The Topology of the Relative Character Varieties of a Quadruply-Punctured Sphere

Robert L. Benedetto and William M. Goldman

CONTENTS

Let M be a quadruply-punctured sphere with boundary components A, B, C, D. The $SL(2,\mathbb{C})$-character variety of M consists of equivalence classes of homomorphisms ρ of $\pi_1(M) \longrightarrow SL(2,\mathbb{C})$ and can be identified with a quartic hypersurface in $\mathbb{C}^7$. For fixed $a, b, c, d \in \mathbb{C}$, the subset $V_{a,b,c,d}$ corresponding to representations ρ with $\mathrm{tr}(\rho(A)) = a$, $\mathrm{tr}(\rho(B)) = b$, $\mathrm{tr}(\rho(C)) = c$, $\mathrm{tr}(\rho(D)) = d$ is a cubic surface in $\mathbb{C}^3$. We determine the singular points of $V_{a,b,c,d}$ and classify its set $V_{a,b,c,d}(\mathbb{R})$ of $\mathbb{R}$-points into six topological types, at least when this set is nonsingular. $V_{a,b,c,d}(\mathbb{R})$ contains a compact component if and only if $-2 < a, b, c, d < 2$. For certain values of (a, b, c, d), this component corresponds to representations in $SL(2,\mathbb{R})$.

1. INTRODUCTION

Moduli spaces of representations of fundamental groups of surfaces arise in many natural algebraic, geometric, and analytic problems. In particular, the classification of geometric structures on manifolds and solutions of the Yang-Mills equations from gauge theory lead to moduli spaces closely related to representations of surface groups. Furthermore these moduli spaces enjoy a rich symmetry due to the large topological symmetry groups of surfaces. The present paper describes in detail the topology of some of these moduli spaces, in the case of surfaces homeomorphic to a quadruply-punctured sphere.

Decompositions of surfaces into subsurfaces provide a technique for analyzing the moduli spaces in terms of moduli spaces for simpler surfaces. For example, every compact connected surface M of genus g with b boundary components decomposes along $3g+b-3$ simple closed curves into $2g+b-2$ pants, provided that

$$0 > \chi(M) = 2 - 2g - b.$$

A pants is a triply-punctured sphere: a surface of genus 0 with three boundary components. In general, we shall refer to a connected surface of genus g

1991 Mathematics Subject Classification: 57M05 (Low-dimensional topology)

Keywords: surface, fundamental group, character variety, representation variety

Goldman gratefully acknowledges partial support from National Science Foundation grants DMS-8613576, DMS-8902619, DMS-9205139, DMS-9504764, the Alfred P. Sloan Foundation, and the University of Maryland Institute for Advanced Computer Studies.

with b boundary components as a *b-punctured surface of genus g,* although, strictly speaking the ends of a "punctured surface" are noncompact and do not have boundary. Similarly, a *b-punctured torus* is a surface of genus 1 with b boundary components. Surfaces of small negative Euler characteristic are the building blocks for all surfaces of negative Euler characteristic and their moduli spaces are the building blocks for moduli spaces of more complicated surfaces. This decomposition technique has been used in [Goldman 1997] to study the automorphism group of surfaces and in [Goldman 1988a] to analyze moduli spaces of geometric structures on surfaces.

In this paper we are concerned with the moduli space of representations of a quadruply-punctured sphere into $\mathrm{SL}(2,\mathbb{C})$. This moduli space is naturally parametrized by the moduli of the images of the boundary components, which are described by four scalar parameters (a,b,c,d). For fixed values of (a,b,c,d), the moduli space is a surface $V_{a,b,c,d}$. We determine for which (a,b,c,d) is $V_{a,b,c,d}$ nonsingular.

Theorem 1.1. *For fixed $a,b,c,d \in \mathbb{C}$, the surface $V_{a,b,c,d}$ is singular if and only if either at least one of a,b,c,d equals ± 2 or there is a reducible representation ρ with boundary traces a,b,c,d. The latter case occurs when $\Delta(a,b,c,d)=0$ where $\Delta(a,b,c,d)$ is the polynomial*

$$\Delta(a,b,c,d) = \big(2(a^2+b^2+c^2+d^2) - abcd - 16\big)^2 \\ -(4-a^2)(4-b^2)(4-c^2)(4-d^2).$$

Thus $V_{a,b,c,d}(\mathbb{C})$ is a cubic surface in $\mathbb{C}^3$. Its projective completion $\bar{V}_{a,b,c,d}(\mathbb{C})$ is smooth at infinity. Furthermore $\bar{V}_{a,b,c,d}(\mathbb{C})\setminus V_{a,b,c,d}$ consists of three lines in general position. Using the classification of cubic surfaces over $\mathbb{C}$, the set of $\mathbb{C}$-points of $V_{a,b,c,d}$ deformation retracts to a bouquet of five 2-spheres.

Theorem 1.1 should be compared to Weil's original sufficient criterion [Weil 1964] for the smoothness of the representation variety. (Some of Weil's results are also expounded in [Raghunathan 1972, § VI].) For the analogous case of representations of fundamental groups of *closed* surfaces, an $\mathrm{SL}(2,\mathbb{C})$-representation is a smooth point of R if and only if it is irreducible. In the present context there are irreducible characters (namely when one of the boundary traces equals ± 2) which define singular points of the relative character variety.

By contrast, the topologies of the *real* varieties are much more varied. For $(a,b,c,d) \in \mathbb{R}$, the set $V_{a,b,c,d}(\mathbb{R})$ of $\mathbb{R}$-points of $V_{a,b,c,d}$ — at least when nonsingular — belongs to one of the six topological types depicted in Figure 1, depending on the boundary parameters a,b,c,d. Let n denote the number of boundary traces a,b,c,d which lie in the interval $(-2,2)$:

Theorem 1.2. *Let $a,b,c,d \in \mathbb{R}$ be such that $V_{a,b,c,d}$ is nonsingular. Then the set $V_{a,b,c,d}(\mathbb{R})$ of $\mathbb{R}$-points is homeomorphic to one of the following:*

1. *A quadruply-punctured sphere if $n=0$ and $abcd<0$;*
2. *A disjoint union of a triply-punctured torus and a disc if $n=0$ and $abcd>0$;*
3. *A disjoint union of a triply-punctured sphere and a disc if $n=1$;*
4. *A disjoint union of an annulus and two discs if $n=2$;*
5. *A disjoint union of four discs if $n=3$;*
6. *A disjoint union of four discs and a sphere if $n=4$.*

Corollary 1.3. *For a,b,c,d and n as above, the Euler characteristic of $V_{a,b,c,d}(\mathbb{R})$ is equal to $2n-2$.*

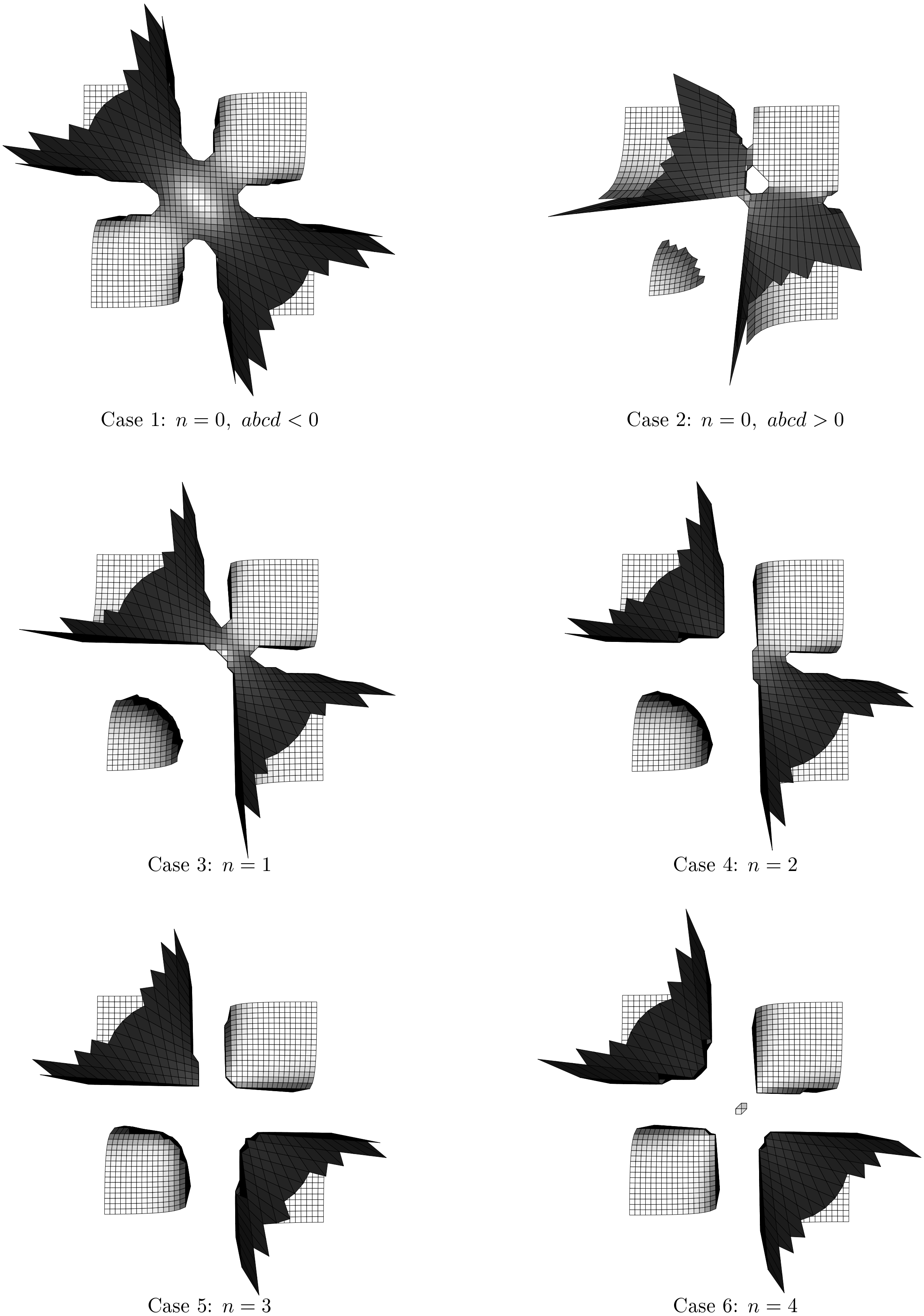
The points of $V_{a,b,c,d}(\mathbb{R})$ correspond to representations into either $\mathrm{SU}(2)$ or $\mathrm{SL}(2,\mathbb{R})$. Components of spaces of representations into $\mathrm{SU}(2)$ are always compact, and in the analogous case where M is a once-punctured torus, the only compact components of relative real character varieties of M correspond to $\mathrm{SU}(2)$-representations. When $-2 < a,b,c,d < 2$, the variety $V_{a,b,c,d}(\mathbb{R})$ contains a compact component, and we determine conditions when this compact component corresponds to representations in $\mathrm{SU}(2)$ or $\mathrm{SL}(2,\mathbb{R})$.

Proposition 1.4. *Suppose $a,b,c,d \in (-2,2)$ and $V_{a,b,c,d}$ is smooth. Then $V_{a,b,c,d}(\mathbb{R})$ contains a compact component, which consists of characters of $\mathrm{SL}(2,\mathbb{R})$-representations if and only if $\Delta(a,b,c,d)>0$ and*

$$16 - abcd - 2(a^2+b^2+c^2+d^2) > 0.$$

Otherwise, each element in the component is the character of an $\mathrm{SU}(2)$-representation.

The component of $V_{a,b,c,d}(\mathbb{R})$ homeomorphic to a disc in case 2 of Theorem 1.2 corresponds to the "Teichmüller space" of M and consists of characters of holonomy representations of hyperbolic structures on M with geodesic boundary. In all of the last four cases of Theorem 1.2, the components which are homeomorphic to a disc admit similar interpretations as deformation spaces of hyperbolic structures with conical singularities; see [Goldman 1988a].



Case 1: $n = 0,\ abcd < 0$

Case 2: $n = 0,\ abcd > 0$

Case 3: $n = 1$

Case 4: $n = 2$

Case 5: $n = 3$

Case 6: $n = 4$

FIGURE 1. Possible topological types of $V_{a,b,c,d}(\mathbb{R})$ (see Theorem 1.2). The type is determined by the number n of boundary traces a, b, c, d that lie in the interval $(-2, 2)$ and (if $n = 0$) by the sign of $abcd$.

