THE SHARP SOBOLEV INEQUALITY AND THE BANCHOFF-POHL INEQUALITY ON SURFACES

RALPH HOWARD
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF SOUTH CAROLINA
COLUMBIA, S.C. 29208
HOWARD@MATH.SC.EDU

Abstract. Let (M, g) be a complete two dimensional simply connected Riemannian manifold with Gaussian curvature $K \leq -1$. If f is a compactly supported function of bounded variation on M then f satisfies the Sobolev inequality

$$4\pi \int_M f^2\, dA + \left(\int_M |f|\, dA \right)^2 \leq \left(\int_M \|\nabla f\|\, dA \right)^2.$$

Conversely letting f be the characteristic function of a domain $D \subset M$ recovers the sharp form $4\pi A(D) + A(D)^2 \leq L(\partial D)^2$ of the isoperimetric inequality for simply connected surfaces with $K \leq -1$. Therefore this is the Sobolev inequality "equivalent" to the isoperimetric inequality for this class of surfaces. This is a special case of a result that gives the equivalence of more general isoperimetric inequalities and Sobolev inequalities on surfaces.

Under the same assumptions on (M, g) if $c\colon [a, b] \to M$ is a closed curve and $w_c(x)$ is the winding number of c about x then the Sobolev inequality implies

$$4\pi \int_M w_c^2\, dA + \left(\int_M |w_c|\, dA \right)^2 \leq L(c)^2$$

which is an extension of the Banchoff-Pohl inequality to simply connected surfaces with curvature ≤ -1. As a final application of the Sobolev inequality it is shown if $D \subset M$ has finite area then there is a refinement of the McKean inequality for the first eigenvalue of a negatively curved surface:

$$\lambda(D) \geq \frac{1}{4} + \frac{\pi}{A(D)}.$$

Both this form of the Banchoff-Pohl inequality and the eigenvalue estimate are special cases of slightly more general results.

Contents

1991 *Mathematics Subject Classification.* Primary 53C42 Secondary 53A04, 53C65.

Key words and phrases. Isoperimetric inequalities, Sobolev inequalities, Banchoff-Pohl inequality.

1. Introduction

Let (M, g) be a two dimensional Riemannian manifold and for any domain D with compact closure in M (write this as $D \Subset M$) let $A(D)$ be the area of D and $L(\partial D)$ be the length of the boundary ∂D of D. Then it is well known that the isoperimetric inequality

$$4\pi A(D) \le L(\partial D)^2 \quad \text{for all} \quad D \Subset M$$

holds if and only if the Sobolev inequality

$$4\pi \int_M f^2 \, dA \le \left(\int_M \|\nabla f\| \, dA \right)^2 \tag{1.1}$$

holds for all compactly supported real valued functions of bounded variation on M (see §2.1 below for a short discussion of functions of bounded variation). For (M, g) Euclidean space this is due to Federer and Fleming [5] and Yau [16] extended their proof to Riemannian manifolds.

Moreover in the case of $(M, g) = (\mathbf{R}^2, dx^2 + dy^2)$ the standard plane if c is a closed curve in $\mathbf{R}^2$, $w_c(x, y)$ is the winding number of c about the point (x, y), and $L(c)$ the length of c then Osserman [9, p. 1194] made the observation the Sobolev inequality (1.1) can be used to prove the wonderful inequality

$$4\pi \int_{\mathbf{R}^2} w_c^2 \, dA \le L(c)^2$$

of Banchoff and Pohl [1].

In the hyperbolic plane with constant Gaussian curvature -1 the sharp isoperimetric inequality is

$$4\pi A(D) + A(D)^2 \le L(\partial D)^2 \tag{1.2}$$

for all domains D with compact closure. In this note we find the Sobolev inequality equivalent to this isoperimetric inequality and use it to give the form of the Banchoff-Pohl inequality in the class of simply connected complete surfaces that have a negative upper bound on the curvature.

