EQUAL COMPOSITIONS OF RATIONAL FUNCTIONS

KENZ KALLAL, MATTHEW LIPMAN, AND FELIX WANG

Abstract. We study the following questions:

(1) What are all solutions to $f \circ \hat{f} = g \circ \hat{g}$ in complex rational functions $f, g \in \mathbb{C}(X)$ and meromorphic functions $\hat{f}, \hat{g}$ on the complex plane?
(2) For which rational functions $f(X)$ and $g(X)$ with coefficients in an algebraic number field K does the equation $f(a) = g(b)$ have infinitely many solutions with $a, b \in K$?

We utilize various algebraic, geometric and analytic results in order to resolve both questions in the case that the numerator of $f(X) - g(Y)$ is an irreducible polynomial in $\mathbb{C}[X, Y]$ of sufficiently large degree. Our work answers a 1973 question of Fried in all but finitely many cases, and makes significant progress towards answering a 1924 question of Ritt and a 1997 question of Lyubich and Minsky.

Date: December 31, 2015.

1. Introduction

Throughout the history of number theory, many mathematicians have studied special cases of the following question:

Question 1.1. *For which rational functions $f, g \in \mathbb{Q}(X)$ does the equation $f(a) = g(b)$ have infinitely many solutions in rational numbers a and b?*

For example, Archimedes studied an instance of the Pell equation $a^2 = db^2 + 1$; we now know that there are infinitely many integers a, b satisfying this equation for any prescribed nonsquare positive integer d [20, p. 184]. More recently, Wiles [38] proved that Fermat's equation $c^n = d^n + e^n$ has no solutions in nonzero integers c, d, e, n with $n > 2$. Upon division by e^n, this result shows that $a^n = b^n + 1$ has no solutions in nonzero rational numbers a, b. Another prominent equation in modern number theory is the Weierstrass equation $Y^2 = X^3 + cX + d$, where c and d are fixed rational numbers such that $4c^3 \neq -27d^2$. This equation defines an elliptic curve, and has infinitely many solutions in rational numbers if and only if the corresponding elliptic curve has positive rank; this rank is the key quantity in the Birch and Swinnerton–Dyer conjecture, which is one of the most important open problems in mathematics [9]. This last example is enlightening, because although there has been progress on describing "how often" such an equation has infinitely many rational solutions [4–6], it seems that there is no hope of finding all pairs (c, d) for which the equation has infinitely many solutions. However, each equation of this form has infinitely many solutions in some algebraic number field K, by which we mean a field which is a finite-dimensional $\mathbb{Q}$-vector space. It is thus natural to modify Question 1.1 as follows:

Question 1.2. *For which rational functions $f, g \in K(X)$, where K is an algebraic number field, does the equation $f(a) = g(b)$ have infinitely many solutions in K?*

We prove the following result:

Theorem 1.3. *For any number field K and any rational functions $f, g \in K(X)$ such that the numerator of $f(X) - g(Y)$ is an irreducible polynomial in $\mathbb{C}[X, Y]$ of (total)* *degree at least* 10^5*, if the equation* $f(a) = g(b)$ *has infinitely many solutions in* K *then one of these holds:*

(1.3.1) *at least one of the extensions* $K(X)/K(f(X))$ *or* $K(X)/K(g(X))$ *has Galois closure of genus* 0 *or* 1*,*

(1.3.2) $f = \mu \circ X^c(X-1)^d \circ \nu_1$ *and* $g = \mu \circ \gamma X^c(X-1)^d \circ \nu_2$ *for some coprime positive integers* c, d*, some* $\gamma \in K \setminus \{0, 1\}$*, and some degree-one* $\mu, \nu_1, \nu_2 \in K(X)$*.*

Conversely, if (1.3.2) holds then $f(X) - g(Y)$ is irreducible in $\mathbb{C}[X, Y]$ and $f(a) = g(b)$ has infinitely many solutions in K. These conclusions are sometimes satisfied when (1.3.1) holds, but not always. However, for each $f(X) \in K(X)$ such that $K(X)/K(f(X))$ has Galois closure of genus 0 or 1, there exist rational functions $g(X) \in \hat{K}(X)$ of arbitrarily large degree (with coefficients in a number field $\hat{K}$ containing K) for which these conclusions are satisfied over $\hat{K}$.

Since automorphism groups of function fields of genus 0 or 1 are well-understood, condition (1.3.1) lets us give a precise description of either f or g. For instance, if the Galois closure of $K(X)/K(f(X))$ has genus 0 and $\deg(f) > 60$ then $f(X)$ is either X^m or $X^m + X^{-m}$ or a Chebyshev polynomial $T_m(X)$, up to composition on both sides with degree-one rational functions. Furthermore, when (1.3.1) holds we describe both $f(X)$ and $g(X)$: for instance, if $f(X) = X^m$ with $m > 6$ then there is some degree-one $\nu \in K(X)$ for which $g \circ \nu$ is $X^c h(X)^m$ with $h \in K(X)$ and c coprime to m. For more results when (1.3.1) holds, see Theorem 4.1.

Most of the previous work on Question 1.1 addresses the much easier problem of determining the polynomials $f, g \in \mathbb{Z}[X]$ for which $f(a) = g(b)$ has infinitely many solutions in integers a, b. This was solved by Bilu and Tichy [7], building on previous work by Davenport, Fried, Lewis, Schinzel, Siegel, and others [12,15,36,37]. It is easy to reduce this question to the case that $f(X) - g(Y)$ is irreducible in $\mathbb{C}[X, Y]$. Question 1.1 for rational solutions has also been studied by several authors. The most general published result was proved by Avanzi and Zannier [2], and addresses the case that f and g are polynomials of coprime degrees. Very recently, Carney et al. have extended this to the case of arbitrary polynomials f and g [10, 11]. Our further extension to rational functions (under some hypotheses) requires completely different methods than were used previously.

The second main topic of this paper is functional equations, and specifically the following questions:

Question 1.4. *What are all solutions to $f \circ \hat{f} = g \circ \hat{g}$ in rational functions $f, \hat{f}, g, \hat{g} \in \mathbb{C}(X)$?*

Question 1.5. *What are all solutions to $f \circ \hat{f} = g \circ \hat{g}$ in rational functions $f, g \in \mathbb{C}(X)$ and meromorphic functions $\hat{f}, \hat{g}$ on the complex plane?*

Here a *meromorphic function* is a ratio h_1/h_2 where h_1, h_2 are entire functions with $h_2 \neq 0$, and an *entire function* is a function $\mathbb{C} \to \mathbb{C}$ defined by a single power series $\sum_{i=0}^{\infty} \alpha_i X^i$ with infinite radius of convergence. For instance, e^X is entire, as are all polynomials, and all rational functions are meromorphic. Hence Question 1.4 is a more restricted version of Question 1.5.

We prove the following result:

Theorem 1.6. *For any $f, g \in \mathbb{C}(X)$ such that the numerator of $f(X) - g(Y)$ is an irreducible polynomial in $\mathbb{C}[X, Y]$ of degree at least 10^5, if there are nonconstant meromorphic functions $\hat{f}, \hat{g}$ on the complex plane such that $f \circ \hat{f} = g \circ \hat{g}$ then one of these holds:*

(1.6.1) *at least one of the extensions $\mathbb{C}(X)/\mathbb{C}(f(X))$ or $\mathbb{C}(X)/\mathbb{C}(g(X))$ has Galois closure of genus 0 or 1*

(1.6.2) *$f = \mu \circ X^c(X-1)^d \circ \nu_1$ and $g = \mu \circ \gamma X^c(X-1)^d \circ \nu_2$ for some coprime positive integers c, d, some $\gamma \in K \setminus \{0, 1\}$, and some degree-one $\mu, \nu_1, \nu_2 \in \mathbb{C}(X)$.*

Conversely, if (1.6.2) holds then the meromorphic functions $\hat{f}, \hat{g}$ satisfying $f \circ \hat{f} = g \circ \hat{g}$ are given by

$$\hat{f} = \nu_1^{-1} \circ \frac{\gamma^b X^c - 1}{\gamma^{a+b} X^{c+d} - 1} \circ h \quad \text{and} \quad \hat{g} = \nu_2^{-1} \circ \frac{\gamma^{a+b} X^{c+d} - \gamma^a X^d}{\gamma^{a+b} X^{c+d} - 1} \circ h$$

where h is meromorphic and a, b are integers such that $bd - ac = 1$. Also, to some extent we can describe all choices for f, g, $\hat{f}$ and $\hat{g}$ when (1.6.1) holds.

