J. reine angew. Math. **829** (2025), 1–36
DOI 10.1515/crelle-2025-0067 Journal für die reine und angewandte Mathematik

Equations involving the modular j-function and its derivatives

By *Vahagn Aslanyan* at Manchester, *Sebastian Eterović* at Wien
and *Vincenzo Mantova* at Leeds

Abstract. We show that, for any polynomial $F(X, Y_0, Y_1, Y_2) \in \mathbb{C}[X, Y_0, Y_1, Y_2]$, the equation $F(z, j(z), j'(z), j''(z)) = 0$ has a Zariski dense set of solutions in the hypersurface $F(X, Y_0, Y_1, Y_2) = 0$, unless F is in $\mathbb{C}[X]$ or it is divisible by Y_0, $Y_0 - 1728$, or Y_1. Our methods establish criteria for finding solutions to more general equations involving periodic functions. Furthermore, they produce a qualitative description of the distribution of these solutions.

1. Introduction

The problem of determining which (systems of) equations involving certain classical transcendental functions of a complex variable have solutions is a natural question at the intersection between complex geometry, model theory, and number theory. In complex geometry, it is a form of analytic Nullstellensatz for the given functions; in model theory, it plays an important role in the definability properties of the functions involved; and in number theory, it is related to Schanuel's conjecture and its analogues (given by special cases of the Grothendieck–André generalised period conjecture). Often, the function under consideration is of arithmetic importance. Examples of such classical functions are the exponential functions of semi-abelian varieties and Fuchsian automorphic functions. In this paper, we focus on the modular j-function and its derivatives.

The first conjecture in this area arose from Zilber's work on the model theory of complex exponentiation [16–18]. It is now referred to as the *Exponential (Algebraic) Closedness* conjecture or *Zilber's Nullstellensatz*, and predicts when systems of equations involving addition, multiplication, and complex exponentiation have solutions in the complex numbers. We refer to the general version of the problem as *Existential Closedness*, or EC for short. An EC

The corresponding author is Vahagn Aslanyan.

Vahagn Aslanyan was supported by Leverhulme Trust Early Career Fellowship ECF-2022-082 at the University of Leeds (where most of this work was done) and by EPSRC Fellowship EP/X009823/1 and DKO Fellowship at the University of Manchester. Sebastian Eterović and Vincenzo Mantova were supported by EPSRC Fellowship EP/T018461/1 at the University of Leeds.

conjecture for the j-function was proposed in [4, §1]; in geometric terms, it states that any algebraic variety $V \subseteq \mathbb{C}^{2n}$ satisfying geometric conditions known as *freeness* and *broadness* intersects the n-fold graph of the j-function. The definition of these geometric conditions is long and will not be used in the present work, so we refer the interested reader to [4, §2.2], but informally, freeness and broadness ensure that the equations defining V do not break any functional properties of j coming from the linear-fractional action of $\mathrm{GL}_2^+(\mathbb{Q})$ (where $+$ denotes positive determinant) on the upper half-plane, as well as not contradicting a conjecture on transcendental values of the j-function analogous to Schanuel's conjecture for exponentiation (see [4, Conjecture 1.1] and [3, §6.3] for the statement of this conjecture).

If one somehow knows that an algebraic variety V does intersect the graph of j, a very natural next question is to determine how these intersection points are distributed within V. For instance, one may ask whether these points are Zariski dense in V. We remark that if $V \subseteq \mathbb{C}^{2n}$ satisfies the above-mentioned geometric conditions of freeness and broadness, then for any Zariski open subset $V' \subseteq V$, it is possible to construct an algebraic variety $W \subseteq \mathbb{C}^{2(n+1)}$ which is also free and broad and projects onto V'. Thus, if we assume EC then, applying it to W, we deduce that V' intersects the graph of j. Since V' was an arbitrary Zariski open subset of V, we conclude that the intersection of V with the graph of j is Zariski dense in V.

In the same work [4], the authors also proposed an extension of the conjecture incorporating the derivatives of j (see [4, Conjecture 1.6]). This version of EC is often referred to as *Existential Closedness with Derivatives*, or ECD for short. This time the variety V in question is a subset of $\mathbb{C}^{4n}$, and the conjecture states that if V satisfies analogous geometric notions of freeness and broadness (again related to a form of Schanuel's conjecture, now involving j and its derivatives), then V intersects the n-fold graph of the map $z \mapsto (j(z), j'(z), j''(z))$. For the definitions and precise statements of these conjectures, see [1, 2, 4]. Note that we do not consider the third and higher derivatives of j as these are rational over j, j', j'' (see (2.2)). As with EC, ECD implies that if $V \subseteq \mathbb{C}^{4n}$ satisfies the geometric conditions of freeness and broadness, then V has a Zariski dense set of points of the desired form.

Very few cases of ECD have been proven, in comparison to EC where various families of varieties in $\mathbb{C}^{2n}$ have been shown to satisfy the conjecture. Prior to the present work, only very special cases of ECD had been solved, proving solvability of some simple equations involving just j' (so not combining it with j or j''); see [9, 10]. An ECD statement for "blurrings" (certain multi-valued twists) of j was obtained in [4]. All of these papers mostly focused on, and established stronger results for, the EC conjecture for the j-function (without derivatives). These results are analogous to their exponential counterparts, namely, [5–7, 11, 12, 16]. Although different methods have been used across these works, one common feature is that they all exploit in some way the periodicity of exp or the $\mathrm{SL}_2(\mathbb{Z})$-invariance of j. Incorporating the derivatives of j into the equations presents then a significant new challenge, as j' and j'' are no longer $\mathrm{SL}_2(\mathbb{Z})$-invariant. It is also worth noting that a differential version of ECD was obtained in [2], and that it is so far the only setting where a full Existential Closedness statement is proved for j with derivatives, and the same method also applies to exp.

In this article, we prove the ECD conjecture when $n = 1$, which precisely states that any algebraic variety $V \subseteq \mathbb{C}^4$ of dimension 3 without constant coordinates contains (a Zariski dense set of) points of the form $(z, j(z), j'(z), j''(z))$. This amounts to checking exactly which equations of one complex variable involving only $z, j(z), j'(z), j''(z)$ have solutions, and whether these solutions are Zariski dense. Note that, for $n = 1$, broadness of V just means $\dim V \geq 3$; hence the only non-trivial case is $\dim V = 3$. On the other hand, freeness for

$n = 1$ means V has no constant coordinates. Nevertheless, we will even be able to decide what happens when V does have a constant coordinate.

Our first main result establishes the existence of solutions in all non-trivial cases.

Theorem 1.1. *Let $F(X, Y_0, Y_1, Y_2) \in \mathbb{C}[X, Y_0, Y_1, Y_2] \setminus \mathbb{C}[X]$. Then the equation*

$$F(z, j(z), j'(z), j''(z)) = 0$$

has infinitely many solutions.

The proof of Theorem 1.1 is based on a generalisation of the methods of [9], which use Rouché's theorem from classical complex analysis to establish some cases of EC for j (without derivatives). Theorem 1.1 can be seen as an analogue of the classical fact that every irreducible polynomial $p(X, Y) \in \mathbb{C}[X, Y]$ which depends on Y has infinitely many zeroes of the form $(z, \exp(z))$, unless $p = cY$ for some $c \in \mathbb{C}^{\times}$.

Throughout the paper, all algebraic subvarieties of $\mathbb{C}^4$ will be defined by polynomials in the ring $\mathbb{C}[X, Y_0, Y_1, Y_2]$.

Our main goal is to obtain a much stronger version of Theorem 1.1. We show that, for any polynomial $F(X, Y_0, Y_1, Y_2)$, the set

$$\{(z, j(z), j'(z), j''(z)) \in \mathbb{H} \times \mathbb{C}^3 : F(z, j(z), j'(z), j''(z)) = 0\}$$

is Zariski dense in the hypersurface $F(X, Y_0, Y_1, Y_2) = 0$, unless F is divisible by an explicit (finite) list of polynomials. In this case, we say that the equation $F(z, j(z), j'(z), j''(z)) = 0$ has a *Zariski dense set of solutions* (see Definition 2.1), that is, by a solution of such an equation, we understand a tuple $(z_0, j(z_0), j'(z_0), j''(z_0))$ rather than just z_0.

We remind the reader that this is equivalent to establishing certain cases of ECD for subvarieties of $\mathbb{C}^8$: given a hypersurface $V \subseteq \mathbb{C}^4$ and a Zariski open dense $V' \subseteq V$, there is $W \subseteq \mathbb{C}^8$ free and broad which projects onto V' such that W intersects the graph of j and its derivatives if and only if V' does. For instance, the system $\{j''(z) = 0,\ j(z) \neq 0\}$ has a solution if and only if $\{j''(z_1) = 0,\ j(z_1)z_2 = 1\}$ does.

The bulk of the paper is focused on proving the Zariski density of the set of solutions, which the proof of Theorem 1.1 does not provide. For instance, the solutions of the equation

$$zj''(z) + (z^3 + 1)j'(z)^2 + j'(z)j(z)^7 = 0$$

found via the proof Theorem 1.1 are the $\mathrm{SL}_2(\mathbb{Z})$-conjugates of $\rho = -\frac{1}{2} + \frac{\sqrt{3}}{2}i$. These are obviously not Zariski dense, for it is well known that $j(\gamma\rho) = j'(\gamma\rho) = j''(\gamma\rho) = 0$ for every $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ (see [13, p. 40]). Indeed, Zariski density requires at least that the solutions are not contained in finitely many $\mathrm{SL}_2(\mathbb{Z})$-orbits, except for when the equation is of the form $\prod_k (j(z) - u_k) = 0$ for some $u_k \in \mathbb{C}$.

The zeroes of j' are in fact problematic: observe that, for all $z \in \mathbb{H}$,

$$j(z)(j(z) - 1728) = 0 \iff j'(z) = 0.$$

This immediately gives that the three equations $j(z) = 0$, $j(z) - 1728 = 0$, and $j'(z) = 0$ do not have Zariski dense sets of solutions. Our main result shows that these are essentially the only non-examples.

Theorem 1.2. *For any polynomial $F(X, Y_0, Y_1, Y_2) \in \mathbb{C}[X, Y_0, Y_1, Y_2] \setminus \mathbb{C}[X]$ which is coprime to $Y_0(Y_0 - 1728)Y_1$, the equation $F(z, j(z), j'(z), j''(z)) = 0$ has a Zariski dense set of solutions, i.e. the set*

$$\{(z, j(z), j'(z), j''(z)) \in \mathbb{H} \times \mathbb{C}^3 : F(z, j(z), j'(z), j''(z)) = 0\}$$

is Zariski dense in the hypersurface $F(X, Y_0, Y_1, Y_2) = 0$.

Remark 1.3. For every rational function $G(X, Y_0, Y_1, Y_2) \in \mathbb{C}(X, Y_0, Y_1, Y_2)$, Theorem 1.2 implies that the function $G(z, j(z), j'(z), j''(z))$ has a zero unless G is of the form

$$\frac{Y_0^s (Y_0 - 1728)^t Y_1^\ell}{H(X, Y_0, Y_1, Y_2)},$$

where H is a polynomial and $s, t, \ell \in \mathbb{N}$.

A special case of Theorem 1.2 is that the equation $j''(z) = 0$ has a Zariski dense set of solutions.[1)] In this case, even proving that there is a solution outside the $\mathrm{SL}_2(\mathbb{Z})$-orbit of ρ is highly non-trivial; see Section 7.1.

Remark 1.4. In [8], the author studies the problem of finding *generic* solutions to equations involving j (and its derivatives) under the assumption that the system has a Zariski dense set of solutions. In particular, combining Theorem 1.2 with [8, Theorem 6.5], we get that there is a countable field $C_j \subseteq \mathbb{C}$ such that, for any irreducible hypersurface $V \subset \mathbb{C}^4$ satisfying the conditions of Theorem 1.2, if V is not definable over C_j, then for any finitely generated subfield $K \subset \mathbb{C}$ over which V can be defined, there is a point of the form $(z, j(z), j'(z), j''(z)) \in V$ such that $\operatorname{tr.deg.}_K K(z, j(z), j'(z).j''(z)) = \dim V = 3$.

To prove Theorem 1.2, we establish general criteria for the solvability of certain equations involving periodic functions (see Section 4). The following proposition is a special case of those criteria.

Definition 1.5. A meromorphic function $f : \mathbb{H} \to \mathbb{C}$ is 1-*periodic* if $f(z+1) = f(z)$ for every $z \in \mathbb{H}$. Every such function induces a meromorphic function $\widetilde{f}(q)$ on the punctured unit disc by performing the change of variable $q = \exp(2\pi i z)$. We say that f is *meromorphic at $i\infty$* if $\widetilde{f}$ is meromorphic at 0.

Proposition 1.6. *Let $f_0, \dots, f_n : \mathbb{H} \to \mathbb{C}$ be 1-periodic functions which are meromorphic on $\mathbb{H} \cup \{i\infty\}$. Suppose that, for some k, one of the following conditions is satisfied:*

- *there is $\tau \in \mathbb{H}$ such that $f_k(z)/f_n(z) \to \infty$ as $z \to \tau \in \mathbb{H}$, or*
- *$f_k(z)/f_n(z) \to \infty$ as $\mathrm{Im}(z) \to +\infty$.*

Then there is a sequence of points $\{z_m\}_{m \in \mathbb{N}} \subseteq \mathbb{H}$ with $z_m \neq \tau$ and $z_m \to \tau$ in the first case, or $\mathrm{Im}(z_m) \to +\infty$ and $0 \le \mathrm{Re}(z_m) \le 2$ in the second case, such that, for all sufficiently large m, the point $z_m + m$ is a solution to the equation

$$f_n(z)z^n + f_{n-1}(z)z^{n-1} + \cdots + f_0(z) = 0.$$

---

1) In particular, the ramification points of j' are not contained in finitely many $\mathrm{SL}_2(\mathbb{Z})$-orbits.

Let us consider an example illustrating how we apply Proposition 1.6 in practice. It also gives an idea of our approach in the general case.

Example 1.7. Consider the equation

$$j'(z)^2 + p(j(z)) = 0, \tag{1.1}$$

where either $p(j(z)) = j(z)(j(z) - 1728)$ or $p(j(z)) = j(z)^2(j(z) - 1728)$. First, we want to get an equivalent equation which is written as a sum of powers of z with periodic coefficients. To that end, we apply the $\mathrm{SL}_2(\mathbb{Z})$-transformation $z \mapsto -\frac{1}{z}$ and, using the identities $j(-\frac{1}{z}) = j(z)$, $j'(-\frac{1}{z}) = z^2 j'(z)$, we get

$$z^4 j'(z)^2 + p(j(z)) = 0. \tag{1.2}$$

Thus we obtain an equation in a suitable form for using Proposition 1.6, where $f_4 = (j')^2$, $f_3 = f_2 = f_1 = 0$ and $f_0 = p(j)$. When $p(j) = j(j - 1728)$, the ratio

$$\frac{f_0}{f_4} = \frac{j(j-1728)}{(j')^2}$$

has a pole at $\tau = \rho$. When $p(j) = j^2(j - 1728)$, the ratio

$$\frac{f_0}{f_4} = \frac{j^2(j-1728)}{(j')^2}$$

has no finite poles, but it has limit ∞ as $\mathrm{Im}(z) \to +\infty$.

Thus, by Proposition 1.6, there is a sequence z_m with $z_m \to \tau$ and $z_m \neq \tau$ in the first case, or $\mathrm{Im}(z_m) \to +\infty$ and $0 \le \mathrm{Re}(z_m) \le 2$ in the second case, such that, for all sufficiently large m, the point $z_m + m$ is a solution to equation (1.2). This already implies that the solutions of (1.2) intersect infinitely many $\mathrm{SL}_2(\mathbb{Z})$-orbits.

To deduce Zariski density of these solutions, suppose that all of them are also solutions of another independent equation $G(z, j(z), j'(z), j''(z)) = 0$. Combining this and (1.2), we can eliminate z and end up with an equation $H(j, j', j'') = 0$. Now, our assumption means that $H(j, j', j'')$ vanishes at $z_m + m$, hence also at z_m by periodicity. This is not possible, for a 1-periodic holomorphic function, meromorphic at $i\infty$, cannot have infinitely many zeroes with real part bounded from above and below and imaginary part bounded from below. This then implies the Zariski density of solutions of (1.1).

We also note that our criteria can be applied to more general periodic functions, beyond polynomials of j, j', j'', such as exp or the Weierstrass $\wp$-function. For instance, Proposition 1.6 implies that the function $j'(z)z + \exp(2\pi i z)$ has infinitely many zeroes around the points $i + m$, where m is a large integer.

1.1. Structure of the paper. In Section 2, we go over some basic preliminaries about the j-function and its derivatives. We also give the definition of Zariski density used in Theorem 1.2. In Section 3, we prove Theorem 1.1 by extending the methods of [9], which are based on Rouché's theorem. In Section 4, we prove criteria for the existence and distribution of solutions of equations involving periodic functions, which combined imply Proposition 1.6 (but are significantly more general). The approach used here involves Rouché's theorem, the Argument

Principle, and some elementary methods from valuation theory. These methods do not appear in later sections of the paper. In Section 5, we use the results of the previous section to obtain concrete criteria for proving Zariski density of equations of the form $F(z, j(z), j'(z), j''(z)) = 0$. These criteria are about the presence of poles in quotients of certain polynomials only in j, j', j''. In Section 6, we produce zero estimates for polynomials in z, j, j', j'' in order to determine when the quotients mentioned above have poles. In Section 7, we first prove Zariski density for **j**-*homogeneous* equations (Definition 7.1), which only involve $j(z)$, $j'(z)$ and $j''(z)$, including the equation $j''(z) = 0$ (Section 7.1). Finally, in Section 7.2, we prove Theorem 1.2 in full generality.