1.1. Theorem (Sharp Sobolev Inequality)**.** *Let (M, g) be a noncompact two dimensional Riemannian manifold (which need not be complete) and assume there are constants $a > 0$ and b so that for every domain $D \Subset M$ the isoperimetric inequality*

$$aA(D) + bA(D)^2 \le L(\partial D)^2 \tag{1.3}$$

holds. If $b < 0$ also assume

$$A(M) \le \frac{a}{2|b|}.$$

Then for every compactly supported f of bounded variation on M

$$a \int f^2\,dA + b\left(\int |f|\,dA\right)^2 \le \left(\int \|\nabla f\|\,dA\right)^2. \tag{1.4}$$

If equality holds then, up to a set of measure zero, f is a constant multiple of the characteristic function of a domain $D \Subset M$ and D makes equality hold in the isoperimetric inequality (1.3). Conversely if the inequality (1.4) holds for all compactly supported functions of bounded variation then the isoperimetric inequality (1.3) holds for all D with compact closure in M.

Elementary consequences of this inequality is a new lower bound for the first eigenvalue of a domain is a simply connected surface and a generalization of the Banchoff-Pohl inequality.

1.2. Theorem. *Let (M, g) be a connected domain in a complete simply connected surface with curvature ≤ -1. Then the first Dirichlet eigenvalue satisfies*

$$\lambda_1(M) \ge \frac{1}{4} + \frac{\pi}{A(M)}.$$

1.3. Theorem (Generalized Banchoff-Pohl Inequality)**.** *Let (M, g) be a noncompact two dimensional simply connected Riemannian (which is not assumed to be complete) and K_0 a constant. Assume the Gaussian curvature of (M, g) satisfies*

$$K \le K_0, \qquad \text{and if } K_0 > 0 \text{ then } A(M) \le \frac{2\pi}{K_0}.$$

If $c\colon [a, b] \to M$ is a closed curve and $w_c(P)$ the winding number of c about $P \in M$ then

$$4\pi \int_M w_c^2\,dA - K_0\left(\int_M |w_c|\,dA\right)^2 \le L(c)^2. \tag{1.5}$$

Equality holds if and only c is the boundary (possibly transversed more than once) of a domain in M isometric to a geodesic disk in the simply connected space of constant curvature K_0.

A simply connected noncompact surface is diffeomorphic to the plane $\mathbf{R}^2$ so that the winding number w_c can be defined in the usual manner. These results apply to simply connected domains in the sphere S^2 with area $\le 2\pi$. As these domains are not complete assuming completeness is not natural in Theorem 1.3.

Our reason for working with functions of bounded variation is that it simplifies the proofs of when equality holds in the inequalities. In many proofs that a Sobolev inequality like (1.4) implies an isoperimetric inequality like (1.3) it is usual to approximate a characteristic function χ_D of by smooth (or Lipschitz) functions f in (1.3) and then take limits (cf. [5, rmk 6.6 p.487], [16], [9, p. 1194], [3, p. 97], [17, p. 81]). As with most proofs of inequalities by approximation this makes understanding the case of equality difficult. The advantage of working with functions of bounded variation in this setting is that if D is a domain with compact closure in M with and so that the boundary ∂D has finite length then the characteristic function of D is of bounded variation and its total variation is given by $\int_M \|\chi_D\|\,dA = L(\partial D)$. Thus in the class of functions of bounded variation the isoperimetric inequality (1.3) can be proven by directly letting $f = \chi_D$ in the Sobolev inequality (1.4). This makes understanding the case of equality more or less straightforward. While using functions of bounded variation in problems of this type is certainly not a new idea, it deserves to be better known.

Under the assumptions $K_0 \le 0$ and (M, g) is simply connected and complete B. Süssmann has independently given a proof of the inequality (1.5). His proof uses the very ingenious idea of studying the effect of the flow of the curve shortening equation on the inequality. When (M, g) is the hyperbolic plane Teufel [11] has given anther generalization of the Banchoff-Pohl inequality: $4\pi \int_M w_c^2\,dA + \left(\int_M w_c\,dA\right)^2 \le L(c)^2$. While this inequality is sharp in that equality holds exactly when c is the boundary of a geodesic disk (possibly transversed more than once), if w_c changes sign on M, then the inequality (1.5) gives a better lower bound on $L(\partial D)^2$. For other extensions of the Banchoff-Pohl inequality to curved surfaces see [4, 7, 11, 12, 13, 14, 15].