Questions 1.4 and 1.5 are of interest for several reasons. First, Nevanlinna showed that if nonconstant meromorphic functions $\hat{f}, \hat{g}$ satisfy $\hat{f}^{-1}(\alpha) = \hat{g}^{-1}(\alpha)$ for five distinct values of $\alpha \in \mathbb{C}$, then we must have $\hat{f} = \hat{g}$ [26]. Subsequent authors have sought analogous results when the values α are replaced by finite sets of complex numbers, and more generally when there are several pairs of finite sets (S_i, T_i) such that $\hat{f}^{-1}(S_i) = \hat{g}^{-1}(T_i)$. If there are nonconstant rational functions f, g for which $f \circ \hat{f} = g \circ \hat{g}$, then $\hat{f}^{-1}(f^{-1}(U)) = \hat{g}^{-1}(g^{-1}(U))$ for any $U \subset \mathbb{C}$, so in this case there are infinitely many pairs (S_i, T_i) of finite subsets of $\mathbb{C}$ for which $\hat{f}^{-1}(S_i) = \hat{g}^{-1}(T_i)$. Conversely, it is conceivable that such an infinitude of pairs (S_i, T_i) only exists when there exist such rational functions f, g. Thus Question 1.5 is a fundamental question about the distribution of preimages of meromorphic functions. We note that quite special cases of Question 1.5 have themselves been major results, for instance the case that f, g are polynomials and b, d are entire [30]. Furthermore, Theorem 1.6 answers a question of Fried [15, Problem 1] when $\max(\deg f, \deg g)$ is sufficiently large, thereby reducing Fried's question to a finite (albeit lengthy) computation. Question 1.4 was originally studied by Ritt [35]; Theorem 1.6 comprises significant progress towards a solution of both Ritt's question and a question of Lyubich and Minsky [22, p. 83] on laminations in holomorphic dynamics.

In the special case that $f, \hat{f}, g, \hat{g}$ are polynomials, Question 1.4 was solved by Ritt [34]. His result has been used to prove important theorems in algebra [39], algebraic geometry [23], differential equations [8, 32], dynamical systems [3, 17, 18], logic [23], topology [28], transcendental number theory [27], and other topics. Solutions to Questions 1.4 or 1.5 would yield vast improvements to all of these theorems. Prior to our work, these polynomial results had been extended only slightly, to cases of Question 1.4 which were not too far from the polynomial case; however, we note that already such extensions required significant effort [29, 31, 40]. Our Theorem 1.6 goes far beyond all these previous results.

This paper is organized as follows. In the next section we show that Theorems 1.3 and 1.6 are consequences of another result (Theorem 2.3), and present several important tools. We use these tools in Section 3 in order to prove Theorem 2.3, and then in Section 4 we refine the conclusions of these three theorems. We conclude in Section 5 with a discussion of future avenues of research.

2. Ramification and genus

In this section we show that the number-theoretic Theorem 1.3 and the analytic Theorem 1.6 are both consequences of a single geometric theorem, and then present several tools we will use to prove this theorem. We begin with some notation.

Definition 2.1 (Ramification Index). *The ramification index $e_f(P)$ of a rational function $f(X)$ at a point $P \in \mathbb{C} \cup \{\infty\}$ is the local degree of $f(X)$ near $X = P$. Concretely, if $P, f(P) \in \mathbb{C}$ then $e_f(P)$ is the multiplicity of $X = P$ as a root of $f(X) - f(P)$, and in other cases $e_f(P)$ can be defined by changing variables to reduce to this case.*

Definition 2.2 (Ramification Multiset). *The ramification multiset $E_f(Q)$ of a rational function f at a point Q is the multiset of all values of $e_f(P)$ for $P \in f^{-1}(Q)$.*

We can now state our main geometric result. Here and elsewhere, the expression $[a^c, b^d, \ldots]$ denotes the multiset containing c copies of a, d copies of b, and so on. Also, by the genus of a plane curve we mean the genus of the corresponding function field.

Theorem 2.3 (LCM Theorem). *Let $f, g \in \mathbb{C}(X)$ have degrees $m, n > 0$, with $n \geq m > 1176$ or $n \geq 84m$. Let $Q_1, \ldots, Q_r$ be the points in $\mathbb{C} \cup \{\infty\}$ for which either $E_f(Q_i) \neq [1^m]$ or $E_g(Q_i) \neq [1^n]$. If the numerator of $f(X) - g(Y)$ defines an irreducible curve of genus 0 or 1, then $F_i := E_f(Q_i)$ and $G_i := E_g(Q_i)$ satisfy one of the following:*

(2.3.1) $\sum_{i=1}^{r}(1 - \frac{1}{\operatorname{lcm}(F_i)}) \leq 2$

(2.3.2) $\sum_{i=1}^{r}(1 - \frac{1}{\operatorname{lcm}(G_i)}) \leq 2$

(2.3.3) *$m = n$, $r = 4$, and (after relabeling the Q_i's) we have $F_1 = G_1 = [m]$, $F_2 = G_2 = [c, m - c]$ for some c coprime to m, $F_3 = G_4 = [1^{m-2}, 2]$, and $F_4 = G_3 = [1^m]$.*

Proof that Theorem 2.3 implies Theorems 1.3 and 1.6. By theorems of Faltings [13] and Picard [33], if the hypotheses of either Theorem 1.3 or Theorem 1.6 hold then the numerator of $f(X)-g(Y)$ defines a curve of genus 0 or 1. The hypotheses of either theorem imply that $m:=\deg(f)$ and $n:=\deg(g)$ satisfy $m+n\ge 10^5$, so that either $m\ge 84n$ or $n\ge 84m$ or $m,n\ge \dfrac{10^5}{85}>1176$. By interchanging f and g if necessary (which does not affect the truth of the conclusions of Theorems 1.3 and 1.6), we may assume that $n\ge m$, so that the hypotheses of Theorem 2.3 are satisfied. Hence Theorem 2.3 implies that one of (2.3.1)–(2.3.3) holds. Let $\mathcal{N}$ be the Galois closure of $\mathbb{C}(X)/\mathbb{C}(f(X))$, and let d be the degree of the extension $\mathcal{N}/\mathbb{C}(f(X))$. Then Q_i lies under $\frac{d}{\operatorname{lcm}(F_i)}$ points of $\mathcal{N}$, each of which has ramification index $\operatorname{lcm}(F_i)$ in $\mathcal{N}/\mathbb{C}(f(X))$. Thus, by the Riemann–Hurwitz formula, if (2.3.1) or (2.3.2) holds then (1.3.1) and (1.6.1) hold. Finally, suppose that (2.3.3) holds. Upon replacing f and g by $f\circ\nu_1$ and $g\circ\nu_2$ for suitable degree-one $\nu_i\in\mathbb{C}(X)$, we may assume that $f(\infty)=Q_1=g(\infty)$ and $f^{-1}(Q_2)=\{0,1\}=g^{-1}(Q_2)$ where $e_f(0)=c=e_g(0)$. Upon replacing f and g by $\mu\circ f$ and $\mu\circ g$ for a suitable degree-one $\mu\in\mathbb{C}(X)$, we may assume that $Q_1=\infty$, $Q_2=0$, and the numerator and denominator of f have the same leading coefficient. It follows that $f=X^c(X-1)^{m-c}$ and $g=\gamma X^c(X-1)^{m-c}$ for some $\gamma\in\mathbb{C}^*$, and the reducibility hypothesis ensures that $\gamma\ne 1$. Hence the original f and g satisfy (1.3.2) and (1.6.2). □

Our proof of Theorem 2.3 proceeds by showing that if f,g satisfy the hypotheses of Theorem 2.3 then the multisets $F_i:=E_f(Q_i)$ and $G_i:=E_g(Q_i)$ satisfy several numerical conditions, and then solving the combinatorial problem of determining all collections of multisets of positive integers which satisfy these conditions. We present these numerical conditions in the remainder of this section, and then prove Theorem 2.3 in the next section. We conclude in Section 4 by determining f and g when (2.3.1) or (2.3.2) holds.