2. Preliminaries

Let $\mathbb{H}$ denote the complex upper half-plane $\{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$. The group $\mathrm{GL}_2^+(\mathbb{R})$ of 2×2 real invertible matrices with positive determinant acts on $\mathbb{H}$ via linear fractional transformations

$$gz := \frac{az + b}{cz + d} \quad \text{for } g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2^+(\mathbb{R}).$$

This action can be seen as a restriction of the action of $\mathrm{GL}_2(\mathbb{C})$ on the Riemann sphere $\mathbb{C} \cup \{\infty\}$. The *modular group* is defined as

$$\mathrm{SL}_2(\mathbb{Z}) := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2^+(\mathbb{R}) : a, b, c, d \in \mathbb{Z} \text{ and } ad - bc = 1 \right\}.$$

As a group, $\mathrm{SL}_2(\mathbb{Z})$ is generated by two elements $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ and $\left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$, which correspond to the actions $z \mapsto z + 1$ and $z \mapsto -\frac{1}{z}$, respectively.

The *modular* j*-function* can be defined as the unique $\mathrm{SL}_2(\mathbb{Z})$-automorphic function $j : \mathbb{H} \to \mathbb{C}$ satisfying $j(\rho) = 0$ (recall that $\rho := \exp(\frac{2\pi i}{3})$; this notation will be kept throughout the paper), $j(i) = 1728$, and $j(\infty) = \infty$ (this last condition should be understood as $\lim_{\mathrm{Im}(z) \to +\infty} j(z) = \infty$). In particular, this means that j satisfies $j(\gamma z) = j(z)$ for every γ in $\mathrm{SL}_2(\mathbb{Z})$ and every z in $\mathbb{H}$, and it is in particular 1-periodic, and by assumption meromorphic at $i\infty$. Its Fourier expansion (also known as a q-expansion) is of the form

$$j(z) = q^{-1} + 744 + \sum_{k=1}^{\infty} a_k q^k, \quad \text{with } q := \exp(2\pi i z) \text{ and } a_k \in \mathbb{C}. \tag{2.1}$$

In fact, $a_k \in \mathbb{Z}$ for every $k \in \mathbb{N}$. The j-function induces an analytic isomorphism of Riemann surfaces $\mathrm{SL}_2(\mathbb{Z}) \setminus \mathbb{H} \simeq \mathbb{C}$ (see [13, Chapter 3, §3]).

Since j is invariant under the action of $\mathrm{SL}_2(\mathbb{Z})$, we can study the behaviour of j by looking at fundamental domains of the action of $\mathrm{SL}_2(\mathbb{Z})$ on $\mathbb{H}$. The *standard fundamental domain* is the set

$$\mathbb{F} := \left\{ z \in \mathbb{C} : -\frac{1}{2} \leq \mathrm{Re}(z) < \frac{1}{2},\ |z| \geq 1,\ \left(|z| = 1 \Rightarrow -\frac{1}{2} \leq \mathrm{Re}(z) \leq 0 \right) \right\}.$$

We let $\overline{\mathbb{F}}$ denote the Euclidean closure of $\mathbb{F}$ (within the Riemann sphere). A diagram of the standard fundamental domain along with some of its $\mathrm{SL}_2(\mathbb{Z})$-translates is given in Figure 1. When we refer to a *fundamental domain*, we will always mean a set of the form $\gamma\mathbb{F}$ for some $\gamma \in \mathrm{SL}_2(\mathbb{Z})$.

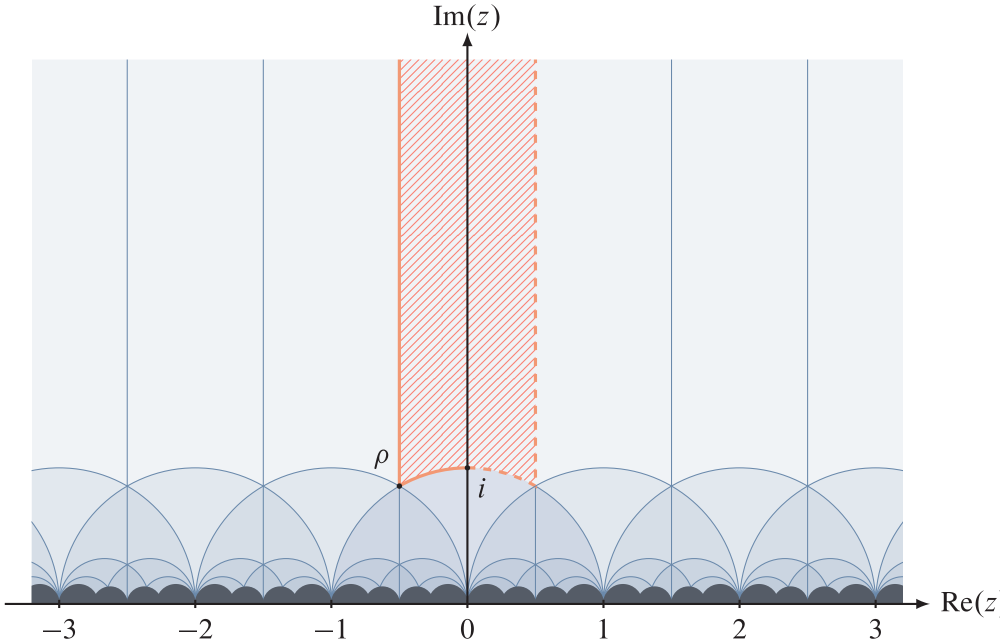

Figure 1. The fundamental domains of the action by $\mathrm{SL}_2(\mathbb{Z})$, where $\mathbb{F}$ is highlighted by the striped background.

It is well known that j satisfies the following third-order differential equation (and none of lower order [15]):[2)]

$$0 = \frac{j'''}{j'} - \frac{3}{2}\left(\frac{j''}{j'}\right)^2 + \frac{j^2 - 1968j + 2654208}{2j^2(j-1728)^2}(j')^2. \tag{2.2}$$

This shows that the derivatives of j of order at least 3 are rational over j, j', and j''. Mahler's result [15] implies that j, j', and j'' are algebraically independent over $\mathbb{C}$.

The functions j, j', and j'' are all 1-periodic and meromorphic at $i\infty$, and by differentiating (2.1), we can obtain the following q-expansions of j' and j'':

$$j'(z) = -\frac{2\pi i}{q} + 2\pi i \sum_{k\geq 1} a_k k q^k, \quad j''(z) = -\frac{4\pi^2}{q} - 4\pi^2 \sum_{k\geq 1} a_k k^2 q^k.$$

Observe that $\mathbb{Q} \cup \{\infty\}$ forms a single orbit under the action of $\mathrm{SL}_2(\mathbb{Z})$. We call these elements the *cusps* of j. Given a fundamental domain $\gamma\mathbb{F}$, the *cusp of* $\gamma\mathbb{F}$ is the unique element of $\mathbb{Q} \cup \{\infty\}$ contained in the Euclidean closure of $\gamma\mathbb{F}$ (where the closure is taken in the Riemann sphere).

Using (2.1) and the q-expansions of j' and j'', it is easy to see that, for every $x \in \mathbb{R}$, we have that each of the expressions $j(x+iy)$, $j'(x+iy)$, and $j''(x+iy)$ grows exponentially to ∞ as $y \to +\infty$. In this case, we sometimes write $z \to i\infty$ to emphasise that z approaches ∞ by increasing its imaginary part, while the real part remains bounded. A similar behaviour takes place when z approaches a rational number from within a fixed fundamental domain containing that rational number in its Euclidean closure. We will give a precise description of this behaviour in Section 4.

[2)] Observe that the denominators in the equation correspond to the polynomials that must be omitted in Theorem 1.2.

We finish this section with the definition of what we mean by finding a Zariski dense set of solutions to an equation.

Definition 2.1. Let $F(X, Y_0, Y_1, Y_2)$ be a polynomial over $\mathbb{C}$. We say the equation

$$F(z, j(z), j'(z), j''(z)) = 0$$

has a *Zariski dense set of solutions* if, for any polynomial $G(X, Y_0, Y_1, Y_2)$ which is not divisible by some irreducible factor of F, there is $z_0 \in \mathbb{H}$ such that $F(z_0, j(z_0), j'(z_0), j''(z_0)) = 0$ and $G(z_0, j(z_0), j'(z_0), j''(z_0)) \neq 0$.

Clearly, it suffices to prove Zariski density for irreducible polynomials to obtain Theorem 1.2, so from now on, we will reduce to the case where F is irreducible.

3. Existence of solutions

We start with the *Rouché method* for proving the existence of solutions, but not yet their Zariski density. We first recall the crucial theorem.

Theorem 3.1 (Rouché; see e.g. [14, Chapter VI, §1, Theorem 1.6]). *Let f, g be meromorphic functions on a complex domain Ω. Let C denote a simple closed curve which is homologous to 0 in Ω and such that f has no zeroes or poles on C. If the inequality*

$$|g(z)| < |f(z)|$$

holds for all z on C, then the difference between the numbers of zeroes and poles in the interior of C for the functions $f + g$ and f is the same.

It is well known that the Euclidean closure of the $\mathrm{SL}_2(\mathbb{Z})$-orbit of any point in $\mathbb{H}$ (where the closure is taken within the Riemann sphere) only accumulates at the boundary of $\mathbb{H}$, that is, at $\mathbb{R} \cup \{\infty\}$. The following lemma will help us choose convenient sequences within any given orbit converging to points in $\mathbb{R}$.

Lemma 3.2. *Let $z \in \mathbb{H}$ and $u \in \mathbb{R}$.*

(i) *If $u \in \mathbb{R} \setminus \mathbb{Q}$ and the sequence $\gamma_k = \left(\begin{smallmatrix} a_k & b_k \\ c_k & d_k \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ is such that $\gamma_k z \to u$ as $k \to +\infty$, then $|c_k| \to +\infty$ and $\frac{a_k}{c_k} \to u$ as $k \to +\infty$.*

(ii) *If $u = \frac{a}{c} \in \mathbb{Q}$ with $\gcd(a, c) = 1$, then there is a sequence $\gamma_k = \left(\begin{smallmatrix} a & b_k \\ c & d_k \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ such that $|b_k|, |d_k| \to +\infty$ and $\gamma_k z \to u$ as $k \to +\infty$.*

Proof. (i) We first show that $|c_k| \to +\infty$. Since every subsequence of $\gamma_k z$ tends to u, it suffices to show that c_k is unbounded. Assume it is bounded; then we can choose a subsequence where the value of c_k is constant. So assume $c_k = c$ is a constant sequence.

Conversely, assume now that $\gamma_k z \to u$. Let $z = x + iy$. If d_k is also bounded, then we may assume it is constant. Then $a_k z + b_k = (a_k x + b_k) + a_k y i$ must be convergent; hence a_k must be convergent, and so constant. Then b_k is also constant, for $a_k d_k - b_k c_k = 1$, a contradiction.

Thus we may assume $|d_k| \to +\infty$. Then

$$\frac{\frac{a_k}{d_k}z + \frac{b_k}{d_k}}{\frac{c}{d_k}z + 1} \to u.$$

Therefore,

$$\frac{a_k}{d_k}z + \frac{b_k}{d_k} = \left(\frac{a_k}{d_k}x + \frac{b_k}{d_k}\right) + \frac{a_k}{d_k}yi \to u.$$

This implies

$$\frac{a_k}{d_k} \to 0, \quad \frac{b_k}{d_k} \to u.$$

On the other hand, $a_k d_k - b_k c = 1$; hence $a_k - \frac{b_k}{d_k}c = \frac{1}{d_k} \to 0$. Thus $a_k \to uc$, which means $a_k = a \in \mathbb{Z}$ is constant. But then $u = \frac{a}{c} \in \mathbb{Q}$.

Since $|c_k| \to +\infty$, then using that

$$\frac{a_k z + b_k}{c_k z + d_k} \cdot \frac{c_k}{a_k} = \frac{a_k z + b_k}{a_k z + b_k + \frac{1}{c_k}} \to 1,$$

we see that $\gamma_k z$ and $\frac{a_k}{c_k}$ must have the same limit.

(ii) As $u = \frac{a}{c}$ with $\gcd(a, c) = 1$, there are integers m, l such that $am - cl = 1$. Choose $b_k = l + ka$, $d_k = m + kc$. Then

$$\lim_{k\to+\infty} \frac{az + b_k}{cz + d_k} = \lim_{k\to+\infty} \frac{\frac{a}{d_k}z + \frac{b_k}{d_k}}{\frac{c}{d_k}z + 1} = \lim_{k\to+\infty} \frac{b_k}{d_k} = \lim_{k\to+\infty} \frac{l + ak}{m + ck} = \frac{a}{c} = u. \qquad \square$$

In order to ease notation, we will start using bold-faced letters to denote vectors, so we set $\mathbf{Y} := (Y_0, Y_1, Y_2)$ and $\mathbf{j} := (j, j', j'')\colon \mathbb{H} \to \mathbb{C}^3$.

We are now ready to prove Theorem 1.1 which we restate below for convenience.

Theorem 1.1. *For every $F(X, \mathbf{Y}) \in \mathbb{C}[X, \mathbf{Y}] \setminus \mathbb{C}[X]$, the equation $F(z, \mathbf{j}(z)) = 0$ has infinitely many solutions.*

Proof. If F does not depend on Y_1 and Y_2, then we are done by the results of [9]. So assume F depends on Y_1 or Y_2. The argument below is a generalisation of the method of [9].

Let $r(X) := -F(X, 0, 0, 0)$ and $G(X, \mathbf{Y}) := F(X, \mathbf{Y}) + r(X)$. Further, let

$$f(z) := G(z, \mathbf{j}(z)).$$

Then we want to solve the equation

$$f(z) = r(z).$$

Notice that $f(\rho) = 0$, for $j(\rho) = j'(\rho) = j''(\rho) = 0$. Let $B \subseteq \mathbb{C}$ be a closed disc centred at ρ with sufficiently small radius such that $j'(z) \neq 0$ and $j''(z) \neq 0$ for $z \in B \setminus \{\rho\}$.

Pick a point $u = \frac{a}{c} \in \mathbb{Q}$, and choose a sequence $\gamma_k \in \mathrm{SL}_2(\mathbb{Z})$ as in Lemma 3.2 (ii) such that $\gamma_k z \to u$ as $k \to +\infty$ for any $z \in \mathbb{H}$. Let $B_k := \gamma_k B$. By compactness of B, the function $r(\gamma_k z)$ tends to $r(u)$ uniformly for $z \in B$.

We have

$$\begin{aligned} f(\gamma_k z) &= G\big(\gamma_k z, j(z), (cz+d_k)^2 j'(z), (cz+d_k)^4 j''(z) + 2c(cz+d_k)^3 j'(z)\big) \\ &= G\Big(\gamma_k z, j(z), d_k^2\Big(\frac{c}{d_k}z+1\Big)^2 j'(z), \\ &\qquad d_k^4\Big(\Big(\frac{c}{d_k}z+1\Big)^4 j''(z) + 2\frac{c}{d_k}\Big(\frac{c}{d_k}z+1\Big)^3 j'(z)\Big)\Big). \end{aligned}$$

Consider the polynomial $G(X, Y_0, T^2 Y_1, T^4 Y_2)$ as a polynomial of T. It clearly has positive degree, for otherwise G (and hence F) would not depend on Y_1 nor Y_2. Let its leading term be $H(X, \mathbf{Y}) \cdot T^m$. Since $\frac{c}{d_k} \to 0$, we see that

$$f(\gamma_k z) = d_k^m \cdot H(u, \mathbf{j}(z)) + o(d_k^m) \quad \text{as } k \to +\infty.$$

We can now shrink B to make sure that $H(u, \mathbf{j}(z)) \neq 0$ on ∂B, so it is uniformly bounded away from 0 for $z \in \partial B$. In particular, $f(\gamma_k z)$ approaches infinity as $k \to +\infty$ uniformly for $z \in \partial B$. So, for sufficiently large k, the inequality $|f(z)| > |r(z)|$ holds for all

$$z \in \partial B_k = \partial(\gamma_k B) = \gamma_k \partial B,$$

and we can apply Rouché's theorem to these functions. Since f has a zero in B_k, namely $\gamma_k \rho$, so does $f - r$. □

Remark 3.3. The following more general statement can be proven by the same argument. *Let $F(X, \mathbf{Y}) \in \mathbb{C}[X, \mathbf{Y}] \setminus \mathbb{C}[X]$. Let $U \subseteq \mathbb{C}$ be an open set such that $U \cap \mathbb{R} \neq \emptyset$ and let $f: U \to \mathbb{C}$ be a holomorphic function. Then the equation $F(z, \mathbf{j}(z)) = f(z)$ has infinitely many solutions.*

As mentioned in Section 1, the proof of Theorem 1.1 does not guarantee a Zariski dense set of solutions. For instance, if $F(X, 0, 0, 0) \equiv 0$, that is, F has no term depending only on X, then the only solutions yielded by the method above are the $\mathrm{SL}_2(\mathbb{Z})$-conjugates of ρ. In order to establish Zariski density, we will look at a refinement of the procedure, where we transform the equation by convenient elements of $\mathrm{SL}_2(\mathbb{Z})$. This will be done starting in Section 6, but first, in the next section, we will develop some tools to study equations involving periodic functions.

4. Solvability of certain equations involving periodic functions

In this section, we establish some general criteria for the solvability of equations involving periodic functions and, in particular, prove Proposition 1.6. We remark that this section is independent in many ways from the rest of the paper as the results we prove make no reference to j or its derivatives, and in particular, the methods developed here will not reappear in the following sections.

