

## Two-generator free Kleinian groups and hyperbolic displacements

İLKER S YÜCE

The log 3 theorem, proved by Culler and Shalen, states that every point in the hyperbolic 3–space  $\mathbb{H}^3$  is moved a distance at least  $\log 3$  by one of the noncommuting isometries  $\xi$  or  $\eta$  of  $\mathbb{H}^3$  provided that  $\xi$  and  $\eta$  generate a torsion-free, discrete group which is not cocompact and contains no parabolic. This theorem lies in the foundations of many techniques that provide lower estimates for the volumes of orientable, closed hyperbolic 3–manifolds whose fundamental groups have no 2–generator subgroup of finite index and, as a consequence, gives insights into the topological properties of these manifolds.

Under the hypotheses of the log 3 theorem, the main result of this paper shows that every point in  $\mathbb{H}^3$  is moved a distance at least  $\log \sqrt{5 + 3\sqrt{2}}$  by one of the isometries  $\xi$ ,  $\eta$  or  $\xi\eta$ .

14E20, 54C40; 46E25, 20C20

### 1 Introduction

Let  $M$  be a closed orientable hyperbolic 3–manifold. Anderson, Canary, Culler and Shalen proved in [3] that the volume of  $M$  is bounded below by 3.08 under the assumptions that the first Betti number of  $M$  is at least 4 and  $\pi_1(M)$  has no subgroup isomorphic to the fundamental group of a genus-two surface. In [6], Culler and Shalen showed that the volume of  $M$  is at least 0.92 provided that the first Betti number of  $M$  is at least 3 and  $\pi_1(M)$  has no two-generator subgroup of finite index. Later Culler, Hersensky and Shalen improved the previous volume estimate to 0.94 in [5]. These deep results are among a number of theorems stated by Culler and Shalen alone [6; 8] and together with their collaborators Agol [2], Anderson and Canary [3] and Hersensky [5] that relate the topology of hyperbolic 3–manifolds to their geometry.

The common denominator in all of the volume estimates listed above is that they are consequences of one of the fundamental results in the study of Kleinian groups, the so-called log 3 theorem proved by Culler and Shalen [6] and generalized by Anderson, Canary, Culler and Shalen [3]. This seminal result can be stated as follows:

**Log 3 theorem** Let  $\xi$  and  $\eta$  be noncommuting isometries of  $\mathbb{H}^3$ . Suppose that  $\xi$  and  $\eta$  generate a torsion-free, topologically tame, discrete group which is not cocompact and contains no parabolic. Let  $\Gamma_1$  and  $\alpha_1$  denote the set of isometries  $\{\xi, \eta\}$  and the real number 9, respectively. Then, for any  $z_0 \in \mathbb{H}^3$ , we have

$$e^{(2 \max_{\gamma \in \Gamma_1} \{\text{dist}(z_0, \gamma \cdot z_0)\})} \geq \alpha_1.$$

The log 3 theorem and its generalization imply that  $\frac{1}{2} \log 5$  and  $\frac{1}{2} \log 3$  are Margulis numbers for hyperbolic 3-manifolds that satisfy the conditions in the cases where the first Betti numbers are at least 4 or 3, respectively. The lower bounds for the volumes of such manifolds computed in [3; 5; 6] follow. Although the bounds given in [3; 6] are superseded by the recent works of Gabai, Meyerhoff and Milley [9; 10] and Milley [12] using a newer approach based on Mom technology, it is conceivable that an improvement in the lower bound for the displacements under the isometries described in the log 3 theorem will lead to improved Margulis numbers and lower bounds for the volumes of the classes of hyperbolic 3-manifolds mentioned above through the ideas introduced in [3; 5; 6]. Motivated by this, we prove the following in this paper:

**Theorem 5.1** Let  $\xi$  and  $\eta$  be noncommuting isometries of  $\mathbb{H}^3$ . Suppose that  $\xi$  and  $\eta$  generate a torsion-free discrete group which is not cocompact and contains no parabolic. Let  $\Gamma_{\dagger}$  denote the set of isometries  $\{\xi, \eta, \xi\eta\}$  and  $\alpha_{\dagger}$  be the real number  $5 + 3\sqrt{2}$ . Then, for any  $z_0 \in \mathbb{H}^3$ , we have

$$e^{(2 \max_{\gamma \in \Gamma_{\dagger}} \{\text{dist}(z_0, \gamma \cdot z_0)\})} \geq \alpha_{\dagger}.$$

This is proved in Section 5.

An orientable hyperbolic 3-manifold may be regarded as the quotient of the hyperbolic 3-space  $\mathbb{H}^3$  by a discrete group  $\Gamma$  of orientation-preserving isometries of  $\mathbb{H}^3$ . If  $\Gamma$  is a torsion-free Kleinian group and  $M = \mathbb{H}^3 / \Gamma$ , then  $\Gamma$  is called *topologically tame* if  $M$  is homeomorphic to the interior of a compact 3-manifold. Agol [1] and Calegari and Gabai [4] proved that every finitely generated Kleinian group is topologically tame. Therefore, we may drop the tameness hypothesis from Theorem 5.1.

The proof of Theorem 5.1 requires the ingredients introduced in [6] to prove the log 3 theorem. We review these ingredients briefly in the following subsections of the introduction. In particular, we summarize the proof of the log 3 theorem in Section 1.1 with an emphasis on the calculations required to obtain the number log 3. In Section 1.2, we propose an alternative technique to perform these calculations which makes it possible to extend Culler and Shalen's arguments in [6] to determine a lower bound for the displacements under any given set of isometries in  $\Gamma = \langle \xi, \eta \rangle$  as long as the

hypotheses of the log 3 theorem are satisfied. We describe this extension and summarize its application to the set  $\Gamma_{\dagger} = \{\xi, \eta, \xi\eta\} \subset \Gamma$  to achieve the lower bound stated in Theorem 5.1 in Section 1.3.

In the rest of this manuscript the boundary of the canonical compactification  $\overline{\mathbb{H}^3}$  of  $\mathbb{H}^3$  will be denoted by  $S_{\infty}$ , which is homeomorphic to  $S^2$ . The notation  $\Lambda_{\Gamma,z}$  will denote the limit set of  $\Gamma$ -orbit of  $z \in \mathbb{H}^3$  on  $S_{\infty}$ . By  $\text{dist}(z, \gamma \cdot z)$  we will mean the hyperbolic displacement of  $z \in \mathbb{H}^3$  under the action of the isometry  $\gamma: \mathbb{H}^3 \rightarrow \mathbb{H}^3$ . Any isometry  $\gamma$  of  $\mathbb{H}^3$  extends to a conformal automorphism  $\bar{\gamma}: \overline{\mathbb{H}^3} \rightarrow \overline{\mathbb{H}^3}$ . The conformal automorphism of  $S_{\infty}$  obtained by restricting  $\bar{\gamma}$  will be denoted by  $\gamma_{\infty}$ .

**Acknowledgements** The author would like to extend his sincerest thanks to the anonymous referee whose recommendations led to a much better exposition of the ideas in this paper, shortened the proofs substantially and made this text much more readable as a result. He is deeply grateful to Peter B Shalen for setting the course of this research and for very helpful discussions. He is also grateful to Marc E Culler for his corrections of an earlier version of this work.

### 1.1 A decomposition of $\Gamma = \langle \xi, \eta \rangle$ and the proof of the log 3 theorem

Let  $\xi$  and  $\eta$  be two noncommuting isometries of  $\mathbb{H}^3$ . Suppose that  $\xi$  and  $\eta$  generate a torsion-free discrete group which is not cocompact and contains no parabolic. Then  $\Gamma = \langle \xi, \eta \rangle$  is a free group of rank 2; see [6, Proposition 9.2]. This fact allows one to decompose  $\Gamma$  as disjoint union of subsets of reduced words. In particular, the decomposition

$$(1) \quad \Gamma = \{1\} \cup \bigcup_{\psi \in \Psi^1} J_{\psi}$$

is used in the proof of the log 3 theorem, which is carried out by considering two cases:

- (I) When  $\Gamma$  is geometrically infinite; that is,  $\Lambda_{\Gamma,z} = S_{\infty}$  for every  $z \in \mathbb{H}^3$ .
- (II) When  $\Gamma$  is geometrically finite.

In (1) each  $J_{\psi}$  is defined as the set of all nontrivial reduced words in  $\Gamma$  that have the initial letter  $\psi \in \Psi^1 = \{\xi, \eta, \eta^{-1}, \xi^{-1}\}$ .

In case (I), Culler and Shalen first prove that the Patterson density, a  $\Gamma$ -invariant conformal density  $(\mu_z)_{z \in \mathbb{H}^3}$  constructed by Patterson [14] and extensively studied by

Sullivan [15; 16], is the area density  $(A_z)_{z \in \mathbb{H}^3}$ , whose support is  $S_\infty$ ; see [6, Propositions 6.9 and 3.9]. Then, using the decomposition (1) together with its group-theoretical properties

$$(2) \quad \psi J_{\psi^{-1}} = \Gamma - J_\psi$$

for  $\psi \in \Psi^1$ , they construct a decomposition of the area density  $(A_z)_{z \in \mathbb{H}^3}$ , which in turn gives a decomposition of the area measure  $A_{z_0}$  based at  $z_0 \in \mathbb{H}^3$  into a finite sum of four measures  $\nu_\xi, \nu_\eta, \nu_{\eta^{-1}}, \nu_{\xi^{-1}}$  so that each measure  $\nu_{\psi^{-1}}$  is transformed to the complement of  $\nu_\psi$  for  $\psi \in \Psi^1$ ; see [6, Proposition 4.2(ii) and Lemma 5.3(ii) and (iii)]. In other words, they obtain the following:

**Theorem 1.1** *Let  $\Gamma = \langle \xi, \eta \rangle$  be a free, geometrically infinite Kleinian group without parabolics. For any  $z_0 \in \mathbb{H}^3$ , let  $A_{z_0}$  be the area measure based at  $z_0$ . There is a family of Borel measures  $\{\nu_\psi\}_{\psi \in \Psi^1}$  on  $S_\infty$  for  $\Psi^1 = \{\xi, \eta, \eta^{-1}, \xi^{-1}\}$  such that:*

$$(1) \quad A_{z_0}(S_\infty) = \sum_{\psi \in \Psi^1} \nu_\psi(S_\infty), \quad \text{where } A_{z_0} \text{ is normalized so that } A_{z_0}(S_\infty) = 1.$$

$$(2) \quad \int_{S_\infty} (\lambda_{\psi, z_0})^2 d\nu_{\psi^{-1}} = 1 - \int_{S_\infty} d\nu_\psi \quad \text{for each } \psi \in \Psi^1.$$

Furthermore, if  $z_0$  is on the common perpendicular  $\ell(\xi, \eta)$  of the isometries  $\xi$  and  $\eta$ , then

$$(3) \quad \int_{S_\infty} d\nu_{\xi^{-1}} = \int_{S_\infty} d\nu_\xi \quad \text{and} \quad \int_{S_\infty} d\nu_{\eta^{-1}} = \int_{S_\infty} d\nu_\eta.$$

Theorem 1.1 is not explicitly stated in [6]. But, as summarized above, it follows from [6, Lemma 5.3] using the conclusions of [6, Propositions 4.2, 6.9 and 3.9]. The function  $\lambda_{\psi, z_0}$  in part (2) is the conformal expansion factor of  $\psi_\infty$  measured in the round metric centered at  $z_0$  (see [6, Section 2.4] for details). The common perpendicular  $\ell(\xi, \eta)$  mentioned in Theorem 1.1 is the fixed locus of the involution  $\tau \in \text{Isom}^+(\mathbb{H}^3)$  that conjugates  $\xi$  to  $\xi^{-1}$  and  $\eta$  to  $\eta^{-1}$  when  $\xi_\infty$  and  $\eta_\infty$  have no common fixed point on  $S_\infty$ ; see [6, Proposition 1.8].

Next, Culler and Shalen relate the masses of the measures  $\nu_\xi, \nu_\eta, \nu_{\eta^{-1}}$  and  $\nu_{\xi^{-1}}$  to the hyperbolic displacements  $\text{dist}(z_0, \xi \cdot z_0), \text{dist}(z_0, \eta \cdot z_0), \text{dist}(z_0, \eta^{-1} \cdot z_0)$  and  $\text{dist}(z_0, \xi^{-1} \cdot z_0)$ . In particular, they prove the statement below:

**Lemma 1.2** [6, Lemma 5.5; 8, Lemma 2.1] *Let  $a$  and  $b$  be numbers in  $[0, 1]$  which are not both equal to 0 and are not both equal to 1. Let  $\gamma$  be a loxodromic isometry*

of  $\mathbb{H}^3$  and let  $z_0$  be a point in  $\mathbb{H}^3$ . Suppose that  $\nu$  is a measure on  $S_\infty$  such that  $\nu \leq A_{z_0}$ ,  $\nu(S_\infty) \leq a$  and  $\int_{S_\infty} (\lambda_{\gamma, z_0})^2 d\nu \geq b$ . Then we have  $a > 0$ ,  $b < 1$  and

$$\text{dist}(z_0, \gamma \cdot z_0) \geq \frac{1}{2} \log \frac{b(1-a)}{a(1-b)}.$$

Then, using Theorem 1.1 and Lemma 1.2, they calculate the lower bound  $\log 3$  when  $\Gamma$  is geometrically infinite as follows [6, Proposition 5.2]: By the geometric fact  $\max_{\gamma \in \{\xi^{-1}, \eta^{-1}\}} \{\text{dist}(z_0, \gamma \cdot z_0)\} \geq \max_{\gamma \in \{\xi^{-1}, \eta^{-1}\}} \{\text{dist}(z_1, \gamma \cdot z_1)\}$  for any  $z_0 \in \mathbb{H}^3$  and the orthogonal projection  $z_1$  of  $z_0$  onto  $\ell(\xi, \eta)$ , we may assume, without loss of generality, that  $z_0 \in \ell(\xi, \eta)$ . Parts (1) and (3) of Theorem 1.1 imply  $1 = 2\nu_\xi(S_\infty) + 2\nu_\eta(S_\infty)$ , which in turn shows either  $\nu_\xi(S_\infty) \leq \frac{1}{4}$  or  $\nu_\eta(S_\infty) \leq \frac{1}{4}$ . If  $\nu_\xi(S_\infty) \leq \frac{1}{4}$ , parts (2) and (3) of Theorem 1.1 give that

$$\int_{S_\infty} (\lambda_{\xi^{-1}, z_0})^2 d\nu_\xi = 1 - \nu_\xi \geq \frac{3}{4}.$$

Since  $\nu_\xi \leq A_{z_0}$  by Theorem 1.1(1), it is possible to apply Lemma 1.2 with the choices  $a = \frac{1}{4}$ ,  $b = \frac{3}{4}$ ,  $\nu = \nu_\xi$  and  $\gamma = \xi^{-1}$  to conclude that  $\text{dist}(z_0, \xi^{-1} \cdot z_0) \geq (\log 9)/2$ . If  $\nu_\eta(S_\infty) \leq \frac{1}{4}$ , an analogous calculation with the choices  $a = \frac{1}{4}$ ,  $b = \frac{3}{4}$ ,  $\nu = \nu_\eta$  and  $\gamma = \eta^{-1}$  shows that  $\text{dist}(z_0, \eta^{-1} \cdot z_0) \geq (\log 9)/2$ . Because  $\text{dist}(z_0, \xi \cdot z_0) = \text{dist}(z_0, \xi^{-1} \cdot z_0)$  and  $\text{dist}(z_0, \eta \cdot z_0) = \text{dist}(z_0, \eta^{-1} \cdot z_0)$ , the  $\log 3$  theorem follows in the case (I).

In the case (II), Culler and Shalen define the function

$$f_{z_0}^1(\xi, \eta) = \max\{\text{dist}(z_0, \xi \cdot z_0), \text{dist}(z_0, \eta \cdot z_0)\}$$

for a fixed  $z_0 \in \mathbb{H}^3$  on the character variety  $\mathfrak{X} = \text{Isom}^+(\mathbb{H}^3) \times \text{Isom}^+(\mathbb{H}^3)$  of the free group on two generators. It is easy to show that  $f_{z_0}$  is proper and continuous on the closure of the set  $\mathfrak{G}\mathfrak{F}$  consisting of  $(\xi, \eta)$  such that  $\langle \xi, \eta \rangle$  is free on the generators  $\xi$  and  $\eta$ , geometrically finite and without any parabolic. They prove that  $f_{z_0}^1$  has no minimum in  $\mathfrak{G}\mathfrak{F}$ ; see [6, Lemma 9.5]. Since  $\mathfrak{G}\mathfrak{F}$  is open in  $\mathfrak{X}$  (see Marden [11, Theorem 8.1]), the function  $f_{z_0}^1$  achieves its minimum on the boundary  $\overline{\mathfrak{G}\mathfrak{F}} - \mathfrak{G}\mathfrak{F}$ . Then they show that the set of  $(\xi, \eta)$  such that  $\langle \xi, \eta \rangle$  is free, geometrically infinite and without any parabolic is dense in  $\overline{\mathfrak{G}\mathfrak{F}} - \mathfrak{G}\mathfrak{F}$  [6, Proposition 8.2]. Therefore, that every  $(\xi, \eta) \in \mathfrak{X}$  so that  $\langle \xi, \eta \rangle$  is free and contains no parabolic is in  $\overline{\mathfrak{G}\mathfrak{F}}$  (see [6, Proposition 9.3]) reduces the geometrically finite case to geometrically infinite case, proving the  $\log 3$  theorem.

### 1.2 An alternative technique to calculate the number $\log 3$

When  $\Gamma = \langle \xi, \eta \rangle$  mentioned in the  $\log 3$  theorem is geometrically infinite, the calculation of the lower bound  $\log 3$  in Section 1.1 follows from the application of

Lemma 1.2 with the choices  $\gamma = \psi^{-1}$ ,  $\nu = \nu_\psi$ ,  $a = \frac{1}{4}$  and  $b = \frac{3}{4}$  for  $\psi \in \Gamma_1 = \{\xi, \eta\}$ . These choices of  $a$  and  $b$  are derived from the equalities  $\nu_\xi(S_\infty) = \nu_{\xi^{-1}}(S_\infty)$  and  $\nu_\eta(S_\infty) = \nu_{\eta^{-1}}(S_\infty)$  in Theorem 1.1(3) which requires the use of a point  $z_0$  on the common perpendicular  $\ell(\xi, \eta)$  of  $\xi$  and  $\eta$ . Such a geodesic perpendicular to the axes of a collection of more than two isometries in  $\Gamma = \langle \xi, \eta \rangle$  does not exist in general.

We calculate  $\log 3$  without referring to  $\ell(\xi, \eta)$  as follows: If  $\nu_{\xi^{-1}}(S_\infty) = 0$ , we get  $\nu_\xi(S_\infty) = 1$  by Theorem 1.1(2). Then we obtain  $\nu_\eta(S_\infty) = 0$  and  $\nu_{\eta^{-1}}(S_\infty) = 0$  by Theorem 1.1(1). Part (2) of Theorem 1.1 applied to  $\psi = \eta^{-1}$  gives a contradiction. Similar arguments for  $\xi^{-1}$ ,  $\eta$ ,  $\eta^{-1}$  show that  $0 < \nu_\psi(S_\infty) < 1$  for  $\psi \in \Psi^1$ . In particular, we derive that

$$0 < \int_{S_\infty} \lambda_{\psi, z_0}^2 d\nu_{\psi^{-1}} = 1 - \int_{S_\infty} d\nu_\psi < 1$$

for every  $\psi \in \Psi^1$  by Theorem 1.1(2). Hence we can apply Lemma 1.2 with the choices

$$\begin{aligned} \gamma = \xi, \quad \nu = \nu_{\xi^{-1}}, \quad a = \nu_{\xi^{-1}}(S_\infty), \quad b &= \int_{S_\infty} (\lambda_{\xi, z_0})^2 d\nu_{\xi^{-1}}, \\ \gamma = \eta, \quad \nu = \nu_{\eta^{-1}}, \quad a = \nu_{\eta^{-1}}(S_\infty), \quad b &= \int_{S_\infty} (\lambda_{\eta, z_0})^2 d\nu_{\eta^{-1}}, \\ \gamma = \eta^{-1}, \quad \nu = \nu_\eta, \quad a = \nu_\eta(S_\infty), \quad b &= \int_{S_\infty} (\lambda_{\eta^{-1}, z_0})^2 d\nu_\eta, \\ \gamma = \xi^{-1}, \quad \nu = \nu_\xi, \quad a = \nu_\xi(S_\infty), \quad b &= \int_{S_\infty} (\lambda_{\xi^{-1}, z_0})^2 d\nu_\xi. \end{aligned}$$

Then Lemma 1.2 produces the inequality

$$(3) \quad e^{2 \operatorname{dist}(z_0, \gamma \cdot z_0)} \geq \frac{(1 - \int_{S_\infty} d\nu_{\gamma^{-1}})(1 - \int_{S_\infty} d\nu_\gamma)}{\int_{S_\infty} d\nu_{\gamma^{-1}} \cdot \int_{S_\infty} d\nu_\gamma}$$

for every  $\gamma \in \Psi^1$ . We consider the constants on the right-hand side of the inequalities in (3) as the values of the functions

$$\begin{aligned} f_1^1(x_1, x_2, x_3, x_4) &= \frac{1-x_4}{x_4} \cdot \frac{1-x_1}{x_1}, & f_2^1(x_1, x_2, x_3, x_4) &= \frac{1-x_3}{x_3} \cdot \frac{1-x_2}{x_2}, \\ f_3^1(x_1, x_2, x_3, x_4) &= \frac{1-x_2}{x_2} \cdot \frac{1-x_3}{x_3}, & f_4^1(x_1, x_2, x_3, x_4) &= \frac{1-x_1}{x_1} \cdot \frac{1-x_4}{x_4} \end{aligned}$$

at  $(\int_{S_\infty} d\nu_\xi, \int_{S_\infty} d\nu_\eta, \int_{S_\infty} d\nu_{\eta^{-1}}, \int_{S_\infty} d\nu_{\xi^{-1}}) \in \mathbb{R}^4$  with  $\sum_{\psi \in \Psi^1} \int_{S_\infty} d\nu_\psi = 1$ .

Although the total masses of the measures  $\nu_\psi$  for  $\psi \in \Psi^1$  may change for a different  $z_0 \in \mathbb{H}^3$ , Theorem 1.1(1) still applies for the same decomposition in (1). Therefore,

the calculation of the number  $\log 3$  in the proof of the  $\log 3$  theorem becomes a consequence of the statement

$$\inf_{\mathbf{x} \in \Delta^3} \{ \max(f_1^1(\mathbf{x}), f_2^1(\mathbf{x}), f_3^1(\mathbf{x}), f_4^1(\mathbf{x})) \} = 9,$$

where  $\Delta^3 = \{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4 : x_1 + x_2 + x_3 + x_4 = 1, x_i > 0, 1 \leq i \leq 4\}$  (see Lemma 2.1). The functions  $f_1^1, f_2^1, f_3^1$  and  $f_4^1$  will be referred to as *displacement functions* for the decomposition of  $\Gamma$  in (1).

When combined with the arguments developed by Culler and Shalen in [6] to prove the  $\log 3$  theorem, this calculation technique extends to a process to find a lower bound for the maximum of the displacements under the isometries determined by a decomposition of  $\Gamma = \langle \xi, \eta \rangle$ . This extension is summarized in the next section.

### 1.3 Decompositions of $\Gamma = \langle \xi, \eta \rangle$ and hyperbolic displacements

For any nonelementary discrete subgroup  $\Gamma$  of the isometries of the hyperbolic  $n$ -space  $\mathbb{H}^n$ , there exists a  $\Gamma$ -invariant conformal density  $(\mu_z)_{z \in \mathbb{H}^n}$  for  $\mathbb{H}^n$ , constructed by Patterson [14], whose support is the limit set of  $\Gamma$ . In the case (I) of the proof of the  $\log 3$  theorem, it is required to decompose the Patterson density for  $\mathbb{H}^3$ , which is the area density  $(A_z)_{z \in \mathbb{H}^3}$ , whose support is  $S_\infty$  (see [6, Propositions 3.9 and 6.9]), corresponding to the decomposition of  $\Gamma = \langle \xi, \eta \rangle$  in (1). To this purpose, Culler and Shalen prove a more general statement (see [6, Proposition 4.2]) which establishes the existence of a family of  $D$ -conformal densities  $(\mathcal{M}_V)_{V \in \mathcal{V}}$ ,  $D \in [0, n - 1]$ , for  $\mathbb{H}^n$  indexed by a countable collection  $\mathcal{V}$  of subsets of an infinite, uniformly discrete subset  $W$  of  $\mathbb{H}^n$  with certain conditions, one of which is that  $\mathcal{M}_W$  can be decomposed as a sum of finitely many  $D$ -conformal densities  $\mathcal{M}_{V_i} = (\mu_{V_i, z})_{z \in \mathbb{H}^3}$  provided that  $W = \bigcup_{i=1}^m V_i \in \mathcal{V}$  for disjoint sets  $V_i \in \mathcal{V}$  for  $1 \leq i \leq m$ ; see [6, Proposition 4.2(ii)]. In particular, [6, Proposition 4.2(ii)] is applied to the disjoint union

$$W^1 = \{z_0\} \cup \bigcup_{\psi \in \Psi^1} \{\gamma \cdot z_0 : \gamma \in J_\psi\} \subset \mathbb{H}^3,$$

which leads to the decomposition of the area density  $(A_z)_{z \in \mathbb{H}^3}$  and, consequently, the decomposition of the area measure  $A_{z_0}$  based at  $z_0$  into a finite sum of Borel measures as stated in Theorem 1.1(1).