The first two numerical conditions satisfied by ramification multisets are

$$\sum_{P\in f^{-1}(Q)} e_f(P)=\deg(f)\qquad\text{for each } Q\in\mathbb{C}\cup\{\infty\} \tag{2.4}$$

$$\sum_{Q\in\mathbb{C}\cup\{\infty\}} \big(\deg(f)-|E_f(Q)|\big)=2\deg(f)-2. \tag{2.5}$$

Equation (2.5) is the Riemann–Hurwitz genus formula for the extension of function fields $\mathbb{C}(X)/\mathbb{C}(f(X))$. If $f, g \in \mathbb{C}(X)$ have degrees $m, n > 0$, and the numerator of $f(X) - g(Y)$ is irreducible, then this numerator defines a curve of genus $\mathfrak{g}$ where

$$2\mathfrak{g} - 2 = -2m + \sum_{Q \in \mathbb{C} \cup \{\infty\}} \sum_{a \in E_f(Q)} \sum_{b \in E_g(Q)} \big(a - \gcd(a, b)\big). \tag{2.6}$$

Equation (2.6) is a version of the Riemann–Hurwitz genus formula for the function field extension $\mathbb{C}(X, Y)/\mathbb{C}(Y)$ (where $f(X) = g(Y)$), and was proved by Ritt [34]. In particular, if $\mathfrak{g} \in \{0, 1\}$ then

$$\sum_{Q \in \mathbb{C} \cup \{\infty\}} \sum_{a \in E_f(Q)} \sum_{b \in E_g(Q)} \big(a - \gcd(a, b)\big) \in \{2m - 2, 2m\}. \tag{2.7}$$

The next two lemmas give new types of constraints on the F_i's and G_i's which are crucial for our work.

Lemma 2.8. *If all elements of $F_1 \cup F_2$ are even then, for each $i > 2$, the multiset F_i can be written as the union of two submultisets each having sum $\frac{m}{2}$.*

Proof. Upon replacing f by $\mu \circ f$ for some degree-one $\mu \in \mathbb{C}(X)$, we may assume that $Q_1 = 0$ and $Q_2 = \infty$, so by hypothesis $f(X) = h(X)^2$ for some $h \in \mathbb{C}(X)$. Then for $i > 2$ we have $E_f(Q_i) = E_h(\sqrt{Q_i}) \cup E_h(-\sqrt{Q_i})$, which implies the result by (2.4). □

Lemma 2.9. *If the numerator of $f(X) - g(Y)$ is irreducible then both of these hold:*

(2.9.1) *For any distinct i, j we have* $\gcd(F_i \cup F_j \cup G_i \cup G_j) = 1$.

(2.9.2) *For any distinct i, j, k such that $F_i \cup F_j$ and $G_i \cup G_j$ each contain at most two odd indices, we must have* $\gcd(F_k \cup G_k) \le 2$.

Proof. We prove the contrapositive. If (2.9.1) fails then, by replacing f and g by $\mu \circ f$ and $\mu \circ g$ for some degree-one $\mu \in \mathbb{C}(X)$, we may assume that $Q_i = 0$ and $Q_j = \infty$. Since $d := \gcd(F_i \cup F_j \cup G_i \cup G_j)$ divides $\gcd(F_i, F_j)$, we can write $f = X^d \circ \hat{f}$ for some $\hat{f} \in \mathbb{C}(X)$, and likewise $g = X^d \circ \hat{g}$. Therefore $f(X) - g(Y) = \prod_{\zeta^d=1} \big(\hat{f}(X) - \zeta \hat{g}(Y)\big)$ is reducible.

Henceforth suppose that (2.9.2) fails. Again we may assume $Q_i = -1$, $Q_j = 1$ and $Q_k = \infty$. First suppose there is an odd prime p which divides $\gcd(F_k \cup G_k)$. Then the degree-p Chebyshev polynomial $T_p(X)$ satisfies $E_{T_p}(Q_i) = E_{T_p}(Q_j) = [1, 2^{(p-1)/2}]$, $E_{T_p}(Q_k) = [p]$, and $E_{T_p}(Q_\ell) = [1^p]$ for $\ell \notin \{i, j, k\}$. Hence

$$\sum_{S\in\mathbb{C}\cup\infty} \; \sum_{a\in E_{T_p}(S)} \; \sum_{b\in E_a(S)} \big(a - \gcd(a,b)\big) \le p-1 < 2p-2,$$

so by (2.6) the numerator of $T_p(X) - f(Y)$ must be reducible, since otherwise it would define a curve having negative genus. Then [14, Prop. 2] implies that $f = f_1 \circ f_2$ for some $f_1, f_2 \in \mathbb{C}(X)$ such that the numerators of $T_p(X) - z$ and $f_1(X) - z$ have the same splitting field as one another over $\mathbb{C}(z)$, where z is transcendental over $\mathbb{C}$. Since the splitting field $\mathcal{M}$ of $T_p(X) - z$ over $\mathbb{C}(z)$ is $\mathbb{C}(y)$ where $y^p + y^{-p} = 2z$, the Galois group of $\mathcal{M}/\mathbb{C}(z)$ is dihedral of order $2p$, so that each non-Galois extension of $\mathbb{C}(z)$ contained in $\mathcal{M}$ has the form $\mathbb{C}(x)$ where $T_p(x) = z$. Hence $f_1 = T_p \circ h$ for some $h \in \mathbb{C}(X)$, so $f = T_p \circ \hat{f}_2$ where $\hat{f}_2 := h \circ f_2$. Likewise $g = T_p \circ g_2$, so that $f(X) - g(Y)$ equals $T_p(f_2(X)) - T_p(g_2(Y))$, whose numerator is reducible since it is divisible by the numerator of $f_2(X) - g_2(Y)$.

The proof is similar but lengthier when $\gcd(F_k \cup G_k)$ is a power of 2, so we just sketch the argument. The main difference is that the ramification of $T_2(X)$ is slightly different from that of $T_p(X)$ for odd p, so that the above argument does not imply that the numerator of $T_2(X) - f(Y)$ is reducible. However, the above argument does imply that the numerator of either $T_2(X) - f(Y)$ or $T_2(X) + f(Y)$ is reducible, so $f = \pm T_2 \circ f_2$. Similarly, $f = \pm T_4 \circ f_2$ and $g = \pm T_4 \circ g_2$, and since both $T_4(X) - T_4(Y)$ and $T_4(X) + T_4(Y)$ are reducible it follows that the numerator of $f(X) - g(Y)$ is reducible. □

3. Proof of LCM Theorem

In this section we prove Theorem 2.3. We first give a quick proof in case $n \geq 84m$.

Proposition 3.1 (Fixed m, Large n)**.** *Using the notation and assumptions of Theorem 2.3, if $n \geq 84m$ then (2.3.1) holds.*

Proof. For each i we write D_i for the multiset of elements of G_i which are not divisible by $\operatorname{lcm}(F_i)$, and let $d_i := |D_i|$ and $c_i := \sum_{b\in D_i} b$. Then

$$|G_i| = d_i + \sum_{\substack{b\in G_i\\ \operatorname{lcm}(F_i)\mid b}} 1 \le d_i + \frac{n-c_i}{\operatorname{lcm}(F_i)} \le d_i + \frac{n-d_i}{\operatorname{lcm}(F_i)}.$$

By (2.7), we have

$$2m \ge \sum_{i=1}^{r} \sum_{b\in S_i} \sum_{a\in F_i} \bigl(a - \gcd(a,b)\bigr) \ge \sum_{i=1}^{r} \sum_{b\in S_i} 1 = \sum_{i=1}^{r} d_i.$$

Letting ℓ be the largest of all the values $\operatorname{lcm}(F_i)$ for $1 \le i \le r$, it follows that

$$\sum_{i=1}^{r} |G_i| \le \sum_{i=1}^{r} \left(\frac{n}{\operatorname{lcm}(F_i)} + d_i \cdot \left(1 - \frac{1}{\operatorname{lcm}(F_i)}\right)\right) \le \sum_{i=1}^{r} \frac{n}{\operatorname{lcm}(F_i)} + 2m \cdot \left(1 - \frac{1}{\ell}\right).$$