Notation and Terminology. By smooth we mean of class C^∞. A domain in a manifold is an open set which we do not assume is connected. If D is a domain in M then $D \Subset M$ means that the closure of D in M is compact. By convention we assume simply connected domains are connected.

2. Proofs

2.1. **Functions of bounded variation and the coarea formula.** Let (M, g) be an oriented n dimensional Riemannian manifold and let dV the the volume form on M. Then for a smooth function $f\colon M \to \mathbf{R}$ let ∇f be the gradient of f, that is ∇f is the vector field so that for all tangent vectors V there holds $df(V) = \langle \nabla f, V\rangle$. Let $C_0^\infty(M, T(M))$ be the space of compactly supported smooth vector fields in on M with the usual inductive limit topology (that is a $\Phi_\ell \to \Phi$ iff there is a compact set that contains the supports of all the Φ_ℓ and moreover the sequence $\{\Phi_\ell\}_{\ell=1}^\infty$ and all its partial derivatives converge uniformly to the corresponding partial derivatives of Φ). If f is a locally integrable function then define a continuous linear functional on $C_0^\infty(M, T(M))$ by

$$\Lambda_f(\Phi) := -\int_M f \operatorname{div}(\Phi)\,dV.$$

If f is C^1 then by the divergence theorem

$$\Lambda_f(\Phi) = \int_M \langle \nabla f, \Phi \rangle \, dV$$

and so when f is sufficiently smooth the linear functional Λ_f is represented by integration against the classical gradient ∇f of f. In general Λ_f can be viewed as the distributional gradient of f. A function is of **bounded variation** iff the linear functional Λ_f is represented by measures of finite total variation. That is if and only if in a local coordinate system $(x^1, \dots, x^n)$ on M defined on an open set U of M there are Borel measures of finite total variation $\mu_1, \dots, \mu_n$ so that for any smooth vector field $\Phi = \sum \varphi^i \partial/\partial x^i$ supported in U

$$\Lambda_f(\Phi) = \int_U \sum_{i,j} g^{ij} \varphi_i \, d\mu_j.$$

In this case case μ_j is the distributional derivative $\partial f/\partial x^j$. A function of bounded variation need not be continuous. If $D \Subset M$ then with Lipschitz boundary then the characteristic function χ_D is of bounded variation (cf. [17, p. 229]). More generally a set $E \Subset M$ is of **finite parimeter** iff the the characteristic function χ_E is of bounded variation. For our purposes all that matters about sets of finite parimeter is that a set E of finite parimeter has a generalized boundary $\partial^* E$ cf. [17, p. 240] (which agrees with the usual topological boundary when E is a domain with C^1 boundary) and $\mathcal{H}^{n-1}(\partial^* E) < \infty$ where $\mathcal{H}^{n-1}$ is $n-1$ dimensional Hausdorff measure.

If Φ is a vector field on M let $|\Phi(x)| = \sqrt{\langle \Phi(x), \Phi(x) \rangle}$. The the **total variation measure** $\|\nabla f\| \, dV$ of a function of bounded variation is defined first on non-negative real valued continous functions u by

$$\int_M u \|\nabla f\| \, dV := \sup\{\Lambda_f(\Phi) : \Phi \in C_0^\infty(M, T(M)), |\Phi(x)| \le u(x)\}$$

and then extended to arbitrary continuous functions by linearity (cf. [17, p. 221]). There is anther characterization of the total variation measure of f by the version of the coarea formula due to Fleming and Rishel [6] (or cf. [17, thm 5.4.4 p. 231 and thm 5.8.1 p. 247]) which gives an integral formula for the total variation of f

$$\int_M \|\nabla f\| \, dV = \int_0^\infty \mathcal{H}^{n-1}(\partial^* \{x : |f(x)| \ge t\}) \, dt. \tag{2.1}$$

(One of the conclusions of [17, thm 5.4.4 p. 231] is that $\{x : |f(x)| \ge t\}$ is of finite parimeter for almost all $t \in \mathbf{R}$ so the integral on the right makes sense.)

