PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 94, Number 2, June 1985

TOTAL CURVATURES AND MINIMAL AREAS OF COMPLETE OPEN SURFACES[1]

KATSUHIRO SHIOHAMA

ABSTRACT. Minimal areas for certain classes of finitely connected complete open surfaces are obtained by using a Bonnesen-style isoperimetric inequality for large balls on the surfaces. In particular, the minimal area of Riemannian planes whose Gaussian curvatures are bounded above by 1 is 4π.

Introduction. The present work was inspired by a fruitful paper of Gromov [**3**]. Throughout let M be a 2-dimensional connected, oriented, noncompact manifold without boundary. Let $\mathfrak{M}_0(M)$ be the set of all complete metrics on M such that, for every g in $\mathfrak{M}_0(M)$, the Gaussian curvature K_g with respect to g satisfies $|K_g| \leqslant 1$. Gromov proved (see [**3**, Appendix 1]) that the infimum of areas $A(R^2, g)$ over all $g \in \mathfrak{M}_0(R^2)$ is greater than $4\pi + 0.01$ and not greater than $(2 + 2\sqrt{2})\pi$. He also proved that if the Euler characteristic $\chi(M)$ of M is nonpositive, then $\inf_{g \in \mathfrak{M}_0(M)} A(M', g) = 2\pi|\chi(M)|$. Moreover, if $g \in \mathfrak{M}_0(M)$ satisfies $A(M, g) < \infty$, then the total curvature $c(M, g) = \int_M K_g\, dA_g = 2\pi\chi(M)$. Here dA_g denotes the area element of (M, g) with respect to g.

We want to provide a partial result for the minimal areas by using a Bonnesen-style isoperimetric inequality. Such an inequality was first shown by Fiala [**2**] for analytic metrics on R^2 and later by Hartman [**4**] for smooth metrics on R^2. The Hartman theorem (see Theorem 7.1 in [**4**]) applies to the distance function to a fixed point p on (R^2, g) as follows. For every $t > 0$ let $S(t) := \{x \in (R^2, g);\ d(p, x) = t\}$ and $B(t) := \{x \in (R^2, g);\ d(p, x) < t\}$, where d is the distance function induced from g. Let $A(t)$ be the area of $\overline{B}(t)$ and $L(t)$ the length of $S(t)$ ($S(t)$ becomes a piecewise smooth curve for almost all $t > 0$). If $\int_{R^2}|K_g|\, dv_g < \infty$, then

$$\lim_{t \to \infty} L^2(t)/A(t) = 2\big(2\pi - c(R^2, g)\big).$$

Now, M is called *finitely connected* if there exists a closed 2-manifold N and finite points $p_1, \ldots, p_m$ on N such that M is homeomorphic to $N - \{p_1, \ldots, p_m\}$. Such an M is said to have m endpoints.

An essential improvement of the Hartman theorem is obtained here by a thorough consideration of geometric significances on the existence of total curvature on a

Received by the editors April 19, 1984 and, in revised form, July 2, 1984.
1980 *Mathematics Subject Classification*. Primary 53C20.
Key words and phrases. Complete manifolds, Gaussian curvature, geodesics, isoperimetric inequality.

[1] Dedicated to Professor Wilhelm Klingenberg on his 60th birthday

finitely generated connected and complete (M, g). Indeed, the existence of $c(M, g)$ in $[-\infty, 2\pi\chi(M)]$ imposes strong restrictions on the distance function from a fixed point and on the topology of $S(t)$ for all sufficiently large t.

The crucial point of our improvement of the Gromov theorems is based on the following generalization of the Hartman theorem.

THEOREM A. *Let (M, g) be complete and finitely connected. Let p be a fixed point on M, and let $S(t) := \{x \in M;\ d(x, p) = t\}$ and $B(t) := \{x \in M;\ d(x, p) < t\}$. If (M, g) admits the total curvature, then*

$$\lim_{t \to \infty} \frac{L(t)}{t} = 2\pi\chi(M) - c(M, g)$$

and

$$\lim_{t \to \infty} \frac{A(t)}{2t^2} = 2\pi\chi(M) - c(M, g),$$

where $L(t)$ and $A(t)$ are the length of $S(t)$ and the area of $\overline{B}(t)$, respectively.

