Mostow Rigidity

Rich Schwartz

September 24, 2013

1 The Proof Modulo some Analysis

1.1 Statement of the Result

Given any metric space (X, d), a map $H : X \to X$ is *BL* (bi-lipshitz) if there is some constant $K > 0$ such that

$$d(H(x), H(y)) \in [K^{-1}, K]\ d(x, y), \qquad \forall x, y \in X. \tag{1}$$

Here is a limited form of Mostow Rigidity:

Theorem 1.1 (Mostow) *Suppose that M_1 and M_2 are both compact hyperbolic 3-manifolds. If there is a BL map $f : M_1 \to M_2$ then there is an isometry $g : M_1 \to M_2$.*

One can press on the proof to yield the stronger statement that f and g are homotopic maps. Also, if one is willing to work with quasi-isometries in place of BL maps, one can just assume that the map f is a homotopy equivalence. I'll leave these matters to the interested reader. Note that if f is a diffeomorphism then f is automatically BL. Hence

Corollary 1.2 *If two closed hyperbolic 3 manifolds are diffeomorphic then they are isometric. Hence, the hyperbolic structure on a compact hyperbolic 3-manifold is unique.*

My proof is, in a certain sense, the standard one, but I figured out a good way to do it which relies on a lot less real analysis. The rest of this chapter assembles the ingredients of the proof and the last section of the chapter puts them together. The second and third chapters do the needed analysis.

1.2 QC Maps

Let $S = \boldsymbol{C} \cup \infty$ be the Riemann sphere. We often think of S as the ideal boundary of $\boldsymbol{H}^3$, hyperbolic 3-space.

Precompact Triples: $\{(a_n, b_n, c_n)\}$ denote a sequence of triples of points in S. We call this sequence *precompact* if there are pairwise disjoint compact subsets $A, B, C \subset S$ such that $a_n \in A$ and $b_n \in B$ and $c_n \in C$ for all n. We say that a sequence $\{h_n\}$ of homeomorphisms of S is *normalized* if there is a precompact sequence $\{(a_n, b_n, c_n)\}$ such that the sequence $\{(h_n(a_n), h_n(b_n), h_n(c_n))\}$ is also precompact.

Auxilliary Sequences: Given a single homeomorphism $h : S \to S$, we say that an *auxilliary sequence based on* h is a sequence of the form $\{h_n\}$ where

$$h_n = f_n h g_n, \qquad f_n, g_n \in PSL_2(\boldsymbol{C}). \tag{2}$$

QC Maps: A homeomorphism $h : S \to S$ is QC if every normalized auxilliary sequence $\{h_n\}$ based on h converges, uniformly on a subsequence, to a homeomorphism. We call such sequences *subconvergent*.

Asterisks: A point $z \in \boldsymbol{C}$ is an *asterisk* for h if the directional derivative $D_v h(z)$ exists and is nonzero for every rational direction v.

The following results are proved in §2 and §3 respectively.

Theorem 1.3 *A BL map of $\boldsymbol{H}^3$ extends continuously to a QC map of S.*

Theorem 1.4 *Every QC map of S has an asterisk.*

Remarks:
(i) My definition of QC maps is nonstandard, though certainly known. The Disk Theorem below relates my definition to the standard definition.
(ii) In e.g. Lehto and Virtanen, one proves the much stronger result that a QC map is a.e. nonsingularly differentiable. However, the proof of Mostow Rigidity here avoids the need for this hard analytic result.
(iii) Many readers will recognize that Theorem 1.3 remains true if H is just a quasi-isometry.

1.3 Zooming

Let h be a QC map. We write $h \to h'$ if there is an auxilliary sequence $\{h_n\}$ based on h such that $h_n \to h'$ on a subsequence. We insist that $h_n(\infty) = \infty$ as well. An easy exercise ("the diagonal trick") shows that h' is also QC. We write $h \Rightarrow h'$ if we have a finite chain $h = h_0 \to h_1 ... \to h_m = h'$.

We say that a line L is *good* for h if $h(L)$ is a straight line. We say that a direction is good for h if every line in that direction is good for h. We make the following easy observation. Suppose a direction D is good for h and $h \to h'$. Then D is also good for h'.