The notion of  $D$ -conformal densities,  $D \in [0, n - 1]$ , for  $\mathbb{H}^n$  was introduced by Sullivan [15; 16] as a generalization of  $\Gamma$ -invariant conformal densities for  $\mathbb{H}^n$  [14]. Interested readers may refer to [6; 14; 15; 16] and Nicholls [13], for the basics of this subject. In this paper, their use will be limited to the application of [6, Proposition 4.2] to a carefully chosen infinite, uniformly discrete subset of  $\mathbb{H}^3$  in Section 3 (see

Lemma 3.3). Therefore, constructions and properties of conformal densities will be assumed without explanation in the rest of this text. Unless stated otherwise, we shall assume that  $\Gamma = \langle \xi, \eta \rangle$  satisfies the hypotheses given in the log 3 theorem.

The organization of the rest of this paper is as follows. Let  $z_0$  be a fixed point in  $\mathbb{H}^3$ . In Section 2, we give the necessary calculations in detail to obtain the number log 3 for the log 3 theorem by using the approach outlined in Section 1.2. In particular, we show that the infimum of the maximum of  $f_1^1(\mathbf{x})$ ,  $f_2^1(\mathbf{x})$ ,  $f_3^1(\mathbf{x})$  and  $f_4^1(\mathbf{x})$  over the simplex  $\Delta^3$  is  $\alpha_1 = 9$ . This follows from the facts

- (a)  $\inf_{\mathbf{x} \in \Delta^3} \{\max(f_1^1(\mathbf{x}), f_2^1(\mathbf{x}))\} = \min_{\mathbf{x} \in \Delta^3} \{\max(f_1^1(\mathbf{x}), f_2^1(\mathbf{x}))\}$ ,
- (b)  $\min_{\mathbf{x} \in \Delta^3} \{\max(f_1^1(\mathbf{x}), f_2^1(\mathbf{x}))\} = f_1^1(\mathbf{x}^*)$  for a point  $\mathbf{x}^* \in \Delta_3 \subset \Delta^3$

proved in Lemma 2.1 in Section 2, where  $\Delta_3 = \{\mathbf{x} \in \Delta^3 : f_1^1(\mathbf{x}) = f_2^1(\mathbf{x})\}$ .

When  $\Delta^3$  is considered as a submanifold of  $\mathbb{R}^3$ , the statement  $\mathbf{x}^* \in \Delta_3$  is deduced from the observation that  $f_1^1$  and  $f_2^1$  are smooth functions in an open neighborhood of  $\Delta^3$  with no local minima. The coordinates of  $\mathbf{x}^*$  are calculated using the conclusion that  $\mathbf{x}^*$  is a solution of a certain Lagrange multipliers problem and satisfies  $f_1^1(\mathbf{x}) = f_2^1(\mathbf{x})$ . The lower bound log 3 is obtained by evaluating  $f_1^1$  at the point  $\mathbf{x}^*$ .

Let  $\Psi^\dagger = \{\xi\eta, \xi^2, \xi\eta^{-1}, \eta, \xi^{-1}, \eta^{-1}\xi^{-1}, \eta^{-2}, \eta^{-1}\xi\}$  and  $\Psi_r^\dagger = \{\xi, \eta^{-1}\}$ . In Section 3, we introduce the decomposition

$$(4) \quad \Gamma_{\mathcal{D}^\dagger} = \{1\} \cup \Psi_r^\dagger \cup \bigcup_{\psi \in \Psi^\dagger} J_\psi$$

of  $\Gamma$  which contains the isometries in  $\Gamma_\dagger = \{\xi, \eta, \xi\eta\} \subset \Psi^\dagger \cup \Psi_r^\dagger$ . We apply [6, Proposition 4.2] to the infinite, uniformly discrete subset

$$(5) \quad W^\dagger = \{z_0\} \cup \{\gamma \cdot z_0 : \gamma \in \Psi_r^\dagger\} \cup \bigcup_{\psi \in \Psi^\dagger} \{\gamma \cdot z_0 : \gamma \in J_\psi\}$$

of  $\mathbb{H}^3$  to construct the decomposition of the area measure  $A_{z_0}$  based at  $z_0$  corresponding the decomposition  $\Gamma_{\mathcal{D}^\dagger}$  in Theorem 3.4, an analog of Theorem 1.1 for  $\Psi^\dagger$ , in Section 3. Using Theorem 3.4 and Lemma 1.2, we determine the displacement functions for the displacements under the isometries in  $\Psi^\dagger \cup \Psi_r^\dagger$  in Section 3. There



are 18 displacement functions  $\{f_i\}_{i=1}^8$ ,  $\{g_j\}_{j=1}^6$  and  $\{h_k\}_{k=1}^4$  for  $\Gamma_{\mathcal{D}^\dagger}$  with formulas

$$\begin{aligned}
 f_1(\mathbf{x}) &= \frac{1-x_4-x_5-x_6}{x_4+x_5+x_6} \cdot \frac{1-x_1}{x_1}, & f_2(\mathbf{x}) &= \frac{1-x_4-x_5-x_6-x_7-x_8}{x_4+x_5+x_6+x_7+x_8} \cdot \frac{1-x_2}{x_2}, \\
 f_4(\mathbf{x}) &= \frac{1-x_1-x_2-x_3}{x_1+x_2+x_3} \cdot \frac{1-x_4}{x_4}, & f_3(\mathbf{x}) &= \frac{1-x_1-x_2-x_3-x_7-x_8}{x_1+x_2+x_3+x_7+x_8} \cdot \frac{1-x_3}{x_3}, \\
 f_7(\mathbf{x}) &= \frac{1-x_4-x_5-x_6}{x_4+x_5+x_6} \cdot \frac{1-x_7}{x_7}, & f_5(\mathbf{x}) &= \frac{1-x_1-x_2-x_3-x_7-x_8}{x_1+x_2+x_3+x_7+x_8} \cdot \frac{1-x_5}{x_5}, \\
 f_8(\mathbf{x}) &= \frac{1-x_1-x_2-x_3}{x_1+x_2+x_3} \cdot \frac{1-x_8}{x_8}, & f_6(\mathbf{x}) &= \frac{1-x_4-x_5-x_6-x_7-x_8}{x_4+x_5+x_6+x_7+x_8} \cdot \frac{1-x_6}{x_6}, \\
 g_1(\mathbf{x}) &= \frac{x_7}{1-x_7} \cdot \frac{1-x_1}{x_1}, & g_2(\mathbf{x}) &= \frac{x_6}{1-x_6} \cdot \frac{1-x_2}{x_2}, \\
 g_3(\mathbf{x}) &= \frac{x_5}{1-x_5} \cdot \frac{1-x_3}{x_3}, & g_4(\mathbf{x}) &= \frac{x_8}{1-x_8} \cdot \frac{1-x_4}{x_4}, \\
 g_5(\mathbf{x}) &= \frac{x_3}{1-x_3} \cdot \frac{1-x_5}{x_5}, & g_6(\mathbf{x}) &= \frac{x_2}{1-x_2} \cdot \frac{1-x_6}{x_6}, \\
 h_1(\mathbf{x}) &= \frac{1-x_6}{x_6} \cdot \frac{1-x_8}{x_8}, & h_2(\mathbf{x}) &= \frac{1-x_2}{x_2} \cdot \frac{1-x_8}{x_8}, \\
 h_3(\mathbf{x}) &= \frac{1-x_5}{x_5} \cdot \frac{1-x_7}{x_7}, & h_4(\mathbf{x}) &= \frac{1-x_3}{x_3} \cdot \frac{1-x_7}{x_7},
 \end{aligned}$$

for  $\mathbf{x} \in \Delta^7 = \left\{ (x_1, x_2, \dots, x_8) \in \mathbb{R}^8 : \sum_{i=1}^8 x_i = 1, x_i > 0, 1 \leq i \leq 8 \right\}$ .

In this paper we will not be concerned with the functions  $\{h_k\}_{k=1}^4$ , because they provide information about displacements under  $\eta^{-2}$ ,  $\eta^{-1}\xi$ ,  $\xi\eta^{-1}$  and  $\xi^2$ . Only the functions  $\{f_i\}_{i=1}^8$  and  $\{g_j\}_{j=1}^6$  are related to the displacements under  $\xi$ ,  $\eta$  and  $\xi\eta$ . Furthermore, it is possible to show that just the first eight  $\{f_i\}_{i=1}^8$  are significant for finding a lower bound for the maximum of the displacements  $\text{dist}(z_0, \xi \cdot z_0)$ ,  $\text{dist}(z_0, \eta \cdot z_0)$  and  $\text{dist}(z_0, \xi\eta \cdot z_0)$ .

We consider  $\Delta^7$  as a submanifold of  $\mathbb{R}^8$ . Let  $I^\dagger = \{1, 2, 3, 4, 5, 6, 7, 8\}$ . Each function  $f_i$  for  $i \in I^\dagger$  is smooth in an open neighborhood of  $\Delta^7$ . Then the calculation of the infimum of the maximum of the functions  $\{f_j\}_{j \in I^\dagger}$  follows from the statements

- (c)  $\inf_{\mathbf{x} \in \Delta^7} \{ \max(f_1(\mathbf{x}), \dots, f_8(\mathbf{x})) \} = \min_{\mathbf{x} \in \Delta^7} \{ \max(f_1(\mathbf{x}), \dots, f_8(\mathbf{x})) \},$
- (d)  $\min_{\mathbf{x} \in \Delta^7} \{ \max(f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_8(\mathbf{x})) \} = f_1(\mathbf{x}^*) \quad \text{for } \mathbf{x}^* \in \Delta_7 \subset \Delta^7,$

proved in Proposition 4.8 in Section 4.2, where  $\Delta_7 = \{ \mathbf{x} \in \Delta^7 : f_j(\mathbf{x}) = f_k(\mathbf{x}), j, k \in I^\dagger \}$ . Although the proof of the observation that  $\mathbf{x}^* \in \Delta_7$  also uses the fact that none of

the functions  $\{f_j\}_{j \in I^\dagger}$  has a local minimum in  $\Delta^7$ , it is more involved. It requires a case-by-case study of the values of the functions  $\{f_j\}_{j \in I^\dagger}$  at the point  $\mathbf{x}^*$ .

In Section 4.1, we use the first-order partial derivatives of the functions  $\{f_j\}_{j \in I^\dagger}$  to show that there are certain direction vectors in the tangent space  $T_{\mathbf{x}^*} \Delta^7$  so that moving along these vectors reduces the number of possible cases for the values of the functions  $\{f_j\}_{j \in I^\dagger}$  at the point  $\mathbf{x}^*$  to only five main cases

- (I)  $f_j(\mathbf{x}^*) < \alpha_*$  for  $j \in \{4, 5, 6, 7, 8\}$  and  $f_j(\mathbf{x}^*) = \alpha_*$  for  $j \in \{1, 2, 3\}$ ,
- (II)  $f_j(\mathbf{x}^*) < \alpha_*$  for  $j \in \{1, 2, 3, 4, 5, 6\}$  and  $f_j(\mathbf{x}^*) = \alpha_*$  for  $j \in \{7, 8\}$ ,
- (III)  $f_j(\mathbf{x}^*) = \alpha_*$  for  $j \in \{1, 2, 3, 4, 5, 6\}$  and  $f_j(\mathbf{x}^*) < \alpha_*$  for  $j \in \{7, 8\}$ ,
- (IV)  $f_j(\mathbf{x}^*) = \alpha_*$  for  $j \in \{1, 2, 3, 7, 8\}$  and  $f_j(\mathbf{x}^*) < \alpha_*$  for  $j \in \{4, 5, 6\}$ ,
- (V)  $f_j(\mathbf{x}^*) = \alpha_*$  for  $j \in I^\dagger$ ,

where  $\alpha_* = \min_{\mathbf{x} \in \Delta^7} \{\max(f_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_8(\mathbf{x}))\}$ . In each of the cases (I), (II), (III) and (IV), we prove in Section 4.2 that there exists a piece of curve in  $\Delta^7$  passing through the point  $\mathbf{x}^*$  such that moving along this curve produces a point in  $\Delta^7$  at which a smaller minimum for the maximum of the functions  $\{f_j\}_{j \in I^\dagger}$  is attained. This leaves only the case  $\mathbf{x}^* \in \Delta_7$  which suggests a method to find the coordinates of the point  $\mathbf{x}^*$ . Then we evaluate one of the displacement functions in  $\{f_j\}_{j \in I^\dagger}$  at  $\mathbf{x}^*$  to calculate  $\alpha_*$ . In other words, we prove the following statement:

**Theorem 4.14** *Let  $F^\dagger: \Delta^7 \rightarrow \mathbb{R}$  be the function defined by  $\mathbf{x} \mapsto \max\{f_i(\mathbf{x}) : i \in I^\dagger\}$ . Then we have  $\inf_{\mathbf{x} \in \Delta^7} F^\dagger(\mathbf{x}) = 5 + 3\sqrt{2}$ .*

This is given in Section 4.2, and provides the main estimate of Theorem 5.1 in the geometrically infinite case.

In Section 5, we show that the proper and continuous function defined for a fixed point  $z_0 \in \mathbb{H}^3$  on the character variety  $\mathfrak{X}$  defined by the formula

$$f_{z_0}^\dagger(\xi, \eta) = \max\{\text{dist}(z_0, \xi \cdot z_0), \text{dist}(z_0, \eta \cdot z_0), \text{dist}(z_0, \xi\eta \cdot z_0)\}$$

has no local minimum in  $\mathfrak{G}\mathfrak{F}$  to reduce the geometrically finite case to the geometrically infinite case, completing the proof of Theorem 5.1. Note that an analogous process applies to a broader class of decompositions.

As summarized above, when  $\Gamma = \langle \xi, \eta \rangle$  is geometrically infinite, the infimum of the maximum of the displacement functions  $f_1, \dots, f_8, g_1, \dots, g_6$ , determined by the decomposition  $\Gamma_{\mathcal{D}^\dagger}$  of  $\Gamma$ , over  $\Delta^7$  provides a lower bound for the displacements under

the isometries  $\xi$ ,  $\eta$  and  $\xi\eta$  in Theorem 5.1. Notice that, similar to the displacement functions  $f_1^1, f_2^1, f_3^1$  and  $f_4^1$  given by the decomposition  $\Gamma_{\mathcal{D}^1}$ , the infimum of the maximum of the displacement functions  $g_1, f_2, f_3, g_4, f_5, f_6, f_7$  and  $f_8$  over  $\Delta^7$  provides a lower bound for the displacements under the isometries  $\xi$  and  $\eta$ . Although we have

$$\inf_{\mathbf{x} \in \Delta^3} \{\max(f_1^1(\mathbf{x}), f_2^1(\mathbf{x}), f_3^1(\mathbf{x}), f_4^1(\mathbf{x}))\} = \inf_{\mathbf{x} \in \Delta^7} \{\max(g_1(\mathbf{x}), f_2(\mathbf{x}), \dots, f_8(\mathbf{x}))\},$$

for  $\Gamma_{\mathcal{D}^1}$  and  $\Gamma_{\mathcal{D}^+}$  by Lemma 2.1 (see Section 2) and the fact that  $g_1(\mathbf{x}) = \frac{11}{4}, f_2(\mathbf{x}) = 9, f_3(\mathbf{x}) = 9, g_4(\mathbf{x}) = \frac{11}{4}, f_5(\mathbf{x}) = 9, f_6(\mathbf{x}) = 9, f_7(\mathbf{x}) = 9$  and  $f_8(\mathbf{x}) = 9$  for the point  $\mathbf{x} = (\frac{1}{12}, \frac{1}{12}, \frac{1}{12}, \frac{1}{12}, \frac{1}{12}, \frac{1}{12}, \frac{1}{4}, \frac{1}{4}) \in \Delta^7$ , it may be possible to obtain a larger lower bound than  $\log 3$  introduced in the  $\log 3$  theorem by examining a carefully chosen decomposition or a sequence of decompositions of  $\Gamma$ . If a larger lower bound for the displacements under the isometries  $\xi$  and  $\eta$  can be achieved, all the results involving the  $\log 3$  theorem in [2; 3; 8] and Culler and Shalen’s [7] can be improved accordingly.

## 2 The log 3 theorem revisited

In this section, we calculate the number  $\log 3$  using the view point proposed in Section 1.2 in case (I) of the proof of the  $\log 3$  theorem.

Let  $\mathcal{F}^1$  be the set of functions  $\{f_1^1, f_2^1, f_3^1, f_4^1\}$  introduced in Section 1.2. Let us define a continuous function  $G^1: \Delta^3 \rightarrow \mathbb{R}$  by  $G^1(\mathbf{x}) = \max\{f(\mathbf{x}) : f \in \mathcal{F}^1\}$ .

We aim to find  $\inf_{\mathbf{x} \in \Delta^3} G^1(\mathbf{x})$ . To this purpose we shall consider  $\Delta^3$  as a submanifold of  $\mathbb{R}^4$ . The tangent space  $T_{\mathbf{x}} \Delta^3$  of  $\Delta^3$  consists of the vectors in  $\mathbb{R}^4$  whose entries sum to 0 at any  $\mathbf{x} \in \Delta^3$ . Note that each displacement function  $f_i^1$  for  $i = 1, 2, 3, 4$  is smooth in an open neighborhood of  $\Delta^3$ . Therefore the directional derivative of  $f_i^1$  in the direction of any  $\vec{v} \in T_{\mathbf{x}} \Delta^3$  is given by  $\nabla f_i^1(\mathbf{x}) \cdot \vec{v}$  for any  $i = 1, 2, 3, 4$ .

We introduce the function  $\sigma: (0, 1) \rightarrow (0, \infty)$  defined by  $\sigma(x) = 1/x - 1$ , where  $\sigma'(x) = -1/x^2 < 0$  for every  $x \in (0, 1)$ . We have the equality

$$\inf_{\mathbf{x} \in \Delta^3} G^1(\mathbf{x}) = \inf_{\mathbf{x} \in \Delta^3} \{\max(f_1^1(\mathbf{x}), f_2^1(\mathbf{x}))\}$$

because  $f_1^1(\mathbf{x}) = f_4^1(\mathbf{x})$  and  $f_2^1(\mathbf{x}) = f_3^1(\mathbf{x})$  for every  $\mathbf{x} \in \Delta^3$ . In other words, it is enough to prove the following:

**Lemma 2.1** *If  $f_1^1(\mathbf{x}) = \sigma(x_4)\sigma(x_1)$  and  $f_2^1(\mathbf{x}) = \sigma(x_3)\sigma(x_2)$  are the displacement functions defined in Section 1.2 for  $\mathbf{x} \in \Delta^3$ , then  $\inf_{\mathbf{x} \in \Delta^3} \{\max(f_1^1(\mathbf{x}), f_2^1(\mathbf{x}))\} = 9$ .*

**Proof** Let  $F^1: \Delta^3 \rightarrow \mathbb{R}$  be the function defined by  $\mathbf{x} \mapsto \max(f_1^1(\mathbf{x}), f_2^1(\mathbf{x}))$ . Since  $f_1^1$  and  $f_2^1$  are continuous on  $\Delta^3$ ,  $F^1$  is also continuous. The number  $\inf_{\mathbf{x} \in \Delta^3} F^1(\mathbf{x})$  exists and is greater than or equal to 1 because the inequalities  $f_1^1(\mathbf{x}) > 1$  and  $f_2^1(\mathbf{x}) > 1$  hold for every  $\mathbf{x} \in \Delta^3$ . We aim to show that  $\inf_{\mathbf{x} \in \Delta^3} F^1(\mathbf{x}) = 9$ . First we prove that  $\inf_{\mathbf{x} \in \Delta^3} F^1(\mathbf{x}) = \min_{\mathbf{x} \in \Delta^3} F^1(\mathbf{x})$ .

Let  $\Delta_n = \{(x_1, x_2, x_3, x_4) \in \Delta^3 : 1/n \leq x_i \leq 1 - 1/(2n) \text{ for } i = 1, 2, 3, 4\}$  for  $n \geq 2$ ; then  $F^1$  has an absolute minimum  $\alpha_n$  at a point  $\mathbf{x}_n$  in  $\Delta_n$  as  $\Delta_n$  is compact. The sequence  $(\alpha_n)_{n=1}^\infty$  is decreasing because  $\Delta_n \subset \Delta_{n+1}$ . In other words, the sequence  $(F^1(\mathbf{x}_n))_{n=1}^\infty$  tends to an infimum of  $F^1$ . Assume that  $\mathbf{x}_n$  approaches to a point  $\mathbf{x}_0 = (b_1, b_2, b_3, b_4) \in \partial\Delta^3$  as  $n$  approaches to infinity. Then we get  $b_i = 0$  for some  $i = 1, 2, 3, 4$ . Suppose that  $b_1 = 0$ . By the definition of  $f_1^1$ , we must have  $b_4 = 1$ . Otherwise  $(F^1(\mathbf{x}_n))_{n=1}^\infty$  would approach to infinity instead of tending to an infimum. Then we conclude that  $b_3 = 0$  and  $b_2 = 0$ . But, the function  $(F^1(\mathbf{x}_n))_{n=1}^\infty$  tends to infinity by the definition of  $f_2^1$  in this case, a contradiction. Thus, we get  $b_1 \neq 0$ . Similar contradictions arise under the assumption  $b_i = 0$  for any  $i \in \{2, 3, 4\}$  implying that  $b_i \neq 0$  for every  $i \in \{2, 3, 4\}$ . Therefore  $(\mathbf{x}_n)_{n=1}^\infty$  cannot have a limit point on the boundary of  $\Delta^3$ .

We claim that there exists a positive integer  $n_0$  so that  $\mathbf{x}_n = \mathbf{x}_{n+1}$  for every  $n \geq n_0$ . Let us assume otherwise that we have a subsequence  $(\mathbf{x}_{n_j})_{j=1}^\infty$  of the sequence  $(\mathbf{x}_n)_{n=1}^\infty$  so that  $\mathbf{x}_{n_j} \in \Delta_{n_{j+1}} - \Delta_{n_j}$  for every integer  $j > 0$ . Since we have  $\bigcup_{n=2}^\infty \Delta_n = \Delta^3$ , there exists a subsequence of  $(\mathbf{x}_{n_j})_{j=1}^\infty$  which has a limit point on  $\partial\Delta^3$ , a contradiction. In other words the absolute minimum of  $F^1$  is attained at a point  $\mathbf{x}^* = (x_1^*, x_2^*, x_3^*, x_4^*) \in \Delta^3$  so that  $F^1(\mathbf{x}^*) = \inf_{\mathbf{x} \in \Delta^3} F^1(\mathbf{x})$ .

Let  $\Delta_3 = \{\mathbf{x} \in \Delta^3 : f_1^1(\mathbf{x}) = f_2^1(\mathbf{x})\}$ . We claim that  $\mathbf{x}^* \in \Delta_3$ . Assume otherwise that  $f_1^1(\mathbf{x}^*) > f_2^1(\mathbf{x}^*)$ . Then there exists a neighborhood  $U$  of  $\mathbf{x}^*$  such that  $f_1^1(\mathbf{x}) > f_2^1(\mathbf{x})$  for every  $\mathbf{x} \in U$ . In particular we get  $F^1(\mathbf{x}) = f_1^1(\mathbf{x})$  on  $U$ . Since  $F^1(\mathbf{x}^*)$  is the minimum value, the function  $f_1^1$  must have a critical point at  $\mathbf{x}^*$ . This is a contradiction. Because the derivative of  $f_1^1$  in the direction of  $\vec{v} = \langle 1, -1, 0, 0 \rangle \in T_{\mathbf{x}^*}\Delta^3$  is calculated as

$$\nabla f_1^1(\mathbf{x}) \cdot \vec{v} = -\frac{1-x_4}{x_1^2 x_4} < 0,$$

the function  $f_1^1$  decreases at any  $\mathbf{x} \in \Delta^3$  in the direction of  $\vec{v}$ . This argument also applies, mutatis mutandis, to show that the assumption  $f_2^1(\mathbf{x}^*) > f_1^1(\mathbf{x}^*)$  leads to a contradiction since the directional derivative  $\nabla f_2^1(\mathbf{x}) \cdot \vec{v} > 0$  at any  $\mathbf{x} \in \Delta^3$ . Hence we obtain that  $\mathbf{x}^* \in \Delta_3$ .

