

# On rational functions with more than three branch points

Jijian Song<sup>1</sup> and Bin Xu<sup>2</sup>

December 30, 2019

<sup>1</sup> Center for Applied Mathematics, Tianjin University.

No. 135 Yaguan Road, Tianjin 300350, P. R. China.

<sup>2</sup>Wu Wen-Tsun Key Laboratory of Mathematics, USTC, Chinese Academy of Sciences

School of Mathematical Sciences, University of Science and Technology of China.

No. 96 Jinzhai Road, Hefei, Anhui Province 230026 P. R. China.

<sup>1</sup>✉smath@mail.ustc.edu.cn

<sup>2</sup>bxu@ustc.edu.cn

**Abstract:** Let  $\Lambda$  be a collection of partitions of a positive integer  $d$  of the form

$$(\mathbf{a}_1, \dots, \mathbf{a}_p), (\mathbf{b}_1, \dots, \mathbf{b}_q), (\mathbf{m}_1 + 1, 1, \dots, 1), \dots, (\mathbf{m}_l + 1, 1, \dots, 1),$$

where  $(\mathbf{m}_1, \dots, \mathbf{m}_l)$  is a partition of  $p + q - 2 > 0$ . We prove that there exists a rational function on the Riemann sphere  $\overline{\mathbb{C}}$  with branch data  $\Lambda$  if and only if

$$\max(\mathbf{m}_1, \dots, \mathbf{m}_l) < \frac{d}{\text{GCD}(\mathbf{a}_1, \dots, \mathbf{a}_p, \mathbf{b}_1, \dots, \mathbf{b}_q)}.$$

As an application, we give a new class of branch data which can be realized by Belyi functions on the Riemann sphere.

**Keywords.** branch data, Realizability Problem, Belyi function, Riemann's existence theorem

**2010 Mathematics Subject Classification.** Primary 20B35; secondary 14H30

## 1 Introduction

Let  $X$  and  $Y$  be two compact connected Riemann surfaces, and let  $f : X \rightarrow Y$  be a holomorphic branched covering of degree  $d$ . For each point  $q$  in  $Y$ , there is a partition  $\lambda(q) = (k_1, \dots, k_r)$  of  $d$  associated to  $q$  such that, over a suitable neighborhood of  $q$  in  $Y$ ,  $f$  is equivalent to the map

$$\{1, \dots, r\} \times \mathbb{D} \rightarrow \mathbb{D}, \quad (j, z) \mapsto z^{k_j}, \quad \text{where } \mathbb{D} := \{z \in \mathbb{C} : |z| < 1\},$$

with  $q$  corresponding to  $0$  in  $\mathbb{D}$ . For any partition  $\lambda = (k_1, k_2, \dots, k_r)$  of  $d$ , we define its length  $\text{Len}(\lambda) = r$ . We call the partition  $\lambda$  of  $d$  *non-trivial* if  $\text{Len}(\lambda) < d$ . For the branched covering  $f: X \rightarrow Y$ , We call a point  $q$  in  $Y$  a *branch point* of  $f$  if and only if  $\lambda(q)$  is non-trivial, and we call the set of branch points of  $f$  *the branch set* of  $f$ , denoted by  $B_f$ . The collection  $\Lambda = \{\lambda(q) : q \in B_f\}$  (with repetitions allowed) is called the *branch data* of  $f$  and

$$v(f) := \sum_{q \in B(f)} (d - \text{Len}(\lambda(q)))$$

the *total branching order* of  $f$ . By the Riemann-Hurwitz theorem, we have that

$$v(f) = 2g(X) - 2 - d(2g(Y) - 2)$$

where  $g(X)$  (resp.  $g(Y)$ ) denotes the genus of  $X$  (resp.  $Y$ ). Therefore, the total branching order  $v(f)$  is an even non-negative integer.

The following problem was first proposed by Edmonds-Kulkarni-Stong [4] and we can trace its history to Hurwitz [9].

**Realizability Problem.** Given a compact connected Riemann surface  $Y$  and a collection  $\Lambda = \{\lambda_1, \dots, \lambda_k\}$  of non-trivial partitions of a positive integer  $d$ , does there exist another compact connected Riemann surface  $X$  together with a branched covering  $f: X \rightarrow Y$  such that  $\Lambda$  is its branch data? If it does, we call the collection  $\Lambda$  *realizable* or *realized by a branched covering*.

See the classical [2, 4, 6, 7, 8, 10, 11, 13, 14, 22, 23] and the more recent [1, 12, 15, 16, 17, 18, 19, 20, 21, 24] about this problem. Here we only review some necessary background and a small part of known results which are closely related to our discussions in the sequel.

Recall that in order to be realizable, a collection  $\Lambda$  should satisfy the condition that its *total branching order*

$$v(\Lambda) := \sum_{j=1}^k (d - \text{Len}(\lambda_j))$$

is even. We call such a collection *compatible*. It is proved in [10, Theorem 9] and [4, Section 3] that a compatible collection is always realizable if  $g(Y) > 0$ . Hence, we always assume that  $Y$  is the Riemann sphere  $\overline{\mathbb{C}}$  in the sequel. It turns out that a compatible collection is not always realizable in this case. We call a compatible collection an *exception* if it is not realizable. Zheng [24] found by computer all the exceptions of degree  $\leq 22$ . Pervova-Petronio [20, 21] used a variety of techniques to give some new infinite series of exceptions, and they used dessins d'enfants to make a theoretical explanation of part of the exceptions given by Zheng [24]. Besides constructing some exceptions, Edmonds-Kulkarni-Stong [4] proposed the so-called *prime degree conjecture*, which says that each compatible collection with prime degree is always realizable, and they reduced it in the same paper to the collections with exactly three partitions. In [18, 19], Pascali-Petronio proved some results which provide strong support to this conjecture.

Characterizing branch data of all rational functions is a very deep and difficult problem, which seems far from being accessible nowadays. Hence, it is meaningful to find reasonably

simple, sufficient conditions for a collection to be realizable. Besides the theorems in [18, 19, 20, 21], some of the other known results are as follows: Thom [23] showed that a compatible collection is realizable if one partition in it has length one. Edmonds-Kulkarni-Stong [4, Theorem 5.4] proved that a compatible collection with degree  $d \neq 4$  is realizable when its total branching order  $\geq 3(d-1)$ . In addition, the exceptions with  $d=4$  are precisely those with partitions  $(2,2), \dots, (2,2), (3,1)$ . Moreover, Boccara [2] obtained a complete determination of the realizability of the collection  $\Lambda$  which consists of the following three partitions of  $d$ :

$$(a_1, \dots, a_p), (b_1, \dots, b_q), (m+1, 1, \dots, 1).$$

He proved that  $\Lambda$  is realizable if and only if it satisfies one of the following two conditions:

- $v(\Lambda) \geq 2d$  is even.
- $v(\Lambda) = 2d - 2$  and  $m < \frac{d}{\text{GCD}(a_1, \dots, a_p, b_1, \dots, b_q)}$ . Note that  $m = p + q - 2$  in this case.

Generalizing the second part of Boccara's result, we show the following

**Theorem 1.1 (Main Theorem).** *Let  $d$  and  $l$  be two positive integers. Consider a collection*

$$\Lambda = \{(a_1, \dots, a_p), (b_1, \dots, b_q), (m_1 + 1, 1, \dots, 1), \dots, (m_l + 1, 1, \dots, 1)\}$$

*consisting of  $l+2$  partitions of  $d$  such that  $(m_1, \dots, m_l)$  is a partition of  $p+q-2 > 0$ . Then there exists a rational function on  $\overline{\mathbb{C}}$  with  $\Lambda$  as its branch data if and only if*

$$\max(m_1, \dots, m_l) < \frac{d}{\text{GCD}(a_1, \dots, a_p, b_1, \dots, b_q)}.$$

**Remark 1.2.** Recently, A. Eremenko [5] applied the main theorem to the investigation of conformal metrics of positive constant curvature on  $\overline{\mathbb{C}}$  which have finitely many conical singularities and co-axial holonomy. In particular, he characterized the conical angles of such metrics and used the main theorem in the solution of Question 2, which is the rational case of Question 1, namely the main problem in [5].

We call a rational function on a compact Riemann surface a *Belyi function* if it has at most three branch points. As an application of the main theorem, we have:

**Theorem 1.3.** *Let  $d$  and  $r$  be two positive integers, and*

$$\Lambda = \{(a_1, \dots, a_p), (b_1, \dots, b_q), (c_1 + 1, \dots, c_r + 1, 1, \dots, 1)\}$$

*be a collection consisting of partitions of  $d$  such that  $(c_1, \dots, c_r)$  is a partition of  $p+q-2 > 0$ . If*

$$\max(c_1, \dots, c_r) < \frac{d}{\text{GCD}(a_1, \dots, a_p, b_1, \dots, b_q)}$$

*then the modified collection*

$$\tilde{\Lambda} := \{(ra_1, \dots, ra_p), (rb_1, \dots, rb_q), (c_1 + 1, \dots, c_r + 1, 1, \dots, 1)\}$$

*of partitions of  $dr$  can be realized by a Belyi function on  $\overline{\mathbb{C}}$ .*

In the remaining of the article, we give proofs of the two theorems. At the very end, we propose a conjecture concerning rational functions on a Riemann surface of positive genus, which is a possible generalization of our main theorem.