Note that the right side of these equations is nonnegative by a well-known theorem due to Cohn-Vossen [**1**]. It is not known whether Theorem A holds for infinitely connected M.

Theorem A was already proved in the simplest case where M is homeomorphic to R^2 (see [**7**, Theorem D]). A generalization of the Hartman theorem is a direct consequence of Theorem A, as stated: If M is finitely connected and if the total curvature of (M, g) exists, then

$$\lim_{t \to \infty} \frac{L^2(t)}{A(t)} = 2(2\pi\chi(M) - c(M, g)).$$

Furthermore, the following is a straightforward consequence of Theorem A.

COROLLARY. *If M is finitely connected, (M, g) is complete $A(M, g) < \infty$, and the total curvature of (M, g) exists, then $c(M, g) = 2\pi\chi(M)$.*

Note also that if M is not finitely connected and if the total curvature of complete (M, g) exists, then a well-known theorem due to Huber [**5**] states that $c(M, g) = -\infty$, and the Corollary holds in this case.

Our result on minimal areas is

THEOREM B. *For a finitely connected M let $\mathfrak{M}(M)$ be the set of all complete Riemannian metrics on M such that for each g in $\mathfrak{M}(M)$, $K_g \leqslant 1$ if $\chi(M) \geqslant 0$ and $K_g \geqslant -1$ if $\chi(M) < 0$. Then*

$$\inf_{g \in \mathfrak{M}(M)} A(M, g) = \begin{cases} 4\pi & \text{if } \chi(M) = 1, \\ 0 & \text{if } \chi(M) = 0, \\ 2\pi|\chi(M)| & \text{if } \chi(M) < 0. \end{cases}$$

However, it is not yet known that the minimal area for $\mathfrak{M}_0(R^2)$ is $(2 + 2\sqrt{2})\pi$, and this problem seems to be very hard.

1. The proof of Theorem A. The proof of Theorem A is obtained by the following Facts 1 and 2, which have been established by the author in [7]. Some notations are needed to state them.

Let (R^2, g) be complete and let $\mathfrak{C}$ be a smooth regular curve on R^2. Let M_1 be the domain with boundary $\mathfrak{C}$ and homeomorphic to the closed half cylinder $S^1 \times [0, \infty)$, and let $\rho\colon M_1 \to R$ be the distance function on M_1 to $\mathfrak{C}$, e.g.

$$\rho(x) := \inf\{d(x, y); y \in \mathfrak{C}\},$$

where d is the distance function on R^2 induced from g. Let $\tilde{S}(t) := \{x \in M_1; \rho(x) = t\}$ and $\tilde{B}(t) := \{x \in M_1; \rho(x) < t\}$. The cut locus $C(\mathfrak{C})$ of $\mathfrak{C}$ (in M_1) was completely determined by Hartmann [4] as follows: For almost t in $[0, \infty)$, $\tilde{S}(t)$ intersects $C(\mathfrak{C})$ at finite points $x_1(t), \ldots, x_k(t)$, and each point $x_i(t)$ is joined to $\mathfrak{C}$ by exactly two distinct minimizing geodesics with length t, along each geodesic of which $x_i(t)$ is not a focal point to $\mathfrak{C}$. Thus, $C(\mathfrak{C})$ forms a smooth curve in a small neighborhood of each $x_i(t)$, and $\tilde{S}(t)$ becomes a piecewise smooth regular curve. Such a t is called *nonexceptional.* The length $\tilde{L}(t)$ of $\tilde{S}(t)$ at each nonexceptional t is differentiable. However, $\tilde{L}(t)$ is not, in general, continuous. It is not known how many components of $\tilde{S}(t)$ there are.

The existence of total curvature of (R^2, g) makes it possible to show the continuity of $\tilde{L}(t)$ and the connectivity of $\tilde{S}(t)$ for all sufficiently large t. The existence of total curvature also makes it possible to prove a sharp estimate for the derivative of $\tilde{L}(t)$ for all sufficiently large nonexceptional t, and this estimate is required for the proof of Theorem A.

