Zeitschrift: Commentarii Mathematici Helvetici

Herausgeber: Schweizerische Mathematische Gesellschaft

Band: 51 (1976)

Artikel: The Cut Locus of Noncompact Finitely Connected Surfaces Without Conjugate Points

Autor: Eberlein, Patrick

DOI: https://doi.org/10.5169/seals-39427

Comment. Math. Helvetici 51 (1976) 23–44 Birkhäuser Verlag, Basel

The Cut Locus of Noncompact Finitely Connected Surfaces Without Conjugate Points

by PATRICK EBERLEIN (University of North Carolina
Chapel Hill, North Carolina 27514, U.S.A.)

Introduction

In this paper we obtain a characterization and derive some implications of the condition that a complete two dimensional Riemannian manifold without conjugate points have finitely generated fundamental group. The characterization in terms of fundamental domains is a classical result in the case of Gaussian curvature $K\equiv -1$. See for example [7] and [11].

Let M denote an arbitrary complete surface without conjugate points along any geodesic, H the universal Riemannian covering of M and D the deckgroup of the covering. Given a point p in H we define the *canonical fundamental domain* for D with *center p* to be the set $R_p \subseteq H$ given by

$$R_p = \bigcap_{\substack{\phi\in D\\ \phi\neq 1}} E^+(p, \phi p)$$

where $E^+(p, \phi p) = \{q\in H: d(p,q)\leqslant d(\phi p, q)\}$. The set $E(p,\phi p) = \{q\in H: d(p,q) = d(\phi p, q)\}$ is a *bounding side* for R_p if R_p is a proper subset of $\bigcap_{\substack{\psi\in D\\ \psi\neq\phi,\, \psi\neq 1}} E^+(p, \psi p)$.

THEOREM A. *Let $M=H/D$ be a complete nonsimply connected surface without conjugate points. Then the following are equivalent.*

1) *$\pi_1(M)$ is finitely generated.*

2) *For some $p\in H$ the fundamental domain R_p has only a finite number of bounding sides.*

3) *For every $p\in H$ the fundamental domain R_p has only a finite number of bounding sides.*

As a corollary we obtain

This research was supported in part by NSF Grant GP-43246.

THEOREM B. *Let M be a complete nonsimply connected surface without conjugate points. For each $p \in M$ let $G_2(p)$ be the set of points q in M for which there are exactly two shortest geodesics from p to q. Then $G_2(p)$ is nonempty for each p, and each connected component of $G_2(p)$ is an open differentiable arc. Moreover the following statements are equivalent.*

1) *$\pi_1(M)$ is finitely generated.*
2) *For some $p \in M$, $G_2(p)$ has a finite number of connected components.*
3) *For every $p \in M$, $G_2(p)$ has a finite number of connected components.*

I am grateful to the referee for pointing out that the attempt to generalize theorems A and B to arbitrary dimensions fails in dimension 3 by theorem 1 of [8]. See also page 410 of [12].

A further consequence of theorem A is

THEOREM C. *Let M be a complete nonsimply connected surface without conjugate points and with finitely generated fundamental group. Then for each point p in M there are at most a finite number of points q in M for which there exist three or more shortest geodesics joining p to q.*

We do not know if the converse to theorem C is true. We remark that there exists at least one shortest geodesic joining any two distinct points p, q of M since M is complete. If q lies in the cutlocus of p, then there are at least two but at most a finite number of shortest connecting geodesics since M has no conjugate points.

In each of the theorems A, B and C it suffices to consider the case that $M = H/D$ is noncompact. If M is compact then $\pi_1(M)$ is finitely generated and each fundamental domain $R_p \subseteq H$ has only a finite number of bounding sides. A proof of the second assertion is contained in the discussion in section two. The first assertion follows from the second in view of the proof of the statement 2) $\to$1) in theorem A.

The paper is organized as follows. Section 1 contains basic definitions and notation. For convenience we assume that all manifolds M and Riemannian metrics g are C^∞. Section 2 contains the statements of basic properties of the fundamental domains R_p. The proofs of these statements form the hardest part of the paper and because of their length are found in the appendix, section 4. In section 3 we prove theorems A, B and C. Theorems B and C follow quickly from theorem A and the facts from section two. The proof of theorem A is reminiscent of the method used by Marden to prove theorem 2 of [11]. In fact, theorem A can also be derived from that result in the orientable case. See the remark at the end of section 3.

§1. Preliminaries

In this section we establish notation and list some basic facts. M will always denote a complete connected Riemannian manifold with Riemannian structure $\langle\ ,\ \rangle$, Riemannian metric $d(\ ,\ )$ and sectional curvature K. Let TM denote the tangent bundle of M and $T_p(M)$ the tangent space to M at p. If $v \in TM$ is given let $\gamma_v : \mathbf{R} \to M$ be the geodesic such that $\gamma_v'(0)=v$. The map $\exp_p : T_p(M) \to M$ given by $\exp_p(v)=\gamma_v(1)$ is the exponential map at p. In this paper all geodesics are assumed to have unit speed and to be defined on the entire real line unless otherwise indicated. A *geodesic segment* is a geodesic defined on a compact interval $[a, b]$. A *geodesic ray* is a geodesic defined on $[0, \infty]$.

A manifold M is said to have no *conjugate points* if there exists no nontrivial Jacobi vector field that vanishes twice on some geodesic γ of M. If M is simply connected, then there is a unique geodesic joining any two distinct points of M. In the sequel H will denote a simply connected and M an arbitrary complete two dimensional manifold without conjugate points. M can be written as a quotient surface H/D, where H is the universal Riemannian cover of M and D is a freely acting, properly discontinuous group of isometries of H. D will always denote such a group. Each non-identity element of D has infinite order since Z_p does not act properly discontinuously on $\mathbf{R}^n$ for any prime p and any integer $n \geqslant 1$ [10].

DEFINITION 1.1. If p and q are distinct points of H, let γ_{pq} denote the unique geodesic such that $\gamma_{pq}(0)=p$ and $\gamma_{pq}(a)=q$, where $a=d(p, q)$. Let $V(p, q)$ denote the unit vector $\gamma_{pq}'(0)$.

Since $H \cong \mathbf{R}^2$ is two dimensional one may define the left and right half planes determined by a geodesic γ of H. Since H is orientable we may assign an orientation to each tangent space $T_p(H)$ that varies continuously with p. A basis $\{v_1, v_2\}$ of $T_p(H)$ is *positively oriented* if the orientation of $T_p(H)$ that it determines agrees with the given orientation for $T_p(H)$ and is *negatively oriented* otherwise. $H-\gamma$ consists of two connected components for any maximal geodesic γ of H. Each of these components is convex in the sense that it contains the unique geodesic segment between any two of its points.

DEFINITION 1.2. Let γ be a maximal geodesic of H. A point p in $H-\gamma$ lies to the *right* (left) of γ if for some number t the pair of unit vectors $\{V(\gamma t, p), \gamma'(t)\}$ is positively (negatively) oriented relative to the fixed orientation of H. The points lying to the right (left) of γ constitute the right (left) *half plane* determined by γ.

Note that the orientation of the pairs $\{V(\gamma t, p), \gamma'(t)\}$ is continuous in t hence constant.

DEFINITION 1.3. An *end* of a Hausdorff space X is a function ε that assigns to each compact subset K' of X a connected component $\varepsilon(K')$ of $X-K'$ with the further requirement that $\varepsilon(K') \supseteq \varepsilon(L)$ if $K' \subseteq L$. A subset U of X is a *neighborhood* of an end ε if U contains $\varepsilon(K')$ for some K'. A sequence of points p_n converges to an end ε if each neighborhood of ε contains p_n for sufficiently large n.

In a noncompact Hausdorff space X a divergent curve $\gamma: [0, \infty) \to X$ determines an end ε of X; for each compact subset K of M define $\varepsilon(K)$ to be the connected component of $X-K$ that contains a terminal segment of γ. A curve γ is *divergent* if for any compact set $C \subseteq X$ there exists a number $T=T(C)>0$ such that $\gamma t \in X-C$ for $t>T$.

If M is a noncompact surface with finitely generated fundamental group, then it is known [1], [9] that M is homeomorphic to a compact surface with a finite number of points removed. Each end of M corresponds to one of these missing points and has a neighborhood U homeomorphic to a punctured disk or equivalently a half cylinder $S^1 x(0, \infty)$.

§2. Fundamental Domains

In this section we define for every point p in H and for every freely acting, properly discontinuous group D of isometries of H, a *canonical fundamental domain* for D with *center* p. We derive basic properties of fundamental domains that are well known if H is the hyperbolic plane but which require more discussion in this general case. We also relate the fundamental domain with center p to the cut locus of $\pi(p)$ in the quotient surface $M=H/D$.

DEFINITION 2.1. Let D be a freely acting, properly discontinuous group of isometries of H. For any point p in H the *canonical fundamental domain* for D with *center* p, denoted R_p, $=\{q \in H: d(p, q) \leqslant d(\phi p, q)$ for all ϕ in $D\}$.