Substituting this into (2.5) yields

$$2n - 2 = \sum_{i=1}^{r} (n - |G_i|) \ge \sum_{i=1}^{r} \left(n - \frac{n}{\operatorname{lcm}(F_i)}\right) + \frac{2m}{\ell} - 2m,$$

so that

$$n\left(\sum_{i=1}^{r}\left(1 - \frac{1}{\operatorname{lcm}(F_i)}\right) - 2\right) \le 2m - \frac{2m}{\ell} - 2 < 2m \le \frac{n}{42}.$$

Thus $\sum_{i=1}^{r}(1 - \frac{1}{\operatorname{lcm}(F_i)})$ is less than $2 + \frac{1}{42}$, which implies by Lemma 3.2 that in fact the sum is at most 2, whence (2.3.1) holds. □

Lemma 3.2. *If $d_1, \dots, d_q$ is a finite sequence of integers greater than 1, then $S := \sum_{i=1}^{q}\bigl(1 - \frac{1}{d_i}\bigr)$ lies in $\{0\} \cup [\frac{1}{2}, 1] \cup [\frac{7}{6}, 2] \cup [2 + \frac{1}{42}, \infty)$. Furthermore, we have $S \le 2$ if and only if either $q \le 2$ or the multiset of d_i's is one of the following: $[2^4]$, $[3^3]$, $[2, 4^2]$, $[2, 3, \ell]$ with $2 \le \ell \le 6$, or $[2^2, k]$ with $k > 1$.*

Proof. Write D for the multiset of d_i's. Note that $S = 2$ when D is $[2^4]$, $[3^3]$, $[2, 4^2]$, or $[2, 3, 6]$. Since the value of S becomes strictly larger if we either append a 2 to D or increase some element of D by 1, and by starting with each of the above four D's and repeatedly applying these operations we obtain every D with $q > 2$ except $[2, 3, \ell]$ with $\ell < 6$ and $[2^2, k]$ with $k > 1$, this implies the last assertion in the result. Moreover, the smallest value of S larger than 2 must occur when D arises from a single such operation, so the smallest such S is $2 + \frac{1}{42}$ which occurs for $D = [2, 3, 7]$. Likewise, if $D = \emptyset$ or

$D = [2^2]$ then $S = 0$ or $S = 1$, so by the same argument the smallest values of S greater than 0 or 1 occur when $D = [2]$ or $D = [2,3]$, respectively, and are $S = \frac{1}{2}$ and $\frac{7}{6}$. □

It remains to prove Theorem 2.3 when $1176 < m \le n \le 84m$, so we assume these inequalities for the rest of this section. Our next result provides a crucial constraint on the multisets F_i and G_i.

Proposition 3.3. *Suppose that f and g satisfy the hypotheses of Theorem 2.3, and also $1176 < m \le n \le 84m$. For any i, put $F := E_f(Q_i)$ and $G := E_g(Q_i)$, and let f_a and g_a be the numbers of copies of the integer a in F and G, respectively. For any integer c such that $0 \le c \le 6$, one of the following holds:*

(3.3.1) *There is a positive integer $d \le c$ such that $f_d > \frac{m}{d} - (2d+3)$.*
(3.3.2) *$f_a, g_a \le 4$ for $1 \le a \le c$.*

Proof. We prove Proposition 3.3 by induction on c. The base case is $c = 0$, where (3.3.2) is vacuously true. For the inductive step it is enough to prove that if $f_a, g_a \le 4$ for $1 \le a \le c-1$ then either $f_c > \frac{m}{c} - (2c+3)$ or $f_c, g_c \le 4$. By condition (2.7), we have

$$2m \ge g_c \sum_{a=1}^{\infty} f_a \cdot \Big(a - \gcd(a,c)\Big) \ge g_c \sum_{a=c+1}^{\infty} f_a \cdot \frac{a}{2} \ge g_c \cdot \frac{1}{2}\Big(m - cf_c - \sum_{a=1}^{c-1} af_a\Big),$$

where we used the facts that $m = \sum_a af_a$ and if $a > c$ then $\gcd(a,c) \le \frac{a}{2}$. The above inequality then implies that

$$(3.4)\quad 4m \ge g_c\Big(m - cf_c - \sum_{a=1}^{c-1} af_a\Big) \ge g_c\Big(m - cf_c - \sum_{a=1}^{c-1} 4a\Big) \ge g_c\Big(m - cf_c - 2c(c-1)\Big).$$

Similarly,

$$(3.5)\quad 4n \ge f_c\Big(n - cg_c - \sum_{a=1}^{c-1} ag_a\Big) \ge f_c\Big(n - cg_c - \sum_{a=1}^{c-1} 4a\Big) \ge f_c\Big(n - cg_c - 2c(c-1)\Big).$$

Assume that $5 \le f_c \le \frac{m}{c} - (2c+3)$; we now show that this leads to a contradiction. Here $m - cf_c - 2c(c-1) > 0$ and $f_c > 0$, so we may combine (3.4) and (3.5) to get

$$\frac{4m}{m - cf_c - 2c(c-1)} \ge g_c \ge \frac{1}{c}\Big(n - 2c(c-1) - \frac{4n}{f_c}\Big).$$

By clearing denominators we obtain $h(f_c) \geq 0$, where $h(X)$ is the polynomial

$$cX^2\Big(n-2c(c-1)\Big)+X\left(4mc-4nc-\Big(m-2c(c-1)\Big)\Big(n-2c(c-1)\Big)\right)+4n\Big(m-2c(c-1)\Big).$$

It is easy to check that $h(X)$ is negative when X is either 5 or $\frac{m}{c}-(2c+3)$. Since $h(X)$ has degree at most 2, and the coefficient of X^2 in $h(X)$ is nonnegative, it follows that $h(X)$ is negative for all X with $5 \leq X \leq \frac{m}{c}-(2c+3)$. This yields the contradiction $h(f_c) < 0$, so our assumption was incorrect and thus either $f_c \leq 4$ or $f_c > \frac{m}{c}-(2c+3)$.

If $f_c > \frac{m}{c}-(2c+3)$ then we are done. If $f_c \leq 4$ then (3.4) implies that

$$4m \geq g_c\Big(m - cf_c - 2c(c-1)\Big) \geq g_c\big(m-4c-2c(c-1)\big) = g_c\big(m-2c(c+1)\big);$$

hence $g_c \leq \frac{4m}{m-2c(c+1)} < 5$, which completes the proof. □

We can improve Proposition 3.3 by strengthening the inequalities used in its proof. In particular, we can replace (3.4) by

$$4m \geq \sum_{j=1}^{c} g_j\Big(m - jf_j - \sum_{a=1}^{j-1} af_a\Big), \tag{3.6}$$

we can make a similar improvement to (3.5), and also for each fixed c we can improve the inequality $\gcd(a,c) \leq \frac{a}{2}$ by using the actual value if $a \leq 2c$ and otherwise using the bound $\gcd(a,c) \leq c$. Applying these improvements requires the separate treatment of a large number of cases, depending on the values of f_a and g_a for several choices of a, and was done with the assistance of a computer program. This yields the following result.

