The main term is

$$8X^5 \sum_{M<p\leq\sqrt{2}X^3} (1/p^2),$$

which after dividing by $4X^5$ and letting $X \to \infty$ is $2\sum_{p>M} 1/p^2$, which tends to 0 as $M \to \infty$ as required.

Each of the remaining terms is of size $O(X^2)$; as the number of terms is at most $\pi(\sqrt{2}X^3) = O(X^3/\log X)$ (by the Prime Number Theorem), their sum is $O(X^5/\log X)$. Dividing by $4X^5$ and letting $X \to \infty$, we see that the contribution of these error terms is negligible.

If the weights are (k_1, k_2) with $k_1/k_2 > 2/3$, then the total contribution of the error terms in the last part of the proof is no longer negligible.

Remark. Although the proof we have given here for the density of square-free values of $a^3 - b^2$ does not work with equal weights, the result also holds in this case, but the proof is considerably deeper. We are grateful to Manjul Bhargava for explaining this to us.

Instead of $a^3 - b^2$ we consider square-free values of $-4a^3 - 27b^2$, the discriminant of the cubic $x^3 + ax + b$. Embed the space of such cubics with integer coefficients into the larger space of all binary cubic forms over $\mathbb{Z}$, on which $\mathrm{GL}_2(\mathbb{Z})$ acts, leaving the discriminant invariant. In this larger space, ordering cubic forms by their height (the maximum absolute value of the coefficients), one can show that the density of those with square-free discriminant is the expected product of local densities, by showing that the associated tail densities tend to zero. Finally, the number of solutions to the Thue equation $F(x, y) = 1$ for a binary cubic form F over $\mathbb{Z}$ is bounded by 10 (Evertse gave the bound 12 in 1983 in [17], and this was improved to 10 by Bennett in 2001 in [5]). Hence each $\mathrm{GL}_2(\mathbb{Z})$-orbit of binary cubic forms contains at most 10 with leading coefficient 1, and possibly fewer with coefficients of the form $1, 0, a, b$, so the tail density estimates for binary cubic forms also apply to square-free discriminants of cubic polynomials $x^3 + ax + b$.

4. Global densities for elliptic curves

We now apply the results of the previous section in dimension $d = 5$, together with the local densities determined in Section 2, to determine global densities of integral Weierstrass equations satisfying certain combinations of local conditions.

Recall from Section 2 that $\mathcal{W}(\mathbb{Z}) = \mathbb{Z}^5$ is the space of all Weierstrass equations with coefficients in $\mathbb{Z}$, and now consider elliptic curves over $\mathbb{Q}$ defined by long integral Weierstrass equations $E_{\boldsymbol{a}}$ for $\boldsymbol{a} = (a_1, a_2, a_3, a_4, a_6) \in \mathcal{W}(\mathbb{Z})$. In common with other work on density results for elliptic curves, we use weighted densities with weights

$$\mathbf{k} = (1/12, 2/12, 3/12, 4/12, 6/12) = (1/12, 1/6, 1/4, 1/3, 1/2),$$

so for $X > 0$ we define

$$\mathcal{E}(X) = \{\boldsymbol{a} \in \mathcal{W}(\mathbb{Z}) \mid |a_i| \leq X^{i/12} \text{ for } i = 1, 2, 3, 4, 6\}.$$

We have $\#\mathcal{E}(X) \sim 32X^{4/3}$, as the sum of the weights is $4/3$.

For any subset $U \subseteq \mathcal{W}(\mathbb{Z})$ recall that the (weighted) density $\rho^{\mathbf{k}}(U)$ of U was defined (see (5)) as

$$\rho^{\mathbf{k}}(U) = \lim_{X \to \infty} \frac{\#\mathcal{E}(X) \cap U}{\#\mathcal{E}(X)}.$$

By homogeneity, the density is unchanged if we use the weight vector $(1, 2, 3, 4, 6)$ instead of $(1/12, 1/6, 1/4, 1/3, 1/2)$, as we did in the Introduction (1). Since the weight vector will remain fixed throughout this section, we simplify notation by writing $\rho(U)$ for $\rho^{\mathbf{k}}(U)$ in what follows. However, apart from the result about square-free discriminants (Theorem 4.6), it is not hard to see that the results of this section are independent of the weights, since we impose no condition at the infinite place and otherwise rely only on Proposition 3.5.

4.1. Global densities with a condition at a single prime

Fix a prime p. For each local type $T(p)$ of elliptic curves over $\mathbb{Q}_p$, let $\mathcal{W}_{T(p)}(\mathbb{Z}) = \mathcal{W}_{T(p)}(\mathbb{Z}_p) \cap \mathcal{W}(\mathbb{Z})$ and $\mathcal{W}^M_{T(p)}(\mathbb{Z}) = \mathcal{W}^M_{T(p)}(\mathbb{Z}_p) \cap \mathcal{W}(\mathbb{Z})$. The *global density* $\rho^{\mathbb{Z}}_{T(p)}$ of type $T(p)$ can now be defined as the density of $\mathcal{W}_{T(p)}(\mathbb{Z})$. There are two versions, the second one including the minimality condition at p.

Definition 5. *Set* $\rho^{\mathbb{Z}}_{T(p)} = \rho(\mathcal{W}_{T(p)}(\mathbb{Z}))$ *and* $\rho^{\mathbb{Z}M}_{T(p)} = \rho(\mathcal{W}^M_{T(p)}(\mathbb{Z}))$.

In words, $\rho^{\mathbb{Z}}_{T(p)}$ is the density of the set of integral Weierstrass equations defining elliptic curves with reduction type $T(p)$ at the prime p, while $\rho^{\mathbb{Z}M}_{T(p)}$

is the density of the set of integral Weierstrass equations which are minimal at the prime p and models of elliptic curves with reduction type $T(p)$.

These global densities are equal to the corresponding p-adic densities, both with and without the minimal condition at p:

Theorem 4.1. *Let $T(p)$ be one of the finite p-adic types (as listed in Proposition 2.2, depending only on $\boldsymbol{a}$ modulo p^6). Then*

$$\rho^{\mathbb{Z}}_{T(p)} = \rho_{T(p)}$$

and

$$\rho^{\mathbb{Z}M}_{T(p)} = \rho^{M}_{T(p)}.$$

Proof. Write $T = T(p)$. The first statement follows from Proposition 3.2, with $U_p = \mathcal{W}_T(\mathbb{Z}_p)$ and $U_q = \mathcal{W}(\mathbb{Z}_q)$ for all primes $q \neq p$.

By definition we have $\rho^{\mathbb{Z}M}_T = \rho(\mathcal{W}^M_T(\mathbb{Z}))$, and the latter is equal to ρ^M_T by Proposition 3.1 with modulus p^6, giving the second statement. □

Example 5. The density of elliptic curves over $\mathbb{Q}$ with good reduction at 2 (with no restrictions at any other primes) is

$$(1 - 2^{-1})/(1 - 2^{-10}) = 2^9/(2^{10} - 1) = 512/1023 \approx 50.0\%.$$

Example 6. The density of elliptic curves over $\mathbb{Q}$ with additive reduction of type III* at 5 (with no restrictions at any other primes) is

$$(5^2 - 5)/(5^{10} - 1) = 5/2441406.$$

4.2. Global densities with conditions at finitely many primes

Let S be a finite set of primes, and for each $p \in S$ fix a finite reduction type $T(p)$. Applying Proposition 2.7 with Proposition 3.1 and Proposition 3.2 we immediately obtain the following.

Theorem 4.2. *Let S be any finite set of primes, and for each $p \in S$ let $T(p)$ be a finite reduction type.*

1) *The density of integral Weierstrass equations which for all $p \in S$ are minimal at p with reduction type $T(p)$ is $\prod_{p\in S} \rho^M_{T(p)}$.*
2) *The density of elliptic curves over $\mathbb{Q}$ whose reduction type at p is $T(p)$ for all $p \in S$ is $\prod_{p\in S} \rho_{T(p)}$.*

Example 7. The density of elliptic curves over $\mathbb{Q}$ with good reduction at both 2 and 3 (with no restrictions at any other primes) is

$$2^9(2-1)3^9(3-1)/(2^{10}-1)(3^{10}-1) = 839808/2516921 \approx 33.37\%.$$

Example 8. Let p_1, p_2 and p_3 be distinct primes. The density of elliptic curves over $\mathbb{Q}$ with good reduction at p_1, multiplicative reduction at p_2 and additive reduction at p_3 (with no restrictions at any other primes) is

$$\left(\frac{1-p_1^{-1}}{1-p_1^{-10}}\right)\left(\frac{p_2^{-1}-p_2^{-2}}{1-p_2^{-10}}\right)\left(\frac{p_3^{-2}-p_3^{-10}}{1-p_3^{-10}}\right).$$

4.3. Global densities with conditions at infinitely many primes

To obtain density results with conditions at infinitely many primes, we may use Proposition 3.4, provided that the excluded sets $U_p \subset \mathcal{W}(\mathbb{Z}_p)$ form an admissible family. The previous subsection dealt with the simplest case where almost all U_p were empty.