Lemma 1.5 *Suppose that k directions are good for h then $h \Rightarrow h_2$, where $k+1$ directions are good for h_2.*

Proof: We rotate so that the horizontal direction is not one of the k good directions for h. Theorem 1.4 tells us that h has an asterisk. Translating, we can assume that 0 is an asterisk for h. We define

$$h_n(z) = h(nz)/n. \tag{3}$$

The constant sequence $\{(-1,0,1)\}$ is precompact. Furthermore, the image sequence $\{(h_n(-1), h_n(0), h_n(1)\}$ is also precompact because it converges to $(-z, 0, z)$, where $z = \partial_x h(0)$. Hence $h \to h_1$ for some h_1.

Let L be some rational line through the origin. By the definition of differentiability, $h_1|L$ is multiplication by the directional derivative in the direction of L. In particular, $h_1(L)$ is a line. Hence, all rational lines through the origin are good for h_1. But then, by continuity, all lines through the origin are good for h_1.

Define

$$g_n(z) = h_1(z-n) - h_1(-n). \tag{4}$$

By construction $g_n(0) = 0$ and the action of g_n near 0 looks like the action of h_1 near n. When n is large, a large neighborhood of n is foliated by nearly horizontal lines which are good for h_1. Hence, when n is large, a large neighborhood of 0 is foliated by nearly horizontal lines which are good for g_n. Note that the restriction of g to the x-axis is a linear map. Hence g_n is the same linear map on the x-axis for all n. Hence $\{g_n\}$ is subconvergent. By construction the horizontal direction is good Hence $h_1 \to h_2$, and the horizontal direction is good for h_2. From our observation above, the other k directions remain good for h_2. ♠

Lemma 1.6 *Suppose that $f : \boldsymbol{R}^2 \to \boldsymbol{R}^2$ is a homeomorphism and 3 directions are good for f. Then f is real affine.*

Proof: We compose with affine maps so that the directions $(1,0)$, $(0,1)$ and $(1,1)$ are good and indeed preserved by f. Say that a *special square* is a square with sides parallel to the coordinate axes. The conditions on f guarantee that f maps special squares to special squares. But then f is a similarity. ♠

Corollary 1.7 *Suppose that h is QC. Then $h \Rightarrow h_6$ where h_6 is real affine, orientation preserving, and fixes 0 and 1.*

Proof: Applying Lemma 1.5 three times, we get $h \Rightarrow h_6$, where three directions are good for h_6. Lemma 1.6 says that h_6 is real affine. Now we can normalize by similarities to get the remaining properties. ♠

1.4 Hausdorff Convergence

Let X be any metric space. Given compact $K_1, K_2 \subset X$, we define $d(K_1, K_2)$ to be the infimal ϵ such that K_1 is contained in the ϵ neighborhood of K_2, and vice versa. This notion of distance is known as *Hausdorff distance*. We say that a sequence of closed subsets $\{S_n\}$ in X converges to $S \subset X$ if, for any compact $K \subset X$, we have $d(S_n \cap K, S \cap K) \to 0$. This notion of convergence is called *Hausdorff convergence*.

For our application, we will take X to be $PSL_2(\boldsymbol{C})$ equipped with any left-invariant metric and we will consider subsets $\Gamma \subset PSL_2(\boldsymbol{C})$ where Γ is a lattice acting on $\boldsymbol{H}^3$ with compact quotient. One basic fact we use is that the sequence of lattices $f_n \Gamma f_n^{-1}$ converges in the Hausdorff topology, on a subsequence, no matter how $f_n \in PSL_2(\boldsymbol{C})$ are chosen. Moreover, the limit lattice is conjugate to Γ. The point simply is that we can always adjust by elements of Γ so that $f_n \Gamma f_n^{-1} = f_n' \Gamma (f_n')^{-1}$ where when $\{f_n'\}$ is bounded.

1.5 Equivariance

Suppose that Γ_1 and Γ_2 are hyperbolic lattices acting with compact quotient. We say that a hmeomorphism F (on a suitable domain) is (Γ_1, Γ_2)-*equivavariant* if $F \Gamma_1 F^{-1} = \Gamma_2$.