In what follows we will only be interested in the two dimensional case. Then the volume measure dV will be replaced by the area measure dA and we will denote the one dimensional Hausdorff measure of E by $L(E)$ as in the case E is a curve $\mathcal{H}^1(E)$ is just the length of E. We also simplify the notation in (2.1) and use ∂ for ∂^*. This should not lead to any confusion.

With this notation the coarea formula for functions of bounded variation on a surface becomes

$$\int_M \|\nabla f\| \, dA = \int_0^\infty L(\partial\{f(x) : |f(x)| \geq t\}) \, dt.$$

Note this form of the coarea formula makes it clear that if $D \Subset M$ has a rectifiable boundary then

$$\int_M \|\nabla \chi_D\| \, dA = L(\partial D).$$

Finally if f is in $W^{1,1}(M)$ (that is the distributional first derivatives of f exist and are Lebesgue integrable) then f is of bounded variation, the total variation measure is absolutely continuous with respect to the area measure on M and is given by

$$\|\nabla f\| \, dA = |\nabla f| \, dA$$

where $|\nabla f(x)| = \sqrt{\langle \nabla f(x), \nabla f(x)\rangle}$.

2.2. **Proof of Theorem 1.1.** Let (M, g) be a noncompact two dimensional Riemannian manifold as in the statement of Theorem 1.1 and so that the isoperimetric inequality (1.3) holds. We use the notation

$$A(t) := A\{x \in M : |f(x)| \geq t\}, \quad L(t) := L(\partial\{x \in M : |f(x)| \geq t\}).$$

(That is $A(t)$ is the Lebesgue measure of the set $\{x \in M : |f(x)| \geq t\}$ and $L(t)$ is the one dimensional Hausdorff measure of $\partial\{x \in M : |f(x)| \geq t\}$.) By a standard result from real analysis for any measurable function u on M

$$\int_M |u| \, dA = \int_0^\infty A\{x : |u(x)| \geq t\} \, dt.$$

Applying this to f and f^2 gives $\displaystyle\int_M |f| \, dA = \int_0^\infty A(t) \, dt$ and

$$\begin{aligned}
\int_M |f|^2 \, dA &= \int_0^\infty A\{x : f(s)^2 \geq s\} \, ds \\
&= \int_0^\infty A\{x : f(x)^2 \geq t^2\} \, 2t \, dt = 2\int_0^\infty A(t) t \, dt.
\end{aligned}$$

By the coarea formula and the isometric inequality (1.3)

$$\int_M \|\nabla f\| \, dA = \int_0^\infty L(t) \, dt \geq \int_0^\infty \sqrt{aA(t) + bA(t)^2} \, dt.$$

So it is enough to prove

$$2a\int_0^\infty A(t) t \, dt + b\left(\int_0^\infty A(t) \, dt\right)^2 \leq \left(\int_0^\infty \sqrt{aA(t) + bA(t)^2} \, dt\right)^2. \tag{2.2}$$

The proof now splits into two cases.

Case 1: $b \le 0$. (This case follows closely the ideas in the papers of Federer and Fleming [5] and Yau [16].) Set

$$\begin{aligned} F(s) &:= 2a \int_0^s A(t)t\,dt + b\left(\int_0^s A(t)\,dA\right)^2, \\ G(s) &:= \left(\int_0^s \sqrt{aA(t)+bA(t)^2}\,dt\right)^2. \end{aligned}$$

So that

$$\begin{aligned} F'(s) &= 2aA(s)s + 2b\int_0^s A(t)\,dtA(s), \\ G'(s) &= 2\int_0^s \sqrt{aA(t)+bA(t)^2}\,dy\sqrt{aA(s)+bA(s)^2}. \end{aligned}$$