## 2 Proof of the main theorem

In Subsection 2.1 we prove the necessary part of the main theorem and observe that the sufficient part can be reduced to the case of  $\text{GCD}(a_1, \dots, a_p, b_1, \dots, b_q) = 1$ . Moreover, we recall that by the Riemann existence theorem ([3, Theorem 2, p.49]) the sufficient part is equivalent to the existence of certain permutations in the symmetry group  $S_d := S_{\{1,2,\dots,d\}}$  associated with the collection  $\Lambda$ . For completeness, we give in Subsection 2.2 a proof in our own strategy for the case  $l = 1$  of the main theorem which was also proved by Boccara [2]. We prove Case  $l \geq 2$  of the main theorem in Subsection 2.3.

### 2.1 Riemann's existence theorem

At first, we prove the necessary part of the main theorem.

*Proof.* Suppose that there exists a rational function  $f$  on  $\overline{\mathbb{C}}$  realizing the branch data  $\Lambda$ . Using suitable Möbius transformations if necessary, we can assume that  $f$  has the form

$$f(z) = \frac{(z - z_1)^{a_1} \dots (z - z_p)^{a_p}}{(z - w_1)^{b_1} \dots (z - w_q)^{b_q}} \quad (1)$$

where  $z_1, \dots, z_p, w_1, \dots, w_q$  are  $(p + q)$  distinct complex numbers. Let

$$k = \text{GCD}(a_1, \dots, a_p, b_1, \dots, b_q).$$

Then we can write  $f$  as  $f = F^k$  for some rational function on  $\overline{\mathbb{C}}$ . And  $F$  has branch data of the form

$$\{(a_1/k, \dots, a_p/k), (b_1/k, \dots, b_q/k), (m_1 + 1, 1, \dots, 1), \dots, (m_l + 1, 1, \dots, 1)\}.$$

Since  $F$  has degree  $d/k$ , we have  $\max(m_1, \dots, m_l) < d/k$  and complete the proof.  $\square$

On the other hand, we claim that *if the collection*

$$\{(a_1/k, \dots, a_p/k), (b_1/k, \dots, b_q/k), (m_1 + 1, 1, \dots, 1), \dots, (m_l + 1, 1, \dots, 1)\}$$

*is realized by a rational function, then so is*

$$\{(a_1, \dots, a_p), (b_1, \dots, b_q), (m_1 + 1, 1, \dots, 1), \dots, (m_l + 1, 1, \dots, 1)\}.$$

Actually, by using Riemann's existence theorem and some Möbius transformations, we could assume that there exists a rational function

$$F(z) = \frac{(z - z_1)^{a_1/k} \dots (z - z_p)^{a_p/k}}{(z - w_1)^{b_1/k} \dots (z - w_q)^{b_q/k}},$$

where  $z_1, \dots, z_p, w_1, \dots, w_q$  are  $(p+q)$  distinct complex numbers, such that there exist exactly  $l$  branch points of  $F$  lying in  $\overline{\mathbb{C}} \setminus \{0, \infty\}$ , say  $y_1, \dots, y_l$ , which satisfy that  $y_1^k, \dots, y_l^k$  are mutually distinct and  $\lambda(y_j) = (m_j + 1, 1, \dots, 1)$  for  $1 \leq j \leq l$ . Since the power function  $y \mapsto y^k$  on  $\overline{\mathbb{C}}$  does not branch on  $\overline{\mathbb{C}} \setminus \{0, \infty\}$ , the collection

$$\{(a_1, \dots, a_p), (b_1, \dots, b_q), (m_1 + 1, 1, \dots, 1), \dots, (m_l + 1, 1, \dots, 1)\}.$$

is realized by the rational function  $F^k$ .

Hence, *in order to show the sufficient part of the main theorem, we may assume*  $\text{GCD}(a_1, \dots, b_q) = 1$ . We need to prepare some notions before giving the proof.

**Definition 2.1.** Let  $m$  be a non-negative integer. A vector  $\alpha = (a_1, a_2, \dots, a_{m+2})$  in  $\mathbb{Z}^{m+2}$  is called a *residue vector* with  $(m+2)$  components if  $a_1 + a_2 + \dots + a_{m+2} = 0$  and  $a_1 a_2 \dots a_{m+2} \neq 0$ . Two residue vectors  $\alpha = (a_1, \dots, a_{m+2})$  and  $\beta = (b_1, \dots, b_{m+2})$  are called *equivalent*, denoted by  $\alpha \sim \beta$ , if there is a nonzero rational number  $\mu$  and a permutation  $\sigma$  in the symmetry group  $S_{m+2}$  such that

$$\mu \cdot \alpha = (\mu a_1, \dots, \mu a_{m+2}) = \sigma(\beta) := (b_{\sigma(1)}, \dots, b_{\sigma(m+2)}).$$

This is an equivalence relation in the set of residue vectors with  $(m+2)$  components. The degree of the residue vector  $\alpha$  is defined to be

$$\deg(\alpha_1, \dots, \alpha_{m+2}) = \frac{\sum_{a_j > 0} a_j}{\text{GCD}(a_1, \dots, a_{m+2})}.$$

We call a residue vector  $\alpha = (a_1, \dots, a_{m+2})$  *primitive* if  $\text{GCD}(a_1, \dots, a_{m+2}) = 1$ . Clearly, the degree of a primitive residue vector equals the sum of all its positive components. Observe also that the logarithmic differential  $d(\log f) = \frac{df}{f}$  of a rational function  $f$  in (1) has residues  $a_1, \dots, a_p, -b_1, \dots, -b_q$ , which form a residue vector with degree  $d/\text{GCD}(a_1, \dots, a_p, b_1, \dots, b_q)$ .

**Definition 2.2.** Let  $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_l)$  be a partition of a positive integer  $n$ . The *weight* of  $\lambda$  is defined to be

$$\text{wt}(\lambda) = \max(\lambda_1, \dots, \lambda_l)$$

Use the notions in the main theorem and denote by  $\lambda$  the partition  $(m_1, \dots, m_l)$  of  $m = p + q - 2 > 0$  and by  $\alpha$  the residue vector  $(a_1, \dots, a_p, -b_1, \dots, -b_q)$ . Then the condition in the theorem can be concisely re-expressed as

$$\deg \alpha > \text{wt}(\lambda).$$

By the Riemann existence theorem [3, Theorem 2, p.49], the sufficient part of the main theorem is equivalent to the following

**Theorem 2.3.** *Under the assumptions of the main theorem, if  $\deg \alpha > \text{wt}(\lambda)$ , then there exist  $(l+2)$  permutations  $\tau_1, \tau_2, \sigma_1, \dots, \sigma_l$  in the symmetry group  $S_d = S_{\{1, 2, \dots, d\}}$  of  $\{1, 2, \dots, d\}$  such that*

- $\tau_1\tau_2\sigma_1\cdots\sigma_l = e$ , where  $e$  is the unit in  $S_d$  and permutations are multiplied from right to left;
- $\tau_1$  has the type of  $a_1^1a_2^1\cdots a_p^1$ ,  $\tau_2$  of  $b_1^1b_2^1\cdots b_q^1$  and  $\sigma_k$  of  $(1+m_k)^11^{d-m_k-1}$  for all  $k=1,\dots,l$ ;
- The subgroup  $\langle\tau_1,\tau_2,\sigma_1,\dots,\sigma_l\rangle$  of  $S_d$  acts transitively on  $\{1,2,\dots,d\}$ .

The following lemma will be useful later, which follows from the Riemann existence theorem and the argument in the first three paragraphs of this subsection.

**Lemma 2.4.** *For each  $m \geq 0$ , proving Theorem 2.3 is equivalent to proving its variant where  $\alpha$  is primitive. We call the latter the primitive version of Theorem 2.3.*

We shall prove Theorem 2.3 and its primitive version simultaneously by induction on  $m = p + q - 2$  in the sequel of this section. The proof will be divided into two parts regarding  $l = 1$  or  $l > 1$ .

Without loss of generality, we may assume the residue vector satisfies the following order assumption:

**Order Assumption (OA)**  $a_1 \leq a_2 \leq \cdots \leq a_p$ ,  $b_1 \geq \cdots \geq b_q$  and  $1 \leq p \leq q$ .

## 2.2 Branch data with three partitions

For the completeness of the manuscript, we include in this subsection the proof of the well known case where  $l = 1$  of Theorem 2.3 (see [2]), whose strategy we also use while proving in Subsection 2.3 the  $l > 1$  case of the theorem.

Here we first make a recall of the case  $l = 1$  of Theorem 2.3.

**Proposition 2.5** (Case  $l = 1$  of Theorem 2.3). *Let  $\alpha = (a_1, \dots, a_p, -b_1, \dots, -b_q)$  be a residue vector and  $\lambda = (m)$  be a partition of  $m = p + q - 2 > 0$  such that  $\deg \alpha > \text{wt}(\lambda) = m$ . Then there exist three permutations  $\tau_1, \tau_2, \sigma_1$  in  $S_d$  satisfying the following three properties:*

- $\tau_1\tau_2\sigma_1 = e$ ;
- $\tau_1$  has the type of  $a_1^1a_2^1\cdots a_p^1$ ,  $\tau_2$  of  $b_1^1b_2^1\cdots b_q^1$  and  $\sigma_1$  of  $(1+m)^11^{d-m-1}$ ;
- The subgroup  $\langle\tau_1,\tau_2,\sigma_1\rangle$  of  $S_d$  acts transitively on  $\{1,2,\dots,d\}$ .