Lemma 1.8 *If h is (Γ_1, Γ_2)-equivariant and $h \to h'$, then h' is (Γ'_1, Γ'_2)-equivariant, where Γ'_j is conjugate to Γ_j.*

Proof: Let $\{h_n\}$ be a convergent auxilliary sequence based on h. Then h_n is equivariant with respect to $(\Gamma_{1,n}, \Gamma_{2,n})$ such that $\Gamma_{j,n}$ is conjugate to Γ_j. From the remarks on Hausdorff convergence above, we may pass to a subsequence so that $\Gamma_{j,n} \to \Gamma'_j$, a lattice conjugate to Γ_j. Let $g'_1 \in \Gamma'_1$ be some element. There is some $g_{1,n} \in \Gamma_{1,n}$ such that $g_{1,n} \to g'_n$. Let $g_{2,n} = hg_{1,n}h^{-1}$. Since everything in sight converges, the sequence $\{g_{2,n}\}$ converges to $g'_2 \in \Gamma'_2$, and $g'_2 = h'g'_1(h')^{-1}$, as desired. Hence $h'\Gamma'(h')^{-1} \subset \Gamma_2$. Reversing the roles of the two lattices, we get equality in the last containment. ♠

Lemma 1.9 *The map h_6 in Corollary 1.7 is the identity if it is equivariant.*

Proof: Choose $g_1 \in \Gamma_1$ which does not fix ∞ and let $g_2 = h_6 g_1 h_6^{-1}$. We have $h_6 = g_2^{-1} h_6 g_1$. Let L_0 be some line so that $L_1 = g_1(L_0)$ is a circle. If h_6 is not the identity then $L_2 = h_6(L_1)$ is a non-circular ellipse and $L_3 = g_2^{-1}(L_2)$ is not a line. But $L_3 = h_6(L_0)$ is a line. Contradiction. ♠

1.6 The Main Argument

Suppose $f : M_1 \to M_2$ is BL. Let Γ_j be the deck group for M_j. The map f has a bilipshitz (Γ_1, Γ_2)-equivariant lift $H : \boldsymbol{H}^3 \to \boldsymbol{H}^3$. The QC extension h guaranteed by Theorem 1.3 is likewise equivariant. We have $h \Rightarrow h_6$ where h_6 is as in Corollary 1.7. The map h_6 is (Γ'_1, Γ'_2)-equivariant (Lemma 1.8) with respect to lattices which are conjugate to the originals. Hence, by Lemma 1.9, h_6 is the identity. Since the identity map is (Γ'_1, Γ'_2)-equivariant, $\Gamma'_1 = \Gamma'_2$. Hence Γ_1 and Γ_2 are conjugate. Hence M_1 and M_2 are isometric.

2 Boundary Extensions

2.1 Bi-Lipschitz Paths

This chapter is devoted to proving Theorem 1.3. Let $H : \boldsymbol{H}^3 \to \boldsymbol{H}^3$ be a BL map.

Lemma 2.1 *Let γ_1 be a geodesic in $\boldsymbol{H}^3$. Then there is a unique geodesic γ_2 such that $H(\gamma_1)$ stays within a bounded neighborhood of γ_2. The size of the neighborhood only depends on the BL constant of H.*

Proof: Uniqueness follows immediately from existence and from divergence properties of hyperbolic geodesics. So, it suffices to prove existence. Existence for the infinite geodesic follows from existence (with a uniform constant) for geodesic segments.

Let K_1 be the BL constant. Suppose that α_1 is a segment of γ_1 and consider $\beta = H(\alpha_1)$. Let γ_2 be the geodesic through the endpoints of β. We want to see that β lies in a uniformly small tubular neighborhood of γ_2. We will assume that this is false and derive a contradiction.

Let N_r denote the r-neighborhood of γ_2. Assume $r > 2$. Let $R = r^2$. Suppose that β exits N_R. Then there are points $p, q \in \beta \cap N_r$ such that the arc β' of β joining p to q remains outside N_r and exits N_R. The length L of β' is at least $2R - 2r$, which exceeds R. In short $L > r^2$.

Let $\pi : \boldsymbol{H}^3 - N_r \to \gamma_2$ be the radial contraction. Thanks to the exponentially divergent nature of geodesics in hyperbolic space, there is some constant K_2 so that π decreases distances by a factor of at least $K_2 \exp(-r)$. The length of $\pi(\beta')$ is at most $K_2 \exp(-r)L$. Hence, we can join p to q by heading directly towards γ, using $\pi(\beta')$, then going directly out to q. This path has length less than $K_2 \exp(-r)L + 2r$.