As $A(\cdot)$ is a decreasing function $\int_0^s A(t)\,dt \ge sA(s)$. Also $b \le 0$ so

$$F'(s) \le 2asA(s) + 2bsA(s)^2. \tag{2.3}$$

Again using that $A(\cdot)$ is decreasing and that the function $A \mapsto \sqrt{aA+bA^2}$ is increasing on $[0, a/(2|b|)]$ (and by one of our assumptions $A(t) \le a/(2|b|)$)

$$G'(s) \ge 2s(aA(s) + bA(s)^2).$$

Therefore $F'(s) \le G'(s)$ and $F(0) = G(0)$ so $F(s) \le G(s)$. Letting $s \to \infty$ completes the proof that the required inequality (2.2) holds and completes the proof that (1.3) implies (1.4) when $b \le 0$. If equality holds in (1.4) then equality must hold in (2.3) for almost all $s > 0$. If s_0 is a point where equality holds and $A(s_0) > 0$ then $A(s) = A(s_0)$ for all $s \in [0, s_0]$. If s_0 is a point where $A(s_0) = 0$ then A is non-negative and monotone decreasing so $A(s) = 0$ for $s > s_0$. Thus for some constants $c_1, c_2 > 0$ the function $A(\cdot)$ is given by $A(s) = c_1\chi_{[0,c_2]}(s)$. Then a farther chase through the definitions shows for some domain $D \Subset M$ that $f = \pm c_2\chi_D$ where D is a domain with $A(D) = c_1$. As equality holds in (1.4) it follows that $aA(D) + bA(D)^2 = L(\partial D)^2$.

Case 2: $b \ge 0$. Set

$$H(\lambda) := \left(\int_0^\infty \sqrt{aA(t)+\lambda A(t)^2}\,dt\right)^2.$$

Then using the Cauchy-Schwartz inequality we estimate the derivative of $H(\cdot)$ from below:

$$\begin{aligned} H'(\lambda) &= \left(\int_0^\infty \sqrt{aA(t)+\lambda A(t)^2}\,dt\right)\left(\int_0^\infty \frac{A(t)^2}{\sqrt{aA(t)+\lambda A(t)^2}}\,dt\right) \\ &\ge \left(\int_0^\infty A(t)\,dt\right)^2. \end{aligned}$$

Noting that the argument we used in proving (2.2) in the case $b \leq 0$ only used that $A(\cdot)$ was decreasing so we can let $b = 0$ in that inequality to get

$$H(0) = \left(\int_0^\infty \sqrt{aA(t)}\, dt \right)^2 \geq 2a \int_0^\infty A(t)t\, dt. \tag{2.4}$$

This implies for all $\lambda \geq 0$ that

$$H(\lambda) = \left(\int_0^\infty \sqrt{aA(t) + \lambda A(t)}\, dt \right)^2 \geq 2a \int_0^\infty A(t)t\, dt + \lambda \left(\int_0^\infty A(t)\, dt \right)^2.$$

Letting $\lambda = b$ in this inequality gives that (2.2) holds and completes the proof of the inequality in the case $b \geq 0$. If equality holds then equality must hold in (2.4). But as this was proven by the same method that was used in the case $b \leq 0$ the same analysis shows that equality in (1.4) implies $f = c\chi_D$ where $D \Subset M$ makes equality hold in (1.3).

Conversely if (M, g) is so that the Sobolev inequality (1.4) holds for all compactly supported f of bounded variation then for a $D \Subset M$ with ∂D rectifiable the characteristic function χ_D will have bounded variation and so letting $f = \chi_D$ in (1.4) gives the isoperimetric inequality (1.3) and completes the proof. □

2.3. **Proof of Theorem 1.3.** Recall that by our convention a simply connected domain is also connected. The full force of the following lemma is not needed in the proof of Theorem 1.3, but it is of interest for its own sake. It is not hard to give examples of complete simply connected surfaces where the domain of least parimeter for a given area is either disconnected or connected but not simply connected. Thus the conclusion of the lemma that in some cases the "isoperimetric" domains must be simply connected in not vacuous.