Recall the standard invariants $c_4, c_6 \in \mathbb{Z}[a_1, a_2, a_3, a_4, a_6]$ of a Weierstrass model $E_{\boldsymbol{a}}$. As elements of $\mathbb{Z}[a_1, a_2, a_3, a_4, a_6]$, they are both irreducible (being linear in a_4 and a_6 respectively), and coprime. The results in this subsection follow from the following.

Lemma 4.3. *For each prime p, define $U_p \subseteq \mathcal{W}(\mathbb{Z}_p)$ by*

$$U_p = \{\boldsymbol{a} \in \mathcal{W}(\mathbb{Z}_p) \mid c_4(\boldsymbol{a}) \equiv c_6(\boldsymbol{a}) \equiv 0 \pmod{p}\}.$$

Then the family $\mathcal{U} = \prod U_p$ is admissible.

Proof. Since c_4 and c_6 are coprime, we may apply Proposition 3.5. □

For $p \geq 5$, the condition $c_4 \equiv c_6 \equiv 0 \pmod{p}$ is equivalent to the Weierstrass model $E_{\boldsymbol{a}}$ being non-minimal or of bad additive reduction[6], so, for $p \geq 5$, we have $U_p = U_p'$ where

$$U_p' = \mathcal{W}(\mathbb{Z}_p) \setminus \left(\mathcal{W}^M_{\mathrm{I}_0} \cup \mathcal{W}^M_{\mathrm{I}_{\geq 1}}\right).$$

Since $\mu(U_p') = 1 - (\rho^M_{I_0} + \rho^M_{I_{\geq 1}}) = 1/p^2$ for all primes p, it follows that $\mu(U_p) = 1/p^2$ for all $p \geq 5$. One may also check that $\mu(U_p) = 1/p$ for $p = 2, 3$,

[6]Note that for $p = 2$ and $p = 3$ one can have good reduction when $p \mid c_4$ and $p \mid c_6$, for example 11a1 for $p = 2$ and 17a1 for $p = 3$. Also, c_4 and c_6 are not coprime as polynomials over $\mathbb{F}_2$ or $\mathbb{F}_3$.

using $c_4 \equiv a_1^4$ and $c_6 \equiv a_1^6 \pmod 2$, and $c_4 \equiv (a_1^2 + a_2)^2$ and $c_6 \equiv -(a_1^2 + a_2)^3 \pmod 3$, but we will not need these values.

Recall that an elliptic curve is called *semistable at a prime* p if its reduction type is either good (type I_0) or multiplicative (type $\mathrm{I}_{\geq 1}$), and *semistable* if it is semistable at all primes.

Theorem 4.4.

1) *The density of integral Weierstrass equations which are minimal models of semistable elliptic curves is* $1/\zeta(2) \approx 60.79\%$.
2) *The density of semistable elliptic curves over* $\mathbb{Q}$ *is* $\zeta(10)/\zeta(2) \approx 60.85\%$.

Proof. Let $\mathcal{U} = \prod U_p$ and $\mathcal{U}' = \prod U_p'$ be as above. Since $\mathcal{U}$ is admissible by Lemma 3.5, so is $\mathcal{U}'$ by Lemma 3.3. Also, $\mu(\mathcal{U}_p') = 1/p^2$ for all p, by Proposition 2.2 (as noted above). Taking $S = \emptyset$ in Proposition 3.4 gives the density stated, since $\prod_p (1 - 1/p^2) = 1/\zeta(2)$.

For the second part we let U_p'' be the set of Weierstrass models of curves with additive reduction. This is a subset of U_p, since U_p includes not only these models but also non-minimal models of curves with good or multiplicative reduction. Now the local density of curves with good or multiplicative reduction is $(1 - p^{-2})/(1 - p^{-10})$, so $\mu(U_p'') = 1 - (1 - p^{-2})/(1 - p^{-10})$. Applying Proposition 3.4 again yields the desired density as $\prod_p (1 - \mu(U_p'')) = \prod_p (1 - p^{-2})/(1 - p^{-10}) = \zeta(10)/\zeta(2)$. □

We can obtain further global density results by changing the local conditions at any finite set of primes, provided that we know the associated local densities. The only constraint on results provable in this way is therefore that, at all but finitely many primes, the condition we impose is that of semistability, *i.e.*, good or multiplicative reduction. As in the two parts of Theorem 4.4, if we also impose conditions of minimality at all primes, this will not affect the convergence criteria, merely dividing the global density by $\prod_p (1 - p^{-10})^{-1} = \zeta(10) = \pi^{10}/93555 \approx 1.000994575$. This establishes the following.

Theorem 4.5. *Let* S *be any finite set of primes, and for each* $p \in S$ *let* $T(p)$ *be a finite reduction type.*

1) *The density of integral Weierstrass equations which are global minimal models of elliptic curves over* $\mathbb{Q}$ *whose reduction type at* p *is* $T(p)$ *for*

all $p \in S$, and which are semistable at all other primes, is

$$\zeta(2)^{-1} \prod_{p \in S} \rho_{T(p)}/(1 - p^{-2}).$$

2) *The density of elliptic curves over $\mathbb{Q}$ whose reduction type at p is $T(p)$ for all $p \in S$ and which are semistable at all other primes is*

$$\zeta(10)\zeta(2)^{-1} \prod_{p \in S} \rho_{T(p)}/(1 - p^{-2}).$$

4.4. Curves with square-free discriminant

A Weierstrass equation has square-free discriminant if and only if it is minimal and of reduction type I_0 or I_1. These have local density $1 - 1/p$ and $(p-1)^2/p^3$, by Propositions 2.2 and 2.5 respectively, so the local density of those with square-free discriminant is $1 - 2/p^2 + 1/p^3$. Hence the set U_p of Weierstrass equations with discriminant divisible by p^2 has local density $2/p^2 - 1/p^3$. By comparison with the case of square-free discriminants of monic cubic polynomials (see Example 4 in the previous section), we expect $\mathcal{U} = \prod U_p$ to be admissible. This is indeed the case, provided that we use appropriate weights, as specified at the start of this section.

Theorem 4.6. 1) *The density of integral Weierstrass equations whose discriminant is square-free is*

$$\prod_p \left(1 - \frac{2}{p^2} + \frac{1}{p^3}\right) \approx 42.89\%.$$

2) *The density of elliptic curves over $\mathbb{Q}$ whose minimal discriminant is square-free is*

$$\zeta(10) \prod_p \left(1 - \frac{2}{p^2} + \frac{1}{p^3}\right) \approx 42.93\%.$$

For $p \neq 2$, the local density $(1 - 2/p^2 + 1/p^3)$ is exactly the same as that of monic cubic polynomials over $\mathbb{Z}_p$ with square-free discriminant (see [9] and [1, Theorem 6.8]). Hence, by [9, Theorem 1.1], the theorem states that the probability that a random integral Weierstrass equation has square-free discriminant is (after taking the discrepancy for $p = 2$ into account) equal to $5/4$ times the probability that a random monic integral cubic polynomial has square-free discriminant.

Proof. The proof follows the argument given in Example 4 above, taking the additional variables into account. We again write U_p as a disjoint union $U_p = U'_p \cup U''_p$, where

$$U'_p = \mathcal{W}(\mathbb{Z}_p) \setminus \mathcal{W}^M_{\mathrm{I}_{\geq 0}},$$

is the set of Weierstrass equations with bad additive reduction at p or non-minimal at p, and

$$U''_p = \mathcal{W}^M_{\mathrm{I}_{\geq 2}}$$

is the set of Weierstrass equations with multiplicative reduction at p of Type I_m for some $m \geq 2$. Admissibility of $\mathcal{U}' = \prod U'_p$ has already been established in the proof of Theorem 4.4, so we consider admissibility of $\mathcal{U}'' = \prod U''_p$. Ignoring $p = 2$ and 3, as we may by Lemma 3.3, the condition for belonging to U''_p is that $p \nmid c_4, c_6$ but $p^2 \mid \Delta$, or equivalently $p^2 \mid c_4^3 - c_6^2$. This is a "mod p^2 condition", in contrast to membership of U'_p which is a "mod p condition".

Regarding Δ as a polynomial in a_6 with coefficients in $\mathbb{Z}[a_1, a_2, a_3, a_4]$, it has degree 2 with leading coefficient $-432 = -2^4 3^3$ and discriminant c_4^3. (Note that $c_4 \in \mathbb{Z}[a_1, a_2, a_3, a_4]$ does not depend on a_6.) Hence for each fixed $(a_1, a_2, a_3, a_4) \in \mathbb{Z}^4$ with $p \nmid c_4$, there are at most 2 solutions for $a_6 \pmod p$ to the congruence $\Delta \equiv 0 \pmod p$, each of which lifts to a unique solution to $\Delta \equiv 0 \pmod{p^2}$.