Let  $f_1(x) = \sigma(x_4)\sigma(x_1)$  and  $f_2(x) = \sigma(x_3)\sigma(x_2)$  be the extensions of  $f_1^1$  and  $f_2^1$  to the open set  $C = (0, 1) \times (0, 1) \times (0, 1) \times (0, 1)$ . Let us consider the Lagrange

multipliers problem of finding the minimum of  $f_1$  subject to the constraints  $G_1(x) = 0$  and  $G_2(x) = 0$  for  $x = (x_1, x_2, x_3, x_4) \in C$ , where  $G_1$  and  $G_2$  are the functions defined by

$$G_1(x) = f_1(x) - f_2(x) \quad \text{and} \quad G_2(x) = x_1 + x_2 + x_3 + x_4 - 1.$$

Since we have  $f_1(\mathbf{x}^*) = \min_{\mathbf{x} \in \Delta^3} F^1(\mathbf{x})$  and  $\mathbf{x}^* \in \Delta_3$ , the point  $\mathbf{x}^*$  is a solution to this problem. Therefore the gradient vector  $\nabla f_1(\mathbf{x}^*)$  is in the subspace of  $\mathbb{R}^4$  spanned by the vectors  $\nabla G_1(\mathbf{x}^*)$  and  $\nabla G_2(\mathbf{x}^*)$ . Equivalently the matrix

$$\begin{bmatrix} \nabla f_1(\mathbf{x}^*) \\ \nabla G_1(\mathbf{x}^*) \\ \nabla G_2(\mathbf{x}^*) \end{bmatrix} = \begin{bmatrix} \sigma'(x_1^*)\sigma(x_4^*) & 0 & 0 & \sigma(x_1^*)\sigma'(x_4^*) \\ \sigma'(x_1^*)\sigma(x_4^*) & -\sigma'(x_2^*)\sigma(x_3^*) & -\sigma(x_2^*)\sigma'(x_3^*) & \sigma(x_1^*)\sigma'(x_4^*) \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

has rank less than 3. By applying the row reduction operations  $R_2 \mapsto (-1)R_1 + R_2$ ,  $R_1 \mapsto (-\sigma(x_1^*)\sigma'(x_4^*))R_3 + R_1$  and  $R_1 \mapsto (\sigma(x_1^*)\sigma'(x_4^*)) / (\sigma(x_2^*)\sigma'(x_3^*))R_2 + R_1$  simultaneously, we see that the matrix above is row equivalent to:

$$\begin{bmatrix} \frac{(x_1^*-x_4^*)(1-x_1^*-x_4^*)}{(x_1^*x_4^*)^2} & \frac{(x_2^*-x_3^*)(1-x_2^*-x_3^*)\sigma(x_1^*)\sigma'(x_4^*)}{(x_2^*x_3^*)^2\sigma(x_2^*)\sigma'(x_3^*)} & 0 & 0 \\ 0 & -\sigma'(x_2^*)\sigma(x_3^*) & -\sigma(x_2^*)\sigma'(x_3^*) & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

The reduced matrix above has rank less than 3 if and only if  $x_1^* = x_4^*$  and  $x_2^* = x_3^*$ . Then it follows from the equation  $f_1(\mathbf{x}^*) = f_2(\mathbf{x}^*)$  that  $x_1^* = x_2^*$ . As a result we find that  $x_i^* = \frac{1}{4}$  for every  $i \in \{1, 2, 3, 4\}$ . Finally a simple evaluation shows that  $\inf_{\mathbf{x} \in \Delta^3} F^1(\mathbf{x}) = 9$ . □

It is worth emphasizing a few key points used in the proof of Lemma 2.1. The functions  $f_1^1$  and  $f_2^1$  have no local minimum on  $\Delta^3$ . This fact implies that  $\mathbf{x}^* \in \Delta_3$ . The proof of Lemma 2.1 shows that the main ingredients used to calculate the number  $\log 3$  are:

- (a)  $\inf_{\mathbf{x} \in \Delta^3} F^1(\mathbf{x}) = \min_{\mathbf{x} \in \Delta^3} F^1(\mathbf{x})$ .
- (b) There exists a point  $\mathbf{x}^* \in \Delta_3 \subset \Delta^3$  such that  $\min_{\mathbf{x} \in \Delta^3} F^1(\mathbf{x}) = f_1^1(\mathbf{x}^*)$ .

Lemma 2.1 also establishes that the point  $\mathbf{x}^* \in \Delta^3$  is unique. If we assume the uniqueness of the point  $\mathbf{x}^*$  a priori together with (a), it is possible to suggest an alternative way of finding the coordinates of the point  $\mathbf{x}^*$ : let  $T_1, T_2: \mathbb{R}^4 \rightarrow \mathbb{R}^4$  be the maps defined by  $(x_1, x_2, x_3, x_4) \mapsto (x_2, x_1, x_4, x_3)$  and  $(x_1, x_2, x_3, x_4) \mapsto (x_4, x_3, x_2, x_1)$ , respectively. We have  $T_i(\mathbf{x}) \in \Delta^3$  and  $f_i^1(T_1(\mathbf{x})) = f_i^1(\mathbf{x})$  for every

$\mathbf{x} \in \Delta^3$ . Let  $H_1: \Delta^3 \rightarrow \mathbb{R}$  be the function so that  $H_1(\mathbf{x}) = \max\{(f_i^1 \circ T_1)(\mathbf{x}) : i = 1, 2\}$ . We see that  $F^1(\mathbf{x}) = H_1(\mathbf{x})$  for every  $\mathbf{x} \in \Delta^3$ . Since  $F^1$  takes its minimum value at the point  $\mathbf{x}^*$ , the function  $H_1$  takes its minimum value at the point  $T_1^{-1}(\mathbf{x}^*)$ . Therefore we obtain  $T_1^{-1}(\mathbf{x}^*) = \mathbf{x}^*$  which implies that  $x_1^* = x_2^*$ , and  $x_3^* = x_4^*$ .

Let  $H_2: \Delta^3 \rightarrow \mathbb{R}$  be the function defined by  $H_2(\mathbf{x}) = \max\{(f_i^1 \circ T_2)(\mathbf{x}) : i = 1, 2\}$ . We repeat an analog of the argument above for  $H_2$  to derive that  $x_1^* = x_4^*$  and  $x_2^* = x_3^*$ . As a result we find  $x_i^* = x_j^*$  for every  $i, j = 1, 2, 3, 4$ . In this calculation method, we do not refer to the statement  $\mathbf{x}^* \in \Delta_3$  given in (b).

### 3 Decomposition of $\Gamma = \langle \xi, \eta \rangle$ for the isometries $\xi, \eta$ and $\xi\eta$

Let  $\Gamma$  be a group which is free on a finite generating set  $\Xi$ . Let  $\Xi^{-1} = \{\psi^{-1} : \psi \in \Xi\}$ . For  $m \geq 2$ , every element  $\gamma$  of  $\Gamma$  can be written uniquely as a reduced word  $\psi_1 \cdots \psi_m$ , where each  $\psi_i$  is an element of  $\Xi \cup \Xi^{-1}$  and  $\psi_{i+1} \neq \psi_i^{-1}$  for  $i = 1, \dots, m - 1$ . If  $k \leq m$  is a positive integer and  $\gamma \neq 1$ , we shall call  $\psi_1 \cdots \psi_k$  the *initial word of length  $k$*  of  $\gamma$ .

Let  $\Psi^*$  be a finite set of words in  $\Gamma$ . For each word  $\psi \in \Psi^*$ , let  $J_\psi$  denote the set of nontrivial elements of  $\Gamma$  that have initial word  $\psi$ . Depending on the number of elements in  $\Xi$  and lengths of words in  $\Psi^*$  there may be a set of words which are not contained in any of  $J_\psi$ . Let us call this set the *residue set* of  $\Psi^*$  and denote it by  $\Psi_r^*$ .

**Definition 3.1** For a given pair  $\mathcal{D}^* = (\Psi^*, \Psi_r^*)$  of finite, disjoint sets of words  $\Psi^*$  and  $\Psi_r^*$  in  $\Gamma$ , if  $\Gamma = \{1\} \cup \Psi_r^* \cup \bigcup_{\psi \in \Psi^*} J_\psi$ , then  $\Gamma_{\mathcal{D}^*}$  will be called a *decomposition* of  $\Gamma$ .

**Definition 3.2** A decomposition  $\Gamma_{\mathcal{D}^*}$  with  $\mathcal{D}^* = (\Psi^*, \Psi_r^*)$  is *symmetric* if  $\Psi^*$  and  $\Psi_r^*$  are preserved by every bijection of  $\Xi \cup \Xi^{-1}$ .

We know that  $\Gamma = \langle \xi, \eta \rangle$  described in the log 3 theorem is a free group on  $\Xi = \{\xi, \eta\}$  by [6, Proposition 9.2]. For  $\mathcal{D}^1 = (\Psi^1, \Psi_r^1 = \emptyset)$ , the decomposition  $\Gamma_{\mathcal{D}^1}$  corresponds to the symmetric decomposition of  $\Gamma$  in (1).

We introduce another decomposition of  $\Gamma$  that contains the set  $\Gamma_\dagger = \{\xi, \eta, \xi\eta\}$ . Let  $\Psi^\dagger = \{\xi\eta, \xi^2, \xi\eta^{-1}, \eta, \xi^{-1}, \eta^{-1}\xi^{-1}, \eta^{-2}, \eta^{-1}\xi\}$ . Then it is straightforward to see that

$$(6) \quad \Gamma = \{1\} \cup \Psi_r^\dagger \cup \bigcup_{\psi \in \Psi^\dagger} J_\psi,$$

where the residue set is  $\Psi_r^\dagger = \{\xi, \eta^{-1}\}$ . Note that  $\Gamma_{\mathcal{D}^\dagger}$  for  $\mathcal{D}^\dagger = (\Psi^\dagger, \Psi_r^\dagger)$  is not symmetric. We shall use  $\Gamma_{\mathcal{D}^\dagger}$  in the proof of Theorem 5.1. In particular, we aim to prove Theorem 3.4, an analog of Theorem 1.1, for  $\Psi^\dagger$ .

We first state a more general result Lemma 3.3, an analog of [6, Lemma 5.3], for  $\Psi^\dagger$ . Although Lemma 3.3 follows directly from the same arguments used in the proof of [6, Lemma 5.3], its proof is included for the sake of completeness. The main tool is [6, Proposition 4.2]. In the following,  $J_\Psi$  denotes the disjoint union  $\bigcup_{\psi \in \Psi} J_\psi$  for  $\Psi \subset \Psi^\dagger$ :

**Lemma 3.3** *Let  $\Gamma$  be a Kleinian group which is free on a generating set  $\{\xi, \eta\}$ . Let  $z_0$  be any point of  $\mathbb{H}^3$ . Then there exists a number  $D \in [0, 2]$ , a  $\Gamma$ -invariant  $D$ -conformal density  $\mathcal{M} = (\mu_z)$  for  $\mathbb{H}^3$  and a family  $\{\nu_\psi\}_{\psi \in \Psi^\dagger}$  of Borel measures on  $S_\infty$  such that:*

- (1)  $\mu_{z_0}(S_\infty) = 1$
- (2)  $\mu_{z_0} = \sum_{\psi \in \Psi^\dagger} \nu_\psi$
- (3) (a)  $\int_{S_\infty} (\lambda_{\xi, z_0})^D dv_{\xi^{-1}} = 1 - \int_{S_\infty} dv_{\xi\eta} - \int_{S_\infty} dv_{\xi^2} - \int_{S_\infty} dv_{\xi\eta^{-1}}$   
 (b)  $\int_{S_\infty} (\lambda_{\xi^{-1}, z_0})^D dv_{\xi\eta^{-1}} = \int_{S_\infty} dv_{\eta^{-1}\xi^{-1}} + \int_{S_\infty} dv_{\eta^{-1}\xi} + \int_{S_\infty} dv_{\eta^{-2}}$   
 (c)  $\int_{S_\infty} (\lambda_{\xi^{-1}, z_0})^D dv_{\xi^2} = \int_{S_\infty} dv_{\xi\eta} + \int_{S_\infty} dv_{\xi^2} + \int_{S_\infty} dv_{\xi\eta^{-1}}$   
 (d)  $\int_{S_\infty} (\lambda_{\xi^{-1}, z_0})^D dv_{\xi\eta} = \int_{S_\infty} dv_\eta$
- (4) (a)  $\int_{S_\infty} (\lambda_{\eta^{-1}, z_0})^D dv_\eta = 1 - \int_{S_\infty} dv_{\eta^{-1}\xi} - \int_{S_\infty} dv_{\eta^{-2}} - \int_{S_\infty} dv_{\eta^{-1}\xi^{-1}}$   
 (b)  $\int_{S_\infty} (\lambda_{\eta, z_0})^D dv_{\eta^{-2}} = \int_{S_\infty} dv_{\eta^{-1}\xi^{-1}} + \int_{S_\infty} dv_{\eta^{-1}\xi} + \int_{S_\infty} dv_{\eta^{-2}}$   
 (c)  $\int_{S_\infty} (\lambda_{\eta, z_0})^D dv_{\eta^{-1}\xi} = \int_{S_\infty} dv_{\xi\eta} + \int_{S_\infty} dv_{\xi^2} + \int_{S_\infty} dv_{\xi\eta^{-1}}$   
 (d)  $\int_{S_\infty} (\lambda_{\eta, z_0})^D dv_{\eta^{-1}\xi^{-1}} = \int_{S_\infty} dv_{\xi^{-1}}$
- (5) (a)  $\int_{S_\infty} (\lambda_{\eta^{-1}\xi^{-1}, z_0})^D dv_{\xi\eta} = 1 - \int_{S_\infty} dv_{\eta^{-1}\xi} - \int_{S_\infty} dv_{\eta^{-2}} - \int_{S_\infty} dv_{\eta^{-1}\xi^{-1}}$   
 (b)  $\int_{S_\infty} (\lambda_{\eta^{-1}\xi^{-1}, z_0})^D dv_{\xi\eta^{-1}} = \int_{S_\infty} dv_{\eta^{-2}}$   
 (c)  $\int_{S_\infty} (\lambda_{\eta^{-1}\xi^{-1}, z_0})^D dv_{\xi^2} = \int_{S_\infty} dv_{\eta^{-1}\xi}$   
 (d)  $\int_{S_\infty} (\lambda_{\xi\eta, z_0})^D dv_{\eta^{-1}\xi^{-1}} = 1 - \int_{S_\infty} dv_{\xi\eta} - \int_{S_\infty} dv_{\xi^2} - \int_{S_\infty} dv_{\xi\eta^{-1}}$   
 (e)  $\int_{S_\infty} (\lambda_{\xi\eta, z_0})^D dv_{\eta^{-1}\xi} = \int_{S_\infty} dv_{\xi^2}$   
 (f)  $\int_{S_\infty} (\lambda_{\xi\eta, z_0})^D dv_{\eta^{-2}} = \int_{S_\infty} dv_{\xi\eta^{-1}}$
- (6) (a)  $\int_{S_\infty} (\lambda_{\eta^{-1}\xi, z_0})^D dv_{\xi^{-1}} = 1 - \int_{S_\infty} dv_{\eta^{-1}\xi}$   
 (b)  $\int_{S_\infty} (\lambda_{\xi\eta^{-1}, z_0})^D dv_\eta = 1 - \int_{S_\infty} dv_{\xi\eta^{-1}}$   
 (c)  $\int_{S_\infty} (\lambda_{\xi^2, z_0})^D dv_{\xi^{-1}} = 1 - \int_{S_\infty} dv_{\xi^2}$   
 (d)  $\int_{S_\infty} (\lambda_{\eta^{-2}, z_0})^D dv_\eta = \int_{S_\infty} dv_{\eta^{-2}}$

**Proof** Since  $\Gamma$  acts freely on  $\mathbb{H}^3$  and it can be decomposed as in (6), the orbit  $W = \Gamma \cdot z_0$  is a disjoint union

$$(7) \quad W^\dagger = \{z_0\} \cup V_0 \cup \bigcup_{\psi \in \Psi^\dagger} V_\psi,$$

where  $V_0 = \{\gamma \cdot z_0 : \gamma \in \Psi_r^\dagger\}$  and  $V_\psi = \{\gamma \cdot z_0 : \gamma \in J_\psi\}$ . Note that  $V_0$  is the finite set  $V_\xi \cup V_{\eta^{-1}} = \{\xi \cdot z_0\} \cup \{\eta^{-1} \cdot z_0\}$ . Let  $\mathcal{V}$  denote the finite collection of all sets of the form  $\bigcup_{\psi \in \Psi} V_\psi$ ,  $V_0 \cup \bigcup_{\psi \in \Psi} V_\psi$ ,  $\{z_0\} \cup \bigcup_{\psi \in \Psi} V_\psi$  or  $\{z_0\} \cup V_0 \cup \bigcup_{\psi \in \Psi} V_\psi$  for  $\Psi \subset \Psi^\dagger$ . We apply [6, Proposition 4.2] to  $W$  and  $\mathcal{V}$ .

Let  $D$  be a number in  $[0, 2]$  and  $(\mathcal{M}_V)_{V \in \mathcal{V}}$  be a family of conformal densities, for which [6, Proposition 4.2(i)–(iv)] are satisfied. Let  $\mathcal{M}_V = (\mu_{V,z})_{z \in \mathbb{H}^3}$ . We set  $\mathcal{M} = \mathcal{M}_W$  and  $\nu_\psi = \mu_{V_\psi, z_0}$  for each  $\psi \in \Psi^\dagger$ . By [6, Proposition 4.2(iii)],  $\mathcal{M}$  is  $\Gamma$ -invariant. By [6, Proposition 4.2(i)] and the definition of a conformal density, we have  $\mu_{z_0}(S_\infty) = \mu_{W, z_0}(S_\infty) \neq 0$ . Therefore, we may assume that  $\mu_{z_0}$  has total mass 1 after normalization, which gives Lemma 3.3(1).

By [6, Proposition 4.2(iv)], we have  $\mu_{\{z_0\}, z_0} = 0$  and  $\mu_{V_0, z_0} = 0$ . Applying [6, Proposition 4.2(ii)] to the disjoint union in (7), we obtain

$$\mu_{z_0} = \mu_{\{z_0\}, z_0} + \mu_{V_0, z_0} + \sum_{\psi \in \Psi^\dagger} \mu_{V_\psi, z_0}.$$

Hence, we get conclusion (2) of Lemma 3.3. In order to complete parts (3)–(6), we need to determine all of the group-theoretical relations between the sets of words  $J_\psi$  for  $\psi \in \Psi^\dagger$ . We know that  $\xi^{-1}\eta \in J_{\xi^{-1}}$ . Therefore, we have  $1 \in \eta^{-1}\xi J_{\xi^{-1}}$ . Let  $w$  be a word in  $J_{\xi^{-1}}$ . Then we have  $w = \xi^{-1}w_1$  for some  $w_1 \in \Gamma$ . We compute that  $\eta^{-1}\xi w = \eta^{-1}w_1$ . The first letter of  $w_1$  cannot be  $\xi$ . But it can be either  $\eta$ ,  $\eta^{-1}$  or  $\xi^{-1}$ . Assume that it is  $\eta$  and  $w_1 \neq \eta$ . Then we have  $w_1 = \eta w_2$  for some word  $w_2 \in \Gamma$ . The first letter of  $w_2$  cannot be  $\eta^{-1}$ , but it can be either  $\eta$ ,  $\xi^{-1}$  or  $\xi$ . Since we get  $\eta^{-1}\xi w = w_2$ , we derive that  $\{1\} \cup J_\eta \cup J_{\xi\eta^{-1}} \cup J_{\xi\eta} \cup J_{\xi^2} \cup J_{\xi^{-1}} \subset \eta^{-1}\xi J_{\xi^{-1}}$ . If the first letter of  $w_1$  is  $\eta^{-1}$ , then we get  $w_1 = \eta^{-1}w_2$  for some  $w_2 \in \Gamma$ . We see that  $\eta^{-1}\xi w = \eta^{-2}w_2$ . This means that  $J_{\eta^{-2}} \subset \eta^{-1}\xi J_{\xi^{-1}}$ . If the first letter of  $w_1$  is  $\xi^{-1}$ , then we get  $w_1 = \xi^{-1}w_2$  for some  $w_2 \in \Gamma$  which implies that  $\eta^{-1}\xi w = \eta^{-1}\xi^{-1}w_2$ . Therefore, we find that  $J_{\eta^{-1}\xi^{-1}} \subset \eta^{-1}\xi J_{\xi^{-1}}$ . In other words,  $\eta^{-1}\xi J_\eta$  contains every word in  $\Gamma$  except the ones start with  $\eta^{-1}\xi$ . Hence,  $\eta^{-1}\xi J_{\xi^{-1}} = \Gamma - J_{\eta^{-1}\xi}$ .

Similar computations show that  $\eta^{-1}\xi J_{\xi^{-1}} = \Gamma - J_{\xi\eta^{-1}}$ ,  $\xi^2 J_{\xi^{-1}} = \Gamma - J_{\xi^2}$  and  $\eta^{-2} J_\eta = \Gamma - J_{\eta^{-2}}$ .

It follows from the discussion above and definitions of the sets  $J_\psi$  for each  $\psi \in \Psi^\dagger$  that, for each row  $\gamma$ ,  $s(\gamma)$  and  $S(\gamma)$  of Table 1, the decomposition  $\Gamma_{\mathcal{D}^\dagger}$  of  $\Gamma = \langle \xi, \eta \rangle$



	$\gamma$	$s(\gamma)$	$S(\gamma)$
(3)	(a)	$\xi$	$\{\xi\eta, \xi^2, \xi\eta^{-1}\}$
	(b)	$\xi^{-1}$	$\{\xi\eta, \xi^2, \xi\eta^{-1}, \eta, \xi^{-1}\}$
	(c)	$\xi^{-1}$	$\{\eta, \xi^{-1}, \eta^{-1}\xi^{-1}, \eta^{-2}, \eta^{-1}\xi\}$
	(d)	$\xi^{-1}$	$\{\xi\eta, \xi^2, \xi\eta^{-1}, \xi^{-1}, \eta^{-1}\xi^{-1}, \eta^{-2}, \eta^{-1}\xi\}$
(4)	(a)	$\eta^{-1}$	$\{\eta^{-1}\xi^{-1}, \eta^{-2}, \eta^{-1}\xi\}$
	(b)	$\eta$	$\{\xi\eta, \xi^2, \xi\eta^{-1}, \eta, \xi^{-1}\}$
	(c)	$\eta$	$\{\eta, \xi^{-1}, \eta^{-1}\xi^{-1}, \eta^{-2}, \eta^{-1}\xi\}$
	(d)	$\eta$	$\{\xi\eta, \xi^2, \xi\eta^{-1}, \eta, \eta^{-1}\xi^{-1}, \eta^{-2}, \eta^{-1}\xi\}$
(5)	(a)	$\eta^{-1}\xi^{-1}$	$\{\eta^{-1}\xi^{-1}, \eta^{-2}, \eta^{-1}\xi\}$
	(b)	$\eta^{-1}\xi^{-1}$	$\{\xi\eta, \xi^2, \xi\eta^{-1}, \eta, \xi^{-1}, \eta^{-1}\xi^{-1}, \eta^{-2}\}$
	(c)	$\eta^{-1}\xi^{-1}$	$\{\xi\eta, \xi^2, \xi\eta^{-1}, \eta, \xi^{-1}, \eta^{-1}\xi^{-1}, \eta^{-1}\xi\}$
	(d)	$\xi\eta$	$\{\xi\eta, \xi^2, \xi\eta^{-1}\}$
	(e)	$\xi\eta$	$\{\xi\eta, \xi\eta^{-1}, \eta, \xi^{-1}, \eta^{-1}\xi^{-1}, \eta^{-2}, \eta^{-1}\xi\}$
	(f)	$\xi\eta$	$\{\xi\eta, \xi^2, \eta, \xi^{-1}, \eta^{-1}\xi^{-1}, \eta^{-2}, \eta^{-1}\xi\}$
(6)	(a)	$\eta^{-1}\xi$	$\{\eta^{-1}\xi\}$
	(b)	$\xi\eta^{-1}$	$\{\xi\eta^{-1}\}$
	(c)	$\xi^2$	$\{\xi^2\}$
	(d)	$\eta^{-2}$	$\{\eta^{-2}\}$

Table 1: Group-theoretical properties of the decomposition  $\Gamma_{\mathcal{D}^\dagger}$

has the group-theoretical property

$$(8) \quad \gamma J_{S(\gamma)} = \Gamma - J_{S(\gamma)}.$$

Let  $V_\Psi$  denote the union  $\bigcup_{\gamma \in \Psi} V_\gamma$  where  $\Psi$  is a subset of  $\Psi^\dagger \cup \Psi_r^\dagger$ . Using the group-theoretical relations in (8), we derive the relations

$$(9) \quad \gamma V_{s(\gamma)} = W - V_{S(\gamma)}$$

between the orbits  $V_{s(\gamma)}$  and  $V_{S(\gamma)}$ . Since  $W - V_{S(\gamma)} = V_{s(\gamma)} \in \mathcal{V}$ , [6, Proposition 4.2(iii)] gives  $\mathcal{M}_{V_{s(\gamma)}} = \gamma_\infty^*(\mathcal{M}_{W - V_{S(\gamma)}})$ . On the other hand, by [6, Proposition 4.2(ii)], we get  $\mathcal{M} = \mathcal{M}_{W - V_{S(\gamma)}} + \mathcal{M}_{V_{S(\gamma)}}$ . We combine the last two equalities to obtain  $\mathcal{M}_{V_{s(\gamma)}} = \gamma_\infty^*(\mathcal{M} - \mathcal{M}_{V_{S(\gamma)}})$ , which implies that

$$(10) \quad d\mu_{V_{s(\gamma)}, \gamma \cdot z_0} = d\left(\gamma_\infty^*\left(\mu_{z_0} - \sum_{\psi \in S(\gamma)} \nu_\psi\right)\right).$$

Since  $\mathcal{M}_{V_{s(\gamma)}}$  is a  $D$ -conformal density and  $d\mu_{V_{s(\gamma)}, \gamma \cdot z_0} = \lambda_{\gamma, z_0}^D d\mu_{V_{s(\gamma)}}$  (see [6, Proposition 2.4]), we obtain the equality

$$\int_{S_\infty} (\lambda_{\gamma, z_0})^D d\mu_{V_{s(\gamma)}} = 1 - \sum_{\psi \in \mathcal{S}(\gamma)} \int_{S_\infty} dv_\psi$$

for every row of Table 1 by equating the total masses of both sides of (10), which provides parts (3)–(6) of the lemma. □

The following is an analog of Theorem 1.1 for the set  $\Psi^\dagger \subset \Gamma = \langle \xi, \eta \rangle$ . Notice that Theorem 3.4 has no analog of part (3) of Theorem 1.1.

**Theorem 3.4** *Let  $\Gamma = \langle \xi, \eta \rangle$  be a free, geometrically infinite Kleinian group without parabolics. For any  $z_0 \in \mathbb{H}^3$ , let  $A_{z_0}$  be the area measure based at  $z_0$ . There is a family of Borel measures  $\{\nu_\psi\}_{\psi \in \Psi^\dagger}$  for  $\Psi^\dagger = \{\xi\eta, \xi^2, \xi\eta^{-1}, \eta, \xi^{-1}, \eta^{-1}\xi^{-1}, \eta^{-2}, \eta^{-1}\xi\}$  on  $S_\infty$  such that:*

- (1)  $A_{z_0} = \sum_{\psi \in \Psi^\dagger} \nu_\psi$ , where  $A_{z_0}$  is normalized so that  $A_{z_0}(S_\infty) = 1$ .
- (2)  $\int_{S_\infty} (\lambda_{\gamma, z_0})^2 dv_{s(\gamma)} = 1 - \sum_{\psi \in \mathcal{S}(\gamma)} \int_{S_\infty} dv_\psi$  for each row of Table 1.