Since β is a K_1 bi-lipschitz path, our new path cannot be more than K_1^2 times shorter than β'. That is

$$K_2 \exp(-r)L + 2r \geq \frac{L}{K_1^2}. \tag{5}$$

Since $L > r^2$, this is false for large r. ♠

2.2 The Extension

We use the ball model for $\boldsymbol{H}^3$ and normalize H so that H fixes the origin. Let $p \in S$ be some point on the boundary. Let γ_1 be the geodesic ray connecting the origin to p. We define $h(p)$ to be the relevant endpoint of γ_2, where γ_2 is such that $H(\gamma_1)$ stays within a bounded neighborhood of γ_2.

Lemma 2.2 *The extension map h is a homeomorphism.*

Proof: Firt we show that h is continuous. Let $\{p_n\}$ be a sequence in S which converges to p. Let $\gamma_{1,n}$ be the geodesic ray connecting the origin to p_n. Clearly $\gamma_{1,n} \to \gamma_1$. Let $\gamma_{2,n}$ relate to $\gamma_{1,n}$ as γ_2 relates to γ_1. Let $K_3 = 2K_2 + K_1$, where K_1 is the BL constant of H and K_2 is the constant from Lemma 2.1. By Lemma 2.1 and the fact that $\gamma_{1,n} \to \gamma_1$ on compact subsets, there is some M_n such that the K_3-neighborhood of $\gamma_{2,n}$ contains at least M_n units of γ_2, and $\lim M_n = \infty$. This forces $\gamma_{2,n} \to \gamma_2$. Hence $h(p_n) \to h(p)$.

To show that h is a homeomorphism, we re-run the same argument for H^{-1} to show that h^{-1} exists and is continuous. ♠

Lemma 2.3 *The extension h is QC.*

Proof: Let $\{h_n\}$ be a normalized auxilliary sequence based on h. We have precompact sequences $\{(a_n, b_n, c_n)\}$ and $\{(h_n(a), h_n(b_n), h_n(c_n))\}$. Now, passing to a subsequence and then adjusting by a convergent sequence in $PSL_2(\boldsymbol{C})$, we can assume that h_n fixes three points a, b, c independent of n.

By construction h_n is the extension of some BL map H_n. Given Lemma 2.2, it suffices to prove that H_n converges to some new BL map of $\boldsymbol{H}^3$. Let Γ be the ideal triangle in $\boldsymbol{H}^3$ whose endpoints are a, b, c. For any K the set of points which are within K units of points on all 3 sides of Γ is bounded in $\boldsymbol{H}^3$. Let O be any point of $\boldsymbol{H}^3$. By Lemma 2.1, and the triangle inequality, there is a constant K such that $H_n(O)$ remains K units from points on all 3 sides of Γ. Hence $H_n(O)$ lies in a compact subset of $\boldsymbol{H}^3$, independent of n. But this property, together with the uniform BL constant, implies that $\{H_n\}$ converges on a subsequence. ♠

3 QC Maps

3.1 Reduction to a Technical Lemma

In this section we reduce Theorem 1.4 to the following lemma, which we then prove in subsequent sections.

Lemma 3.1 *Let Q be the unit square. Then there is a full measure subset $\Omega_1 \subset Q$ and a positive measure subset $\Omega_2 \subset \Omega_1$ such that the partial derivative dh/dx exists on Ω_1 and is nonzero on Ω_2.*

Corollary 3.2 *There is a full measure subset Ω_3 of the plane such that every rational directional derivative exists in Ω_3.*

Proof: Applying Lemma 3.1 to the map $z \to h(z-v)$ we see that the unit square with vertex v also has a full measure subset on which dh/dx exists. Letting v range through all the Gaussian integers, we get see that dh/dx exists on a full measure subset of $\boldsymbol{C}$. By symmetry, the directional derivative in any given direction exists on a full measure subset of $\boldsymbol{C}$. The intersection Ω_3 of the countably many subsets corresponding to the rational directions has the desired properties. ♠

Corollary 3.3 *There exists a point $z \in \boldsymbol{C}$ such that the directional derivative of h at z exists in all rational directions and $dh/dx(z) \neq 0$.*

Proof: We just take any point in the intersection $\Omega_2 \cap \Omega_3$. This intersection is nonempty because Ω_2 has positive measure and Ω_3 has full measure. ♠

We normalize by a complex affine map so that $z = 0$ and $h(0) = 0$ and $dh/dx(0) = 1$. Here, of course, we are using complex notation, so that "1" means the unit vector in the positive real direction.