It is easy to see that the interior of R_p, denoted $\mathrm{Int}(R_p)$, $=\{q \in R_p: d(p, q) < d(\phi p, q)$ for every $\phi \neq 1$ in $D\}$. Also ∂R_p, the boundary of R_p, $=\{q \in R_p: d(p, q) = d(\phi p, q)$ for some $\phi \neq 1$ in $D\}$. Hence ∂R_p is contained in the union of the equidistant sets $E(p, \phi p)$, $\phi \in D$, where $E(p, \phi p) = \{q \in H: d(p, q) = d(\phi p, q)\}$. Now, for each ϕ in D and each point p in H define $E^+(p, \phi p)$ to be $\{q \in H: d(p, q) \leqslant d(\phi p, q)\}$. By definition then

$$R_p = \bigcap_{\substack{\phi \in D \\ \phi \neq 1}} E^+(p, \phi p).$$

We remark that R_p is starshaped relative to p; that is if $q \in R_p$ then the geodesic segment γ_{pq} is contained in R_p. This assertion follows from the fact that for each $\phi \in D$ the function $r \to d(p, r) - d(\phi p, r)$ is nondecreasing on geodesics starting at p, which implies that each set $E^+(p, \phi p)$ is starshaped relative to p.

DEFINITION 2.2. We say that an equidistant set $E(p, \phi p)$ is a *bounding side* for R_p if R_p is a proper subset of

$$\bigcap_{\substack{\psi \neq \phi \\ \psi \in D,\, \psi \neq 1}} E^+(p, \psi p)$$

The definitions and discussion so far apply to a manifold H of arbitrary dimension. The next definition is motivated by the fact that in dimension two the sets $E(p, \phi p)$ are differentiable curves in H that meet transversally if at all.

DEFINITION 2.3. A point $\tilde{q}$ in R_p is a *vertex* of R_p if it lies on the intersection of two distinct bounding sides of R_p.

The proper discontinuity of D implies that only finitely many of the sets $E(p, \phi p)$, $\phi \in D$, meet any given compact subset of H. Since two distinct sets $E(p, \phi p)$ and $E(p, \psi p)$ intersect in at most one point by Proposition 2.8 below, it follows that only finitely many vertices of R_p lie in any given compact subset of H.

We next briefly describe the cut locus at a point p of an arbitrary complete Riemannian manifold M. If M has no conjugate points, then we relate the cut locus at p to the canonical fundamental domain for D with center $\tilde{p}$ in H, where $M = H/D$ and $\pi\tilde{p} = p$, $\pi: H \to M$.

Let M be a complete Riemannian manifold of arbitrary dimension, and let p be a point of M. If $S(p)$ denotes the sphere of unit vectors in $T_p(M)$ let $f: S(p) \to [0, \infty]$ be given by $f(v) = \sup\{t \geqslant 0: d(p, \exp_p(tv)) = t\}$. The function f is known to be continuous on the extended real numbers, and hence it has a positive lower bound on $S(p)$ [3].

The cut locus at p, denoted $C(p)$, is defined to be $\{\exp_p(f(v)\cdot v): v \in S(p) \subseteq T_p(M)\}$. The cut locus at p is a closed subset of M, and $f(v)$ measures the distance from p to $C(p)$ in the direction v.

DEFINITION 2.3′. A point $q \in C(p)$ is a *vertex* of $C(p)$ if there are at least three distinct shortest geodesics from p to q.

We shall show later in corollary 2.15 that a point $\tilde{q} \in R_p$ is a vertex of R_p if and only if $q = \pi\tilde{q} \in C(\pi p)$ and q is a vertex of $C(\pi p)$. If $q \in C(p)$, then it is known [3] that either there are at least two distinct shortest geodesics from p to q, or q is conjugate to p along some shortest geodesic segment from p to q. If M has no conjugate points, then the second case does not occur, and furthermore there are only finitely many shortest geodesics from p to q. If $M = H/D$, where H is the universal Riemannian cover of M and D the deckgroup of M, then it is straightforward to verify the following assertions.

1) For any point p in H, a point q lies in the interior of R_p if and only if there is a unique shortest geodesic in M from πp to πq.

2) $\pi: H \to M$ maps the interior of R_p onto its image in a one-one fashion.

3) $\pi: H \to M$ maps R_p onto M and maps ∂R_p onto $C(\pi p)$.

4) If $q \in R_p$ is a vertex of R_p, then πq lies in $C(\pi p)$, and πq is a vertex of $C(\pi p)$.

For a more refined study of the cut locus of a compact surface with curvature $K \leqslant 0$ see [2].

We return to a study of the fundamental domain R_p, especially its boundary. To do this we will need to establish certain properties of the equidistant sets $E(p, q) = \{r \in H: d(p, r) = d(q, r)\}$ for any pair of distinct points p and q in H. The set $E(p, q)$ is a geodesic if H is the hyperbolic plane. In the general case $E(p, q)$ is no longer a geodesic but retains some properties of a geodesic. First, $E(p, q)$ is a C^∞ one dimensional submanifold of H since it is the zero level set of the function $\bar{g}(r) = d(p, r) - d(q, r)$, which is C^∞ on $H - \{p \cup q\}$. Note that the gradient of $\bar{g}$ is nonzero at any point r in $E(p, q)$ since the gradients of $r \to d(p, r)$ and $r \to d(q, r)$ point radially outward from p and q respectively if $r \neq p$ and $r \neq q$. Precisely, these gradients are $-V(r, p)$ and $-V(r, q)$.

In the remainder of this section we omit the proofs of the results to make reading easier. The proofs may be found in section 4, the appendix.

We first define the *canonical parametrization* of $E(p, q)$. Actually there are two such parametrizations; if α is one then $\alpha^*: t \to \alpha(-t)$ is the other. This parametrization has also been used in [6].

PROPOSITION 2.4. *Let p and q be distinct points in H. Then there exists a continuous, one-one map* $\alpha: \mathbf{R} \to H$ *such that* $\alpha(\mathbf{R}) = E(p, q)$, $\alpha(0)$ *is the midpoint of the segment* γ_{pq} *and* $d(p, \alpha t) = |t| + t_0/2$ *for every* $t \in \mathbf{R}$, *where* $t_0 = d(p, q)$.

PROPOSITION 2.5. *The canonical parametrization* α *of an equidistant set* $E(p, q)$ *is* C^∞ *at every number* $t \neq 0$.

We now describe some of the properties of geodesics of H that are retained by the equidistant sets $E(p, q)$.

PROPOSITION 2.6. *Let p and q be distinct points in H. Then* $H - E(p, q)$ *consists of two connected components. The components containing p and q are starshaped relative to p and q respectively. Any maximal geodesic containing p or q meets* $E(p, q)$ *at most once.*

PROPOSITION 2.7. *Let p and q be distinct points in H and let* α *be the canonical parametrization of* $E(p, q)$. *Then* $\lim_{t \to \infty} V(p, \alpha t)$ *and* $\lim_{t \to -\infty} V(p, \alpha t)$ *exist and are distinct. If* γ_1 *and* γ_2 *are the geodesics whose initial velocities are these limits, then the maximal geodesics* γ_1 *and* γ_2 *do not intersect* $E(p, q)$. *The same assertions hold if p is replaced by q.*

PROPOSITION 2.8. *Let p, q and r be distinct points in H. Then* $E(p, q) \cap E(q, r)$ *contains at most one point.*

The results above prepare one to study the properties of the bounding sides of a fundamental domain R_p in H with center p, relative to a freely acting, properly discontinuous group D of isometries of H.

PROPOSITION 2.9. *The boundary of a fundamental domain R_p for D with center p is contained in the union of the bounding sides.*

COROLLARY 2.10. *R_p is the intersection of those sets $E^+(p, \phi p)$ such that $E(p, \phi p)$ is a bounding side for R_p.*

PROPOSITION 2.11. *Let $E(p, \phi p)$ be a bounding side of R_p. Then $E(p, \phi p) \cap R_p$ is nonempty and consists of a subarc, finite or infinite, of the arc $E(p, \phi p)$. If q is an interior point of $E(p, \phi p) \cap R_p$, then q is not a vertex of R_p. If $E(p, \phi p) \cap R_p$ is nonempty for some $\phi \neq 1$ such that $E(p, \phi p)$ is not a bounding side, then $E(p, \phi p) \cap R_p$ is a single point.*

The next result shows that the bounding sides of a canonical fundamental domain may be identified in pairs.

PROPOSITION 2.12. *Let $E(p, \phi p)$ be a bounding side for R_p. Then $R_p \cap E(p, \phi^{-1} p) = \phi^{-1}\{R_p \cap E(p, \phi p)\}$. In particular $E(p, \phi^{-1} p)$ is also a bounding side for R_p.*

The next results characterize the vertices of R_p.