For a point $x \in M_1$ consider all minimizing geodesics joining x to points on $\mathfrak{C}$ having the same length $\rho(x)$. If there are at least two distinct minimizing geodesics joining x to points on $\mathfrak{C}$ with the same length $\rho(x)$, then there is a compact domain E_x in M_1 homeomorphic to a closed 2-disk which is bounded by the two geodesics with length $\rho(x)$ and the subarc $\mathfrak{C}$ having the same endpoints as theirs, where the two geodesics on the boundary of E_x are chosen in such a way that if σ is a minimizing geodesic with length $\rho(x)$ joining x to a point on $\mathfrak{C}$, then σ lies on E_x. If there is a unique minimizing geodesic from x to a point on $\mathfrak{C}$ with length $\rho(x)$, then E_x consists of all points on the geodesic. Let $\beta(x)$ be the angle at x between the two vectors tangent to the geodesics lying in the boundary of E_x which is measured with respect to E_x. $\beta(x) > 0$ if there are two distinct geodesics in the boundary of E_x, and $\beta(x) = 0$ otherwise. The following facts have been established in [7].

FACT 1. *There exists a $T > 0$ such that $\tilde{S}(t)$ is homeomorphic to a circle for all $t > T$.*

FACT 2. *For any positive ε there exists a $T_\varepsilon > 0$ such that if $t > T_\varepsilon$, then $\sum_{x \in \tilde{S}(t)} \beta(x) < \varepsilon$.*

THE PROOF OF THEOREM A. Choose a large number T such that T is a nonexceptional value and such that $M - B(T)$ has exactly m unbounded components. Let M_1' be a fixed unbounded component of $M - B(T)$. Then the boundary of M_1' is a piecewise smooth regular curve and homeomorphic to a circle. M_1' is homeomorphic to $S^1 \times [0, \infty)$. Since the angle at each nondifferentiable point on the boundary curve is less than π (measured with respect to M_1'), there exists a small positive number δ and a smooth regular curve $\mathfrak{C}$ in M_1' with the following properties:

(1) If $x_1(T),\dots,x_k(T)$ are all nondifferentiable points on the boundary curve of M_1' and if for each $i = 1,\dots,k$, $B_i(\delta)$ is an open δ-ball around $x_i(T)$, then $\partial M_1' - \bigcup_{i=1}^k B_i(\delta) \subset \mathfrak{C}$.

(2) If $M_1 \subset M_1'$ is the domain bounded by $\mathfrak{C}$ and homeomorphic to $S^1 \times [0,\infty)$, and if $\rho\colon M_1 \to R$ is the distance function to $\mathfrak{C}$, then $\rho(x) + T = d(p, x)$ for all $x \in M_1$ with $\rho(x) > 1$.

Property (2) is checked as follows. Since T is nonexceptional and the set of all nonexceptional values is open in $[0,\infty)$, there is a $T' > T$ such that every t in $[T, T')$ is nonexceptional. For every $t \in [T, T')$, $S(t) \cap C(p) \cap M_1'$ consists of k points $x_i(t),\dots,x_k(t)$, and, for each $i = 1,\dots,k$, $t \to x_i(t)$ is a smooth regular curve bisecting two minimizing geodesics joining p to $x_i(t)$. For each $t \in [T, T')$ and $i = 1,\dots,k$, let $\sigma_{t,i}$, $\tau_{t,i}$: $[0, t] \to M_1'$ be the two geodesics with $\sigma_{t,i}(0) = \tau_{t,i}(0) = p$ and $\sigma_{t,i}(t) = \tau_{t,i}(t) = x_i(t)$, and let

$$\delta(t) := \max\{ d(x_i(T), \sigma_{t,i}(T)), d(x_i(T), \tau_{t,i}(T)); i = 1,\dots,k\}.$$