Lemma 3.4 *$D_v h(0) \neq 0$ for any rational direction v.*

Proof: For ease of exposition, we will show that $dh/dy(0) \neq 0$. The general case is the same. We will assume the contrary and derive a contradiction. Let $h_n(z) = nh(z/n)$. By the definition of one-dimensional derivatives, the triple $\{(h_n(-1), h_n(0), h_n(1))\}$ converges to $(-1, 0, 1)$. Hence $\{h_n\}$ is subconvergent. But then $\{h_n(i)\}$ converges to some nonzero point in $\boldsymbol{C}$. This is not compatible with $dh/dy(0) = 0$. ♠

3.2 Absolute Continuity on Lines

The rest of these notes are devoted to proving Lemma 3.1. We use the notation from Lemma 3.1.

Suppose $f : [0,1] \to \boldsymbol{R}$ is a continuous map. Let $I = \{I_1, ..., I_n\}$ denote a finite list of intervals of $[0,1]$ having pairwise disjoint interiors. We call I a *partial partition*. We define $|I| = \sum |I_k|$, the total length of I. We define I'_k to be the interval bounded by the endpoints of $f(\partial I_k)$. We define $I' = \{I'_1, ..., I'_n\}$ and $|I'| = \sum |I'_k|$. Note that the intervals in I' might overlap. Here is a slightly nonstandard definition of what is meant by maps which are absolutely continuous in measure.

Definition: The map f is AC if the following property holds. For all $\epsilon > 0$ there is some $\delta > 0$ such that $|I| < \delta$ implies that $|I'| < \epsilon$.

The following result relates our definition of QC maps to the standard definition.

Theorem 3.5 (Disk) *Let h be a QC map. There is a constant K, depending only on h, with the following property. Let Δ be any disk such that $h(\Delta) \subset \boldsymbol{C}$. There are disks D_1, D_2 so that $D_1 \subset h(\Delta) \subset D_2$ and* $\text{diam}(D_2)/\text{diam}(D_1) < K$.

Proof: Suppose $\{\Delta_n\}$ is a sequence where the best ratio for $h(\Delta_n)$ tends to 0. Composing with Mobius transformations, we get an auxilliary sequence $\{h_n\}$ such that the best ratio tends to 0 on $h_n(\Delta)$, where Δ is unit disk, and $h_n(\pm 1) = \pm 1$. Note that $\{h_n\}$ cannot be subconvergent.

There is some point $u_n \in \Delta$ such that $|h_n(u_n) - \pm 1| \in [1,3]$. Let g_n be the Mobius transformation mapping $(-1, u_n, 1)$ to $(-1, i, 1)$. Let $\widehat{h}_n = h_n \circ g_n^{-1}$. Note that $g_n(\Delta) = \Delta$, and hence $\widehat{h}_n(\Delta) = h_n(\Delta)$. Hence, we can replace $\{h_n\}$ by $\{\widehat{h}_n\}$ and we still have sequence of counterexamples. But this latter sequence is normalized and hence subconvergent. This is a contradiction. ♠

The following idea is in Lehto and Virtanen's book. Suppose that Q is the unit square and $h : Q \to \boldsymbol{C}$ is QC. For each $y \in [0,1]$ let $A(y)$ denote the area of $h([0,1] \times [0,y])$. The function A is monotone and hence almost everywhere differentiable. Let $L_y \subset Q$ be the horizontal segment of height y.

Theorem 3.6 *Suppose that $A'(y)$ exists. Let $\pi : \boldsymbol{R}^2 \to \boldsymbol{R}$ be any linear projection. Then $\pi \circ h$ is AC on L_y.*

Proof: Let $Q(\epsilon) = [0,1] \times [y-\epsilon, y+\epsilon]$. This set has area ϵ. Set $K_1 = 4|A'(y)|$. Once ϵ is sufficiently small,

$$\text{area}(h(Q(\epsilon)) = A(y+\epsilon) - A(y-\epsilon) < K_1\epsilon. \tag{6}$$

Let $f = \pi \circ h|L_y$. Suppose f is not AC. Then we can scale the picture so that there is a sequence of partial partitions $\{I^n\}$ with $|I^n| < 1/n$ and $|I'_n| \geq 1$.