2.1. Lemma. *Let (M, g) be a compact simply connected two dimensional Riemannian so that every simply connected domain $D \Subset M$ satisfies the isoperimetric inequality (1.3) and if $b < 0$ also assume $A(M) \leq a/2|b|$. Then every $D \Subset M$ satisfies this inequality. If D is a domain so that equality holds in the inequality, then D is simply connected.*

Proof. By the classification of surfaces M is diffeomorphic to the plane $\mathbf{R}^2$. Let $D_1 \Subset M$ be a connected domain in M. Let D be the domain obtained from D_1 by filling in the holes of D_1. To be precise a point x of M is in D if and only if there is a closed curve c in D_1 so that the winding number of c about x is non-zero. (As M is diffeomorphic to $\mathbf{R}^2$ the winding number can be defined in the usual manner.) Then D is also a bounded domain in M and it is simply connected. This D satisfies the given isoperimetric inequality. But $D_1 \subseteq D$ and $\partial D \subseteq \partial D_1$ so $A(D_1) \leq A(D)$ and $L(\partial D) \leq L(\partial D_1)$. Therefore

$$aA(D_1) + bA(D_1)^2 \leq aA(D) + bA(D)^2 \leq L(D)^2 \leq L(D_1)^2$$

as the function $A \mapsto aA + bA^2$ is increasing on the interval $[0, A(M)]$ (this is where the assumption $A(M) \le a/2|b|$ for $b < 0$ is used). This shows that any connected domain $D_1 \Subset M$ satisfies the required inequality. Moreover as $A(D_1) = A(D)$ if and only if $D_1 = D$ we see that equality holds for a connected domain D_1 if and only if $D_1 = D$, that is if and only if D_1 is simply connected.

It is an elementary exercise to show that for positive real numbers A_1, A_2, L_1, L_2 with both $aA_1 + bA_1^2$ and $aA_2 + bA_2^2$ nonnegative that the implication

$$\begin{aligned} aA_1 + bA_1^2 \le L_1^2 \quad &\text{and} \quad aA_2 + bA_2^2 \le L_2^2 \\ \text{implies} \quad a(A_1 + A_2) &+ b(A_1 + A_2)^2 < (L_1 + L_2)^2 \end{aligned} \tag{2.5}$$

holds. This and induction shows that the required inequality holds for all domains $D_2 \Subset M$ that are finite unions of connected domains. As any domain $D_3 \Subset M$ is a countable union of connected domains the general case follows by an easy limit argument.

If D is so that the equality $aA(D) + bA(D)^2 = L(D)^2$ holds then D must be connected as otherwise D would be the disjoint union of two subdomains D' and D'' each of which satisfies the inequality (1.3). But then the implication (2.5) would imply $aA(D) + bA(D)^2 < L(D)^2$ contrary to the assumption that equality holds. But if D is connected then as remarked above in the "filling in the holes" argument equality in the isoperimetric inequality implies D is simply connected. This completes the proof. □

To prove Theorem 1.3 we first note if (M, g) satisfies the hypothesis of the theorem and $D \Subset M$ is simply connected, then the Euler characteristic of D is $\chi(D) = 1$. By the form of the isoperimetric inequality in the book of Burago and Zalgaller [2, thm 2.2.1 p. 11] the domain D satisfies

$$4\pi\chi(D)A(D) - K_0 A(D)^2 = 4\pi A(D) - K_0 A(D)^2 \le L(\partial D)^2.$$

Therefore by the lemma this inequality holds for all $D \Subset M$. Now let $c\colon [a, b] \to M$ be a rectifiable curve. Then the function $x \mapsto w_c(x)$ is of bounded variation on M and, as in [9, pp. 1194–1195],

$$\int_M \|\nabla w_c\| \, dA = L(c).$$

The inequality (1.5) of Theorem 2 now follows by letting $f = w_c$ and using Theorem 1.1.