Proposition 3.7. *Under the hypotheses of Proposition 3.3, if c is an integer with $1 \leq c \leq 6$ then one of the following holds:*

(3.7.1) *There is a positive integer $d \leq c$ such that $f_d > \frac{m-w_d}{d}$, where $w_1 = 5$, $w_2 = 12$, $w_3 = 15$, $w_4 = w_5 = 24$, and $w_6 = 36$*

(3.7.2) *$\sum_{a\leq c} f_a \leq 4$ and $\sum_{a\leq c} g_a \leq 4$.*

Propositions 3.3 and 3.7 show that, for each i, either there is some (necessarily unique) integer d_i with $1 \leq d_i \leq 6$ for which the sum of the elements of F_i different from d_i is bounded by an absolute constant, or else F_i contains a bounded number of elements smaller than 7 (in which case we define $d_i := \infty$). In our proof of Theorem 2.3, we combine this information across all the different points Q_i in order to determine the possibilities for the multiset D consisting of all d_i's greater than 1. We first give a heuristic argument illustrating our approach. If $d_i \le 6$ then $|F_i| \approx \frac{m}{d_i}$, and if $d_i = \infty$ then $|F_i|$ is at most $\frac{m}{7}+c$ for some small constant c. By (2.5), we have $2m-2 = \sum_{i=1}^{r}(m-|F_i|)$, so that

$$2m - \sum_{i\colon d_i\le 6} m\Big(1-\frac{1}{d_i}\Big) \approx \sum_{i\colon d_i=\infty} (m-|F_i|), \tag{3.8}$$

where each summand on the right side is between $\frac{6m}{7} - c$ and m. By Lemma 3.2, the quantity $\sum_{i\colon d_i\le 6}(1-\frac{1}{d_i})$ is either 0 or an element of one of the intervals $[\frac{1}{2}, 1]$ or $[\frac{7}{6}, \infty)$, so the left side of (3.8) is either $2m$ or an element of $[m, \frac{3m}{2}]$ or $(-\infty, \frac{5m}{6}]$. Since the right side of (3.8) is a sum of elements of $[\frac{6m}{7} - c, m]$, the only possibility is that each summand on the right side is approximately m, whence $\sum_{d\in D}(1-\frac{1}{d}) = 2$. This equation implies that D is one of the multisets

$$[2,2,2,2],\ [2,4,4],\ [3,3,3],\ [2,3,6],\ [2,2,\infty],\ [\infty,\infty].$$

Below we prove Theorem 2.3 via a rigorous version of this heuristic argument, first restricting the possibilities for the F_i's and then deducing the desired conclusion.

In what follows, we write $f_{i,a}$ for the number of copies of a in the multiset F_i.

Lemma 3.9. *If $f_{i,2} > \frac{m}{2} - 6$ for $1 \le i \le 4$, then $\bigcup_{i=1}^{4} G_i = [1^4, 2^{2n-2}]$ and $G_i = [1^n]$ for $i > 4$. In particular, (2.3.2) holds.*

Proof. Let k be the number of odd elements in $\bigcup_{i=1}^{4} G_i$. If $k \ge 5$ then

$$2m-2 \ge \sum_{i=1}^{4}\sum_{a\in F_i}\sum_{b\in G_i}\big(a-\gcd(a,b)\big) \ge \sum_{i=1}^{4}\sum_{\substack{a\in F_i\\ a=2}}\sum_{\substack{b\in G_i\\ b \text{ odd}}} 1 > 5\Big(\frac{m}{2}-6\Big) > 2m-2,$$

a contradiction. Hence $k \le 4$, so by (2.5) we have

$$2n-2 = \sum_{i=1}^{r}\big(n-|G_i|\big) \ge \sum_{i=1}^{4}\big(n-|G_i|\big) \ge 4n - \Big(k + \frac{4n-k}{2}\Big) = \frac{4n-k}{2} \ge 2n-2.$$

Thus this chain of inequalities must consist of equalities; proceeding from left to right, it follows that if $i > 4$ then $|G_i| = n$ (and hence $G_i = [1^n]$); if $i \leq 4$ then G_i contains only 1's and 2's; and finally, $k = 4$. This yields the desired conclusion. □

Proof of Theorem 2.3 when $f_{i,1} \leq 4$ for at least four i's. Without loss of generality, we may assume that $f_{i,1} \leq 4$ for $1 \leq i \leq 4$, so that $|F_i| \leq 4 + \frac{m-4}{2} = \frac{m}{2} + 2$ and $m - |F_i| \geq \frac{m}{2} - 2$ for $1 \leq i \leq 4$. If $f_{1,1} + f_{1,2} \leq 4$ then $|F_1| \leq 4 + \frac{m-4}{3} = \frac{m+8}{3}$ and therefore $m - |F_1| \geq \frac{2m-8}{3}$. Then $\sum_{1 \leq i \leq 4}(m - |F_i|) \geq 3(\frac{m}{2} - 2) + \frac{2m-8}{3} = \frac{13m}{6} - \frac{26}{3} > 2m - 2$, a contradiction. Thus $f_{1,1}+f_{1,2} > 4$, and then by Proposition 3.7 we must have $f_{1,2} > \frac{m}{2} - 6$ and similarly $f_{i,2} > \frac{m}{2} - 6$ for $2 \leq i \leq 4$. Lemma 3.9 yields the desired conclusion. □

Lemma 3.10. *If $|F_i| = 1$ then $\gcd(F_i, G_i) = m$ or $\sum_{a \in F_i} \sum_{b \in G_i} \big(a - (a, b)\big) \geq \frac{m}{2}$. If $|F_i| = 2$ and $n \leq m + 4$ then $\gcd(F_i, G_i) = \frac{m}{2}$ or $\sum_{a \in F_i} \sum_{b \in G_i} \big(a - (a, b)\big) \geq \frac{m}{4}$. If $|F_i| = 3$ and $n \leq m + 4$ then $\gcd(F_i, G_i) \in \{\frac{m}{3}, \frac{m}{4}, \frac{m}{6}\}$ or $\sum_{a \in F_i} \sum_{b \in G_i} \big(a - (a, b)\big) \geq \frac{m}{6}$.*

Proof. We prove Lemma 3.10 when $|F_i| = 1$ and $F_i = [m]$. If m divides each element of G_i then $\gcd(F_i, G_i) = m$. If G_i contains an element c which is not divisible by m, then $\sum_{a \in F_i} \sum_{b \in G_i} \big(a - (a, b)\big) \geq m - (m, c) \geq \frac{m}{2}$. The proofs of the other two assertions are a bit more complicated, but are based on similar ideas. □

Proof of Theorem 2.3 when at most two i's satisfy $f_{i,1} \leq 4$. By Proposition 3.7, if $f_{i,1} > 4$ then $f_{i,1} \geq m - 5$, so that $|F_i| \geq m - 4$ and therefore $m - |F_i| \leq 4$. By (3.5) it follows that $g_{i,1} \geq n - \frac{4n}{f_{i,1}} \geq n - \frac{4n}{m-4} \geq n - 4 \cdot 85 = n - 340$ because $n \leq 84m$. Hence there are at most two i's for which $4 < f_{i,1} < m$, since otherwise

$$2m = \sum_{i=1}^{r} \sum_{a \in F_i} \sum_{b \in G_i} \big(a - (a, b)\big) \geq 3(n - 340) > 2m.$$

Since each such i satisfies $m - |F_i| \leq 4$, it follows that the sum of the corresponding values of $(m - |F_i|)$ is at most 8, so (2.5) implies that $f_{i,1} \leq 4$ for at least two (hence exactly two) values of i.

We may assume that $f_{i,1} \leq 4$ if and only if $i \leq 2$. Then $g_{i,1} \geq n - 340$ for $i \geq 3$, so

$$2m \geq \sum_{i \geq 3} g_{i,1} \sum_{a \in F_i} (a - 1) \geq \sum_{i \geq 3} g_{i,1}(m - |F_i|) \geq (n - 340) \sum_{i \geq 3} (m - |F_i|).$$

This implies $\sum_{i\geq 3}(m-|F_i|) \leq \frac{2m}{n-340} < 3$ because $n \leq 84m$. If $\sum_{i\geq 3}(m-|F_i|) = 2$, then by (2.5), $2m-2 = m-|F_1|+m-|F_2|+\sum_{i\geq 3}(m-|F_i|)$ so $|F_1|+|F_2| = 4$. In particular, $1 \leq |F_1| \leq 3$. We must also have

$$\sum_{i=1}^{2}\sum_{a\in F_i}\sum_{b\in G_i}(a-(a,b)) \leq 2m-2(n-\frac{4n}{m-5}) < 9;$$

this also implies $n \leq m+4$. If $|F_1| = 1$ and $|F_2| = 3$ then by Lemma 3.10 we can conclude that $\gcd(F_1, G_1) = m$, since $\frac{m}{2} > 9$, and that $\gcd(F_2, G_2) \in \{\frac{m}{3}, \frac{m}{4}, \frac{m}{6}\}$, since $\frac{m}{6} > 9$. Then $\gcd(F_1, F_2, G_1, G_2) > 1$, which contradicts (2.9.1). A similar argument demonstrates that when $|F_1| = |F_2| = 2$ or $|F_1| = 3$ and $|F_2| = 1$, we must again have $\gcd(F_1, F_2, G_1, G_2) > 1$, a contradiction. Thus $\sum_{i\geq 3}(m-|F_i|) \leq 1$, and then by (2.5), $2m-2 = m-|F_1|+m-|F_2|+\sum_{i\geq 3}(m-|F_i|)$, so $|F_1|+|F_2| \leq 3$. If $|F_1|+|F_2| = 2$, we must have $F_1 = F_2 = [m]$ and then (2.3.1) holds. If $|F_1|+|F_2| = 3$, we can write $F_1 = [m]$ and $F_2 = [c, m-c]$. Now an analysis of the multisets G_i, using (2.5) and (2.7), yields (2.3.3). □