2. CHARACTER VARIETIES FOR SURFACES

Let $\Bbbk$ denote a field (usually either $\mathbb{R}$ or $\mathbb{C}$) and let $G = \mathrm{SL}(2,\Bbbk)$. For the moment let π denote an arbitrary finitely generated group. We denote the set of homomorphisms $\pi \longrightarrow G$ by $\mathrm{Hom}(\pi, G)$; since G is an affine algebraic group defined over $\Bbbk$, $\mathrm{Hom}(\pi, G)$ naturally enjoys the structure of the set of $\Bbbk$-points of an affine algebraic set defined over $\mathbb{Z}$. (Compare [Lubotzky and Magid 1985, § 1].) A central problem is to determine the global topology and geometric structure on the space $\mathrm{Hom}(\pi, G)$.

We would like to consider the space of equivalence classes of such representations under the natural action of $\mathrm{SL}(2,\mathbb{C})$ by conjugation, but the quotient space is generally not well-behaved. By enlarging the equivalence relation slightly, one can define a *categorical quotient* $\mathrm{Hom}(\pi, G)//G$ (in the sense of algebraic geometry) which does have the structure of an affine algebraic set. For brevity denote $\mathrm{Hom}(\pi, G)$ by R. The guiding principle is that the regular functions on the quotient $R//G$ should be the G-invariant regular functions on R, and thus the quotient $R//G$ is *defined* as $\mathrm{Spec}(\Bbbk[R]^G)$. The points of $R//G$ correspond to equivalence classes, where two orbits are considered equivalent if they cannot by distinguished by G-invariant regular functions on R. In general there is a G-invariant Zariski-dense subset R^- such that the restriction of the quotient map $R \longrightarrow R//G$ to R^- defines a bijection of the orbit space R^-/G to $R//G$. In practice, we can take (assuming $\Bbbk$ is algebraically closed) R^- to be the set of *completely reducible* (or *semisimple*) representations $\pi \longrightarrow G$. We denote the quotient $R//G$ by X. For a careful treatment of this material, see [Lubotzky and Magid 1985, § 1].

2A. Cyclic Groups

When π is a cyclic group, then a homomorphism $\pi \longrightarrow G$ is completely determined by the image of a generator of π. Therefore the representation variety R is isomorphic to G. Every regular function on G invariant under conjugation is a polynomial in the trace function

$$\begin{aligned} G &\longrightarrow \Bbbk \\ A &\longmapsto \mathrm{tr}(A). \end{aligned}$$

Thus $X \cong \Bbbk$ under the trace mapping. Note that the trace function has critical values ± 2, with critical points at the center $\pm I$.

More generally, suppose that π is a free group freely generated by $x_1, \dots, x_n$. Then $\mathrm{Hom}(\pi, G)$ is identified with the Cartesian product G^n. Procesi [1976] proved that every G-invariant function on G^n is a polynomial in finitely many functions of the form

$$\begin{aligned} G^n &\longrightarrow \Bbbk \\ (A_1, \dots, A_n) &\longmapsto \mathrm{tr}\, w(A_1, \dots, A_n), \end{aligned}$$

where $w(x_1, \dots, x_n)$ is a word in $x_1, \dots, x_n$, that is, an element of π.

For example, for $n = 2$ and $\Bbbk = \mathbb{C}$, classical work dating at least as far back as Fricke and Klein [1897, Section II.2, page 285] implies the following fundamental result:

Theorem 2.1. *The invariant function*

$$\begin{aligned} \chi : \mathrm{Hom}(\pi, G) &\longrightarrow \mathbb{C}^3 \\ \rho &\longmapsto \begin{bmatrix} \mathrm{tr}(\rho(x_1)) \\ \mathrm{tr}(\rho(x_2)) \\ \mathrm{tr}(\rho(x_1x_2)) \end{bmatrix} \end{aligned}$$

defines an isomorphism $X \longrightarrow \mathbb{C}^3$. That is, every function of pairs $(A_1, A_2) \in \mathrm{SL}(2,\mathbb{C}) \times \mathrm{SL}(2,\mathbb{C})$ invariant under simultaneous conjugation

$$(A_1, A_2) \mapsto (TA_1T^{-1}, TA_2T^{-1})$$

can be expressed is a function of $\mathrm{tr}(A_1)$, $\mathrm{tr}(A_2)$, *and* $\mathrm{tr}(A_1A_2)$. *Furthermore, if A_1, A_2 generate an irreducible representation on $\mathbb{C}^2$, then for any other pair A_1', A_2' with*

$$\begin{aligned} \mathrm{tr}(A_1) &= \mathrm{tr}(A_1'), \\ \mathrm{tr}(A_2) &= \mathrm{tr}(A_2'), \\ \mathrm{tr}(A_1A_2) &= \mathrm{tr}(A_1'A_2'), \end{aligned}$$

there exists $T \in \mathrm{SL}(2,\mathbb{C})$ with $A_1' = TA_1T^{-1}$, $A_2' = TA_2T^{-1}$.

We call the triple $\chi(\rho)$ the *character* of ρ, since from it one can deduce the trace $\mathrm{tr}(\rho(A))$ for any $A \in \pi$. Compare [Magnus 1980].

The criterion for reducible representations is somewhat more complicated. If ρ is reducible, then it is conjugate to a representation by upper-triangular matrices

$$\rho(\gamma) = \begin{bmatrix} a(\gamma) & b(\gamma) \\ 0 & a(\gamma)^{-1} \end{bmatrix}.$$

It and its *semisimplification*

$$\rho^s(\gamma) = \begin{bmatrix} a(\gamma) & 0 \\ 0 & a(\gamma)^{-1} \end{bmatrix}$$

have the same character, but in general will not be conjugate. Indeed, ρ^s lies in the closure of the orbit of ρ. Two reducible representations have the same character if and only if their semisimplifications are conjugate, or equivalently, if their orbit closures are equal.

For example, the G-invariant function on $G \times G$ defined by

$$f : (A_1, A_2) \mapsto \operatorname{tr}(A_1 A_2 A_1^{-1} A_2^{-1})$$

is given by the expression

$$f(A_1, A_2) = \kappa\big(\operatorname{tr}(A_1),\ \operatorname{tr}(A_2),\ \operatorname{tr}(A_1 A_2)\big),$$

where $\kappa \in \mathbb{C}[x, y, z]$ is the cubic polynomial defined by

$$\kappa(x, y, z) = x^2 + y^2 + z^2 - xyz - 2. \qquad (2\text{–}1)$$

A slight variant of this polynomial will be useful later on. Let $a, b \in \Bbbk$. Then $\kappa_{a,b}(x) \in \Bbbk[x]$ is the *quadratic* polynomial defined by

$$\kappa_{a,b}(x) = \kappa(a, b, x) - 2 = x^2 - abx + (a^2 + b^2 - 4).$$

The reducible representations occur for characters (x, y, z) with

$$\kappa(x, y, z) = 2,$$

where $\kappa : \mathbb{C}^3 \longrightarrow \mathbb{C}$ is defined by (2–1). We call such a character *reducible*. See [Goldman 1988b] for details.

The set of $\mathbb{R}$-points in X equals $X(\mathbb{R}) = \mathbb{R}^3$. By [Morgan and Shalen 1984, Prop. III.1.1, p. 458], real characters correspond to representations taking values in the real forms $\mathrm{SU}(2)$ and $\mathrm{SL}(2, \mathbb{R})$ of $\mathrm{SL}(2, \mathbb{C})$. A point $(x, y, z) \in \mathbb{R}^3$ corresponds to an $\mathrm{SU}(2)$-representation if and only if

$$-2 \le x, y, z \le 2 \qquad \text{and} \qquad \kappa(x, y, z) \le 2.$$

Otherwise $(x, y, z) \in \mathbb{R}^3$ corresponds to an $\mathrm{SL}(2, \mathbb{R})$-representation. A character $(x, y, z) \in \mathbb{R}^3$ corresponds both to an $\mathrm{SU}(2)$-representation and to an $\mathrm{SL}(2, \mathbb{R})$-representation if and only if it corresponds to a representation in

$$\mathrm{SO}(2) = \mathrm{SU}(2) \cap \mathrm{SL}(2, \mathbb{R});$$

such a representation is necessarily reducible. See [Goldman 1988b] for details.

2B. The Triply-Punctured Sphere

We interpret Theorem 2.1 in terms of the the triply-punctured sphere. For any compact manifold, set

$$\mathrm{X}(M) = \coprod_i \operatorname{Hom}(\pi_1(M_i), G)/\!/G,$$

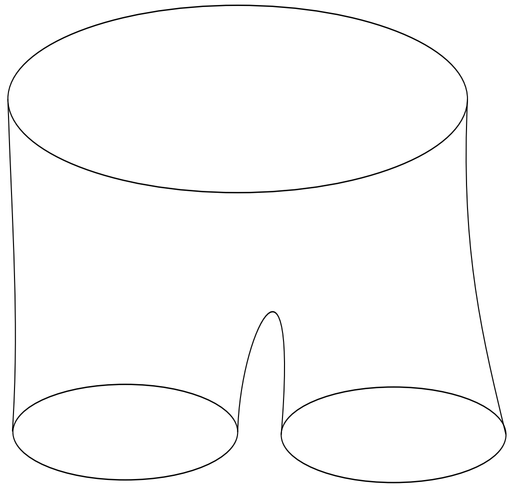


FIGURE 2. A triply-punctured sphere and its three boundary components.

where

$$M = \coprod_i M_i$$

is the decomposition of M into connected components. When M is a compact surface with boundary

$$\partial M = \coprod_{i=1}^{m} \partial_i M$$

each component $\partial_i M$ has fundamental group

$$\pi_1(\partial_i M) \cong \mathbb{Z}.$$

Thus each $\mathrm{X}(\partial_i M) \cong \mathbb{C}$ and $\mathrm{X}(\partial M) \cong \mathbb{C}^m$. The homomorphism of fundamental groups

$$\pi_1(\partial_i M) \longrightarrow \pi_1(M)$$

induces a restriction map

$$\mathrm{X}(M) \longrightarrow \mathrm{X}(\partial_i M) \cong \mathbb{C}.$$

Taking the product, we obtain

$$\partial_* : \mathrm{X}(M) \longrightarrow \mathbb{C}^m.$$

Let M denote a triply-punctured sphere. The fundamental group of M admits the redundant geometric presentation

$$\pi = \langle X, Y, Z \mid XYZ = I \rangle,$$

where X, Y, Z are simple loops corresponding to the three boundary components. In this case the boundary map ∂_* defines an isomorphism $\mathrm{X}(M) \approx \mathbb{C}^3$, where

$$\partial_*([\rho]) = \big(\operatorname{tr}(\rho(X)),\ \operatorname{tr}(\rho(Y)),\ \operatorname{tr}(\rho(Z))\big).$$

2C. Properties of $\kappa(x, y, z)$

The polynomial κ and its variant $\kappa_{a,b}$ satisfy many remarkable properties related to their natural occurrence as invariants of elements of $\mathrm{SL}(2,\mathbb{C})$.

We use the following lemma concerning $\kappa_{a,b}$ and $\mathbb{R}$-points extensively in later sections. Its proof is straightforward.

Lemma 2.2. *Fix $a, b \in \mathbb{R}$. Then $\kappa_{a,b}$ has:*

- *one real double root if either of a or b is ± 2;*
- *no real roots if $|a| > 2$ and $|b| < 2$ or vice versa;*
- *two real roots, both greater than or equal to 2, if $|a|, |b| > 2$ and $ab > 0$;*
- *two real roots, both less than or equal to -2, if $|a|, |b| > 2$ and $ab < 0$;*
- *two real roots, both in $[-2, 2]$, if $|a|, |b| < 2$.*

This lemma illustrates the properties of reducible characters, since x is a root of $\kappa_{a,b}$ if and only if (a, b, x) is a reducible character. Indeed, we may assume that (a, b, x) is the character of a representation with abelian image. For example, the second case occurs when a is the trace of a hyperbolic element and b is the trace of an elliptic element, and in $\mathrm{SL}(2,\mathbb{R})$ hyperbolic elements and elliptic elements cannot commute. The other cases can be discussed similarly.