We fix n for now and let $I = I^n$ be one of these partial partitions. We can subdivide the partition (without destroying the basic property) so that all the intervals of I have the same size up to a factor of 2. We write $I = \{I_1, ..., I_k\}$. Here k depends on n in some way. Let $\epsilon = \max |I_j|$. Since $\epsilon < 2\min |I_j|$, we have $k\epsilon < 2|I|$. Hence

$$\epsilon < 2/(kn) \tag{7}$$

Let Δ_j be the disk having I_j as a diameter. We have $\Delta_j \subset H(\epsilon)$. Hence

$$\sum_{j=1}^{k} \text{area } h(\Delta_j) < \text{area}(h(Q(\epsilon)) < K_1\epsilon < \frac{2K_1}{kn}. \tag{8}$$

Let K_2 be the constant from the Disk Theorem. Let $D_{1,j}$ and $D_{2,j}$ be disks such that $D_{1,j} \subset h(\Delta_j) \subset D_{2,j}$ and $\text{diam}(D_{2,j})/\text{diam}(D_{1,j}) < K_2$. We have

$$\sum_{j=1}^{k} \text{diam}(D_{2,j}) \geq 1, \qquad \sum_{j=1}^{k} \text{diam}(D_{1,j}) \geq \frac{1}{K_2}. \tag{9}$$

The first equation, which comes from $|I'| \geq 1$, implies the second equation.

The sum of the areas of the disks $\{D_{1,j}\}$ would be minimized if they all had the same size, namely some common radius $r \geq 1/(2K_2k)$. Therefore,

$$\frac{2K_1}{kn} > \sum_{j=1}^{k} \text{area } h(Q_k) \geq \sum_{j=1}^{k} \text{area}(D_{1,j}) \geq k\pi r^2 \geq \frac{\pi}{4K_2^2 k}. \tag{10}$$

This is a contradiction for large n. ♠

3.3 Bounded Variation

Let $f : [0,1] \to \boldsymbol{R}$ be some continuous function. We say that f is BV (bounded variation) if there is some N such that $|I'| < N$ for all choices of partial partition I.

Lemma 3.7 *If f is AC, then f is BV.*

Proof: Suppose f is not BV. Then, for any N, there is a partition I such that $|I'| > N$. We can subdivide the intervals in I so that they all have the same length up to a factor of 2, and all have length less than $1/N$. This subdivision does not decrease $|I'|$. Now we can split I into partial partitions, all having total length between $1/N$ and $2/N$. There are at most N of these partial partitions. Hence, one of these partitial partitions J will be such that $|J'| > 1$. In summary, $|J| < 2/N$ and $|J'| > 1$. Hence f is not AC. ♠

It is well known that BV functions are a.e. differentiable. For the sake of completeness, I'll include the proof

Lemma 3.8 *If f is BV, then $f = f_+ - f_-$, where f_+ and f_- are monotone increasing functions. Hence f is a.e. differentiable.*

Proof: Let $s \in (0,1]$. Given a partition I of $[0,x]$, define $|I'|_\pm = \sum |I_k'|$, where the sum is taken over the intervals I_k such that $\pm f$ maps the endpoints of I_k into $\boldsymbol{R}$ in order. In case I is a partition of $[0,x]$, we have

$$f(x) = |I'|_+ - |I'|_-. \tag{11}$$

Define $f_\pm(x) = \sup |I'|_\pm$, where the supremum is taken over all partitions of $[0,x]$. These two functions are well defined (since f is BV) and monotone.

If we take a partition I and refine it, it does not decrease $|I'|_\pm$. For this reason, we can find a single partition I of $[0,x]$ such that $f_\pm(x) - |I'|_\pm < \epsilon$. Plugging this into Equation 11, we get $|f(x) - (f_+(x) - f_-(x))| < 2\epsilon$. But x and ϵ are both arbitrary. Hence $f = f_+ - f_-$. ♠

Corollary 3.9 *If f is AC, then f is a.e. differentiable.*

3.4 Nonzero Derivatives

Lemma 3.10 *Suppose $f : [0,1] \to \boldsymbol{R}$ is AC and $f(0) \neq f(1)$. Then there is some positive measure subset $C \subset [0,1]$ such that f' exists and is nonzero for all $x \in C$.*

Proof: We suppose that the conclusion is false and derive a contradiction. If $f' = 0$ a.e. then we have $[0,1] = A \cup B$ where B has measure 0 and $f'(x) = 0$ for all $x \in A$. Lemma 3.11 below shows that $f(B)$ has measure 0.