Proof of Theorem 2.3 when exactly three i's satisfy $f_{i,1} \leq 4$. We split this case into four subcases based on how many of the three points satisfy $f_{i,2} > 4$. Three of these four subcases are resolved via the methods used to treat the case when $f_{i,1} \leq 4$ for at least four i's. The fourth subcase, when exactly two i's satisfy $f_{i,2} > 4$, is more difficult. We resolve this subcase by combining several tools, including Lemmas 2.8 and 2.9, a more general version of Lemma 3.10, and a series of computer programs. □

We have indicated the main arguments in all cases, which concludes our proof of Theorem 2.3.

4. From the LCM Theorem to Rational Functions

In this section we describe the possibilities for the ramification multisets F_i and G_i which satisfy the hypotheses of Theorem 2.3 when (2.3.1) or (2.3.2) holds. By symmetry it suffices to do this when (2.3.1) holds, that is, $\sum_{i=1}^{r}(1-\frac{1}{\operatorname{lcm}(F_i)}) \leq 2$. It is easy to determine all corresponding possibilities for the sequence of $\operatorname{lcm}(F_i)$'s. For each such sequence, we use (2.5) to find the F_i's, and then we use (2.7) to determine the G_i's. In some cases we go further and describe the corresponding rational functions $f(X)$ and $g(X)$. To simplify the description of these functions, we often change variables by replacing f and g by $\mu \circ f \circ \nu_1$ and $\mu \circ g \circ \nu_2$ for some degree-one $\mu, \nu_1, \nu_2 \in \mathbb{C}(X)$; such a change does not affect whether the numerator of $f(X) - g(Y)$ defines an irreducible curve of genus 0 or 1, or whether (2.3.1) holds. Our result when $m > 60$ is as follows; the result when $m \le 60$ is similar but longer.

Theorem 4.1. *If (2.3.1) holds, $m > 60$, and the numerator of $f(X) - g(Y)$ defines an irreducible curve of genus 0 or 1, then we can relabel the Q_i's so that one of these occurs:*

(4.1.1) $F_1 = F_2 = [m]$ *and* $F_i = [1^m]$ *for all* $i > 2$. *Up to a change of variables, in this case* $f(X) = X^m$ *and* $g(X) = X^a h(X)^m$ *for some integer* a *coprime to* m *and some* $h \in \mathbb{C}(X)$.

(4.1.2) $F_1 \cup F_2 = [1^2, 2^{m-1}]$ *and* $F_3 = [m]$, *while* $F_i = [1^m]$ *for all* $i > 3$. *Up to a change of variables, in this case* $f(X) = T_m(X)$, $Q_1 = 1$, $Q_2 = -1$, $Q_3 = \infty$, *and one of the following holds, where* $G'_i := [\gcd(b, \operatorname{lcm}(F_i)) : b \in G_i, \operatorname{lcm}(F_i) \nmid b]$*:*

- $G'_1 \cup G'_2 = [1^{2c}], G'_3 = [1^{2-c}]$, *where* $c \in \{0, 1, 2\}$ *and if* m *even then* $|G'_1| = |G'_2|$
- m *is even,* $G'_1 = G'_3 = \emptyset$, $G'_2 = [1^4]$
- m *is even,* $G'_1 = \emptyset$, $G'_2 = [1^2]$, $G'_3 = [2]$.

(4.1.3) $F_1 = F_2 = [2^{m/2}]$ *and* $F_3 = [\left(\frac{m}{2}\right)^2]$, *while* $F_i = [1^m]$ *for all* $i > 3$. *Up to a change of variables, in this case* $f(X) = X^{m/2} + X^{-m/2}$, $Q_1 = 2$, $Q_2 = -2$, $Q_3 = \infty$, $G'_1 \cup G'_2 = [1^{2c}]$, *and* $G'_3 = [1^{2-c}]$ *for some* $c \in \{1, 2\}$ *(where we use the same notation as in (4.1.2)).*

(4.1.4) $\bigcup_{i=1}^4 F_i = [1^4, 2^{2m-2}]$, $\bigcup_{i=1}^4 G_i = [1^4, 2^{2n-2}]$, *and* $r = 4$.

(4.1.5) $\bigcup_{i=1}^3 F_i = [1^3, 3^{m-1}]$, $\bigcup_{i=1}^3 G_i = [1^3, 3^{n-1}]$, *and* $r = 3$.

(4.1.6) $F_1 = [1^{a_1}, 2^{\frac{m-a_1}{2}}]$, $G_1 = [1^{a_2}, 2^{\frac{m-a_2}{2}}]$, $F_2 \cup F_3 = [1^{b_1}, 2^{c_1}, 4^{\frac{2m-b_1-2c_1}{4}}]$, $G_2 \cup G_3 = [1^{b_2}, 2^{c_2}, 4^{\frac{2m-b_2-2c_2}{4}}]$, *and* $r = 3$, *where* a_j, b_j, c_j *are nonnegative integers satisfying* $2(a_j + c_j) + 3b_j = 8$ *for* $j \in \{1, 2\}$.

(4.1.7) $F_1 = [1^{a_1}, 2^{\frac{m-a_1}{2}}]$, $G_1 = [1^{a_2}, 2^{\frac{m-a_2}{2}}]$, $F_2 = [1^{b_1}, 3^{\frac{m-b_1}{3}}]$, $G_2 = [1^{b_2}, 3^{\frac{m-b_2}{3}}]$, $F_3 = [1^{c_1}, 2^{d_1}, 3^{e_1}, 6^{\frac{m-c_1-2d_1-3e_1}{6}}]$, $G_3 = [1^{c_2}, 2^{d_2}, 3^{e_2}, 6^{\frac{m-c_2-2d_2-3d_2}{6}}]$, *and* $r = 3$, *where* a_j, b_j, c_j, d_j, e_j *are nonnegative integers satisfying* $3(a_j + e_j) + 4(b_j + d_j) + 5c_j = 12$.

Proof. We first determine the possibilities for the F_i's. Let C denote the multiset of all values $\operatorname{lcm}(F_i)$, and let D be the multiset of elements in C which are at least 2. By (2.3.1) we have

$$2 \geq \sum_{c \in C} \Bigl(1 - \frac{1}{c}\Bigr) = \sum_{c \in D} \Bigl(1 - \frac{1}{c}\Bigr).$$

Thus Lemma 3.2 implies that either $|D| \leq 2$ or D is one of $[2^4]$, $[3^3]$, $[2, 4^2]$, $[2, 3, \ell]$ with $2 \leq \ell \leq 6$, or $[2^2, k]$ with $k > 1$. We analyze each case in succession. For instance, if $|D| = 2$ then F_i equals $[1^m]$ for all but two values of i, which we may assume are $i = 1$ and $i = 2$. Then (2.5) implies that $\sum_{i=1}^{2}(m - |F_i|) = 2m - 2$, so $\sum_{i=1}^{2} |F_i| = 2$ and thus $F_1 = F_2 = [m]$, as in (4.1.1). Likewise, if $D = [2^4]$ then $F_i = [1^{c_i}, 2^{d_i}]$ for some nonnegative integers c_i, d_i, where we may assume that $d_i = 0$ when $i > 4$. Since $c_i + 2d_i = m$, we have $|F_i| = \frac{m+c_i}{2}$, so that (2.5) implies that $\sum_{i=1}^{4} c_i = 4$, as in (4.1.4). Via similar arguments, we can solve for the F_i's in every case, obtaining F_i's as in one of (4.1.1)–(4.1.7) in all cases except when $F_1 = F_2 = [2^{m/2}]$ and $|F_3| = 2$; in that case Lemma 2.8 implies that $F_3 = [(\frac{m}{2})^2]$, so that the F_i's are as in (4.1.3).