An important property of κ is that its restriction to every coordinate plane and coordinate line is quadratic. In particular its values at ± 2 are very special.

Lemma 2.3. $\kappa_{a,b}(\pm 2) = (a \pm b)^2$.

The discriminant of the quadratic polynomial $\kappa_{a,b}(x)$ equals

$$(a^2 - 4)(b^2 - 4).$$

Furthermore the critical points of $\kappa : \mathbb{C}^3 \longrightarrow \mathbb{C}$ are the four points

$$(2, 2, 2), \quad (2, -2, -2), \quad (-2, 2, -2), \quad (-2, -2, 2)$$

with critical value $+2$ and the point $(0, 0, 0)$ with critical value -2.

3. THE QUADRUPLY-PUNCTURED SPHERE

Let M be a quadruply-punctured sphere, and let its boundary components be A, B, C, D. Denoting the corresponding elements of $\pi = \pi_1(M)$ by the same names, the fundamental group $\pi = \pi_1(M)$ has the redundant geometric presentation

$$\pi = \langle A, B, C, D \mid ABCD = I\rangle;$$

compare Figure 3. However, in the present context there is an even more concrete description. A representation $\rho : \pi \longrightarrow \mathrm{SL}(2,\mathbb{C})$ corresponds to a quadruple

$$\alpha = \rho(A), \quad \beta = \rho(B),$$
$$\gamma = \rho(C), \quad \delta = \rho(D)$$

in $\mathrm{SL}(2,\mathbb{C})^4$, where

$$\alpha\beta\gamma\delta = I.$$

Using this equation to eliminate δ, such a quadruple merely refers to an arbitrary triple $(\alpha, \beta, \gamma) \in \mathrm{SL}(2,\mathbb{C})^3$.

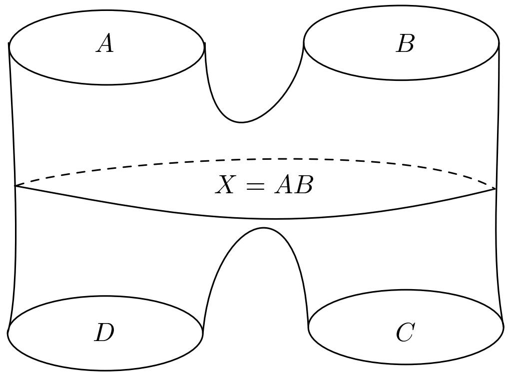


FIGURE 3. A quadruply-punctured sphere.

3A. Amalgamating Representations

Let ξ denote the simple loop on M which separates boundary components A, B from C, D. The corresponding element of the fundamental group satisfies:

$$X = AB = (CD)^{-1}.$$

Denote by $M|\xi$ the compact surface obtained by *splitting* M along ξ: each of the two components M_1 and M_2 of $M|\xi$ are homeomorphic to a triply-punctured sphere. We express the character variety $\mathrm{X}(M)$ in terms of the character variety $\mathrm{X}(M|\xi)$.

The fundamental group of M_1 has presentation

$$\pi_1(M_1) = \langle A, B, \bar{X} \mid AB\bar{X} = I\rangle$$

and the fundamental group of M_2 has presentation

$$\pi_1(M_2) = \langle X, C, D \mid XCD = I\rangle.$$

The fundamental group of M is the amalgamated free product

$$\pi_1(M) = \pi_1(M_1) \amalg_{\mathbb{Z}} \pi_1(M_2)$$

amalgamated by the monomorphisms taking a generator of $\mathbb{Z}$ to $\bar{X}^{-1} \in \pi_1(M_1)$ and $X \in \pi_1(M_2)$. A representation

$$\rho : \pi_1(M) \longrightarrow \mathrm{SL}(2,\mathbb{C})$$

restricts to two representations

$$\rho_i : \pi_1(M_i) \longrightarrow \mathrm{SL}(2,\mathbb{C})$$

satisfying

$$\rho_1(\bar{X})\,\rho_2(X) = I. \tag{3–1}$$

Conversely, two representations ρ_1, ρ_2 that satisfy (3–1) define a unique representation ρ. Compare [Goldman 1988a].

3B. Defining Equations

The boundary map $\partial_* : \mathrm{X}(M) \longrightarrow \mathbb{C}^4$ is defined by four functions

$$\begin{aligned} a([\rho]) &= \mathrm{tr}(\rho(A)), & b([\rho]) &= \mathrm{tr}(\rho(B)),\\ c([\rho]) &= \mathrm{tr}(\rho(C)), & d([\rho]) &= \mathrm{tr}(\rho(D)). \end{aligned}$$

The elements $X = AB$, $Y = BC$ and $Z = CA$ of π are represented by simple loops on M. The corresponding traces x, y, z, together with the boundary traces a, b, c, d, define coordinates on $\mathrm{X}(M)$ that are subject to the relation

$$\begin{aligned} &x^2 + y^2 + z^2 + xyz\\ &\quad = (ab + cd)x + (ad + bc)y + (ac + bd)z\\ &\qquad\quad +(4 - a^2 - b^2 - c^2 - d^2 - abcd). \end{aligned} \tag{3–2}$$

Compare [Magnus 1980]. Unlike the case in which π is free of rank two, $\mathrm{X}(M)$ is not an affine space, but rather the hypersurface $V \subset \mathbb{C}^7$ consisting of points whose coordinates (a, b, c, d, x, y, z) satisfy (3–2). For each $(a, b, c, d) \in \mathbb{C}^4$, the set

$$\partial_*^{-1}(a, b, c, d) = V_{a,b,c,d}$$

consists of all $(x, y, z) \in \mathbb{C}^3$ satisfying (3–2). Rewrite the basic defining equation (3–2) in terms of x as

$$\begin{aligned} &\frac{2{+}x}{4}\left((y{+}z) - \frac{(a{+}b)(d{+}c)}{2{+}x}\right)^2\\ &\qquad + \frac{2{-}x}{4}\left((y{-}z) - \frac{(a{-}b)(d{-}c)}{2{-}x}\right)^2\\ &= \frac{(x^2{-}abx{+}a^2{+}b^2{-}4)(x^2{-}cdx{+}c^2{+}d^2{-}4)}{4{-}x^2}. \end{aligned} \tag{3–3}$$

We may write (3–3) as follows. For fixed a, b, c, d, x, the expression

$$\begin{aligned} Q_x^{(a,b,c,d)}(y,z) &= \frac{2+x}{4}\left((y+z) - \frac{(a+b)(d+c)}{2+x}\right)^2\\ &\qquad + \frac{2-x}{4}\left((y-z) - \frac{(a-b)(d-c)}{2-x}\right)^2 \end{aligned}$$

is an inhomogeneous quadratic function of (y, z). Equation (3–3) is equivalent to

$$Q_x^{(a,b,c,d)}(y,z) = \frac{\kappa_{a,b}(x)\kappa_{c,d}(x)}{4 - x^2},$$

which describes a conic $V_{a,b,c,d}(x)$ of center (y_0, z_0), where

$$\begin{aligned} y_0 + z_0 &= \frac{(a+b)(d+c)}{2+x},\\ y_0 - z_0 &= \frac{(a-b)(d-c)}{2-x}. \end{aligned}$$

3C. Symmetries

A group of 192 symmetries of the family $\mathrm{X}(M)$ acts by simple linear transformations on the parameters (a, b, c, d). This action is generated by an action of the *central characters,* which multiply an even number of the generators by $-I$, and the *permutations* of the four generators A, B, C, D.

The symmetric group S_4 consisting of permutations of $\{A, B, C, D\}$ acts on the family $\mathrm{X}(M)$ of relative character varieties $V_{a,b,c,d}$. The symmetries are apparent from the defining equation (3–2). For example, the three products of disjoint transpositions act trivially on the (x, y, z) coordinates, while nontrivially permuting (a, b, c, d). However, the transpositions $A \leftrightarrow B$, $B \leftrightarrow C$, $C \leftrightarrow D$, act respectively as follows:

$$\begin{bmatrix} a\\ b\\ c\\ d\\ x\\ y\\ z \end{bmatrix} \mapsto \begin{bmatrix} b\\ a\\ c\\ d\\ x\\ z\\ y \end{bmatrix}, \quad \begin{bmatrix} a\\ b\\ c\\ d\\ x\\ y\\ z \end{bmatrix} \mapsto \begin{bmatrix} a\\ c\\ b\\ d\\ z\\ y\\ x \end{bmatrix}, \quad \begin{bmatrix} a\\ b\\ c\\ d\\ x\\ y\\ z \end{bmatrix} \mapsto \begin{bmatrix} a\\ b\\ d\\ c\\ x\\ z\\ y \end{bmatrix}$$

Furthermore every character variety admits an action by *central characters*, as described in [Lubotzky and Magid 1985, § 5] (where this operation is called *twisting* by a character). If $Z \subset G$ is central, then $\mathrm{Hom}(\pi, Z)$ is a *group,* which acts on $\mathrm{Hom}(\pi, G)$ by pointwise multiplication. Since the center of $G = \mathrm{SL}(2,\mathbb{C})$ equals $\{\pm I\} \cong \mathbb{Z}/2$ and the fundamental group π of a quadruply-punctured sphere is free of rank 3, the group

$$\mathrm{Hom}(\pi, \{\pm I\}) \cong \mathbb{Z}/2 \oplus \mathbb{Z}/2 \oplus \mathbb{Z}/2$$

acts on $V_{a,b,c,d}$ as follows. Such a central character maps the images of an even number of the generators A, B, C, D to their negatives: for example,

$$\begin{bmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{bmatrix} \mapsto \begin{bmatrix} -\alpha \\ -\beta \\ \gamma \\ \delta \end{bmatrix}$$

is the action on generators corresponding to

$$(1, 1, 0) \in \mathbb{Z}/2 \oplus \mathbb{Z}/2 \oplus \mathbb{Z}/2$$

and whose action on the trace coordinates is given by

$$\begin{bmatrix} a \\ b \\ c \\ d \\ x \\ y \\ z \end{bmatrix} \mapsto \begin{bmatrix} -a \\ -b \\ c \\ d \\ x \\ -y \\ -z \end{bmatrix}.$$

S_4 normalizes this $\mathrm{Hom}(\pi, \{\pm I\})$-action. Since the induced S_4-action on $\mathrm{Hom}(\pi, \{\pm I\})$ is transitive, the action of every central character is obtained by conjugating the example above by a permutation. These considerations of symmetry immediately imply the corresponding relative character varieties are isomorphic:

Lemma 3.1. *Let $(a, b, c, d) \in \mathbb{R}^4$. Let $\sigma : \mathbb{R}^4 \to \mathbb{R}^4$ be any permutation of coordinates, and let $\tau : \mathbb{R}^4 \to \mathbb{R}^4$ be any map which acts by reversing signs of two of the coordinates and leaving the other two coordinates unchanged. Then $V_{\sigma(a,b,c,d)}$ and $V_{\tau(a,b,c,d)}$ are each diffeomorphic to $V_{a,b,c,d}$.*

4. SINGULAR POINTS AND REDUCIBLE CHARACTERS

4A. A Criterion for Singularity

Theorem 4.1. *$V_{a,b,c,d}$ is singular if and only if at least one of these conditions holds:*

- *One of a, b, c, d equals ± 2.*
- *Some reducible representation has boundary traces (a, b, c, d).*

The latter case occurs if and only if there exists $x \in \mathbb{C}$ such that $\kappa_{a,b}(x) = \kappa_{c,d}(x) = 0$.

4B. Solutions of Quadratic Equations

We will need the following technical lemmas on discriminant loci for the proof of Theorem 4.1 and for later results.

Lemma 4.2. *Let $n \geq 1$ be an integer, and let $\Bbbk$ be a field of characteristic not equal to 2. Pick $B(x)$ and $C(x)$ in $\Bbbk[x_1, \dots, x_n]$. Let*

$$Y = \{(x_1, \dots, x_n, z) \in \Bbbk^{n+1} \mid z^2 + B(x)z + C(x) = 0\}$$

and

$$Z = \{(x_1, \dots, x_n) \in \Bbbk^n \mid B(x)^2 - 4C(x) = 0\}.$$

If Y is singular at $(x_1, \dots, x_n, z)$, then Z is singular at $(x_1, \dots, x_n)$. If Z is singular at $(x_1, \dots, x_n)$, then there exists $z \in \Bbbk$ such that Y is singular at $(x_1, \dots, x_n, z)$.