**Proof** By the conclusions of [6, Propositions 6.9 and 3.9] and tameness [1; 4], we have that every  $\Gamma$ -invariant  $D$ -conformal density  $\mathcal{M}$  is a constant multiple of the area density  $\mathcal{A}$ , ie  $D = 2$ . By Lemma 3.3(1), we get  $\mathcal{M} = \mathcal{A}$ . Then (2) follows from Lemma 3.3(3)–(6). □

We shall use Theorem 3.4 together with Lemma 1.2 to produce the displacement functions for the decomposition  $\Gamma_{\mathcal{D}^\dagger}$ . In the rest of this paper, we will use the bijection  $p: \Psi^\dagger \rightarrow I^\dagger$  defined by

$$(11) \quad \begin{aligned} \xi\eta &\mapsto 1, & \xi^2 &\mapsto 2, & \xi\eta^{-1} &\mapsto 3, & \eta &\mapsto 7, \\ \eta^{-1}\xi^{-1} &\mapsto 4, & \eta^{-2} &\mapsto 5, & \eta^{-1}\xi &\mapsto 6, & \xi^{-1} &\mapsto 8, \end{aligned}$$

to enumerate the displacement functions and their variables. We have:

**Proposition 3.5** *Let  $\Gamma = \langle \xi, \eta \rangle$  be a free Kleinian group. For any  $z_0 \in \mathbb{H}^3$  and for each  $\gamma \in \{\xi, \eta, \xi^{-1}, \eta^{-1}, \xi\eta, \eta^{-1}\xi^{-1}\}$ , the expression  $e^{2 \text{dist}(z_0, \gamma \cdot z_0)}$  is bounded below by  $f_i(\mathbf{x})$  or  $g_j(\mathbf{x})$  for at least one of  $f_i$  or  $g_j$  for  $i \in I^\dagger = \{1, 2, 3, 4, 5, 6, 7, 8\}$  and  $j \in \{1, 2, 3, 4, 5, 6\}$  in the list*

$$\begin{aligned}
 f_1(\mathbf{x}) &= \frac{1-x_4-x_5-x_6}{x_4+x_5+x_6} \cdot \frac{1-x_1}{x_1}, & f_2(\mathbf{x}) &= \frac{1-x_4-x_5-x_6-x_7-x_8}{x_4+x_5+x_6+x_7+x_8} \cdot \frac{1-x_2}{x_2}, \\
 f_4(\mathbf{x}) &= \frac{1-x_1-x_2-x_3}{x_1+x_2+x_3} \cdot \frac{1-x_4}{x_4}, & f_3(\mathbf{x}) &= \frac{1-x_1-x_2-x_3-x_7-x_8}{x_1+x_2+x_3+x_7+x_8} \cdot \frac{1-x_3}{x_3}, \\
 f_7(\mathbf{x}) &= \frac{1-x_4-x_5-x_6}{x_4+x_5+x_6} \cdot \frac{1-x_7}{x_7}, & f_5(\mathbf{x}) &= \frac{1-x_1-x_2-x_3-x_7-x_8}{x_1+x_2+x_3+x_7+x_8} \cdot \frac{1-x_5}{x_5}, \\
 f_8(\mathbf{x}) &= \frac{1-x_1-x_2-x_3}{x_1+x_2+x_3} \cdot \frac{1-x_8}{x_8}, & f_6(\mathbf{x}) &= \frac{1-x_4-x_5-x_6-x_7-x_8}{x_4+x_5+x_6+x_7+x_8} \cdot \frac{1-x_6}{x_6}
 \end{aligned}$$

(the functions above are produced from rows (3)(a)–(c), (4)(a)–(c), (5)(a) and (5)(d) of Table 1) and

$$\begin{aligned}
 g_1(\mathbf{x}) &= \frac{x_7}{1-x_7} \cdot \frac{1-x_1}{x_1}, & g_2(\mathbf{x}) &= \frac{x_6}{1-x_6} \cdot \frac{1-x_2}{x_2}, \\
 g_3(\mathbf{x}) &= \frac{x_5}{1-x_5} \cdot \frac{1-x_3}{x_3}, & g_4(\mathbf{x}) &= \frac{x_8}{1-x_8} \cdot \frac{1-x_4}{x_4}, \\
 g_5(\mathbf{x}) &= \frac{x_3}{1-x_3} \cdot \frac{1-x_5}{x_5}, & g_6(\mathbf{x}) &= \frac{x_2}{1-x_2} \cdot \frac{1-x_6}{x_6}
 \end{aligned}$$

(these come from rows (3)(d), (4)(d), (5)(b), (5)(c), (5)(e) and (5)(f) of Table 1) for some

$$\mathbf{x} = (x_1, \dots, x_8) \in \Delta^7 = \left\{ \mathbf{x} \in \mathbb{R}_+^8 : \sum_{i=1}^8 x_i = 1 \right\}.$$

Under the same hypothesis on  $\Gamma$ , for any  $z_0 \in \mathbb{H}^3$  and for each  $\gamma \in \{\xi^2, \eta^{-2}, \xi\eta^{-1}, \eta\xi\}$ , the expression  $e^{2 \operatorname{dist}(z_0, \gamma \cdot z_0)}$  is bounded below by  $h_i(\mathbf{x})$  for at least one of  $h_i$  from the list

$$\begin{aligned}
 h_1(\mathbf{x}) &= \frac{1-x_6}{x_6} \cdot \frac{1-x_8}{x_8}, & h_2(\mathbf{x}) &= \frac{1-x_2}{x_2} \cdot \frac{1-x_8}{x_8}, \\
 h_3(\mathbf{x}) &= \frac{1-x_5}{x_5} \cdot \frac{1-x_7}{x_7}, & h_4(\mathbf{x}) &= \frac{1-x_3}{x_3} \cdot \frac{1-x_7}{x_7},
 \end{aligned}$$

for some  $i \in \{1, 2, 3, 4\}$  and  $\mathbf{x} \in \Delta^7$  (the functions  $h_i$  are produced from rows (6)(a)–(d) of Table 1).

**Proof** By Lemma 3.3(1), we have  $0 \leq \nu_\psi(S_\infty) \leq 1$  for every  $\psi \in \Psi^\dagger$ . We aim to show that  $0 < \nu_\psi(S_\infty) < 1$  for any  $\psi \in \Psi^\dagger$ . First assume to the contrary that  $\nu_{\xi^{-1}}(S_\infty) = 0$ . Applying Theorem 3.4(2) to row (6)(a) of Table 1 gives  $\nu_{\eta^{-1}\xi}(S_\infty) = 1$ . By Theorem 3.4(1), we see that  $\nu_\psi(S_\infty) = 0$  for every  $\psi \in \Psi^\dagger - \{\eta^{-1}\xi\}$ . Using that  $\nu_\eta(S_\infty) = 0$  and applying Theorem 3.4(2) to row (6)(b) of Table 1 shows that  $\nu_{\xi\eta^{-1}}(S_\infty) = 1$ , a contradiction. A similar argument can be repeated for  $\nu_\eta(S_\infty)$  by exchanging the roles of  $\xi^{-1}$  and  $\eta$  above. Therefore, we have  $\nu_{\xi^{-1}}(S_\infty) \neq 0$  and  $\nu_\eta(S_\infty) \neq 0$ .

Assume that  $\nu_{\psi_0}(S_\infty) = 0$  for a given  $\psi_0 \in \{\xi\eta, \xi\eta^{-1}, \xi^2, \eta^{-1}\xi, \eta^{-1}\xi^{-1}, \eta^{-2}\}$ . Consider the lists

$$\begin{aligned} &(\xi^{-1}, \xi\eta, \Psi^\dagger - \{\eta\}, \xi^{-1}), & (\xi^{-1}, \xi\eta^{-1}, \Psi^\dagger - \{\eta^{-1}\xi, \eta^{-1}\xi^{-1}, \eta^{-2}\}, \eta), \\ &(\xi^{-1}, \xi^2, \Psi^\dagger - \{\xi\}, \xi^{-1}), & (\eta, \eta^{-2}, \Psi^\dagger - \{\eta^{-1}\xi, \eta^{-1}\xi^{-1}, \eta^{-2}\}, \xi^{-1}), \\ &(\eta, \eta^{-1}\xi^{-1}, \Psi^\dagger - \{\xi^{-1}\}, \eta), & (\eta, \eta^{-1}\xi, \Psi^\dagger - \{\xi\eta^{-1}, \xi\eta, \xi^2\}, \xi^{-1}), \end{aligned}$$

where each entry in a list is assigned for  $\gamma_0, \psi_0, S(\gamma_0), \psi_1$ , respectively. By applying Theorem 3.4(2) to Table 1 with  $\psi_0 = s(\gamma_0)$ , we get  $\sum_{\psi \in S(\gamma_0)} \nu_\psi = 1$ . We have  $\psi_1 \notin S(\gamma_0)$ . Therefore, we obtain  $\nu_{\psi_1}(S_\infty) = 0$  for some  $\psi_1 \in \{\xi^{-1}, \eta\}$ , a contradiction. As a result, we conclude that  $0 < \nu_\psi(S_\infty) < 1$  for every  $\psi \in \Psi^\dagger$ . Since we have  $\psi = s(\gamma)$  for some  $\gamma$  in Table 1 and  $S(\gamma) \subset \Psi^\dagger$ , we also conclude

$$0 < \int_{S_\infty} (\lambda_{\gamma, z_0})^2 d\mu_{V_{s(\gamma)}} = 1 - \sum_{\psi \in S(\gamma)} \int_{S_\infty} d\nu_\psi < 1$$

by Theorem 3.4(2). In other words,  $\nu_{s(\gamma)}$  and  $\int_{S_\infty} \lambda_{\gamma, z_0}^2 d\mu_{V_{s(\gamma)}}$  satisfy the hypothesis of Lemma 1.2 for every  $\gamma$  in Table 1.

We apply Lemma 1.2 to every row of Table 1 with  $\nu = \nu_{s(\gamma)}$ ,  $a = \nu_{s(\gamma)}(S_\infty)$  and  $b = \int_{S_\infty} \lambda_{\gamma, z_0}^2 d\mu_{V_{s(\gamma)}}$ . Using Theorem 3.4(2), we calculate the lower bounds as

$$(12) \quad e^{2 \text{dist}(z_0, \gamma \cdot z_0)} \geq \frac{(1 - \sum_{\psi \in S(\gamma)} m_p(\psi)) \cdot (1 - m_p(s(\gamma)))}{(\sum_{\psi \in S(\gamma)} m_p(\psi)) \cdot m_p(s(\gamma))},$$

where  $\int_{S_\infty} d\nu_\psi = m_p(s(\gamma))$  for the bijection  $p$  in (11). Upon replacing each constant  $m_p(s(\gamma))$  appearing in (12) with the variable  $x_{p(s(\gamma))}$  we obtain the functions listed in the proposition. □

Note that we have 18 lower bounds given in the expression (12) for the displacements under the isometries in  $\Psi_r^\dagger \cup \Psi^\dagger$  because there are 18 group-theoretical relations listed in (8). Since we are interested in the displacements under the isometries in  $\Gamma_\dagger = \{\xi, \eta, \xi\eta\}$ , we will concentrate on the first 14 displacement functions  $f_1, f_2, \dots, f_8, g_1, g_2, \dots, g_6$  in Proposition 3.5 for the proofs of Lemmas 4.2, 4.3, 4.4, 4.5, 4.6, 4.7, 4.9, 4.11, 4.12, 4.13 and Theorems 4.14, 4.1 and 5.1.

### 4 Lower bound for $\max\{\text{dist}(z, \gamma \cdot z) : \gamma \in \Gamma_\dagger\}$ when $\Lambda_{\Gamma \cdot z} = S^2$

Let  $\mathcal{F}^\dagger = \{f_1, f_2, \dots, f_8, g_1, g_2, \dots, g_6\}$ . The constants on the right-hand side of the inequalities in (12) can be considered as the values of the functions in  $\mathcal{F}^\dagger$  at the point  $m_\dagger = (m_1, m_2, \dots, m_8) \in \Delta^7 = \{(x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8) \in \mathbb{R}_+^8 : \sum_{i=1}^8 x_i = 1\}$ .

When  $\Gamma = \langle \xi, \eta \rangle$  is geometrically infinite, the lower bound given in Theorem 5.1 for the displacements under the isometries in  $\Gamma^\dagger = \langle \xi, \eta, \xi\eta \rangle$  follows from the calculation of the infimum of the maximum of the functions in  $\mathcal{F}^\dagger$  over the simplex  $\Delta^7$ . Therefore, in this section, we aim to prove the statement below:

**Theorem 4.1** *Let  $G^\dagger: \Delta^7 \rightarrow \mathbb{R}$  be the function defined by  $x \mapsto \max\{f(x) : f \in \mathcal{F}^\dagger\}$ . Then  $\inf_{x \in \Delta^7} G^\dagger(x) = 5 + 3\sqrt{2}$ .*

To this purpose, we shall show that it is enough to calculate the infimum of the maximum of the first eight  $f_1, f_2, \dots, f_8$  of the displacement functions in  $\mathcal{F}^\dagger$ . Let  $I^\dagger = \{1, 2, 3, 4, 5, 6, 7, 8\}$ . Then we first state the following:

**Lemma 4.2** *Let  $F^\dagger: \Delta^7 \rightarrow \mathbb{R}$  be the function defined by  $x \mapsto \max\{f_i(x) : i \in I^\dagger\}$ . Then  $\alpha_* = \inf_{x \in \Delta^7} F^\dagger(x)$  is attained in  $\Delta^7$  and satisfies  $9 \leq \alpha_* \leq 5 + 3\sqrt{2}$ .*

**Proof** It is clear that  $\inf_{x \in \Delta^7} \{\max(f_7(x), f_8(x))\} \leq \inf_{x \in \Delta^7} F^\dagger(x)$ . We apply the substitution  $X_1 = x_7, X_2 = x_8, X_3 = x_1 + x_2 + x_3, X_4 = x_4 + x_5 + x_6$ . Then we see that  $f_7(x) = \sigma(X_4)\sigma(X_1)$  and  $f_8(x) = \sigma(X_3)\sigma(X_2)$ , where  $\sum_{i=1}^4 X_i = 1$  and  $\sigma(x) = 1/x - 1$  for  $x \in (0, 1)$ . By Lemma 2.1, we obtain that  $9 \leq \alpha_*$ .

Let  $\Delta_n = \{x \in \Delta^7 : 1/n \leq x_i \leq 1 - 1/(2n) \text{ for } i \in I^\dagger\}$  of  $\Delta^7$  for every  $n \geq 2$ . Note that  $\Delta_{n+1} \subset \Delta_n$ . The function  $F^\dagger$  has an absolute minimum  $F^\dagger(x_n)$  at some point  $x_n \in \Delta_n$ . The sequence  $(F^\dagger(x_n))_{n=1}^\infty$  tends to an infimum because it is a decreasing sequence which is bounded below by 9.

We claim that the sequence  $(x_n)_{n=1}^\infty$  cannot have a limit point on the boundary of  $\Delta^7$ . Assume on the contrary that  $x_n \rightarrow \mathbf{b} \in \partial\Delta^7$  as  $n \rightarrow \infty$ . If  $(b_1, b_2, \dots, b_8)$  denotes the coordinates of the point  $\mathbf{b}$ , then  $b_i = 0$  for some  $i \in I^\dagger$ . Let us assume that  $b_i = 0$  for some  $i \in \{1, 7\}$ . Then using the function  $f_i$ , we conclude that  $b_4 + b_5 + b_6 = 1$  because otherwise  $(f_i(x_n))_{n=1}^\infty$  would tend to infinity. But it is supposed to be tending to an infimum of  $F^\dagger$ . Therefore, we must have  $b_j = 0$  for every  $j \in \{2, 3, 8\}$ . Then, we get that  $b_4 + b_5 + b_6 + b_7 + b_8 = 1, b_1 + b_2 + b_3 + b_7 + b_8 = 1$  and  $b_1 + b_2 + b_3 = 1$ , because otherwise  $(f_j(x_n))_{n=1}^\infty$  would tend to infinity when it is supposed to tend to an infimum of  $F^\dagger$ . In any case, we obtain that  $b_4 + b_5 + b_6 + b_i > 1$  for some  $i \in \{1, 2, 3, 7, 8\}$ . This is a contradiction. Therefore,  $b_1 \neq 0$  and  $b_7 \neq 0$ . Similar arguments with suitably chosen displacement functions show that  $b_i \neq 0$  for every  $i \in \{2, 3, 4, 5, 6, 8\}$ . Hence, the sequence  $(x_n)_{n=1}^\infty$  cannot have a limit point on the boundary of  $\Delta^7$ . Then there exists a positive integer  $n_0$  so that  $x_n = x_{n+1}$  for every  $n \geq n_0$ . Otherwise we would have a subsequence  $(x_{n_j})_{j=1}^\infty$  of the sequence  $(x_n)_{n=1}^\infty$  so that  $x_{n_j} \in \Delta_{n_{j+1}} - \Delta_{n_j}$  for every integer  $j > 0$ . Since we have  $\bigcup_{n=2}^\infty \Delta_n = \Delta^7$ ,

there exists a subsequence of  $(\mathbf{x}_{n_j})_{j=1}^\infty$  which has a limit point on  $\partial\Delta^7$ , a contradiction. As a result,  $\inf_{\mathbf{x} \in \Delta^7} F^\dagger(\mathbf{x})$  is attained at some point in  $\Delta^7$ , ie  $\alpha_* = \min_{\mathbf{x} \in \Delta^7} F^\dagger(\mathbf{x})$ . Let  $x_i = (\sqrt{2} - 1)/2$  for  $i = 1, 4, 7, 8$  and  $x_i = (3 - 2\sqrt{2})/4$  for  $i = 2, 3, 5, 6$ . Then  $\bar{\mathbf{x}} = (x_i)_{i \in I^\dagger}$  is a point in  $\Delta^7$  such that  $f_i(\bar{\mathbf{x}}) = 5 + 3\sqrt{2}$  for every  $i \in I^\dagger$ . Therefore, we get  $F^\dagger(\bar{\mathbf{x}}) = 5 + 3\sqrt{2} \geq \alpha_*$ .  $\square$

In the rest of this section, we will consider  $\Delta^7$  as a submanifold of  $\mathbb{R}^8$ . The tangent space  $T_{\mathbf{x}}\Delta^7$  at any  $\mathbf{x} \in \Delta^7$  consists of vectors whose coordinates sum to 0. Note that each displacement function  $f_i$  for  $i \in I^\dagger$  is smooth in an open neighborhood of  $\Delta^7$ . Therefore, the directional derivative of  $f_i$  in the direction of any  $\vec{v} \in T_{\mathbf{x}}\Delta^7$  is given by  $\nabla f_i(\mathbf{x}) \cdot \vec{v}$  for any  $i \in I^\dagger$ .

We shall use the identity  $\sum_{i=1}^8 x_i = 1$  to rewrite the formulas of the functions  $f_i$  given in Proposition 3.5 in various ways in the proofs of lemmas below to suit our purposes. Although they do not take the same values on all of  $\mathbb{R}^8$ , we will abuse notation and call the rewritten functions by  $f_i$ , which agree with the originals on  $\Delta^7$ .

### 4.1 Relationships between the displacement functions $f_1, f_2, \dots, f_8$

By Lemma 4.2, we know that  $\alpha_*$  is attained by a displacement function  $f_i$  for some  $i \in I^\dagger$ . In fact, it is possible to see that more than one function in  $\{f_1, \dots, f_8\}$  attain the value  $\alpha_*$ . In other words, we have:

**Lemma 4.3** *Let  $\alpha_* = \inf_{\mathbf{x} \in \Delta^7} \max\{f_1(\mathbf{x}), \dots, f_8(\mathbf{x})\}$ , where  $f_i$  for  $i \in I^\dagger$  are as in Proposition 3.5. At any  $\mathbf{x}^* \in \Delta^7$  such that  $F^\dagger(\mathbf{x}^*) = \alpha_*$ , there exist at least two functions  $f_i, f_j$  such that  $f_i(\mathbf{x}^*) = f_j(\mathbf{x}^*)$  for  $i \neq j$ , where  $i, j \in I^\dagger = \{1, 2, 3, 4, 5, 6, 7, 8\}$ .*

**Proof** Observe that for each function  $f_i$  for  $i \in I^\dagger$  there is a variable  $x_j$  such that the first-order partial derivative of  $f_i$  with respect to  $x_j$  at  $\mathbf{x}$  is 0 for every  $\mathbf{x} \in \Delta^7$ . But the first-order partial derivatives of  $f_i$  with respect to  $x_i$  are strictly negative at every  $\mathbf{x} \in \Delta^7$ . These facts imply that  $\nabla f_i$  is not a scalar multiple of the perpendicular  $\langle 1, 1, \dots, 1 \rangle$  to  $T_{\mathbf{x}}\Delta^7$  for any  $i \in I^\dagger$ . Therefore, none of the functions  $f_1, f_2, \dots, f_8$  has a local extremum on  $\Delta^7$ .

If  $f_i(\mathbf{x}^*) \neq f_j(\mathbf{x}^*)$  for every  $i \neq j$ , then the set  $\{f_1(\mathbf{x}^*), f_2(\mathbf{x}^*), \dots, f_8(\mathbf{x}^*)\}$  has a unique largest element. By renumbering the functions, we may assume that  $f_1(\mathbf{x}^*)$  is the largest value, ie  $f_1(\mathbf{x}^*) = \alpha_*$ . By the continuity of  $F^\dagger$ , there exists a neighborhood  $U$  of  $\mathbf{x}^*$  contained in  $\Delta^7$  so that  $F^\dagger(\mathbf{x}) = f_1(\mathbf{x})$  for every  $\mathbf{x} \in U$ . Since  $F^\dagger$  has a minimum at  $\mathbf{x}^*$ , then  $f_1$  must have a local minimum at  $\mathbf{x}^*$ , a contradiction. The lemma follows.  $\square$

Next, we will consider the cases in which  $f_i$  and  $f_j$  in Lemma 4.3 are in the sets  $\{f_1, f_2, f_3\}$ ,  $\{f_4, f_5, f_6\}$  and  $\{f_7, f_8\}$ , respectively:

**Lemma 4.4** *Let  $\alpha_* = \inf_{\mathbf{x} \in \Delta^7} \max\{f_1(\mathbf{x}), \dots, f_8(\mathbf{x})\}$ , where  $f_i$  for  $i \in I^\dagger$  are as in Proposition 3.5. At any  $\mathbf{x}^* \in \Delta^7$  such that  $F^\dagger(\mathbf{x}^*) = \alpha_*$ , we have either*

- (1)  $f_l(\mathbf{x}^*) = \alpha_*$  for all  $l \in I_1 = \{1, 2, 3\}$  or
- (2)  $f_l(\mathbf{z}) < \alpha_*$  for all  $l \in I_1$  and  $f_j(\mathbf{z}) = f_j(\mathbf{x}^*)$  for all  $j \in \mathcal{I}_1 = \{4, 5, 6, 7, 8\}$  for some  $\mathbf{z} \in \Delta^7$  such that  $F^\dagger(\mathbf{z}) = F^\dagger(\mathbf{x}^*)$ .

**Proof** Assume that part (1) of the lemma does not hold at  $\mathbf{x}^*$ . If  $f_i(\mathbf{x}^*) < \alpha_*$  for every  $i \in I_1$ , the point  $\mathbf{z} = \mathbf{x}^*$  satisfies part (2). Then it is enough to consider the case  $f_i(\mathbf{x}^*) < f_j(\mathbf{x}^*) \leq f_k(\mathbf{x}^*) = \alpha_*$ , and  $f_l(\mathbf{x}^*) \leq \alpha_*$  for  $l \in \mathcal{I}_1$ , where  $i, j, k \in I_1$  such that  $i \neq j, j \neq k, i \neq k$ . Let us define the vectors  $\vec{u}_1^2, \vec{u}_1^3$  and  $\vec{u}_2^3$  as

$$\langle -1, 1, 0, 0, 0, 0, 0 \rangle, \quad \langle -1, 0, 1, 0, 0, 0, 0 \rangle \quad \text{and} \quad \langle 0, -1, 1, 0, 0, 0, 0 \rangle$$

in  $T_{\mathbf{x}^*}\Delta^7$ , respectively. Also let  $\vec{u}_2^1 = -\vec{u}_1^2, \vec{u}_3^1 = -\vec{u}_1^3$  and  $\vec{u}_3^2 = -\vec{u}_2^3$ .

Using the identity  $x_k = 1 - \sum_{n=1, n \neq k}^8 x_n$ , we calculate the directional derivatives of all of the functions  $f_1, f_2, \dots, f_8$  in the direction of the vector  $\vec{u}_i^j$ . Note that none of the functions  $f_4, f_5, \dots, f_8$  contains the variables  $x_1, x_2$  or  $x_3$ . For every  $\mathbf{x} \in \Delta^7$  and for every  $l \in \mathcal{I}_1$  we see that

$$\nabla f_i(\mathbf{x}) \cdot \vec{u}_i^j > 0, \quad \nabla f_j(\mathbf{x}) \cdot \vec{u}_i^j < 0, \quad \nabla f_k(\mathbf{x}) \cdot \vec{u}_i^j = 0, \quad \nabla f_l(\mathbf{x}) \cdot \vec{u}_i^j = 0,$$

which implies that the values of  $f_j$  and  $f_k$  decrease along a line segment in the direction of  $\vec{v} = \vec{u}_i^j + \vec{u}_i^k$ . The values of  $f_l$  are constant along this segment, and for a short distance along  $\vec{v}$  the values of  $f_i$  is smaller than those of  $f_j$  and  $f_k$ . Therefore there exists a point  $\mathbf{z}$  on this line segment satisfying part (2) of the lemma.  $\square$

Analogous results hold for the displacement functions in  $\{f_4, f_5, f_6\}$  and  $\{f_7, f_8\}$ . In particular, we have the following:

**Lemma 4.5** *Let  $\alpha_* = \inf_{\mathbf{x} \in \Delta^7} \max\{f_1(\mathbf{x}), \dots, f_8(\mathbf{x})\}$ , where  $f_i$  for  $i \in I^\dagger$  are as in Proposition 3.5. At any  $\mathbf{x}^* \in \Delta^7$  such that  $F^\dagger(\mathbf{x}^*) = \alpha_*$ , we have either*

- (1)  $f_l(\mathbf{x}^*) = \alpha_*$  for all  $l \in I_2 = \{4, 5, 6\}$  or
- (2)  $f_l(\mathbf{z}) < \alpha_*$  for all  $l \in I_2$  and  $f_j(\mathbf{z}) = f_j(\mathbf{x}^*)$  for all  $j \in \mathcal{I}_2 = \{1, 2, 3, 7, 8\}$  for some  $\mathbf{z} \in \Delta^7$  such that  $F^\dagger(\mathbf{z}) = F^\dagger(\mathbf{x}^*)$ .