We recall that, given a meromorphic function f, not identically 0, and a point z_0, the *order* of f at z_0 is the unique integer n such that $(z - z_0)^{-n} f(z)$ is holomorphic and non-zero at z_0.

Proposition 4.1. *Let $f_0, \dots, f_n: \mathbb{H} \to \mathbb{C}$ be 1-periodic meromorphic functions and let $-\ell$ be the minimum order of f_k/f_n at a fixed $z_0 \in \mathbb{H}$ for $k = 0, \dots, n-1$.*

If $\ell > 0$, then for any sufficiently small disc D centred at z_0 and for every sufficiently large $m \in \mathbb{Z}$, the equation

$$f_n(z)z^n + f_{n-1}(z)z^{n-1} + \cdots + f_0(z) = 0$$

has ℓ solutions, counted with multiplicity, in $m + (D \setminus \{z_0\})$.

Proof. For simplicity, assume that f_0/f_n has a pole at $z_0 \in \mathbb{H}$ of order $\ell > 0$; the same proof will work in the general case with trivial modifications.

Under the above assumptions, $(f_n/f_0)(z_0) = 0$, and moreover, $(f_k/f_0)(z_0) \neq \infty$ for all k. Let $F(z) := f_n(z)z^n + f_{n-1}(z)z^{n-1} + \cdots + f_0(z)$. Consider the functions

$$G(z) := \frac{f_n(z)}{f_0(z)}z^n \quad \text{and} \quad H(z) := \frac{f_{n-1}(z)}{f_0(z)}z^{n-1} + \cdots + \frac{f_1(z)}{f_0(z)}z + 1.$$

Pick a small closed disc D centred at z_0 such that the f_k's have no zeroes nor poles in $D \setminus \{z_0\}$. Since $f_k(z)/f_0(z)$ are periodic and bounded on D, for large enough m, we have

$$|G(z+m)| = \left|\frac{f_n(z)}{f_0(z)}\right| |z+m|^n > |H(z+m)| \quad \text{for } z \in \partial D.$$

By Rouché's Theorem 3.1, the number of zeroes of the functions

$$G(z+m) \quad \text{and} \quad G(z+m) + H(z+m) = f_0(z)^{-1}F(z+m)$$

inside D is the same. Since the former has a zero at z_0 of order ℓ and no other zero, the latter must also have ℓ zeroes in D, counted with multiplicity. Thus $f_0(z)^{-1}F(z+m)$ has ℓ zeroes in D, and so $f_0^{-1}F$ has ℓ zeroes in $m + D$, counted with multiplicity.

Finally, note that $G(z_0 + m) + H(z_0 + m) = 0$ holds for at most $n - 1$ values of m; thus, for m sufficiently large, the above ℓ solutions in $m + D$ are actually in $m + (D \setminus \{z_0\})$.

This finishes the proof of the proposition when f_0/f_n has a pole at z_0 of order greater than or equal to that of f_k/f_n for $k = 1, \ldots, n-1$. For when the maximum order of pole at z_0 is attained by f_k/f_n for some $k \neq 0$, simply divide by $f_k(z)$ rather than $f_0(z)$ when defining G and H. □

Following the notation of the proposition, when the f_k's are polynomials in j, j', j'' and some f_k/f_n has a pole in $\mathbb{H}$, the above proposition applies. Instead, when the functions f_k/f_n have no poles in $\mathbb{H}$, we will rely on the asymptotic behaviour of f_k/f_n towards the boundary of $\mathbb{H}$. Specifically, we will prove an analogue of Proposition 4.1 in the case where

$$\frac{f_k(z)}{f_n(z)} \to \infty \quad \text{as } z \to i\infty.$$

As the following example shows, we may also need to consider non-periodic functions which are *asymptotically periodic*. Dealing with those functions requires a considerably more sophisticated setup than the one in Proposition 1.6, so we first discuss the example in detail to clarify the choices made in the rest of this section.

Example 4.2. Consider the equation

$$(j')^5 + j^2(j-1728)^2(j'')^2 + \alpha j^2(j-1728)(j')^3 = 0,$$

where α is to be determined later. In order to write this equation as a polynomial in z with periodic coefficients, we apply the $z \mapsto -\frac{1}{z}$ transformation[3)] and get

$$z^{10}(j'(z))^5 + z^8 j(z)^2(j(z)-1728)^2(j''(z))^2 + z^7 4j(z)^2(j(z)-1728)^2 j'(z)j''(z) \\ + z^6(4j(z)^2(j(z)-1728)^2(j'(z))^2 + \alpha j(z)^2(j(z)-1728)(j'(z))^3) = 0.$$

In this example, the ratio of the coefficients of z^8 and z^{10} actually has a pole at i. However, checking for poles among such ratios in a general equation requires sufficiently precise zero estimates at the conjugates of ρ and i, which are hard to produce for polynomials involving j'' (see Example 6.7); on the other hand, we can provide sharp zero estimates for polynomials in j, j' only (see Section 6). The latter estimates turn out to be enough: for instance, when the original equation does not contain z, after the $z \mapsto -\frac{1}{z}$ transformation, the coefficient of the lowest power of z does not depend on j''. This is exemplified here by the coefficient of z^6. Hence our strategy hinges on the fact that the ratio between a particular coefficient not involving j'', which we can procure in all cases, and the leading coefficient has a pole or exponential growth in some fundamental domain.

In this particular example, the ratio in question is between the coefficients of z^6 and z^{10}, thus the function

$$f(z) := \frac{4j^2(j-1728)^2 + \alpha j^2(j-1728)j'}{(j')^3}.$$

We claim that $f(z)$ has no pole in $\mathbb{H}$. Indeed, easy calculations show that the orders of the numerator at ρ and i (and their $\mathrm{SL}_2(\mathbb{Z})$-orbits) are equal to 6 and 3 respectively. The denominator has the same orders at these points, so f has no poles. Moreover, choosing $\alpha = \frac{2}{\pi i}$ ensures the leading terms in the q-expansions of the two terms in the numerator cancel out. This then means that $f(z)$ tends to a constant as $z \to i\infty$. Therefore, Proposition 1.6 cannot be applied in this situation. However, $f(z)$ has exponential growth as we approach 0 from within a fundamental domain with a cusp at 0 (in fact, we shall prove that $f(z)$ must have exponential growth in *most* fundamental domains). Indeed, after applying the $z \mapsto -\frac{1}{z}$ transformation, we get

$$g(z) := f\left(-\frac{1}{z}\right) = \frac{4j(z)^2(j(z)-1728)^2 + \alpha z^2 j(z)^2(j(z)-1728)j'(z)}{z^6 j'(z)^3},$$

and because of the extra factor z^2 in the second summand in the numerator, no cancellation is possible, thus guaranteeing that $g(z)$ grows exponentially as $z \to i\infty$. Note however that this function is not periodic, but only *asymptotically periodic* in the sense that

$$\lim_{z \to i\infty} \frac{g(z+1)}{g(z)} = 1.$$

This fact is responsible for the technicalities in the rest of this section.

It is worth mentioning that, for equations of the form $F(j, j') = 0$, where F is a polynomial, the transformation $z \mapsto -\frac{1}{z}$ always guarantees that the ratio of the lowest and highest powers of z has a pole in $\mathbb{H}$ or exponential growth at $i\infty$. In general, Proposition 1.6 is sufficient when we deal with equations of the form $F(z, j(z), j'(z)) = 0$, although the argument is somewhat more complicated.

[3)] The action of the group $\mathrm{SL}_2(\mathbb{Z})$ is generated by the transformations $z \mapsto z+1$ and $z \mapsto -1/z$. Since our functions are invariant under the former, it is natural to apply the latter; while j is invariant under it, j' and j'' are not, and we take advantage of this fact.

The reader may benefit from revisiting this example after reading the rest of the paper, as it will make the above-mentioned phenomena less obscure.

Notation. Let $\mathcal{P}$ denote the field of 1-periodic meromorphic functions on $\mathbb{H}$ which are also meromorphic at $i\infty$ (recall Definition 1.5). We write $\mathcal{P}[w]$ and $\mathcal{P}(w)$ respectively for the polynomial ring and its fraction field generated by the variable w over $\mathcal{P}$. We remark that the functions in $\mathcal{P}$ will also be thought of as meromorphic functions of the variable w.

Also, given an unbounded region $U \subseteq \mathbb{C}$ and two meromorphic functions f, g on U, we write $f \sim g$ *for* $w \to \infty$ *in* U to mean that the limit of the ratio $f(w)/g(w)$ tends to 1 as w approaches infinity from within U.

Lemma 4.3. *Let $f \in \mathcal{P}(w)$. Then there are $\alpha \in \mathbb{C}^\times$, $e, d \in \mathbb{Z}$, and a positive $C \in \mathbb{R}$ such that $f(w) \sim \alpha w^d q^e$ for $w \to \infty$ in the region* $\mathrm{Im}(w) \geq C \log|w|$, *where* $q = \exp(2\pi i w)$.

Proof. It suffices to prove the conclusion for $f \in \mathcal{P}[w]$. Write

$$f(w) = \sum_{k=0}^{n} g_k(w) w^k, \quad \text{with } g_k \in \mathcal{P}.$$

Each $g_k(w)$ has a meromorphic q-expansion $\tilde{g}_k(q)$ which converges on some neighbourhood of $q = 0$. Let e be the minimum order of $\tilde{g}_k(q)$ at $q = 0$ for $k = 0, \ldots, n$. Let $\alpha_k \in \mathbb{C}$ be such that $\tilde{g}_k(q) = q^e(\alpha_k + O(q))$. Let d be the maximum k such that $\tilde{g}_k(q)$ has order e at $q = 0$, namely such that $\alpha_k \neq 0$. In the region $|q| \leq |w|^{-(n+1)}$, we have $O(q) = O(w^{-(n+1)})$, in which case

$$G(w, q) = \sum_{k=0}^{n} \tilde{g}_k(q) w^k = q^e w^d \sum_{k=0}^{n} \big(\alpha_k + O(w^{-(n+1)})\big) w^{k-d} = \alpha_d q^e w^d \big(1 + O(w^{-1})\big).$$

It now suffices to specialise at $q = \exp(2\pi i w)$ and observe that $|q| = e^{-2\pi \,\mathrm{Im}\, w} \leq |w|^{-(n+1)}$ if and only if $\mathrm{Im}(w) \geq \frac{n+1}{2\pi} \log|w|$. □

It follows at once that all functions in $\mathcal{P}(w)$ are "asymptotically periodic" in the sense that $f(w+1) \sim f(w)$ for $w \to \infty$ in the above region. Moreover, the lemma allows us to make the following definition.

Definition 4.4. Call the *order at* $i\infty$ of $f \in \mathcal{P}(w)$, written $\mathrm{ord}_{w=i\infty}(f)$, the pair $(e, d) \in \mathbb{Z}^2$ of integers such that, for some $\alpha \in \mathbb{C}$ and some $C \in \mathbb{R}$, we have

$$f(w) \sim \alpha w^{-d} \exp(2\pi i e w) \quad \text{for } w \to \infty$$

in the region $\mathrm{Im}(w) \geq C \log|w|$.

We say that f has *exponential growth at* $i\infty$ if its order is (e, d) with $e < 0$.

Here we consider $\mathbb{Z}^2$ as a lexicographically ordered group. This makes $(\mathcal{P}(w), \mathrm{ord}_{w=i\infty})$ into a valued field, and we have for instance $f(w) \to \infty$ in a suitable region $\mathrm{Im}(w) \geq C \log|w|$ if and only if $\mathrm{ord}_{w=i\infty} f(w) < (0, 0)$.

We can now set up a generalisation of Proposition 4.1 that will cover our application to j. Let us fix the following data:

- a polynomial $F(z,w) = \sum_{k=0}^{n} z^k f_k(w)$, where each f_k is in $\mathcal{P}(w)$ and $f_n \neq 0$;
- a value s which is either 0 or 1.

We look for the zeroes of functions of the form $F_r(w) := F(r + sw, w)$ for r varying among the real numbers. For each r sufficiently large, we pick a suitable rectangle Ξ_r, pictured in Figure 2 and defined in Proposition 4.8, and integrate the logarithmic derivative F_r'/F_r along the boundary of Ξ_r. Provided that F_r does neither have zeroes nor poles on such a boundary, by the Argument Principle (see Theorem 4.7), the value of the integral counts the difference between the number of zeroes and poles, with multiplicity, inside Ξ_r. We will choose Ξ_r so that the integral has positive value.

We first parametrise the roots of $F(z,w)$ as a polynomial in z in terms of w varying in a suitable region. The resulting functions, which by construction are algebraic over $\mathcal{P}(w)$, admit an order at $i\infty$ which may be a pair of rational numbers, rather than only integers.

Lemma 4.5. *There is a positive $C \in \mathbb{R}$ such that, in the region*

$$U = \{w \in \mathbb{H} : \mathrm{Im}(w) \geq C \log|w|\},$$

there are holomorphic functions $\beta_1, \ldots, \beta_n : U \to \mathbb{C}$ such that $F(\beta_k(w), w) = 0$ for all $w \in U$, and if $F(\beta, w) = 0$, then $\beta = \beta_k(w)$ for some k.

Moreover, there are $\alpha_k \in \mathbb{C}^\times$, $e_k, d_k \in \mathbb{Q}$ such that $\beta_k(w) \sim \alpha_k w^{-d_k} q^{e_k}$ for $w \to \infty$ in U, where $q = \exp(2\pi i w)$ and $w^{-d_k} = \exp(-2\pi i d_k \log(w))$ for some holomorphic branch of $\log(w)$ on U.

Proof. It suffices to prove the conclusion for F irreducible as a polynomial over $\mathcal{P}(w)$.

Let $F^{(z)} = \partial F/\partial z$. Then there are $G, H \in \mathcal{P}(w)[z]$ such that $GF + HF^{(z)} = 1$. We take C large enough that, by Lemma 4.3, the coefficients of F, $F^{(z)}$, G, and H are holomorphic in the region $U = \{w \in \mathbb{H} : \mathrm{Im}(w) \geq C \log|w|\}$. In particular, $F(z, w_0)$ and $F^{(z)}(z, w_0)$ have no common roots for any $w_0 \in U$, and so, by the implicit function theorem, and because U is simply connected, there are m holomorphic functions $\beta_1, \ldots, \beta_m : U \to \mathbb{C}$ such that $F(\beta_t(w), w) = 0$ for every t and $w \in U$, and moreover taking distinct values at all $w \in U$; thus if $F(\beta, w) = 0$, then $\beta = \beta_k(w)$ for some k.

Now fix some k. For every $w \in U$, there is some t such that

$$|\beta_k(w)^t f_t(w)| \geq |\beta_k(w)^\ell f_\ell(w)| \quad \text{for every } \ell,$$

and since $F(\beta_k(w), w) = 0$, there is also $h \neq t$ such that $|\beta_k(w)^h f_h(w)| \geq \frac{1}{n}|\beta_k(w)^t f_t(w)|$. Let $U_{t,h}$ be the region where those inequalities hold. We have in particular

$$1 \geq |\beta_k(w)| \sqrt[h-t]{\left|\frac{f_h(w)}{f_t(w)}\right|} \geq \sqrt[h-t]{\frac{1}{n}}$$

Let

$$(e_k, d_k) = \frac{\mathrm{ord}_{w=i\infty}(f_t/f_h)}{h-t}.$$

By Lemma 4.3 combined with the above inequalities, there is $N > 1$ such that, whenever w is sufficiently large in $U_{t,h}$, we have

$$N \geq \frac{|\beta_k(w)|}{|w^{-d_k} q^{e_k}|} \geq \frac{1}{N}$$

for some fixed determination of w^{-d_k} on U. Choose N so that it works simultaneously for any possible pair t, h. By continuity of β_k, the numbers d_k, e_k do not depend on t, h.

Let S be the set of indices t such that $\mathrm{ord}_{w=i\infty}(f_t) + (te_k, td_k)$ reaches a minimum value (e, d). By construction, S contains at least two elements. Write

$$F(w^{-d_k} q^{e_k} z', w) = q^e w^d (G_0(z') + G_1(z', w)),$$

where now G_0 is a non-trivial polynomial in z' with $|S| \geq 2$ terms and constant coefficients, and G_1 has coefficients that tend to 0 for $w \to \infty$ in $U_{t,h}$. Since

$$\frac{\beta_k(w)}{w^{-d_k} q^{e_k}}$$

is bounded and continuous, it must converge to a non-zero root α_k of $G_0(z')$; thus we have $\beta_k(w) \sim \alpha_k w^{-d_k} q^{e_k}$, as desired. □

We now provide an estimate on the size of $F_r(w)$ that we can use on the boundary of Ξ_r.

Lemma 4.6. *There exist* $0 \leq x_0 < 1$, $C, y_0 \geq 4$, $D, E_0, E_1 \geq 0$ *(with* E_0 *possibly* $+\infty$*),* $\delta > 0$ *such that*

$$|F_r(w)| = |F(r + sw, w)| > \delta \sum_{k=0}^{n-1} |f_k(w)||r + sw|^k$$

for all $w \in \mathbb{H}$, $r \in \mathbb{R}$ *such that* $\mathrm{Im}(w) \geq y_0$, $|r| \geq \mathrm{Im}(w)^D$, *and one of the following holds:*

- $0 \leq \mathrm{Re}(w) \leq 2$, $\mathrm{Im}(w) \leq E_0 \log|r|$, *or*
- $0 \leq \mathrm{Re}(w) \leq 2$, $E_1 \log|r| \leq \mathrm{Im}(w)$, *or*
- $\mathrm{Re}(w) \in \{x_0, x_0 + 1\}$.