Secondly, each term in Δ has weight 12 when we give a_i weight i, so for $\boldsymbol{a}$ bounded by $|a_i| \leq X^{i/12}$, each monomial appearing in Δ is bounded by X, and hence $|\Delta| \leq BX$, where B is the sum of the absolute values of the coefficients of Δ. (In fact, $B = 1714$, but the actual numerical coefficient is unimportant.) It follows that if $\boldsymbol{a}$ satisfies the weighted bounds, $p^2 \mid \Delta$ and $\Delta \neq 0$, then $p \leq (BX)^{1/2}$.

To compute the tail density $\rho^{\mathbf{k}}_M(\mathcal{U}'')$, we must estimate the cardinality of the set

$$\bigcup_{p>M} \{\boldsymbol{a} \in \mathbb{Z}^5 : |a_i| \leq X^{i/12}, p \nmid c_4, p^2 \mid \Delta\},$$

and we may ignore $\boldsymbol{a}$ with $\Delta = 0$ as these have zero density. The pth set in this union is empty unless $p \leq (BX)^{1/2}$. For each p below this bound, the number of 4-tuples (a_1, a_2, a_3, a_4) satisfying the bounds is $O(X^{10/12}) = O(X^{5/6})$, and each 4-tuple determines at most two values of $a_6 \pmod{p^2}$, hence $O(X^{1/2}/p^2) + O(1)$ values of a_6 also satisfying $|a_6| \leq X^{1/2}$. Adding over all p with $M < p \leq (BX)^{1/2}$, the main term is

$$O(X^{5/6}) \sum_{M<p\leq(BX)^{1/2}} O(X^{1/2}/p^2) = O(X^{4/3}) \sum_{M<p\leq(BX)^{1/2}} 1/p^2,$$

contributing at most $\sum_{p>M} 1/p^2$ to the tail density. Each error term is $O(X^{5/6})$ and the number of terms is $O(\pi((BX)^{1/2})) = O(X^{1/2}/\log X)$, so the total error is $O(X^{4/3}/\log X)$ which is $o(X^{4/3})$ and hence negligible.

This completes the proof that $\mathcal{U}''$ is admissible, and the rest of the statement of the Theorem follows as before. □

Remark. It is perhaps worth noting what are the properties of the discriminant polynomial $\Delta(a_1, a_2, a_3, a_4, a_6)$ which ensure that the above proof works.

Firstly, it is *isobaric* with respect to certain positive weights of the variables a_i (meaning that each monomial has the same weight). We use these weights of the variables (scaled by $1/12$ to make the total weight of Δ equal to 1, but that is unimportant) as the weights used to define the density.

Secondly, we used the fact that Δ has degree only 2 in one of the variables, a_6. Careful examination of the proof above reveals that, in order to show that the error terms were negligible, it was crucial that the exponent $1/2$ on the bound for this variable matched the exponent on the bound on p, which in turn came from the fact that our condition was that Δ was square-free.

As with square-free values of the discriminant of a cubic polynomial $x^3 + ax + b$, it is possible that Theorem 4.6 also holds using equal weights on the coefficients a_i, but we have not tried to prove this. One approach might be to embed the space of Weierstrass cubics in the larger set of ternary cubic forms over $\mathbb{Z}$, for which this result is known: see [9].

We would also expect the methods used here to be able to establish the density of monic integer quartic polynomials whose discriminant is *cube-free*, using the natural weights rather than equal weights, since that discriminant has degree 3 in the constant coefficient, but that determining the density of square-free discriminants of quartic (and higher degree) monic integer polynomials would be harder; indeed, the methods used in [9] to evaluate this (in arbitrary degree) are much deeper.

4.5. Curves with prime-power conductor (or discriminant)

Finally in this section, we consider elliptic curves with a single prime of bad reduction.

Fix $X > 0$, and consider first elliptic curves with a single prime $p < X$ of bad multiplicative reduction, good reduction at all other primes $q < X$, and no restriction at primes $q > X$. That is, we consider elliptic curves of conductor $N = pN'$ where N' has no prime factors less than X. This set has density

$$\sum_{p \le X} \left((1/p - 1/p^2) \prod_{q \le X, q \ne p} (1 - 1/q) \right) = \left(\sum_{p \le X} 1/p \right) \left(\prod_{q \le X} (1 - 1/q) \right).$$

As $X \to \infty$, the first factor $\sum_{p \le X}(1/p) \sim \log\log X$, while $\prod_{q \le X}(1 - 1/q) \sim e^{-\gamma}/\log X$, where γ is Euler's constant. Hence the density is $O(\log\log X/\log X)$, and tends to 0 as $X \to \infty$.

Hence the density of elliptic curves with prime discriminant is also zero, as these are a subset of those with prime conductor.

A small modification of this argument applies to curves of prime power conductor (equivalently, prime power discriminant). For each X, the set of curves with precisely one prime $p \le X$ of bad reduction has density

$$\sum_{p \le X} \left(1/p \prod_{q \le X, q \ne p} (1 - 1/q) \right) = \left(\sum_{p \le X} 1/(p-1) \right) \left(\prod_{q \le X} (1 - 1/q) \right).$$

Since $1/(p-1) - 1/p = 1/p(p-1)$ and $\sum 1/p(p-1)$ converges, the asymptotics are unchanged.

5. Local densities II

In this section we extend the local density results of Section 2 to include the distribution of conductor exponents f_p and Tamagawa numbers c_p, for each type of reduction. Consequent global results may be obtained using the methods of Sections 3 and 4.

The results here are all obtained by following in detail the steps of Tate's Algorithm, as originally given in [29]. Our methods are similar to those employed by Papadopoulos in [23], where he establishes congruence conditions on the Weierstrass coefficients a_i for each Kodaira reduction type. As Papadopoulos observes, for $p \ge 5$ the type is completely determined by the valuations of the invariants c_4, c_6 and Δ; for $p = 3$ one can make use of the coefficients b_i, while for $p = 2$ one is forced to consider all the a_i. Since the expression for Δ as a polynomial in the a_i has 26 terms, this would be tiresome to do by hand, and we use computer algebra to assist us. The reader may find `Sage` code to verify the claims made in this section at [14]. The main differences between the results of this section and those of Papadopoulos are that we quantify each step in order to find the p-adic density of each case, while on the other hand Papadopoulos works in the more general context of a local field and not just $\mathbb{Q}_p$ itself.

Throughout, p will denote a fixed prime; in the results and proofs we often need to consider $p = 2$ and $p = 3$ separately.

All curves with good reduction at p have $f_p = 0$ and $c_p = 1$. It is well-known that the density of Weierstrass equations which have good reduction is $1 - 1/p$. The first author first learned the following fact from Hendrik Lenstra (who showed him a different proof from the one which follows), but as we do not know a suitable reference we include a proof here.

Lemma 5.1. *Let q be a prime power. Of the q^5 Weierstrass equations over $\mathbb{F}_q$, precisely q^4 are singular.*

Proof. Weierstrass equations define irreducible cubic curves, and by Bezout's Theorem, they can have at most one singular point, which is not the unique point at infinity, and hence is one of the q^2 points in the affine plane. For each of these, the number of equations having the specified point as its singular point is the same (by translation), so it suffices to count equations for which $P = (0, 0)$ is singular. Now P lies on the curve if and only if $a_6 = 0$, and then P is singular if and only if $a_3 = a_4 = 0$, so there are q^2 equations for which P is singular, and q^4 singular equations in all. □

Recall from Section 2 the notation

$$\mathcal{W}(v_1, v_2, v_3, v_4, v_6) = \{\boldsymbol{a} \in \mathcal{W}(\mathbb{Z}_p) \mid v(a_i) \geq v_i \text{ for } i = 1, 2, 3, 4, 6\}.$$

Proposition 5.2. *The density of Weierstrass equations over $\mathbb{Z}_p$ which have good reduction is*

$$\mu(\mathcal{W}(0, 0, 0, 0, 0 \mid v(\Delta) = 0)) = 1 - 1/p.$$

Proof. Immediate from Lemma 5.1. □

For the bad reduction types the distributions of f_p and c_p are as follows.