**Lemma 2.6.** *Let  $\alpha = (a_1, \dots, a_p, -b_1, \dots, -b_q)$  be a residue vector with  $\deg \alpha > m = p + q - 2$ . If  $m > 0$ , we have  $a_p > b_q$ .*

*Proof.* If  $a_p < b_q$ , then it follows from the order assumption (OA) that  $\sum_{i=1}^p a_i < \sum_{j=1}^q b_j$ , contradicting the definition of residue vector. If  $a_p - b_q = 0$ , then by OA and  $m = p + q - 2 > 0$  we have  $p = q \geq 2$  and  $a_i = b_j$  for all  $i = 1, 2, \dots, p$  and  $j = 1, 2, \dots, q$ . Since  $p = \deg \alpha > m = p + q - 2 = 2p - 2 = 2q - 2$ , we obtain  $p = q = 1$ , which contradicts  $p + q > 2$ .  $\square$

To prove Proposition 2.5, we need to use the following lemma, where we propose a new concept, called *contraction of a residue vector*.

**Lemma 2.7.** *Under the assumptions of Proposition 2.5, there exist  $i_0 \in \{1, 2, \dots, p\}$  and  $j_0 \in \{1, 2, \dots, q\}$  such that*

$$\widehat{\alpha} = \left( a_1, \dots, a_{i_0-1}, a_{i_0} - b_{j_0}, a_{i_0+1}, \dots, a_p, -b_1, \dots, \widehat{(-b_{j_0})}, \dots, -b_q \right)$$

is a residue vector with  $\deg \widehat{\alpha} > m - 1$ , where  $a_{i_0} - b_{j_0} > 0$  and the hat over term  $(-b_{j_0})$  means that  $(-b_{j_0})$  is removed. Note that the number of components of  $\widehat{\alpha}$  is one less than that of  $\alpha$ . We call  $\widehat{\alpha}$  a **contraction** of  $\alpha$ .

*Proof.* Without loss of generality, we assume  $\alpha$  is primitive. If not, we may replace it by a primitive residue vector equivalent to it.

By OA and Lemma 2.6, we have  $q \geq 2$  and  $a_p - b_q > 0$ , and then we obtain another residue vector

$$\beta_1 := (a_1, \dots, a_{p-1}, a_p - b_q, -b_1, \dots, -b_{q-1})$$

with  $(m - 1)$  components. We divide the proof into the following three steps.

Step 1. Assume  $q = 2$ . Then  $m = p + q - 2 = p$  and  $\deg \beta_1 \geq p$ , so  $\beta_1$  is a contraction of  $\alpha$ . We assume that  $q > 2$  in the sequel of the proof.

Step 2. Suppose that  $\beta_1$  is primitive. Then  $\deg \beta_1 > (m - 1) = (p + q - 3)$ , so it gives a contraction of  $\alpha$ . Indeed, if not, then by OA we have

$$2q - 2 \geq p + q - 2 > \deg \beta_1 = -b_q + \sum_{i=1}^p a_i = -b_q + \sum_{j=1}^q b_j = \sum_{j=1}^{q-1} b_j.$$

thus  $b_{q-1} = b_q = 1$  and  $b_{q-2} \leq 2$ . Moreover, if  $b_{q-2} = 2$ , then  $b_1 = \dots = b_{q-2} = 2$ . Therefore

$$m + 1 = p + q - 1 > 1 + \sum_{j=1}^{q-1} b_j = \sum_{j=1}^q b_j = \sum_{i=1}^p a_i = \deg \alpha,$$

which contradicts the assumption on the degree of  $\alpha$ . We assume that  $\beta_1$  is **not** primitive in the left part of the proof.

Step 3. If  $\deg \beta_1 > (m - 1)$ , then we are done. So without loss of generality, suppose  $\deg \beta_1 \leq (m - 1) = (p + q - 3)$ . Let  $D > 1$  be the greatest common divisor of all components of  $\beta_1$ . Then, by OA and the definition of degree, we obtain that

$$2q - 2 \geq p + q - 2 > \deg \beta_1 = \frac{a_p - b_q}{D} + \sum_{i=1}^{p-1} \frac{a_i}{D} = \sum_{j=1}^{q-1} \frac{b_j}{D}.$$

Hence  $b_{q-1} = D$ ,  $D|b_j$  for all  $j = 1, \dots, q-2$ ,  $D|a_i$  for all  $i = 1, \dots, p-1$ , and  $D|(a_p - b_q)$ . Consequently,  $a_p \geq b_q + D > D = b_{q-1}$  and we obtain another residue vector

$$\beta_2 := (a_1, \dots, a_{p-1}, a_p - b_{q-1}, -b_1, \dots, -b_{q-2}, -b_q)$$

with  $(m-1)$  components. If  $\deg \beta_2 > (m-1)$ , then we are done. So suppose  $\deg \beta_2 \leq (m-1)$ . We divide our discussion into two cases:  $\beta_2$  is primitive and otherwise.

- (a) First, suppose that  $\beta_2$  is primitive. Since  $\deg \beta_2 < m$ , we find that  $b_q = 1$  and  $b_{q-2} \leq 2$  by a same argument for  $\beta_1$  in step 2. As a result  $2 \geq b_{q-2} \geq b_{q-1} = D > 1$ , and thus  $b_{q-2} = b_{q-1} = D = 2$ . By the same argument for  $\beta_1$ , we deduce that  $b_1 = \dots = b_{q-1} = D = 2$  and  $a_1, \dots, a_{p-1}$  are even. In particular,  $a_1 - b_q = a_1 - 1 > 0$ . Hence

$$\hat{\alpha} := (a_1 - b_q, a_2, \dots, a_p, -b_1, \dots, -b_{q-1})$$

is a primitive residue vector with  $(m-1)$  components. Moreover,  $\deg \hat{\alpha} = 2q - 2 \geq p + q - 2 = m > m - 1$  and thus  $\hat{\alpha}$  is a contraction of  $\alpha$ .

- (b) Second, if  $\beta_2$  is not primitive. A similar argument as in the  $\beta_1$  not primitive case gives that  $b_q = E$  and  $b_{q-2} = E$  or  $2E$ , where  $E$  is the greatest common divisor of the components of  $\beta_2$ . By OA,  $E = b_q \leq b_{q-1}$ . If  $D = b_{q-1} = b_q = E$ , then  $E | (a_p - b_{q-1})$  implying that  $E$  divides all the components of  $\alpha$ , this contradicts the primitive property of  $\alpha$ . Hence,  $E = b_q < b_{q-1} \leq b_{q-2}$ . Also since  $b_{q-2} = E$  or  $2E$ , we have  $b_{q-2} = 2E$  and  $1 \leq b_{q-2}/b_{q-1} < 2$ . Recall that  $b_{q-1} = D$  and  $D | b_{q-2}$ , so  $D = b_{q-1} = b_{q-2} = 2E = 2b_q$  and  $E | a_p$ , which gives the same contradiction as above.

□

We shall prove Proposition 2.5 by induction on  $m$ . To this end, we need two lemmas.

**Lemma 2.8.** *Let  $\sigma$  and  $\tau$  be two permutations in  $S_d$  for some positive integer  $d$  such that  $\sigma$  is a cycle of length greater than 1 and  $\tau$  has form  $\nu_1 \nu_2 \dots \nu_r$ , where  $\nu_i$ 's are mutually disjoint cycles of length  $d_i$  for  $i = 1, 2, \dots, r$  and  $d_1 + d_2 + \dots + d_r = d$ . We call  $\nu_i$ 's cycle factors of  $\tau$ . Then the following two conditions are equivalent:*

- (i) *The subgroup  $\langle \sigma, \tau \rangle$  of  $S_d$  acts transitively on the set  $\{1, 2, \dots, d\}$ .*
- (ii) *Each cycle factor  $\nu_i$  of  $\tau$  intersects the cycle  $\sigma$  in the sense that the subset of  $\{1, 2, \dots, d\}$  associated with  $\nu_i$  intersects that associated with the cycle  $\sigma$ .*

*Proof.* (i)  $\Rightarrow$  (ii) We prove the intersection by contradiction. Suppose that there exists a cycle factor  $\nu_i$  of  $\tau$  not intersecting  $\sigma$ . Then  $\nu_i$  is a cycle factor of  $\tau\sigma$ . Therefore, the subset associated with  $\nu_i$  forms an orbit under the action of the subgroup  $\langle \tau, \sigma \rangle$  of  $S_d$  on the set  $\{1, 2, \dots, d\}$ . Since this action is transitive, the cycle  $\nu_i$  has length  $d$ . Hence,  $\nu_i$  must intersect  $\sigma$ , contradiction!