Proof. We work in a region $U = \{w \in \mathbb{H} : 0 \leq \mathrm{Re}(w) \leq 2 \wedge \mathrm{Im}(w) \geq y_0\}$ for some y_0 sufficiently large as determined by this proof. We start by taking y_0 large enough that, by Lemma 4.3, f_n has neither zeroes nor poles at w. We also require that $|r| \geq C$ for some C sufficiently large, again as determined by this proof.

By Lemma 4.5, provided y_0 is sufficiently large, there are holomorphic functions

$$\beta_1, \ldots, \beta_n : U \to \mathbb{C}$$

parametrising the roots of $F(z, w) = 0$ as functions of w, and for $w \to \infty$ in U, we have

$$\beta_k \sim \alpha_k w^{-d_k} q^{e_k} \quad \text{for some } \alpha_k \in \mathbb{C}^\times \text{ and } d_k, e_k \in \mathbb{Q}.$$

Since we are assuming $0 \leq \mathrm{Re}(w) \leq 2$, we also have $\mathrm{Im}(w) \leq |w| \leq \mathrm{Im}(w) + 2$. We require that $C \geq 4$, $y_0 \geq 4$, so that we have the following simple inequalities:

$$|r + sw| \geq \max\{|r| - 2, s\,\mathrm{Im}(w)\} \geq \max\left\{\frac{|r|}{2}, s\frac{|w|}{2}\right\} \geq \frac{|r + sw|}{4}.$$

It follows, for instance, that $|w| \sim \mathrm{Im}(w) \to +\infty$ for $w \to \infty$ in U. We shall omit the specification "in U" in the rest of this proof.

We now bound $|r + sw - \beta_k(w)|$, distinguishing multiple cases depending on (e_k, d_k).

If $(e_k, d_k) \geq (0, 0)$, then $\beta_k(w)$ converges to a finite value γ_k for $w \to \infty$. In this case, we require C (if $s = 0$) and y_0 to be large enough that

$$|r + sw| \geq 4|\gamma_k| \geq 2|\beta_k(w)|.$$

This ensures that

$$|r + sw - \beta_k(w)| \geq \frac{|r + sw|}{2} \geq \frac{|r + sw| + |\beta_k(w)|}{4}.$$

Suppose that $e_k = 0$ and $d_k < 0$. For $|r| \geq \mathrm{Im}(w)^{2\max\{1,-d_k\}}$, we have both $|r| \geq 4|sw|$ and $|r| \geq 4|\beta_k(w)|$ for w sufficiently large, and so, for y_0 large, we get

$$|r + sw - \beta_k(w)| \geq \frac{|r|}{2} \geq \frac{|r + sw| + |\beta_k(w)|}{8}.$$

If $e_k < 0$, then $sw - \beta_k(w) \sim -\beta_k(w)$ for $w \to \infty$. Since $\mathrm{Re}(w)$ is bounded, we get

$$\mathrm{Re}(\log(sw - \beta_k(w))) = \log|sw - \beta_k(w)| \sim \log|\beta_k(w)| \sim -2\pi e_k \,\mathrm{Im}(w).$$

We first give bounds when $|r|$ is roughly at least $|\beta_k|^2$, and when $|\beta_k|$ is roughly at least $|r|^2$. More precisely, for y_0 sufficiently large, we have

$$|r + sw - \beta_k(w)| > \frac{|r|}{2} \geq \frac{|r + sw| + |\beta_k(w)|}{8} \qquad \text{for } |r| \geq e^{-4\pi e_k \,\mathrm{Im}(w)},$$
$$|r + sw - \beta_k(w)| > \frac{|\beta_k(w)|}{2} \geq \frac{|r + sw| + |\beta_k(w)|}{8} \quad \text{for } |r| \leq e^{-\pi e_k \,\mathrm{Im}(w)}.$$

For $w \to \infty$, we also have

$$\mathrm{Im}(\log(sw - \beta_k(w))) \sim \arg(\alpha_k) + \pi + 2\pi e_k \mathrm{Re}(w) \bmod 2\pi.$$

Now choose x_0 so that $\arg(\alpha_k) + 2\pi e_k x_0$ and $\arg(\alpha_k) + 2\pi e_k(x_0 + 1)$ are not in $\pi\mathbb{Z}$, and so are different from $\arg(r) \in \pi\mathbb{Z}$. Note that we can do this simultaneously for all k such that $e_k < 0$. In particular, there is $\delta_k > 0$ such that, after taking y_0 sufficiently large, we have

$$|r + sw - \beta_k(w)| > \delta_k(|r + sw| + |\beta_k(w)|) \quad \text{for } \mathrm{Re}(w) \in \{x_0, x_0 + 1\}.$$

Now, if $e_k = 0$ for some k, let D be the maximum between the values $-2d_k$ and 2 for such k; otherwise, we can take $D = 0$. If $e_k < 0$ for some k, let E_0 be the maximum of $-4\pi e_k$ and let E_1 be the minimum of $-\pi e_k$ for such k; otherwise, let $E_0 = +\infty$ and $E_1 = 0$. Under the above choices, there is $\delta > 0$ such that

$$|F_r(w)| > \delta |f_n(w)| \prod_k (|r + sw| + |\beta_k(w)|) \geq \delta \sum_{k=0}^{n} |f_k(w)||r + sw|^k$$

for any $w \in \mathbb{H}$, $r \in \mathbb{R}$ satisfying the requirements in the conclusion. □

We now recall the Argument Principle, which plays a key role in the proof of Proposition 4.8.

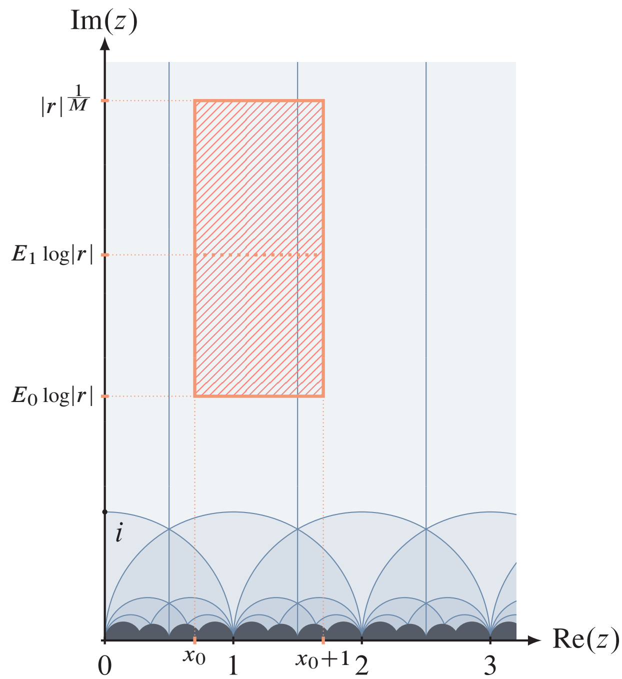

Figure 2. The region Ξ_r highlighted by the striped background.

Theorem 4.7 (Argument Principle; see e.g. [14, Chapter VI, §1, Theorem 1.5]). *Let f be a meromorphic function on a complex domain Ω. Let C be a simple closed curve (positively oriented) which is homologous to 0 in Ω and such that f has no zeroes or poles on C. Let Z and P respectively denote the number of zeroes and poles (counted with multiplicity) of f in the interior of C. Then*

$$2\pi i(Z - P) = \oint_C \frac{f'(z)}{f(z)}\,dz = \oint_{f\circ C} \frac{dz}{z}.$$

In the proof of the following proposition, we will integrate F_r'/F_r along the boundary of Ξ_r and use the above estimates to find a positive lower bound, thus proving the existence of zeroes of F_r within Ξ_r. For a more geometric description, integrating F_r'/F_r computes how many times the image $F_r(z)$ winds around 0 while z moves along $\partial\Xi_r$. The bounds below will determine a rough picture of $F_r(\partial\Xi_r)$, as in Figure 3, and in turn determine the number of zeroes, counted with multiplicity.

Proposition 4.8. *Let (e, d) be the minimum order of f_k/f_n at $i\infty$ for $k = 0, \ldots, n-1$ and suppose that $e < 0$. We work under the notation of Lemma 4.6. Then, for all $r \in \mathbb{R}$ sufficiently large, the function $F(r + sw, w)$ has $-e$ zeroes, counted with multiplicity, within the region (Figure 2)*

$$\Xi_r = \{w \in \mathbb{H} : x_0 < \mathrm{Re}(w) < x_0 + 1,\ E_0 \log|r| < \mathrm{Im}(w) < |r|^{\frac{1}{M}}\},$$

where $M = \lceil D \rceil$ if $D > 0$ and $M = 1$ otherwise. Moreover, for $w \in \partial\Xi_r$, we have

$$|F(r + sw, w)| \geq \delta \sum_{k=0}^{n} |f_k(w)||r + sw|^k.$$

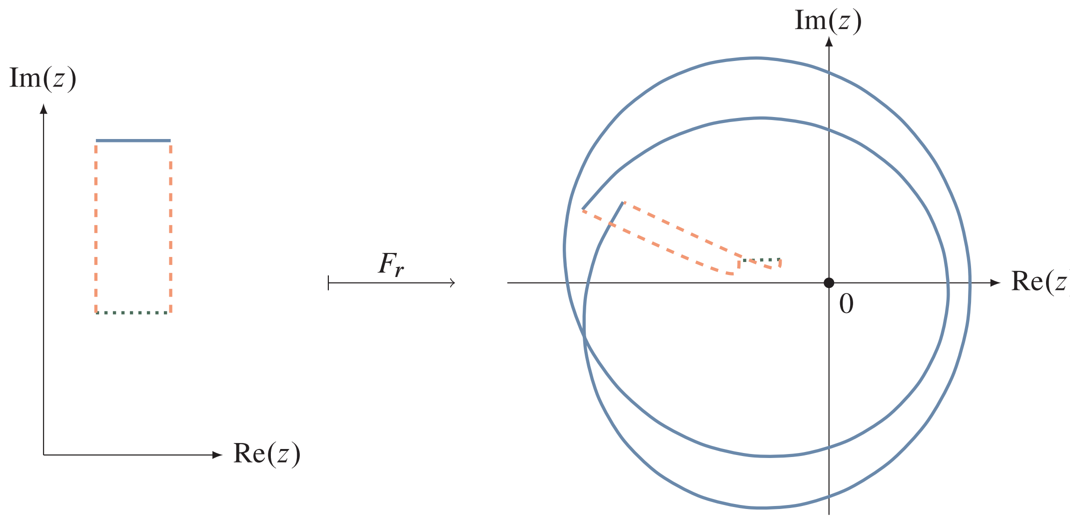

Figure 3. Visual representation of the action of F_r on the boundary of a typical rectangle Ξ_r for F of the form $Az^2 + Bz + Cj(z)^2 + Dj(z) + E$. The term j^2 has lowest order $(-2, 0)$, so F_r winds around 0 twice while following the top side of the rectangle.

Proof. Recall that $F_r(w) = F(r + sw, w)$. First, we apply Lemma 4.6 to $F_r(w)$ and find relevant constants $x_0, y_0, \delta, C, D, E_0, E_1$. Let $M = \lceil D \rceil$ if $D > 0$ and $M = 1$ otherwise. Let r be large enough so that $|r| \geq C$, $E_0 \log|r| \geq y_0$, and $E_1 \log|r| \leq |r|^{\frac{1}{M}}$; hence

$$|F_r(w)| \geq \delta \sum_{k=0}^{n} |f_k(w)||r + sw|^k$$

for $w \in \partial\Xi_r$. It follows, for instance, that $F_r(w)$ has order (e, d) at $i\infty$. We also take r sufficiently large so that each $f_k(w)$ does not have poles in Ξ_r, so in particular $F_r(w)$ is holomorphic on Ξ_r.

Note that $F_r(w)$ is never zero on the boundary of Ξ_r; thus its logarithmic derivative F_r'/F_r is holomorphic there. We shall now compute the integral of F_r'/F_r of F_r along such a boundary.

Vertical sides. We show that the images of the vertical sides of Ξ_r under F_r must be close to each other, and so their contributions cancel out as they are taken with opposite orientations. For instance, if F_r happens to be 1-periodic (as a function of w, for instance, when $s = 0$ and the coefficients of F are 1-periodic), then the images of these vertical sides are identical.

First, we observe that

$$F_r(w + 1) - F_r(w) = \sum_{k=0}^{n} f_k(w)(r + sw)^k \left(\frac{f_k(w + 1)}{f_k(w)} \left(\frac{r + sw + 1}{r + sw} \right)^k - 1 \right).$$

Since $f_k \in \mathcal{P}(w)$, we have that

$$f_k(w + 1) \sim f_k(w) \quad \text{as } \operatorname{Im}(w) \to +\infty.$$

Likewise, $r + sw + 1 \sim r + sw$ for $r + sw \to \infty$. Therefore, we can choose r large enough so that the last factor on the right-hand side has modulus less than $\frac{\delta}{2}$ for all k and for any w on the boundary of Ξ. We then have

$$\left|\frac{F_r(w+1)}{F_r(w)} - 1\right| = \left|\frac{F_r(w+1) - F_r(w)}{F_r(w)}\right| \leq \frac{\delta}{2|F_r(w)|}\left(\sum_{k=0}^{n} |f_k(w)||r+sw|^k\right) < \frac{1}{2} \tag{4.1}$$

for $w \in \partial\Xi$. We may now choose a branch of log in the disc around 1 of radius $\frac{1}{2}$ and estimate the integral along the vertical sides as[4)]

$$\left|\int_{E_0 \log|r|}^{|r|^{\frac{1}{M}}} \left(\frac{F_r'(x_0+1+iy)}{F_r(x_0+1+iy)} - \frac{F_r'(x_0+iy)}{F_r(x_0+iy)}\right) dy\right| = \left|\log\left(\frac{F_r(w+1)}{F_r(w)}\right)\Big|_{w=x_0+iE_0\log|r|}^{w=x_0+i|r|^{\frac{1}{M}}}\right| < \log\left(\frac{3}{2}\right) - \log\left(\frac{1}{2}\right) = \log(3) < 2.$$

Bottom side. We now show that the image of the bottom side is away from 0 and cannot wind much. Using (4.1), for r sufficiently large,

$$\left|\int_{x_0}^{x_0+1} \frac{F_r'(x+iE_0\log|r|)}{F_r(x+iE_0\log|r|)}\, dx\right| = \left|\log\left(\frac{F_r(x+iE_0\log|r|+1)}{F_r(x+iE_0\log|r|)}\right)\right| < \log\left(\frac{3}{2}\right) < 1.$$

Top side. We show that F_r behaves like $\exp(2\pi iew)$ on the top side of Ξ_r, and so the image under F_r is roughly a circle traversed approximately e times.

We now constraint w to the region $\mathrm{Im}(w) = |r|^{\frac{1}{M}}$, $x_0 \leq \mathrm{Re}(w) \leq x_0 + 1$. We have

$$\delta \max_k |f_k(w)||r+sw|^k \leq |F_r(w)| \leq n \max_k |f_k(w)||r+sw|^k.$$

By construction, $r \sim \zeta w^M$ for some power ζ of i depending on M and the sign of r. For simplicity, fix the sign of r, and assume that $\zeta = 1$, so that $r \sim w^M$. Then

$$F(w^M + sw, w) - F_r(w) = \sum_{k=0}^{n} f_k(w)(r+sw)^k \left(\left(\frac{w^M + sw}{r+sw}\right)^k - 1\right).$$

In particular, by Lemma 4.6, we find that $F(w^M + sw, w) - F_r(w) = o(F_r(w))$ for $r \to +\infty$. Since $F(w^M + sw, w)$ is in $\mathcal{P}(w)$, it has an order (e', d') at $i\infty$, and in fact, $e' = e$ because the term w^M cannot alter the exponential growth.

Therefore, we find that there is $\alpha \in \mathbb{C}^\times$ such that, for r large enough,

$$\left|\frac{F_r(w)}{\alpha w^{-d'} \exp(2\pi iew)} - 1\right| < \frac{1}{4}.$$

Observe that, for r large enough, we have

$$\left|\int_{x_0}^{x_0+1} \frac{((x+i|r|^{\frac{1}{M}})^{-d'} \exp(2\pi ie(x+i|r|^{\frac{1}{M}})))'}{(x+i|r|^{\frac{1}{M}})^{-d'} \exp(2\pi ie(x+i|r|^{\frac{1}{M}}))}\, dx - 2\pi ie\right| = \left|\int_{x_0}^{x_0+1} \frac{d'}{x+i|r|^{\frac{1}{M}}}\, dx\right| < 1.$$

4) Recall that $(fg)'/(fg) = f'/f + g'/g$, and that $\int_\gamma f'/f = \log(f(\gamma(1))) - \log(f(\gamma(0)))$ whenever $\gamma\colon [0,1] \to \mathbb{C}$ is a path and $f \circ \gamma$ takes values in a simply connected region on which we have fixed a continuous branch of log.