Theorem 5.3 (Distribution of conductor exponents and Tamagawa numbers by reduction type). *Within each bad reduction type, whose density is given by Proposition 2.2, the relative densities of each possible conductor exponent and Tamagawa number are as follows. Where two possibilities are given for the Tamagawa number, the density is split equally between them.*

1) *Multiplicative reduction types, all p:*

Type		f_p	c_p	*relative density*	*absolute density*
I_m	*each* $m \geq 1$	1		$(p-1)/p^m$	$(p-1)^2/p^{m+2}$
I_m	*split*	1	m	$1/2$	$(p-1)/(2p^2)$ *(total, all* m*)*
	non-split, m *even*	1	2	$1/(2(p+1))$	$(p-1)/(2p^2(p+1))$
	non-split, m *odd*	1	1	$p/(2(p+1))$	$(p-1)/(2p(p+1))$

2) *Additive reduction types:*
$p \geq 5$:

Type	f_p	c_p	*relative density*
$\mathrm{II},\mathrm{II}^*$	2	1	1
$\mathrm{III},\mathrm{III}^*$	2	2	1
$\mathrm{IV}, \mathrm{IV}^*$	2	1 *or* 3	1
I_0^*	2	1	$(p+1)/(3p)$
	2	2	$1/2$
	2	4	$(p-2)/(6p)$
I_m^*	2	2 *or* 4	1

$p = 3$:

Type	f_p	c_p	*relative density*
$\mathrm{II},\mathrm{II}^*$	3	1	$2/3$
	4	1	$2/9$
	5	1	$1/9$
$\mathrm{III},\mathrm{III}^*$	2	2	1
$\mathrm{IV},\mathrm{IV}^*$	3	1 *or* 3	$2/3$
	4	1 *or* 3	$2/9$
	5	1 *or* 3	$1/9$
I_0^*	2	1	$4/9$
	2	2	$1/2$
	2	4	$1/18$
I_m^*	2	2 *or* 4	1

$p = 2$:

Type	f_p	c_p	*relative density*
II	4	1	1/2
	6	1	3/8
	7	1	1/8
II*	3	1	1/2
	4	1	1/4
	6	1	1/4
III,III*	3	2	1/2
	5	2	1/4
	7	2	1/8
	8	2	1/8
IV,IV*	2	1 *or* 3	1
I$_0^*$	4	1 *or* 2	1/2
	5	1 *or* 2	1/4
	6	1 *or* 2	1/4
I$_m^*$	3	2 *or* 4	1/2
	4	2 *or* 4	1/4
	5	2 *or* 4	1/16
	6	2 *or* 4	1/8
	7	2 *or* 4	1/16

In the proof we use the following elementary counting lemmas; the second is Lemma 3 in [7].

Lemma 5.4. *Let q be a prime power. Of the q^2 monic quadratics $f \in \mathbb{F}_q[X]$,*

- *q have a double root;*
- *$q(q-1)/2$ have distinct roots in $\mathbb{F}_q$;*
- *$q(q-1)/2$ have conjugate roots in $\mathbb{F}_{q^2}$.*

Lemma 5.5. *Let q be a prime power. Of the q^3 monic cubics $g \in \mathbb{F}_q[X]$,*

- *q^2 have a multiple root, of which*
  - · *q have a triple root (necessarily in $\mathbb{F}_q$);*
  - · *$q(q-1)$ have a double root and a single root (both in $\mathbb{F}_q$);*
- *$q^3 - q^2$ have distinct roots, of which*
  - · *$q(q-1)(q-2)/6$ have distinct roots in $\mathbb{F}_q$;*
  - · *$q^2(q-1)/2$ have one root in $\mathbb{F}_q$ and two conjugate roots in $\mathbb{F}_{q^2}$;*
  - · *$q(q^2-1)/3$ have conjugate roots in $\mathbb{F}_{q^3}$.*

5.1. Proof of Theorem 5.3

During the course of the proof, we will fill in details which were only sketched in the proof of Proposition 2.2.

We follow the steps of Tate's Algorithm. Recall from Section 2 the notation $\mathcal{T} = \mathcal{T}(\mathbb{Z}_p) = \{\tau(r,s,t) \mid r,s,t \in \mathbb{Z}_p\}$. We also set $n = v(\Delta)$.

Initially there are no conditions except integrality of the coefficients, so we start in $\mathcal{W}(0,0,0,0,0)$. At each step, we either exit the algorithm based on a divisibility test; or, we divide into subcases. The exit criteria always occur with probability $1/p$. The division into subcases is always into p subcases, except at the beginning where there are p^2 subcases, one for each possibility for the singular point mod p. The subcases occur with equal probabilities, and the relative densities within each subcase are independent of the specific subcase: for example, when there is bad reduction, each of the p^2 points in the affine $\mathbb{F}_p$-plane is equally likely to be the unique singular point, and the densities of each bad reduction type do not depend on which point is singular.

Good reduction. The exit condition is $n = 0$: then $f_p = 0$ and $c_p = 1$. This occurs with probability $1 - 1/p$, by Proposition 5.2. Otherwise (with probability $1/p$), we divide into p^2 equiprobable subcases, proceeding with the case where the point $(0,0)$ is singular (modulo p). So now $\boldsymbol{a} \in \mathcal{W}(0,0,1,1,1)$.

Since Δ is invariant under the whole translation group $\mathcal{T}$, the exit condition is well-defined. We claim that the stabiliser of $\mathcal{W}(0,0,1,1,1)$ in $\mathcal{T}$ is $\mathcal{T}_{1,0,1}$. In one direction this is obvious from the transformation formulas for $\tau(r,s,t)$, which for convenience we recall here:

$$\begin{aligned}
a_1' - a_1 &= R_1 = 2s \\
a_2' - a_2 &= R_2 = -sa_1 + 3r - s^2 \\
a_3' - a_3 &= R_3 = ra_1 + 2t \\
a_4' - a_4 &= R_4 = -sa_3 + 2ra_2 - (t+rs)a_1 + 3r^2 - 2st \\
a_6' - a_6 &= R_6 = ra_4 + r^2a_2 + r^3 - ta_3 - t^2 - rta_1.
\end{aligned}$$

If p divides all of r, t, a_3, a_4, a_6, then it divides a_3', a_4', a_6' also. Conversely, suppose that $\tau(r,s,t)$ preserves $\mathcal{W}(0,0,1,1,1)$. Then $R_3 \equiv R_4 \equiv R_6 \equiv 0$, and

$$r^3 \equiv (rs - t)R_3 + rR_4 - 2R_6 \pmod{p}$$

implies $r \equiv 0$; then $-t^2 \equiv R_6 \equiv 0$ implies $t \equiv 0$.

Multiplicative reduction. Given $\boldsymbol{a} \in \mathcal{W}(0,0,1,1,1)$, the exit condition $v(b_2) = 0$ is that $f = y^2 + a_1 y - a_2$ has distinct roots modulo p. By Lemma 5.4, this occurs with probability $1 - 1/p$, so

$$(9) \quad \mu(\mathcal{W}(0,0,1,1,1 \mid v(b_2) = 0) = \frac{p-1}{p}\mu(\mathcal{W}(0,0,1,1,1) = (p-1)/p^4.$$

Note that this condition is invariant under $\mathcal{T}_{1,0,1}$, since $b_2' = b_2 + 12r \equiv b_2 \pmod{p}$. In this case, $f_p = 1$ and the type is I_m where $m = n$ $(= v(\Delta))$, while the value of c_p depends on the parity of m and on whether the reduction type is split or non-split, which in turn depends on whether or not the roots of f lie in $\mathbb{F}_p$.

In the split case, $c_p = m$, with density $\frac{1}{2}(p-1)^2/p^{m+2}$, for each $m \geq 1$. Relative to the total density of Type $\mathrm{I}_{\geq 1}$, this is $(p-1)/2p^m$.

In the non-split case, $c_p = 1$ for odd m, with total density $\frac{1}{2}\sum_{k=0}^{\infty}(p-1)^2/p^{2k+3} = (p-1)/2p(p+1)$, while $c_p = 2$ for even m, with total density $\frac{1}{2}\sum_{k=1}^{\infty}(p-1)^2/p^{2k+2} = (p-1)/2p^2(p+1)$. Relative to the total density of Type $\mathrm{I}_{\geq 1}$, these are $p/2(p+1)$ and $1/2(p+1)$ respectively.

Otherwise, $v(b_2) \geq 1$ and we move on to the types of additive reduction; after another transformation taking the double root of $f \pmod{p}$ to 0, we have $\boldsymbol{a} \in \mathcal{W}(1,1,1,1,1)$. This translation has the form $\tau(0,s,0) \in \mathcal{T}_{1,0,1}$ with s unique modulo p, so the stabiliser of $\mathcal{W}(1,1,1,1,1)$ is cut down from $\mathcal{T}_{1,0,1}$ to $\mathcal{T}_{1,1,1}$.

For $\boldsymbol{a} \in \mathcal{W}(v_1,v_2,v_3,v_4,v_6)$ we follow Tate's notation in [29] and write $a_{i,v_i} = p^{-v_i}a_i$. In the course of the proof, there are many claims of the form $\Delta \equiv * \pmod{p^k}$, where the right-hand side is in $\mathbb{Z}[a_1,a_2,a_3,a_4,a_6]$, and the claim is made under the assumption that $p^{v_i} \mid a_i$ for $1 \leq i \leq 6$; such claims can all be verified by expanding the difference of both sides and checking that every term has valuation at least k. In all cases, the coefficients of every term are not divisible by any primes other than 2 and 3, which explains why these primes often need separate treatment. While it would be possible to give a simpler proof for $p \geq 5$ only, in term of the invariants c_4 and c_6, we will treat all primes in as uniform a way as possible, for clarity. All these claims may be checked using the `Sage` code at [14].