Each point $x \in A$ is the midpoint of an open interval $U = U_x$ such that $\text{diam}(f(U)) < \epsilon \ \text{diam}(U)$. We can choose U_x so that it has diameter 2^{-k}. So, $\{U_x\}$ is a nice cover of A, in the sense of Lemma 3.12 below. By Lemma 3.12, there is some subcover of total length at most 2. But then $f(A)$ has measure less than 2ϵ. Since ϵ is arbitrary, $f(A)$ has measure 0.

Since $f(A)$ and $f(B)$ both have measure 0, the connected set $f([0,1])$ has measure 0. But then $f([0,1])$ is a single point. This is a contradiction. ♠

Now we take care of the unfinished business in Lemma 3.10.

Lemma 3.11 *If f is AC and $B \subset [0,1]$ has measure 0, then $f(B)$ has measure 0.*

Proof: We can find an open set U having length less than δ such that $B \subset U$. The set U is a countable union of open intervals. Let U^n denote the union of the first n intervals of U.

There is a partition $I = I^n$ of closure(U^n) so that $f(U^n) \subset \bigcup I'_k$. The idea is as follows. For each compact connected component K of closure(U^n), we include in I^n an interval connecting a point of K where where $f|K$ takes on a minimum value to a point of K where $f|K$ takes on a maximum value.

By construction $|I^n| < \delta$. Choosing δ small enough, we can guarantee that $|(I^n)'| < \epsilon$. Setting $B_n = B \cap U^n$, we see that $f(B_n) \subset \bigcup_k (I^n_k)'$ has measure less than ϵ. But $f(B) = \bigcup f(B_n)$. Hence $f(B)$ has measure at most ϵ. But ϵ is arbitrary. ♠

Now we deal with the second half of the unfinished business. Say that a *nice cover* of $A \subset [0,1]$ is a cover in which every point of A is the midpoint of some interval in the cover, and each interval has length 2^{-k} for some k (depending possibly on the interval.)

Lemma 3.12 *Any nice cover has a subcover of total length at most* 2.

Proof: This is essentially the Besicovich Covering Lemma. Let I^* be the interval obtained by shrinking an interval I by a factor of 2 about its midpoint. We produce a sequence of intervals $I_1, I_2, ...$ as follows. Assuming that the first k of these intervals have been chosen, let I_{k+1} be any largest interval such that I^*_{k+1} is disjoint from I^*_j for $j = 1, ..., k$. This algorithm produces a countable (perhaps finite) list whose total length is at most 2.

We claim that $A \subset \bigcup I_j$. If not, let $x \in A$ be some uncovered point and let J be an interval in the original cover that has x as a midpoint. By construction, x is at least $|I^*_j|$ away from the endpoints of I^*_j. But there are only finitely many intervals $I_1, ..., I_k$ not smaller than J, and our algorithm would have placed J on the list right after I_k, a contradiction. ♠

3.5 The End of the Proof

Now we prove Lemma 3.1. Let $Y \subset [0,1]$ be the full measure set where the area function A is differentiable. Let $\pi : \boldsymbol{C} \to \boldsymbol{R}$ be any linear projection. Let $L_y \subset Q$ be the unit horizontal segment of height y. For each $y \in Y$, the map $\pi \circ h | L_y$ is AC and hence a.e. differentiable on L_y.

Using real notation, we write $h = (h_1, h_2)$. Taking π to be projection onto the real and imaginary axes, we see that $\partial h_j / \partial x$ exists at almost every point of almost every L_y. So, by Fubini's Theorem, $\partial h / \partial x$ exist a.e. in Q. This gives us Ω_1 in Lemma 3.1.

For any $y \in Y$, choose π so that $\pi \circ h$ does not identify the endpoints of L_y. Lemma 3.10 tells us that $\pi \circ h | L_y$ has nonzero derivative on a positive set $S_y \subset L_y$. But then $\partial h / \partial x$ is nonzero on S_y. The union $\Omega_2 = \bigcup_y S_y \cap \Omega_1$ has positive measure, by Fubini's Theorem, and is the set in Lemma 3.1.