(If $n = 1$, then Z is singular at x if x is a multiple root of $B(x)^2 - 4C(x)$.)

Proof. Denote by $dB(x)$ the differential of $B(x)$ with respect to the x_i-variables *only*, and similarly for $dC(x)$. Let $(x_1, \dots, x_n) \in \Bbbk^n$. Now Z is singular at $(x_1, \dots, x_n)$ if and only if

$$\left.\begin{aligned} B(x)^2 - 4C(x) &= 0, \\ 2B(x)\, dB(x) - 4\, dC(x) &= 0, \end{aligned}\right\} \tag{4–1}$$

and Y is singular at $(x_1, \dots, x_n, z)$ if and only if

$$\left.\begin{aligned} z^2 + B(x)z + C(x) &= 0, \\ 2z + B(x) &= 0, \\ z\, dB(x) + dC(x) &= 0. \end{aligned}\right\} \tag{4–2}$$

We must show that (4–1) is equivalent to the existence of $z \in \Bbbk$ satisfying (4–2). Note that

$$B(x)^2 - 4C(x) = 0$$

if and only if

$$z^2 + B(x)\, z + C(x)$$

has a double root $z \in \Bbbk$; this root must be

$$z = -\tfrac{1}{2} B(x).$$

Substituting $z = -\frac{1}{2} B(x)$ into the final equation of (4–2) and multiplying by -4 yields the second equation of (4–1). □

Lemma 4.3. *Let $\Bbbk$ be an algebraically closed field of characteristic not equal to 2. Suppose that $A_1(x)$, $B_1(x)$, $C_1(x)$ are elements of $\Bbbk[x]$ such that $A_1(x)$ has no multiple roots. Let*

$$Y = \{(x, y) \in \Bbbk^2 \mid A_1(x)y^2 + B_1(x)y + C_1(x) = 0\}$$

and

$$Z = \{x \in \Bbbk \mid B_1(x)^2 - 4A_1(x)C_1(x) = 0\}.$$

If Y is singular at (x,y), then Z is singular at x. If Z is singular at x, then there exists $y \in k$ such that Y is singular at (x,y).

Proof. The proof follows that of Lemma 4.2, modified in the case that $A_1(x) = 0$. Suppose that $A_1(x) = 0$. Then Z is singular at x if and only if

$$\left.\begin{aligned} B_1(x)^2 &= 0, \\ 2B_1(x)B_1'(x) - 4C_1(x)A_1'(x) &= 0, \end{aligned}\right\} \tag{4–3}$$

and Y is singular at (x,y) if and only if there exists $y \in \Bbbk$ such that

$$\left.\begin{aligned} yB_1(x) + C_1(x) &= 0, \\ B_1(x) &= 0, \\ y^2A_1'(x) + yB_1'(x) + C_1'(x) &= 0. \end{aligned}\right\} \tag{4–4}$$

Clearly (4–3) is equivalent to $B_1(x) = C_1(x) = 0$, which (4–4) implies. Since $A_1'(x) \neq 0$ (by hypothesis), and $\Bbbk$ is algebraically closed, some $y \in \Bbbk$ satisfies the last equation in (4–4), concluding the proof of Lemma 4.3. □

Lemma 4.4. *Let $\Bbbk = \mathbb{R}$. Suppose that $A_1(x)$, $B_1(x)$, $C_1(x)$ are elements of $\mathbb{R}[x]$ such that $A_1(x)$ has no multiple roots. Set*

$$Y = \{(x,y) \in \mathbb{R}^2 \mid A_1(x)y^2 + B_1(x)y + C_1(x) = 0\},$$
$$Z = \{x \in \mathbb{R} \mid B_1(x)^2 - 4A_1(x)C_1(x) = 0\},$$

and let

$$\delta(x) = B_1'(x)^2 - 4A_1'(x)C_1'(x). \tag{4–5}$$

If Y is singular at (x,y), then Z is singular at x. Suppose for some $x \in \mathbb{R}$ that

$$A_1(x) = B_1(x) = C_1(x) = 0$$

and

$$\delta(x) \geq 0.$$

If Z is singular at x, then there exists $y \in \mathbb{R}$ such that Y is singular at (x,y).

Proof. The only point where the proof of Lemma 4.3 fails for $\Bbbk = \mathbb{R}$ is the final step, where the last equation in (4–4) may not have a solution $y \in \mathbb{R}$. Since the discriminant of this quadratic equation equals $\delta(x)$, our assumption $\delta(x) \geq 0$ guarantees the existence of a solution $y \in \mathbb{R}$. □

4C. Singularities of $\mathrm{V}_{a,b,c,d}$

Proof of Theorem 4.1. Fix $(a,b,c,d) \in \mathbb{C}^4$, and rewrite (3–2) as

$$z^2 + B(x,y)z + C(x,y) = 0,$$

where

$$\begin{aligned} B(x,y) &= xy - (ac+bd), \\ C(x,y) &= x^2 + y^2 - (ab+cd)x - (ad+bc)y \\ &\qquad - (4 - a^2 - b^2 - c^2 - d^2 - abcd). \end{aligned}$$

By Lemma 4.2, $V_{a,b,c,d}$ is singular if and only if the curve C in the xy-plane defined by the discriminant

$$f(x,y) = B(x,y)^2 - 4C(x,y) = 0$$

is singular. Now write $f(x,y)$ as a quadratic function in y:

$$f(x,y) = A_1(x)y^2 + B_1(x)y + C_1(x),$$

where

$$\begin{aligned} A_1(x) &= (x^2-4), \\ B_1(x) &= 4(ad+bc) - 2(ac+bd)x, \\ C_1(x) &= -4x^2+4(ab+cd)x+a^2c^2+b^2d^2 \\ &\qquad -2abcd-4a^2-4b^2-4c^2-4d^2+16 = 0. \end{aligned} \tag{4–6}$$

Since $A_1(x) = (x-2)(x+2)$ has simple zeroes, Lemma 4.3 implies that $V_{a,b,c,d}$ is singular if and only if the discriminant $g(x)$ of $f(x,y)$ as a quadratic function of y vanishes. This discriminant is easily computed to be

$$\begin{aligned} g(x) &= B_1(x)^2 - 4A_1(x)C_1(x) \\ &= 16(x^2-abx+a^2+b^2-4)(x^2-cdx+c^2+d^2-4) \\ &= 16(\kappa_{a,b}(x))(\kappa_{c,d}(x)). \end{aligned}$$

Thus $V_{a,b,c,d}$ is singular if and only if there exists $x \in \Bbbk$ satisfying one of the following three conditions:

1. x is a double zero of $\kappa_{a,b}$;
2. x is a double zero of $\kappa_{c,d}$;
3. $\kappa_{a,b}(x) = \kappa_{c,d}(x) = 0$.

Now the first case occurs if and only if the discriminant of $\kappa_{a,b}$ is zero, that is, if and only if

$$(a^2-4)(b^2-4) = 0,$$

that is, if and only if $a = \pm 2$ or $b = \pm 2$. Similarly, the second case occurs if and only if $c = \pm 2$ or $d = \pm 2$.

Consider the remaining case,

$$\kappa_{a,b}(x) = \kappa_{c,d}(x) = 0. \tag{4–7}$$

We have to find a reducible representation having traces $(a, b, c, d; x)$.

Decompose M along ξ as above into two triply-punctured spheres M_1 and M_2. (Compare Section 3A.) Then there exist representations

$$\rho_i : \pi_1(M_i) \longrightarrow \mathrm{SL}(2, \mathbb{C})$$

having characters (a, b, x) and (x, c, d) respectively. By (4–7), both ρ_1 and ρ_2 are reducible.

We distinguish two cases, depending on whether $x = \pm 2$ or not. If $x = \pm 2$, choose representations ρ_1, ρ_2 so that

$$\rho_1(\bar{X}) \neq \pm I \quad \text{and} \quad \rho_2(X) \neq \pm I.$$

Then $\rho_1(\bar{X})$ and $\rho_2(X)$ necessarily have unique fixed points, say p_1 and p_2, in $\mathbb{CP}^1$. There exists $g \in G$ such that

$$g(p_1) = p_2, \qquad g\rho_1(\bar{X}^{-1})g^{-1} = \rho_2(X). \tag{4–8}$$

Define ρ by

$$\begin{aligned} \rho(A) &= \rho_1(A), & \rho(B) &= \rho_1(B), \\ \rho(C) &= g\rho_2(C)g^{-1}, & \rho(D) &= g\rho_1(D)g^{-1}. \end{aligned} \tag{4–9}$$

Since the image of ρ fixes p_1, it is reducible and its character lies in $V_{a,b,c,d}$. (Alternatively, one can choose

$$\rho(\bar{X}) = \pm I$$

and $\rho(C), \rho(D)$ to be diagonal matrices.)

Now suppose $x \neq \pm 2$. Then $\rho_1(\bar{X}^{-1})$ and $\rho_2(X)$ are elements of G having trace x and are thus conjugate. Indeed, they are diagonalizable with distinct eigenvalues λ, λ^{-1} where $x = \lambda + \lambda^{-1}$. Let $p_1 \in \mathbb{CP}^1$ be the fixed point of $\rho_1(\bar{X}^{-1})$ corresponding to the λ-eigenspace of $\rho_1(\bar{X}^{-1})$ and $p_2 \in \mathbb{CP}^1$ be the fixed point of $\rho_2(X^{-1}))$ corresponding to the λ-eigenspace of $\rho_2(X^{-1}))$. Then as above there exists g satisfying (4–8). The representation ρ defined by (4–9) fixes p_1 and its character lies in $V_{a,b,c,d}$. □

4D. A Discriminant Polynomial

Given $(a, b, c, d) \in \mathbb{C}^4$, the existence of an $x \in \mathbb{C}$ satisfying condition (4–7) is equivalent to the vanishing of the *resultant* $\Delta(a, b, c, d)$ of the polynomials $\kappa_{a,b}$ and $\kappa_{c,d}$, which (remarkably) is a symmetric polynomial in a, b, c, d. This resultant is the sextic polynomial

$$\begin{aligned} \Delta(a,b,c,d) &= \begin{vmatrix} 1 & -ab & a^2+b^2-4 & 0 \\ 0 & 1 & -ab & a^2+b^2-4 \\ 1 & -cd & c^2+d^2-4 & 0 \\ 0 & 1 & -cd & c^2+d^2-4 \end{vmatrix} \\ &= \bigl(2(a^2+b^2+c^2+d^2) - abcd - 16\bigr)^2 \\ &\qquad - (4-a^2)(4-b^2)(4-c^2)(4-d^2). \end{aligned}$$

The vanishing of $\Delta(a, b, c, d)$ implies only that at least one character in $V_{a,b,c,d}$ is reducible. In general, a character

$$(a, b, c, d, x, y, z) \in V_{a,b,c,d}$$

is reducible if and only if

$$\begin{aligned} &\kappa_{a,b}(x) = \kappa_{c,d}(x) = \\ &\kappa_{a,d}(y) = \kappa_{b,c}(y) = \\ &\kappa_{a,c}(z) = \kappa_{b,d}(z) = 0. \end{aligned}$$

5. TOPOLOGY OF THE SET OF $\mathbb{R}$-POINTS

5A. Singular Points

For $(a, b, c, d) \in \mathbb{R}^4$, the variety $V_{a,b,c,d}$ is defined over $\mathbb{R}$; from now on we will only consider its set $V_{a,b,c,d}(\mathbb{R})$ of $\mathbb{R}$-points. By [Morgan and Shalen 1984, Prop. III.1.1, p. 458], $V_{a,b,c,d}(\mathbb{R})$ consists of characters corresponding to representations in either SU(2) or $\mathrm{SL}(2, \mathbb{R})$. Since the leading term of the cubic polynomial defining $V_{a,b,c,d}$ is xyz, its closure in $\mathbb{RP}^3$ meets the plane at infinity in three lines in general position (the singular cubic defined by $xyz = 0$ in $\mathbb{RP}^2$). Thus for any values of $(a, b, c, d) \in \mathbb{R}^4$, the surface $V_{a,b,c,d}(\mathbb{R})$ has four ends. (See Figure 4.)