Additive reduction, Type II. Given $\boldsymbol{a} \in \mathcal{W}(1,1,1,1,1)$, the exit condition for Type II is $v(a_6) = 1$. This is well-defined since for $\boldsymbol{a} \in \mathcal{W}(1,1,1,1,1)$ and $\tau \in \mathcal{T}_{1,1,1}$ we have $v(a_6' - a_6) \geq 2$.

In this case we have $c_p = 1$ and $f_p = n$. Given $\boldsymbol{a} \in \mathcal{W}(1,1,1,1,1)$, we find that $\Delta \equiv -2^4 3^3 a_6^2 \pmod{p^3}$; so $n \geq 2$, and when the exit condition holds (so that $v(a_6) = 1$), we have $n = 2$, provided that $p \geq 5$.

For $p = 3$, we have $\Delta \equiv -a_4^3 \pmod{3^4}$, so $n \geq 3$, with $n = 3 \iff v(a_4) = 1$, which has relative probability $2/3$. Otherwise, $\boldsymbol{a} \in \mathcal{W}(1,1,1,2,{=}1)$, with $\Delta \equiv -3a_2^2 a_6 \pmod{3^5}$, so $n \geq 4$, with $n = 4 \iff v(a_2) = 1$, since $v(a_6) = 1$; this case happens with relative probability $(1/3)(2/3) = 2/9$. Otherwise, $\boldsymbol{a} \in \mathcal{W}(1,2,1,2,{=}1)$, with $\Delta \equiv -3^3 a_6^2 \pmod{3^6}$, so $n = 5$ with the remaining relative probability $1/9$.

For $p = 2$, we have $\Delta \equiv a_3^4 \pmod{2^5}$, so $n \geq 4$, and $n = 4 \iff v(a_3) = 1$, which has relative probability $1/2$. Otherwise, $\boldsymbol{a} \in \mathcal{W}(1,1,2,2,{=}1)$, with $\Delta \equiv a_1^4 a_4^2 - 2^4 a_6^2 \equiv a_1^4 a_4^2 - 2^6 \pmod{2^7}$, so $n \geq 6$, with $n = 6 \iff v(a_1^4 a_4^2) = 6$; this case happens when either $v(a_1) \geq 2$ or $v(a_4) \geq 3$, so with relative probability $(1/2)(3/4) = 3/8$. Assuming that both $v(a_1) = 1$ and $v(a_4) = 2$, we find that $\Delta \equiv 2^7 \pmod{2^8}$, so $n = 7$ with the remaining relative probability $1/8$.

Otherwise, we have $\boldsymbol{a} \in \mathcal{W}(1,1,1,1,2)$, with unchanged stabiliser $\mathcal{T}_{1,1,1}$.

Additive reduction, Type III. Given $\boldsymbol{a} \in \mathcal{W}(1,1,1,1,2)$, the exit condition for Type III is $v(a_4) = 1$. This is well-defined since for $\boldsymbol{a} \in \mathcal{W}(1,1,1,1,2)$ and $\tau \in \mathcal{T}_{1,1,1}$ we have $v(a_4' - a_4) \geq 2$.

In this case we have $c_p = 2$ and $f_p = n - 1$. Now we have $\Delta \equiv -2^6 a_4^3 \pmod{p^4}$, so $n = 3$ and $f_p = 2$ for $p \geq 3$, since $v(a_4) = 1$.

For $p = 2$ and $\boldsymbol{a} \in \mathcal{W}(1,1,1,{=}1,2)$, we have $\Delta \equiv a_3^4 \pmod{2^5}$, so $n \geq 4$, with $n = 4 \iff v(a_3) = 1$, which happens with relative probability $1/2$. Otherwise, $\boldsymbol{a} \in \mathcal{W}(1,1,2,{=}1,2)$, and we have $\Delta \equiv 2^2 a_1^4 \pmod{2^7}$, so $n \geq 6$, with $n = 6 \iff v(a_1) = 1$; this case has relative probability $(1/2)(1/2) = 1/4$. Otherwise, $\boldsymbol{a} \in \mathcal{W}(2,1,2,{=}1,2)$, and we have $\Delta \equiv 2^8(a_{2,1}^2 + a_{3,2}^4 + a_{6,2}^2) \pmod{2^9}$, so $n \geq 8$, with $n = 8 \iff 2 \nmid a_{2,1} + a_{3,2} + a_{6,2}$; this case has relative probability $(1/4)(1/2) = 1/8$. Finally, assuming that $a_6 \equiv 2a_2 + a_3 \pmod{8}$ (so that $a_{6,2} \equiv a_{2,1} + a_{3,2} \pmod{2}$), we find that $\Delta \equiv 2^9 \pmod{2^{10}}$, so $n = 9$ with the remaining relative probability $1/8$.

The relative probabilities for $n = 4, 6, 8$, and 9 (respectively, $f_2 = 3, 5, 7$, and 8) are therefore $1/2, 1/4, 1/8$, and $1/8$.

Otherwise, we have $\boldsymbol{a} \in \mathcal{W}(1,1,1,2,2)$, with unchanged stabiliser $\mathcal{T}_{1,1,1}$.

Additive reduction, Type IV. Given $\boldsymbol{a} \in \mathcal{W}(1,1,1,2,2)$, the exit condition for Type IV is that the quadratic $f = y^2 + a_{3,1}y - a_{6,2}$ has distinct roots modulo p, or equivalently that $v(b_6) = 2$. This condition is well-defined, since for $\tau \in \mathcal{T}_{1,1,1}$ we have $v(b_6' - b_6) \geq 3$. Using Lemma 5.4 again, we have

$$(10) \quad \mu(\mathcal{W}(1,1,1,2,2 \mid v(b_6) = 2)) = \frac{p-1}{p}\mu(\mathcal{W}(1,1,1,2,2)) = (p-1)/p^8.$$

Now $f_p = n-2$, and $c_p = 1$ or 3, according to whether the roots of f are in $\mathbb{F}_p$ or not, which have relative probability 1/2 each; it remains to determine the possible values of the discriminant valuation n and their relative densities.

For $\boldsymbol{a} \in \mathcal{W}(1,1,1,2,2 \mid v(b_6)=2)$, we have $\Delta \equiv -3^3 b_6^2 \pmod{p^5}$, so for $p \neq 3$ we have $n=4$ and $f_p = 2$.

For $p=3$, we have $\Delta \equiv -a_2^3 b_6 \pmod{3^6}$, so $n \geq 5$, with $n=5 \iff v(a_2)=1$. Otherwise, $\boldsymbol{a} \in \mathcal{W}(1,2,1,2,2 \mid v(b_6)=2)$, and we have $\Delta \equiv b_4^3 \pmod{3^7}$. Note that $b_4 = a_1 a_3 + 2a_4$, so $v(b_4) \geq 2$. Hence $n \geq 6$, and $n=6 \iff v(b_4)=2 \iff a_4 \not\equiv a_1 a_3 \pmod{3^3}$. Assuming that $a_4 \equiv a_1 a_3 \pmod{3^3}$, so that $v(b_4) \geq 3$, we find that $\Delta \equiv -3^3 b_6^2 \pmod{3^8}$, so $n=7$. Thus for $p=3$, we have $n = 5, 6$, or 7 and $f_3 = 3, 4$ or 5 with relative probabilities 2/3, 2/9, and 1/9 respectively.

Otherwise, $v(b_6) \geq 3$, so the quadratic $y^2 + a_{3,1}y - a_{6,2}$ has a repeated root. A transformation τ in a unique coset of $\mathcal{T}_{1,1,2}$ in $\mathcal{T}_{1,1,1}$ takes the root to 0 and hence the coefficients into $\mathcal{W}(1,1,2,2,3)$, with stabiliser $\mathcal{T}_{1,1,2}$.

Additive reduction, Type I_0^*. Given $\boldsymbol{a} \in \mathcal{W}(1,1,2,2,3)$, the exit condition for Type I_0^* is $v(\operatorname{disc}(g)) = 6$, where $g = x^3 + a_2 x^2 + a_4 x + a_6$. Equivalently, the condition is that $g_1(x) = g(px)/p^3 = x^3 + a_{2,1}x^2 + a_{4,2}x + a_{6,3}$ should have distinct roots in $\overline{\mathbb{F}}_p$, since $v(\operatorname{disc}(g_1)) = v(\operatorname{disc}(g)) - 6$. Note that after transforming the equation by $\tau_{r,s,t} \in \mathcal{T}_{1,1,2}$, $g_1(x)$ becomes $g_1(x + r/p)$, so the condition is well-defined.