Now we compute the possibilities for the G_i's and, when possible, for the functions f and g. First suppose that the F_i's are as in (4.1.1). Then (2.7) becomes

$$\sum_{i=1}^{2} \sum_{b \in G_i} \bigl(m - (m, b)\bigr) \in \{2m - 2, 2m\}.$$

Letting D be the multiset of values $\frac{m}{(m,b)}$ where $b \in G_1 \cup G_2$ and $m \nmid b$, we see that D is a finite multiset of integers greater than 1 and $\sum_{d \in D}(1 - \frac{1}{d}) \leq 2$. By Lemma 3.2, either $|D| \leq 2$ or D is one of $[2^4]$, $[3^3]$, $[2, 4^2]$, $[2^2, k]$ with $k > 1$, or $[2, 3, \ell]$ with $2 \leq \ell \leq 6$. By (2.9.1), the least common multiple of the elements of D must be m, so either $|D| \leq 2$ or $D = [2^2, k]$. In the latter case, at least one of G_1 and G_2 (say G_1) consists of elements divisible by $\frac{m}{2}$, so $\frac{m}{2}$ divides the sum of the elements in G_1, which is n. Hence $\frac{m}{2}$ also divides the sum of the elements in G_2; since $\frac{m}{2}$ divides all but at most one element of G_2, it must divide all elements of G_2, so $\frac{m}{2} \mid \frac{m}{k}$ and thus $k \mid 2$, contrary to the condition

$\operatorname{lcm}(D) = m$. Thus $|D| \le 2$, so since $\sum_{d\in D}\bigl(m - \frac{m}{d}\bigr) \ge 2m-2$ we must have $D = [m^2]$. By replacing f and g by $f\circ\nu_1$ and $g\circ\nu_2$ for some degree-one $\nu_i \in \mathbb{C}(X)$, we may assume that $f(0) = Q_1$, $f(\infty) = Q_2$, and that the two points P in $g^{-1}(\{Q_1, Q_2\})$ for which $m \nmid e_g(P)$ are 0 and ∞. By replacing f and g by $\mu\circ f$ and $\mu\circ g$ for some degree-one $\mu \in \mathbb{C}(X)$, we may assume that $Q_1 = 0$, $Q_2 = \infty$, and the numerator and denominator of f have the same leading coefficient. It follows that $f(X) = X^m$ and $g(X) = X^a h(X)^m$ for some a coprime to m and some $h \in \mathbb{C}(X)$, so that (4.1.1) holds.

Next suppose that the F_i's are as in (4.1.2). Then the numbers of 1's in F_1 and F_2 are d and $2-d$, where we may assume that $d > 0$ by interchanging Q_1 and Q_2 if necessary. Putting $p := |G_1'|$, $q := |G_2'|$, and $D := [\frac{m}{b} \colon b \in G_3']$, (2.7) becomes

$$\text{(4.2)} \qquad \frac{m-d}{2}p + \frac{m+d-2}{2}q + \sum_{b\in D}\Bigl(m - \frac{m}{b}\Bigr) \in \{2m-2, 2m\}.$$

Hence $\frac{m-2}{2}(p+q) \le 2m$, so that $p+q < 5$. Since the sum of the elements in $G_1 \cup G_2$ is $2n$, and this sum is congruent mod 2 to $p+q$, it follows that $p+q$ is even, so $p+q \in \{0,2,4\}$. If $p+q = 4$ then (4.2) implies that $D = \emptyset$, and that if in addition m is even (so that $d = 2$) then $p \in \{0,2\}$. Now assume $p+q \le 2$, so by (2.9.2) we have $\operatorname{lcm}(D) \ge \frac{m}{2}$. Since $2 > \sum_{b\in D}\bigl(1 - \frac{1}{b}\bigr)$, Lemma 3.2 implies that either $|D| \le 2$ or $D = [2^2, k]$ with $k \ge \frac{m}{4}$. If $p+q = 0$ then by (4.2) we have either $D = [1^2]$ or $D = [2^2, \frac{m}{2}]$, but the latter possibility cannot occur since exactly one element of G_3 is not divisible by $\frac{m}{2}$, contradicting Lemma 2.8. If $p+q = 2$ then (4.2) rules out $D = [2^2, k]$, and one easily checks that (since $\operatorname{lcm}(D) \ge \frac{m}{2}$) the only possibilities are either $D = [\frac{m}{2}]$ with $d = 2$ and $p = 0$, or $D = [m]$ with $1 \in \{d, p\}$. Finally, by replacing f and g by $\mu\circ f\circ\nu_1$ and $\mu\circ g$ for some degree-one $\mu, \nu_1 \in \mathbb{C}(X)$, we may assume that $f(1) = 1 = Q_1$, $f(\infty) = \infty = Q_3$, $Q_2 = -1$, and $e_f(1) = 1 = e_f(-1)$; since the F_i's are as in (4.1.2), these conditions imply that $f(X) = T_m(X)$ [34], so that (4.1.2) holds.

Lemma 3.9 implies the result when the F_i's are as in (4.1.4), and we have used similar arguments to prove Theorem 4.1 in the remaining cases.

□

Theorem 4.1 describes a short list of candidate ramification types for f and g. However, it is not always true that each such candidate ramification type is actually the ramification type of a rational function. For instance, the theory of elliptic curves, along with an analytic result due to Milnor [24, Thm. 3.1], implies that there exists $f \in \mathbb{C}(X)$ with ramification as in (4.1.5) if and only if m can be written as $c^2 + cd + d^2$ for some nonnegative integers c and d. In general, the existence of a rational function with prescribed ramification type can be determined by means of the following result of Hurwitz, which reduces the question to examining tuples of elements in a finite symmetric group (which is a finite problem for any prescribed degree):

Theorem 4.3 (Hurwitz)**.** *For any positive integer m, any multisets $A_1, \ldots, A_r$ consisting of positive integers, and any distinct $Q_1, \ldots, Q_r \in \mathbb{C} \cup \{\infty\}$, the following are equivalent:*

(4.3.1) *there exists a degree-m rational function $f(X) \in \mathbb{C}(X)$ such that $E_f(Q_i) = A_i$ and $E_f(Q) = [1^m]$ for each $Q \notin \{Q_1, \ldots, Q_r\}$*

(4.3.2) *all three of the following hold:*

- $\sum_{a \in A_i} a = m$ *for each i with $1 \le i \le r$*
- $\sum_{i=1}^{r} (m - |A_i|) = 2m - 2$
- *there are elements $g_1, \ldots, g_r \in S_m$ such that the multiset of cycle lengths of g_i is A_i, the product $g_1 g_2 \ldots g_r$ is the identity permutation, and the subgroup of S_m generated by $g_1, \ldots, g_r$ is transitive.*

Furthermore, Theorem 4.3 can be combined with (2.7) and Fried's reducibility theorem [14] in order to give a similar characterization of the ramification types of pairs of rational functions (f, g) for which the numerator of $f(X) - g(Y)$ defines an irreducible curve of genus 0 or 1. Thus it is a finite problem to determine the ramification types of all such pairs of rational functions having prescribed degrees. In particular, it is a finite problem to determine all solutions to Questions 1.4 or 1.5 which do not satisfy the hypotheses of Theorem 1.6.