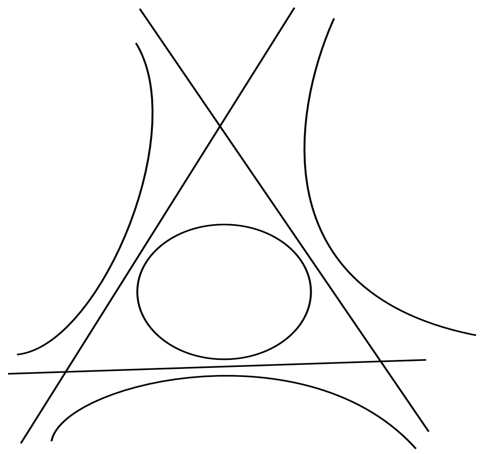


FIGURE 4. What $V_{a,b,c,d}(\mathbb{R})$ looks like near infinity.

Theorem 5.1. *Let $(a, b, c, d) \in \mathbb{R}^4$. Then $V_{a,b,c,d}(\mathbb{R})$ contains a singular point if and only if one of the following conditions holds*:

1. *One of a, b, c, d equals ± 2.*
2. *There is a reducible $\mathrm{SU}(2)$-representation having boundary traces (a, b, c, d), and $a, b, c, d \in [-2, 2]$.*
3. *There is a reducible $\mathrm{SL}(2, \mathbb{R})$-representation having boundary traces (a, b, c, d), and $a, b, c, d \in \mathbb{R} \setminus (-2, 2)$.*

The second condition is equivalent to

$$\Delta(a, b, c, d) = 0 \quad \textit{and} \quad a, b, c, d \in [-2, 2];$$

the third condition is equivalent to

$$\Delta(a, b, c, d) = 0 \quad \textit{and} \quad a, b, c, d \in \mathbb{R} \setminus (-2, 2).$$

We first verify the final statements. Suppose $x \in \mathbb{R}$ satisfies

$$\kappa_{a,b}(x) = \kappa_{c,d}(x) = 0$$

and $a, b, c, d \in [-2, 2]$. By Section 4D, we have $\Delta(a, b, c, d) = 0$. Conversely, given $\Delta(a, b, c, d) = 0$ with $a, b, c, d \in [-2, 2]$, there exists a reducible $\mathrm{SL}(2, \mathbb{C})$-representation ρ with the given boundary traces. Furthermore,

$$x = \mathrm{tr}(\rho(X)) \in [-2, 2]$$

by Lemma 2.2. By the results of Section 2B, the triples (a, b, x) and (c, d, x) correspond to $\mathrm{SU}(2)$-representations. Amalgamating the representations as in the proof of Theorem 4.1 produces a $\mathrm{SU}(2)$-representation ρ of $\pi_1(M)$. The equivalence of the third condition is proved similarly.

The proof of the rest of Theorem 5.1 is analogous to the proof of Theorem 4.1. The only difference is that we must use Lemma 4.4 instead of Lemma 4.3. Thus, it suffices to show:

Lemma 5.2. *Let $x, a, b, c, d \in \mathbb{R}$ and define $A_1(x)$, $B_1(x)$, $C_1(x)$ by (4–6). Then $\delta(x) \geq 0$, where $\delta(x)$ is defined in (4–5).*

Proof. By (4–6), we have

$$\begin{aligned} A_1'(x) &= 2x, \\ B_1'(x) &= -2(ac + bd), \\ C_1'(x) &= -8x + 4(ab + cd), \end{aligned}$$

and

$$\begin{aligned} \delta(x) &= B_1'(x)^2 - 4A_1'(x)C_1'(x) \\ &= 4(ac + bd)^2 + 64x^2 - 32x(ab + cd) \\ &= 4\left(4x - (ac + bd)\right)^2 \geq 0. \end{aligned}$$ □

5B. Topological Classification of the Smooth Real Varieties

Assume that $V_{a,b,c,d}(\mathbb{R})$ is nonsingular. We now classify the possible topologies for $V_{a,b,c,d}(\mathbb{R})$. The six topological types are depicted in Figure 1.

Theorem 5.3. *If $(a, b, c, d) \in \mathbb{R}^4$ and $V_{a,b,c,d}(\mathbb{R})$ is nonsingular, then $V_{a,b,c,d}(\mathbb{R})$ is diffeomorphic to one of the following*:

1. *a quadruply-punctured sphere, if $a, b, c, d \in \mathbb{R} \setminus [-2, 2]$ with $abcd < 0$;*
2. *the disjoint union of a triply-punctured torus and a disk, if $a, b, c, d \in \mathbb{R} \setminus [-2, 2]$ with $abcd > 0$;*
3. *the disjoint union of a triply-punctured sphere and a disk, if one of a, b, c, d is in $(-2, 2)$ and the rest are in $\mathbb{R} \setminus [-2, 2]$;*
4. *the disjoint union of an annulus and two disks, if two of a, b, c, d are in $(-2, 2)$ and two are in $\mathbb{R} \setminus [-2, 2]$;*
5. *the disjoint union of four disks, if three among a, b, c, d are in $(-2, 2)$ and one is in $\mathbb{R} \setminus [-2, 2]$;*
6. *the disjoint union of four disks and a sphere, if $a, b, c, d \in (-2, 2)$.*

In case 2, the component of $V_{a,b,c,d}$ homeomorphic to a disc is the “Teichmüller space” of M, consisting of hyperbolic structures with geodesic boundary on M. Compare [Goldman 1988a; 1988b].

5C. Reduction to a Plane Quartic Curve

To prove Theorem 5.3, reduce the dimension by taking discriminants, as in the determination of the singularities. Again, difficulties arise because the discriminant $f(x, y)$ is not monic in y. The set $V_{a,b,c,d}(\mathbb{R})$ contains two points for every $(x, y) \in \mathbb{R}^2$ with $f(x, y) > 0$, one point for $f(x, y) = 0$, and none for $f(x, y) < 0$. Define

$$\mathrm{S} = \{(x, y) \in \mathbb{R}^2 : f(x, y) > 0\}$$

and

$$\mathrm{C} = \{(x, y) \in \mathbb{R}^2 : f(x, y) = 0\} = \partial \mathrm{S}.$$

The projection

$$(x, y) : V_{a,b,c,d}(\mathbb{R}) \longrightarrow \mathbb{R}^2$$

expresses $V_{a,b,c,d}(\mathbb{R})$ as a union of two copies of S identified along their boundary C. This description as a “double branched cover” determines $V_{a,b,c,d}(\mathbb{R})$ up to homeomorphism (and hence diffeomorphism).

Fixing $x \in \mathbb{R}$ with $x \neq \pm 2$, for $|y|$ large, we have

$$f(x, y) \sim (x^2 - 4)y^2,$$

so $f(x,y)$ is positive if $|x|>2$ and negative if $|x|<2$. More precisely:

1. For any x with $|x|<2$, there exists $M>0$ such that for $|y|>M$, $(x,y)\notin \mathrm{S}$.
2. For any x with $|x|>2$, there exists $M>0$ such that for $|y|>M$, $(x,y)\in \mathrm{S}$.

To study C, we reduce dimensions once more via discriminants. For fixed $x\neq\pm 2$, the tangent of C at (x,y) for some $y\in\mathbb{R}$ is vertical if and only if the discriminant

$$g(x)=16\kappa_{a,b}(x)\kappa_{c,d}(x)=0.$$

Lemma 5.4. *Let $x\in\mathbb{R}\setminus\{\pm 2\}$. Then the set of $y\in\mathbb{R}$ such that $(x,y)\in \mathrm{C}$ consists of*:

1. *exactly two points if $g(x)>0$;*
2. *exactly one point if $g(x)=0$;*
3. *no points if $g(x)<0$.*

(Recall the definition of $A_1(x), B_1(x), C_1(x)$ from (4–6).)

Lemma 5.5. *Suppose $x=\pm 2$ and $g(x)\neq 0$. Then $B_1(x)\neq 0$, and there exists exactly one $y\in\mathbb{R}$ such that $(x,y)\in \mathrm{C}$.*

Lemma 5.6. *Suppose that $x=\pm 2$ and $g(x)=0$. Then $B_1(x)=0$, and the set of $y\in\mathbb{R}$ such that $(x,y)\in \mathrm{C}$ is empty, if $C_1(x)\neq 0$ or $\mathbb{R}$, if $C_1(x)=0$.*

Proof of the Lemmas. We write $f(x,y)$ as

$$f(x,y)=A_1(x)y^2+B_1(x)y+C_1(x).$$

Lemma 5.4 follows because $f(x,y)$ is a quadratic polynomial in y with discriminant $g(x)$. Lemma 5.5 follows because $A_1(x)=0$ and

$$g(x)=B_1(x)^2-4A_1(x)C_1(x).$$

Finally, Lemma 5.6 follows because its hypotheses imply $f(x,y)=C_1(x)$. $\square$

Lemma 5.7. *Suppose that $a,b,c,d\in\mathbb{R}\setminus\{2,-2\}$.*

1. *If $B_1(2)=C_1(2)=0$, then $a=b$ and $c=d$.*
2. *If $B_1(-2)=C_1(-2)=0$, then $a=-b$ and $c=-d$.*

Furthermore, in either of these cases, $V_{a,b,c,d}(\mathbb{R})$ is singular if and only if a, b, c, and d are either all in $(-2,2)$ or all in $\mathbb{R}\setminus[-2,2]$.

Proof. $B_1(\pm 2)=0$ factors as $(b\mp a)(c\mp d)=0$, so without loss, $a=\pm b$. Substituting, $C_1(\pm 2)=0$ becomes $(b^2-4)(c\mp d)^2=0$. Since $b\neq\pm 2$, we have $c=\pm d$. One calculates $\Delta(\pm b,b,\pm d,d)=0$. Now apply Theorem 5.1. $\square$

Lemma 5.8. *Suppose $V_{a,b,c,d}(\mathbb{R})$ contains no singular points but $g(x)$ has a real double root. Then this root is ± 2, and $a=\pm b$ and $c=\pm d$. One of these two possibilities must occur*:

1. *$a\in(-2,2)$ and $c\in\mathbb{R}\setminus[-2,2]$;*
2. *$c\in(-2,2)$ and $a\in\mathbb{R}\setminus[-2,2]$.*

Furthermore C *contains the line $x=\pm 2$.*

Proof. By Theorem 5.1, this situation can only arise if g has a double root and yet some of a,b,c,d are in $(-2,2)$ and some are in $\mathbb{R}\setminus[-2,2]$. Note that for $\kappa_{a,b}(x)$ to have a double root, one of a or b must be ± 2, but this does not happen since $V_{a,b,c,d}(\mathbb{R})$ is nonsingular. Thus, $\kappa_{a,b}(x)$ and $\kappa_{c,d}(x)$ have a common root. If this root lies in $(-2,2)$, then Lemma 2.2 implies that all of a,b,c,d lie in $(-2,2)$. Similarly, if this root lies in $\mathbb{R}\setminus[-2,2]$, then all of a,b,c,d lie in $\mathbb{R}\setminus[-2,2]$. In either case, Theorem 5.1 implies that $V_{a,b,c,d}(\mathbb{R})$ is singular. Therefore, the root is ± 2, implying $a=\pm b$ and $c=\pm d$. In addition, $B_1(\pm 2)=C_1(\pm 2)=0$, and so Lemma 5.6 implies that C contains the line $x=\pm 2$. $\square$

We continue the analysis of C. The solution of the quadratic equation $f(x,y)=0$ for y when $x\neq\pm 2$ is $y=\varphi_{\pm}(x)$ where

$$\varphi_{\pm}(x)=\frac{-B_1(x)\pm\sqrt{g(x)}}{2(x^2-4)}.$$

Note that $B_1(x)$ is linear in x and $g(x)\sim 16x^4$; so as $|x|\to\infty$,

$$\varphi_{\pm}(x)\to\pm 2.$$

Similarly, fix $y\neq\pm 2$ and define $x=\psi_{\pm}(y)$ by solving $f(x,y)=0$. Then, as $|y|\to\infty$,

$$\psi_{\pm}(y)\to\pm 2.$$

Thus, for any $\varepsilon>0$, there is $M>2+\varepsilon$ for which

$$|x|>M\implies\bigl(|\varphi_+(x)-2|<\varepsilon \text{ and } |\varphi_-(x)+2|<\varepsilon,\bigr)$$

and for which

$$|y|>M\implies\bigl(|\psi_+(y)-2|<\varepsilon \text{ and } |\psi_-(y)+2|<\varepsilon.\bigr)$$

Fix $0<\varepsilon<1$, and let Q be the square $(-M,M)^2$. Then $\mathrm{C}\cap(\mathbb{R}^2\setminus Q)$ is homeomorphic to a disjoint union of eight copies of $[M,\infty)$. For example, one of the eight homeomorphisms would map $t\in[M,\infty)$ to $(-t,\varphi_+(-t))$. Note that C intersects ∂Q at the images of M under the eight homeomorphisms. C is smooth and hence a 1-manifold, so $\overline{Q} \cap \mathrm{C}$ is a compact 1-manifold with boundary $\partial\overline{Q} \cap \mathrm{C}$ consisting of eight points.