δ is continuous and $\delta(T) = 0$, and, hence, there is a $t_0 \in [T, T')$ such that $\delta(t_0) =: \delta$ is less than the convexity radius of the compact set $\bar{B}(T') - B(T)$ and such that $t_0 - T < 1$. For each i let $\mathfrak{C}_i$ be a smooth regular curve joining $\sigma_{t_0,i}(T)$ to $\tau_{t_0,i}(T)$ such that it is tangent to $S(T)$ at its endpoints and contained entirely in $M_1' \cap B_\delta(x_i(T)) - B_{t_0 - T}(x_i(t_0))$, where $B_r(q)$ is the open metric r-ball around q. The desired curve $\mathfrak{C}$ in a neighborhood of each $x_i(T)$ is obtained as $\mathfrak{C}_i$. Let $y \in M_1$ with $\rho(y) > 1$. Let γ: $[0, \rho(y)] \to M_1$ be a minimizing geodesic joining y to a point z on $\mathfrak{C}$. Suppose $z \in \mathfrak{C}_i$ for some i. Then γ intersects at a point q on $\sigma_{t_0,i}([T, t_0])$ (or on $\tau_{t_0,i}([T, t_0])$). Since

$$\rho(x_i(t_0)) = t_0 - T = \text{length}\left(\sigma_{t_0,i}|[T, t_0]\right),$$

and since q is an interior of $\sigma_{t_0,i}([T, t_0])$, $\rho(q)$ realizes at a unique point $\sigma_{t_0,i}(T)$. This implies $\rho(q) < d(q, z)$ and, hence,

$$d(y, z) > d(y, q) + d\left(q, \sigma_{t_0,i}(T)\right) > d\left(y, \sigma_{t_0,i}(T)\right),$$

contradicting the choice of Z. Hence, $z \in \mathfrak{C} - \bigcup_{i=1}^k \mathfrak{C}_i$ and the triangle inequality implies $\rho(y) + T \geqslant d(p, y)$; the reversed inequality is obvious. This proves $\rho(y) + T = d(p, y)$ for y with $\rho(y) > 1$.

The above argument shows that if the distance function from p is restricted to $M_1 - B(T+1)$, then it is $\rho + T$, and Facts 1 and 2 can be applied to it. There exists a $T_0 > T + 1$ such that if $t > T_0$, then $S(t)$ has exactly m components and each component is homeomorphic to a circle, and $L(t)$ is continuous in t. Here the continuity of $L(t)$ follows from $\lim_{h \downarrow 0} S(t-h) = S(t) = \lim_{h \downarrow 0} S(t+h)$.

Now if $t > T_0$ is nonexceptional, then the derivative of $L(t)$ is given as follows. Set $c(\bar{B}(t)) := \int_{\bar{B}(t)} K_g \, dA_g$.

$$\frac{dL(t)}{dt} = 2\pi\chi(M) - c(\bar{B}(t)) - \sum_{x \in S(t)} \left[2 \tan \frac{\beta(x)}{2} - \beta(x)\right].$$

It follows from Fact 2 and the finite connectivity of M that if $\varepsilon > 0$ is arbitrarily given, then there is a $T(\varepsilon)'$ such that

$$\sum_{x \in S(t)} \left[2 \tan \frac{\beta(x)}{2} - \beta(x) \right] < \varepsilon$$

holds for all $t > T(\varepsilon)'$, and hence

$$2\pi\chi(M) - c(\bar{B}(t)) - \varepsilon \leqslant \frac{dL(t)}{dt} \leqslant 2\pi\chi(M) - c(\bar{B}(t)).$$

On the other hand, the area $A(t)$ of $\bar{B}(t)$ is given as

$$A(t) - A(T) = \int_T^t L(u)\, du.$$

If $c(M, g) = -\infty$, then the limit of the derivative of $L(t)$ as $t \to \infty$ is $-\infty$, and the proof of Theorem A in this case is obvious by L'Hospital's theorem. If $c(M, g) > -\infty$, then for any $\varepsilon > 0$ there is a $T(\varepsilon)''$ such that $|c(\bar{B}(t)) - c(M, g)| < \varepsilon$ for all $t > T(\varepsilon)''$. The proof is completed by the following inequalities:

$$2\pi\chi(M) - c(M, g) - 2\varepsilon \leqslant \lim_{t \to \infty} \frac{L(t)}{t} \leqslant 2\pi\chi(M) - c(M, g) + \varepsilon,$$

$$2\pi\chi(M) - c(M, g) - 2\varepsilon \leqslant \lim_{t \to \infty} \frac{A(t)}{2t^2} \leqslant 2\pi\chi(M) - c(M, g) + \varepsilon.$$