Thus, for sufficiently large r, we have

$$\left|\int_{x_0}^{x_0+1} \frac{F_r'(x+i|r|^{\frac{1}{N}})}{F_r(x+i|r|^{\frac{1}{N}})}\,dx - 2\pi i e\right| \le \left|\log\left(\frac{F_r(w)}{\alpha w^{-d'}\exp(2\pi i e w)}\right)\Big|_{w=x_0+i|r|^{\frac{1}{N}}}^{w=x_0+1+i|r|^{\frac{1}{N}}}\right| + 1 < 2.$$

Conclusion. Using the above estimates, we can now find the winding number of $F_r(\partial\Xi_r)$ at 0. Summing up the contributions from all sides, with the appropriate orientations, we obtain

$$\left|\frac{1}{2\pi i}\oint_{\partial\Xi_r}\frac{F_r(w)'}{F_r(w)}\,dw + e\right| < \frac{1}{2\pi}(2+1+2) < 1.$$

By the Argument Principle, the integral on the left-hand side must be the difference between the number of zeroes and poles of $F_r(w)$ inside Ξ (in particular an integer) multiplied by $2\pi i$. Since $F_r(w)$ is holomorphic on Ξ_r, it must have $-e$ zeroes in Ξ_r, counted with multiplicity. □

Proof of Proposition 1.6. This follows from combining Propositions 4.1 and 4.8, where we use $r = m$ a large integer and $s = 1$. Note that if $f_k(z)/f_n(z) \to \infty$ as $z \to i\infty$, then $f_k(z)/f_n(z)$ must have exponential growth at $i\infty$, for it is periodic and so has a q-expansion. □

5. Some criteria for Zariski density

Recall that $\mathbf{Y} := (Y_0, Y_1, Y_2)$ and $\mathbf{j} := (j, j', j'')\colon \mathbb{H} \to \mathbb{C}^3$.

5.1. Generic transforms. Given $p \in \mathbb{C}[X, \mathbf{Y}]$, or more generally $p \in K[X, \mathbf{Y}]$ for some field K, we define the *generic* $\mathrm{SL}_2(\mathbb{Z})$*-transform of* p as the polynomial $\Gamma(p) \in K[Z, W, C, \mathbf{Y}]$ given by

$$\Gamma(p)(Z, W, C, \mathbf{Y}) := p(Z, Y_0, W^2Y_1, W^4Y_2 + 2CW^3Y_1).$$

In particular, we have $\deg_X(p) = \deg_Z(\Gamma(p))$. By construction, for any $\gamma = \left(\begin{smallmatrix} a & b \\ c & d\end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$, if $n = \deg_X(p)$, we have

$$p(\gamma z, \mathbf{j}(\gamma z)) = \frac{p^\gamma(z, \mathbf{j}(z))}{(cz+d)^n}, \quad \text{where } p^\gamma(X, \mathbf{Y}) := (cX+d)^n\,\Gamma(p)\Big(\frac{aX+b}{cX+d}, cX+d, c, \mathbf{Y}\Big).$$

Note that $p^\gamma \in K[X, \mathbf{Y}]$.

We make the following observations.

(O1) The map $\Gamma\colon K[X, \mathbf{Y}] \to K[Z, W, C, \mathbf{Y}]$ defined above is a K-algebra homomorphism with left inverse $p(X, \mathbf{Y}) = \Gamma(p)(X, 1, 0, \mathbf{Y})$.

(O2) For each $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, the map $p \mapsto p^\gamma$ is multiplicative, that is, $(p_1p_2)^\gamma = p_1^\gamma p_2^\gamma$ for any $p_1, p_2 \in K[X, \mathbf{Y}]$. Indeed, this follows at once from the fact that Γ is a homomorphism (O1) and $\deg_X(p_1p_2) = \deg_X(p_1) + \deg_X(p_2)$.

(O3) For any $p \in K[X, \mathbf{Y}]$ and any $\gamma_1, \gamma_2 \in \mathrm{SL}_2(\mathbb{Z})$, there is $r \in \mathbb{Z}[X]$ such that

$$(p^{\gamma_1})^{\gamma_2} = r(X)p^{\gamma_1\gamma_2}.$$

Indeed, let $n = \deg_X(p)$ and $m = \deg_X(p^{\gamma_1})$. By construction, $m = n + \deg_{Y_0}(p) \ge n$. Now write

$$\gamma_t = \begin{pmatrix} a_t & b_t \\ c_t & d_t\end{pmatrix} \quad \text{for } t \in \{1, 2\}, \quad \text{and} \quad \gamma_1\gamma_2 = \begin{pmatrix} \tilde{a} & \tilde{b} \\ \tilde{c} & \tilde{d}\end{pmatrix}.$$

Thus

$$\frac{p^{\gamma_1\gamma_2}(z,\mathbf{j}(z))}{(\widetilde{c}z+\widetilde{d})^n} = p(\gamma_1\gamma_2 z, \mathbf{j}(\gamma_1\gamma_2 z)) = \frac{p^{\gamma_1}(\gamma_2 z, \mathbf{j}(\gamma_2 z))}{(c_1\gamma_2 z + d_1)^n}$$
$$= \frac{(p^{\gamma_1})^{\gamma_2}(z,\mathbf{j}(z))}{(c_1\gamma_2 z + d_1)^n (c_2 z + d_2)^m}.$$

Since $m \geq n$ and $(c_1\gamma_2 z + d_1)(c_2 z + d_2) = \widetilde{c}z + \widetilde{d}$, we get

$$r(X) := \frac{(c_2 X + d_2)^m}{(\widetilde{c}X + \widetilde{d})^n}\Big(c_1 \frac{a_2 X + b_2}{c_2 X + d_2} + d_2\Big)^n = (c_2 X + d_2)^{m-n} \in \mathbb{Z}[X].$$

(O4) For any $p \in K[X, \mathbf{Y}]$ and any $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, if we consider p and p^γ as polynomials in the variables $\mathbf{Y}$ with coefficients in $K(X)$, then p is irreducible in $K(X)[\mathbf{Y}]$ if and only if so is p^γ.

Indeed, note that if p is not a unit (meaning it contains one of the variables Y_0, Y_1, Y_2), then p^γ is also not a unit. It follows by (O2) that if p is reducible in $K(X)[\mathbf{Y}]$, thus a product of two non-units, then so is p^γ. Likewise, if p^γ is reducible, then $(p^\gamma)^{\gamma^{-1}}$ is reducible too, and $(p^\gamma)^{\gamma^{-1}} = r(X)p$ for some $r(X) \in \mathbb{C}[X]$ by (O3); since $r(X)$ is a unit, it follows that p is reducible.

Proposition 5.1. *Let p be an irreducible polynomial in $\mathbb{C}[X, \mathbf{Y}] \setminus \mathbb{C}[X]$ and $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. Let h be the irreducible factor of p^γ that is not in $\mathbb{C}[X]$. Then the equation $p(z, \mathbf{j}(z)) = 0$ has a Zariski dense set of solutions if and only if the equation $h(z, \mathbf{j}(z)) = 0$ has a Zariski dense set of solutions.*

Proof. By (O3) and (O4), $p^\gamma = r(X)h$ and $h^{\gamma^{-1}} = s(X)p$ for some $r, s \in \mathbb{C}[X]$. It follows that $p(z, \mathbf{j}(z)) = 0$ and $h(z, \mathbf{j}(z)) = 0$ have the same solutions except possibly for the zeroes of $r(z)$ and $s(z)$. Since those are only finitely many, the solutions of the former equation are Zariski dense if and only if so are the solutions of the latter. □

5.2. Density criteria. We now apply the results of Section 4 to establish some useful criteria for Zariski density of solutions of equations involving z, $j(z)$, $j'(z)$, $j''(z)$.

Definition 5.2. Given a function $g \in \mathbb{C}(z, \mathbf{j}(z))$ and $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, we say that $g(z)$ *has exponential growth in $\gamma\mathbb{F}$* if $g(\gamma^{-1}z)$ has exponential growth at $i\infty$. Furthermore, if r is the cusp of $\gamma\mathbb{F}$ (that is, $r \in \mathbb{Q} \cup \{\infty\}$ is in the Euclidean closure of $\gamma\mathbb{F}$), then we define the *order of $g(z)$ in $\gamma\mathbb{F}$ at r* as $\mathrm{ord}_{z=i\infty}(g(\gamma^{-1}z))$.

Proposition 5.3. *Let $F(X, \mathbf{Y}) = \sum_{k=0}^n X^k p_k(\mathbf{Y})$ be a polynomial. Assume that, for some k and some $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, the function $p_k(\mathbf{j}(z))/p_n(\mathbf{j}(z))$ has exponential growth in $\gamma\mathbb{F}$. Then there are $\ell > 0$, $0 \leq x_0 < 1$, $M > 0$, $E_0 > 0$ such that, for all $m \in \mathbb{Z}$ sufficiently large, the function $F(z, \mathbf{j}(z))$ has ℓ zeroes, counted with multiplicity, within the region $m + \gamma\Xi_m$, where*

$$\Xi_m = \{z \in \mathbb{H} : x_0 < \mathrm{Re}(z) < x_0 + 1,\ E_0 \log|m| < \mathrm{Im}(z) < |m|^{\frac{1}{M}}\}.$$

Proof. Fix some $s \in \{0, 1\}$, $t \in \mathbb{R}$ to be determined later. For $m \in \mathbb{Z}$, let

$$F_m(z) := F(m + t + sz, \mathbf{j}(\gamma z)).$$

Likewise, set $G(z) := F(z, \mathbf{j}(z))$. By Proposition 4.8 applied to $F(z, \mathbf{j}(\gamma z))$, $F_m(z)$ has $\ell > 0$ zeroes in a certain region Ξ_m and is suitably bounded from below for $z \in \partial \Xi_m$, as long as m is sufficiently large.

If γ is upper triangular, namely $\gamma z = z + k$ for some k, we choose $s = 1$, $t = 0$, and observe that, since the functions of $\mathbf{j}$ are 1-periodic, $G(z+m) = F(z+m, \mathbf{j}(\gamma z)) = F_m(z)$; thus $G(z)$ has ℓ zeroes in each region $m + \Xi_m$.

Otherwise, let $s = 0$ and let t be the limit of γz as $z \to \infty$ (where in fact $t \in \mathbb{Q}$). In particular,

$$\begin{aligned} G(m + \gamma z) - F_m(z) &= F(m + \gamma z, \mathbf{j}(\gamma z)) - F(m + t, \mathbf{j}(\gamma z)) \\ &= \sum_{k=0}^{n} p_k(\mathbf{j}(\gamma z))(m+t)^k \Big(\Big(\frac{m + \gamma z}{m + t}\Big)^k - 1\Big). \end{aligned}$$

Thus, as soon as z is sufficiently large, the last factor on the right-hand side has modulus less than $\frac{1}{2}$ independently of m. Then pick m large enough so that this happens whenever $\operatorname{Im}(z) > E_0 \log|m|$, and so

$$|G(m + \gamma z) - F_m(z)| < \frac{1}{2} |F_m(z)|.$$

Therefore, by Rouché's Theorem 3.1, $G(m + \gamma z)$ and $F_m(z)$ have the same number of zeroes in Ξ_m, counted with multiplicity. It follows that $G(z)$ has ℓ zeroes in the region $m + \gamma \Xi_m$. □

Proposition 5.4. *Let $F(X, \mathbf{Y}) = \sum_{k=0}^{n} X^k p_k(\mathbf{Y})$ be irreducible. Suppose for some k the function $P(z) = p_k(\mathbf{j}(z))/p_n(\mathbf{j}(z))$ satisfies one of the following:*

(i) *$P(z)$ has a pole in $\mathbb{H}$, or*

(ii) *$P(z)$ has exponential growth in some fundamental domain.*

Then the equation $F(z, \mathbf{j}(z)) = 0$ has a Zariski dense set of solutions.

Proof. Suppose by contradiction that all the solutions of $F(z, \mathbf{j}(z)) = 0$ lie on a further hypersurface $G = 0$, where $G \in \mathbb{C}[X, \mathbf{Y}]$ is a non-constant polynomial not divisible by F. In particular, the algebraic subset of $\mathbb{C}^4$ defined by $\{F = G = 0\}$ has dimension two, so its projection onto the variables Y_0, Y_1, Y_2 has dimension at most two, meaning that the solutions satisfy an equation $H(\mathbf{j}(z)) = 0$ for some non-constant polynomial $H \in \mathbb{C}[\mathbf{Y}]$. The assumption on F implies that F depends on the variable X (i.e. $n \geq 1$); thus F and H are coprime.

By Propositions 4.1 and 5.3, for $m \in \mathbb{Z}$ large, there are regions Ξ_m such that the original equation has solutions in $m + \Xi_m$, and moreover the real part of each Ξ_m is bounded from above and below and the imaginary part is bounded away from 0 uniformly in m. Each solution can be in $m + \Xi_m$ for at most finitely many $m \in \mathbb{Z}$. This implies that, for some m sufficiently large, the function $H(\mathbf{j}(z))$ has infinitely many zeroes in the region $\bigcup_{|k|>|m|} \Xi_k$. If this union is bounded, then we conclude that H is constantly zero by the identity theorem from complex analysis, but this contradicts the algebraic independence of j, j', j''. So we assume that the union is unbounded, but by Lemma 4.3, there exist $\alpha \in \mathbb{C}^\times$, $d, e \in \mathbb{Z}$, and $C > 0$ such that $H(z) \sim \alpha z^d \exp(e 2\pi i z)$ in the region $U := \{z \in \mathbb{H} : \operatorname{Im}(z) \geq C \log|z|\}$. Then the only way H can have infinitely many zeroes in $\bigcup_{|k|>|m|} \Xi_k$ is if H is constantly zero, again contradicting the algebraic independence of j, j', j''. This completes the proof. □

Corollary 5.5. *Let $F(X, \mathbf{Y}) = \sum_{k=0}^{n} X^k p_k(\mathbf{Y})$ be irreducible. Suppose that p_n has a factor h such that the equation $h(\mathbf{j}(z)) = 0$ has a Zariski dense set of solutions. Then the equation $F(z, \mathbf{j}(z)) = 0$ has a Zariski dense set of solutions.*

Proof. If $n = 0$, then $F = p_0$ which, by irreducibility, is equal to a constant multiple of h, and so the result is immediate. If instead $n > 0$, then by irreducibility of F for some $k \in \{0, \ldots, n-1\}$, p_k is non-zero and not divisible by h. Then $p_k(\mathbf{j}(z))/p_n(\mathbf{j}(z))$ will have poles in $\mathbb{H}$ at those solutions of $h(\mathbf{j}(z)) = 0$ which satisfy $p_k(\mathbf{j}(z)) \neq 0$ (which exist since we are assuming Zariski density of the solutions of $h(\mathbf{j}(z)) = 0$). So now the corollary follows from Proposition 5.4. □

6. Zero estimates

In view of the results in the previous section, we will now look at the poles of quotients of the form $p_k(\mathbf{j})/p_n(\mathbf{j})$, with $p_k, p_n \in \mathbb{C}[\mathbf{Y}]$. We keep using the notation introduced in Section 5.1.

Given $p \in \mathbb{C}[X, \mathbf{Y}]$, we will prove a few estimates on the order of $p(z, \mathbf{j}(z))$ at different points, distinguishing three cases. Before that, we note that specialising the variables Y_1 and Y_2 of $\Gamma(p)$ at some complex values will almost always return an "obfuscated" copy of the original polynomial, which for instance cannot be constant unless p itself was. More precisely, we note the following trivial identity.

Lemma 6.1. *For every $\alpha, \beta \in \mathbb{C}$ with $\alpha \neq 0$, we have*

$$\Gamma(p)\Big(X, \alpha^{-1}U_1, \alpha\frac{U_2 - \alpha^{-4}\beta U_1^4}{2U_1^3}, Y_0, \alpha^2, \beta\Big) = p(X, Y_0, U_1^2, U_2).$$

Proof. Immediate. □

First, we consider unramified points of j, that is, points $\tau \in \mathbb{H}$ such that $j'(\tau) \neq 0$. If $(Y_0 - j(\tau))^s$ divides p, then $p(z, \mathbf{j}(z))$ has obviously order at least s at all conjugates of τ. We show that the order is exactly s for most conjugates.

Proposition 6.2. *Let $p \in \mathbb{C}[X, \mathbf{Y}]$ be non-zero, $\tau \in \mathbb{H}$. Suppose that $j'(\tau) \neq 0$ and let s be the maximum integer such that $(Y_0 - j(\tau))^s$ divides p. Then, for all γ in a Zariski open dense subset of $\mathrm{SL}_2(\mathbb{Z})$, the function $p(z, \mathbf{j}(z))$ has order s at $z = \gamma\tau$.*

Proof. Let $f \in \mathbb{C}[X, \mathbf{Y}]$ be such that $p = (Y_0 - j(\tau))^s f$. Then we need to show that $f(z, \mathbf{j}(z))$ has order 0 at $z = \gamma\tau$, for all γ in a Zariski open dense subset of $\mathrm{SL}_2(\mathbb{Z})$. In other words, it suffices to prove the proposition for the case when p is not divisible by $Y_0 - j(\tau)$ (and hence $s = 0$).

From now on, we assume $s = 0$. By Lemma 6.1, at $\alpha^2 = j'(\tau) \neq 0$, $\beta = j''(\tau)$, $U_1^2 = Y_1$, $U_2 = Y_2$, there are $V_1, V_2 \in \mathbb{C}(U_1, U_2)$ such that $\Gamma(p)(X, V_1, V_2, \mathbf{j}(\tau)) = p(X, j(\tau), Y_1, Y_2)$. Since p is not divisible by $Y_0 - j(\tau)$, $p(X, j(\tau), Y_1, Y_2)$ is a non-zero polynomial; hence, in particular, $r(Z, W, C) := \Gamma(p)(Z, W, C, \mathbf{j}(\tau))$ is also non-zero.

Now take $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ and write

$$p^{\gamma}(\tau, \mathbf{j}(\tau)) = (c\tau + d)^n r(\gamma\tau, c\tau + d, c) = (c\tau + d)^n p(\gamma\tau, \mathbf{j}(\gamma\tau)),$$

where $n = \deg_X(p)$. In particular, $p(\gamma\tau, \mathbf{j}(\gamma\tau)) = 0$ if and only if $r(\gamma\tau, c\tau + d, c) = 0$.