(ii)  $\Rightarrow$  (i) For each  $1 \leq i \leq r$ , we choose a number  $x_i$  lying in both  $\nu_i$  and  $\sigma$ . Consider the action of the subgroup  $\langle \tau, \sigma \rangle$  on the set  $\{1, \dots, d\}$ . Then  $x_1, \dots, x_r$  lie in the same orbit of this action by assumption. Moreover, for each  $1 \leq i \leq r$ ,  $x_i$  belongs to the same orbit with any other numbers in the cycle  $\nu_i$ , and the numbers in  $\nu_1, \dots, \nu_r$  exhaust all the numbers in  $\{1, \dots, d\}$ . Therefore, the action has a single orbit. □

**Remark 2.9.** We consider a more general case of the part of (i)  $\Rightarrow$  (ii) in the above lemma by replacing  $\sigma$  by  $\sigma_1, \sigma_2, \dots, \sigma_l$ , all of which are cycles in  $S_d$  of length greater than 1. Suppose that the subgroup  $\langle \sigma_1, \sigma_2, \dots, \sigma_l, \tau \rangle$  of  $S_d$  acts transitively on the set  $\{1, 2, \dots, d\}$ . Then it follows from a similar argument as in the preceding proof that each cycle factor of  $\tau$  must intersect some  $\sigma_k$  for  $1 \leq k \leq l$ . However, for  $l \geq 2$ , the converse fails in general. For example, the subgroup of  $S_4$  generated by (12), (34) and (12)(34) acts on the set  $\{1, 2, 3, 4\}$  with the two orbits of  $\{1, 2\}$  and  $\{3, 4\}$ .

**Lemma 2.10.** Under the assumptions of Proposition 2.5, suppose that there exist permutations  $\tau_1, \tau_2, \sigma_1$  in  $S_d$  such that  $\tau_1 \tau_2 \sigma_1 = e$  and they have types of  $a_1^1 a_2^1 \cdots a_p^1$ ,  $b_1^1 b_2^1 \cdots b_q^1$  and  $(1+m)^1 1^{d-m-1}$ , respectively. Then the subgroup  $\langle \tau_1, \tau_2, \sigma_1 \rangle$  acts transitively on the set  $\{1, 2, \dots, d\}$  if and only if each cycle factor of  $\tau_1$  and  $\tau_2$  intersects the  $(m+1)$ -cycle  $\sigma_1$ .

*Proof.* Since  $\tau_1 \tau_2 \sigma_1 = e$ , the following three subgroups coincide with each other:

$$\langle \tau_1, \sigma_1 \rangle = \langle \tau_1, \tau_2, \sigma_1 \rangle = \langle \tau_2, \sigma_1 \rangle.$$

The result follows from Lemma 2.8.  $\square$

Now we give the proof of Proposition 2.5.

*Proof.* We argue by induction on  $m = p + q - 2 \geq 0$ . It holds trivially as  $m = 0$ , which is equivalent to  $p = q = 1$ . Assume  $p + q > 2$  in what follows. By Lemma 2.7, there exists a contraction  $\widehat{\alpha}$  of  $\alpha$ . Without loss of generality, we assume that the contraction  $\widehat{\alpha}$  has the form

$$\widehat{\alpha} = (a_1, \dots, a_{p-1}, a_p - b_q, -b_1, \dots, -b_{q-1}).$$

By the induction hypothesis, there exist in  $S_{d-b_q}$  a permutation  $\tau_2 = \nu_1 \nu_2 \cdots \nu_{q-1}$  of type  $b_1^1 b_2^1 \cdots b_{q-1}^1$  and a cycle  $\sigma_1$  of length  $(p+q-2)$  such that  $\tau_2 \sigma_1 = \nu_1 \nu_2 \cdots \nu_{q-1} \sigma_1$  has type of  $a_1^1 \cdots a_{p-1}^1 (a_p - b_q)^1$  and the subgroup generated by  $\tau_2$  and  $\sigma_1$  acts transitively on the set  $\{1, 2, \dots, d - b_p\}$ , where  $\nu_j$  is a cycle factor of length  $b_j$  of  $\tau_2$  for all  $j = 1, 2, \dots, q-1$ . In addition, by the induction hypothesis,  $\tau_2 \sigma_1$  has the form

$$\nu_1 \nu_2 \cdots \nu_{q-1} \sigma_1 = \mu_1 \mu_2 \cdots \mu_p$$

where  $\mu_i$ 's are mutually disjoint cycles for  $1 \leq i \leq p$ , the length of  $\mu_p$  is  $(a_p - b_q)$  and  $\mu_k$  has the length  $a_k$  for  $1 \leq k \leq p-1$ .

By Lemma 2.10, we can choose an integer  $1 \leq x \leq (d - b_q)$  lying in both  $\mu_p$  and  $\sigma_1$ . Choose in  $S_d$  a cycle  $\nu_q$  of length  $b_q$  such that  $\nu_q$  does not intersect  $\nu_j$  for all  $j = 1, \dots, q-1$ , for example  $\nu_q = (d, d-1, \dots, d-b_q+1)$  and pick an integer  $y$  in  $\nu_q$ . Then  $\nu_1 \nu_2 \cdots \nu_{q-1} \nu_q$  has type  $b_1^1 b_2^1 \cdots b_q^1$  and  $\widetilde{\sigma}_1 := \sigma_1(x, y)$  is a cycle of length  $(p+q-1)$ . The subgroup  $\langle \nu_1 \nu_2 \cdots \nu_{q-1} \nu_q, \widetilde{\sigma}_1 \rangle$  of  $S_d$  acts transitively on the set  $\{1, 2, \dots, d\}$  by Lemma 2.10. We also observe that  $\widetilde{\mu}_p := \nu_q \mu_p(x, y)$  is a cycle of length  $a_p$  since  $\nu_q$  does not

intersect  $\mu_p$  and  $x$  and  $y$  lie in  $\mu_p$  and  $\nu_q$ , respectively. Moreover, since  $\nu_q$  does not intersect  $\mu_i$  for all  $1 \leq i \leq p-1$ ,  $\tilde{\mu}_p$  does not intersect  $\mu_i$  for all  $1 \leq i \leq p-1$ . So we have

$$\begin{aligned} (\nu_1 \nu_2 \cdots \nu_{q-1} \nu_q) \widetilde{\sigma}_1 &= \nu_q (\nu_1 \nu_2 \cdots \nu_{q-1} \sigma_1)(x, y) = \nu_q (\mu_1 \mu_2 \cdots \mu_p)(x, y) \\ &= (\mu_1 \mu_2 \cdots \mu_{p-1}) (\nu_q \mu_p(x, y)) = \mu_1 \mu_2 \cdots \mu_{p-1} \tilde{\mu}_p. \end{aligned}$$

Hence, we see that the three permutations  $(\mu_1 \mu_2 \cdots \mu_{p-1} \tilde{\mu}_p)^{-1}$ ,  $\nu_1 \nu_2 \cdots \nu_{q-1} \nu_q$  and  $\widetilde{\sigma}_1$  in  $S_d$  satisfy the three properties listed in Proposition 2.5.  $\square$

### 2.3 Branch data with more than three partitions

We prove Case  $l \geq 2$  of Theorem 2.3. At first we deal with residue vectors with components only  $\pm 1$ .

**Proposition 2.11.** *Let  $D$  be a positive integer. Assume  $\alpha = (\underbrace{1, 1, \dots, 1}_{D+1}, \underbrace{-1, -1, \dots, -1}_{D+1})$ .*

*Then for each partition  $\lambda = (m_1, m_2, \dots, m_l)$  of  $2D$  such that  $\text{wt}(\lambda) < \deg \alpha = D + 1$ , there exist  $l$  permutations  $\sigma_1, \dots, \sigma_l$  in  $S_{D+1}$  such that the following properties hold*

- $\sigma_1 \cdots \sigma_l = e$ ;
- $\sigma_j$ 's are cycles of length  $(m_j + 1)$ ;
- $\langle \sigma_1, \dots, \sigma_l \rangle$  acts transitively on the set  $\{1, \dots, D + 1\}$ .

*Proof.* It is easy to see that  $l \geq 2$ . We divide the proof by considering three cases.

Case 1 If  $l = 2$ , we know that  $m_1 = m_2 = D$  since  $2D = m_1 + m_2$  and  $m_1, m_2 \leq D$ . Then we are done by choosing

$$\sigma_1 = (1, 2, \dots, D + 1), \quad \sigma_2 = \sigma_1^{-1}.$$

Case 2 If  $l = 3$ . Since  $m_1, m_2, m_3 \leq D$  and  $m_1 + m_2 + m_3 = 2D$ , We have  $m_1 + m_2 \geq D$ . Choosing

$$\begin{aligned} \sigma_1 &= (1, 2, \dots, m_1 + 1) \quad \text{and} \\ \sigma_2 &= (1, \underbrace{m_1 + 1, m_1, m_1 - 1, \dots, m_1 + m_3 - D + 2}_{m_1 + m_2 - D}, \underbrace{m_1 + 2, m_1 + 3, \dots, D + 1}_{D - m_1}), \end{aligned}$$

we obtain

$$\sigma_1 \sigma_2 = \begin{cases} (\underbrace{m_1 + 2, m_1 + 3, \dots, D + 1}_{D - m_1}, \underbrace{2, 3, \dots, m_1 + m_3 - D + 2}_{m_1 + m_3 - D + 1}), & \text{if } m_3 < D, \\ (\underbrace{m_1 + 2, m_1 + 3, \dots, D + 1}_{D - m_1}, \underbrace{2, 3, \dots, m_1 + 1, 1}_{m_1 + 1}), & \text{if } m_3 = D. \end{cases}$$

Then the permutations  $\sigma_1, \sigma_2$  and  $(\sigma_1 \sigma_2)^{-1}$  satisfy the three properties.