**Lemma 4.6** Let  $\alpha_* = \inf_{x \in \Delta^7} \max\{f_1(x), \dots, f_8(x)\}$ , where  $f_i$  for  $i \in I^\dagger$  are as in Proposition 3.5. At any  $x^* \in \Delta^7$  such that  $F^\dagger(x^*) = \alpha_*$ , we have either

- (1)  $f_l(x^*) = \alpha_*$  for all  $l \in I_3 = \{7, 8\}$  or
- (2)  $f_l(z) < \alpha_*$  for all  $l \in I_3$  and  $f_j(z) = f_j(x^*)$  for all  $j \in \mathcal{I}_3 = \{1, 2, 3, 4, 5, 6\}$  for some  $z \in \Delta^7$  such that  $F^\dagger(z) = F^\dagger(x^*)$ .

The proof of Lemma 4.4 applies, mutatis mutandis, to prove Lemmas 4.5 and 4.6. In particular, using each identity  $x_k = 1 - \sum_{n=1, n \neq k}^8 x_n$  for  $k \in I_2 \cup I_3$ , we perturb in the directions of the vectors  $\vec{u}_4^5 = \langle 0, 0, 0, -1, 1, 0, 0, 0 \rangle$ ,  $\vec{u}_4^6 = \langle 0, 0, 0, -1, 0, 1, 0, 0 \rangle$  and  $\vec{u}_5^6 = \langle 0, 0, 0, 0, -1, 1, 0, 0 \rangle$  for Lemma 4.5, and perturb in the direction of the vectors  $\vec{u}_7^8 = \langle 0, 0, 0, 0, 0, 0, -1, 1 \rangle$  and  $\vec{u}_8^7 = \langle 0, 0, 0, 0, 0, 0, 1, -1 \rangle$  for Lemma 4.6. Lemmas 4.4, 4.5 and 4.6 imply the following:

**Lemma 4.7** Let  $\alpha_* = \inf_{x \in \Delta^7} \max\{f_1(x), \dots, f_8(x)\}$ , where  $f_i$  for  $i \in I^\dagger$  are as in Proposition 3.5. There exists a point  $x^* \in \Delta^7$  which satisfies one of the cases (I), (II), (III), (IV) or (V), where

- (I)  $f_j(x^*) < \alpha_*$  for  $j \in \mathcal{I}_1 = \{4, 5, 6, 7, 8\}$  and  $f_j(x^*) = \alpha_*$  for  $j \in I_1 = \{1, 2, 3\}$ ,
- (II)  $f_j(x^*) < \alpha_*$  for  $j \in \mathcal{I}_3 = \{1, 2, 3, 4, 5, 6\}$  and  $f_j(x^*) = \alpha_*$  for  $j \in I_3 = \{7, 8\}$ ,
- (III)  $f_j(x^*) = \alpha_*$  for  $j \in \mathcal{I}_3 = \{1, 2, 3, 4, 5, 6\}$  and  $f_j(x^*) < \alpha_*$  for  $j \in I_3 = \{7, 8\}$ ,
- (IV)  $f_j(x^*) = \alpha_*$  for  $j \in \mathcal{I}_2 = \{1, 2, 3, 7, 8\}$  and  $f_j(x^*) < \alpha_*$  for  $j \in I_2 = \{4, 5, 6\}$ ,
- (V)  $f_j(x^*) = \alpha_*$  for  $j \in I^\dagger = \{1, 2, 3, 4, 5, 6, 7, 8\}$ .

**Proof** Let  $x \in \Delta^7$  be a point such that  $F^\dagger(x) = \alpha_*$ . First, assume that  $f_i(x) < \alpha_*$  for some  $i \in I_1$ . By Lemma 4.4, there exists a point  $z_1 \in \Delta^7$  with  $f_i(z_1) < \alpha_*$  for all  $i \in I_1$  and  $f_j(z_1) \leq \alpha_*$  for all  $j \in \mathcal{I}_1$  with  $F^\dagger(z_1) = \alpha_*$ .

If  $f_j(z_1) < \alpha_*$  for some  $j \in I_2$ , there exists a point  $z_2 \in \Delta^7$  with  $f_i(z_1) = f_i(z_2) < \alpha_*$  for all  $i \in I_1$ ,  $f_j(z_2) < \alpha_*$  for all  $j \in I_2$ ,  $f_k(z_1) = f_k(z_2)$  for all  $k \in I_3$  and  $F^\dagger(z_2) = \alpha_*$  by Lemma 4.5. We must have  $f_k(z_2) = \alpha_*$  for all  $k \in I_3$  by Lemma 4.3. Thus  $x^* = z_2$  satisfies case (II). Assume that  $f_j(z_1) = \alpha_*$  for all  $j \in I_2$ . Let  $T_1: \Delta^7 \rightarrow \Delta^7$  be the transformation

$$(13) \quad \begin{aligned} x_1 &\mapsto x_4, & x_2 &\mapsto x_5, & x_3 &\mapsto x_6, & x_4 &\mapsto x_1, \\ x_5 &\mapsto x_2, & x_6 &\mapsto x_3, & x_7 &\mapsto x_8, & x_8 &\mapsto x_7. \end{aligned}$$

If  $f_k(z_1) = \alpha_*$  for all  $k \in I_3$ , then  $x^* = T_1(z_1)$  satisfies case (IV). Otherwise, there exists a point  $z_2 \in \Delta^7$  such that  $f_i(z_2) = f_i(z_1)$  for  $i \in I_1$ ,  $f_j(z_2) = f_j(z_1)$  for  $j \in I_2$  and  $f_k(z_2) < \alpha_*$  for  $k \in I_3$  by Lemma 4.6. Thus,  $x^* = T_1(z_2)$  satisfies case (I).



Consider the case  $f_i(\mathbf{x}) = \alpha_*$  for all  $i \in I_1$ . If also  $f_j(\mathbf{x}) = \alpha_*$  for all  $j \in I_2$ , then either  $\mathbf{x}^* = \mathbf{x}$  satisfies case (V) or there exists a point  $\mathbf{x}^* = \mathbf{z}_1$  obtained by Lemma 4.6 satisfying case (III). Therefore, assume  $f_j(\mathbf{x}) < \alpha_*$  for some  $j \in I_2$ . By Lemma 4.5, there exists a point  $\mathbf{z}_1 \in \Delta^7$  with the property that  $f_i(\mathbf{z}_1) = \alpha_*$  for all  $i \in I_1$ ,  $f_j(\mathbf{z}_1) < \alpha_*$  for all  $j \in I_2$  and  $f_k(\mathbf{z}_1) = f_k(\mathbf{x})$  for all  $k \in I_3$ . Then either  $\mathbf{x}^* = \mathbf{z}_1$  satisfies case (IV), or there exists a point  $\mathbf{x}^* = \mathbf{z}_2 \in \Delta^7$  satisfying case (I) by Lemma 4.6.  $\square$

### 4.2 Calculations of the infima

Let  $\Delta_7 = \{\mathbf{y} \in \Delta^7 : f_j(\mathbf{y}) = f_k(\mathbf{y}), j, k \in I^\dagger\} \subset \Delta^7$ . Note that  $\bar{\mathbf{x}} \in \Delta_7$  (see Lemma 4.2). We aim to prove the following proposition:

**Proposition 4.8** *The infimum  $\alpha_* = \min_{\mathbf{x} \in \Delta^7} F^\dagger(\mathbf{x}) \in [9, 5 + 3\sqrt{2}]$  is attained at some point  $\mathbf{x}^* \in \Delta_7$ .*

To this purpose, we need to show that cases (I), (II), (III) and (IV) are not possible at a point  $\mathbf{x} \in \Delta^7$  so that  $F^\dagger(\mathbf{x}) = \alpha_*$ . We start with case (I).

**Lemma 4.9** *Let  $\alpha_* = \inf_{\mathbf{x} \in \Delta^7} \max\{f_1(\mathbf{x}), \dots, f_8(\mathbf{x})\}$ , where  $f_i$  for  $i \in I^\dagger$  are as in Proposition 3.5. At any  $\mathbf{x}^* \in \Delta^7$  satisfying  $f_i(\mathbf{x}^*) = f_j(\mathbf{x}^*)$  for every  $i, j \in \{1, 2, 3\}$  and  $f_2(\mathbf{x}^*) > f_6(\mathbf{x}^*)$ , there exists  $\vec{v} \in T_{\mathbf{x}^*}\Delta^7$  such that each of  $f_1, f_2$  and  $f_3$  decreases in the direction of  $\vec{v}$ .*

**Proof** Using the identity  $x_8 = 1 - \sum_{n=1}^7 x_n$ , we rewrite  $f_1, f_2$  and  $f_3$  as

$$f_1(\mathbf{x}) = \sigma(\Sigma_2(\mathbf{x}))\sigma(x_1), \quad f_2(\mathbf{x}) = \frac{\sigma(x_2)}{\sigma(\Sigma_1(\mathbf{x}))}, \quad f_3(\mathbf{x}) = \frac{\sigma(x_3)}{\sigma(\Sigma_2(\mathbf{x}))},$$

where  $\Sigma_1(\mathbf{x}) = x_1 + x_2 + x_3$  and  $\Sigma_2(\mathbf{x}) = x_4 + x_5 + x_6$  for every  $\mathbf{x} \in \Delta^7$ . These functions are each well-defined and smooth on an open neighborhood of  $\Delta^7$  in  $\mathbb{R}^8$ . Because  $\sigma(x)$  decreases in  $x$  and  $\Sigma_2$  is constant in all variables but  $x_4, x_5$  and  $x_6$ , it is clear that

$$\begin{aligned} (1) \quad \frac{\partial f_1}{\partial x_1} \Big|_{\mathbf{x}^*} &= -\frac{\sigma(\Sigma_2(\mathbf{x}^*))}{(x_1^*)^2} < 0, & \frac{\partial f_1}{\partial x_2} \Big|_{\mathbf{x}^*} &= 0, & \frac{\partial f_1}{\partial x_3} \Big|_{\mathbf{x}^*} &= 0, \\ (2) \quad \frac{\partial f_3}{\partial x_3} \Big|_{\mathbf{x}^*} &= -\frac{1}{(x_3^*)^2 \sigma(\Sigma_2(\mathbf{x}^*))} < 0, & \frac{\partial f_3}{\partial x_1} \Big|_{\mathbf{x}^*} &= 0, & \frac{\partial f_3}{\partial x_2} \Big|_{\mathbf{x}^*} &= 0, \\ (3) \quad \frac{\partial f_1}{\partial x_k} \Big|_{\mathbf{x}^*} &= 0, & \frac{\partial f_2}{\partial x_k} \Big|_{\mathbf{x}^*} &= 0, & \frac{\partial f_3}{\partial x_k} \Big|_{\mathbf{x}^*} &= 0, \end{aligned}$$

for every  $k \in I_3$ . These facts imply that at any  $\mathbf{x}^* \in \Delta^7$  such that the equation

$$(14) \quad \left. \frac{\partial f_2}{\partial x_2} \right|_{\mathbf{x}^*} = 0$$

does not hold, there exists some  $r \in \mathbb{R}$  such that each of  $f_1, f_2$  and  $f_3$  decreases in the direction of the vector  $\vec{v}_r = \langle 1, r, 1, 0, 0, 0, -r - 2, 0 \rangle$ . Note that  $\vec{v}_r \in T_{\mathbf{x}^*} \Delta^7$  since its coordinates sum to 0. Thus it only remains to consider the case in which the equality in (14) holds.

A computation gives that

$$\left. \frac{\partial f_2}{\partial x_2} \right|_{\mathbf{x}^*} = \frac{(x_2^* - \Sigma_1(\mathbf{x}^*))(1 - x_2^* - \Sigma_1(\mathbf{x}^*))}{(x_2^*)^2 (\Sigma_1(\mathbf{x}^*))^2},$$

which vanishes if and only if  $\Sigma_2(\mathbf{x}^*) + x_2^* = 1$ . Since  $\Sigma_1(\mathbf{x}^*) = x_1^* + x_2^* + x_3^*$ , we conclude that equation in (14) holds if and only if  $x_2^* = (1 - x_1^* - x_3^*)/2$ . By the identity  $\sum_{i=1}^8 x_i^* = 1$ , this is in turn equivalent to

$$x_2^* = x_4^* + x_5^* + x_6^* + x_7^* + x_8^*.$$

Therefore, we find that  $x_6^* < x_2^*$ . Then the lemma follows, because by the definitions of  $f_2$  and  $f_6$ , we obtain  $f_2(\mathbf{x}^*) = (\sigma(x_2^*))^2 < \sigma(x_2^*)\sigma(x_6^*) = f_6(\mathbf{x}^*)$ ; a contradiction. □

Before we proceed to cases (II), (III) and (IV), we shall first prove the following statement:

**Lemma 4.10** *For  $1 \leq k \leq n - 1$ , let  $f_1, f_2, \dots, f_k$  be smooth functions on an open neighborhood  $U$  of the  $(n - 1)$ -simplex  $\Delta^{n-1}$  in  $\mathbb{R}^n$ . If at some  $\mathbf{x} \in \Delta^{n-1}$  the collection  $\{\nabla f_1(\mathbf{x}), \nabla f_2(\mathbf{x}), \dots, \nabla f_k(\mathbf{x}), (1, 1, \dots, 1)\}$  of vectors in  $\mathbb{R}^n$  is linearly independent, then there exists a vector  $\vec{u} \in T_{\mathbf{x}} \Delta^{n-1}$  such that each  $f_i$  for  $i = 1, \dots, k$  decreases in the direction of  $\vec{u}$  at  $\mathbf{x}$ .*

**Proof** Let  $\mathcal{B} = \{\vec{v}_1, \dots, \vec{v}_n\}$  be a collection of  $n$  linearly independent vectors in  $\mathbb{R}^n$ . We claim that there exists a vector  $\vec{u} \in \mathbb{R}^n$  such that  $\vec{u} \cdot \vec{v} < 0$  for every  $\vec{v} \in \mathcal{B}$ . The assertion is clear for  $n = 1$ . For  $n > 1$ , assume that there exists a vector  $\vec{u}_0 \in \text{Span}\{\vec{v}_1, \dots, \vec{v}_{n-1}\}$  such that  $\vec{u}_0 \cdot \vec{v} < 0$  for every  $v \in \mathcal{S} = \{\vec{v}_1, \dots, \vec{v}_{n-1}\}$  by induction.

There is a nonzero vector  $\vec{v}_0 \in \mathbb{R}^n$  orthogonal to each vector in  $\mathcal{S}$ . If we have  $\vec{v}_0 \cdot \vec{v}_n = 0$ , then  $\vec{v}_n$  is in the space  $\vec{v}_0^\perp$  of vectors perpendicular to  $\vec{v}_0$ . Since  $\dim \vec{v}_0^\perp = n - 1$ , the set  $\mathcal{S}$  spans  $\vec{v}_0^\perp$ . The set  $\mathcal{B}$  is linearly independent therefore, we get  $\vec{v}_n \cdot \vec{v}_0 \neq 0$ . Let  $\vec{u} = \vec{u}_0 - c\vec{v}_0$  for  $c = (\vec{u}_0 \cdot \vec{v}_n + 1)/\vec{v}_0 \cdot \vec{v}_n$ . Then we see that  $\vec{u} \cdot \vec{v} < 0$  for every  $\vec{v} \in \mathcal{B}$ , which proves the claim.

Let  $\vec{w} = (1, 1, \dots, 1) \in \mathbb{R}^n$ . Complete the set  $\{\nabla f_1(\mathbf{x}), \nabla f_2(\mathbf{x}), \dots, \nabla f_k(\mathbf{x}), \vec{w}\}$  to a basis  $\mathcal{B} = \{\nabla f_1(\mathbf{x}), \nabla f_2(\mathbf{x}), \dots, \nabla f_k(\mathbf{x}), \vec{u}_{k+1}, \dots, \vec{u}_{n-1}, \vec{w}\}$  for  $\mathbb{R}^n$ . If we declare  $\vec{v}_i = \text{proj}_{\vec{w}^\perp} \nabla f_i(\mathbf{x})$  for  $i = 1, \dots, k$  and  $\vec{v}_j = \text{proj}_{\vec{w}^\perp} \vec{u}_j$  for  $j = k+1, \dots, n-1$ , then  $\{\vec{v}_1, \dots, \vec{v}_{n-1}, \vec{w}\}$  is linearly independent. This is because  $\mathcal{B}$  is linearly independent. Let  $\mathcal{S} = \{\vec{v}_1, \dots, \vec{v}_{n-1}\}$ . Since  $\vec{w}^\perp$  has dimension  $n-1$  and  $\mathcal{S}$  is linearly independent, we have  $\text{Span } \mathcal{S} = \vec{w}^\perp$ . By the fact above, there exists a vector  $\vec{u} \in \text{Span } \mathcal{S}$  so that  $\vec{u} \cdot \vec{v} < 0$  for every  $\vec{v} \in \mathcal{S}$ . In particular, we get  $\vec{u} \cdot \vec{v}_i = \vec{u} \cdot \nabla f_i(\mathbf{x}) < 0$  for  $i = 1, \dots, k$ . Since  $T_{\mathbf{x}} \Delta^{n-1}$  consists of vectors whose entries sum to 0, we have  $\vec{w}^\perp = T_{\mathbf{x}} \Delta^{n-1}$ , which completes the proof.  $\square$

Lemmas 4.11, 4.12 and 4.13 below show respectively that cases (II), (III) and (IV) are not possible at a point at which  $F^\dagger$  takes its minimum value:

**Lemma 4.11** *Let  $\alpha_* = \inf_{\mathbf{x} \in \Delta^7} \max\{f_1(\mathbf{x}), \dots, f_8(\mathbf{x})\}$ , where  $f_i$  for  $i \in I^\dagger$  are as in Proposition 3.5. At any  $\mathbf{x}^* \in \Delta^7$  such that  $f_7(\mathbf{x}^*) = f_8(\mathbf{x}^*)$  and  $f_4(\mathbf{x}^*) < f_8(\mathbf{x}^*)$ , there exists a vector  $\vec{v} \in T_{\mathbf{x}^*} \Delta^7$  so that  $f_7$  and  $f_8$  decrease in the direction of  $\vec{v}$ .*

**Proof** We aim to apply Lemma 4.10. Therefore, we need to show that the set  $\{\nabla f_7(\mathbf{x}^*), \nabla f_8(\mathbf{x}^*), \vec{w}\}$  is linearly independent, where  $\vec{w} = \langle 1, 1, 1, 1, 1, 1, 1, 1 \rangle$ . It is enough to show that the matrix

$$(15) \quad \begin{bmatrix} \nabla f_7 \\ \nabla f_8 \\ \vec{w} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & \frac{\partial f_7}{\partial x_4} & \frac{\partial f_7}{\partial x_5} & \frac{\partial f_7}{\partial x_6} & \frac{\partial f_7}{\partial x_7} & 0 \\ \frac{\partial f_8}{\partial x_1} & \frac{\partial f_8}{\partial x_2} & \frac{\partial f_8}{\partial x_3} & 0 & 0 & 0 & 0 & \frac{\partial f_8}{\partial x_8} \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

has full rank at any  $\mathbf{x}^* \in \Delta^7$  which satisfies the hypotheses of the lemma. We have

- (1)  $\frac{\partial f_7}{\partial x_4} \Big|_{\mathbf{x}^*} = -\frac{\sigma(x_7^*)}{(\Sigma_2(\mathbf{x}^*))^2}, \quad \frac{\partial f_7}{\partial x_4} \Big|_{\mathbf{x}^*} = \frac{\partial f_7}{\partial x_5} \Big|_{\mathbf{x}^*}, \quad \frac{\partial f_7}{\partial x_5} \Big|_{\mathbf{x}^*} = \frac{\partial f_7}{\partial x_6} \Big|_{\mathbf{x}^*},$
- (2)  $\frac{\partial f_8}{\partial x_1} \Big|_{\mathbf{x}^*} = -\frac{\sigma(x_8^*)}{(\Sigma_1(\mathbf{x}^*))^2}, \quad \frac{\partial f_8}{\partial x_1} \Big|_{\mathbf{x}^*} = \frac{\partial f_8}{\partial x_2} \Big|_{\mathbf{x}^*}, \quad \frac{\partial f_8}{\partial x_2} \Big|_{\mathbf{x}^*} = \frac{\partial f_8}{\partial x_3} \Big|_{\mathbf{x}^*},$
- (3)  $\frac{\partial f_7}{\partial x_7} \Big|_{\mathbf{x}^*} = -\frac{\sigma(\Sigma_2(\mathbf{x}^*))}{(x_7^*)^2} \neq 0, \quad \frac{\partial f_8}{\partial x_8} \Big|_{\mathbf{x}^*} = -\frac{\sigma(\Sigma_1(\mathbf{x}^*))}{(x_8^*)^2} \neq 0.$

Let  $A = (f_7)_4(\mathbf{x}^*)$ ,  $B = (f_7)_7(\mathbf{x}^*)$ ,  $C = (f_8)_1(\mathbf{x}^*)$  and  $D = (f_8)_8(\mathbf{x}^*)$ . We apply two row operations: first  $R_2 \mapsto -D \cdot R_3 + R_2$ , then  $R_1 \mapsto (B/D) \cdot R_2 + R_1$ . Then we obtain the row equivalent matrix:

$$\begin{bmatrix} \frac{B}{D}(C-D) & \frac{B}{D}(C-D) & \frac{B}{D}(C-D) & A-B & A-B & A-B & 0 & 0 \\ C-D & C-D & C-D & -D & -D & -D & -D & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

The matrix above has full rank if and only if  $A - B \neq 0$  or  $C - D \neq 0$ , where

$$A - B = \frac{(\Sigma_2(\mathbf{x}^*) - x_7^*)(1 - \Sigma_2(\mathbf{x}^*) - x_7^*)}{(\Sigma_2(\mathbf{x}^*))^2(x_7^*)^2},$$

$$C - D = \frac{(\Sigma_1(\mathbf{x}^*) - x_8^*)(1 - \Sigma_1(\mathbf{x}^*) - x_8^*)}{(\Sigma_1(\mathbf{x}^*))^2(x_8^*)^2}.$$

Equivalently, it has full rank at  $\mathbf{x}^* \in \Delta^7$  unless  $x_7^* = \Sigma_2(\mathbf{x}^*)$  and  $x_8^* = \Sigma_1(\mathbf{x}^*)$ .

At any  $\mathbf{x}^* \in \Delta^7$  such that  $x_7^* = x_4^* + x_5^* + x_6^*$ ,  $x_8^* = x_1^* + x_2^* + x_3^*$  and  $f_7(\mathbf{x}^*) = f_8(\mathbf{x}^*)$ , we derive  $(\sigma(x_7^*))^2 = (\sigma(x_8^*))^2$ , which implies that  $x_7^* = x_8^*$ . Using  $\sum_{n=1}^8 x_n^* = 1$ , we find that

$$\Sigma_1(\mathbf{x}^*) = \frac{1}{4}, \quad \Sigma_2(\mathbf{x}^*) = \frac{1}{4}, \quad x_7^* = \frac{1}{4} \quad \text{and} \quad x_8^* = \frac{1}{4}.$$

Since we have  $\Sigma_2(\mathbf{x}^*) = \frac{1}{4}$ ,  $x_5^* > 0$  and  $x_6^* > 0$ , we get  $x_4^* < \frac{1}{4}$ , which implies that  $f_4(\mathbf{x}^*) = \sigma(\Sigma_1(\mathbf{x}^*))\sigma(x_4^*) > f_8(\mathbf{x}^*) = \sigma(\Sigma_1(\mathbf{x}^*))\sigma(x_8^*)$ , a contradiction. As a result, the matrix in (15) has full rank. By Lemma 4.10, the conclusion of the lemma follows. □

**Lemma 4.12** *Let  $\alpha_* = \inf_{\mathbf{x} \in \Delta^7} \max\{f_1(\mathbf{x}), \dots, f_8(\mathbf{x})\}$ , where  $f_i$  for  $i \in I^\dagger$  are as in Proposition 3.5. For any  $\mathbf{x}^* \in \Delta^7$  such that  $f_i(\mathbf{x}^*) = f_j(\mathbf{x}^*)$  for every  $i, j \in I_1 \cup I_2$ , there exists a vector  $\vec{v} \in T_{\mathbf{x}^*}\Delta^7$  such that  $f_i$  decreases in the direction of  $\vec{v}$  for each  $i \in I_1 \cup I_2 = \{1, 2, 3, 4, 5, 6\}$ .*

**Proof** Using the identity  $x_8 = 1 - \sum_{n=1}^7 x_n$  we rewrite the formulas of  $f_1, f_2, \dots, f_6$  as

$$f_1(\mathbf{x}) = \sigma(\Sigma_2(\mathbf{x}))\sigma(x_1), \quad f_2(\mathbf{x}) = \frac{\sigma(x_2)}{\sigma(\Sigma_1(\mathbf{x}))}, \quad f_3(\mathbf{x}) = \frac{\sigma(x_3)}{\sigma(\Sigma_2(\mathbf{x}))},$$

$$f_4(\mathbf{x}) = \sigma(\Sigma_1(\mathbf{x}))\sigma(x_4), \quad f_5(\mathbf{x}) = \frac{\sigma(x_5)}{\sigma(\Sigma_2(\mathbf{x}))}, \quad f_6(\mathbf{x}) = \frac{\sigma(x_6)}{\sigma(\Sigma_1(\mathbf{x}))}.$$

All of these functions are smooth on an open neighborhood of  $\Delta^7$  in  $\mathbb{R}^8$  and do not depend on  $x_7$  or  $x_8$ . Since  $f_2(\mathbf{x}^*) = f_6(\mathbf{x}^*)$  and  $f_3(\mathbf{x}^*) = f_5(\mathbf{x}^*)$ , we get  $x_2^* = x_6^*$  and  $x_3^* = x_5^*$ . Then using  $f_1(\mathbf{x}^*) = f_4(\mathbf{x}^*)$ , we obtain  $x_1^* = x_4^*$ . As a consequence, in particular, we find  $\Sigma_1(\mathbf{x}^*) = \Sigma_2(\mathbf{x}^*)$ . Using this fact together with  $f_2(\mathbf{x}^*) = f_3(\mathbf{x}^*)$  yields  $x_2^* = x_3^*$ . Let  $\Sigma_1^*$  denote  $\Sigma_1(\mathbf{x}^*)$ .