The map $\upsilon \colon \mathrm{SL}_2(\mathbb{C}) \to \mathbb{C}^3$ given by $\gamma \mapsto (\gamma\tau, c\tau + d, c)$ is injective, and since

$$\dim \mathrm{SL}_2(\mathbb{C}) = 3 = \dim \mathbb{C}^3,$$

by the fibre-dimension theorem, υ is also dominant. As r is a non-zero polynomial, there is a non-empty Zariski open subset U_1 of $\mathbb{C}^3$ such that r never vanishes on U. This then gives a Zariski open subset U_0 of $\mathrm{SL}_2(\mathbb{C})$ such that, for every $\gamma \in U_0$, $\upsilon(\gamma) \in U_1$. This implies that, for $\gamma \in U_0$, $p(z, \mathbf{j}(z))$ does not vanish at $z = \gamma\tau$, and hence $p(z, \mathbf{j}(z))$ has order $s = 0$ at $z = \gamma\tau$. □

With this proposition, we obtain the following special case of Theorem 1.2.

Corollary 6.3. *The equation $j(z) - u = 0$ has a Zariski dense set of solutions if and only if $u \notin \{0, 1728\}$.*

Proof. If $u \notin \{0, 1728\}$, then for any $\tau \in \mathbb{H}$ satisfying $j(\tau) = u$, we have $j'(\tau) \neq 0$. We need to show that the solutions of the equation $j(z) = u$ are Zariski dense. So suppose that $p(X, \mathbf{Y}) \in \mathbb{C}[X, \mathbf{Y}]$ is such that its zero locus contains all the solutions of $j(z) = u$. The solutions of $j(z) = u$ are precisely $\mathrm{SL}_2(\mathbb{Z})\tau$, where $j(\tau) = u$, so $p(\gamma\tau, \mathbf{j}(\gamma\tau)) = 0$ for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. Therefore, $p(z, \mathbf{j}(z))$ has positive order at $\gamma\tau$ for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, so by Proposition 6.2, $Y_0 - u$ divides p, thus proving Zariski density.

On the other hand, if $u \in \{0, 1728\}$, then for any $\tau \in \mathbb{H}$ satisfying $j(\tau) = u$, we have that $j'(\tau) = 0$, so the solutions of $j(z) = u$ lie in the proper Zariski closed subset given by $Y_1 = 0$. □

Now we consider the behaviour of $p(z, \mathbf{j}(z))$ towards the cusps of the fundamental domains. We write $T\mathbf{Y} := (TY_0, TY_1, TY_2)$. One can easily see that the order at the cusp of any given fundamental domain is at least $(-e, -N)$ for some N, where $e = \deg_T(p(X, T\mathbf{Y}))$. This is not far from the actual behaviour at most cusps.

Proposition 6.4. *Let $p \in \mathbb{C}[X, \mathbf{Y}]$. Then there is $0 \leq M \leq \deg_W(\Gamma(p))$ such that, for all γ in a Zariski open dense subset of $\mathrm{SL}_2(\mathbb{Z})$, the function $p(z, \mathbf{j}(z))$ has order $(-e, -M)$ at the cusp of $\gamma\mathbb{F}$, where e is the degree of $p(X, T\mathbf{Y})$ in T.*

Proof. Let $n = \deg_X(p)$, $e = \deg_T(p(X, T\mathbf{Y}))$, and write

$$p(X, T\mathbf{Y}) = \sum_{k=0}^{e} T^k p_k(X, \mathbf{Y}),$$

where each p_k is homogeneous of degree k in the variables $\mathbf{Y}$, namely

$$p_k(X, T\mathbf{Y}) = T^k p_k(X, \mathbf{Y}).$$

Since $\Gamma(p)(Z, W, C, T\mathbf{Y}) = \Gamma(p(X, T\mathbf{Y}))$, we have

$$\Gamma(p)(Z, W, C, T\mathbf{Y}) = \sum_{k=0}^{e} T^k \Gamma(p_k)(Z, W, C, \mathbf{Y}),$$

and so each $\Gamma(p_k)$ is still homogeneous of degree k in $\mathbf{Y}$.

Recall that, for $\mathrm{Im}(z) \to +\infty$, letting $q = \exp(2\pi i z)$, we have

$$j(z) = q^{-1} + O(1), \quad j'(z) = -2\pi i q^{-1} + O(1), \quad j''(z) = -4\pi^2 q^{-1} + O(1).$$

Now fix some arbitrary $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$. We have

$$p^{\gamma}(z, \mathbf{j}(z)) = q^{-e}(cz + d)^n r(\gamma z, cz + d, c) + O(q^{-e+1} z^{n+N})$$

for $\mathrm{Im}(z) \to +\infty$ in the standard fundamental domain, where N is the degree of $\Gamma(p)$ in the variable W, and $r(Z, W, C) := \Gamma(p_e)(Z, W, C, 1, -2\pi i, -4\pi^2)$. More precisely, when $c \neq 0$, letting $M = \deg_W(r)$, we can also say

$$p^{\gamma}(z, \mathbf{j}(z)) = q^{-e}(cz)^n r\Big(\frac{a}{c}, cz, c\Big) + O(q^{-e} z^{n+M-1}).$$

Therefore, the order of $p^{\gamma}(z, \mathbf{j}(z))$ at $i\infty$ is $(-e, -(n + M))$, and so the order of $p(\gamma z, \mathbf{j}(\gamma z))$ is $(-e, -M)$, unless the leading coefficient of $r(\frac{a}{c}, W, c)$ vanishes or $c = 0$.

To conclude, since the map $\gamma \mapsto (a, c)$ from $\mathrm{SL}_2(\mathbb{C})$ to $\mathbb{C}^2$ is dominant, it suffices to prove that r is non-trivial (like we did in the proof of Proposition 6.2). By Lemma 6.1, at $\alpha^2 = -2\pi i$, $\beta = -4\pi$, $U_1^2 = Y_1 Y_0^{-1}$, $U_2 = Y_2 Y_0^{-1}$, there are $V_1, V_2 \in \mathbb{C}(U_1, U_2)$ such that

$$r(X, V_1, V_2) = \Gamma(p_e)(X, V_1, V_2, 1, -2\pi i, -4\pi^2) = p_e\Big(X, \frac{Y_0}{Y_0}, \frac{Y_1}{Y_0}, \frac{Y_2}{Y_0}\Big) = Y_0^{-e} p_e(X, \mathbf{Y}),$$

where the last equality follows from the homogeneity of p_e. Since by assumption $p_e \neq 0$, this shows that r is non-trivial, as desired. □

Example 6.5. It is easy to construct examples where a function has no exponential growth at *some* cusp. For instance, if $p(\mathbf{Y}) = 4\pi^2 Y_0 + Y_2$, then $p(\mathbf{j}(z)) = 4\pi^2 j(z) + j''(z)$ is bounded for $z \to i\infty$ in the standard fundamental domain. However, after the transformation $z \mapsto -\frac{1}{z}$, one gets $4\pi^2 j(z) + z^4 j''(z) + 2z^3 j'(z)$, which has order $(-1, -4)$ at $i\infty$, since $z^4 j''(z)$ is the dominant term. Note that here $4 = \deg_W(\Gamma(p))$.

There are also simple examples where the dominant terms cancel out at all cusps. Take $h(\mathbf{Y}) = Y_1^2 - Y_0 Y_2$. Then

$$h(\mathbf{j}(\gamma z)) = (cz + d)^4 j'(z)^2 - (cz + d)^4 j(z) j''(z) - 2c(cz + d)^3 j(z) j'(z)$$

has order $(-2, -3)$ for $c \neq 0$, because $j(z)j''(z) \sim (j'(z))^2 \sim -4\pi^2 q^{-2}$, and it has order $(-2, 0)$ for $c = 0$. On the other hand, $\deg_W(\Gamma(h)) = 4$. Proposition 6.4 guarantees that, even though these cancellations may occur at all cusps, some term of maximal exponential growth is not cancelled, at least generically.

Finally, we compute the order of $p(z, \mathbf{j}(z))$ at the points τ for which $j'(\tau) = 0$, namely the orbits of ρ and i. Here, the order is at least the maximum ν such that T^{ν} divides respectively $p(X, T^3 Y_0, T^2 Y_1, T Y_2)$ (for the conjugates of ρ) and $p(X, T^2 Y_0 + 1728, T Y_1, Y_2)$ (for the conjugates of i). This estimate is not sharp when p depends on Y_2, as we show in an example below, so we restrict to $p \in \mathbb{C}[X, Y_0, Y_1]$.

Proposition 6.6. *Let $p \in \mathbb{C}[X, Y_0, Y_1]$, $\tau \in \mathbb{H}$, $u = j(\tau)$, and let μ be the order of $j(z) - u$ at $z = \tau$. Then, for all γ in a Zariski open dense subset of $\mathrm{SL}_2(\mathbb{Z})$, the order of $p(z, \mathbf{j}(z))$ at $z = \gamma\tau$ is the highest power of T dividing $p(X, T^{\mu}Y_0 + u, T^{\mu-1}Y_1)$.*

Proof. The proof is very similar to that of Proposition 6.4. Write

$$p(X, T^{\mu}Y_0 + u, T^{\mu-1}Y_1) = \sum_{k=\nu}^{m} T^k p_k(X, \mathbf{Y}),$$

where each p_k satisfies the homogeneity condition

$$p_k(X, T^{\mu}Y_0, T^{\mu-1}Y_1) = T^k p_k(X, Y_0, Y_1)$$

and $p_\nu \neq 0$. Since

$$\Gamma(p)(X, T^{\mu}Y_0 + u, T^{\mu-1}Y_1) = \Gamma(p(X, T^{\mu}Y_0 + u, T^{\mu-1}Y_1)),$$

we have

$$\Gamma(p)(T^{\mu}Y_0 + u, T^{\mu-1}Y_1) = \sum_{k=\nu}^{m} T^k \Gamma(p_k)(X, \mathbf{Y})$$

and each $\Gamma(p_k)$ satisfies the above homogeneity condition.

Fix $\alpha_0 \in \mathbb{C}$ such that $j(z) - u \sim \alpha_0(z-\tau)^{\mu}$, and so also $j'(z) \sim \mu\alpha_0(z-\tau)^{\mu-1}$, for $z \to \tau$, where by assumption $\mu\alpha_0 \neq 0$. For $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$, we have

$$p^{\gamma}(z, \mathbf{j}(z)) = (c\tau + d)^n r(\gamma\tau, c\tau + d, c)(z-\tau)^{\nu} + O((z-\tau)^{\nu-1})$$

as $z \to \tau$, where

$$r(Z, W, C) := \Gamma(p_\nu)(Z, W, C, \alpha_0, \mu\alpha_0).$$

It follows that the order of $p^{\gamma}(z, \mathbf{j}(z))$ at $z = \tau$ is ν as long as r does not vanish. Since $z \mapsto \gamma z$ is a diffeomorphism, this coincides with the order of $p(z, \mathbf{j}(z))$ at $z = \gamma\tau$.

To conclude, since the map $\gamma \mapsto (\gamma\tau, c\tau + d, c)$ is injective on $\mathrm{SL}_2(\mathbb{C})$, hence dominant on $\mathbb{C}^3$, we only need to show that r is non-trivial. Pick β_0 such that $\beta_0^{2\mu} = \alpha_0^{-1}$ and U_0 such that $U_0^{2\mu} = Y_0$. By Lemma 6.1, at $\alpha^2 = \mu\alpha_0$, $\beta = 0$, $U_1^2 = (\beta_0 U_0)^{-2(\mu-1)}Y_1$, there are $V_1, V_2 \in \mathbb{C}(U_1, U_2)$ such that

$$\begin{aligned} r(X, V_1, V_2) = \Gamma(p_\nu)(X, V_1, V_2, \alpha_0, \mu\alpha_0) &= p_k\Big(X, \alpha_0, \frac{Y_1}{(\beta_0 U_0)^{2(\mu-1)}}\Big) \\ &= p_\nu\Big(X, \frac{U_0^{2\mu}}{(\beta_0 U_0)^{2\mu}}, \frac{Y_1}{(\beta_0 U_0)^{2(\mu-1)}}\Big) \\ &= (\beta_0 U_0)^{-2k} p_\nu(X, Y_0, Y_1), \end{aligned}$$

where the last equality is implied by the homogeneity condition. It follows that r is non-trivial, as desired. □

Example 6.7. The above method fails when p depends on Y_2. Indeed, to compute say the order of $p(\mathbf{j}(\gamma z))$ at ρ, we would look at the maximum power of T dividing

$$p^{\gamma}(X, T^3Y_0, T^2Y_1, TY_2).$$

However, if p contains Y_2, then

$$\Gamma(p)(X, T^3Y_0, T^2Y_1, TY_2) \neq \Gamma(p(X, T^3Y_0, T^2Y_1, TY_2)),$$

breaking the very first steps of the argument.

The order can indeed be higher than expected at all the conjugates of ρ or i. For instance, for the polynomial

$$p(\mathbf{Y}) = Y_0Y_2 - \frac{2}{3}Y_1^2,$$

the maximum power of T dividing $p(T^3Y_0, T^2Y_1, TY_2)$ is 4; however, the function $p(\mathbf{j}(z))$ has order at least 5 at all conjugates of ρ: if $j(z) \sim \alpha(z-\gamma\rho)^3$ for $z \to \gamma\rho$, then

$$p(\mathbf{j}(z)) = \alpha^2(z-\gamma\rho)^4\Big(6 - \frac{2}{3}3^2 + O(z-\gamma\rho)\Big) = O((z-\gamma\rho)^5).$$

Corollary 6.8. *For all $p \in \mathbb{C}[Y_0, Y_1]$, $h \in \mathbb{C}[Y_0]$, and $\ell \geq \deg_{Y_1}(p)$, if the function*

$$P(z) := \frac{p(j(z), j'(z))}{h(j(z))j'(z)^\ell}$$

is non-constant, then it has a pole in $\mathbb{H}$ or it has exponential growth in some fundamental domains.

Proof. Suppose that $P(z)$ has no poles in $\mathbb{H}$ and has no exponential growth in any fundamental domain. Without loss of generality, we may assume that p and h are coprime.

Let β be a root of $h(Y_0)$ such that $\beta \notin \{0, 1728\}$, and let τ be such that $j(\tau) = \beta$. In particular, $Y_0 - \beta$ does not divide p. By Proposition 6.2, $p(\gamma\tau, j(\gamma\tau)) \neq 0$ for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$ except for some proper Zariski closed set, and so P has a pole at $\gamma\tau$, a contradiction. Therefore, $h(Y_0)$ is of the form $\alpha Y_0^s(Y_0 - 1728)^t$ for some $\alpha \in \mathbb{C}$.

Now $h(j(z))j'(z)^\ell$ has order $3s + 2\ell$ at all the conjugates of ρ and $2t + \ell$ at all the conjugates of i; thus $p(j(z), j'(z))$ must have at least the same order at those points. On writing $p = \sum_{\ell'} Y_1^{\ell'} p_{\ell'}(Y_0)$, Proposition 6.6 implies that each $p_{\ell'}(Y_0)$ is divisible by

$$Y_0^{s+\lceil\frac{2(\ell-\ell')}{3}\rceil} \quad \text{and} \quad (Y_0 - 1728)^{t+\lceil\frac{\ell-\ell'}{2}\rceil}.$$

In particular, whenever $p_{\ell'} \neq 0$, we have

$$\deg(p) \geq \deg(p_{\ell'}) \geq s + t + \left\lceil\frac{2(\ell-\ell')}{3}\right\rceil + \left\lceil\frac{\ell-\ell'}{2}\right\rceil + \ell' \geq s + t + \ell' + \frac{7}{6}(\ell - \ell'),$$

with strict inequality if $p_{\ell'}$ is not a constant multiple of

$$Y_0^{s+\lceil\frac{2(\ell-\ell')}{3}\rceil}(Y_0 - 1728)^{t+\lceil\frac{\ell-\ell'}{2}\rceil}.$$

On the other hand, by Proposition 6.4, on a Zariski open dense set of fundamental domains, the denominator has order at most $(-(s+t+\ell), 0)$ at the cusp, while the numerator has order at most $(-\deg(p), 0)$; thus, whenever $p_{\ell'} \neq 0$, we also have

$$s + t + \ell' + \frac{7}{6}(\ell - \ell') \leq \deg(p) \leq s + t + \ell.$$

As $\ell' \leq \deg_{Y_1}(p) \leq \ell$ and $\frac{7}{6} > 1$, it follows at once that $\ell = \ell'$ and that $p = p_\ell$ is a constant multiple of $Y_0^s(Y_0 - 1728)^t Y_1^\ell$, and so that $P(z)$ is constant. □

7. The main result

7.1. j-homogeneous equations. Before tackling the general case of Theorem 1.2, we look at equations of the form $F(\mathbf{j}(z)) = 0$, where $F \in \mathbb{C}[\mathbf{Y}]$ satisfies the homogeneity condition below.

Definition 7.1. The $\mathbf{j}$-*degree* of $F \in \mathbb{C}[X, \mathbf{Y}]$ is the degree of $F(X, Y_0, T^2 Y_1, T^4 Y_2)$ in T, which we denote $\deg_{\mathbf{j}}(F)$. We say that F is $\mathbf{j}$-*homogeneous* if $F(X, Y_0, T^2 Y_1, T^4 Y_2)$ is homogeneous in the variable T.

One of the easiest examples of a $\mathbf{j}$-homogeneous polynomial is $F = Y_2$, which has $\mathbf{j}$-degree 4. For the sake of exposition, we first sketch the proof of Zariski density for this F, namely for the equation $j''(z) = 0$.