Case 3 Suppose  $l > 3$ . Since  $m_1, \dots, m_l \leq D$ , we can choose  $1 < r \leq l$  such that  $m_1 + m_2 + \dots + m_{r-1} \leq D$  and  $m_1 + m_2 + \dots + m_r > D$ .

Subcase 3.1 Suppose that  $r < l$ . Choosing

$$\begin{aligned}\sigma_1 &= (1, 2, \dots, m_1 + 1), \\ \sigma_2 &= (m_1 + 1, m_1 + 2, \dots, m_1 + m_2 + 1), \\ &\dots \\ \sigma_{r-1} &= (m_1 + \dots + m_{r-2} + 1, \dots, m_1 + \dots + m_{r-1} + 1),\end{aligned}$$

we obtain

$$\tau_1 := \sigma_1 \sigma_2 \dots \sigma_{r-1} = (1, 2, 3, \dots, m_1 + \dots + m_{r-1} + 1)$$

By Case 2, there exist two cycles  $\tau_2$  and  $\tau_3$  which have length  $1 + m_r$  and  $m_{r+1} + \dots + m_l + 1 < D + 1$ , respectively, such that  $\tau_1 \tau_2 \tau_3 = e$ . As the construction of  $\sigma_1, \dots, \sigma_{r-1}$ , we can find  $\sigma_{r+1}, \sigma_{r+2}, \dots, \sigma_l$  directly such that  $\sigma_j$  has the type of  $(1 + m_j)^{1 + D - m_j}$  for  $r + 1 \leq j \leq l$  and  $\sigma_{r+1} \dots \sigma_l = \tau_3$ . Therefore the  $l$  cycles of  $\sigma_1, \dots, \sigma_{r-1}, \sigma_r := \tau_2, \sigma_{r+1}, \dots, \sigma_l$  satisfy the three properties.

Subcase 3.2 Suppose  $r = l > 3$ . Since  $m_1 + \dots + m_l = 2D$  and  $\max(m_1, \dots, m_l) \leq D$ , we have  $m_1 + \dots + m_{l-1} = m_l = D$ . Then the problem can be reduced to Case 1 by a similar argument as above.

□

To complete the proof of Theorem 2.3, we need the following lemma and its two corollaries.

**Lemma 2.12.** *Let  $\Gamma$  be a subgroup of  $S_d = S_{\{1, 2, \dots, d\}}$  for some integer  $d > 1$  and  $\theta \in S_d$  a cycle of length greater than 1. Assume that the subgroup  $G$  generated by  $\Gamma$  and  $\theta$  acts transitively on the set  $\{1, 2, \dots, d\}$ . Then, for each number  $1 \leq x \leq d$  not contained in  $\theta$ , the  $\Gamma$ -orbit  $\Gamma x$  of  $x$  intersects  $\theta$ .*

*Proof.* We argue by contradiction. Suppose that the orbit  $\Gamma x$  does not intersect  $\theta$ . Take an arbitrary permutation  $\xi$  in  $G$ . We can express it as

$$\xi = \pi_1 \pi_2 \dots \pi_s$$

where either  $\pi_i = \theta$  or  $\pi_i \in \Gamma$ . Let  $\xi'$  be the permutation obtained from the product  $\pi_1 \pi_2 \dots \pi_s$  by removing all those  $\pi_i$ 's satisfying  $\pi_i = \theta$ . Then  $\xi' \in \Gamma$ . Since each number not contained in  $\theta$  is a fixed point of  $\theta$ , by the hypothesis of the contradiction argument, we find that  $\xi(x) = \xi'(x)$  is not contained in  $\theta$ . Since  $\xi \in G$  has been chosen arbitrarily, the orbit  $Gx$  does not intersect  $\theta$ , which contradicts that  $G$  acts transitively on  $\{1, 2, \dots, d\}$ . □

As an application of the above lemma, we have

**Corollary 2.13.** *Let  $\gamma_1, \gamma_2, \dots, \gamma_l, \theta$  be  $(l+1)$  permutations in  $S_{d-1}$  for some integer  $d > 2$  and  $\theta = (x_1, x_2, \dots, x_n)$  a cycle in  $S_{d-1}$  of length  $n > 1$ . Suppose that the subgroup  $\langle \gamma_1, \gamma_2, \dots, \gamma_l, \theta \rangle$  acts transitively on the set  $\{1, 2, \dots, d-1\}$ . Then so does the subgroup  $\langle \gamma_1, \gamma_2, \dots, \gamma_l, \tilde{\theta} \rangle$  on  $\{1, 2, \dots, d\}$ , where  $\tilde{\theta} := (x_1, \dots, x_n, d)$  is a cycle of length  $(n+1)$  in  $S_d$ .*

*Proof.* By Lemma 2.12, we can see that for each number  $1 \leq x \leq d-1$  not contained in  $\theta$ , there exists  $\gamma \in \Gamma := \langle \gamma_1, \dots, \gamma_l \rangle$  such that  $\gamma(x)$  is contained in  $\theta$ . Hence, the action of  $\langle \gamma_1, \gamma_2, \dots, \gamma_l, \tilde{\theta} \rangle$  on  $\{1, 2, \dots, d\}$  has only one orbit.  $\square$

Similarly, we obtain

**Corollary 2.14.** *Let  $\gamma_1, \gamma_2, \dots, \gamma_l, \theta$  be permutations in  $S_d$  for some integer  $d > 2$  and  $\theta = (x_1, x_2, \dots, x_n)$  a cycle of length  $1 < n < d$ . Suppose that the subgroup  $\langle \gamma_1, \gamma_2, \dots, \gamma_l, \theta \rangle$  acts transitively on the set  $\{1, 2, \dots, d\}$ . Then so does the subgroup  $\langle \gamma_1, \gamma_2, \dots, \gamma_l, \tilde{\theta} \rangle$  of  $S_d$  on the set  $\{1, 2, \dots, d\}$ , where  $\tilde{\theta} = (x_1, \dots, x_n, y) \in S_d$  is a cycle of length  $(n+1)$  with  $y \in \{1, 2, \dots, d\} \setminus \{x_1, \dots, x_n\}$ .*

Now we arrive at proving the case  $l \geq 2$  of Theorem 2.3.

*Proof.* By Lemma 2.4 we could also assume that the residue vector  $\alpha = (a_1, \dots, a_p, -b_1, \dots, -b_q)$  is primitive so that  $\deg \alpha = d = a_1 + \dots + a_p = b_1 + \dots + b_q$ .

**Part I** Suppose  $d \geq m+1 = p+q-1$ . By Proposition 2.5 there exist  $\tau_1, \tau_2, \sigma$  such that

- (1)  $\tau_1 \tau_2 \sigma = e$ ;
- (2)  $\tau_1$  has type of  $a_1^1 a_2^1 \dots a_p^1$ ,  $\tau_2$  of  $b_1^1 b_2^1 \dots b_q^1$ ,  $\sigma$  of  $(1+m)^1 1^{d-m-1}$ ;
- (3) the subgroup  $\langle \tau_1, \tau_2, \sigma \rangle$  acts transitively on  $\{1, 2, \dots, d\}$ .

Assume that  $\sigma = (1, 2, \dots, m+1)$  for simplicity of notion. We are done by choosing

$$\begin{aligned} \sigma_1 &= (1, 2, \dots, m_1 + 1), \\ \sigma_2 &= (m_1 + 1, m_1 + 2, \dots, m_1 + m_2 + 1), \\ &\dots\dots \\ \sigma_l &= (m_1 + \dots + m_{l-1} + 1, m_1 + \dots + m_{l-1} + 2, \dots, m_1 + \dots + m_l + 1), \end{aligned}$$

**Part II** Suppose  $d = \deg \alpha \leq m = p+q-2$ . We first reduce the problem to the two cases that  $l = 2$  and  $l = 3$ , then we prove these two cases by using the contraction argument and the induction argument similar as the proof of Proposition 2.5. The details given as follows form the left part of this section.

By OA, we have that  $d = \sum_{j=1}^q b_j \geq q \geq \frac{p+q}{2}$ . Since  $d = \deg \alpha > \text{wt}(\lambda) = \max(m_1, \dots, m_l)$ , the partition  $\lambda$  of  $m = (p+q-2)$  has at least two components, i.e.  $l > 1$ .