PROPOSITION 2.13. *If $q \in \partial R_p$ is a vertex of R_p, then q lies on the intersection of exactly two bounding sides of R_p.*

PROPOSITION 2.14. *A point $q \in R_p$ is a vertex of R_p if q lies in the intersection of any two distinct equidistant sets $E(p, \phi p)$ and $E(p, \psi p)$ that are not necessarily bounding sides of R_p.*

COROLLARY 2.15. *Let $q \in \partial R_p$ be a point such that $\pi(q)$ is a vertex of $C(\pi p)$, the cut locus at $\pi(p)$ in $M = H/D$. Then q is a vertex of R_p.*

These last results show that there exists an element $\phi \neq 1$ in D such that $E(p, \phi p) \cap R_p$ is a single point q if and only if for some point q in R_p there are at least four distinct shortest geodesics from $\pi(p)$ to $\pi(q)$ in $M = H/D$. If there exist at least four shortest geodesics from $\pi(p)$ to $\pi(q)$ in M, then there exist at least three distinct, nonidentity elements ϕ_1, ϕ_2 and ϕ_3 in D such that $q \in E(p, \phi_i p)$ for $i = 1, 2, 3$. One of these equidistants sets cannot be a bounding side of R_p by proposition 2.13, and therefore it intersects the set R_p in exactly the point q. Conversely if $E(p, \phi p) \cap R_p$ is a single point q for some $\phi \neq 1$ in D, then $q \in \partial R_p$ and q lies in some bounding side $E(p, \psi p)$ by proposition 2.9. By proposition 2.14 q is a vertex of R_p and since $E(p, \phi p)$ is not a bounding side of R_p there exists by proposition 2.13 a third element $\xi \neq 1$ in D such that $E(p, \xi p)$ is a bounding side and $q \in E(p, \xi p)$. Therefore there are at least four shortest geodesics from πp to πq in $M = H/D$. For a discussion of this possibility in the case that M is compact with curvature $K \equiv -1$ see [2].

§ 3. The Main Results

In this section we prove the theorems A, B, and C stated in the introduction. For the proof of theorem A we shall need the following result in which we make no assumption about conjugate points.

LEMMA. *Let M be a complete, noncompact Riemannian manifold of dimension two. Let U be an unbounded open set in M that is homeomorphic to $S^1 \times (0, \infty)$, and let $p \in M - \bar{U}$ be given. Let p_n be a divergent sequence of points contained in U for large n for which there exist distinct shortest geodesics γ_n and σ_n joining p to p_n. Then infinitely many of the loops at p given by $\alpha_n = \sigma_n^{-1}\gamma_n$ are homotopic.*

Proof. Passing to a subsequence we may assume that $p_n \in U$ for all n and there exist geodesics γ and σ (possibly equal) such that $\gamma_n'(0) \to \gamma'(0)$ and $\sigma_n'(0) \to \sigma'(0)$ as $n \to \infty$. The geodesics γ and σ start at p and are distance minimizing on $[0, \infty)$. Denote $\partial U = S^1 \times \{0\}$ by C. We may assume that C is parametrized as a nonsingular C^∞ curve; merely replace U by $S^1 \times (1, \infty)$ and replace $S^1 \times \{1\}$ by a nonsingular C^∞ curve from the same homotopy class that lies in $S^1 \times (0, \infty)$. Since C is compact we may define $t_0 > 0 = \sup\{t > 0 : \gamma t \in C\}$ and $s_0 > 0 = \sup\{t > 0 : \sigma t \in C\}$. By further altering C if necessary we may assume that γ and σ meet C transversally at t_0 and s_0 respectively. Letting $c_n = d(p, p_n)$ we define $t_n = \sup\{0 < t \leqslant c_n : \gamma_n t \in C\}$ and $s_n = \sup\{0 < t \leqslant c_n : \sigma_n t \in C\}$. Since γ and σ meet C transversally it follows that $t_n \to t_0$ and $s_n \to s_0$ as $n \to \infty$. Moreover $\gamma_n t \in U$ for $t_n < t \leqslant c_n$ and $\sigma_n t \in U$ for $s_n < t \leqslant c_n$ since $p_n \in U$. Note that $c_n \to +\infty$ since p_n is a divergent sequence. Finally, $\gamma t \in U$ for $t > t_0$ and $\sigma t \in U$ for $t > s_0$.

Parametrize C on $[0, 2\pi]$ and let a_n, b_n be those numbers in $[0, 2\pi]$ such that $\gamma_n(t_n) = C(a_n)$ and $\sigma_n(s_n) = C(b_n)$. The points $\gamma_n(t_n)$ and $\sigma_n(s_n)$ are distinct for large n since γ_n and σ_n are minimizing on $[0, c_n]$. By passing to a subsequence and relabeling if necessary we may assume that $a_n < b_n$ for every n. Let C_n denote the restriction of C to $[a_n, b_n]$. Let γ_n^*, $\bar{\gamma}_n$ denote the restrictions of γ_n to $[0, t_n]$ and $[t_n, c_n]$ respectively. Let σ_n^*, $\bar{\sigma}_n$ denote the restrictions of σ_n to $[0, s_n]$ and $[s_n, c_n]$. Let γ^* and σ^* denote the restrictions of γ and σ to $[0, t_0]$ and $[0, s_0]$. We may write the loop $\alpha_n = \sigma_n^{-1}\gamma_n$ as a product of two loops at p, $\alpha_n = [(\sigma)_n^{*-1} B_n \sigma_n^*] A_n$, where $A_{nn}^* = (\sigma)^{-1} C_n \gamma_n^*$ and $B_n = (\bar{\sigma}_n)^{-1} (\bar{\gamma}_n) C_n^{-1}$. The curve B_n is a simple closed curve since $\bar{\gamma}_n$ and $\bar{\sigma}_n$ intersect C only at t_n and s_n by the definition of these numbers and intersect each other only at $p_n = \gamma_n(c_n) = \sigma_n(c_n)$ since γ_n and σ_n are minimizing on $[0, c_n]$. Since B_n lies in $\bar{U}$, a closed half cylinder, and is a simple closed curve an application of the Jordan curve theorem shows that B_n is homotopic either to a point or to the curve C^{-1}, which wraps around the cylinder exactly once. Passing to a subsequence the loops $(\sigma_n^*)^{-1} B_n \sigma_n^*$ are homotopic either to a point for all n or to the loop $(\sigma^*)^{-1} C^{-1} \sigma^*$ for all n. For large n the loop A_n is homotopic to the loop $(\sigma^*)^{-1} C^* \gamma^*$, where C^* is a point if

$\gamma=\sigma$ and is the restriction of C to $[a, b]$ if $\gamma \neq \sigma$, where $C(a)=\gamma(t_0)$ and $C(b)=\sigma(s_0)$. Therefore the loops α_n are homotopic to each other for large n.

We now begin the proof of theorem A. The assertion 3)→2) is obvious. We show that 2)→1). Let $S=\{\phi \in D: \phi(R_p) \cap R_p \text{ is nonempty}\}$, where R_p is a fixed fundamental domain in H with a finite number of bounding sides. We assert that S is a finite set, and assuming that this has been proved we apply theorem 29.4 (i) of [4, p. 184] to conclude that S is a generating set for D.

Suppose that S is an infinite set. By removing a finite number of elements from S we may assume that for each $\phi \in S$, $E(p, \phi p)$ is not a bounding side of R_p. By proposition 2.11 $\phi(R_p) \cap R_p = R_p \cap E(p, \phi p)$ is a single point q for $\phi \in S$. Since $q \in \partial R_p$ it follows that q lies in some bounding side of R_p by proposition 2.9. Hence q must be a vertex of R_p by proposition 2.14. The proper discontinuity of D implies that only finitely many $\phi \in S$ determine the same vertex $\phi(R_p) \cap R_p$, and therefore R_p has infinitely many vertices. However, R_p has only finitely many vertices since any two distinct bounding sides intersect at most once by proposition 2.8. This contradiction shows that S is a finite set and completes the proof. One may also show that if $\phi_1, \ldots, \phi_k$ are those elements of D such that $E(p, \phi_i p)$, $1 \leqslant i \leqslant k$, are the bounding sides of R_p, then the elements $\phi_1, \ldots, \phi_k$ are a set of generators for D.

We now prove that 1)→3). Let $\tilde{p} \in H$ be given and suppose $R_{\tilde{p}}$ has an infinite number of bounding sides $E(\tilde{p}, \phi_n \tilde{p})$, $n=1, 2, \ldots$. Choose a point $\tilde{p}_n \in E(\tilde{p}, \phi_n \tilde{p}) \cap R_{\tilde{p}}$, which is possible by proposition 2.11, and let $p_n=\pi(\tilde{p}_n)$. By the choice of $\tilde{p}_n$ the geodesic segments $\sigma_n=\pi \circ \gamma_{\phi_n \tilde{p}, \tilde{p}_n}$ and $\gamma_n=\pi \circ \gamma_{\tilde{p} \tilde{p}_n}$ are distinct shortest geodesics in M from $p=\pi\tilde{p}$ to p_n. By elementary covering space facts no two loops $\alpha_n=\sigma_n^{-1}\gamma_n$ and $\alpha_m=\sigma_m^{-1}\gamma_m$ are homotopic if $m \neq n$ since $\phi_n \neq \phi_m$. If we pass to a subsequence the points p_n converge to some end A of M since the sequence p_n is divergent. Since $\pi_1(M)$ is finitely generated it is known that M is homeomorphic to a compact surface with a finite number of points removed. For a surface of this type each end A has a neighborhood U homeomorphic to a punctured disk or equivalently to $S^1 \times (0, \infty)$. Applying the lemma above we obtain a contradiction.