First, we observe that, for $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$,

$$j''(\gamma z) = j''(z)c^4\left(\left(z + \frac{d}{c}\right)^4 + 2\left(z + \frac{d}{c}\right)^3 \frac{j'(z)}{j''(z)}\right) = j''(z)c^4 h\left(z + \frac{d}{c}, z\right), \tag{7.1}$$

where

$$h(X, W) = X^4 + 2\frac{j'(W)}{j''(W)} X^3.$$

If we can find $\tau \in \mathbb{H}$ such that $j''(\tau) = 0 \neq j'(\tau)$, we are done by Proposition 5.4, so suppose by contradiction that this does not happen. In this case, $j'(z)/j''(z)$ is bounded on the standard fundamental domain $\mathbb{F}$: by construction, it cannot have poles in $\mathbb{H}$, and by looking at the q-expansions, it is also bounded for $\mathrm{Im}(z) \to +\infty$. This also implies that $h(\tau + \frac{d}{c}, \tau) \neq 0$ for any $\frac{d}{c} \in \mathbb{Q}$, $\tau \in \mathbb{H}$ except possibly when $j'(\tau) = 0$.

Second, under the above assumptions, we shall verify that $h(\tau + r, \tau) \neq 0$ for all $r \in \mathbb{R}$, $\tau \in \mathbb{H}$ (Claim 7.3.1), and in turn deduce that

$$\left|h\left(z + \frac{d}{c}, z\right)\right| \geq \varepsilon \left|z + \frac{d}{c}\right|^4$$

for all $z \in \mathbb{F}$, for some $\varepsilon > 0$ (Claim 7.3.2).

In particular, for all $z \in \mathbb{F}$, we have

$$\left|\frac{j'(\gamma z)}{j''(\gamma z)}\right| \leq \left|\frac{j'(z)}{\varepsilon j''(z)}\right| |cz + d|^{-2} = \left|\frac{j'(z)}{\varepsilon j''(z)}\right| \frac{\mathrm{Im}(\gamma z)}{\mathrm{Im}(z)} \leq \frac{2M\,\mathrm{Im}(\gamma z)}{\sqrt{3}\varepsilon}$$

where M is a bound for $|j'(z)/j''(z)|$ on $\mathbb{F}$, and so $j'(z)/j''(z) \to 0$ as $\mathrm{Im}(z) \to 0$. By the Schwarz Reflection Principle, $j'(z)/j''(z)$ extends to a holomorphic function on $\mathbb{C}$ that vanishes for all $z \in \mathbb{R}$, and is thus constantly 0, a contradiction.

For a general $\mathbf{j}$-homogeneous F, we just need to find an appropriate generalisation of equation (7.1) and fill the details in the above sketch.

Lemma 7.2. *Let $F \in \mathbb{C}[\mathbf{Y}]$ be $\mathbf{j}$-homogeneous. Then there are polynomials $p_k \in \mathbb{C}[\mathbf{Y}]$ and $h \in \mathbb{C}[X, \mathbf{Y}]$ such that, for all $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$ with $c \neq 0$, we have*

$$F(\mathbf{j}(\gamma z)) = F(\mathbf{j}(z))c^N\left(\sum_{k=k_0}^{N} \frac{p_k(\mathbf{j}(z))}{F(\mathbf{j}(z))}\left(z + \frac{d}{c}\right)^k\right) = F(\mathbf{j}(z))c^N h\left(z + \frac{d}{c}, \mathbf{j}(z)\right),$$

where $N = \deg_{\mathbf{j}}(F)$, $p_N = F$, and $0 \neq p_{k_0} \in Y_1^{\ell} \cdot \mathbb{C}[Y_0]$ with $2\ell \leq k_0$.

Proof. Let F be as in the hypothesis. By $\mathbf{j}$-homogeneity, we have that

$$\begin{aligned}\Gamma(F)(Z, W, C, \mathbf{Y}) &= F(Z, Y_0, W^2 Y_1, W^4 Y_2 + 2CW^3 Y_1) \\ &= C^N F\Big(Z, Y_0, \frac{W^2}{C^2} Y_1, \frac{W^4}{C^4} Y_2 + 2\frac{W^3}{C^3} Y_1\Big).\end{aligned}$$

Therefore, we can write $\Gamma(F)$ as

$$\Gamma(F) = C^N \sum_{k=0}^{N} p_k(\mathbf{Y}) \frac{W^k}{C^k},$$

where $N = \deg_{\mathbf{j}}(F)$. Let k_0 be the least integer such that $p_{k_0} \neq 0$.

By a further application of $\mathbf{j}$-homogeneity, we also have

$$\Gamma(F) = W^N Y_1^{\frac{N}{2}} F\Big(Z, Y_0, 1, \frac{Y_2}{Y_1^2} + \frac{2C}{WY_1}\Big).$$

It follows at once that the terms of maximum degree in W, which make up $(W^N/C^N)p_N$, are found by discarding $2C/(WY_1)$, and in particular, we discover that $p_N(\mathbf{Y}) = F(\mathbf{Y})$. Similarly, the terms of lowest degree in W/C are obtained by specialising at $Y_2 = 0$ and taking the least power of Y_1, and so $p_{k_0} \in Y_1^{\ell} \cdot \mathbb{C}[Y_0]$. Moreover, if t is the degree of F in Y_2, we have that $k_0 = N - t$, $\ell = \frac{N}{2} - t$; thus $2\ell \leq k_0$. □

Theorem 7.3. *For any irreducible* $\mathbf{j}$*-homogeneous polynomial* $F(\mathbf{Y}) \notin \mathbb{C}[X, Y_0, Y_1]$, *the equation* $F(\mathbf{j}(z)) = 0$ *has a Zariski dense set of solutions.*

Proof. Let F be as in the hypothesis and fix the polynomials p_k, h as in the conclusion of Lemma 7.2.

If some $p_k(\mathbf{j}(z))/F(\mathbf{j}(z))$ has a pole in $\mathbb{H}$ or exponential growth in some fundamental domain, we are done by Proposition 5.4. Therefore, we shall assume that this is not the case. In particular, we assume that $h(z + \frac{d}{c}, \mathbf{j}(z))$ has no pole in $\mathbb{H}$ for any $\frac{d}{c}$.

With this additional assumption, if $F(\mathbf{j}(\tau)) = 0$, then $F(\mathbf{j}(\gamma\tau)) = 0$ for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. Since F is not divisible by $Y_0 - j(\tau)$, Proposition 6.2 implies that $j'(\tau) = 0$. In particular, for any $\frac{d}{c}$, $h(\tau + \frac{d}{c}, \mathbf{j}(\tau)) = 0$ implies that $j'(\tau) = 0$.

Claim 7.3.1. *For every* $u \in \mathbb{R}$, *the function* $h(z + u, \mathbf{j}(z))$ *has no zero in* $\mathbb{H}$.

Proof of the claim. Suppose $h(\tau + u, \mathbf{j}(\tau)) = 0$ for some $(\tau, u) \in \mathbb{H} \times \mathbb{R}$. As $h(Z, \mathbf{Y})$ is monic in Z, the analytic map $(z, u) \mapsto (h(z + u, \mathbf{j}(z)), u)$ has finite fibres, in particular of dimension zero. Then the Open Mapping Theorem implies that the image of any ball around (τ, u) contains an open neighbourhood of $(0, u)$. In particular, for every rational number r arbitrarily close to u, there is τ_r close to τ such that $h(\tau_r + r, \mathbf{j}(\tau_r)) = 0$.

Since h is monic in Z, the polynomial $h(\tau + Z, \mathbf{j}(\tau))$ is not identically zero; thus, for r sufficiently close to u, we have $h(\tau + r, \mathbf{j}(\tau)) \neq 0$, and so $\tau_r \neq \tau$. Therefore, the τ_r's can be chosen to accumulate at τ for $r \to u$. However, our assumptions imply that $j'(\tau_r) = 0$, so the τ_r's lie in a closed discrete subset of $\mathbb{H}$ (namely, the orbits of ρ and i), a contradiction. □

Claim 7.3.2. *There is $\varepsilon > 0$ such that*

$$|h(z+u, \mathbf{j}(z))| \geq \varepsilon |z+u|^N$$

for all $z \in \mathbb{F}$ and $u \in \mathbb{R}$.

Proof of the claim. Our current assumptions imply that each $p_k(\mathbf{j}(z))/F(\mathbf{j}(z))$ has neither a pole in $\overline{\mathbb{F}}$ nor exponential growth in $\mathbb{F}$. Since these functions are in $\mathcal{P}$ rather than $\mathcal{P}(w)$, their order at $i\infty$ is of the form $(e, 0)$, thus at least $(0, 0)$, and so they are bounded in $\overline{\mathbb{F}}$, say by $M > 0$. In particular, since h is monic in Z, we have $|h(z+u, \mathbf{j}(z))| > \frac{1}{2}|z+u|^N$ as soon as $|z+u| > 2NM$. On the other hand, $|z+u| \leq 2NM$ defines a compact subset K of $\overline{\mathbb{F}} \times \mathbb{R}$, and so the function

$$\left|\frac{h(z+u, \mathbf{j}(z))}{(z+u)^N}\right|$$

attains some minimum $\varepsilon' > 0$ on K, since it does not vanish by Claim 7.3.1. The conclusion follows on taking $\varepsilon = \min\{\varepsilon', \frac{1}{2}\}$. □

Therefore, for $z \in \mathbb{F}$, we find that

$$|F(\mathbf{j}(\gamma z))| \geq \varepsilon |F(\mathbf{j}(z))| c^N \left|z + \frac{d}{c}\right|^N = \varepsilon |F(\mathbf{j}(z))| |cz+d|^N,$$

and in particular, for some $M > 0$ independent of z, we have

$$\left|\frac{p_{k_0}(\mathbf{j}(\gamma z))}{F(\mathbf{j}(\gamma z))}\right| = \frac{|p_{k_0}(\mathbf{j}(z))||cz+d|^{2\ell}}{|F(\mathbf{j}(\gamma z))|} \leq \frac{|p_{k_0}(\mathbf{j}(z))|}{\varepsilon |F(\mathbf{j}(z))|}|cz+d|^{2\ell - N} \leq M \operatorname{Im}(\gamma z)^{\ell - \frac{N}{2}},$$

where we have used that $2\ell < N$ and that, for $z \in \mathbb{F}$, $p_{k_0}(\mathbf{j}(z))/F(\mathbf{j}(z))$ is bounded while also

$$|cz+d|^2 = \frac{\operatorname{Im}(z)}{\operatorname{Im}(\gamma z)} \geq \frac{\sqrt{3}}{2\operatorname{Im}(\gamma z)} > 0.$$

Therefore,

$$\frac{p_{k_0}(\mathbf{j}(z))}{F(\mathbf{j}(z))} \to 0 \quad \text{as } \operatorname{Im}(z) \to 0;$$

thus, by Schwarz's reflection principle, $p_{k_0}(\mathbf{j}(z))/F(\mathbf{j}(z))$ has a holomorphic extension to $\mathbb{C}$ that is constantly zero on $\mathbb{R}$, thus constantly zero on $\mathbb{C}$; hence $p_{k_0} = 0$, a contradiction. □

Remark 7.4. Let $S = \{z \in \mathbb{H} : F(\mathbf{j}(z)) = 0\}$ for some F as in Theorem 7.3. As observed in the proof, every $\tau \in S$ that is not in the $\mathrm{SL}_2(\mathbb{Z})$-orbit of ρ or i must also be a pole of some coefficient $p_k(\mathbf{j}(z))/F(\mathbf{j}(z))$ appearing in Lemma 7.2. Combining this information with Proposition 4.1, it follows that $j(S)$ is infinite and every point of $j(S) \setminus \{0, 1728\}$ is an accumulation point of $j(S)$.

7.2. Proof of Theorem 1.2. For the rest of the section, fix some $F \in \mathbb{C}[X, \mathbf{Y}]$ and let $p_k \in \mathbb{C}[Z, C, \mathbf{Y}]$ be polynomials such that

$$\Gamma(F)(Z, W, C, \mathbf{Y}) = \sum_{k=0}^{N} p_k(Z, C, \mathbf{Y}) W^k,$$

where $N = \deg_W(\Gamma(F))$. Given a polynomial $\alpha \in \mathbb{C}[X, \mathbf{Y}]$, we will use the notation[5)]

$$\alpha^{\mathbb{N}} := \{\alpha^s : s \in \mathbb{N}\}.$$

Given different polynomials $\alpha_1, \dots, \alpha_\ell \in \mathbb{C}[X, \mathbf{Y}]$, we use $\alpha_1^{\mathbb{N}} \cdots \alpha_\ell^{\mathbb{N}}$ to denote the set of all products between elements of different $\alpha_t^{\mathbb{N}}$.

Proposition 7.5. *The polynomial p_N is the sum of the terms of maximum $\mathbf{j}$-degree in $F(Z, Y_0, W^2Y_1, W^4Y_2)$. In particular, $N = \deg_{\mathbf{j}}(F)$, p_N is $\mathbf{j}$-homogeneous, and p_N does not depend on C.*

Proof. Let $X^\alpha Y_0^{\beta_0} Y_1^{\beta_1} Y_2^{\beta_2}$ denote a monomial, so $\alpha, \beta_0, \beta_1, \beta_2 \in \mathbb{N}$. Observe that

$$\Gamma(X^\alpha Y_0^{\beta_0} Y_1^{\beta_1} Y_2^{\beta_2}) = Z^\alpha Y_0^{\beta_0} Y_1^{\beta_1} W^{2\beta_1+3\beta_2}(WY_2 + 2CY_1)^{\beta_2}.$$

Hence

$$\deg_W(\Gamma(X^\alpha Y_0^{\beta_0} Y_1^{\beta_1} Y_2^{\beta_2})) = 2\beta_1 + 4\beta_2$$

and the term accompanying $W^{2\beta_1+4\beta_2}$ is $Z^\alpha Y_0^{\beta_0} Y_1^{\beta_1} Y_2^{\beta_2}$.

Now write $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)$ and

$$F(X, \mathbf{Y}) = \sum_{(\alpha, \boldsymbol{\beta}) \in \mathbb{N}^4} c_{\alpha,\boldsymbol{\beta}} X^\alpha \mathbf{Y}^{\boldsymbol{\beta}},$$

where $\mathbf{Y}^{\boldsymbol{\beta}} = Y_0^{\beta_0} Y_1^{\beta_1} Y_2^{\beta_2}$ and $c_{\alpha,\boldsymbol{\beta}} \in \mathbb{C}$. Then, since Γ is a homomorphism by (O1),

$$\Gamma(F) = \sum_{(\alpha, \boldsymbol{\beta}) \in \mathbb{N}^4} c_{\alpha,\boldsymbol{\beta}} \Gamma(X^\alpha \mathbf{Y}^{\boldsymbol{\beta}}).$$

Since $N = \deg_W(\Gamma(F))$, then

$$p_N = \sum_{(\alpha, \boldsymbol{\beta}) \in \mathbb{N}^4 : 2\beta_1 + 4\beta_2 = N} c_{\alpha,\boldsymbol{\beta}} Z^\alpha Y_0^{\beta_0} Y_1^{\beta_1} Y_2^{\beta_2}.$$

From this, we see that p_N does not depend on C and that p_N is $\mathbf{j}$-homogeneous. This expression also gives us that p_N is the coefficient of W^N in $F(Z, Y_0, W^2Y_1, W^4Y_2)$, and so

$$N = \deg_W(F(Z, Y_0, W^2Y_1, W^4Y_2)) = \deg_{\mathbf{j}}(F). \qquad \square$$

Definition 7.6. The $\mathbf{j}$-*order* of F, denoted by $\mathrm{ord}_{\mathbf{j}}(F)$, is the maximum power of T dividing $F(X, Y_0, T^2Y_1, T^3Y_2)$.

Proposition 7.7. *Let k_0 be minimum such that $p_{k_0} \neq 0$. Then p_{k_0} is the sum of the terms of minimum degree in W of $F(Z, Y_0, W^2Y_1, 2CW^3Y_1)$. In particular,*

$$k_0 = \mathrm{ord}_{\mathbf{j}}(F) \geq 2 \deg_{Y_1}(p_{k_0})$$

and p_{k_0} does not depend on Y_2.

5) We remark that, for us, $0 \in \mathbb{N}$.

Proof. We proceed as in the proof of Proposition 7.5. From

$$\Gamma(X^\alpha Y_0^{\beta_0} Y_1^{\beta_1} Y_2^{\beta_2}) = Z^\alpha Y_0^{\beta_0} Y_1^{\beta_1} W^{2\beta_1+3\beta_2}(WY_2 + 2CY_1)^{\beta_2},$$

we see that the smallest power of W appearing in this expression is $2\beta_1 + 3\beta_2$, and it is accompanied by $2^{\beta_2} C^{\beta_2} Z^\alpha Y_0^{\beta_0} Y_1^{\beta_1+\beta_2}$, which does not depend on Y_2. Hence

$$p_{k_0} = \sum_{(\alpha,\boldsymbol{\beta})\in\mathbb{N}^4 : 2\beta_1+3\beta_2=k_0} c_{\alpha,\boldsymbol{\beta}} 2^{\beta_2} C^{\beta_2} Z^\alpha Y_0^{\beta_0} Y_1^{\beta_1+\beta_2},$$

which also shows that p_{k_0} is the coefficient of W^{k_0} in $F(Z, Y_0, W^2 Y_1, 2CW^3 Y_1)$. Using the change of variables $C = Y_2/(2Y_1)$, we conclude that $k_0 = \deg_{\mathbf{j}}(F)$, and since

$$k_0 = 2\beta_1 + 3\beta_2 \geq 2(\beta_1 + \beta_2),$$

we have $k_0 \geq 2 \deg_{Y_1}(p_{k_0})$, concluding the proof. □

Corollary 7.8. *If $F \notin Y_1^{\mathbb{N}} \mathbb{C}[X, Y_0]$, then* $\deg_{\mathbf{j}}(F) > 2 \deg_{Y_1}(p_{\mathrm{ord}_{\mathbf{j}}(F)})$.