Now, c_p is equal to one more than the number of roots of g_1 in $\mathbb{F}_p$. By Lemma 5.5, this number is 0, 1 or 3 with relative probabilities $(p+1)/(3p)$, 1/2 and $(p-2)/(6p)$ respectively. We have $\Delta \equiv 16 \operatorname{disc}(g) \pmod{p^7}$, so for $p \neq 2$, the exit condition implies $n=6$, and then $f_p = n - 4 = 2$.

Now let $p=2$. For $\boldsymbol{a} \in \mathcal{W}(1,1,2,2,3)$, we have $\operatorname{disc}(g) \equiv a_6^2 - a_2^2 a_4^2 \pmod{2^7}$, so the exit condition implies $a_6 \not\equiv a_2 a_4 \pmod{2^7}$. Assuming that $a_6 \not\equiv a_2 a_4 \pmod{2^7}$, we find that $\Delta \equiv a_1^4 a_4^2 - a_3^4 \pmod{2^9}$, so $n \geq 8$, with $n = 8 \iff a_{3,2} \not\equiv a_{1,1}a_{4,2} \pmod{2}$. Assuming further that $a_{3,2} \equiv a_{1,1}a_{4,2} \pmod{2}$, we have $\Delta \equiv 2^8 a_1 \pmod{2^{10}}$, so $n \geq 9$, with $n=9$ if and only if $v(a_1) = 1$. When $v(a_1) \geq 2$, then also $v(a_3) \geq 3$, and these imply that $\Delta \equiv 2^{10} \pmod{2^{11}}$, giving $n = 10$.

The preceding analysis shows that for type I_0^* curves when $p=2$ we have $n = v(\Delta) = 8$, 9, or 10, and respectively $f_2 = 4$, 5, or 6, with relative probabilities 1/2, 1/4, and 1/4.

Additive reduction, Type I_m^*, $m \geq 1$. The exit condition for type I_0^* fails when $g_1(x)$ has a repeated root modulo p. We can move this root to zero using a transform in a unique coset of $\mathcal{T}_{2,1,2}$ in $\mathcal{T}_{1,1,2}$, after which

$\boldsymbol{a} \in \mathcal{W}(1,1,2,3,4)$, with stabiliser now $\mathcal{T}_{2,1,2}$. The condition for type I_m^* is that the repeated root is only a double root, which (after the transform) is that $v(a_2) = 1$.

Looking at the details of Tate's algorithm in this case, it proceeds in a sequence of substeps: at each substep the value of m is incremented; there is an exit condition that a monic quadratic mod p has distinct roots; and the value of c_p depends on whether this quadratic has roots in $\mathbb{F}_p$ (in which case $c_p = 4$) or not ($c_p = 2$). So, overall, each of these two values occurs in half the cases, by Lemma 5.4. Moreover, the stabiliser index increases by a factor of p at each stage, since when the quadratic has a double root we can move it to 0 with a transform in a uniquely determined coset of an index p subgroup of the current stabiliser.

We have $f_p = n - m - 4$. We treat separately the cases $p \geq 3$, where we will see that $f_p = 2$ always, and $p = 2$. The case $p \geq 3$ is well-known (see Kraus [22] or Kobayashi [20]), but we include the details here since the analysis is similar to that required for $p = 2$.

Write

$$\mathcal{W}_{\mathrm{odd}}(k) = \mathcal{W}(1, {=}1, k+1, k+2, 2k+2),$$
$$\mathcal{W}_{\mathrm{even}}(k) = \mathcal{W}(1, {=}1, k+2, k+2, 2k+3),$$

so initially, $\boldsymbol{a} \in \mathcal{W}(1, {=}1, 2, 3, 4) = \mathcal{W}_{\mathrm{odd}}(1)$. The exit conditions are:

- for $\boldsymbol{a} \in \mathcal{W}_{\mathrm{odd}}(k)$: that $y^2 + a_{3,k+1}y - a_{6,2k+2}$ has distinct roots over $\mathbb{F}_p$, or equivalently that $v(b_6) = 2k+2$, and for $p = 2$ to $v(a_3) = k+1$;
- for $\boldsymbol{a} \in \mathcal{W}_{\mathrm{even}}(k)$: that $x^2 + a_{4,k+2}x + a_{6,2k+3}$ have distinct roots over $\mathbb{F}_p$, equivalently that $v(b_8) = 2k+4$, or that $v(a_4) = k+2$ when $p = 2$.

First assume that $p \neq 2$. For $\boldsymbol{a} \in \mathcal{W}_{\mathrm{odd}}(k)$ we have $\Delta \equiv -2^4 a_2^3 b_6 \pmod{p^{2k+6}}$; since $v(a_2) = 1$, when the exit condition $v(b_6) = 2k+2$ holds, for $p \neq 2$ we have $n = 2k+5$ exactly. Hence $n = m+6$ and $f_p = 2$. Otherwise, after shifting the double root to 0 by a suitable translation, we arrive in $\mathcal{W}_{\mathrm{even}}(k)$, where $\Delta \equiv -2^4 b_8 a_2^2 \pmod{p^{2k+7}}$; when the exit condition $v(b_8) = 2k+4$ holds, we have $v(2^4 b_8 a_2^2) = 2k+6$, so $n = 2k+6$. Again, $n = m+6$ and $f_p = 2$. Otherwise, after another shift we arrive in $\mathcal{W}_{\mathrm{odd}}(k+1)$, so we increment k and repeat.

Hence for $p \geq 3$, we always have $f_p = 2$.

Now let $p = 2$. Again, the value of m is initialized to 1 and we proceed recursively; at each stage we either exit (always with relative probability 1/2), or increment m. The recursive steps alternate in nature depending on the parity of m; after the first three cases ($m = 1, 2, 3$) which are slightly different, all the remaining cases may be dealt with generically.

At first, $m = 1$ with $\boldsymbol{a} \in \mathcal{W}_{\text{odd}}(1) = \mathcal{W}(1, {=}1, 2, 3, 4)$, where we have $\Delta \equiv a_3^4 \pmod{2^9}$. When the exit condition $v(a_3) = 2$ holds, we have $n = 8$ and $f_2 = 3$, and Type I_1^*.

Otherwise, $v(a_3) \geq 3$, and we shift y so that the quadratic $y^2 + a_{3,2}^2 y - a_{6,4}$ has double root at $y \equiv 0 \pmod 2$, so that $\boldsymbol{a} \in \mathcal{W}_{\text{even}}(1) = \mathcal{W}(1, {=}1, 3, 3, 5)$, and we increment m to 2.

Now we have $\Delta \equiv a_1^4 a_4^2 \pmod{2^{11}}$, so $n \geq 10$. When the exit condition $v(a_4) = 3$ holds, either $v(a_1) = 1$, giving $n = 10$ and $f_2 = 4$; or $v(a_1) \geq 2$, and $\boldsymbol{a} \in \mathcal{W}(2, {=}1, 3, {=}3, 5)$. In the latter case, $\Delta \equiv a_3^4 + 2^{12} \pmod{2^{14}}$, so $n = 12$ if $v(a_3) \geq 4$, and $n = 13$ if $v(a_3) = 3$. Hence for Type I_2^* we have $n = 10, 12$, or 13 (respectively, $f_2 = 4, 6$, or 7) with relative probabilities $1/2, 1/4$, and $1/4$.

Otherwise, when the exit condition at $m = 2$ fails, we have $v(a_4) \geq 4$, and we shift x so that the quadratic $x^2 + a_{4,3}x + a_{6,5}$ has double root at $x \equiv 0 \pmod 2$, so that $\boldsymbol{a} \in \mathcal{W}_{\text{odd}}(2) = \mathcal{W}(1, {=}1, 3, 4, 6)$ and we increment m to 3.

Now we have $\Delta \equiv 2a_1^4 a_3^2 \pmod{2^{12}}$, so $n \geq 11$. When the exit condition $v(a_3) = 3$ holds, either $v(a_1) = 1$, giving $n = 11$ and $f_2 = 4$; or $v(a_1) \geq 2$, and $\boldsymbol{a} \in \mathcal{W}(2, {=}1, {=}3, 4, 6)$. In the latter case, $\Delta \equiv 2^{12} \pmod{2^{13}}$, so $n = 12$ and $f_2 = 5$. Hence for Type I_3^*, we have $n = 11$ or 12 (respectively, $f_2 = 4$ or 5) with equal probability.