We aim to apply Lemma 4.10. To this purpose, we need to show that the matrix

$$(16) \quad \begin{bmatrix} \nabla f_1 \\ \nabla f_2 \\ \nabla f_3 \\ \nabla f_4 \\ \nabla f_5 \\ \nabla f_6 \\ \vec{w} \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & 0 & 0 & \frac{\partial f_1}{\partial x_4} & \frac{\partial f_1}{\partial x_5} & \frac{\partial f_1}{\partial x_6} & 0 & 0 \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_3} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\partial f_3}{\partial x_3} & \frac{\partial f_3}{\partial x_4} & \frac{\partial f_3}{\partial x_5} & \frac{\partial f_3}{\partial x_6} & 0 & 0 \\ \frac{\partial f_4}{\partial x_1} & \frac{\partial f_4}{\partial x_2} & \frac{\partial f_4}{\partial x_3} & \frac{\partial f_4}{\partial x_4} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{\partial f_5}{\partial x_4} & \frac{\partial f_5}{\partial x_5} & \frac{\partial f_5}{\partial x_6} & 0 & 0 \\ \frac{\partial f_6}{\partial x_1} & \frac{\partial f_6}{\partial x_2} & \frac{\partial f_6}{\partial x_3} & 0 & 0 & \frac{\partial f_6}{\partial x_6} & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

has full rank at  $\mathbf{x}^* \in \Delta^7$ . Using the equality  $f_1(\mathbf{x}^*) = f_2(\mathbf{x}^*)$ , the coordinates of  $\mathbf{x}^*$  and the definitions of  $f_1, f_2, \dots, f_6$ , we calculate that

- (1)  $\frac{\partial f_1}{\partial x_1} \Big|_{\mathbf{x}^*} = -\frac{\sigma(\Sigma_1^*)}{(x_1^*)^2}, \quad \frac{\partial f_2}{\partial x_2} \Big|_{\mathbf{x}^*} = \frac{x_2^*(1-x_2^*)-\Sigma_1^*(1-\Sigma_1^*)}{(1-\Sigma_1^*)^2(x_2^*)^2}, \quad \frac{\partial f_1}{\partial x_4} \Big|_{\mathbf{x}^*} = -\frac{\sigma(x_1^*)}{(\Sigma_1^*)^2},$
- (2)  $\frac{\partial f_3}{\partial x_3} \Big|_{\mathbf{x}^*} = -\frac{1}{\sigma(\Sigma_1^*)(x_2^*)^2}, \quad \frac{\partial f_2}{\partial x_1} \Big|_{\mathbf{x}^*} = -\frac{\partial f_1}{\partial x_4} \Big|_{\mathbf{x}^*}, \quad \frac{\partial f_4}{\partial x_4} \Big|_{\mathbf{x}^*} = \frac{\partial f_1}{\partial x_1} \Big|_{\mathbf{x}^*},$
- (3)  $\frac{\partial f_2}{\partial x_2} \Big|_{\mathbf{x}^*} = \frac{\partial f_5}{\partial x_5} \Big|_{\mathbf{x}^*}, \quad \frac{\partial f_3}{\partial x_3} \Big|_{\mathbf{x}^*} = \frac{\partial f_6}{\partial x_6} \Big|_{\mathbf{x}^*}, \quad \frac{\partial f_2}{\partial x_1} \Big|_{\mathbf{x}^*} = \frac{\partial f_2}{\partial x_3} \Big|_{\mathbf{x}^*}, \quad \frac{\partial f_5}{\partial x_4} \Big|_{\mathbf{x}^*} = \frac{\partial f_5}{\partial x_6} \Big|_{\mathbf{x}^*},$
- (4)  $\frac{\partial f_2}{\partial x_3} \Big|_{\mathbf{x}^*} = \frac{\partial f_3}{\partial x_j} \Big|_{\mathbf{x}^*}, \quad \frac{\partial f_5}{\partial x_6} \Big|_{\mathbf{x}^*} = \frac{\partial f_6}{\partial x_i} \Big|_{\mathbf{x}^*}, \quad \frac{\partial f_4}{\partial x_i} \Big|_{\mathbf{x}^*} = \frac{\partial f_1}{\partial x_j} \Big|_{\mathbf{x}^*}, \quad \frac{\partial f_5}{\partial x_j} \Big|_{\mathbf{x}^*} = \frac{\partial f_6}{\partial x_i} \Big|_{\mathbf{x}^*},$

for every  $i \in I_1$  and  $j \in I_2$ . Let  $A = \sigma(x_1^*), B = \sigma(x_2^*), C = \sigma(\Sigma_1^*), A' = \sigma'(x_1^*), B' = \sigma'(x_2^*)$  and  $C' = \sigma'(\Sigma_1^*)$ . Note that  $A' \neq 0, B' \neq 0$  and  $C \neq 0$ .

We perform simultaneously the following elementary row operations  $R_1 \mapsto R_1 + R_3, R_4 \mapsto R_4 + R_2, R_5 \mapsto (-1)R_3 + R_5, R_6 \mapsto (-1)R_2 + R_6, R_3 \mapsto (ACC')/B' \cdot R_6 + R_3, R_3 \mapsto (ACC')/B' \cdot R_5 + R_3, R_3 \mapsto (AC')/(A'C) \cdot R_4 + R_3, R_2 \mapsto (AC')/(A'C) \cdot R_1 + R_2$  in the matrix in (16) to obtain the matrix:

$$\left[ \begin{array}{c|cc|cccc} A'C & 0 & \frac{B'}{C} & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & \frac{B'}{C} - AC' & \frac{AC'B'}{A'C^2} - AC' & 0 & 0 & 0 & 0 & 0 \\ 0 & \frac{AC'B'}{A'C^2} - AC' & \frac{B'}{C} - AC' & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & \frac{B'}{C} & 0 & A'C & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{B'}{C} & 0 & \frac{B'}{C} & 0 & 0 & 0 \\ 0 & -\frac{B'}{C} & 0 & 0 & 0 & \frac{B'}{C} & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{array} \right]$$

If we mark each partition of the matrix above by  $(i, j)$  reading the first entry from top to bottom and the second entry from left to right for  $i = 1, 2, 3$  and  $j = 1, 2, 3$ , then the matrix in (16) has full rank if and only if the  $(2, 2)$  partition has full rank. Its determinant factors as

$$(17) \quad \frac{B'}{C} \left(1 - \frac{AC'}{A'C}\right) \left(\frac{B'}{C} \left(1 + \frac{AC'}{A'C}\right) - 2AC'\right),$$

where

$$\frac{B'}{C} \left(1 - \frac{AC'}{A'C}\right) = \frac{-\Sigma_1^*}{(1-\Sigma_1^*)(x_2^*)^2} \left(1 - \frac{x_1^*(1-x_1^*)}{\Sigma_1^*(1-\Sigma_1^*)}\right) \neq 0,$$

which follows from the fact that the function  $t \mapsto t(1-t)$  is increasing on  $(0, \frac{1}{2})$  and the inequality  $0 < x_1^* < \Sigma_1^* < \frac{1}{2}$ , an implication of the equality  $2\Sigma_1^* + x_7^* + x_8^* = 1$ .

Let us assume that the determinant in (17) vanishes. Then we must have the expression

$$\frac{-\Sigma_1^*}{(1-\Sigma_1^*)(x_2^*)^2} \left(1 + \frac{x_1^*(1-x_1^*)}{\Sigma_1^*(1-\Sigma_1^*)}\right) + \frac{2(1-x_1^*)}{x_1^*(\Sigma_1^*)^2} = 0.$$

Simplifying the left-hand summand, applying the identity  $f_1(x^*) = f_2(x^*)$  to the right, and finding a common denominator yield

$$\frac{-\Sigma_1^*(1-\Sigma_1^*) - x_1^*(1-x_1^*) + 2x_2^*(1-x_2^*)}{(x_2^*)^2(1-\Sigma_1^*)^2} = 0.$$

We use the fact that  $\Sigma_1^* = x_1^* + 2x_2^*$  in the equality above. Then it simplifies to  $(x_2^*)^2 + 2x_1^*x_2^* - x_1^*(1-x_1^*) = 0$ . The solutions to this quadratic are

$$x_2^* = -x_1^* \pm \sqrt{x_1^*}.$$

Since  $x_2^* > 0$ , we get

$$x_2^* = -x_1^* + \sqrt{x_1^*}.$$

Using this formula we find that  $x_2^* < x_1^*$  if and only if  $x_1^* > \frac{1}{4}$ .

Since  $\sigma(x) = 1/x - 1$  decreases in  $x$  and  $\Sigma_1^* < \frac{1}{2}$ , we get  $\sigma(\Sigma_1^*) > 1$ . The identity  $f_1(x^*) = f_2(x^*)$  implies that  $\sigma(x_2^*) = (\sigma(\Sigma_1^*))^2 \sigma(x_1^*) > \sigma(x_1^*)$ , which in turn gives that  $x_2^* < x_1^*$ . Then we derive that

$$\Sigma_1^* = x_1^* + 2x_2^* = 2\sqrt{x_1^*} - x_1^* > \frac{3}{4},$$

a contradiction. Hence, the matrix in (16) has full rank. By Lemma 4.10, there is a direction in which each  $f_i$  decreases for  $i \in I_1 \cup I_2$ . □

**Lemma 4.13** Let  $\alpha_* = \inf_{\mathbf{x} \in \Delta^7} \max\{f_1(\mathbf{x}), \dots, f_8(\mathbf{x})\}$ , where  $f_i$  for  $i \in I^\dagger$  are as in Proposition 3.5. At any  $\mathbf{x}^* \in \Delta^7$  satisfying  $F^\dagger(\mathbf{x}^*) = \alpha_*$  and  $f_i(\mathbf{x}^*) = f_j(\mathbf{x}^*) > f_k(\mathbf{x}^*)$  for every  $i, j \in I_1 \cup I_3 = \{1, 2, 3, 7, 8\}$  and  $k \in I_2 = \{4, 5, 6\}$ , there exists a vector  $\vec{v} \in T_{\mathbf{x}^*}\Delta^7$  such that  $f_i$  decreases in the direction of  $\vec{v}$  for each  $i \in I_1 \cup I_3$ .

**Proof** Define the function  $\Sigma_{1,3}(\mathbf{x}) = \Sigma_1(\mathbf{x}) + \Sigma_3(\mathbf{x})$  for  $\mathbf{x} \in \Delta^7$ . We use the identity  $x_4 = 1 - \sum_{n=1, n \neq 4}^8 x_n$  to rewrite  $f_i$  as follows:

$$\begin{aligned} f_1(\mathbf{x}) &= \frac{\sigma(x_1)}{\sigma(\Sigma_{1,3}(\mathbf{x}))}, & f_2(\mathbf{x}) &= \frac{\sigma(x_2)}{\sigma(\Sigma_1(\mathbf{x}))}, & f_3(\mathbf{x}) &= \sigma(\Sigma_{1,3}(\mathbf{x})) \cdot \sigma(x_3), \\ f_4(\mathbf{x}) &= \frac{\sigma(\Sigma_1)}{\sigma(x_5 + x_6 + \Sigma_{1,3}(\mathbf{x}))}, & f_5(\mathbf{x}) &= \sigma(\Sigma_{1,3}(\mathbf{x}))\sigma(x_5), \\ f_6(\mathbf{x}) &= \frac{\sigma(x_6)}{\sigma(\Sigma_1(\mathbf{x}))}, & f_7(\mathbf{x}) &= \frac{\sigma(x_7)}{\sigma(\Sigma_{1,3}(\mathbf{x}))}, & f_8(\mathbf{x}) &= \sigma(\Sigma_1(\mathbf{x}))\sigma(x_8), \end{aligned}$$

which are smooth on an open neighborhood of  $\Delta^7$  in  $\mathbb{R}^8$ . The functions  $f_1, f_2, f_3, f_7$  and  $f_8$  do not depend on  $x_4, x_5$  or  $x_6$ . We will show that the matrix

$$(18) \quad \begin{bmatrix} \nabla f_1 \\ \nabla f_2 \\ \nabla f_3 \\ \nabla f_7 \\ \nabla f_8 \\ \vec{w} \end{bmatrix} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \frac{\partial f_1}{\partial x_3} & 0 & 0 & 0 & \frac{\partial f_1}{\partial x_7} & \frac{\partial f_1}{\partial x_8} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \frac{\partial f_2}{\partial x_3} & 0 & 0 & 0 & 0 & 0 \\ \frac{\partial f_3}{\partial x_1} & \frac{\partial f_3}{\partial x_2} & \frac{\partial f_3}{\partial x_3} & 0 & 0 & 0 & \frac{\partial f_3}{\partial x_7} & \frac{\partial f_3}{\partial x_8} \\ \frac{\partial f_7}{\partial x_1} & \frac{\partial f_7}{\partial x_2} & \frac{\partial f_7}{\partial x_3} & 0 & 0 & 0 & \frac{\partial f_7}{\partial x_7} & \frac{\partial f_7}{\partial x_8} \\ \frac{\partial f_8}{\partial x_1} & \frac{\partial f_8}{\partial x_2} & \frac{\partial f_8}{\partial x_3} & 0 & 0 & 0 & 0 & \frac{\partial f_8}{\partial x_8} \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

has full rank at  $\mathbf{x}^* \in \Delta^7$ . Using the coordinates of  $\mathbf{x}^*$ , the definitions of the functions  $f_1, f_2, f_3, f_7, f_8$  and the equalities  $f_1(\mathbf{x}^*) = f_7(\mathbf{x}^*), f_1(\mathbf{x}^*) = f_3(\mathbf{x}^*), f_3(\mathbf{x}^*) = f_7(\mathbf{x}^*)$  and  $f_2(\mathbf{x}^*) = f_8(\mathbf{x}^*)$ , we find the following:

- (1)  $\frac{\partial f_1}{\partial x_1} \Big|_{\mathbf{x}^*} = \frac{x_1^*(1-x_1^*)-\Sigma_{1,3}(\mathbf{x}^*)(1-\Sigma_{1,3}(\mathbf{x}^*))}{(1-\Sigma_{1,3}(\mathbf{x}^*))^2(x_1^*)^2}, \quad \frac{\partial f_1}{\partial x_2} \Big|_{\mathbf{x}^*} = \frac{\sigma(x_1^*)}{(1-\Sigma_{1,3}(\mathbf{x}^*))^2},$
- (2)  $\frac{\partial f_2}{\partial x_1} \Big|_{\mathbf{x}^*} = \frac{\sigma(x_2^*)}{(1-\Sigma_1(\mathbf{x}^*))^2}, \quad \frac{\partial f_2}{\partial x_2} \Big|_{\mathbf{x}^*} = \frac{x_2^*(1-x_2^*)-\Sigma_1(\mathbf{x}^*)(1-\Sigma_1(\mathbf{x}^*))}{(1-\Sigma_1(\mathbf{x}^*))^2(x_2^*)^2},$
- (3)  $\frac{\partial f_3}{\partial x_1} \Big|_{\mathbf{x}^*} = -\frac{\sigma(x_3^*)}{(\Sigma_{1,3}(\mathbf{x}^*))^2}, \quad \frac{\partial f_3}{\partial x_3} \Big|_{\mathbf{x}^*} = \frac{-x_3^*(1-x_3^*)-\Sigma_{1,3}(\mathbf{x}^*)(1-\Sigma_{1,3}(\mathbf{x}^*))}{(\Sigma_{1,3}(\mathbf{x}^*))^2(x_3^*)^2},$
- (4)  $\frac{\partial f_8}{\partial x_8} \Big|_{\mathbf{x}^*} = -\frac{\sigma(\Sigma_1(\mathbf{x}^*))}{(x_8^*)^2}, \quad \frac{\partial f_8}{\partial x_1} \Big|_{\mathbf{x}^*} = -\frac{\partial f_2}{\partial x_1} \Big|_{\mathbf{x}^*}, \quad \frac{\partial f_3}{\partial x_1} \Big|_{\mathbf{x}^*} = -\frac{\partial f_1}{\partial x_3} \Big|_{\mathbf{x}^*},$
- (5)  $\frac{\partial f_3}{\partial x_7} \Big|_{\mathbf{x}^*} = -\frac{\partial f_7}{\partial x_3} \Big|_{\mathbf{x}^*}, \quad \frac{\partial f_2}{\partial x_1} \Big|_{\mathbf{x}^*} = \frac{\partial f_2}{\partial x_3} \Big|_{\mathbf{x}^*}, \quad \frac{\partial f_1}{\partial x_1} \Big|_{\mathbf{x}^*} = \frac{\partial f_7}{\partial x_7} \Big|_{\mathbf{x}^*}, \quad \frac{\partial f_8}{\partial x_1} \Big|_{\mathbf{x}^*} = \frac{\partial f_8}{\partial x_i} \Big|_{\mathbf{x}^*},$
- (6)  $\frac{\partial f_1}{\partial x_2} \Big|_{\mathbf{x}^*} = \frac{\partial f_1}{\partial x_j} \Big|_{\mathbf{x}^*}, \quad \frac{\partial f_3}{\partial x_1} \Big|_{\mathbf{x}^*} = \frac{\partial f_3}{\partial x_k} \Big|_{\mathbf{x}^*}, \quad \frac{\partial f_7}{\partial x_1} \Big|_{\mathbf{x}^*} = \frac{\partial f_7}{\partial x_l} \Big|_{\mathbf{x}^*},$

for every  $i = 2, 3$ ,  $j = 3, 7, 8$ ,  $k = 2, 7, 8$  and  $l = 2, 3, 8$ . Let  $A = (f_1)_1(\mathbf{x}^*)$ ,  $B = (f_1)_2(\mathbf{x}^*)$ ,  $C = (f_2)_1(\mathbf{x}^*)$ ,  $D = (f_2)_2(\mathbf{x}^*)$ ,  $E = (f_3)_3(\mathbf{x}^*)$  and  $F = (f_8)_8(\mathbf{x}^*)$ . Note that  $B \neq 0$ ,  $B - A \neq 0$  and  $D - C \neq 0$ . We also have  $E + B \neq 0$ , which follows from the equality  $f_1(\mathbf{x}^*) = f_3(\mathbf{x}^*)$ .

We simultaneously apply the row operations:  $R_1 \mapsto -A \cdot R_6 + R_1$ ,  $R_2 \mapsto R_5 + R_2$ ,  $R_3 \mapsto R_4 + R_3$ ,  $R_4 \mapsto -B \cdot R_6 + R_4$ ,  $R_5 \mapsto C \cdot R_6 + R_5$ ,  $R_1 \mapsto -R_4 + R_1$ ,  $R_1 \mapsto (1/(B - A))R_1$ ,  $R_2 \mapsto (C - D)R_1 + R_2$ ,  $R_2 \mapsto (1/(C - D))R_2$ ,  $R_3 \mapsto -(E + B)R_2 + R_3$ ,  $R_3 \mapsto -(1/(E + B))R_3$ ,  $R_5 \mapsto (C/B)R_4 + R_5$  and  $R_4 \mapsto B \cdot R_3 + R_4$  to the matrix in (18) to obtain the matrix:

$$(19) \quad \left[ \begin{array}{cccccc|ccc} 0 & 1 & 1 & 1 & 1 & 1 & 2 & & 1 \\ 0 & 0 & 1 & 1 & 1 & 1 & 2 & & \frac{F+C-D}{C-D} \\ 0 & 0 & 0 & 1 & 1 & 1 & \frac{B-A}{E+B} + 2 & & \frac{F+C-D}{C-D} \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & B\left(\frac{B-A}{E+B} + 2\right) + A - B & & B\left(\frac{F+C-D}{C-D}\right) \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{CA}{B} & & F + C \\ \hline 1 & 1 & 1 & 1 & 1 & 1 & 1 & & 1 \end{array} \right]$$

The matrix above has full rank if and only if the (2, 2) partition has nonzero determinant, where

$$B\left(\frac{B-A}{E+B} + 2\right) + A - B = \frac{2x_1^*(1-x_1^*) - x_3^*(1-x_3^*) - \Sigma_{1,3}(\mathbf{x}^*)(1 - \Sigma_{1,3}(\mathbf{x}^*))}{(1 - \Sigma_{1,3}(\mathbf{x}^*))^2(x_3^*)^2},$$

$$F + C = \frac{x_8^*(1-x_8^*) - \Sigma_1(\mathbf{x}^*)(1 - \Sigma_1(\mathbf{x}^*))}{(\Sigma_1(\mathbf{x}^*))^2(x_8^*)^2},$$

$$C - D = \frac{\Sigma_1(\mathbf{x}^*)}{(x_2^*)^2(1 - \Sigma_1(\mathbf{x}^*))^2},$$

calculated by using the facts  $f_1(\mathbf{x}^*) = f_3(\mathbf{x}^*)$  and  $f_2(\mathbf{x}^*) = f_8(\mathbf{x}^*)$ . Assume that the determinant of the (2, 2) partition of the matrix in (19) vanishes. Then the equality

$$(20) \quad \left(B \cdot \frac{B-A}{E+B} + A + B\right)(F + C) = CA\left(\frac{F+C-D}{C-D}\right)$$

must hold at any point  $\mathbf{x}^* \in \Delta^7$  satisfying the hypotheses of the lemma. Let  $\Sigma_1^*$ ,  $\Sigma_2^*$  and  $\Sigma_{1,3}^*$  denote  $\Sigma_1(\mathbf{x}^*)$ ,  $\Sigma_2(\mathbf{x}^*)$  and  $\Sigma_{1,3}(\mathbf{x}^*)$ , respectively.

Since  $f_1(\mathbf{x}^*) = f_7(\mathbf{x}^*)$ , we get  $x_1^* = x_7^*$ . Using the inequalities  $f_2(\mathbf{x}^*) > f_6(\mathbf{x}^*)$ ,  $f_3(\mathbf{x}^*) > f_5(\mathbf{x}^*)$  and  $f_8(\mathbf{x}^*) > f_4(\mathbf{x}^*)$ , we derive that  $x_2^* < x_6^*$ ,  $x_3^* < x_5^*$  and  $x_8^* < x_4^*$ , which implies  $\Sigma_1(\mathbf{x}^*) < \frac{1}{2}$ . Because otherwise we find  $x_5^* + x_6^* + x_1^* > \frac{1}{2}$  contradicting with the fact that  $\mathbf{x}^* \in \Delta^7$ .



Note that  $2x_8^* + x_2^* + x_3^* < x_2^* + x_3^* + x_4^* + x_8^* < 1$  and  $x_1^* + x_7^* + x_2^* + x_3^* = 2x_1^* + x_2^* + x_3^* < 1$ . So we have  $x_1^*, x_8^* \in (0, (1 - x_2^* - x_3^*)/2)$ . From the inequalities  $x_2^* < x_6^*$ ,  $x_3^* < x_5^*$  and  $x_8^* < x_4^*$ , we obtain

$$f_1(\mathbf{x}^*) = \frac{1 - x_4^* - x_5^* - x_6^*}{x_4^* + x_5^* + x_6^*} \cdot \frac{1 - x_1^*}{x_1^*} < \frac{1 - x_8^* - x_2^* - x_3^*}{x_8^* + x_2^* + x_3^*} \cdot \frac{1 - x_1^*}{x_1^*}.$$

By the equality  $f_1(\mathbf{x}^*) = f_8(\mathbf{x}^*)$ , we get that  $\sigma(x_8^* + x_2^* + x_3^*)\sigma(x_1^*) > \sigma(x_1^* + x_2^* + x_3^*)\sigma(x_8^*)$ . Since the function  $\sigma(x)/\sigma(x + x_2^* + x_3^*)$  is decreasing over the interval  $(0, (1 - x_2^* - x_3^*)/2)$ , we find  $x_1^* < x_8^*$ . By the facts  $\Sigma_1(\mathbf{x}^*) < \frac{1}{2}$  and  $(\sigma(\Sigma_1(\mathbf{x}^*)))^2\sigma(x_8^*) = \sigma(x_2^*)$ , which follows from the rearranging of the equality  $f_2(\mathbf{x}^*) = f_8(\mathbf{x}^*)$ , we also find that  $x_2^* < x_8^*$ .

By using the equality  $f_2(\mathbf{x}^*) = f_8(\mathbf{x}^*)$ , we simplify the right-hand side of the equality above to  $(x_1^* - \Sigma_{1,3}^*)(1 - x_1^* - \Sigma_{1,3}^*)(x_8^* - x_2^*)(1 - x_8^* - x_2^*)$ , which is nonzero because  $x_2^* < x_8^*$  and  $1 - x_1^* - \Sigma_{1,3}^* > 0$  by the inequality  $f_1(\mathbf{x}^*) > f_4(\mathbf{x}^*)$ .