Lemma 5.9. *Assume $V_{a,b,c,d}(\mathbb{R})$ is nonsingular. With Q as above, $\overline{Q} \cap \mathrm{C}$ has finitely many connected components. Each component is homeomorphic to either a closed interval $[-1, 1]$, if at least one (and hence exactly two) points of the component lie in ∂Q, or S^1 otherwise. Furthermore, there must be exactly four components homeomorphic to an interval.*

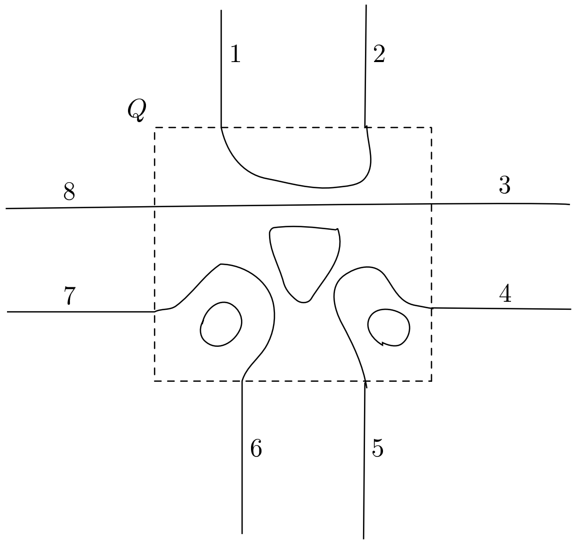


FIGURE 5. How the curve C meets the square Q.

C must look something like Figure 5. Outside Q, all points of C are close to the eight asymptotes. By choosing Q sufficiently large, C will be transverse to the four sides of ∂Q and $\mathrm{C} \cap \partial Q$ is a set of eight points,

$$P = \{p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8\},$$

numbered as in Figure 5. To determine the topologies of C and S, it suffices to decide how these eight points of $\mathrm{C} \cap \partial Q$ are joined pairwise inside Q, and to find the number and positions of any bounded components.

In each of the following lemmas, the set of $\mathbb{R}$-points is assumed nonsingular.

Lemma 5.10. *If p_n is joined by C to p_m, then $n + m$ is odd.*

Proof. Suppose $n + m$ is even. Then on ∂Q, the points p_n and p_m separate the remaining points p_i into two sets P_1 and P_2 of odd cardinality. Since an arc α of $\mathrm{C} \cap Q$ joins p_n and p_m, this arc separates P_1 from P_2 in $\bar{Q}$. Furthermore the points of

$$P - \{p_m, p_n\} = P_1 \coprod P_2$$

are pairwise joined by arcs of $\mathrm{C} \cap Q$ which (since C is smooth) do not meet α. Thus the points of P_1 join to points of P_1, and similarly for P_2. This partitions P_1 into disjoint pairs, contradicting the odd cardinality of P_1. □

Lemma 5.11. 1. *p_3 joins to p_4 if and only if $g(x)$ has a root in $(2, \infty)$.*
2. *p_7 joins to p_8 if and only if $g(x)$ has a root in $(-\infty, -2)$.*

Proof. We'll show the first statement; the second is similar. Suppose first that p_3 joins to p_4. The component joining p_3 and p_4 must have a vertical tangent — corresponding to a root of $g(x)$ — at some point x_0. If $x_0 < 2$, then the component crosses the line $x = 2$ twice, contradicting Lemma 5.5. (If the component crossed the line at the same point each time, there would be a singularity. Alternatively, if $g(2) = 0$, then by Lemma 5.6, C must contain the whole line $x = 2$, and we would have singularities where the component crosses.) If $x_0 = 2$, then, by Lemma 5.6, either C doesn't intersect the line $x = 2$, or else C contains it. But a vertical tangent at $x = 2$ would force at least one point in C on that line; and if C contained the line, a component from the right touching the line at a point would result in a singularity. Thus $x_0 > 2$. Since C has a vertical tangent there, $g(x_0) = 0$.

Conversely, let $x_0 > 2$ be the root in question. Without loss, x_0 is the largest root of g. Thus $g(s) < 0$ for some $s \in (2, x)$, since $g(x) > 0$ for $x > x_0$, and x_0 is not a double root. If p_3 were to join to p_1, p_6, p_7, or p_8, then clearly C would cross the line $x = s$, contradicting Lemma 5.4. It also can't join to p_2 or p_5, since in that case, C would cross $x = s$ somewhere along the approach to the asymptote at $x = 2$. Thus p_3 must join to p_4. □

Lemma 5.12. 1. *p_1 joins to p_6 if and only if g has two roots in $(-\infty, -2]$ and two roots in $[-2, 2]$.*
2. *p_2 joins to p_5 if and only if g has two roots in $[2, \infty)$ and two roots in $[-2, 2]$.*

Proof. We'll show the first statement; the second is similar. To prove the forward direction, suppose p_1 and p_6 are joined by a component. If this component is the line $x = -2$, then Lemmas 5.6 and 5.7 imply that $a = -b$ and $c = -d$. Thus g has a double root at -2. By Lemma 5.8, we can assume without loss that $a, b \in (-2, 2)$ and $c, d \in \mathbb{R} \setminus [-2, 2]$. Therefore, by Lemma 2.2, g has a root in $(-\infty, -2]$, a double root at -2, and a root in $(-2, 2]$, and we are done.

Now assume C does not contain the line $x = -2$. Consider the whole component of C that joins p_1 to p_6, and not just the part lying inside Q; let α be a diffeomorphism mapping $\mathbb{R}$ onto this component. If -2 is a root of g, Lemma 5.6 implies that C does not intersect the line $x = -2$. If the x-value of $\alpha(t)$ increases from -2 as t increases from $-\infty$, then at some point it must decrease back to -2 as t increases to $+\infty$. Thus, there must be a vertical tangent — and hence a root of g — at some x-value greater than -2. This root must be less than 2, for the same reasons as in Lemma 5.11. Meanwhile, p_7 and p_8 must join to each other, since they have nowhere else to go. Lemma 5.11 implies that g has a root in $(-\infty, -2)$. We now know that g has at least three real roots; it therefore has a fourth. By Lemma 2.2, this fourth root is somewhere in $(-\infty, 2]$, and we have the desired conclusion. The argument is similar if the x-value of α decreases and then increases.

On the other hand, if -2 is not a root of g, Lemma 5.4 implies that C must cross the line $x = -2$ at exactly one point. If any component other than the one in question were to cross this line, then it must cross it at least twice (once to get across, and once to get back) or at a singularity or vertical tangent. Thus, the component in question must cross $x = -2$ exactly once. The x-value of α must increase from -2, decrease back down to -2 and below, and then increase back up to -2; as a result, we get at least two vertical tangents, one on each side of -2. By the same argument as above, both roots must be less than 2. Lemma 2.2 implies g must also have two other roots, one on each side of -2, and neither larger than 2.

For the converse, suppose that -2 is a simple root of g, and $x_1 \in (-\infty, -2)$ and $x_2, x_3 \in (-2, 2]$ are also roots. Lemma 5.11 implies that p_7 and p_8 join. Furthermore, g is negative on (x_2, x_3), and so there are no points of C above that interval. Thus p_1 and p_6 are cut off from the other points, and must join to each other. The argument is similar if -2 is a root, with two roots less than -2 and one greater, or if there are two roots less than -2 and two greater. Finally, if -2 is a multiple root of g, then Lemma 5.8 implies $a = -b$ and $c = -d$. As a result, $B_1(-2) = C_1(-2) = 0$. Lemma 5.6 implies that C contains the line $x = -2$, and so p_1 and p_6 join. □

(Note that in the above lemma, the case of a double root at ± 2 is considered as a deformation of the case of one root just above and one just below.)

Lemma 5.13. *If* C *has a compact component, then g has four distinct real roots $x_1 < x_2 < x_3 < x_4$, and the projection of the compact component onto the x-axis is $[x_2, x_3]$.*

Proof. The projection of the component onto the x-axis must be $[x_i, x_j]$, where x_i and x_j are distinct roots of g. Furthermore, g must be strictly positive on (x_i, x_j), and negative for some $x < x_i$ and for some $x > x_j$. Thus, g must have four real roots, and x_i and x_j must be the middle two. □

Corollary 5.14. *There is at most one compact component of* C.

Lemma 5.15. *If g has four real roots $x_1 < x_2 < x_3 < x_4$ and $2, -2 \notin [x_2, x_3]$, then* C *has a compact component projecting onto $[x_2, x_3]$.*

Proof. Since the leading coefficient of f is nonzero throughout $[x_2, x_3]$, and the discriminant g is zero at x_2 and x_3 and positive in between, the part of C that lies above the $[x_2, x_3]$ must consist of two line segments glued at the endpoints. As g is negative just below x_2 and just above x_3, it follows C has a compact component. □

5D. The Six Cases

We now analyze individual cases to find the resulting topologies.

In the first case, we are given $a, b, c, d \in \mathbb{R} \setminus [-2, 2]$ and $abcd < 0$. By applying Lemma 3.1, we can assume that $a, b, c > 2$ and $d < -2$. Thus, by Lemma 2.2, g has two roots in $[2, +\infty)$ and two in $(-\infty, 2]$. Note that by Lemma 5.8, the roots of g are all distinct; in particular, there must be at least one root greater than 2. Now by Lemma 5.11, p_3 and p_4 join, and p_7 and p_8 join. p_1 and p_5 do not join, by Lemma 5.10. If p_1 and p_6 were to join, then by Lemma 5.12, all four roots would be in $(-\infty, 2]$, contradicting the existence of a root greater than 2. Thus, p_1 and p_2 join; finally, p_5 and p_6 must join. Furthermore, there cannot be a compact component, since by Lemma 5.13 such a component would have to project onto $[x_2, x_3]$, with $x_2 \leq -2$ and $x_3 \geq 2$; but such a component would cross the lines $x = \pm 2$ twice, contradicting Lemma 5.5 and Lemma 5.6. Thus, $\mathrm{C} \cap \overline{Q}$ consists of just the four components mentioned above.

Since f is negative for $x = 0$ and y large, we see that S is homeomorphic to a closed disk minus four points of the boundary. (See Figure 6.) Doubling along the boundary, we see that $V_{a,b,c,d}(\mathbb{R})$ is homeomorphic, and hence diffeomorphic, to a quadruply-punctured sphere. (See Case 1 in Figure 1.)

In the second case, all $a, b, c, d \in \mathbb{R} \setminus [-2, 2]$ and $abcd > 0$. Applying Lemma 3.1, we may assume that $a, b, c, d > 2$. Lemma 2.2 implies that $g(x)$ has four roots in $[2, +\infty)$. Lemma 5.11 implies that p_3 and p_4 must join, and p_7 and p_8 cannot join. Lemma 5.12 implies that p_1 and p_6 cannot join, nor can p_2 and p_5. Thus p_1 joins either p_2 or p_8, by Lemma 5.10. If p_1 joins p_2, then p_7 must join p_6 and so p_8 joins p_5; or, if p_1 joins p_8, then p_5 joins p_6 and p_2 joins p_7. Meanwhile, Lemma 5.15 implies that C has a compact component somewhere between the component joining p_3 to p_4 and the line $x = 2$. Thus C consists of a circle, together with one of the two possible sets of four noncompact components; both setups result in the same topology for S.

Since f is negative for $x = 0$ and y large, S is homeomorphic to a closed annulus minus three points of the outer boundary, along with a closed disk minus one point of the boundary. (See Figure 7.) Doubling again, $V_{a,b,c,d}(\mathbb{R})$ is homeomorphic, and hence diffeomorphic, to a triply-punctured torus and a disk. (See Case 2 in Figure 1.)