Now let $m = 2k \geq 4$, with $\boldsymbol{a} \in \mathcal{W}_{\text{even}}(k)$, and exit condition $v(a_4) = k + 2$. Then $\Delta \equiv a_1^4 a_4^2 \pmod{2^{2k+9}}$, so $n \geq v(a_1^4 a_4^2) = 2k + 8$. Assuming that the exit condition $v(a_4) = k + 2$ holds, we have $n = 2k + 8 = m + 8$ and $f_2 = 4$, provided that $v(a_1) = 1$. Otherwise, $v(a_1) \geq 2$, and now $\Delta \equiv 2^{2k+10} \pmod{2^{2k+11}}$, so $n = 2k + 10 = m + 10$ and $f_2 = 6$. Thus for $m = 2k \geq 4$ we have $f_2 = 4$ or $f_2 = 6$, with equal probability.

If the exit condition fails, $v(a_4) \geq k + 3$, and we may shift x so that the quadratic $x^2 + a_{4,k+2}x + a_{6,2k+3}$ has its double root at $x \equiv 0 \pmod 2$, so also $v(a_6) \geq 2k + 4$ and $\boldsymbol{a} \in \mathcal{W}_{\text{odd}}(k + 1)$. Incrementing both k and m so that $m = 2k - 1$, we have $\boldsymbol{a} \in \mathcal{W}_{\text{odd}}(k)$.

Next, $m = 2k - 1 \geq 5$, with $\boldsymbol{a} \in \mathcal{W}_{\text{odd}}(k)$ and exit condition $v(a_3) = k + 1$. Now, $\Delta \equiv 2a_1^4 a_3^2 \pmod{2^{2k+8}}$, so $n \geq v(2a_1^4 a_3^2) = 2k + 7$. Assuming that the exit condition $v(a_3) = k + 1$ holds, we have $n = 2k + 7 = m + 8$ and $f_2 = 4$, provided that $v(a_1) = 1$. Otherwise, $v(a_1) \geq 2$, and now $\Delta \equiv 2^{2k+9} \pmod{2^{2k+10}}$, so $n = 2k + 9 = m + 10$ and $f_2 = 6$. Thus for $m = 2k - 1 \geq 5$, we again have $f_2 = 4$ or $f_2 = 6$ with equal probability.

If the exit condition fails, $v(a_3) \ge k+2$, and we may shift y so that the quadratic $y^2 + a_{3,k+1}y - a_{6,2k+2}$ has double root at $y \equiv 0$, so that $\boldsymbol{a} \in \mathcal{W}_{\text{even}}(k)$, and we increment m to $2k$ and recurse.

Taking all Types I_m^* for $m \ge 1$ together, we find that $f_2 = 3, 4, 5, 6$, or 7 with relative probabilities $1/2, 1/4, 1/16, 1/8$, and $1/16$.

This completes the analysis of type I_m^*.

Additive reduction, Type IV*. The exit condition for type I_m^* fails when the cubic $g(x)$ has a triple root; after the transform moving the root to 0, this means that $v(a_2) \ge 2$, so $\boldsymbol{a} \in \mathcal{W}(1,2,2,3,4)$, with the same stabiliser as for $\mathcal{W}(1,{=}1,2,3,4)$, namely $\mathcal{T}_{2,1,2}$.

The exit condition for type IV* is that $f = y^2 + a_{3,2}y - a_{6,4}$ has distinct roots modulo p, or equivalently $v(b_6) = 4$, which happens with probability $1 - 1/p$. Thus

$$\mu(\mathcal{W}(1,2,2,3,4 \mid v(b_6)=4)) = \frac{p-1}{p}\mu(\mathcal{W}(1,2,2,3,4)) = (p-1)/p^{13}. \tag{11}$$

Now, $c_p = 1$ or $c_p = 3$, depending on whether the roots are in $\mathbb{F}_p$ or not, and these have equal probability by Lemma 5.4.

To compute $f_p = n - 6$, we first note that for $\boldsymbol{a} \in \mathcal{W}(1,2,2,3,4)$ we have $\Delta \equiv -3^3 b_6^2 \pmod{3^9}$; when $v(b_6) = 4$, this implies that $n = 8$ and $f_p = 2$ provided that $p \ne 3$.

Now consider $p = 3$. We have $v(b_2) \ge 2$, $v(b_4) \ge 3$, and $\Delta \equiv b_4^3 \pmod{3^{10}}$, so $n \ge 9$, and $n = 9 \iff v(b_4) = 3$, which is equivalent to $a_{4,3} \not\equiv a_{1,1}a_{3,2} \pmod 3$, so has relative probability $2/3$. Assuming that $a_{4,3} \equiv a_{1,1}a_{3,2} \pmod 3$, we find that $\Delta \equiv -3^4 b_2 b_6 \pmod{3^{10}}$. Hence $n \ge 10$, with $n = 10 \iff v(b_2) = 2 \iff a_{1,1}^2 + a_{2,2} \not\equiv 0 \pmod 3$. Assuming further that $a_{2,2} \equiv -a_{1,1}^2 \pmod 3$, we have $\Delta \equiv -3^3 b_6^2 \pmod{3^{12}}$, so $n = 11$ exactly. Hence for $p = 3$ we have $n = 9, 10$, or 11 (respectively, $f_3 = 3, 4$, or 5) with relative probabilities $2/3, 2/9$, and $1/9$.

Additive reduction, Type III*. When the exit condition for type IV* fails, we move the root of the quadratic to 0 using a transform in a unique coset of $\mathcal{T}_{2,1,3}$ in $\mathcal{T}_{2,1,2}$ to arrive in $\mathcal{W}(1,2,3,3,5)$ with stabiliser $\mathcal{T}_{2,1,3}$.

The exit condition for type III* is $v(a_4) = 3$. In all cases we have $c_p = 2$. To compute $f_p = n - 7$, we first note that for $\boldsymbol{a} \in \mathcal{W}(1,2,3,3,5)$ we have $\Delta \equiv -2^6 a_4^3 \pmod{p^{10}}$; when $v(a_4) = 3$, this implies that $n = 9$ and $f_p = 2$, provided that $p \ne 2$.

Let $p = 2$. For $\boldsymbol{a} \in \mathcal{W}(1,2,3,3,5)$ we now have $\Delta \equiv a_1^4 a_4^2 \pmod{2^{11}}$, so $n \ge 10$, and when the exit condition $v(a_4) = 3$ holds, we have $n =$

$10 \iff v(a_1) = 1$. Assuming that $v(a_1) \geq 2$, so $\boldsymbol{a} \in \mathcal{W}(2,2,3,3,5)$, we have $\Delta \equiv a_3^4 \pmod{2^{13}}$, so $n \geq 12$, with $n = 12 \iff v(a_3) = 3$. Assuming further that $v(a_3) \geq 4$, so $\boldsymbol{a} \in \mathcal{W}(2,2,4,3,5)$, we have $\Delta \equiv 2^4(2^2a_1^4 + 2^6a_2^2 + a_6^2) \equiv 2^{14}(a_{1,2}^4 + a_{2,2}^2 + a_{6,5}^2) \pmod{2^{15}}$, so $n \geq 14$, with $n = 14 \iff a_{6,5} \not\equiv a_{1,2} + a_{2,2} \pmod 2$. Assuming that $a_{6,5} \equiv a_{1,2} + a_{2,2} \pmod 2$, we find that $\Delta \equiv 2^{15} \pmod{2^{16}}$, so that $n = 15$.

Hence for $p = 2$ we have $n = 10, 12, 14$, or 15 (respectively, $f_2 = 3, 5, 7$, or 8) with relative probability $1/2, 1/4, 1/8$ and $1/8$.

Additive reduction, Type II*. When the exit condition for type III* fails we are in $\mathcal{W}(1,2,3,4,5)$ with the same stabiliser $\mathcal{T}_{2,1,3}$, since $\mathcal{T}_{2,1,3}$ preserves the condition $v(a_3) = 3$.

The exit condition for type II* is $v(a_6) = 5$. In all cases we have $c_p = 1$, and $f_p = n - 8$.

For $\boldsymbol{a} \in \mathcal{W}(1,2,3,4,5)$, we have $\Delta \equiv -2^4 3^3 a_6^2 \pmod{p^{11}}$, so when the exit condition holds we have $n = 10$ and $f = 2$ for all $p \geq 5$.

Let $p = 3$. Now, $v(b_2) \geq 2$, and $\Delta \equiv -a_6 b_2^3 \pmod{3^{12}}$; hence $n \geq 11$, with $n = 11 \iff v(b_2) = 2 \iff a_{2,2} \not\equiv -a_{1,1}^2 \pmod 3$. Assuming that $a_{2,2} \equiv -a_{1,1}^2 \pmod 3$, we find that $\Delta \equiv b_4^3 \pmod{3^{13}}$, so $n \geq 12$, with $n = 12 \iff v(b_4) = 3 \iff a_{4,4} \not\equiv a_{1,1}a_{3,3} \pmod 3$. Assuming further that $a_{4,4} \equiv a_{1,1}a_{3,3} \pmod 3$, we find that $\Delta \equiv -3^3 a_6 \pmod{3^{14}}$, so $n = 13$. Hence for $p = 3$ we have $n = 11, 12$, or 13 (respectively, $f_3 = 3, 4$, or 5) with relative probabilities $2/3, 2/9$, and $1/9$.