Similarly, by using the equality  $f_1(\mathbf{x}^*) = f_3(\mathbf{x}^*)$ , we reduce the left-hand side of (20) to  $(x_8^* - \Sigma_1^*)(1 - x_8^* - \Sigma_1^*)((x_1^* - x_3^*)(1 - x_1^* - x_3^*) + (x_1^* - \Sigma_{1,3}^*)(1 - x_1^* - \Sigma_{1,3}^*))$ . We first distribute the factor  $(x_8^* - \Sigma_1^*)(1 - x_8^* - \Sigma_1^*)$  and move the second summand in the resulting expression to the right-hand side of the equation in (20). On the right-hand side of (20) the term  $(x_1^* - \Sigma_{1,3}^*)(1 - x_1^* - \Sigma_{1,3}^*)$  is a common factor. We factor this term and, after simplifications, we obtain

$$(21) \quad -(x_8^* + \Sigma_1^*)(1 - x_1^* - \Sigma_{1,3}^*)(x_1^* + x_3^*)(1 - x_2^* - \Sigma_1^*) < 0.$$

On the left-hand side of (20), we have

$$(22) \quad (x_8^* - \Sigma_1^*)(1 - x_8^* - \Sigma_1^*)(x_1^* - x_3^*)(1 - x_1^* - x_3^*).$$

Since we assume that the expressions in (21) and (22) are equal, there are two cases to consider:

$$\begin{aligned} \Sigma_1^* < x_8^* \quad \text{and} \quad x_1^* < x_3^*, \\ \Sigma_1^* > x_8^* \quad \text{and} \quad x_1^* > x_3^*. \end{aligned}$$

Assume we are in the first case. Note that  $\Sigma_2^* > \frac{1}{2}$  because  $x_1^* < x_3^*$  and  $f_1(\mathbf{x}^*) = f_3(\mathbf{x}^*)$ . We claim that  $\Sigma_1^* < \frac{1}{4} < x_8^*$ . If  $\frac{1}{4} \leq \Sigma_1^* < x_8^*$  holds, then  $\sum_{n=1}^8 x_n^* > 1$ , a contradiction.

If  $\Sigma_1^* < x_8^* \leq \frac{1}{4}$ , we see that  $\sigma(x_8^*) \geq 3$ . We get  $3\sigma(\Sigma_1^*) \leq f_8(\mathbf{x}^*) = \alpha_*$ , which implies  $3/(\alpha_* + 3) \leq \Sigma_1^*$ . By Lemma 4.2, we know that  $9 \leq \alpha_* \leq 5 + 3\sqrt{2}$ . As a consequence, we derive  $(24 - 9\sqrt{2})/46 \leq \Sigma_1^* < x_8^*$ . Since we have  $\Sigma_2^* > \frac{1}{2}$ , we find that  $(71 - 18\sqrt{2})/46 \leq \Sigma_1^* + \Sigma_2^* + x_8^*$ , which in turn gives  $x_7^* \leq (-25 + 18\sqrt{2})/46$ .

For a lower bound for  $\Sigma_2^*$ , we solve the inequality  $\sigma(\Sigma_2^*)(49 + 36\sqrt{2}) < f_7(\mathbf{x}^*) = \alpha_*$ . Using this lower bound, we conclude that  $\Sigma_1^* + \Sigma_2^* = (873 - 221\sqrt{2})/483 > 1$ , a contradiction. The claim follows.

Upon setting  $f_8(\mathbf{x}^*) = \alpha_*$ , substituting  $1 - \sum_{n=1}^7 x_n^*$  for  $x_8^*$  and collecting powers of  $\Sigma_1^*$  in the resulting formula, we see that  $\Sigma_1^*$  is a root of the quadratic

$$q(x) = (\alpha_* - 1)x^2 + (1 - \alpha_*)(1 - \Sigma_2^* - x_7^*)x + (\Sigma_2^* + x_7^*).$$

By the formula for the addition of the roots, we find that  $x_8^*$  is the other root. Since  $\Sigma_1^* \neq x_8^*$ , the discriminant of the quadratic above is strictly positive. By the inequality  $\Sigma_1^* < \frac{1}{4} < x_8^*$ , we find  $q(\frac{1}{4}) < 0$ , which implies that  $\Sigma_2^* + x_7^* < (21 + 18\sqrt{2})/92$ .

By substituting  $1 - \sum_{n=1, n \neq 7}^8 x_n^*$  for  $x_7^*$  in the equality  $f_7(\mathbf{x}^*) = \alpha_*$  and collecting powers of  $\Sigma_2^*$ , we find that  $\Sigma_2^*$  is a root of the quadratic

$$Q(x) = (\alpha_* - 1)x^2 + (1 - \alpha_*)(1 - \Sigma_1^* - x_8^*)x + (\Sigma_1^* + x_8^*).$$

By the formula for the addition of the roots, we derive that  $x_7^*$  is the other root. We obtain  $Q(\frac{1}{2}) < 0$  because we have  $x_1^* = x_7^* < \frac{1}{2} < \Sigma_2^*$ . Then, we calculate that

$$\Sigma_1^* + x_8^* < (1 + \sqrt{2})/6$$

which, in turn, implies that  $\sum_{n=1}^8 x_n^* < 1$ ; a contradiction. Hence the matrix in (19) has full rank in the case (1).

If we are in the second case, consider

$$\begin{aligned} q_\alpha(x) &= (\alpha - 1)x^2 + (1 - \alpha)(1 - \Sigma_2^* - x_7^*)x + (\Sigma_2^* + x_7^*), \\ Q_\alpha(x) &= (\alpha - 1)x^2 + (1 - \alpha)(1 - \Sigma_1^* - x_8^*)x + (\Sigma_1^* + x_8^*) \end{aligned}$$

for  $\alpha \in [9, 5 + 3\sqrt{2}]$ . Note that  $q_{\alpha_*}(x) = q(x)$  and  $Q_{\alpha_*}(x) = Q(x)$ . The solutions of  $q_\alpha(x) + Q_\alpha(x) = 0$  are

$$x_+(\alpha) = \frac{1}{4} + \frac{1}{4} \sqrt{\frac{\alpha - 9}{\alpha - 1}} \quad \text{or} \quad x_-(\alpha) = \frac{1}{4} - \frac{1}{4} \sqrt{\frac{\alpha - 9}{\alpha - 1}}.$$

Using  $x_8^* > x_4^*$ ,  $x_2^* > x_6^*$ ,  $x_3^* > x_5^*$  and  $x_1^* < x_7^*$ , we derive that  $x_7^* < \Sigma_1^* < \Sigma_2^*$ . Then we obtain  $q_{\alpha_*}(\Sigma_1^*) + Q_{\alpha_*}(\Sigma_1^*) = Q_{\alpha_*}(\Sigma_1^*) < 0$ , and  $q_{\alpha_*}(x_8^*) + Q_{\alpha_*}(x_8^*) = Q_{\alpha_*}(x_8^*) < 0$ , which implies that  $x_-(\alpha_*) < x_8^* < \Sigma_1^* < x_+(\alpha_*)$ . We shall use the previous inequality to produce lower and upper bounds for each of the factors in (21) and (22).

Since  $x_-(\alpha)$  is decreasing over  $[9, 5 + 3\sqrt{2}]$ , we have  $a_1 = (\sqrt{2} - 1)/2 < x_8^*$ . By the assumption  $x_8^* < \Sigma_1^*$  and the fact  $9 \leq \alpha_*$ , we find  $x_8^* < b_1 = \frac{1}{4}$ . Otherwise,

we would compute that  $f_8(\mathbf{x}^*) = \alpha_* < 9$ , a contradiction. Using the inequality  $\alpha_* = f_8(\mathbf{x}^*) > \sigma(\Sigma_1^*)\sigma(b_1)$ , we also get  $a_2 < \Sigma_1^* < b_2$ , where  $a_2 = 3/(8 + 3\sqrt{2})$  and  $b_2 = x_+(5 + 3\sqrt{2}) = (2 - \sqrt{2})/2$  as  $x_+(\alpha)$  is an increasing function on  $[9, 5 + 3\sqrt{2}]$ .

We find a lower and an upper bound for  $\Sigma_2^*$  as follows. From the assumption  $x_1^* > x_3^*$  and the equality  $f_1(\mathbf{x}^*) = f_3(\mathbf{x}^*)$ , we have  $\Sigma_2^* < b_3 = \frac{1}{2}$ . By the inequalities  $f_4(\mathbf{x}^*) < \alpha_*$ ,  $f_5(\mathbf{x}^*) < \alpha_*$  and  $f_6(\mathbf{x}^*) < \alpha_*$ , we obtain

$$x_4^* + x_6^* > \frac{1 - \Sigma_1^*}{(\alpha_* - 1)\Sigma_1^* + 1} + \frac{\Sigma_1^*}{\alpha_*(1 - \Sigma_1^*) + \Sigma_1^*}, \quad x_5^* > \frac{\Sigma_2^*}{\alpha_*(1 - \Sigma_2^*) + \Sigma_2^*}.$$

The expression on the right-hand side of the first inequality above is decreasing both in  $\alpha_*$  and  $\Sigma_1^* < \frac{1}{2}$ . So we find  $x_4^* + x_6^* > \frac{1}{4}$  by using the bounds  $5 + 3\sqrt{2}$  and  $b_2$  for  $\alpha_*$  and  $\Sigma_1^*$ , respectively. Then we have  $\Sigma_2^* > x_5^* + \frac{1}{4}$ .

Since the expression on the right-hand side of the second inequality above is decreasing in  $\alpha_*$  and increasing in  $\Sigma_2^*$ , by substituting the bounds  $5 + 3\sqrt{2}$  and  $x_5^* + \frac{1}{4}$  and rearranging, we get  $(4 + 3\sqrt{2})(x_5^*)^2 - (\frac{3}{4})(4 + 3\sqrt{2})x_5^* + \frac{1}{4} < 0$ . Thus  $x_5^*$  is greater than the smaller root  $(3 - 2\sqrt{2})/4$  of the left-hand side quadratic. Then it follows that  $(2 - \sqrt{2})/2 < \Sigma_2^*$ . Next we will consider the two cases

$$\frac{2 - \sqrt{2}}{2} < \Sigma_2^* < \frac{1}{3}, \quad \frac{1}{3} \leq \Sigma_2^* < \frac{1}{2}.$$

Assume we are in the second case. By rearranging the equalities  $f_1(\mathbf{x}^*) = \alpha_*$ ,  $f_2(\mathbf{x}^*) = \alpha_*$  and  $f_3(\mathbf{x}^*) = \alpha_*$ , we derive

$$(23) \quad x_1^* = \frac{1 - \Sigma_2^*}{(\alpha_* - 1)\Sigma_2^* + 1}, \quad x_2^* = \frac{\Sigma_1^*}{\alpha_*(1 - \Sigma_1^*) + \Sigma_1^*}, \quad x_3^* = \frac{\Sigma_2^*}{\alpha_*(1 - \Sigma_2^*) + \Sigma_2^*}.$$

The right-hand side of the expression for  $x_2^*$  is increasing in  $\Sigma_1^*$  and decreasing in  $\alpha_*$ . Therefore we find  $x_2^* < b_4 = (9\sqrt{2} - 10)/62$  by substituting the relevant bounds 9 and  $b_2$  for  $\alpha_*$  and  $\Sigma_1^*$ , respectively. We also find  $x_1^* \leq b_5 = \frac{2}{11}$  by plugging in 9 and  $\frac{1}{3} = a_3 \leq \Sigma_2^*$  because the expression on the right-hand side of the equality for  $x_1^*$  above is decreasing in both  $\Sigma_2^*$  and  $\alpha_*$ . Similarly, since the right-hand side of the equality

$$(24) \quad x_1^* + x_3^* = \frac{1 - \Sigma_2^*}{(\alpha_* - 1)\Sigma_2^* + 1} + \frac{\Sigma_2^*}{\alpha_*(1 - \Sigma_2^*) + \Sigma_2^*}$$

is decreasing in both  $\alpha_*$  and  $\Sigma_2^*$ , we get  $a_4 = (2 - \sqrt{2})/3 < x_1^* + x_3^*$  by substituting  $5 + 3\sqrt{2}$  and  $b_3 = \frac{1}{2}$  for  $\alpha_*$  and  $\Sigma_2^*$ , respectively. The right-hand side of the expression

$$(25) \quad x_1^* - x_3^* = \frac{1 - \Sigma_2^*}{(\alpha_* - 1)\Sigma_2^* + 1} - \frac{\Sigma_2^*}{\alpha_*(1 - \Sigma_2^*) + \Sigma_2^*}$$

is again decreasing in both  $\alpha_*$  and  $\Sigma_2^*$ . So we obtain  $x_1^* - x_3^* < b_6 = \frac{27}{209}$  by plugging in 9 and  $a_3 = \frac{1}{3}$  for  $\alpha_*$  and  $\Sigma_2^*$ , respectively. As a result we have the bounds

$$a_1 > 0.20710, \quad a_2 > 0.24504, \quad a_3 \geq 0.33333, \quad a_4 > 0.19526, \quad b_1 = 0.25000, \\ b_2 < 0.29290, \quad b_3 = 0.50000, \quad b_4 < 0.04400, \quad b_5 < 0.18182, \quad b_6 < 0.12919.$$

Then we compute that  $(a_1 + a_2)a_4(a_3 - b_5)(1 - b_4 - b_2) > 0.00886$ , which implies that the expression in (21) is less than  $-0.00886$ . Similarly, we also calculate that  $(b_2 - a_1)b_6(1 - a_1 - a_2)(1 - a_4) < 0.00489$ , which shows that the expression in (22) is greater than  $-0.00489$ . This is a contradiction. Hence, the determinant of the (2, 2) partition of the matrix in (19) cannot be 0.

Assume that the inequality in the first case holds. In this case, we have  $\Sigma_2^* < b_3 = \frac{1}{3}$ . Using the equality in (24) we get  $(1027 - 480\sqrt{2})/1519 = a_4 < x_1^* + x_3^*$ . Since we have  $\Sigma_1^* = x_1^* + x_2^* + x_3^* > x_2^* + a_4$ , by rearranging the equality for  $x_2^*$  in (23), we derive the inequality

$$(4 + 3\sqrt{2})(x_2^*)^2 - (1 - a_4)(4 + 3\sqrt{2})x_2^* + a_4 < 0.$$

Thus  $x_2^*$  is greater than the smaller root of the left-hand quadratic in the previous inequality. This implies that

$$\frac{2424 + 1698\sqrt{2} - \sqrt{9776852 + 6468345\sqrt{2}}}{1519(4 + 3\sqrt{2})} + \frac{1027 - 480\sqrt{2}}{1519} = a_2 < \Sigma_1^*.$$

Substituting the bounds 9 for  $\alpha_*$  and  $a_3 = (2 - \sqrt{2})/2$  for  $\Sigma_2^*$  in the expression for  $x_1^*$  in (23) we obtain  $x_1^* < b_5 = 1/(9\sqrt{2} - 8)$ . Similarly, using the previous bounds for  $\alpha_*$  and  $\Sigma_2^*$  in (25), we get  $x_1^* - x_3^* < b_6 = (369 - 81\sqrt{2})/1519$ . As a result we have

$$a_1 > 0.20710, \quad a_2 > 0.26716, \quad a_3 > 0.29289, \quad a_4 > 0.22921, \quad b_1 = 0.25000, \\ b_2 < 0.29290, \quad b_3 = 0.33333, \quad b_4 < 0.04400, \quad b_5 < 0.21151, \quad b_6 < 0.16752.$$

Using these estimates we calculate  $(a_1 + a_2)a_4(a_3 - b_5)(1 - b_4 - b_2) > 0.00586$  and  $(b_2 - a_1)b_6(1 - a_1 - a_2)(1 - a_4) < 0.00583$ , a contradiction. Hence, the determinant of the (2,2) partition of the matrix in (19) cannot be 0 in this case as well. Finally by Lemma 4.10, we obtain the conclusion of the lemma.  $\square$

**Proof of Proposition 4.8** It follows from Lemmas 4.7, 4.9, 4.11, 4.12 and 4.13.  $\square$

We use Proposition 4.8 to calculate the infimum of  $G^\dagger$  over the simplex  $\Delta^7$ . In particular, we prove Theorem 4.1. First, we establish the following:

**Theorem 4.14** Let  $F^\dagger: \Delta^7 \rightarrow \mathbb{R}$  be the function defined by  $\mathbf{x} \mapsto \max\{f_i(\mathbf{x}) : i \in I^\dagger\}$ , where  $f_i$  are defined as in Proposition 3.5. Then  $\inf_{\mathbf{x} \in \Delta^7} F^\dagger(\mathbf{x}) = 5 + 3\sqrt{2}$ .

**Proof** By Proposition 4.8, we know that  $\mathbf{x}^* \in \Delta_7$ , ie  $f_i(\mathbf{x}^*) = \alpha_*$  for every  $i \in I^\dagger$ . Using the identities  $f_4(\mathbf{x}^*) = f_8(\mathbf{x}^*)$ ,  $f_1(\mathbf{x}^*) = f_7(\mathbf{x}^*)$ ,  $f_2(\mathbf{x}^*) = f_6(\mathbf{x}^*)$  and  $f_3(\mathbf{x}^*) = f_5(\mathbf{x}^*)$ , we get  $x_4^* = x_8^*$ ,  $x_1^* = x_7^*$ ,  $x_6^* = x_2^*$  and  $x_3^* = x_5^*$ . By the fact  $f_4(\mathbf{x}^*) = f_1(\mathbf{x}^*)$ , we obtain

$$(x_1^* - x_4^*)(x_2^* + x_3^*)(1 - x_1^* - x_2^* - x_3^* - x_4^*) = 0,$$

or  $x_1^* = x_4^*$ . The last equality, in turn, gives that  $\Sigma_1^* = \Sigma_2^*$ .

By the equality  $f_2(\mathbf{x}^*) = f_3(\mathbf{x}^*)$ , we see that  $x_2^* = x_3^*$ . Since  $\sum_{n=1}^8 x_n^* = 1$ , we obtain  $x_2^* = \frac{1}{4} - x_1^*$ . Using the equality  $f_1(\mathbf{x}^*) = f_2(\mathbf{x}^*)$ , we find

$$x_2^* - 3x_2^*x_1^* - (x_1^*)^2x_2^* - 4(x_2^*)^2 + 4(x_2^*)^2x_1^* + 4(x_2^*)^3 - (x_1^*)^3 = 0,$$

which simplifies to  $1 - 4x_1^* - 4(x_1^*)^2 = 0$ . The solutions are  $x_1^* = (-1 \pm \sqrt{2})/2$ . Since  $x_1^* > 0$ , we get  $x_2^* = (3 - 2\sqrt{2})/4$ . In particular, we conclude that  $x_i^* = x_1^*$  for every  $i \in \{4, 7, 8\}$  and  $x_j^* = x_2^*$  for every for  $x_j \in \{3, 5, 6\}$ . Finally, we calculate that

$$\inf_{\mathbf{x} \in \Delta^7} F^\dagger(\mathbf{x}) = \sigma((2 - \sqrt{2})/2)\sigma((\sqrt{2} - 1)/2) = 5 + 3\sqrt{2}. \quad \square$$

**Proof of Theorem 4.1** By the definitions of  $G^\dagger$  and  $F^\dagger$ , we have  $G^\dagger(\mathbf{x}) \geq F^\dagger(\mathbf{x})$  for every  $\mathbf{x} \in \Delta^7$ . A direct computation shows that  $G^\dagger(\mathbf{x}^*) = F^\dagger(\mathbf{x}^*)$ . Then the conclusion of the theorem follows.  $\square$

### 4.3 On the uniqueness of $\mathbf{x}^*$ in Theorem 4.14

It is worth emphasizing the similarities between the statements (a) and (b) listed in Section 2 and the two statements

- (c)  $\inf_{\mathbf{x} \in \Delta^7} F^\dagger(\mathbf{x}) = \min_{\mathbf{x} \in \Delta^7} F^\dagger(\mathbf{x})$ ,
- (d) there exists  $\mathbf{x}^* \in \Delta_7 \subset \Delta^7$  such that  $\min_{\mathbf{x} \in \Delta^7} F^\dagger(\mathbf{x}) = f_i(\mathbf{x}^*)$  for  $i \in I^\dagger$

used in the proof of Theorem 4.14 to calculate the number  $(\frac{1}{2}) \log(5 + 3\sqrt{2})$ . Although it is straightforward to observe (b), it takes considerable calculations to prove (d). Analogous to Lemma 2.1, Theorem 4.14 shows that the point  $\mathbf{x}^*$  is unique. Assuming the uniqueness of the point  $\mathbf{x}^*$  a priori together with (c) suggests an alternative way of finding the coordinates of the point  $\mathbf{x}^* \in \Delta^7$ .

Let  $T_1$  be the transformation defined in (13). Since  $f_i(T_1(\mathbf{x})) = f_j(\mathbf{x})$  for every  $\mathbf{x} \in \Delta^7$  for every pair  $(i, j) \in \{(1, 4), (4, 1), (3, 5), (5, 3), (7, 8), (8, 7), (2, 6), (6, 2)\}$ , we conclude that  $\{f_i : i \in I^\dagger\} = \{f_i \circ T_1 : i \in I^\dagger\}$ . Let us define  $H_1: \Delta^7 \rightarrow \mathbb{R}$ , where  $H_1(\mathbf{x}) = \max\{(f_i \circ T_1)(\mathbf{x}) : i \in I^\dagger\}$ . We see that  $F^\dagger(\mathbf{x}) = H_1(\mathbf{x})$  for every  $\mathbf{x} \in \Delta^7$  and  $\min_{\mathbf{x} \in \Delta^7} F^\dagger(\mathbf{x}) = \min_{\mathbf{x} \in \Delta^7} H_1(\mathbf{x})$ . Since  $F^\dagger$  takes its minimum value at the point  $\mathbf{x}^*$  and  $\{f_i(\mathbf{x}^*) : i \in I^\dagger\} = \{(f_i \circ T_1)(\mathbf{x}^*) : i \in I^\dagger\}$ , the function  $H_1$  takes its minimum value at the point  $T_1^{-1}(\mathbf{x}^*)$ . Then we obtain  $T_1^{-1}(\mathbf{x}^*) = \mathbf{x}^*$  by the uniqueness of  $\mathbf{x}^*$ . This means that  $x_1^* = x_4^*, x_2^* = x_5^*, x_3^* = x_6^*$  and  $x_7^* = x_8^*$ .

Let  $\Delta^3 = \{\mathbf{x} \in \Delta^7 : x_1 = x_4, x_2 = x_5, x_3 = x_6, x_7 = x_8\}$ . Note that  $f_1(\mathbf{x}) = f_4(\mathbf{x}), f_2(\mathbf{x}) = f_5(\mathbf{x}), f_3(\mathbf{x}) = f_6(\mathbf{x})$  and  $f_7(\mathbf{x}) = f_8(\mathbf{x})$  for every  $\mathbf{x} \in \Delta^3$ . Define the continuous function  $F_1: \Delta^3 \rightarrow \mathbb{R}$  such that  $\mathbf{x} \mapsto \max(g_1(\mathbf{x}), g_2(\mathbf{x}), g_3(\mathbf{x}), g_7(\mathbf{x}))$ , where  $g_i = f_i|_{\Delta^3}$  for  $i = 1, 2, 3, 7$ . Then we have  $\min_{\mathbf{x} \in \Delta^7} F^\dagger(\mathbf{x}) = \min_{\mathbf{x} \in \Delta^3} F_1(\mathbf{x})$ .

Consider  $T_2: \Delta^7 \rightarrow \Delta^7$  defined by  $x_2 \mapsto x_3, x_3 \mapsto x_2$  and  $x_i \mapsto x_i$  for every  $i \in I^\dagger - \{2, 3\}$ . The map  $T_2$  preserves  $\Delta^7$  and  $\Delta^3$ . Then we have  $g_i(T_2(\mathbf{x})) = f_j(\mathbf{x})$  for every  $\mathbf{x} \in \Delta^3$  for every pair  $(i, j) \in \{(1, 1), (2, 3), (3, 2), (7, 7)\}$ . An argument similar to the one above for  $H_2: \Delta^7 \rightarrow \mathbb{R}, H_2(\mathbf{x}) = \max\{(f_i \circ T_2)(\mathbf{x}) : i = 1, 2, 3, 7\}$  shows that  $T_2^{-1}(\mathbf{x}^*) = \mathbf{x}^*$ . This means that  $x_2^* = x_3^*$ .

Let  $\Delta^2 = \{\mathbf{x} \in \Delta^3 : x_2 = x_3\}$ . Note that  $g_2(\mathbf{x}) = g_3(\mathbf{x})$  for every  $\mathbf{x} \in \Delta^2$ . Define the functions  $h_i: \Delta^2 \rightarrow \mathbb{R}$  such that  $h_i = g_i|_{\Delta^2}$  for  $i = 1, 2, 7$ . Introduce the continuous function  $F_2: \Delta^2 \rightarrow \mathbb{R}$ , where  $\mathbf{x} \mapsto \max(h_1(\mathbf{x}), h_2(\mathbf{x}), h_7(\mathbf{x}))$ . Then we have  $\min_{\mathbf{x} \in \Delta^3} F_1(\mathbf{x}) = \min_{\mathbf{x} \in \Delta^2} F_2(\mathbf{x})$ . Note that  $\mathbf{x}^* = (x_1^*, x_2^*, x_2^*, x_1^*, x_2^*, x_2^*, x_7^*, x_7^*)$  with  $x_1^* + 2x_2^* + x_7^* = \frac{1}{2}$ .