Proof. One can immediately verify that

$$\begin{aligned}\deg_{\mathbf{j}}(F) = \deg_T(F(X, Y_0, T^2 Y_1, T^4 Y_2)) &\geq \deg_T(F(X, Y_0, T^2 Y_1, T^3 Y_2)) \\ &\geq \mathrm{ord}_{T=0}(F(X, Y_0, T^2 Y_1, T^3 Y_2)) = \mathrm{ord}_{\mathbf{j}}(F) \\ &\geq 2 \deg_{Y_1}(p_{\mathrm{ord}_{\mathbf{j}}(F)}),\end{aligned}$$

where $\mathrm{ord}_{T=0}(P)$ is the maximum power of T dividing P.

If the second inequality is an equality, then

$$F(X, Y_0, T^2 Y_1, T^3 Y_2) = T^m F(X, Y_0, Y_1, Y_2),$$

where $m = \mathrm{ord}_{\mathbf{j}}(F)$. In particular,

$$\deg_T(F(X, Y_0, T^2 Y_1, T^4 Y_2)) = \deg_T(T^m F(X, Y_0, Y_1, TY_2)) = m + \deg_{Y_2}(F).$$

If the first inequality is also an equality, then $\deg_{Y_2}(F) = 0$, and moreover F is homogeneous in Y_1 of degree $\frac{m}{2}$; thus $F \in Y_1^{\mathbb{N}} \mathbb{C}[X, Y_0]$. □

We can now prove the main result of this paper, Theorem 1.2, the statement of which is recalled below for the convenience of the reader.

Theorem 1.2. *For any polynomial $F(X, \mathbf{Y}) \in \mathbb{C}[X, \mathbf{Y}] \setminus \mathbb{C}[X]$ which is coprime to $Y_0(Y_0 - 1728)Y_1$, the equation $F(z, \mathbf{j}(z)) = 0$ has a Zariski dense set of solutions.*

Proof. It suffices to prove the conclusion for F irreducible, not in $\mathbb{C}[X]$, and not a constant multiple of Y_0, $Y_0 - 1728$, or Y_1. Let $n = \deg_X(F)$ and write

$$\begin{aligned}F^\Gamma &= (CX + D)^n \Gamma(F)\Big(\frac{AX + B}{CX + D}, CX + D, C, \mathbf{Y}\Big) \\ &= (CX + D)^n \sum_{k=0}^{N} p_k\Big(\frac{AX + B}{CX + D}, C, \mathbf{Y}\Big)(CX + D)^k \\ &= \sum_{k=0}^{n+N} h_k(A, B, C, D, \mathbf{Y}) X^k.\end{aligned}$$

We recall that $F^{\gamma} = F^{\Gamma}(a,b,c,d,X,\mathbf{Y})$ for any $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$. If $F \in \mathbb{C}[Y_0]$, the conclusion follows by Corollary 6.3, so we may assume that this is not the case, and in particular that $n + N > 0$.

We observe immediately that $h_{n+N} = C^n p_N(\frac{A}{C}, \mathbf{Y})$ using Proposition 7.5. Moreover, by Proposition 7.7,

$$h_0 = D^n \sum_{k=0}^{N} p_k\Big(\frac{B}{D}, C, \mathbf{Y}\Big) D^k = D^n \sum_{k=\mathrm{ord}_{\mathbf{j}}(F)}^{N} p_k\Big(\frac{B}{D}, C, \mathbf{Y}\Big) D^k.$$

We claim that, for $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ in some Zariski open dense subset of $\mathrm{SL}_2(\mathbb{Z})$, we have that $h_{n+N}(a,b,c,d,\mathbf{Y})$ has a factor $r \in \mathbb{C}[\mathbf{Y}]$ such that the equation $r(\mathbf{j}(z)) = 0$ has a Zariski dense set of solutions, or that $h_0/h_{n+N}(a,b,c,d,\mathbf{j}(z))$ has a pole or exponential growth in some fundamental domain. In particular, for any one of those γ's, we find that

$$F^{\gamma}(z, \mathbf{j}(z)) = \sum_{k=0}^{n+N} h_k(a,b,c,d,\mathbf{j}(z)) z^k = 0$$

has a Zariski dense set of solutions (by Corollary 5.5 in the first case, and Proposition 5.4 in the second one); hence so does $F(z, \mathbf{j}(z)) = 0$ (by Proposition 5.1), as desired.

To prove the claim, we distinguish three cases.

Suppose that p_N is not in $Y_0^{\mathbb{N}}(Y_0 - 1728)^{\mathbb{N}} Y_1^{\mathbb{N}} \mathbb{C}[Z]$. Recall that p_N is $\mathbf{j}$-homogeneous by Proposition 7.5. Since $h_{n+N} = C^n p_N(\frac{A}{C}, \mathbf{Y})$, if p_N depends on Y_2, so does $h_{n+N}(\frac{a}{c}, \mathbf{Y})$ for all but finitely many values of $\frac{a}{c}$. If p_N does not depend on Y_2, then $p_N = Y_1^N h(\frac{A}{C}, Y_0)$ for some polynomial $h \in \mathbb{C}[Z, Y_0]$, and by assumption, $h(\frac{A}{C}, Y_0)$ has at least one root distinct from 0 and 1728, seen as a polynomial in Y_0 (in an algebraic closure of $\mathbb{C}(\frac{A}{C})$), in which case so does $h(\frac{a}{c}, Y_0)$ except for finitely many values of $\frac{a}{c}$.

In either case, since the map $\gamma \mapsto \frac{a}{c}$ from $\mathrm{SL}_2(\mathbb{C})$ to $\mathbb{C}$ is dominant, we get that, for all γ except on some proper Zariski closed subset of $\mathrm{SL}_2(\mathbb{Z})$, the polynomial

$$h_{n+N}(a,b,c,d,\mathbf{Y}) = c^n p_N\Big(\frac{a}{c}, \mathbf{Y}\Big)$$

has an irreducible factor $r \in \mathbb{C}[\mathbf{Y}]$ such that $r(\mathbf{j}(z)) = 0$ has a Zariski dense set of solutions (by respectively Theorem 7.3, after noticing that the factors of a $\mathbf{j}$-homogeneous polynomial are $\mathbf{j}$-homogeneous, and Corollary 6.3).

Suppose that F is in $\mathbb{C}[X, Y_0]$. By irreducibility of F, we have that $F = F(X, Y_0)$ is neither divisible by Y_0 nor $Y_0 - 1728$. In particular, $p_N = \Gamma(F) = F(Z, Y_0)$ is also neither divisible by Y_0 nor $Y_0 - 1728$, which puts us back in the previous case.

Suppose that p_N is in $Y_0^{\mathbb{N}}(Y_0 - 1728)^{\mathbb{N}} Y_1^{\mathbb{N}} \mathbb{C}[Z]$ and F is not in $\mathbb{C}[X, Y_0]$. Write $p_N = r(Z) Y_0^s (Y_0 - 1728)^t Y_1^{\ell}$. Since F is irreducible, F is also not in $Y_1^{\mathbb{N}}\mathbb{C}[X, Y_0]$; thus, by Corollary 7.8,

$$2\ell = 2\deg_{Y_1}(p_N) = \deg_{\mathbf{j}}(p_N) = \deg_{\mathbf{j}}(F) > 2\deg_{Y_1}(p_{\mathrm{ord}_{\mathbf{j}}(F)}).$$

Since the map $\gamma \mapsto (\frac{a}{c}, \frac{b}{d}, c)$ from $\mathrm{SL}_2(\mathbb{C})$ to $\mathbb{C}^3$ is dominant, for γ in some Zariski open dense subset of $\mathrm{SL}_2(\mathbb{Z})$, we have

$$\deg_{Y_1}\Big(p_N\Big(\frac{a}{c}, c\Big)\Big) = \deg_{Y_1}(p_N) > \deg_{Y_1}(p_{\mathrm{ord}_{\mathbf{j}}(F)}) = \deg_{Y_1}\Big(p_{\mathrm{ord}_{\mathbf{j}}(F)}\Big(\frac{b}{d}, c, \mathbf{Y}\Big)\Big),$$

in which case the function

$$\frac{p_{\mathrm{ord}_{\mathbf{j}}(F)}(\frac{b}{d}, c, \mathbf{j}(z))}{p_N(\frac{a}{c}, \mathbf{j}(z))}$$

has a pole at some $\tau \in \mathrm{SL}_2(\mathbb{Z})\rho \cup \mathrm{SL}_2(\mathbb{Z})i$ or exponential growth in some fundamental domain by Corollary 6.8 (recall that, by Proposition 7.7, we know that $p_{\mathrm{ord}_{\mathbf{j}}(F)}$ does not depend on Y_2). If we fix some $\frac{b}{d}$, c as above, and a corresponding pole τ or a fundamental domain $\eta\mathbb{F}$ with exponential growth, then the function

$$\frac{h_0(\frac{b}{d}, c, \mathbf{j}(z))}{h_{n+N}(\frac{a}{c}, \mathbf{j}(z))} = \frac{d^n}{c^n r(\frac{a}{c})} \sum_{k=\mathrm{ord}_{\mathbf{j}}(F)}^{N} \frac{p_k(\frac{b}{d}, c, \mathbf{j}(z))}{j(z)^s (j(z)-1728)^t j'(z)^\ell} d^k$$

has neither a pole at τ nor exponential growth in $\eta\mathbb{F}$ only when d satisfies a non-trivial polynomial equation over $\frac{b}{d}$, c. Since $\gamma \mapsto (\frac{b}{d}, c, d)$ is also a dominant map, for γ in a Zariski open dense subset of $\mathrm{SL}_2(\mathbb{Z})$, the above function has a pole at τ or exponential growth in $\eta\mathbb{F}$, as claimed. □

7.3. Two more examples.

Example 7.9. Let us apply Theorem 1.2 to get some information on the zeroes of the function j'''. From the differential equation of the j-function (2.2), we see that

$$j''' = \frac{3}{2} \cdot \frac{(j'')^2}{j'} - \frac{j^2 - 1968j + 2654208}{2j^2(j-1728)^2}(j')^3.$$

By Theorem 1.2, the equation

$$3j^2(j-1728)^2(j'')^2 - (j^2 - 1968j + 2654208)(j')^4 = 0 \tag{7.2}$$

has a Zariski dense set of solutions outside $\mathrm{SL}_2(\mathbb{Z})\rho \cup \mathrm{SL}_2(\mathbb{Z})i$ (this actually follows directly from Theorem 7.3, because the equation is $\mathbf{j}$-homogeneous). Then these are also solutions of $j'''(z) = 0$.

Upon applying the transformation $z \mapsto -\frac{1}{z}$, we get

$$\begin{aligned} z^8\big(3j(z)^2(j(z)-1728)^2 j''(z)^2 - (j(z)^2 - 1968j(z) + 2654208) j'(z)^4\big) \\ + z^7 12 j(z)^2 (j(z)-1728)^2 j'(z) j''(z) \\ + z^6 12 j(z)^2 (j(z)-1728)^2 j'(z)^2 = 0. \end{aligned} \tag{7.3}$$

We can see that the ratios

$$\frac{12j^2(j-1728)^2 j' j''}{3j^2(j-1728)^2(j'')^2 - (j^2-1968j+2654208)(j')^4}, \quad \frac{12j^2(j-1728)^2(j')^2}{3j^2(j-1728)^2(j'')^2 - (j^2-1968j+2654208)(j')^4}$$

are equal to 0 at i and ρ and do not have exponential growth in any fundamental domain. However, we know that (7.2) has a zero $\tau \notin \mathrm{SL}_2(\mathbb{Z})\rho \cup \mathrm{SL}_2(\mathbb{Z})i$. Therefore, the second ratio above has a pole at τ, and so, for all large enough m, equation (7.3) has a zero near $\tau + m$. This means that (7.2), and hence $j''' = 0$, has solutions near $-\frac{1}{\tau+m}$ which accumulate at 0.

Example 7.10. Given $F(z, \mathbf{j}(z)) = 0$, our strategy of the proof of Theorem 1.2 is to apply a *generic* $\mathrm{SL}_2(\mathbb{Z})$-transformation and show that, in the function $F(\gamma z, \mathbf{j}(\gamma z))$, the ratio of the coefficients of the lowest and highest powers of z has a pole at some point in τ or has exponential growth at a cusp. In some cases, e.g. when $F(X, \mathbf{Y})$ does not depend on X, we can keep things simple and just use the good old transformation $z \mapsto -\frac{1}{z}$. Indeed, in this case, it turns out that the coefficient of the lowest power of z does not depend on j'' and so we can apply Corollary 6.8. For instance,

$$j''\Big(-\frac{1}{z}\Big)^2 = z^8 j''(z)^2 + 4z^7 j'(z) j''(z) + 4z^6 j'(z)^2.$$

We give an example to show that the transformation $-\frac{1}{z}$ and even $a - \frac{1}{z}$ for any integer a does not suffice in general (when F depends on X), as the aforementioned coefficient may depend on j''. Consider the function $f(z) = 2j'(z) + zj''(z)$. After a $z \mapsto a - \frac{1}{z}$ transformation, we get

$$\begin{aligned} f\Big(a - \frac{1}{z}\Big) &= 2z^2 j'(z) + \Big(a - \frac{1}{z}\Big)(z^4 j''(z) + 2z^3 j'(z)) \\ &= az^4 j''(z) + z^3(2aj'(z) - j''(z)). \end{aligned}$$

We see that the coefficient of z^3 depends on j'' regardless of the value of a.

Acknowledgement. We thank the referee for a thorough reading of the paper and for numerous suggestions that helped us improve the presentation.

References

[1] *V. Aslanyan*, Adequate predimension inequalities in differential fields, Ann. Pure Appl. Logic **173** (2022), no. 1, Article ID 103030.

[2] *V. Aslanyan*, *S. Eterović* and *J. Kirby*, Differential existential closedness for the j-function, Proc. Amer. Math. Soc. **149** (2021), no. 4, 1417–1429.

[3] *V. Aslanyan*, *S. Eterović* and *J. Kirby*, A closure operator respecting the modular j-function, Israel J. Math. **253** (2023), no. 1, 321–357.

[4] *V. Aslanyan* and *J. Kirby*, Blurrings of the J-function, Q. J. Math. **73** (2022), no. 2, 461–475.

[5] *V. Aslanyan*, *J. Kirby* and *V. Mantova*, A geometric approach to some systems of exponential equations, Int. Math. Res. Not. IMRN **2023** (2023), no. 5, 4046–4081.

[6] *W. D. Brownawell* and *D. W. Masser*, Zero estimates with moving targets, J. Lond. Math. Soc. (2) **95** (2017), no. 2, 441–454.

[7] *P. D'Aquino*, *A. Fornasiero* and *G. Terzo*, Generic solutions of equations with iterated exponentials, Trans. Amer. Math. Soc. **370** (2018), no. 2, 1393–1407.

[8] *S. Eterović*, Generic solutions of equations involving the modular j function, Math. Ann. **391** (2025), no. 4, 6401–6449.

[9] *S. Eterović* and *S. Herrero*, Solutions of equations involving the modular j function, Trans. Amer. Math. Soc. **374** (2021), no. 6, 3971–3998.

[10] *F. P. Gallinaro*, Solving systems of equations of raising-to-powers type, Israel J. Math. (2025), DOI 10.1007/s11856-025-2778-2.

[11] *F. P. Gallinaro*, Exponential sums equations and tropical geometry, Selecta Math. (N. S.) **29** (2023), no. 4, Paper No. 49.

[12] *J. Kirby*, Blurred complex exponentiation, Selecta Math. (N. S.) **25** (2019), no. 5, Paper No. 72.

[13] *S. Lang*, Elliptic functions, 2nd ed., Grad. Texts in Math. **112**, Springer, New York 1987.

[14] *S. Lang*, Complex analysis, 4th ed., Grad. Texts in Math. **103**, Springer, New York 1999.

[15] *K. Mahler*, On algebraic differential equations satisfied by automorphic functions, J. Aust. Math. Soc. **10** (1969), 445–450.

[16] *B. Zilber*, Exponential sums equations and the Schanuel conjecture, J. Lond. Math. Soc. (2) **65** (2002), no. 1, 27–44.
[17] *B. Zilber*, Pseudo-exponentiation on algebraically closed fields of characteristic zero, Ann. Pure Appl. Logic **132** (2005), no. 1, 67–95.
[18] *B. Zilber*, The theory of exponential sums, preprint 2015, https://arxiv.org/abs/1501.03297.

Vahagn Aslanyan, Department of Mathematics, University of Manchester,
Manchester, M13 9PL, United Kingdom
https://orcid.org/0000-0003-4523-6773
e-mail: vahagn.aslanyan@manchester.ac.uk

Sebastian Eterović, Kurt Gödel Research Center, Universität Wien, 1090 Wien, Austria
https://orcid.org/0000-0001-6724-5887
e-mail: sebastian.eterovic@univie.ac.at

Vincenzo Mantova, School of Mathematics, University of Leeds, Leeds, LS2 9JT, United Kingdom
https://orcid.org/0000-0002-8454-7315
e-mail: v.l.mantova@leeds.ac.uk

Eingegangen 15. Dezember 2023, in revidierter Fassung 31. Juli 2025