Finally, let $p = 2$. Now $\Delta \equiv a_1^6 a_6 \pmod{2^{12}}$, so $n \geq 11$, with $n = 11 \iff v(a_1) = 1$. If $v(a_1) \geq 2$, then $\Delta \equiv a_3^4 \pmod{2^{13}}$, so $n \geq 12$, with $n = 12 \iff v(a_3) = 3$. If also $v(a_3) \geq 4$, then $\Delta \equiv 2^{14} \pmod{2^{15}}$, so $n = 14$. Hence for $p = 2$ we have $n = 11, 12$, or 14 (respectively, $f_2 = 3, 4$, or 6) with relative probability $1/2, 1/4$ and $1/4$.

When the exit condition for type II* fails we are in $\mathcal{W}(1,2,3,4,6)$ with the same stabiliser $\mathcal{T}_{2,1,3}$.

This completes the proof of Theorem 5.3. □

5.2. Distribution of conductor exponents

Finally, we collect together the possible conductor exponents over all reduction types, to find the overall density of each. Here we omit non-minimal models, so the densities add up to $1 - 1/p^{10}$.

Theorem 5.6 (Overall distribution of conductor exponents). *The overall densities of conductor exponents f_p for minimal Weierstrass models over $\mathbb{Z}_p$ are as follows:*

1) *Good and multiplicative reduction*

f_p	*density*
0	$1-1/p$
1	$1/p-1/p^2$

2) *Additive reduction.*
$p \geq 5$.

f_p	*density*
2	$1/p^2-1/p^{10}$

$p=3$. *The following densities add up to* $59040/3^{12} = 1/3^2 - 1/3^{10}$*:*

f_p	*density*
2	$15120/3^{12}$
3	$29280/3^{12}$
4	$9760/3^{12}$
5	$4880/3^{12}$

$p=2$. *The following densities add up to* $1020/2^{12} = 1/2^2 - 1/2^{10}$*:*

f_p	*density*
2	$144/2^{12}$
3	$150/2^{12}$
4	$297/2^{12}$
5	$84/2^{12}$
6	$213/2^{12}$
7	$99/2^{12}$
8	$33/2^{12}$

Proof. Immediate from 5.3. □

Acknowledgements:

The authors would like to thank Manjul Bhargava, Adam Harper, Bjorn Poonen, and Michael Stoll for many useful suggestions; Ignazio Longhi for the correction needed in (7); and the anonymous referee, whose comments led to an improved exposition, supported by computer code now available at [14], of the details in Section 5. The first author was supported by EPSRC Programme Grant EP/K034383/1 *LMF: L-Functions and Modular Forms*, the Horizon 2020 European Research Infrastructures project *OpenDreamKit* (#676541), and the Heilbronn Institute for Mathematical Research.

References

[1] Avner Ash, Jos Brakenhoff, and Theodore Zarrabi, *Equality of polynomial and field discriminants*, Experiment. Math. **16** (2007), no. 3, 367–374.

[2] Stephan Baier, *Elliptic curves with square-free Δ*, Int. J. Number Theory **12** (2016), no. 3, 737–764.

[3] Stephan Baier and Tim D. Browning, *Inhomogeneous cubic congruences and rational points on del Pezzo surfaces*, J. Reine Angew. Math. **680** (2013), 69–151.

[4] Ebru Bekyel, *The density of elliptic curves having a global minimal Weierstrass equation*, Journal of Number Theory **109** (2004), no. 1, 41–58.

[5] Michael Bennett, *On the representation of unity by binary cubic forms*, Transactions of the American Mathematical Society **353** (2001), no. 4, 1507–1534.

[6] Manjul Bhargava, *The geometric sieve and the density of squarefree values of invariant polynomials*, Inventiones Mathematicae **228** (2022), 1037–1073.

[7] Manjul Bhargava, John Cremona, and Tom Fisher, *The proportion of plane cubic curves over $\mathbb{Q}$ that everywhere locally have a point*, Int. J. Number Theory **12** (2016), no. 4, 1077–1092.

[8] Manjul Bhargava, John E. Cremona, Tom Fisher, Nick G. Jones, and Jonathan P. Keating, *What is the probability that a random integral quadratic form in n variables has an integral zero?*, Int. Math. Res. Not. IMRN (2016), no. 12, 3828–3848.

[9] Manjul Bhargava, Arul Shankar, and Xiaoheng Wang, *Squarefree values of polynomial discriminants I*, preprint, `arXiv:1611.09806`, 2016.

[10] Armand Brumer and Oisín McGuinness, *The behavior of the Mordell-Weil group of elliptic curves*, Bulletin of the American Mathematical Society **23** (1990), no. 2, 375–382.

[11] Peter J. Cho and Keunyoung Jeong, *Counting elliptic curves with local conditions and its applications*, preprint, `arXiv:2003.09102`, 2020.

[12] J. E. Cremona, *Algorithms for modular elliptic curves*, second ed., Cambridge University Press, 1997, available from `http://www.warwick.ac.uk/staff/J.E.Cremona/book/fulltext/index.html`.

[13] J. E. Cremona, *Computing in component groups of elliptic curves*, Algorithmic Number Theory: 8th International Symposium, ANTS-VIII, Banff, Canada. Springer, 2008, pp. 118–124.

[14] J. E. Cremona and M. Sadek, *Local and global densities for Weierstrass models of elliptic curves*, preprint and auxiliary files of `Sage` code, `arXiv:2003.08454`, 2021.

[15] Luca Demangos and Ignazio Longhi, *Densities on Dedekind domains, completions and Haar measure*, preprint, `arXiv:2009.04229`, 2020.

[16] Torsten Ekedahl, *An Infinite Version of the Chinese Remainder Theorem*, Comment. Math. Univ. St. Paul. **40** (1991), no. 1, 53–59.

[17] J.-H. Evertse, *On the representation of integers by binary cubic forms of positive discriminant*, Inventiones Mathematicae **73** (1983), no. 1, 117–138.

[18] Andrew Granville, *ABC allows us to count squarefrees*, Internat. Math. Res. Notices (1998), no. 19, 991–1009.

[19] C. Hooley, *On the power free values of polynomials*, Mathematika **14** (1967), 21–26.

[20] Shin-ichi Kobayashi, *The local root number of elliptic curves with wild ramification*, Math. Ann. **323** (2002), no. 3, 609–623.

[21] Alain Kraus, *Quelques remarques à propos des invariants c_4, c_6 et Δ d'une courbe elliptique*, Acta Arith. **54** (1989), 75–80.

[22] ______, *Sur le défaut de semi-stabilité des courbes elliptiques à réduction additive*, Manuscripta Math. **69** (1990), no. 4, 353–385.

[23] Ioannis Papadopoulos, *Sur la classification de Néron des courbes elliptiques en caractéristique résiduelle 2 et 3*, J. Number Theory, **44:2** (1993), 119–152.

[24] Bjorn Poonen, *Squarefree values of multivariable polynomials*, Duke Math. J. **118** (2003), no. 2, 353–373.

[25] Bjorn Poonen and Michael Stoll, *The Cassels-Tate pairing on polarized abelian varieties*, Ann. of Math. (2) **150** (1999), no. 3, 1109–1149.

[26] ______, *A local-global principle for densities*, Topics in number theory (University Park, PA, 1997), Math. Appl., vol. 467, Kluwer Acad. Publ., Dordrecht, 1999, pp. 241–244.

[27] Bjorn Poonen and José Felipe Voloch, *Random Diophantine equations*, Arithmetic of higher-dimensional algebraic varieties (Palo Alto, CA, 2002), Progr. Math., vol. 226, Birkhäuser Boston, Boston, MA, 2004, With appendices by Jean-Louis Colliot-Thélène and Nicholas M. Katz, pp. 175–184.

[28] Mohammad Sadek, *Counting elliptic curves with bad reduction over a prescribed set of primes*, preprint, `arXiv:1704.02056`, 2017.

[29] John Tate, *Algorithm for determining the type of a singular fiber in an elliptic pencil*, Modular functions of one variable IV (Springer Lecture Notes in Mathematics 476), Springer 1975, 33–52.

[30] Mark Watkins, *Some heuristics about elliptic curves*, Experimental Mathematics **17** (2008), no. 1, 105–125.

[31] Siman Wong, *On the density of elliptic curves*, Compositio Math. **127** (2001), no. 1, 23–54.

Mathematics Institute, University of Warwick
Coventry CV4 7AL, UK
E-mail address: `j.e.cremona@warwick.ac.uk`

Faculty of Engineering and Natural Sciences, Sabanci University
Tuzla, İstanbul, 34956 Turkey
E-mail address: `mmsadek@sabanciuniv.edu`

Received June 16, 2020
Accepted November 8, 2021