In the third case, three of $a, b, c, d \in \mathbb{R} \setminus [-2, 2]$ and the fourth in $(-2, 2)$. By applying Lemma 3.1, we can assume that $d \in (-2, 2)$ and $a, b, c > 2$. By Lemma 2.2, g has two roots in $[2, +\infty)$ and no other real roots. Thus, by Lemma 5.13, there cannot be a compact component. Furthermore, Lemma 5.11 and Lemma 5.12 imply that p_3 and p_4 must join, but p_7 and p_8 cannot join, p_1 and p_6 cannot join, and p_2 and p_5 cannot join. Thus, we end up with the same arrangement of noncompact components as in Case 2, but no compact component.

In the fourth case, two of $a, b, c, d \in \mathbb{R} \setminus [-2, 2]$ and two in $(-2, 2)$. Applying Lemma 3.1, we can assume that $a, b \in (-2, 2)$ and $c, d > 2$. There are two possibilities.

Suppose first that $g(x)$ has a double root. Lemma 5.8 implies that this double root must be 2; there is also one root in $[-2, 2)$ and one in $(2, \infty)$. By Lemma 5.13, there is no compact component and

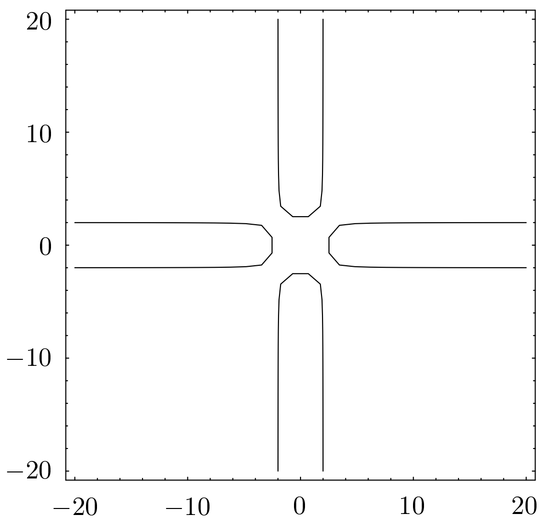


FIGURE 6. Case 1, with $(a, b, c, d) = (2.1, 2.1, 2.1, -2.1)$.

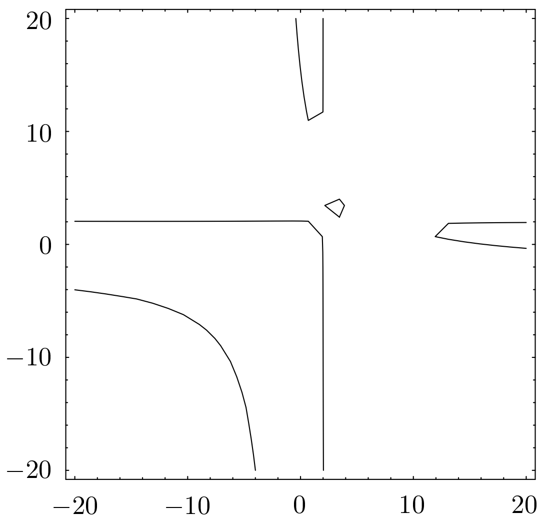


FIGURE 7. Case 2, with $(a, b, c, d) = (2.5, 2.5, 3, 4)$.

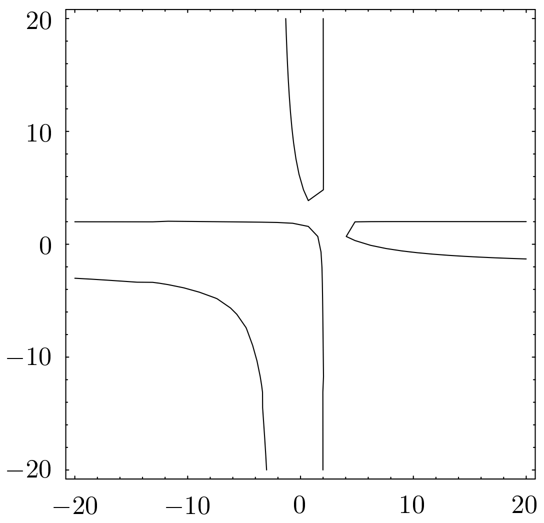


FIGURE 8. Case 3, with $(a, b, c, d) = (2.4, 2.2, 2.3, 1)$.

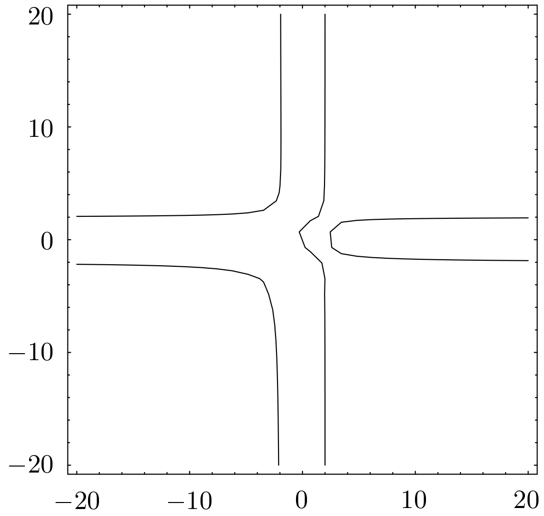


FIGURE 9. Case 4, with $(a, b, c, d) = (2.1, 2.1, 1.7, -1.3)$.

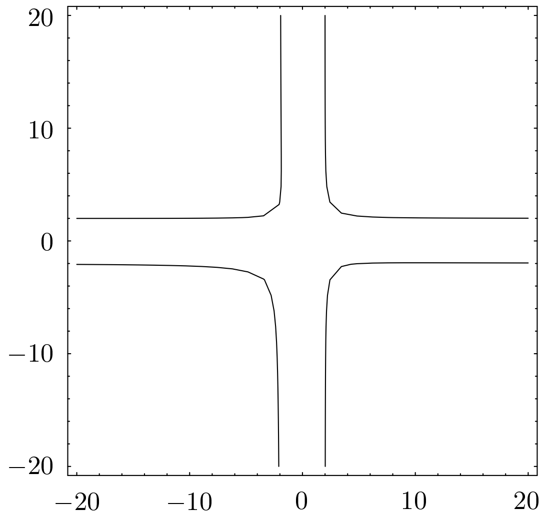


FIGURE 10. Case 5, with $(a, b, c, d) = (2.1, 0.2, 0.3, 0.4)$.

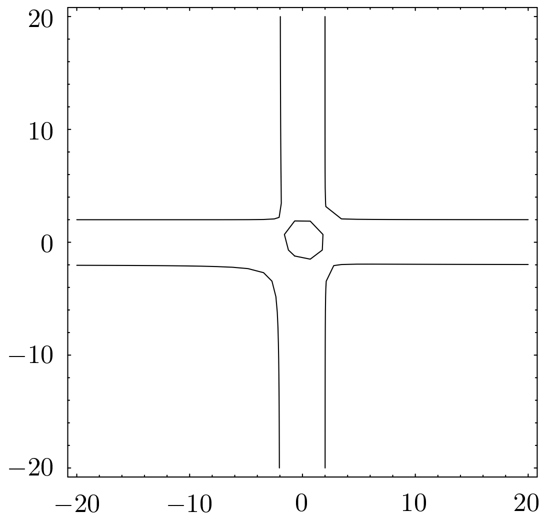


FIGURE 11. Case 6, with $(a, b, c, d) = (1, 0.2, 0.3, 0.4)$.

Lemma 5.8 implies that p_2 and p_5 join by a vertical line. Lemma 5.11 implies that p_3 and p_4 join, but p_7 and p_8 do not. Thus p_1 and p_8 must join, as must p_6 and p_7.

The second possibility is that all of the roots of g are distinct; thus, two distinct roots lie in $[-2, 2]$ and two distinct roots lie in $[2, +\infty)$, and 2 is not a double root. Once again, Lemma 5.13 precludes a compact component, and Lemma 5.11 implies that p_3 and p_4 join. In addition, Lemma 5.12 implies that p_2 and p_5 join, while p_1 and p_6 do not. Thus, p_1 and p_8 join, as do p_6 and p_7.

Both possibilities give the same result, namely, S is homeomorphic to a closed disk minus two points of the boundary, along with two closed disks missing one boundary point each. (See Figure 9.) Doubling again, $V_{a,b,c,d}(\mathbb{R})$ is homeomorphic, and hence diffeomorphic, to a disjoint union of an annulus and two disks. (See Case 4 in Figure 1.)

In the fifth case, one of a, b, c, d lie in $\mathbb{R} \setminus [-2, 2]$ and the other three in $(-2, 2)$. Applying Lemma 3.1, we can assume that $a, b, c \in (-2, 2)$ and $d > 2$. Lemma 2.2 implies g has two roots in $[-2, 2]$ and no other real roots. Lemma 5.13 precludes a compact component. Furthermore, Lemmas 5.11 and 5.12 imply that p_3 can't join to p_4, p_7 can't join to p_8, p_1 can't join to p_6, and p_2 can't join to p_5. Since g is negative for some $x_0 \in (-2, 2)$, none of the points on the left can join to any of the points on the right, by Lemma 5.4. Therefore p_1 joins to p_8, p_2 joins to p_3, p_4 joins to p_5, and p_6 joins to p_7.

From the sign of f, we see that S is homeomorphic to a union of four disjoint closed disks missing one boundary point each. (See Figure 10.) Doubling, we see that $V_{a,b,c,d}(\mathbb{R})$ is homeomorphic, and hence diffeomorphic, to a union of four disjoint disks. (See Case 5 in Figure 1.)

In the sixth case, all of a, b, c, d lie in $(-2, 2)$. Lemma 2.2 implies that g has four roots in $[-2, 2]$. Lemma 5.15 implies that C has a compact component between the lines $x = -2$ and $x = 2$. The rest of the analysis is exactly as in Case 5; the result is that p_1 joins to p_8, p_2 joins to p_3, p_4 joins to p_5, and p_6 joins to p_7.

From the sign of f, we see that S is homeomorphic to a closed disk, along with four closed disks missing one boundary point each. (See Figure 11.) Doubling, we see that $V_{a,b,c,d}(\mathbb{R})$ is homeomorphic, and hence diffeomorphic, to four disks and a sphere. (See Case 6 in Figure 1.)

This concludes our case-by-case analysis and the proof of Theorem 5.3.

6. A COMPACT COMPONENT OF SL(2, ℝ)-REPRESENTATIONS

When $(a, b, c, d) \in \mathbb{R}$ and $V_{a,b,c,d}(\mathbb{R})$ is smooth, then each connected component of $V_{a,b,c,d}(\mathbb{R})$ consists of characters of representations in either $\mathrm{SL}(2, \mathbb{R})$ or $\mathrm{SU}(2)$. Components of spaces of representations in $\mathrm{SU}(2)$ are compact. Conversely, when M is a once-punctured torus, the only compact components of $\mathrm{X}(M)(\mathbb{R})(t)$ are characters of $\mathrm{SU}(2)$-representations. (See [Goldman 1988b] for a description of the topology of the relative character varieties for a punctured torus.) However, for the quadruply punctured sphere, there are compact components of $V_{a,b,c,d}(\mathbb{R})$ for certain values of (a, b, c, d) which do not correspond to $\mathrm{SU}(2)$-representations.

If $V_{a,b,c,d}(\mathbb{R})$ contains a compact component, Theorem 5.3 implies that

$$-2 < a, b, c, d < 2,$$

and conversely, every such $(a, b, c, d) \in (-2, 2)^4$ for which $V_{a,b,c,d}(\mathbb{R})$ is smooth contains a compact component. (Compare the analysis of Case 6.) Here is a criterion for when this compact component consists of $\mathrm{SL}(2, \mathbb{R})$-characters or $\mathrm{SU}(2)$-characters, in terms of the function $\Delta(a, b, c, d)$ defined in Section 4D.