We now prove theorem B. Let $p \in M$ be given, and let $\tilde{p} \in \pi^{-1}(p)$, $\pi: H \to M$, be arbitrarily chosen. By definition $G_2(p)$ equals the cut locus of p, $C(p)$, minus the vertices of $C(p)$. By the discussion following definition 2.3′, $G_2(p)$ is the image under π of the boundary of $R_{\tilde{p}}$ minus the vertices of $R_{\tilde{p}}$. By proposition 2.9 and 2.11 the set $\partial R_{\tilde{p}}$ minus vertices of $R_{\tilde{p}}$ is the union of the interiors of the differentiable arcs $E(\tilde{p}, \phi\tilde{p}) \cap R_{\tilde{p}}$, where $E(\tilde{p}, \phi\tilde{p})$ is a bounding side of $R_{\tilde{p}}$. To show that $G_2(p)$ is a disjoint union of open differentiable arcs it suffices to show that if $E(\tilde{p}, \phi\tilde{p}) \cap R_{\tilde{p}}$ and $E(\tilde{p}, \psi\tilde{p}) \cap R_{\tilde{p}}$ are distinct boundings arcs of $R_{\tilde{p}}$, then the images under π of their interiors are either disjoint or identical. Suppose that these images intersect for some $\phi, \psi \in D$. Then there exists a point $\tilde{q}$ in the interior of $E(\tilde{p}, \phi\tilde{p}) \cap R_{\tilde{p}}$ and an element $\xi \neq 1$ in D such that $\xi\tilde{q}$ lies in the interior of $E(\tilde{p}, \psi\tilde{p}) \cap R_{\tilde{p}}$. Define geodesics segments

$\gamma_1=\pi\circ\gamma_{\psi\tilde{p},\xi\tilde{q}}=\pi\circ\gamma_{\tilde{p},\psi^{-1}\xi\tilde{q}}$, $\gamma_2=\pi\circ\gamma_{\tilde{p}\tilde{q}}$, $\gamma_3=\pi\circ\gamma_{\tilde{p},\xi\tilde{q}}$ and $\gamma_4=\pi\circ\gamma_{\phi\tilde{p}.\tilde{q}}=\pi\circ\gamma_{\tilde{p},\phi^{-1}\tilde{q}}$. By the choice of ϕ, ψ and ξ it follows that these are all shortest geodesics in M from p to $q=\pi(\tilde{q})$. The point $\tilde{q}$ is not a vertex of $R_{\tilde{p}}$ since it lies in the interior of $E(\tilde{p}, \phi\tilde{p})\cap R_{\tilde{p}}$, and therefore q is not a vertex of $C(p)$. It follows that at most two of the geodesics γ_1, γ_2, γ_3, γ_4 are distinct. By inspection $\gamma_2\neq\gamma_3$ and $\gamma_2\neq\gamma_4$. Hence $\gamma_3=\gamma_4$, which implies that $\xi=\phi^{-1}$. Similarly $\gamma_1\neq\gamma_3$ and since $\gamma_2\neq\gamma_3$ it follows that $\gamma_1=\gamma_2$, implying that $\psi^{-1}\xi=1$. Therefore $\psi=\phi^{-1}$ and $E(\tilde{p}, \psi\tilde{p})\cap R_{\tilde{p}}=\phi^{-1}\{E(\tilde{p}, \phi\tilde{p})\cap R_{\tilde{p}}\}$ by proposition 2.12. Therefore the images under π of the interiors of $E(\tilde{p}, \phi\tilde{p})\cap R_{\tilde{p}}$ and $E(\tilde{p}, \psi\tilde{p})\cap R_{\tilde{p}}$ are identical.

Suppose now that M is a complete surface without conjugate points and that $\pi_1(M)$ is finitely generated. Let $p\in M$ be given, and let $\tilde{p}\in H$, $\tilde{p}\in\pi^{-1}(p)$, be chosen arbitrarily. By theorem A $R_{\tilde{p}}$ has a finite number of bounding sides, and by the discussion above it follows that $G_2(p)$ has a finite number of connected components. Thus 1)→3) in theorem B. Clearly 3)→2). The discussion above also shows that each connected component of $G_2(p)$ is the image under π of exactly two bounding arcs of $R_{\tilde{p}}$. If $G_2(p)$ has a finite number k of connected components for some $p\in M$, then $R_{\tilde{p}}$ has $2k$ bounding sides for any $\tilde{p}\in\pi^{-1}(p)$. By theorem A $\pi_1(M)$ is finitely generated and we have proved that 2)→1).

We now prove theorem C. Let M be a complete surface without conjugate points and with finitely generated fundamental group, and let $p\in M$ be given. If $q\in M$ is a point for which there are at least 3 shortest geodesics from p to q, then by definition q lies in the cut locus, $C(p)$, of p and is a vertex of $C(p)$. Let $\tilde{p}\in\pi^{-1}(p)$ be arbitrary. By corollary 2.15 and the discussion in section 2, $q=\pi(\tilde{q})$, where $\tilde{q}$ is a vertex of $R_{\tilde{p}}$. By theorem A $R_{\tilde{p}}$ has only a finite number of bounding sides. It follows that $R_{\tilde{p}}$ has at most a finite number of vertices since a vertex lies in the intersection of two bounding sides, which must be a single point by proposition 2.8. Therefore $C(p)$ has a finite number of vertices, which completes the proof of theorem C.

We do not know if the converse to theorem C is true although we suspect that it is. In principe it might be possible to have a deckgroup D of isometrics of H for which each fundamental domain R_p has an infinite number of bounding sides, only finitely many of which intersect. The quotient surface H/D would then have an infinitely generated fundamental group, but each cut locus $C(p)$ would have only a finite number of vertices.

Remark. Theorem A can be derived from theorem 2 of Marden [11] in the case that M is a noncompact, nonsimply connected orientable surface. Since M is orientable it has the structure of a Riemann surface, and M is therefore diffeomorphic to a quotient Δ/G, where Δ is the open unit disk in the complex plane and G is a freely acting, properly discontinuous group of fractional linear transformations preserving Δ. If g is a metric without conjugate points in M, then the covering map $\pi:\Delta\to M$ induces a metric π^*g without conjugate points in Δ, and the elements of G are isometries of the metric π^*g. The results of section 2 of this paper show that the canonical fundamental domains in Δ determined by G are fundamental regions in the sense of Marden. Therefore if $G \cong \pi_1(M)$ is finitely generated, then each canonical fundamental domain has finitely many bounding sides by Marden's theorem 2. This shows that 1) implies 3) in theorem A and the other assertions follow as above.

§4. Appendix

In this section we give the proofs of the results stated in section 2.

Proof of proposition 2.4. We shall need

LEMMA 2.4. *Let p and q be distinct points of H, and let $t_0 = d(p, q)$. Then there exists a unique point z in $E(p, q)$ such that $d(p, z) = d(q, z) = t_0/2$, and moreover, $d(y, p) \geqslant d(z, p)$ for any $y \in E(p, q)$. If $t^* > t_0/2$ is any number, then there are exactly two points z_1, z_2 in $E(p, q)$ such that $d(p, z_1) = d(p, z_2) = t^*$.*

Proof. Let γ be the maximal geodesic γ_{pq}. If z is the midpoint of the geodesic segment γ_{pq} between p and q, then $d(p, z) = d(q, z) = t_0/2$. The uniqueness of z and the fact that $d(y, p) \geqslant d(z, p)$ for every $y \in E(p, q)$ follow immediately from the triangle inequality. Let H be given a fixed orientation. Given $t^* > t_0/2$, let $\beta : [0, A] \to H$ be a unit speed parametrization of the circle of center p and radius t^* such that $\beta(0)$ and $\beta(A^*)$, $0 < A^* < A$, are the two points of this circle that lie on γ and $\beta : (0, A^*) \to H$ parametrizes the semicircle lying to the left of γ. Let $J(t) = d(q, \beta t)$. Since $\beta(t) \neq q$ for $0 < t < A^*$ it follows from lemma 2.3 of [5] that $J'(t) = -\langle \beta'(t), V(\beta t, q) \rangle$. Since $\beta'(t)$ is orthogonal to the vector $V(\beta t, p)$ for every t and since $V(\beta t, p)$ and $V(\beta t, q)$ are not collinear for $0 < t < A^*$ it follows that $J'(t) \neq 0$ in this interval. Since $J(0) = t_0 + t^*$ and $J(A^*) = |t_0 - t^*| < t^*$ by a suitable choice of β, it follows that there exists a unique number s with $0 < s < A^*$ such that $d(p, \beta s) = d(q, \beta s) = t^*$. Similarly there is a unique number s' with $A^* < s' < A$ such that $d(p, \beta s) = d(q, \beta s') = t^*$. The points βs and $\beta s'$ are on opposite sides of γ.

We now complete the proof of proposition 2.4. Given distinct points p and q in H, let $\alpha(0)$ be the midpoint of the segment γ_{pq}. Let $t > 0$ be given. Relative to a fixed orientation of H, let $\alpha(t)$ be the unique point to the right of γ_{pq} such that $\alpha(t)$ lies in $E(p, q)$ and $d(p, \alpha t) = t + t_0/2$, where $t_0 = d(p, q)$. Let $\alpha(-t)$ denote the unique point in $E(p, q)$ such that $\alpha(-t)$ lies to the left of γ_{pq} and $d(p, \alpha(-t)) = t + t_0/2$. This defines a map $\alpha : \mathbf{R} \to E(p, q)$ which is a homeomorphism. The proof is complete.