In the rest of the discussion, we will consider  $\Delta^2$  as a submanifold of  $\mathbb{R}^8$ . Then the tangent space  $T_{\mathbf{x}} \Delta^2$  at any  $\mathbf{x} \in \Delta^2$  is a subspace of  $T_{\mathbf{x}} \Delta^7$  generated by the vectors  $\vec{u}_1 = \langle 1, 0, 0, 1, 0, 0, -1, -1 \rangle$  and  $\vec{u}_2 = \langle 0, 1, 1, 0, 1, 1, -2, -2 \rangle$ . Note that  $h_1(\mathbf{x}), h_2(\mathbf{x})$  and  $h_7(\mathbf{x})$  are smooth in an open neighborhood of  $\Delta^2$ . Therefore,  $\nabla h_i(\mathbf{x}) \cdot \vec{v}$  is the derivative of  $h_i$  in the direction of  $\vec{v} \in T_{\mathbf{x}} \Delta^2$  for each  $i \in \{1, 2, 7\}$ .

Using the identity  $x_7 = \frac{1}{2} - x_1 - 2x_2$ , we rewrite the formulas of  $h_1(\mathbf{x}), h_2(\mathbf{x})$  and  $h_7(\mathbf{x})$  as follows:  $h_1(\mathbf{x}) = \sigma(x_1 + 2x_2)\sigma(x_1), h_2(\mathbf{x}) = \sigma(x_2)/\sigma(x_1 + 2x_2)$  and  $h_7(\mathbf{x}) = \sigma(x_1 + 2x_2)\sigma(\frac{1}{2} - x_1 - 2x_2)$ . Then we find the partial derivatives

$$\begin{aligned} \frac{\partial h_1}{\partial x_1} &= \frac{2(x_1^2 + 2x_1x_2 - x_1 - x_2 + 2x_2^2)}{x_1^2(x_1 + 2x_2)^2}, & \frac{\partial h_1}{\partial x_2} &= -\frac{2(1 - x_1)}{x_1(x_1 + 2x_2)^2}, \\ \frac{\partial h_2}{\partial x_1} &= \frac{1 - x_2}{x_2(1 - x_1 - 2x_2)^2}, & \frac{\partial h_2}{\partial x_2} &= \frac{4x_1x_2 + 2x_2^2 - x_1 + x_1^2}{x_2^2(1 - x_1 - 2x_2)^2}. \end{aligned}$$

It is clear that  $h_1$  and  $h_2$  have no critical points in  $\Delta^2$ . On the other hand, every point on the set  $L = \{\mathbf{x} \in \Delta^2 : -1 + 4x_1 + 8x_2 = 0\}$  is a critical point for the function  $h_7$  because we have

$$\begin{aligned} \frac{\partial h_7}{\partial x_1} &= \frac{-1 + 4x_1 + 8x_2}{(x_1 + 2x_2)^2(-1 + 2x_1 + 4x_2)^2}, \\ \frac{\partial h_7}{\partial x_2} &= \frac{2(-1 + 4x_1 + 8x_2)}{(x_1 + 2x_2)^2(-1 + 2x_1 + 4x_2)^2}. \end{aligned}$$

Let  $\tilde{h}_i = h_i|_L$  for  $i = 1, 2, 7$ . Then we get

$$\tilde{h}_1(x_2) = 3 \cdot \frac{3 + 8x_2}{1 - 8x_2}, \quad \tilde{h}_2(x_2) = \frac{1}{3} \cdot \frac{1 - x_2}{x_2}, \quad \tilde{h}_7(x_2) = 9,$$

for  $0 < x_2 < \frac{1}{8}$ . Let  $I = (0, \frac{1}{8})$ . Then  $\tilde{h}_1$  and  $\tilde{h}_2$  have no critical points in  $I$ . We see  $\tilde{h}_1(x_2) > \tilde{h}_7(x_2)$  for every  $x_2 \in I$ , because we have

$$\lim_{x_2 \rightarrow 0^+} 3 \cdot \frac{3 + 8x_2}{1 - 8x_2} = 9^+ \quad \text{and} \quad \lim_{x_2 \rightarrow \frac{1}{8}^-} 3 \cdot \frac{3 + 8x_2}{1 - 8x_2} = \infty.$$

Therefore, it is enough to calculate the infimum of the maximum of  $\tilde{h}_1$  and  $\tilde{h}_2$  over  $I$  to calculate the infimum of the maximum of  $h_1, h_2$  and  $h_7$  over  $L$ .

Since  $\tilde{h}_1$  and  $\tilde{h}_2$  have no critical points in  $I$ , the infimum of the maximum of  $\tilde{h}_1$  and  $\tilde{h}_2$  is attained at a point  $x_2^*$  such that

$$\tilde{h}_1(x_2^*) = \tilde{h}_2(x_2^*).$$

In other words, we need to solve  $64(x_2^*)^2 + 36x_2^* - 1 = 0$ . We get  $x_2^* = (-9 \pm \sqrt{97})/32$ . Since  $x_2^*$  is positive,

$$\inf_{x_2 \in I} \{\max(\tilde{h}_1(x_2), \tilde{h}_2(x_2))\} = (17 + 2\sqrt{97})/3.$$

Note that the point  $x_2^*$  at which the infimum of the maximum of  $\tilde{h}_1$  and  $\tilde{h}_2$  over  $L$  is unique.

We claim that there exist  $i, j \in \{1, 2, 7\}$  with  $i \neq j$  such that  $h_i(\mathbf{x}^*) = h_j(\mathbf{x}^*)$ . Assume otherwise that  $h_i(\mathbf{x}^*) \neq h_j(\mathbf{x}^*)$  for every  $i, j \in \{1, 2, 7\}$  for  $i \neq j$ . Then we have either  $h_1(\mathbf{x}^*) > h_i(\mathbf{x}^*)$  for  $i = 2, 7$  or  $h_2(\mathbf{x}^*) > h_i(\mathbf{x}^*)$  for  $i = 1, 7$  or  $h_7(\mathbf{x}^*) > h_i(\mathbf{x}^*)$  for  $i = 1, 2$ . Since  $h_1$  and  $h_2$  have no critical points in  $\Delta^2$ , we cannot have  $h_1(\mathbf{x}^*) > h_i(\mathbf{x}^*)$  for  $i = 2, 7$  or  $h_2(\mathbf{x}^*) > h_i(\mathbf{x}^*)$  for  $i = 1, 7$ .

Assume  $h_7(\mathbf{x}^*) > h_i(\mathbf{x}^*)$  for  $i = 1, 2$ . Since  $h_7$  is continuous on  $\Delta^2$ , there exists a neighborhood  $V$  of  $\mathbf{x}^*$  in  $\Delta^2$  so that  $h_7(\mathbf{x}) > h_i(\mathbf{x})$  for every  $\mathbf{x} \in V$  for  $i = 1, 2$ .

Hence,  $F_2(x) = h_7(x)$  for every  $x \in V$ . Since  $h_7(x^*) = F_2(x^*)$ , the function  $h_7$  has a local minimum at  $x^* \in V$ . This means  $x^* \in L$  so that  $x_1^* = (13 - \sqrt{97})/16$ ,  $x_2^* = (-9 + \sqrt{97})/32$  and  $x_7^* = \frac{1}{4}$ . This is a contradiction, because we know by Lemma 4.2 that  $F^\dagger(\bar{x}) = 5 + 3\sqrt{2} < h_2(x^*)$ . Hence, there exist  $i, j \in \{1, 2, 7\}$  with  $i \neq j$  such that  $h_i(x^*) = h_j(x^*)$ . We need to consider the cases

- (I)  $h_1(x^*) = h_7(x^*) > h_2(x^*)$ ,
- (II)  $h_7(x^*) = h_2(x^*) > h_1(x^*)$ ,
- (III)  $h_1(x^*) = h_2(x^*) > h_7(x^*)$ ,
- (IV)  $h_1(x^*) = h_2(x^*) = h_7(x^*)$ .

Assume we are in case (I). Let  $\vec{v}_1 = \vec{u}_1 + (-\frac{1}{2})\vec{u}_2$ . The first-order partial derivatives of  $h_1$ ,  $h_2$  and  $h_7$  show that  $\nabla h_1(x) \cdot \vec{v}_1 < 0$ ,  $\nabla h_7(x) \cdot \vec{v}_1 = 0$ ,  $\nabla h_2(x) \cdot \vec{v}_1 > 0$  for every  $x \in \Delta^2$ . Therefore,  $h_1$  is decreasing,  $h_2$  is increasing and  $h_7$  is constant along a line segment in the direction of  $\vec{v}_1$ . This means that if we move along the line segment starting at  $x^*$  in the direction of the vector  $\vec{v}_1$  for a sufficiently small amount, we obtain a point  $y \in \Delta^2$  such that  $h_i(y) < \alpha_*$  for  $i = 1, 2$  and  $h_7(y) = \alpha_*$ . This is a contradiction. Therefore, we cannot have  $h_1(x^*) = h_7(x^*) > h_2(x^*)$ .

Assume we are in case (II). Let  $\vec{v}_2 = -\vec{v}_1$ . The first-order partial derivatives of  $h_1$ ,  $h_2$  and  $h_7$  show that  $\nabla h_1(x) \cdot \vec{v} > 0$ ,  $\nabla h_7(x) \cdot \vec{v} = 0$ ,  $\nabla h_2(x) \cdot \vec{v} < 0$  for every  $x \in \Delta^2$ . An argument similar to the above applies, mutatis mutandis, in this case to show that we cannot have  $h_2(x^*) = h_7(x^*) > h_1(x^*)$ . We already know that (III) is not possible at the point  $x^*$ , because this case corresponds to case (III) in Lemma 4.12. As a result, we have  $h_2(x^*) = h_7(x^*)$  and  $h_7(x^*) = h_1(x^*)$ .

Using the equality  $h_1(x^*) = h_2(x^*)$ , we see that

$$\sigma(x_1^*) = \sigma(\frac{1}{2} - x_1^* - 2x_2^*),$$

which implies  $x_1^* = \frac{1}{4} - x_2^*$ . We use  $h_2(x^*) = h_7(x^*)$  to obtain  $16(x_2^*)^2 - 24x_2^* + 1 = 0$  or  $x_2^* = (3 \pm 2\sqrt{2})/4$ . Since  $x_2^*$  is positive, we find

$$x_2^* = (3 - 2\sqrt{2})/4, \quad x_1^* = (\sqrt{2} - 1)/2 \quad \text{and} \quad x_7^* = x_7^*.$$

Finally, we calculate that  $\min_{x \in \Delta^2} F_2(x) = 5 + 3\sqrt{2}$ .

In the discussion above, we don't refer to the statement  $x^* \in \Delta_7$  given in (d). The assumption that the point  $x^*$  is unique reduces the necessary calculations to obtain  $\alpha_*$  considerably. Notice the fact that  $\Gamma_{\mathcal{D}^\dagger}$  is not a symmetric decomposition of  $\Gamma$ . The investigation of the likely conditions such as the convexity properties of the displacement functions for the decompositions  $\Gamma_{\mathcal{D}^1}$  and  $\Gamma_{\mathcal{D}^\dagger}$  that might lead to a proof of the uniqueness of the points  $x^*$  will be left to future studies.



### 5 Proof of the main theorem

In this final section, we present a detailed proof of the main result of this paper:

**Theorem 5.1** *Let  $\xi$  and  $\eta$  be noncommuting isometries of  $\mathbb{H}^3$ . Suppose that  $\xi$  and  $\eta$  generate a torsion-free discrete group which is not cocompact and contains no parabolic. Let  $\Gamma_{\dagger}$  and  $\alpha_{\dagger}$  denote the set of isometries  $\{\xi, \eta, \xi\eta\}$  and the real number  $5 + 3\sqrt{2}$ , respectively. Then for any  $z_0 \in \mathbb{H}^3$  we have*

$$e^{(2 \max_{\gamma \in \Gamma_{\dagger}} \{\text{dist}(z_0, \gamma \cdot z_0)\})} \geq \alpha_{\dagger}.$$

**Proof** By [6, Proposition 9.2], the group  $\langle \xi, \eta \rangle$  is a free group on the generators  $\xi$  and  $\eta$ . Let  $z_0$  be a point in  $\mathbb{H}^3$ . If  $\Gamma = \langle \xi, \eta \rangle$  is geometrically infinite, then Theorem 3.4 and Lemma 1.2 imply that

$$\max_{\gamma \in \Gamma_{\dagger}} \{\text{dist}(z_0, \gamma \cdot z_0)\} \geq \frac{1}{2} \log G^{\dagger}(\mathbf{m}) \geq \frac{1}{2} \log \left( \inf_{x \in \Delta^7} G^{\dagger}(x) \right)$$

for  $\mathbf{m} = (m_p(\psi))_{\psi \in \Psi^{\dagger}} \in \Delta^7$ , where  $p$  and  $m_p(\psi)$  are the bijection and the total measures defined in (11) and Section 3, respectively. The function  $G^{\dagger}$  is defined in Theorem 4.1, which implies the conclusion of the theorem.

Assume that  $\Gamma = \langle \xi, \eta \rangle$  is geometrically finite. Then  $(\xi, \eta)$  is in  $\mathfrak{GF}$ , an open subset of the character variety  $\mathfrak{X} = \text{Isom}^+(\mathbb{H}^3) \times \text{Isom}^+(\mathbb{H}^3)$ , consisting of  $(\xi, \eta)$  such that  $\langle \xi, \eta \rangle$  is free, geometrically finite and without any parabolic. Let  $f_{z_0}^{\dagger}: \mathfrak{X} \rightarrow \mathbb{R}$  be the function defined as

$$(\xi, \eta) \mapsto \max \{ \text{dist}(z_0, \xi \cdot z_0), \text{dist}(z_0, \eta \cdot z_0), \text{dist}(z_0, \xi\eta \cdot z_0) \}.$$

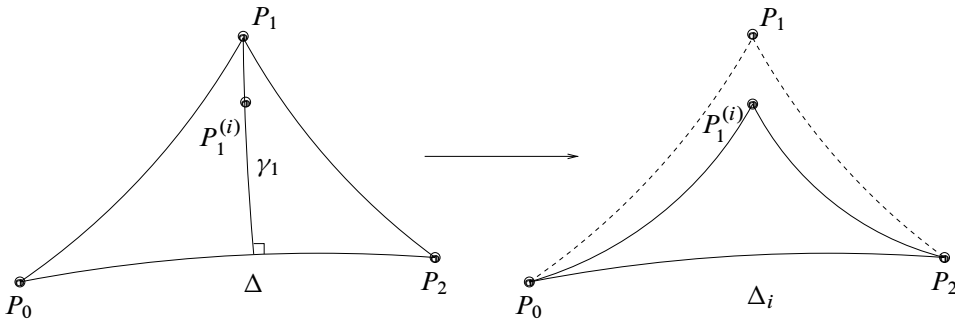
It is straightforward to see that  $f_{z_0}^{\dagger}$  is a proper, continuous and nonnegative-valued function on  $\mathfrak{X}$ . Therefore, it takes a minimum value at some point  $(\xi_0, \eta_0) \in \overline{\mathfrak{GF}}$ . We claim that  $(\xi_0, \eta_0)$  is in  $\overline{\mathfrak{GF}} - \mathfrak{GF}$ .

Assume on the contrary that  $(\xi_0, \eta_0)$  is in  $\mathfrak{GF}$ . Since  $\xi_0, \eta_0$  and  $\xi_0\eta_0$  have infinite orders in  $\langle \xi_0, \eta_0 \rangle$ , we have  $\xi_0 \cdot z \neq z, \eta_0 \cdot z \neq z$  and  $\xi_0\eta_0 \cdot z \neq z$  for every  $z \in \mathbb{H}^3$ . In particular, we get that  $\xi_0 \cdot z_0 \neq z_0, \eta_0 \cdot z_0 \neq z_0$  and  $\xi_0\eta_0 \cdot z_0 \neq z_0$ . Therefore, there exists hyperbolic geodesic segments joining  $z_0$  to  $\xi_0 \cdot z_0, z_0$  to  $\eta_0 \cdot z_0$  and  $z_0$  to  $\xi_0\eta_0 \cdot z_0$ . Note that we have the equalities

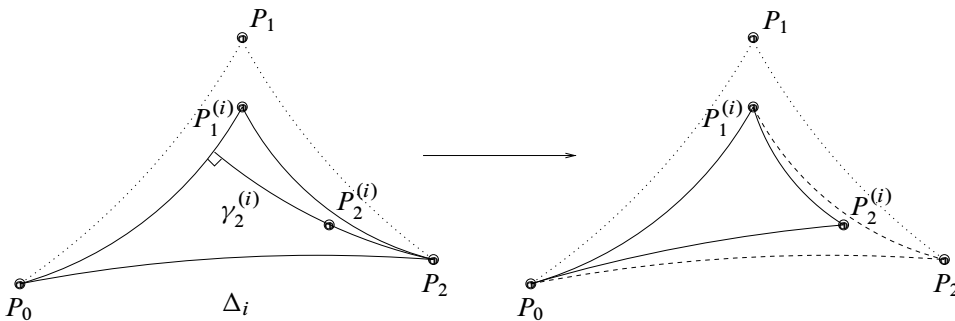
$$\text{dist}(z_0, \xi_0\eta_0 \cdot z_0) = \text{dist}(\xi_0^{-1} \cdot z_0, \eta_0 \cdot z_0) \quad \text{and} \quad \text{dist}(z_0, \xi \cdot z_0) = \text{dist}(z_0, \xi^{-1} \cdot z_0).$$

We consider the geodesic triangle  $\Delta = \Delta_{P_2 P_0 P_1}$ , where  $P_1 = \xi_0^{-1} \cdot z_0, P_0 = z_0$  and  $P_2 = \eta_0 \cdot z_0$ . The value  $f_{z_0}^{\dagger}(\xi_0, \eta_0)$  is the longest side length of  $\Delta$ . There are two cases to consider:  $\Delta$  is acute or  $\Delta$  is not acute.

Assume the latter is the case. Let  $\gamma$  be the unique longest edge of  $\Delta$ . By the hyperbolic law of sines,  $\gamma$  is opposite to the nonacute angle. If  $P_1$  lies in  $\gamma$ , let  $P_1^{(i)}$  be a sequence of points in the interior of  $\gamma$  so that  $P_1^{(i)} \rightarrow P_1$ . Let  $P_j^{(i)} = P_j$  for  $j \in \{0, 2\}$  and  $i \in \mathbb{N}$ . Otherwise, let  $P_2^{(i)}$  be a sequence of points in the interior of  $\gamma$  so that  $P_2^{(i)} \rightarrow P_2$  and define  $P_j^{(i)} = P_j$  for  $j \in \{0, 1\}$  and  $i \in \mathbb{N}$ . Let  $\Delta_i$  be the geodesic triangle contained in  $\Delta$  with vertices  $P_0^{(i)}$ ,  $P_1^{(i)}$  and  $P_2^{(i)}$ . By the construction, the unique longest side  $\gamma_i$  of  $\Delta_i$  is contained in  $\gamma$  for all but finitely many  $i$ . Let  $\{\xi_i\}$  be a sequence of isometries such that  $\xi_i \rightarrow \xi$  and  $\xi_i^{-1} \cdot z_0 = P_1^{(i)}$ . Similarly, let  $\{\eta_i\}$  be a sequence of isometries such that  $\eta_i \rightarrow \eta$  and  $\eta_i \cdot z_0 = P_2^{(i)}$ . Then  $(\xi_i, \eta_i) \in \mathcal{G}\mathcal{F}$  for all but finitely many  $i$  and  $f_{z_0}^\dagger(\xi_i, \eta_i) = l(\gamma_i) < f_{z_0}^\dagger(\xi_0, \eta_0)$ , a contradiction.



Assume  $\Delta$  is acute. Then the perpendicular arc  $\gamma_1$  from  $P_1$  to the geodesic containing  $P_0$  and  $P_2$  meets it in the interior of the edge of  $\Delta$  opposite to  $P_1$ . Let  $P_1^{(i)}$  be a sequence of points in the interior of  $\gamma_1$  so that  $P_1^{(i)} \rightarrow P_1$ . For each  $i$ , we see that  $d(P_1^{(i)}, P_0) < d(P_1, P_0)$  by applying the hyperbolic law of cosines to the right triangle containing  $P_1^{(i)}$ ,  $P_0$  and a subarc of  $\gamma_1$ . Similarly, we have  $d(P_1^{(i)}, P_2) < d(P_1, P_2)$ .



The triangle  $\Delta_i$  with vertices  $P_0$ ,  $P_1^{(i)}$  and  $P_2$  is itself acute because its angles at  $P_0$  and  $P_2$  are less than those of  $\Delta$ , and its angle at  $P_1^{(i)}$  limits to the angle of  $\Delta$  at  $P_1$ .

Thus, the perpendicular arc  $\gamma_2^{(i)}$  from  $P_2$  to the geodesic containing  $P_0$  and  $P_1^{(i)}$  meets this geodesic inside of  $\Delta_i$ . Let  $P_2^{(i)}$  be the point on  $\gamma_2^{(i)}$  at distance  $1/i$  from  $P_2$ . We find that

$$d(P_2^{(i)}, P_0) < d(P_2, P_0) \quad \text{and} \quad d(P_2^{(i)}, P_1^{(i)}) < d(P_2, P_1^{(i)}) < d(P_2, P_1)$$

by the hyperbolic law of cosines. In other words, by the two-step process described above, we obtain a triangle with vertices at  $P_0$ ,  $P_1^{(i)}$  and  $P_2^{(i)}$  so that all edge lengths are less than those of  $\Delta$ . Let  $\{\xi_i\}$  and  $\{\eta_i\}$  be the sequences such that  $\xi_i^{-1} \cdot z_0 = P_1^{(i)}$  and  $\eta_i \cdot z_0 = P_2^{(i)}$ . Then we have  $f_{z_0}^\dagger(\xi_i, \eta_i) < f_{z_0}^\dagger(\xi_0, \eta_0)$  for all but finitely many  $i$ , a contradiction. Hence, we conclude that  $(\xi_0, \eta_0) \in \mathfrak{G}\mathfrak{F}$ .

Finally, since the set of  $(\xi, \eta)$  such that  $\langle \xi, \eta \rangle$  is free, geometrically infinite and without any parabolic is dense in  $\overline{\mathfrak{G}\mathfrak{F}} - \mathfrak{G}\mathfrak{F}$  (see [6, Proposition 8.2]) and every  $(\xi, \eta) \in \mathfrak{X}$  with  $\langle \xi, \eta \rangle$  is free and without any parabolic is in  $\overline{\mathfrak{G}\mathfrak{F}}$  (see [6, Proposition 9.3]), we reduce the geometrically finite case to the geometrically infinite case proving the theorem.  $\square$

## References

- [1] **I Agol**, *Tameness of hyperbolic 3-manifolds* arXiv:math/0405568
- [2] **I Agol, M Culler, P B Shalen**, *Singular surfaces, mod 2 homology, and hyperbolic volume, I*, Trans. Amer. Math. Soc. 362 (2010) 3463–3498 MR2601597
- [3] **J W Anderson, R D Canary, M Culler, P B Shalen**, *Free Kleinian groups and volumes of hyperbolic 3-manifolds*, J. Differential Geom. 43 (1996) 738–782 MR1412683
- [4] **D Calegari, D Gabai**, *Shrinkwrapping and the taming of hyperbolic 3-manifolds*, J. Amer. Math. Soc. 19 (2006) 385–446 MR2188131
- [5] **M Culler, S Hersonsky, P B Shalen**, *The first Betti number of the smallest closed hyperbolic 3-manifold*, Topology 37 (1998) 805–849 MR1607748
- [6] **M Culler, P B Shalen**, *Paradoxical decompositions, 2-generator Kleinian groups, and volumes of hyperbolic 3-manifolds*, J. Amer. Math. Soc. 5 (1992) 231–288 MR1135928
- [7] **M Culler, P B Shalen**, *Betti numbers and injectivity radii*, Proc. Amer. Math. Soc. 137 (2009) 3919–3922 MR2529901
- [8] **M Culler, P B Shalen**, *Margulis numbers for Haken manifolds*, Israel J. Math. 190 (2012) 445–475 MR2956250
- [9] **D Gabai, R Meyerhoff, P Milley**, *Minimum volume cusped hyperbolic three-manifolds*, J. Amer. Math. Soc. 22 (2009) 1157–1215 MR2525782
- [10] **D Gabai, R Meyerhoff, P Milley**, *Mom technology and volumes of hyperbolic 3-manifolds*, Comment. Math. Helv. 86 (2011) 145–188 MR2745279

- [11] **A Marden**, *The geometry of finitely generated Kleinian groups*, Ann. of Math. 99 (1974) 383–462 MR0349992
- [12] **P Milley**, *Minimum volume hyperbolic 3-manifolds*, J. Topol. 2 (2009) 181–192 MR2499442
- [13] **P J Nicholls**, *The ergodic theory of discrete groups*, London Math. Soc. Lecture Note Series 143, Cambridge Univ. Press (1989) MR1041575
- [14] **S J Patterson**, *Lectures on measures on limit sets of Kleinian groups*, from: “Fundamentals of hyperbolic geometry: Selected expositions”, (R D Canary, D Epstein, A Marden, editors), London Math. Soc. Lecture Note Ser. 328, Cambridge Univ. Press (2006) 291–335 MR2235713
- [15] **D Sullivan**, *The density at infinity of a discrete group of hyperbolic motions*, Inst. Hautes Études Sci. Publ. Math. (1979) 171–202 MR556586
- [16] **D Sullivan**, *On the ergodic theory at infinity of an arbitrary discrete group of hyperbolic motions*, from: “Riemann surfaces and related topics: Proceedings of the 1978 Stony Brook Conference”, (I Kra, B Maskit, editors), Ann. of Math. Stud. 97, Princeton Univ. Press (1981) 465–496 MR624833

*Basic Sciences Unit, TED University*

*Ziya Gökalp St., No. 48, Kolej 06420, Çankaya, Ankara, Turkey*

`ilkers.yuce@tedu.edu.tr`

Received: 16 December 2009

Revised: 30 May 2014