Proposition 6.1. *Suppose $a, b, c, d \in (-2, 2)$ and $V_{a,b,c,d}$ is smooth. Then the compact component of $V_{a,b,c,d}(\mathbb{R})$ consists of characters of $\mathrm{SL}(2, \mathbb{R})$-representations if and only if*

$$\left.\begin{aligned} \Delta(a, b, c, d) > 0, \\ 16 - abcd - 2(a^2 + b^2 + c^2 + d^2) > 0. \end{aligned}\right\} \tag{6–1}$$

Otherwise, each element in the component is the character of an $\mathrm{SU}(2)$-representation.

As pointed out by a referee, $\mathrm{SU}(2)$-representations in $V_{a,b,c,d}(\mathbb{R})$ correspond to quadrilaterals in S^3 whose side lengths are determined by the parameters a, b, c, d. (Compare [Kirk and Klassen 1991].) Necessary and sufficient conditions for the existence of such polygons are triangle inequalities in $\mathrm{SU}(2)$, as discussed in [Galitzer 1997; Kapovich and Millson 1995]. Thus conditions (6–1) are just the triangle inequalities in disguise.

Proof. Decompose the quadruply-punctured sphere M along ξ, as in Section 3A, as a union of two triply-punctured spheres. A given representation $\pi_1(M) \longrightarrow \mathrm{SL}(2, \mathbb{C})$ has image in $\mathrm{SU}(2)$ only if its restriction to the fundamental group of each component of $M|\xi$ has image in $\mathrm{SU}(2)$. Furthermore there is an $\mathrm{SU}(2)$ character on M with boundary traces (a, b, c, d) if and only if there is an x such that (a, b, x) and (c, d, x) are $\mathrm{SU}(2)$ characters on triply-punctured spheres. To construct this character, amalgamate the two representations along the common boundary X as in Section 3A; since two elements of $\mathrm{SU}(2)$ have the same trace if and only if they are conjugate, we can ensure that the two representations agree on this boundary.

Recall from Section 2B that (a, b, x) is an $\mathrm{SU}(2)$-character for the triply-punctured sphere if and only if $a, b, x \in [-2, 2]$ and $\kappa_{a,b}(x) \leq 0$. For $a, b \in \mathbb{R}$, define the interval $I_{a,b} \subset \mathbb{R}$ by:

$$I_{a,b} = \{x \in \mathbb{R} \mid \kappa_{a,b}(x) \leq 0\} = [i^-_{a,b}, i^+_{a,b}],$$

where

$$i^-_{a,b} = \frac{ab - \sqrt{(a^2-4)(b^2-4)}}{2},$$

$$i^+_{a,b} = \frac{ab + \sqrt{(a^2-4)(b^2-4)}}{2}.$$

Since $\kappa_{a,b}(\pm 2) \geq 0$, we have

$$I_{a,b} \subset [-2, 2].$$

For given (a, b, c, d) with $-2 < a, b, c, d < 2$, the values of $x \in \mathbb{R}$ such that (a, b, c, d, x) extends to a character of an $\mathrm{SU}(2)$-representation comprise the intersection

$$I_{a,b} \cap I_{c,d} \subset [-2, 2].$$

Furthermore the possible $(y, z) \in \mathbb{R}^2$ comprise a conic, which by (3–3), is an ellipse. Thus characters of $\mathrm{SU}(2)$-representations with boundary traces (a, b, c, d) form a pencil of ellipses over the interval $I_{a,b} \cap I_{c,d}$.

However, if $I_{a,b} \cap I_{c,d} = \varnothing$, there will be a compact component of $V_{a,b,c,d}(\mathbb{R})$ not corresponding to $\mathrm{SU}(2)$-representations. By applying a permutation if necessary, assume that

$$i^+_{a,b} < i^-_{c,d}.$$

Then a component of $V_{a,b,c,d}(\mathbb{R})$ will map under

$$x : V_{a,b,c,d}(\mathbb{R}) \longrightarrow \mathbb{R}$$

to the interval $[i^+_{a,b}, i^-_{c,d}]$. By (3–3), the restriction of x to $x^{-1}([-2,2])$ is proper. If $x_0 \in I_{a,b} \cup I_{c,d}$, then

$$\frac{\kappa_{a,b}(x_0)\kappa_{c,d}(x_0)}{4-x_0^2} < 0$$

and by (3–3), the preimage $x^{-1}(x_0)$ is empty. Thus

$$x^{-1}([i^+_{a,b}, i^-_{c,d}])$$

is a compact component of $V_{a,b,c,d}(\mathbb{R})$ not containing characters of $\mathrm{SU}(2)$-representations. Since real characters are characters of either $\mathrm{SU}(2)$-representations or $\mathrm{SL}(2,\mathbb{R})$-representations, this compact component must consist of $\mathrm{SL}(2,\mathbb{R})$-representations.

Thus, the compact component of $V_{a,b,c,d}(\mathbb{R})$ consists of $\mathrm{SU}(2)$-representations if and only if

$$I_{a,b} \cap I_{c,d} \neq \varnothing$$

and otherwise consists of $\mathrm{SL}(2,\mathbb{R})$-representations. The latter case occurs when

$$i^+_{a,b} < i^-_{c,d}$$

or

$$i^+_{c,d} < i^-_{a,b}.$$

These conditions are equivalent to conditions (6–1) by the following argument:

The condition $i^+_{a,b} < i^-_{c,d}$ is equivalent to

$$\sqrt{(a^2-4)(b^2-4)} + \sqrt{(c^2-4)(d^2-4)} < cd - ab,$$

and $i^+_{c,d} < i^-_{a,b}$ is equivalent to

$$\sqrt{(a^2-4)(b^2-4)} + \sqrt{(c^2-4)(d^2-4)} < ab - cd.$$

The following lemma, whose straightforward proof is omitted, then applies:

Lemma 6.2. *Let p, q, r be nonnegative real numbers. Then*

$$\sqrt{p} + \sqrt{q} < r$$

if and only if

$$\begin{aligned} r^4 - 2r^2(p+q) + (p-q)^2 &> 0, \\ r^2 - p - q &> 0. \end{aligned}$$

Taking

$$\begin{aligned} p &= (a^2-4)(b^2-4), \\ q &= (c^2-4)(d^2-4), \\ r &= |ab - cd|, \end{aligned}$$

and calculating that

$$\begin{aligned} r^4 - 2r^2(p+q) + (p-q)^2 &= 16\Delta(a,b,c,d), \\ r^2 - p - q &= 2(16 - abcd - 2(a^2+b^2+c^2+d^2)), \end{aligned}$$

the compact component corresponds to $\mathrm{SL}(2,\mathbb{R})$-representations if and only if conditions (6–1) hold. □

Here is an explicit example. Let $a, b, c = \frac{3}{2}$ and $d = -\frac{3}{2}$. For example,

$$\alpha = \begin{bmatrix} 1 & 1 \\ -\frac12 & \frac12 \end{bmatrix}, \quad \beta = \begin{bmatrix} 1 & \frac12 \\ -1 & \frac12 \end{bmatrix}, \quad \gamma = \begin{bmatrix} 1 & 1 \\ -\frac12 & \frac12 \end{bmatrix}$$

generates an $\mathrm{SL}(2,\mathbb{R})$-representation with the desired character. Furthermore, its (x,y,z) traces are $(0,0,\frac14)$. Now the equation for the relative character variety at these boundary traces is

$$x^2 + y^2 + z^2 + xyz - \tfrac{1}{16} = 0.$$

Its locus does not intersect the planes $x = \pm 2$, $y = \pm 2$, or $z = \pm 2$ at real points.

The noncompact components are discs and correspond to hyperbolic structures on S^2 with singular points with prescribed cone angles (determined by a, b, c, d). These hyperbolic structures have "fundamental polygons" which are embedded quadrilaterals. Elements of the compact component correspond to nonembedded quadrilaterals, such as the one depicted in Figure 12. The compactness of the configuration space of such polygons reflects geometric bounds on these polygons with fixed vertex angles.

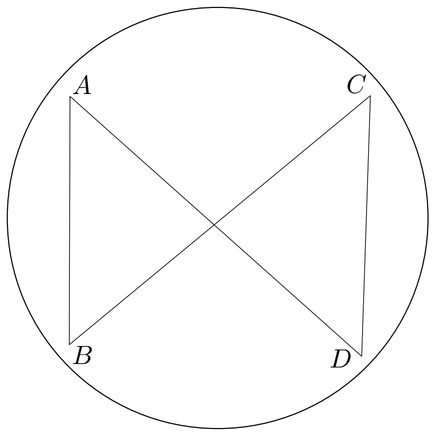


FIGURE 12. A nonembedded quadrilateral in $\mathbf{H}^2$.

ACKNOWLEDGMENTS

This investigation began as an REU (Research Experiences for Undergraduates) project at the University of Maryland in the summers of 1992 and 1993, under Goldman's direction. During this project, Benedetto used Geomview [Phillips et al. 1993] and Mathematica on Silicon Graphics workstations to visualize the components of the real moduli spaces.

Many of the results in this paper were suggested by computer experiments before they were rigorously proved; compare [Benedetto 1993]. The illustrations in this paper were drawn using Geomview and Mathematica software on Silicon Graphics and NeXT workstations.

We wish to thank the referees for carefully reading the manuscript and suggesting several improvements.

REFERENCES

[Benedetto 1993] R. Benedetto, *Relative character varieties of representations of the fundamental group of the four-punctured sphere in* $\mathrm{SL}(2,\mathbb{R})$, Senior thesis, Harvard University, 1993.

[Fricke and Klein 1897] R. Fricke and F. Klein, *Vorlesungen über die theorie der automorphen Functionen,* vol. 1, Teubner, Leipzig, 1897. Reprinted by Johnson Reprint, New York, and Teubner, Stuttgart, 1965.

[Galitzer 1997] A. Galitzer, *The moduli space of polygon linkages in the* 2*-sphere*, doctoral dissertation, University of Maryland, 1997.

[Goldman 1988a] W. M. Goldman, "Geometric structures on manifolds and varieties of representations", pp. 169–198 in *Geometry of group representations* (Boulder, CO, 1987), edited by W. M. Goldman and A. R. Magid, Contemporary Mathematics **47**, Amer. Math. Soc., Providence, RI, 1988.

[Goldman 1988b] W. M. Goldman, "Topological components of spaces of representations", *Invent. Math.* **93**:3 (1988), 557–607.

[Goldman 1997] W. M. Goldman, "Ergodic theory on moduli spaces", *Ann. of Math.* (2) **146**:3 (1997), 475–507.

[Kapovich and Millson 1995] M. Kapovich and J. Millson, "On the moduli space of polygons in the Euclidean plane", *J. Differential Geom.* **42**:1 (1995), 133–164.

[Kirk and Klassen 1991] P. A. Kirk and E. P. Klassen, "Representation spaces of Seifert fibered homology spheres", *Topology* **30**:1 (1991), 77–95.

[Lubotzky and Magid 1985] A. Lubotzky and A. R. Magid, *Varieties of representations of finitely generated groups*, Mem. Amer. Math. Soc. **336**, Amer. Math. Soc., Providence, 1985.

[Magnus 1980] W. Magnus, "Rings of Fricke characters and automorphism groups of free groups", *Math. Z.* **170**:1 (1980), 91–103.

[Morgan and Shalen 1984] J. W. Morgan and P. B. Shalen, "Valuations, trees, and degenerations of hyperbolic structures, I", *Ann. of Math.* (2) **120**:3 (1984), 401–476.

[Phillips et al. 1993] M. Phillips, S. Levy, and T. Munzner, "Geomview: an interactive geometry viewer", *Notices Amer. Math. Soc.* **40** (October 1993), 985–988. See http://www.geom.umn.edu/locate/geomview.

[Procesi 1976] C. Procesi, "The invariant theory of $n\times n$ matrices", *Advances in Math.* **19**:3 (1976), 306–381.

[Raghunathan 1972] M. S. Raghunathan, *Discrete subgroups of Lie groups*, Ergebnisse der Mathematik und ihrer Grenzgebiete **68**, Springer, New York, 1972.

[Weil 1964] A. Weil, "Remarks on the cohomology of groups", *Ann. of Math.* (2) **80** (1964), 149–157.

Robert L. Benedetto, Mathematics Department, University of Rochester, Rochester, NY 14627, USA (benny@math.rochester.edu)

William M. Goldman, Mathematics Department, University of Maryland, College Park, MD 20742, USA (wmg@math.umd.edu)

Received July 31, 1997; accepted in revised form June 8, 1998