Proof of proposition 2.5. The set $E(p, q)$ is a C^∞ one dimensional manifold since it is the zero level set of the function $r \to d(p, r) - d(q, r)$, which is C^∞ on $H - \{p \cup q\}$ and whose gradient is never zero on $E(p, q)$. Given a number $t > 0$ let $\beta : (-\varepsilon, \varepsilon) \to \to E(p, q)$ be a nonsingular C^∞ map such that $\beta(0) = \alpha(t)$. The C^∞ function $\phi(u) = = d(p, \beta u) = d(q, \beta u)$ is nonsingular at $u = 0$. If this were not the case, then the vectors

$V(\beta 0, p)$ and $V(\beta 0, q)$ would both be perpendicular to $\beta'(0)$, which would imply that p, $\beta 0$ and q are collinear, contradicting the fact that $\beta(0)=\alpha(t)$, $t>0$. If $h(s)$ is the inverse function of ϕ, then h is a C^∞ diffeomorphism of some neighborhood J of $t+t_0/2$ onto some neighborhood I of 0, $I\subseteq(-\varepsilon, \varepsilon)$. It follows that $d(p, \beta(hs))=s$ for all s in J, and therefore $\alpha(s-t_0/2)=\beta(h(s))$ for all s in J. Hence α is C^∞ at every number $t>0$.

Proof of proposition 2.6. Let $g:H\to\mathbf{R}$ be given by $g(r)=d(p, r)-d(q, r)$. It is clear that $E(p, q)=g^{-1}(0)$ and that the two components of $H-E(p, q)$ are the sets $A_1=\{r:g(r)>0\}$ and $A_2=\{r:g(r)<0\}$. The set A_1 is starshaped relative to q since g is nonincreasing on every geodesic starting at q, and A_2 is starshaped relative to p since g is nondecreasing on every geodesic starting at p.

Suppose now that γ is a unit speed geodesic with $\gamma(0)=p$ such that γ meets $E(p, q)$ twice at times $s\neq 0$, $t\neq 0$. There are two cases: 1) $q\in\gamma$ and 2) $q\notin\gamma$. In the first case $\gamma=\gamma_{pq}$. Now $d(p, \gamma t)-d(q, \gamma t)$ vanishes for only one value of t since p, q and $\gamma(t)$ are always collinear. Suppose now that $q\notin\gamma$. Then q, γs and γt are not collinear. If s and t both have the same sign with $|t|>|s|$, then $d(p, \gamma t)=d(q, \gamma t)<d(q, \gamma s)+d(\gamma s, \gamma t)=$ $=d(p, \gamma s)+d(\gamma s, \gamma t)=d(p, \gamma t)$, a contradiction. Suppose that s and t have opposite signs. Then $d(\gamma s, \gamma t)<d(\gamma s, q)+d(q, \gamma t)=d(\gamma s, p)+d(p, \gamma t)=d(\gamma s, \gamma t)$, another contradiction. Similarly no geodesic containing q can meet $E(p, q)$ twice.

Proof of proposition 2.7. We prove the assertions only for the point p. The curve $f(t)=V(p, \alpha t)$ is a continuous map of $\mathbf{R}$ into S^1, the unit vectors in $T_p(H)$. Since any maximal geodesic through p meets $E(p, q)$ at most once it follows that f is one-one and $f(\mathbf{R})$ contains no pair of antipodal points. Therefore $f(\mathbf{R})$ is an open arc in S^1 with distinct endpoints $v_1=\lim_{t\to\infty} V(p, \alpha t)$ and $v_2=\lim_{t\to\infty} V(p, \alpha t)$. Suppose now that the maximal geodesic γ_1 intersects $E(p, q)$ at $\gamma_1(s)=\alpha(t)$ for some numbers s and t. If $s>0$ then $v_1=\gamma_1'(0)=V(p, \alpha t)$ is an interior point of $f(\mathbf{R})$, which is impossible. If $s<0$ then $-v_1=V(p, \alpha t)$ is an interior point of $f(\mathbf{R})$, which implies that $f(\mathbf{R})$ contains a pair of antipodal points near $\{v_1, -v_1\}$, a contradiction. Therefore γ_1 does not meet $E(p, q)$. Similarly γ_2 does not meet $E(p, q)$.

Proof of proposition 2.8. We shall need some preliminary results.

LEMMA 2.8a. *Let p, q, r be distinct points in H. Let $h:H\to\mathbf{R}$ be the function given by $a\to d(r, a)-d(q, a)$, and let α be the canonical parametrization of $E(p, q)$. Then $h\circ\alpha$ has at most one relative maximum or minimum point. If $h\circ\alpha$ has a relative maximum or minimum point at $t_0\in\mathbf{R}$ then either*

1) $r=\alpha(t_0)$
2) *$r\notin\alpha$; p, r and $\alpha(t_0)$ are collinear with p and r on the same side of α or*
3) *$r\notin\alpha$; q, r and $\alpha(t_0)$ are collinear with q and r on the same side of α.*

Proof. If r lies in α, say $r=\alpha(t^*)$, then the triangle inequality implies that $h\circ\alpha$ has a strict global minimum at t^*. Assuming that the latter part of the lemma has been proved it follows that in this case $h\circ\alpha$ has no relative maximum or minimum at $t_0\neq t^*$.

Suppose now that $h \circ \alpha$ has a relative maximum or minimum at t_0 and that $r \neq \alpha(t_0)$. Then h is C^∞ in a neighborhood of $\alpha(t_0)$. If $t_0 = 0$ then α is not C^∞ at t_0, but in any case we can find a C^∞ diffeomorphism $\sigma:(-\varepsilon, \varepsilon) \to E(p, q)$ such that $\sigma(0) = \alpha(t_0)$ since $E(p, q)$ is a C^∞ submanifold. Therefore $0 = (h \circ \sigma)'(0) = \langle \sigma'(0), \operatorname{grad} h(\alpha t_0) \rangle$ since $h \circ \alpha$ has a relative maximum or minimum at t_0. Define $g: H \to \mathbf{R}$ by $g(a) = d(p, a) - d(q, a)$. Since $E(p, q) = g^{-1}(0)$, $g \circ \alpha \equiv 0$ and thus $0 = (g \circ \sigma)'(0) = \langle \sigma'(0), \operatorname{grad} g(\alpha t_0) \rangle$. Since $\sigma'(0)$ is nonzero it follows that either

i) $(\operatorname{grad} g)(\alpha t_0) = 0$,

ii) $(\operatorname{grad} h)(\alpha t_0) = 0$, or

iii) $(\operatorname{grad} g)(\alpha t_0)$ and $(\operatorname{grad} h)(\alpha t_0)$ are both nonzero and collinear.

If $a \neq p$ and $a \neq q$, then $(\operatorname{grad} g)(a)$ exists and equals $-V(a, p) + V(a, q)$. In particular $\operatorname{grad} g$ is nonzero at all points of $E(p, q)$ so case i) does not occur. If $a \neq r$ and $a \neq q$, then $(\operatorname{grad} h)(a)$ exists and equals $-V(a, r) + V(a, q)$. The point $\alpha(t_0)$ is neither p nor q, nor r by assumption so that both $\operatorname{grad} g$ and $\operatorname{grad} h$ exist at $\alpha(t_0)$. If $(\operatorname{grad} h)(\alpha t_0) = 0$ as in case ii), then $V(\alpha t_0, r) = V(\alpha t_0, q)$, which implies that q, r and $\alpha(t_0)$ are collinear. Moreover r does not lie on α since no geodesic through q intersects $E(p, q)$ twice. Hence q and r lie on the same side of α. Finally suppose that $(\operatorname{grad} g)(\alpha t_0)$ and $(\operatorname{grad} h)(\alpha t_0)$ are both nonzero and collinear. If the unit vectors $V(\alpha t_0, r)$, $V(\alpha t_0, p)$ and $V(\alpha t_0, q)$ are all distinct, then it is easy to see from the expressions above that $(\operatorname{grad} g)(\alpha t_0)$ and $(\operatorname{grad} h)(\alpha t_0)$ are not collinear, a contradiction. Hence $V(\alpha t_0, r) = V(\alpha t_0, p)$ (implying that $\operatorname{grad} h = \operatorname{grad} g$ at $\alpha(t_0)$). Thus p, r and $\alpha(t_0)$ are collinear with p and r lying on the same side of α. The point r cannot lie on α since no geodesic from p meets $E(p, q)$ twice. This property of geodesics through p or q now implies that $h \circ \alpha$ has at most one relative maximum or minimum upon inspection of the possibilities 1), 2), 3) of the lemma.

LEMMA 2.8b. *Let p, q, r, h and α be as in the previous lemma. Then one of the following must occur:*

1) *$h \circ \alpha$ has a unique global minimum at some number t_0, and $h \circ \alpha$ is strictly monotone on $(-\infty, t_0)$ and (t_0, ∞)*

2) *$h \circ \alpha$ has a unique global maximum at some number t_0, and $h \circ \alpha$ is strictly monotone on $(-\infty, t_0)$ and (t_0, ∞)*

3) *$h \circ \alpha$ is strictly monotone on* $\mathbf{R}$.

Moreover,

1) *occurs if $r = \alpha(t_0)$ or if q, r, $\alpha(t_0)$ are collinear with q and r on the same side of α, and r between q and $\alpha(t_0)$ or if p, r, $\alpha(t_0)$ are collinear with p and r on the same side of α, and r between p and $\alpha(t_0)$,*

2) *occurs if q, r, $\alpha(t_0)$ are collinear with q and r on the same side of α, and q between r and $\alpha(t_0)$ or if p, r, $\alpha(t_0)$ are collinear with p and r on the same side of α, and p between r and $\alpha(t_0)$,*

3) *occurs if* r *and* q *lie on the same side of* α *and* γ_{qr} *does not meet* α *or if* r *and* p *lie on the same side of* α *and* γ_{pr} *does not meet* α

Proof. If $h\circ\alpha$ has no relative maximum or minimum on **R**, then it is one-one and hence strictly monotone on **R**. If $h\circ\alpha$ has a relative maximum or minimum at a number t_0, then by the previous lemma it is one-one hence strictly monotone on the intervals $(-\infty, t_0)$ and (t_0, ∞). Therefore t_0 is a global maximum or minimum.