2. The proof of Theorem B. It should be noted that if $A(M, g) = \infty$ for some $g \in \mathfrak{M}(M)$, then either the total curvature of (M, g) does not exist, or else $c(M, g) < 2\pi\chi(M)$. This fact is an immediate consequence of Theorem A.

The proof of Theorem B in the case $\chi(M) < 0$ is clear from the following inequalities: Let $K_g^- := \min\{K_g, 0\}$. Then

$$c(M, g) \geqslant \int_M K_g^-\, dA_g \geqslant -A(M, g),$$

where equality holds if and only if $K_g \equiv -1$.

The proof of Theorem B when $\chi(M) = 0$ is clear, since for any positive ε there exists a complete surface of revolution in E^3 around the x-axis such that the Gaussian curvature of it is 1 around the origin and is negative away from the origin such that the area of it is less than ε.

The following lemma is useful for the proof of the rest of Theorem B.

LEMMA. *For every $g \in \mathfrak{M}(R^2)$ with $c(R^2, g) = 2\pi$, there exists a point p_0 and an $R \geqslant \pi$ such that the metric R-ball $B_R(p_0)$ around p_0 has area greater than 4π.*

PROOF. Since $c(R^2, g) = 2\pi$, every Busemann function on it is exhaustion, and, in particular, takes minimum. Let p_0 be a point on the minimum set of a Busemann function. Then for every unit vector v at p_0 there exists a ray σ emanating from p_0 such that $\langle v, \dot{\sigma}(0)\rangle \geqslant 0$. The exhaustion property of every Busemann function was proved in [6]. The above fact implies that there are at least two distinct rays emanating from p_0. Consider the set V of all points on all rays emanating from p_0.

Set $\bigcup_{\lambda\in\lambda}U_\lambda = R^2 - V$, where $U_\lambda \cap U_\mu = \varnothing$ for $\lambda \neq \mu$. For each $\lambda \in \Lambda$, $\overline{U}_\lambda - U_\lambda$ consists of two distinct rays, and the angle at p_0 between the two rays measured on U_λ is not greater than π. Each U_λ contains no ray emanating from p_0 and contains a component of $C(p_0)$.

If the injectivity radius $i(p_0)$ of the exponential map at p_0 is not smaller than π, then the conclusion is direct from the Rauch comparison theorem.

If $i(p_0) < \pi$, then there is a point $p' \in C(p_0) \cap U_\lambda$ for some $\lambda \in \Lambda$ such that $d(p_0, p') < \pi$. Let $p_1 \in C(p_0) \cap U_\lambda$ be a point with the property that $d(p_0, p_1) = d(p_0, C(p_0) \cap U_\lambda) =: a_0$. Then there exists a geodesic loop γ_0 at p_0 of length $2a_0$ such that $\gamma_0(a_0) = p_1$ and $\gamma_0((0, 2a_0))$ is contained in U_λ. γ_0 bounds a 2-disk D_0 which is contained in U_λ, and the angle α_0 of γ_0 at p_0 measured on D_0 is less than π. Thus, D_0 is convex. It follows from $a_0 < \pi$ and $\alpha_0 < \pi$ that there exists a point $q \in C(p_1) \cap D_0$ with the property that $d(p_1, q) < d(p_1, p_0)$. Therefore, there is a point p_2 on $D_0 \cap C(p_1)$ such that $d(p_1, p_2) = d(p_1, C(p_1) \cap D_0) < d(p_0, p_1)$. Set $a_1 := d(p_1, p_2)$. There exists a geodesic loop γ_1 at p_1 of length $2a_1$ whose image lies in D_0, and the angle α_1 of γ_1 at p_1 measured on D_1 is less than π. By iterating this procedure, one finally gets a simply closed geodesic γ in D_0 whose length is lim $2a_j$ and γ bounds a 2-disk D contained entirely in D_0. The Gauss-Bonnet theorem implies that $c(D, g) = 2\pi$, and, in particular, $A(D, g) \geqslant 2\pi$ follows from the assumption $K_g \leqslant 1$.