At first we show that

**CLAIM 1:** *the problem can be reduced to the two cases where the partitions of  $m$  have two and three components, respectively.*

**PROOF OF CLAIM 1**

Suppose that Part II holds for each partition of  $m$  which has three components and has weight less than  $d$ . Then we shall prove that so does Part II for each partition  $(m_1, \dots, m_l)$  of  $m$  such that  $l > 3$  and its weight is less than  $d$ . To this end, since  $m_1, \dots, m_l < d$ , we can choose  $1 < r \leq l$  such that

$$m_1 + \dots + m_{r-1} < d \quad \text{and} \quad m_1 + \dots + m_r \geq d.$$

We shall define a new partition, called  $\lambda'$ , with three components as following.

- Suppose  $r < l$ . Then we consider the partition  $\lambda' := (m'_1, m'_2, m'_3)$  of  $m$ , where  $m'_1, m'_2, m'_3$  are defined by

$$m'_1 := m_1 + \dots + m_{r-1} < d, \quad m'_2 := m_r < d, \quad m'_3 := m_{r+1} + \dots + m_l.$$

Moreover, since  $d \geq \frac{p+q}{2}$ , we observe that

$$m'_3 = p + q - 2 - (m'_1 + m'_2) < (p + q - d) - 1 \leq \frac{p+q}{2} - 1 \leq d - 1.$$

- Suppose  $r = l > 3$ . Then we choose the partition  $\lambda' := (m_1 + \dots + m_{l-2}, m_{l-1}, m_l)$  of  $m$ , which has weight less than  $d$ .

Since the partition  $\lambda'$  has weight less than  $d$ , and we have assumed the validity of Case  $l = 3$  for  $\alpha$  and  $\lambda'$ , we can find in  $S_d$  the following five permutations  $\tau_1, \tau_2, \sigma'_1, \sigma'_2, \sigma'_3$  which satisfy the three properties. By a similar construction as Case 3 in the proof of Proposition 2.11, we know that the proposition holds for the partition  $(m_1, \dots, m_l)$ . Therefore, we have justified the claim.

We always assume  $l = 2$  or  $3$  in the left part of the proof.

Recall that OA states that  $1 \leq p \leq q, a_1 \leq a_2 \leq \dots \leq a_p, b_1 \geq b_2 \geq \dots \geq b_q$ . Without loss of generality, we further assume  $m_1 \leq m_2 \leq \dots \leq m_l$  for the partition  $\lambda$ . We call these two assumptions OA in the sequel by an abuse of notation.

By OA, we can see that if  $a_p = 1$ , then  $a_1 = a_2 = \dots = a_p = b_1 = b_2 = \dots = b_q = 1$  and we are done by Proposition 2.11. We may assume that  $a_p > 1$  in the left part of the proof. Since  $d \leq m$ , we have  $\sum_{j=1}^q b_j \leq p + q - 2 \leq 2q - 2$ . Then, we have  $b_{q-1} = b_q = 1$ .

Since  $m_1 \leq \dots \leq m_l$  and  $a_p > 1$ ,  $\deg \alpha > \text{wt}(\lambda)$  and  $l = 2$  or  $3$ , we observe that  $m_l$  is always greater than 1 except when  $\alpha = (1, 2, -1, -1, -1)$  and  $\lambda = (1, 1, 1)$ , for which Part II holds trivially. We may assume that  $m_l > 1$  in the left part of the proof. In order to do induction on  $m$ , we choose the partition

$$\lambda_1 := (m_1, \dots, m_{l-1}, (m_l - 1))$$

of  $(m - 1)$  and the residue vector

$$\widehat{\alpha} = (a_1, a_2, \dots, a_{p-1}, a_p - b_q, -b_1, -b_2, \dots, -b_{q-1})$$

with  $(p + q - 1) = (m + 1)$  components. Then, since  $b_{q-1} = b_q = 1$ ,  $\widehat{\alpha}$  is primitive and

$$\deg \widehat{\alpha} = -1 + \deg \alpha.$$

Then we make the following

CLAIM 2:  $\text{wt}(\lambda_1) < \deg \widehat{\alpha}$ . Hence we may think of  $(\widehat{\alpha}, \lambda_1)$  as a contraction of  $(\alpha, \lambda)$  and we shall use the former to do induction argument.

PROOF OF CLAIM 2

- (i) If  $m_1 \leq \dots \leq m_{l-1} = m_l$ , then  $\text{wt}(\lambda_1) = m_{l-1} = m_l$  and  $2m_l \leq m = p + q - 2 \leq 2q - 2$ , i.e.  $m_l \leq q - 1$ . If  $\deg \widehat{\alpha} > q - 1$ , then we are done. Assume  $\deg \widehat{\alpha} = q - 1$ , which implies that  $b_1 = \dots = b_q = 1$  and  $p < q$  since  $a_p > 1$ . Then  $2m_l \leq m = p + q - 2 < 2q - 2$  and  $m_l < q - 1$ . Hence  $\deg \widehat{\alpha} = q - 1 > m_l = \text{wt}(\lambda_1)$ .
- (ii) If  $m_{l-1} < m_l$ , then  $\text{wt}(\lambda_1) = m_l - 1 = \text{wt}(\lambda) - 1 < \deg \alpha - 1 = \deg \widehat{\alpha}$ . The claim is proved.

Then in the left part of the proof we use the induction on  $m$  to prove that Part II of Theorem 2.3 holds provided that  $l$  equals either 2 or 3. We observe that the initial case of  $m = 2$  holds trivially, where  $\alpha = (1, 1, -1, -1)$  and  $\lambda = (1, 1)$ .

– Suppose  $l = 2$ . We recall our setting as follows. Take a primitive residue vector

$$\alpha = (a_1, \dots, a_p, -b_1, \dots, -b_q)$$

and a partition  $\lambda = (m_1, m_2)$  of  $m = p + q - 2 \geq 3$  such that  $d = \deg \alpha \leq m$ , and  $\alpha$  and  $\lambda$  satisfy OA. Then we have  $2 \leq m_2 = \text{wt}(\lambda) < d$  and  $b_{q-1} = b_q = 1$ .

Then, by claim 2, we could take another primitive residue vector

$$\widehat{\alpha} = (a_1, a_2, \dots, a_{p-1}, a_p - b_q, -b_1, -b_2, \dots, -b_{q-1})$$

and another partition  $\lambda_1 = (m_1, m_2 - 1)$  such that  $\text{wt}(\lambda_1) < \deg \widehat{\alpha} = d - 1$ . By the induction hypothesis, there exist in  $S_{d-1} = S_{\{1, 2, \dots, d-1\}}$  a permutation of type  $b_1^1 \dots b_{q-1}^1$ , called  $\tau_2$ , and two cycles  $\sigma_1, \sigma_2$  of length  $(1 + m_1), m_2$ , respectively, such that the subgroup  $\langle \sigma_1, \sigma_2, \tau_2 \rangle$  of  $S_{d-1}$  acts transitively on  $\{1, 2, \dots, d - 1\}$  and the permutation

$$\tau_1 := \tau_2 \sigma_1 \sigma_2$$

has the type of  $a_1^1 \dots a_{p-1}^1 (a_p - b_q)^1$ . We re-express  $\tau_1$  by  $\tau_1 = \mu_1 \dots \mu_p$  such that  $\mu_j$ 's are its cycle factors and  $\mu_j$  has length  $a_j$  for  $1 \leq j < p$  and  $\mu_p$  has length  $(a_p - b_q)$ . Since  $b_q = 1$ , we can think of  $\tau_2$  as a permutation in  $S_d$  with the type of  $b_1^1 \dots b_{q-1}^1 b_q^1$ . By Remark 2.9, the cycle factor  $\mu_p$  intersects either  $\sigma_1$  or  $\sigma_2$ . We shall divide the left part of the proof of Case  $l = 2$  into the following two steps.

Step 2.1 Suppose that  $\sigma_2$  intersects  $\mu_p$ . Then we pick a number  $x$  in both  $\sigma_2$  and  $\mu_p$  and define  $\widetilde{\sigma}_2 := \sigma_2(x, d)$  and  $\widetilde{\mu}_p := \mu_p(x, d)$ . Then, since  $b_q = 1$ ,  $\widetilde{\sigma}_2$  and  $\widetilde{\mu}_p$  are cycles in  $S_d$  of length  $(1 + m_2)$  and  $a_p$ , respectively. Moreover, we have

$$\tau_2 \sigma_1 \widetilde{\sigma}_2 = \mu_1 \mu_2 \dots \mu_{p-1} \widetilde{\mu}_p.$$

Since  $\langle \tau_2, \sigma_1, \sigma_2 \rangle$  acts transitively on the set  $\{1, 2, \dots, d-1\}$ , the action of  $\langle \tau_2, \sigma_1, \tilde{\sigma}_2 \rangle$  on the set  $\{1, 2, \dots, d\}$  has a single orbit by Corollary 2.13. Therefore, the following four permutations  $(\mu_1 \mu_2 \cdots \mu_{p-1} \tilde{\mu}_p)^{-1}$ ,  $\tau_2$ ,  $\sigma_1$ ,  $\tilde{\sigma}_2$  satisfy the three properties.