We now consider the various cases in which these possibilities occur. We have already observed that 1) occurs if $r=\alpha(t_0)$ for some t_0. Suppose now that r does not lie on α. If q and r lie on the same side of α and γ_{qr} does not meet α, then $h\circ\alpha$ has no relative maximum or minimum by the previous lemma and hence case 3) occurs. Suppose now that q and r lie on the same side of α and γ_{qr} meets α at $\alpha(t_0)$ (only one intersection is possible). If r lies between q and $\alpha(t_0)$, then for any $s\in\mathbf{R}$ $(h\circ\alpha)(s)-(h\circ\alpha)(t_0)=d(r, \alpha s)-d(q, \alpha s)-d(r, \alpha t_0)+d(q, \alpha t_0)=d(r, \alpha s)-d(q, \alpha s)+d(q, r)\geqslant 0$. Hence case 1) occurs. If q lies between r and $\alpha(t_0)$, then $(h\circ\alpha)(s)-(h\circ\alpha)(t_0)=d(r, \alpha s)-d(q, \alpha s)-d(q, r)\leqslant 0$. Hence case 2) occurs. The various cases where p and r lie on the same side of α are handled in a similar fashion. Note that $d(q, \alpha s)=d(p, \alpha s)$ for all s.

We begin the proof of proposition 2.8. We show first that $E(p, q)\cap E(q, r)$ contains at most two points. This is equivalent to showing that $h\circ\alpha$ is zero at most twice. If $h\circ\alpha$ had at least three zeros, however, then it would have at least two relative maxima or minima, which is impossible by lemma 2.8a.

Let β denote the canonical parametrization of $E(r, q)$. Suppose that $E(p, q)\cap E(r, q)$ contains two points $\alpha(t_0)=\beta(\bar{t}_0)$ and $\alpha(t_1)=\beta(\bar{t}_1)$. By replacing α if necessary by the other canonical parametrization of $E(p, q)$, $t\to\alpha(-t)$, we may assume that $t_0<t_1$ and $\bar{t}_0<\bar{t}_1$. Now let S^1 denote the unit vectors in $T_q(H)$. Define continuous curves $\gamma_1:[t_0, t_1]\to S^1$ and $\gamma_2:[\bar{t}_0, \bar{t}_1]\to S^1$ by setting $\gamma_1(t)=V(q, \alpha t)$ and $\gamma_2(t)=V(q, \beta t)$. Let $z_1=\gamma_1(t_0)=\gamma_2(\bar{t}_0)$ and $z_2=\gamma_1(t_1)=\gamma_2(\bar{t}_1)$. Then γ_1 and γ_2 are both one-one arcs in S^1 joining z_1 to z_2, since each geodesic from q meets α or β at most once. Therefore either

1) $\gamma_1\cup\gamma_2=S^1$ or
2) $\gamma_1=\gamma_2$.

Suppose that case 1) holds. Then any geodesic γ starting at q intersects $\alpha[t_0, t_1]\cup\beta[\bar{t}_0, \bar{t}_1]\subseteq E(p, q)\cup E(q, r)$ at least twice, including one intersection point of the form $\gamma(t)$, $t>0$ and one point of the form $\gamma(t')$, $t'<0$. Consider the geodesic γ such that $\gamma'(0)=\lim_{t\to\infty} V(q, \gamma t)$. By proposition 2.7 γ never intersects $\alpha=E(p, q)$, so that γ must intersect $\beta=E(q, r)$ at least twice by the preceding remark. This contradicts the fact that any maximal geodesic through q meets $E(q, r)$ at most once, and hence case 1) is eliminated.

Suppose that case 2) holds. Choose any number $t\in(t_0, t_1)$. By hypothesis the geodesic $\gamma_{q\alpha t}$ meets $\beta(\bar{t}_0, \bar{t}_1)$ in some point $\beta(t^*)$, and $\beta(t^*)\neq\alpha(t)$ since $E(p, q)\cap E(q, r)$ consists of the points $\alpha(t_0)$, $\alpha(t_1)$. If $\beta(t^*)$ lies between q and $\alpha(t)$, then q and $\alpha(t)$ lie on opposite sides of $E(q, r)$. Hence $\alpha(t_0, t_1)$ and q lie on opposite sides of $E(q, r)$. If $\alpha(t)$ lies between q and $\beta(t^*)$, then it follows similarly that $\beta(\tilde{t}_0, \tilde{t}_1)$ and q lie on opposite sides of $E(p, q)$. We shall only obtain a contradiction to the first possibility since the second reduces to the first by interchanging the roles of p and r.

We are given that $\alpha(t_0, t_1)$ and q lie on opposite sides of $E(q, r)$. Define a continuous map $\varrho: [t_0, t_1] \to E(q, r)$ by setting $\varrho(t)$ equal to the unique point of intersection of $\gamma_{q\alpha t}$ with $E(q, r)$. The fact that $\gamma_1 = \gamma_2$, the hypothesis of case 2), implies immediately that the point sets $\varrho[t_0, t_1]$ and $\beta[\tilde{t}_0, \tilde{t}_1]$ are equal. Let $g: H \to \mathbf{R}$ be given by $g(a) = d(p, a) - d(q, a)$. Then $g(\varrho t) > 0$ for any number $t \in (t_0, t_1)$ since $\varrho(t)$ lies between q and $\alpha(t)$ and g is strictly monotone decreasing on the segment $\gamma_{q\alpha t}$. Therefore $(g \circ \beta) > 0$ on $(\tilde{t}_0, \tilde{t}_1)$, which implies that $g \circ \beta$ has a global maximum in $(\tilde{t}_0, t_1)$ by lemma 2.8b and the fact that $g \circ \beta$ vanishes at $\tilde{t}_0$ and $\tilde{t}_1$. Lemma 2.8b implies further that either

i) γ_{pq} meets $E(q, r)$ in a point z and q lies between p and z or

ii) γ_{pr} meets $E(q, r)$ in a point z, and r lies between p and z.

We treat these cases separately.

If i) holds then $g(z) = d(p, z) - d(q, z) = d(p, q) > 0$. Thus $z = \beta(t)$ for some t in $(\tilde{t}_0, \tilde{t}_1)$ since $(g \circ \beta) \leqslant 0$ on $(-\infty, \tilde{t}_0]$ and $[\tilde{t}_1, \infty)$. If $t^* \in (t_0, t_1)$ is that number such that $\beta(t) = \varrho(t^*)$, then γ_{pq} meets $E(p, q)$ twice, once between p and q and once at $\alpha(t^*)$, which is beyond $z = \beta(t) = \varrho(t^*)$. This contradicts the fact that any geodesic from p meets $E(p, q)$ at most once.

Suppose that ii) holds. Then γ_{pr} meets $E(p, r)$ at a point y between p and r and meets $E(r, q)$ at a point z as assumed. The point r is thus an interior point of the segment γ_{yz}. Now if u and v are the points of intersection of $E(p, q)$ and $E(r, q)$ then they are equidistant from p, q and r and hence also lie in $E(p, r)$. Let δ be the canonical parametrization of $E(p, r)$ with $u = \delta(t_0^*)$, $v = \delta(t_1^*)$ and $t_0^* < t_1^*$. As in case 1) earlier we let S^1 denote the unit vectors in $T_r(H)$. Define continuous curves $\delta_1: [t_0^*, t_1^*] \to S^1$ and $\delta_2: [\tilde{t}_0, \tilde{t}_1] \to S^1$ by $\delta_1(t) = V(r, \delta t)$ and $\delta_2(t) = V(r, \beta t)$. Then δ_1 and δ_2 are both one-one arcs in S^1 joining $V(r, u)$ to $V(r, v)$. Either

1) $\delta_1 \cup \delta_2 = S^1$ or

2) $\delta_1 = \delta_2$.

The case 1) is impossible by the same argument used earlier in the proof. Suppose that $\delta_1 = \delta_2$. We show first that $y \in \delta[t_0^*, t_1^*]$ and $z \in \beta[\tilde{t}_0, \tilde{t}_1]$. By the definition of z, $g \circ \beta$ has a global maximum at t^*, where $\beta(t^*) = z$, and $t^* \in (\tilde{t}_0, \tilde{t}_1)$. Moreover, $V(r, z)$ is an interior point of δ_2. Now δ_2 does not contain any pair of antipodal points of S^1; if $V(r, a)$ and $V(r, b) = -V(r, a)$ both lay in δ_2 for points a, b in $\beta[\tilde{t}_0, \tilde{t}_1]$, then the geodesic $\gamma_{ab} = \gamma_{ra} = \gamma_{rb}$ would intersect $E(q, r)$ at a and b, contradicting the fact that any geodesic containing r meets $E(q, r)$ at most once. Therefore δ_2 is an arc in S^1 of length $< \pi$. The fact that $V(r, z)$ is an interior point of δ_2, whose endpoints are $V(r, u)$ and $V(r, v)$, now implies that u and v lie on opposite sides of the maximal geodesic $\gamma_{rz} = \gamma_{pr}$. Now $\delta[t_0^*, t_1^*]$ is a curve joining u to v so δ must intersect γ_{pr} in a point y^*.