5. Conclusion

In this paper, we use combinatorial and algebraic methods to prove geometric results (Theorems 2.3 and 4.1) describing the ramification of large-degree complex rational functions f and g for which the numerator of $f(X) - g(Y)$ defines an irreducible curve of genus 0 or 1. We deduce two consequences: the number theoretic Theorem 1.3 addressing rational functions $f, g \in \mathbb{Q}(X)$ for which $f(\mathbb{Q}) \cap g(\mathbb{Q})$ is infinite, and the analytic Theorem 1.6 about the functional equation $f \circ \hat{f} = g \circ \hat{g}$ with $f, g \in \mathbb{C}(X)$ and $\hat{f}, \hat{g}$ meromorphic on $\mathbb{C}$. The results show that the rational functions satisfying any of these conditions are unexpectedly nice: it must be that either the Galois closure of $\mathbb{C}(X)/\mathbb{C}(f(X))$ has genus 0 or 1 (in which case all corresponding functions are understood), or the analogous condition holds for g, or there is a change of variables turning the equation $f(X) = g(Y)$ into the special equation $X^c(X-1)^d = \gamma Y^c(Y-1)^d$.

In the future, the numerical bounds will hopefully be removed from the hypotheses of Theorems 1.3, 1.6, and 2.3. Another idea is exploring the possibility of using the case where the numerator of $f(X) - g(Y)$ is irreducible as the base case for an inductive approach that resolves the case where the numerator of $f(X) - g(Y)$ is reducible. Finally, Theorem 4.3 and other results can be used to determine which of the ramification types in the conclusion of Theorem 4.1 actually correspond to pairs of rational functions $f, g \in \mathbb{C}(X)$ satisfying the hypotheses of that result.

6. Acknowledgements

We would like to acknowledge several individuals that have supported us throughout MIT-PRIMES 2015. First of all, we would like to thank Professor Michael E. Zieve from the University of Michigan and Thao Do from MIT for suggesting this project and mentoring us throughout the year. We would also like to thank Tanya Khovanova and other leaders at the MIT-PRIMES program for their assistance. Finally, we want to thank MIT-PRIMES program for giving us the opportunity to conduct this incredibly fascinating mathematical research, and of course our parents for their invaluable support throughout the year.

References

[1] M. Aschbacher, *On conjectures of Guralnick and Thompson*, J. Algebra **135** (1990), 277–343.

[2] R. M. Avanzi and U. M. Zannier, *Genus one curves defined by separated variable polynomials and a polynomial Pell equation*, Acta Arith. **99** (2001), 227–256.

[3] M. Baker and L. De Marco, *Special curves and postcritically finite polynomials*, Forum Math. Pi **1** (2013), e3, 35 pp.

[4] M. Bhargava and A. Shankar, *Binary quartic forms having bounded invariants, and the boundedness of the average rank of elliptic curves*, Annals of Math. **181** (2015), 191–242.

[5] M. Bhargava and A. Shankar, *Ternary cubic forms having bounded invariants, and the existence of a positive proportion of elliptic curves having rank* 0, Annals of Math. **181** (2015), 587–621.

[6] M. Bhargava, C. Skinner and W. Zhang, *A majority of elliptic curves over* $\mathbb{Q}$ *satisfy the Birch and Swinnerton-Dyer conjecture*, arXiv:1407.1826 (2014).

[7] Y. F. Bilu and R. F. Tichy, *The Diophantine equation* $f(x) = g(y)$, Acta Arith. **95** (2000), 261–288.

[8] M. Briskin, N. Roytvarf and Y. Yomdin, *Center conditions at infinity for Abel differential equations*, Annals of Math. (2) **172** (2010), 437–483.

[9] J. Carlson, A. Jaffe and A. Wiles (eds.), The Millennium Prize Problems, Clay Mathematics Institute, Cambridge, MA (2006).

[10] A. Carney, T. Do, J. Hallett, Y. Jiang, B. L. Weiss, E. Wells and M. E. Zieve, *Diophantine equations with separated variables, I: the irreducible case*, preprint (2015).

[11] A. Carney, J. Hallett, Q. Sun, B. L. Weiss, Y. Xia and M. E. Zieve, *Diophantine equations with separated variables, II: the reducible case*, preprint (2015).

[12] H. Davenport, D. J. Lewis and A. Schinzel, *Equations of the form* $f(x) = g(y)$, Quart. J. Math. Oxford (2) **12** (1961), 304–312.

[13] G. Faltings, *Endlichkeitssätze für abelsche Varietäten über Zahlkörpern*, Invent. Math. **73** (1983), 349–366.

[14] M. Fried, *The field of definition of function fields and a problem in the reducibility of polynomials in two variables*, Illinois J. Math. **17** (1973), 128–146.

[15] M. D. Fried, *On a theorem of Ritt and related Diophantine problems*, J. Reine Angew. Math. **264** (1973), 40–55.

[16] D. Frohardt and K. Magaard, *Composition factors of monodromy groups*, Annals of Math. **154** (2001), 327–345.

[17] D. Ghioca, T. J. Tucker and M. E. Zieve, *Intersections of polynomial orbits, and a dynamical Mordell–Lang conjecture*, Invent. Math. **171** (2008), 463–483.

[18] D. Ghioca, T. J. Tucker and M. E. Zieve, *Linear relations between polynomial orbits*, Duke Math. J. **161** (2012), 1379–1410.

[19] R. M. Guralnick and J. G. Thompson, *Finite groups of genus zero*, J. Algebra **131** (1990), 303–341.

[20] H. W. Lenstra, Jr., *Solving the Pell equation*, Notices of the Amer. Math. Soc **49** (2002), 182–192.

[21] M. Liebeck and A. Shalev, *Simple groups, permutation groups, and probability*, J. Amer. Math. Soc. **12** (1999), 266–314.

[22] M. Lyubich and Y. Minsky, *Laminations in holomorphic dynamics*, J. Diff. Geom. **47** (1997), 17–94.

[23] A. Medvedev and T. Scanlon, *Invariant varieties for polynomial dynamical systems*, Annals of Math. **179** (2014), 81–177.

[24] J. Milnor, *On Lattès maps*, in: Dynamics on the Riemann sphere, pages 9–43. Eur. Math. Soc., Zürich, 2006.

[25] D. Neftin and M. E. Zieve, *Monodromy groups of minimal covers*, preprint (2015).

[26] R. Nevanlinna, *Einige Eindeutigkeitssätze in der Theorie der Meromorphen Funktionen*, Acta Math. **48** (1926), 367–391.

[27] K. D. Nguyen, *Algebraic independence of local conjugacies and related questions in polynomial dynamics*, Proc. Amer. Math. Soc. **143** (2015), 1491–1499.

[28] F. Pakovich, *On polynomials sharing preimages of compact sets, and related questions*, Geom. Funct. Anal. **18** (2008), 163–183.

[29] F. Pakovich, *Prime and composite Laurent polynomials*, Bull. des Sci. Math. **133** (2009), 693–732.

[30] F. Pakovich, *On the equation $P(f) = Q(g)$ where P, Q are polynomials and f, g are entire functions*, Amer. J. Math. **132** (2010), 1591–1607.

[31] F. Pakovich, *Algebraic curves $P(x) - Q(y) = 0$ and functional equations*, Complex Var. Elliptic Equ. **56** (2011), 199–213.

[32] F. Pakovich and M. Muzychuk, *Solution of the polynomial moment problem*, Proc. London Math. Soc. (3) **99** (2009), 633–657.

[33] E. Picard, *Démonstration d'un théorème général sur les fonctions uniformes liées par une relation algébrique*, Acta Math. **11** (1887), 1–12.

[34] J. F. Ritt, *Prime and composite polynomials*, Trans. Amer. Math. Soc. **23** (1922), 51–66.

[35] J. F. Ritt, *Permutable rational functions*, Trans. Amer. Math. Soc. **25** (1923), 399–448.

[36] A. Schinzel, Selected Topics on Polynomials, The Univ. of Michigan Press, Ann Arbor, 1982.

[37] C. L. Siegel, *The integer solutions of the equation $y^2 = ax^n + bx^{n-1} + \cdots + k$*, J. London Math. Soc. **1** (1926), 66–68.

[38] A. Wiles, *Modular elliptic curves and Fermat's Last Theorem*, Annals of Math. **142** (1995), 443–551.

[39] U. Zannier, *On a functional equation relating a Laurent series $f(x)$ to $f(x^m)$*, Aequat. Math. **55** (1998), 15–43.

[40] M. E. Zieve, *Decompositions of Laurent polynomials*, arXiv:0710.1902 (2007).