Step 2.2 Suppose that  $\sigma_2$  does not intersect  $\mu_p$ . Then  $\sigma_1$  intersects  $\mu_p$ . Choose a number  $x$  in both  $\sigma_1$  and  $\mu_p$ . Then

$$\begin{aligned} \tau_2 \sigma_1 \sigma_2(x, d) &= \mu_1 \cdots \mu_p(x, d) \\ &= \mu_1 \cdots \mu_{p-1} \tilde{\mu}_p \quad \text{where} \quad \tilde{\mu}_p := \mu_p(x, d) \\ &=: \tilde{\tau}_1 \end{aligned}$$

where  $\tilde{\tau}_1$  has the type of  $a_1^1 \cdots a_p^1$ . Meanwhile, since  $\sigma_2$  does not intersect  $(x, d)$ , we have

$$\tilde{\tau}_1 = \tau_2 \sigma_1 \sigma_2(x, d) = \tau_2 \sigma_1(x, d) \sigma_2 =: \tau_2 \sigma'_1 \sigma_2,$$

where  $\sigma'_1 := \sigma_1(x, d)$  is a cycle of length  $(m_1 + 2)$ . Observe that  $\sigma'_1$  intersects  $\sigma_2$ . For, otherwise,  $\sigma_1$  does not intersect  $\sigma_2$  and there are  $(m_1 + m_2 + 1)$  different numbers appearing in both  $\sigma_1$  and  $\sigma_2$ , lying in  $S_{d-1}$ . Hence, we have

$$(d-1) \geq 1 + m_1 + m_2 = 1 + m \geq 1 + d,$$

contradiction!

Take the smallest positive integer  $s$  such that  $y := (\sigma'_1)^{-s}(x)$  is contained in  $\sigma_2$ . Since  $x$  is not contained in  $\sigma_2$ , by the minimal property of  $s$ ,  $(\sigma'_1)(y)$  is not contained in  $\sigma_2$ . We can rewrite  $\sigma'_1$  as

$$\sigma'_1 = (x_1, x_2, \dots, x_{m_1}, y, \sigma'_1(y)) = \tilde{\sigma}_1(y, \sigma'_1(y)) \quad \text{with} \quad \tilde{\sigma}_1 := (x_1, \dots, x_{m_1}, y).$$

Then we have

$$\sigma'_1 \sigma_2 = \tilde{\sigma}_1(y, \sigma'_1(y)) \sigma_2 = \tilde{\sigma}_1 \tilde{\sigma}_2 \quad \text{where} \quad \tilde{\sigma}_2 := (y, \sigma'_1(y)) \sigma_2,$$

and

$$\tilde{\tau}_1 = \tau_2 \sigma'_1 \sigma_2 = \tau_2 \tilde{\sigma}_1 \tilde{\sigma}_2,$$

where  $\tilde{\tau}_1$  and  $\tau_2$  have types of  $a_1^1 \cdots a_p^1$  and  $b_1^1 \cdots b_{q-1}^1 b_q^1$ , respectively, and the two cycles  $\tilde{\sigma}_1$  and  $\tilde{\sigma}_2$  have lengths of  $1 + m_1$  and  $1 + m_2$ , respectively. Since  $\langle \tau_2, \tilde{\tau}_1, \sigma_2 \rangle = \langle \tau_2, \sigma'_1, \sigma_2 \rangle$  acts transitively on the set  $\{1, 2, \dots, d\}$  by Corollary 2.13, so does  $\langle \tau_2, \tilde{\tau}_1, \tilde{\sigma}_2 \rangle$  on the same set by Corollary 2.14. Therefore, the following four permutations  $(\tilde{\tau}_1)^{-1}$ ,  $\tau_2$ ,  $\tilde{\sigma}_1$ ,  $\tilde{\sigma}_2$  satisfy the three properties.

- Suppose  $l = 3$ . We may further assume that  $m_1 + m_2 \geq d$ . Otherwise, replacing  $\lambda$  by  $\lambda' = (m_1 + m_2, m_3)$  and using the similar reduction argument as above, we can reduce the problem to the known case of  $l = 2$ .

By the induction hypothesis, there exist in  $S_{d-1}$  a permutation  $\tau_2$  of type  $b_1^1 \cdots b_{q-1}^1$  and three cycles  $\sigma_1, \sigma_2, \sigma_3$  of length  $1 + m_1, 1 + m_2, m_3$ , respectively, such that the subgroup  $\langle \sigma_1, \sigma_2, \sigma_3, \tau_2 \rangle$  acts transitively on the set  $\{1, 2, \dots, d-1\}$  and

$$\tau_2 \sigma_1 \sigma_2 \sigma_3 = \tau_1 = \mu_1 \cdots \mu_p,$$

where  $\tau_1 = \mu_1 \cdots \mu_p$  has type  $a_1^1 \cdots a_{p-1}^1 (a_p - b_q)^1$ ,  $\mu_j$ 's are the cycle factors of  $\tau_1$  and  $\mu_j$  has length  $a_j$  for  $1 \leq j < p$  and  $\mu_p$  has length  $(a_p - b_q) = (a_p - 1)$ . Since  $m_1 + m_2 \geq d$ , by OA, any two of the three cycles  $\sigma_1, \sigma_2, \sigma_3$  in  $S_{d-1}$  intersect since  $1 + m_i + m_j > d > d - 1$  for  $1 \leq i \neq j \leq 3$ . We divide the left part of the proof into the following three steps.

Step 3.1 If  $\mu_p$  intersects  $\sigma_3$ , we are done by a similar argument as in step 2.1.

Step 3.2 If  $\mu_p$  intersects  $\sigma_2$  but does not intersect  $\sigma_3$ , the proof goes through as Step 2.2 since  $\sigma_2$  intersects  $\sigma_3$ .

Step 3.3 Suppose that neither  $\sigma_2$  nor  $\sigma_3$  intersects  $\mu_p$ . Then, by Remark 2.9,  $\mu_p$  intersects  $\sigma_1$  since  $\langle \tau_1, \sigma_1, \sigma_2, \sigma_3 \rangle$  acts transitively on the set  $\{1, 2, \dots, d-1\}$ . Choosing a number  $x$  in both  $\mu_p$  and  $\sigma_1$  and denoting  $\tilde{\mu}_p := \mu_p(x, d)$  and  $\sigma'_1 := \sigma_1(x, d)$ , we obtain

$$\tau_2 \sigma'_1 \sigma_2 \sigma_3 = \tau_2 \sigma_1 \sigma_2 \sigma_3(x, d) = \mu_1 \cdots \mu_{p-1} \tilde{\mu}_p =: \tilde{\tau}_1$$

where  $\sigma'_1$  is a cycle of length  $(m_1 + 2)$ ,  $\tilde{\mu}_p$  is a cycle factor of  $\tilde{\tau}_1$ , and  $\tilde{\tau}_1$  has the type of  $a_1^1 \cdots a_{p-1}^1 a_p^1$ .

Step 3.3.A Suppose that  $\sigma_2$  contains a number which is not contained in  $\sigma_3$ . Then we could find in  $S_d$  the three cycles of  $\tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{\sigma}_3$  with length  $1 + m_1, 1 + m_2, 1 + m_3$ , respectively, such that

$$\tau_2 \tilde{\sigma}_1 \tilde{\sigma}_2 \tilde{\sigma}_3 = \tilde{\tau}_1$$

and  $\langle \tau_2, \tilde{\tau}_1, \tilde{\sigma}_1, \tilde{\sigma}_3 \rangle$  acts transitively on  $\{1, 2, \dots, d\}$ . Indeed, rewrite  $\sigma'_1$  as

$$\sigma'_1 = \tilde{\sigma}_1(b_1, z_1)$$

where  $b_1$  lies in the intersection of  $\sigma_1$  and  $\sigma_2$ , the number  $z_1$  is contained in  $\sigma'_1$  but not in  $\sigma_2$ , and  $\tilde{\sigma}_1$  is a cycle of length  $(1 + m_1)$ . Choose  $\sigma'_2 := (b_1, z_1)\sigma_2$  and rewrite it as

$$\sigma'_2 = \tilde{\sigma}_2(b_2, z_2),$$

where  $b_2$  lies in both  $\sigma'_2$  and  $\sigma_3$ , the number  $z_2$  is contained in  $\sigma'_2$  but not in  $\sigma_3$ ,  $\tilde{\sigma}_2$  is a  $(1 + m_2)$ -cycle. Hence  $\tilde{\sigma}_3 := (b_2, z_2)\sigma_3$  is a  $(1 + m_3)$ -cycle and the equality  $\tau_2 \tilde{\sigma}_1 \tilde{\sigma}_2 \tilde{\sigma}_3 = \tilde{\tau}_1$  holds. Then all the following subgroups

$$\langle \tau_2, \sigma'_1, \sigma_2, \sigma_3 \rangle = \langle \tau_2, \tilde{\tau}_1, \sigma_2, \sigma_3 \rangle, \quad \langle \tau_2, \tilde{\tau}_1, \sigma'_2, \sigma_3 \rangle = \langle \tau_2, \tilde{\tau}_1, \tilde{\sigma}_1, \sigma_3 \rangle$$

and  $\langle \tau_2, \tilde{\tau}_1, \tilde{\sigma}_1, \tilde{\sigma}_3 \rangle$  of  $S_d$  act transitively on  $\{1, 2, \dots, d\}$  by Corollaries 2.13 and 2.14. Therefore, the following permutations

$$(\tilde{\tau}_1)^{-1}, \tau_2, \tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{\sigma}_3$$

satisfy the three properties.