Since γ_{pr} meets $E(p, r)$ in the points y and y^* it follows that $y=y^*$. Thus $y\in\delta\,[t_0^*, t_1^*]$. The previous work has shown that $V(r, y)\in\delta_1$ and $V(r, z)=-V(r, y)\in\delta_2$. Since $\delta_1=\delta_2$ by hypothesis, the geodesic $\gamma_{yz}=\gamma_{rz}=\gamma_{ry}$ meets $E(p, r)$ (and $E(q, r)$) twice, on opposite sides of the point r. This contradiction completes the proof of the proposition.

Proof of proposition 2.9. We have already observed that ∂R_p is contained in the union of the sets $E(p, \phi p)$, $\phi\in D$. Let a point $q\in\partial R_p$ be given. The proper discontinuity of D implies that q is contained in only finitely many equidistant sets $E(p, \phi_i p)$, $1\leqslant i\leqslant n$. Since only finitely many of the sets $E(p, \phi p)$ meet any compact neighborhood of q, it follows that for a sufficiently small open set 0 containing q and any $\phi\in D-\{\phi_1,\ldots, \phi_n\}$, $0\subseteq \operatorname{Int} E^+(p, \phi p)=\{r\in H: d(p, r)<d(\phi p, r)\}$. Choose $x\in 0-R_p$; this can be done since $q\in\partial R_p$. Moreover, let p, q and x be noncollinear. The geodesic segment γ_{px} meets ∂R_p, and in fact if γ_{px} intersects $E(p, \phi p)$ then $\phi\in\{\phi_1,\ldots, \phi_n\}$ by the way in which 0 was chosen. Now γ_{px} meets each of the sets $E(p, \phi_i p)$ at most once so there are a finite number $k\leqslant n$ of intersections of γ_{px} with $\bigcup_{i=1}^n E(p, \phi_i p)$. Let $\gamma_{px}(t_i)$, $1\leqslant i\leqslant k$, be these intersections, where $t_i<t_{i+1}$, and let r be that integer such that $\gamma_{px}(t_1)\in E(p, \phi_r p)$. Note that $\gamma_{px}(t_1)$ lies in exactly one of the sets $E(p, \phi_i p)$ since the unique point of intersection of any two sets $E(p, \phi_i p)$, $E(p, \phi_j p)$ is q.

We assert that $E(p, \phi_r p)$ is a bounding side. Let $t^*\in(t_1, t_2)$ be arbitrary. If $z=\gamma_{px}(t^*)$ then the geodesic segment γ_{pz} intersects $E(p, \phi_r p)$ but not $E(p, \phi p)$ if $\phi\neq\phi_r$, $\phi\neq 1$. Therefore $d(p, z)<d(\phi p, z)$ if $\phi\neq\phi_r$, $\phi\neq 1$ and $z\in\bigcap_{\substack{\phi\neq\phi_r\\ \phi\neq 1}} E^+(p, \phi p)$. However $z\notin R_p$ since γ_{pz} intersects $E(p, \phi_r p)$, implying that $d(p, z)<d(\phi_r p, z)$. This proves that $E(p, \phi_r p)$ is a bounding side.

Proof of corollary 2.10. If A denotes the intersection of these sets it is clear from the definition of R_p that $R_p\subseteq A$. If R_p were a proper subset of A, then for any point q in $A-R_p$ the geodesic segment γ_{pq} would meet ∂R_p in a point q^* in the interior of γ_{pq}. By the preceding result q^* lies in some bounding side $E(p, \phi p)$, which implies that points on γ_{pq} beyond q^*, in particular q, lie in $H-E^+(p, \phi p)$. This contradicts the hypothesis that $q\in A\subseteq E^+(p, \phi p)$.

Proof of proposition 2.11. We first establish the following

LEMMA 2.11. *If $R_p\cap E(p, \phi p)$ is nonempty, then $R_p\cap E(p, \phi p)$ is an arc connected subset of $E(p, \phi p)$.*

Proof. We may assume that $R_p\cap E(p, \phi p)$ contains at least two points, for otherwise the result is vacuously true. Let q_1 and q_2 be two points in $E(p, \phi p)\cap R_p$. Giving $E(p, \phi p)$ the canonical parametrization α, we know that $q_1=\alpha(s)$ and $q_2=\alpha(t)$ for some numbers s and t. We may assume that $s<t$. If $\alpha(u)\in H-R_p$ for some number u with $s<u<t$, then $\alpha(u)\in H-E^+(p, \psi p)$ for some nonidentity element $\psi\neq\phi$. Let $f: H\to \mathbf{R}$ be the function given by $f(r)=d(p, r)-d(\psi p, r)$. Now $f(\alpha s)\leqslant 0$ and $f(\alpha t)\leqslant 0$ since $\alpha(s)$ and $\alpha(t)$ lie in R_p. On the other hand $f(\alpha u)>0$ by hypothesis.

Hence $f\circ\alpha$ equals zero at some points s^* and t^* with $s \leqslant s^* < u < t^* \leqslant t$. Therefore $E(p, \phi p)\cap E(p, \psi p)$ contains the distinct points $\alpha(s^*)$ and $\alpha(t^*)$, contradicting proposition 2.8. Therefore $\alpha[s, t]\subseteq R_p\cap E(p, \phi p)$.

We begin the proof of proposition 2.11. Suppose that $E(p, \phi p)\cap R_p$ is nonempty but $E(p, \phi p)$ is not a bounding side of R_p. Assuming that $E(p, \phi p)\cap R_p$ contains more than one point we see by the previous lemma that $E(p, \phi p)\cap R_p$ consists of an entire subarc of $E(p, \phi p)$. Let q be an interior point of $E(p, \phi p)\cap R_p$, and choose a number $\varepsilon>0$ such that the set $A=\overline{B_\varepsilon(q)}\cap E(p, \phi p)$ is a compact subarc of $E(p, \phi p)$ that is contained in R_p. In particular $A\subseteq\partial R_p$. ($\overline{B_\varepsilon(q)}$ denotes the closed ball of radius ε and center q in H). Let $\psi_1, \dots, \psi_k$ be those elements of D such that $E(p, \psi_i p)$, $1\leqslant i\leqslant k$, are the only bounding sides of R_p that intersect A. The element ϕ is not equal to ψ_i for any i since $E(p, \phi p)$ is not a bounding side of R_p. Hence $E(p, \phi p)\cap E(p, \psi_i p)$ is at most one point for each i by proposition 2.8. Therefore the set $A-\bigcup_{i=1}^k E(p, \psi_i p)$ is an infinite set. Let r be an arbitrary point of point of $A-\bigcup_{i=1}^k E(p, \psi_i p)$. Proposition 2.9 implies that r lies in some bounding side of R_p since $r\in\partial R_p$. However, the only bounding sides of R_p that meet A are $E(p, \psi_i p)$, $1\leqslant i\leqslant k$, contradicting the choice of r. Therefore $E(p, \phi p)\cap R_p$ is a single point.

Next let $E(p, \phi p)$ be a bounding side of R_p. We show first that $E(p, \phi p)\cap R_p$ is nonempty. By definition there exists a point q in $\bigcap_{\substack{\psi\neq 1\\ \psi\neq\phi}} E^+(p, \psi p)-R_p$. Let q^* be the unique point of intersection of the geodesic segment γ_{pq} with $E(p, \phi p)$. We claim that $q^*\in E(p, \phi p)\cap R_p$. For every ψ in D the set $E^+(p, \psi p)$ is starshaped relative to p, and any geodesic from p that meets $E(p, \psi p)$ leaves $E^+(p, \psi p)$ after intersecting $E(p, \psi p)$. If $\psi\neq\phi$ it follows that $q^*\in E^+(p, \psi p)$ but not in $E(p, \psi p)$ since $q\in E^+(p, \psi p)$. Since $q^*\in E(p, \phi p)\subseteq E^+(p, \phi p)$ by the choice of q^* it follows that $q^*\in R_p\cap E(p, \phi p)$.

We show that R_p contains a set U such that $q^*\in U\subseteq E(p, \phi p)$ and U is open in $E(p, \phi p)$. If this were not the case, then we could find a sequence of points $q_n^*\subseteq(H-R_p)\cap E(p, \phi p)$ that converges to q^*. Therefore we could find a sequence $\phi_n\subseteq D$ such that $\phi_n\neq\phi$ and $d(\phi_n p, q_n^*)<d(p, q_n^*)$ for every n. There are only finitely many distinct elements ϕ_n by the proper discontinuity of D since the points $\phi_n(p)$ are a bounded sequence in H. Passing to a subsequence we may assume that $\phi_n=\psi\neq\phi$ for every n. Since $E(p, \psi p)$ is closed $q^*\in E(p, \psi p)$, which contradicts the fact proved above that $q^*\notin E(p, \psi p)$ if $\psi\neq\phi$. Since $E(p, \phi p)\cap R_p$ contains more than one point it consists of an entire subarc of R_p by the previous lemma.