The above argument shows that if there is a point q on $C(p_0) \cap U_\lambda$ such that $d(p_0, q) < \pi$, then there is an $R \geqslant \pi$ such that $A(B_R(p_0) \cap U_\lambda, g) > 2\pi$. Therefore, $A(B_R(p_0), g) > 4\pi$ holds for some $R \geqslant \pi$ if there are at least two points q and q' on $C(p_0)$ such that $q \in U_\lambda$, $q' \in U_\mu$ with $\lambda \neq \mu$, and such that $d(p_0, q)$ and $d(p_0, q')$ are less than π. If there is a unique $\lambda \in \Lambda$ such that every point $q \in C(p_0)$ with $d(p_0, q) < \pi$ lies in U_λ, then the Rauch theorem implies that $A(B_\pi(p_0) - U_\lambda, g) = 2(2\pi - \tau)$, where θ is the angle of U_λ at p_0. Thus,

$$A(B_R(p_0), g) > 2\pi + 2(2\pi - \tau) \geqslant 4\pi$$

holds for some $R \geqslant \pi$, and the proof is complete.

The rest of the proof of Theorem B is achieved by showing that for any positive ε there exists a $g_\varepsilon \in \mathfrak{M}(R^2)$ such that $A(R^2, g) < 4\pi + \varepsilon$. Let $y = f(x)$, $x \geqslant 0$, be the equation of a tractrix with $f(0) = 1$. For a given positive ε there is a small positive η such that the area of the surface of revolution in E^3 around the x-axis, whose profile curve is given by $y = \eta f(x)$, has are less than $\varepsilon/2$. Let S^2 be the unit sphere in E^3 around the origin and remove from S^2 a small ball around the point $(1, 0, 0)$. Then attach a portion of the surface of revolution to the hole such that the total area of the resulting C^0-surface is less than $4\pi + 2\varepsilon/3$. This surface is approximated by smooth surfaces whose induced metrics have Gaussian curvature not greater than 1 and whose area is less than $4\pi + \varepsilon$. This completes the proof of Theorem B.

Note that if $\{\varepsilon_j\}$ is a monotone decreasing sequence with $\lim_{j\to\infty}\varepsilon_j = 0$ and if $g_j \in \mathfrak{M}(R^2)$ is obtained in the above construction for ε_j, then $\lim_{j\to\infty} A(R^2, g_j) = 4\pi$, and $\inf_{j\to\infty} \inf_{R^2} K_{g_j} = -\infty$. Hence, $g_j \notin \mathfrak{M}_0(R^2)$ for all large j.

REFERENCES

1. S. Cohn-Vossen, *Kürzeste Wege and Totalkrümmung auf Flächen*, Compositio Math. **2** (1935) 69–133.
2. F. Fiala, *Le problème isopérimètres sur les surfaces ouvertes à courbure positive*, Comment. Math. Helv. **13** (1940), 293–346.
3. M. Gromov, *Volume and bounded cohomology*, Publ. Math. I.H.E.S. **56** (1982), 5–100.
4. P. Hartman, *Geodesic parallel coordinates in the large*, Amer. J. Math. **86** (1964), 705–727.
5. A. Huber, *On subharmonic functions and differential geometry, in the large*, Comment. Math. Helv. **32** (1957), 13–72.
6. K. Shiohama, *A role of total curvature on complete noncompact Riemannian 2-manifolds*, Illinois J. Math. **28** (1984), 597–620.
7. ______, *Cut locus and parallel circles of a closed curve on a Riemannian plane admitting total curvature*, Comment. Math. Helv. (to appear).

DEPARTMENT OF MATHEMATICS, FACULTY OF SCIENCE, KYUSHU UNIVERSITY, FUKUOKA 812, JAPAN