Step 3.3.B Suppose that every number in the cycle  $\sigma_2$  is contained in  $\sigma_3$ . Rewrite  $\sigma'_1$  as

$$\sigma'_1 = \tilde{\sigma}_1(b, z)$$

where  $\tilde{\sigma}_1$  is a  $(1 + m_1)$ -cycle,  $b$  lies in both  $\sigma'_1$  and  $\sigma_3$ ,  $z$  is not contained in  $\sigma_3$  and then it is not contained in  $\sigma_2$ , either.

• Suppose that  $b$  is not contained in  $\sigma_2$ . Then

$$\begin{aligned} \tau_2 \sigma'_1 \sigma_2 \sigma_3 &= \tau_2 \tilde{\sigma}_1(b, z) \sigma_2 \sigma_3 \\ &= \tau_2 \tilde{\sigma}_1 \sigma_2((b, z) \sigma_3) \\ &= \tau_2 \tilde{\sigma}_1 \sigma_2 \tilde{\sigma}_3, \end{aligned}$$

where  $\tilde{\sigma}_3 := (b, z) \sigma_3$  is a  $(1 + m_3)$ -cycle. By Corollary 2.14, the subgroup  $\langle \tau_2, \tilde{\tau}_1, \sigma_2, \tilde{\sigma}_3 \rangle$  acts transitively on the set  $\{1, 2, \dots, d\}$  since the subgroup  $\langle \tau_2, \tilde{\tau}_1, \sigma_2, \sigma_3 \rangle = \langle \tau_2, \sigma'_1, \sigma_2, \sigma_3 \rangle$  has the same property. Therefore, the following five permutations  $(\tilde{\tau}_1)^{-1}, \tau_2, \tilde{\sigma}_1, \sigma_2, \tilde{\sigma}_3$  satisfy the three properties.

• Suppose that  $b$  is contained in  $\sigma_2$ . Then we have

$$\tau_2 \sigma'_1 \sigma_2 \sigma_3 = \tau_2 \tilde{\sigma}_1((b, z) \sigma_2) \sigma_3 = \tau_2 \tilde{\sigma}_1 \sigma'_2 \sigma_3 \quad \text{where} \quad \sigma'_2 := (b, z) \sigma_2.$$

Then applying a similar argument in Step 2.2 to  $\sigma'_2$  and  $\sigma_3$ , we obtain the two cycles  $\tilde{\sigma}_2$  and  $\tilde{\sigma}_3$  with length of  $(1 + m_2), (1 + m_3)$ , respectively, such that

$$\sigma'_2 \sigma_3 = \tilde{\sigma}_2 \tilde{\sigma}_3 \quad \text{and} \quad \tau_2 \tilde{\sigma}_1 \tilde{\sigma}_2 \tilde{\sigma}_3 = \tilde{\tau}_1.$$

Moreover, by Corollaries 2.13 and 2.14, the action of  $\langle \tau_2, \tilde{\tau}_1, \tilde{\sigma}_1, \tilde{\sigma}_3 \rangle$  on the set  $\{1, 2, \dots, d\}$  is transitive. Therefore, the following five permutations  $(\tilde{\tau}_1)^{-1}, \tau_2, \tilde{\sigma}_1, \tilde{\sigma}_2, \tilde{\sigma}_3$  satisfy the three properties.

□

### 3 Proof of Theorem 1.3

We prove Theorem 1.3 in this subsection.

*Proof.* Consider the collection

$$\Lambda^* = \{(a_1, \dots, a_p), (b_1, \dots, b_q), (c_1 + 1, 1, \dots, 1), \dots, (c_r + 1, 1, \dots, 1)\}$$

of degree  $d$  and with total branching order  $2d - 2$ . By Theorem 2.3 and the Riemann existence theorem, there exists a rational function  $f$  of degree  $d$  on the Riemann sphere such that

- (i)  $f$  branches over  $0, \infty, \zeta, \dots, \zeta^r$ , where  $\zeta = \exp(2\pi\sqrt{-1}/r)$ .
- (ii) The partitions of the above branch points coincide with  $(a_1, \dots, a_p), (b_1, \dots, b_q), (c_1 + 1, 1, \dots, 1), \dots, (c_r + 1, 1, \dots, 1)$ , respectively.

Then  $f^r$  is the Belyi function as desired. □

## 4 A conjecture

In order to generalize the first part of Boccara's result, we make the following

**Conjecture** *Let  $d, g$  and  $l$  be three positive integers. Suppose that the collection  $\Lambda$  consists of  $l+2$  partitions of  $d$  and has form*

$$\Lambda = \{(a_1, \dots, a_p), (b_1, \dots, b_q), (m_1 + 1, 1, \dots, 1), \dots, (m_l + 1, 1, \dots, 1)\},$$

where  $(m_1, \dots, m_l)$  is a partition of  $p + q - 2 + 2g$ . Then there always exists a rational function on some compact Riemann surface of genus  $g$  with branch data  $\Lambda$ .

**Acknowledgments.** The authors would like to express their deep gratitude to Professor Qing Chen, Professor Mao Sheng and Mr Bo Li for their valuable discussions during the course of this work. The second author is indebted to Mr Weibo Fu for informing him [2, 4, 24].

The first author is supported in part by the National Natural Science Foundation of China (grant no. 11426236). The second author is supported in part by the National Natural Science Foundation of China (grant no. 11571330) and the Fundamental Research Funds for the Central Universities.

## References

- [1] K. Baránski. On realizability of branched coverings on the sphere. *Topol. Appl.* **116** (2001), 279-291.
- [2] G. Boccara. Cycles comme produit de deux permutations de classes données. *Discrete Math.* **58** (1982), 129-142.
- [3] S. Donaldson. *Riemann surfaces*. Oxford University Press, Oxford, 2011.
- [4] A. L. Edmonds, R. S. Kulkarni and R. E. Stong. Realizability of branched coverings of surfaces. *Trans. Amer. Math. Soc.* **282**(2) (1984), 773-790.
- [5] A. Eremenko. Co-axial monodromy. *preprint arXiv:1706.04608v1*, 2017.

- [6] C. L. Ezell. Branch point structure of covering maps onto nonorientable surface. *Trans. Am. Math. Soc.* **243** (1978), 123-133.
- [7] G. Francis. Assembling Riemann surfaces with prescribed boundary curves and branch points. *Illinois J. Math.* **20** (1976), 198-217.
- [8] S. M. Gersten. On branched coverings of the 2-sphere by the 2-sphere. *Proc. Amer. Math. Soc.* **101** (1987), 761-766.
- [9] A. Hurwitz. Über Riemann'sche Flächen mit gegebenen Verzweigungspunkten. *Math. Ann.* **103** (1891), 1-60.
- [10] D. H. Husemoller. Ramified coverings of Riemann surfaces. *Duke Math. J.* **29** (1962), 167-174.
- [11] A. G. Khovanskii and S. Zdravkovska. Branched covers of  $S^2$  and braid groups. *J. Knot Theory Ramif.* **5** (1995), 55-75.
- [12] S. K. Lando and A. K. Zvonkin. *Graphs on Surfaces and their Applications, With an appendix by Don B. Zagier*. Springer-Verlag, Berlin, 2004.
- [13] A. D. Mednykh. Nonequivalent coverings of Riemann surfaces with a given ramification type. *Sib. Math. Zh.* **25** (1984), 120-142.
- [14] A. D. Mednykh. Branched covering of Riemann surfaces whose branch orders coincide with the multiplicities. *Comm. Algebra* **18** (1990), 1517-1533.
- [15] S. Monni, J. S. Song and Y. S. Song. The Hurwitz enumeration problem of branched covers and Hodge integrals. *J. Geom. Phys.* **50** (2004), 223-256.
- [16] A. Okounkov and R. Pandharipande. Gromov-Witten theory, Hurwitz theory and complete cycles. *Ann. Math.* **163** (2006), 517-560.
- [17] F. Pakovich. Solution of Hurwitz problem for Laurent polynomials. *J. Knot Theory Ramif.* **18** (2009), 271-302.
- [18] M. A. Pascali and C. Petronio. Surface branched covers and geometric 2-orbifolds. *Trans. Amer. Math. Soc.* **361** (2009), 5885-5920.
- [19] M. A. Pascali and C. Petronio. Branched covers of the sphere and the prime-degree conjecture. *Annali di Matematica.* **191** (2012), 563-594.
- [20] E. Pervova and C. Petronio. On the existence of branched coverings between surfaces with prescribed branch data, I. *Algebra Geom. Topol.* **6** (2006), 1957-1985.
- [21] E. Pervova and C. Petronio. On the existence of branched coverings between surfaces with prescribed branch data, II. *J. Knot Theory Ramif.* **17** (2008), 787-816.
- [22] D. Singerman. Subgroups of Fuchsian groups and finite permutation groups. *Bull. Lond. Math. Soc.* **2** (1970), 319-323.
- [23] R. Thom. L'équivalence d'une fonction différentiable et d'un polinôme. *Topology* **3**(suppl. 2) (1965), 297-307.
- [24] H. Zheng. Realizability of branched covering of  $S^2$ . *Topol. Appl.* **153** (2006), 2123-2134.