Finally we show that no interior point of $E(p, \phi p)\cap R_p$ can be a vertex of R_p. Suppose that q is a vertex of R_p and also an interior point of $E(p, \phi p)\cap R_p$ for some bounding side $E(p, \phi p)$. By the definition of vertex there exists an element $\psi\neq\phi$ in D such that $E(p, \psi p)$ is a bounding side for R_p and $q\in E(p, \psi p)$. Relative to canonical parametrizations α, β for $E(p, \phi p)$, $E(p, \psi p)$ we can write $q=\alpha(t_0)=\beta(s_0)$. There exists by hypothesis an $\varepsilon>0$ such that $\alpha t\in E(p, \phi p)\cap R_p$ for $|t-t_0|<\varepsilon$. Since

$E(p, \psi p) \cap R_p$ is an arc there exists some $\delta > 0$ such that $\beta(s) \in R_p$ for all $s \in [s_0 - \delta, s_0]$ or all $s \in [s_0, s_0 + \delta]$. Hence there exists $t \neq t_0$ such that $|t - t_0| < \varepsilon$ and $\gamma_{p\alpha t}$ intersects $E(p, \psi p) \cap R_p$ in a point q^*. However $\alpha t \in \delta R_p$ also and since any geodesic from p meets δR_p at most once it follows that $\alpha t = q^*$. This implies that both q and q^* lie in $E(p, \phi p) \cap E(p, \psi p)$, contradicting proposition 2.8.

Proof of proposition 2.12. We show first that if $q \in E(p, \phi p) \cap R_p$ then $\phi^{-1} q \in R_p \cap E(p, \phi^{-1} p)$. Let such a point q be given. Clearly $\phi^{-1} q \in E(p, \phi^{-1} p)$ so it suffices to show that $\phi^{-1} q \in R_p$. For any nonidentity element ψ in D we know that $d(\psi p, \phi^{-1} q) = d(\phi \psi p, q) \geqslant d(p, q) = d(\phi p, q) = d(p, \phi^{-1} q)$. Therefore $\phi^{-1} q \in R_p$, and this implies that $\phi^{-1} \{E(p, \phi p) \cap R_p\} \subseteq E(p, \phi^{-1} p) \cap R_p$. Reversing the roles of ϕ and ϕ^{-1} we see that $\phi \{E(p, \phi^{-1} p) \cap R_p\} \subseteq E(p, \phi p) \cap R_p$, which implies that $E(p, \phi^{-1} p) \cap R_p \subseteq \subseteq \phi^{-1} \{E(p, \phi p) \cap R_p\}$ and proves that $E(p, \phi^{-1} p) \cap R_p = \phi^{-1} \{E(p, \phi p) \cap R_p\}$. The set $E(p, \phi p) \cap R_p$ is an arc by the preceding result, hence $E(p, \phi^{-1} p) \cap R_p$ is an arc. This implies that $E(p, \phi^{-1} p)$ is a bounding side for R_p, again by the previous result.

Proof of proposition 2.13. Let q be a vertex of R_p and suppose that q lies on three distinct bounding sides $L_1 = E(p, \phi_1 p)$, $L_2 = E(p, \phi_2 p)$ and $L_3 = E(p, \phi_3 p)$. Fix an orientation of the tangent space $T_p(H)$, and set $v = V(p, q)$. For each positive number ε we let $B_\varepsilon^+(v) = \{w \in T_p(H) : \|w\| = 1$ and $0 \leqslant \sphericalangle(v, w) < \varepsilon\}$ and $B_\varepsilon^-(v) = \{w \in T_p(H) : \|w\| = 1$ and $-\varepsilon < \sphericalangle(v, w) \leqslant 0\}$. By $\sphericalangle(v, w) > 0$ (respectively < 0) we mean that the pair $\{v, w\}$ is positively (respectively negatively) oriented relative to the given orientation of $T_p(H)$. Since L_i is a bounding side of R_p for each i the point q is an endpoint of some arc β_i contained in $L_i \cap R_p$. Therefore for each $i = 1, 2, 3$ we can find a number $\varepsilon_i > 0$ such that one of the following two possibilities occurs:

i) For any vector $w \in B_{\varepsilon_i}^+(v)$, γ_w intersects $L_i \cap R_p$.

ii) For any vector $w \in B_{\varepsilon_i}^-(v)$, γ_w intersects $L_i \cap R_p$.

Since we have three bounding sides L_1, L_2, L_3 we can find an $\varepsilon > 0$ such that one of the half neighborhoods of v, say $B_\varepsilon^+(v)$, corresponds to two of the bounding sides. Denote these sides by L and L'. Now $L \cap L' = \{q\}$ by proposition 2.8 so that if $w \in B_\varepsilon^+(v)$ is not equal to v, then γ_w meets $L \cap R_p$ and $L' \cap R_p$ in distinct points r and r'. The points r and r' both lie in ∂R_p. However, any geodesic γ from p intersects ∂R_p in at most one point q, for if $q \in E(p, \phi p)$ for some $\phi \neq 1$ in D, then all points on γ_{pq} beyond q lie in $H - E^+(p, \phi p)$. We have obtained a contradiction to our assumption that q lies in three bounding sides of R_p.

Proof of proposition 2.14. Since $q \in \partial R_p$, q lies in some bounding side $E(p, \xi p)$, $\xi \in D$, by proposition 2.9. One of the elements $\{\phi, \psi\}$, say ϕ, is not equal to ξ. If $E(p, \phi p)$ is a bounding side of R_p, then we are done so we may suppose that $E(p, \phi p)$ is not a bounding side of R_p. By proposition 2.11 $E(p, \phi p) \cap R_p$ is the single point q. Let α be the canonical parametrization of $E(p, \phi p)$, and let $f : H \to \mathbf{R}$ be the function $r \to d(\xi p, r) - d(p, r)$. Now $q = \alpha(t_0)$ for some number t_0, and hence $(f \circ \alpha) = 0$. Since $\xi p \neq \phi p$ lemma 2.8a implies that $f \circ \alpha$ is nonzero at any relative maximum or minimum point. Therefore $(f \circ \alpha)$ has no relative maximum or minimum at t_0, and lemma 2.8b implies that $f \circ \alpha$ is strictly monotone in some neighborhood U of t_0. Therefore either

i) $(f \circ \alpha)(t) > 0$ for $t > t_0$, $t \in U$ or

ii) $(f \circ \alpha)(t) > 0$ for $t < t_0$, $t \in U$.

Without loss of generality we may assume that i) occurs. Let t_n be any sequence of numbers such that $t_n > t_0$ for each n and $t_n \to t_0$ as $n \to \infty$. If $q_n = \alpha(t_n)$, then $q_n \in H - R_p$ since $E(p, \phi p) \cap R_p = q = \alpha(t_0)$. By assumption $(f \circ \alpha)(t_n) = d(\xi p, q_n) - d(p, q_n) > 0$ for large n. The geodesic segment γ_{pq_n} meets ∂R_p in a point r_n, and by the triangle inequality $d(p, r_n) \leqslant d(p, r_n) + d(r_n, q_n) + d(r_n, \xi p) - d(q_n, \xi p) = d(p, q_n) - d(\xi p, q_n) + d(r_n, \xi p) < d(\xi p, r_n)$. Hence r_n does not lie in $E(p, \xi p)$ for large n. Since r_n is a bounded sequence in the boundary of R_p, proposition 2.9 and the proper discontinuity of D imply that by passing to a subsequence we can find an element $\psi \neq \xi$ in D such that $r_n \in E(p, \psi p)$ for all n and $E(p, \psi p)$ is a bounding side of R_p. Passing to a further subsequence we may assume that r_n converges to a point r in $E(p, \psi p)$. Since r_n lies on the geodesic γ_{pq_n} for every n it follows that r lies on the geodesic γ_{pq}. Hence γ_{pq} meets ∂R_p at both q and r, and this implies that $r = q$ since a geodesic starting at p can meet ∂R_p at most once. Therefore q lies on the distinct bounding sides $E(p, \psi p)$ and $E(p, \xi p)$, which by definition means that q is a vertex of R_p.

REFERENCES

[1] Ahlfors L. and Sario, L., *Riemann Surfaces*, Princeton University Press, Princeton, New Jersey, 1960.

[2] Bleecker, D., *Ph.D. thesis*, University of California, Berkeley, 1973.

[3] Bishop, R. and Crittenden, R., Geometry of Manifolds, Academic Press, New York, New York 1964, pp. 237–241.

[4] Busemann, H., *The Geometry of Geodesics*, Academic Press, New York, New York, 1955.

[5] Eberlein, P., *Geodesics and Ends in Certain Surfaces Without Conjugate Points*, submitted.

[6] Grant, A., *Surfaces of Negative Curvature and Permanent Regional Transitivity*, Duke J. Math. *5* (1939), 207–229.

[7] Greenberg, L., *Fundamental Polygons for Fuchsian Groups*, Journ. d'Analyse Math., Jerusalem, 1967.

[8] ——, *Fundamental Polyhedra for Kleinian Groups*, Annals of Math., *84* (3), (1966), 433–441.

[9] v. Kerekjarto, B., *Vorlesungen über Topologie*, Springer Verlag, Berlin 1923.

[10] Lefschetz, S., *Algebraic Topology*, Amer. Math. Soc. Colloq., 27, Appendix B, New York, New York, 1942.

[11] Marden, A., *On Finitely Generated Fuchsian Groups*, Comment. Math. Helv. *42* (2), (1967), 81–85.

[12] ——, *The Geometry of Finitely Generated Kleinian Groups*, Annals of Math., *99* (1974), 383–462.

Received May 6/August 13, 1975