If equality holds in (1.5) then by using when equality holds in Theorem 1.1 there is a constant C and a domain $D \Subset M$ so that $w_c = C\chi_D$ and D makes equality hold in the isoperimetric inequality (1.2). By the lemma this implies D is simply connected and therefore ∂D is connected. But then c must be ∂D transversed one or more times in the same direction. But equality holds in the isoperimetric inequality for a simply connected domain $D \Subset M$ if and only if D is isometric to a disk in the simply connected complete surface of constant curvature K_0 (cf. [2, thm 4.3.1 p. 33]). This completes the proof. □

3. A Lower Bound for the First Eigenvalue on a Surface

Let (M, g) be a two dimensional noncompact Riemannian manifold which is not assume to be complete. Let $C_0^\infty(M)$ be the space of smooth compactly supported real valued functions on M. Then the first eigenvalue $\lambda_1(M)$ of (M, g) is defined by

$$\lambda_1(M) := \inf_{0 \neq f \in C_0^\infty(M)} \frac{\int_M |\nabla f|^2 \, dA}{\int_M f^2 \, dA}$$

If M is a domain with compact closure and smooth boundary in a complete surface then $\lambda_1(M)$ is the first eigenvalue of the Laplace-Beltrami operator with Dirichlet boundary values. Now assume that for every $D \Subset M$ the isoperimetric inequality (1.3) holds and if $b < 0$ assume $A(M) \leq a/2|b|$. Then by Theorem 1.1 for any compactly supported Lipschitz function u

$$a \int_M u^2 \, dA + b \left(\int_M |u| \, dA \right)^2 \leq \left(\int_M |\nabla u| \, dA \right)^2 . \tag{3.1}$$

Let f be a smooth function supported in M and let $1 \leq p < \infty$. Let $q := p/(p-1)$ be the conjugate exponent (i.e. $1/p + 1/q = 1$). Then for any smooth compactly supported function f on M by Hölder's inequality

$$\begin{aligned}
\left(\int_M |\nabla f^p| \, dA \right)^2 &= p^2 \left(\int_M |f|^{p-1} |\nabla f| \, dA \right)^2 \\
&\leq p^2 \left(\int_M |f|^p \, dA \right)^{\frac{2(p-1)}{p}} \left(\int_M |\nabla f|^p \right)^{\frac{2}{p}}
\end{aligned}$$

and by the Cauchy-Schwartz inequality

$$\frac{1}{A(M)} \left(\int_M |f|^p \, dA \right)^2 \leq \int_M |f|^{2p} \, dA.$$

If we make the substitution $u = f^p$ in (3.1) and use the last two inequalities

$$\left(\frac{a}{A(M)} + b \right) \left(\int_M |f|^p \, dA \right)^{\frac{2}{p}} \leq p^2 \left(\int_M |\nabla f|^p \, dA \right)^{\frac{2}{p}} .$$

Letting $p = 2$ leads to:

3.1. Theorem. *Let (M, g) be a two dimensional noncompact Riemannian manifold so that the isoperimetric inequality (1.3) holds for all $D \Subset M$ and if $b < 0$ assume $A(M) \leq a/2|b|$. Then the first eigenvalue satisfies*

$$\lambda_1(M) \geq \frac{1}{4} \left(\frac{a}{A(M)} + b \right) . \quad \square$$

3.2. Corollary. *Let (M, g) be a connected domain in a complete simply connected surface with curvature ≤ -1. Then*

$$\lambda_1(M) \geq \frac{1}{4} + \frac{\pi}{A(M)}.$$

Proof. This follows from the last theorem and that in a simply connected complete surface with curvature ≤ -1 the isoperimetric inequality $4\pi A(D)+A(D)^2 \leq L(\partial D)^2$ holds for all domains with compact closure. $\square$

By a well known theorem of McKean [8] (or cf. [3, p. 70]) if (M,g) is a domain in a surface with curvature bounded above by -1 then $\lambda_1(M) \geq 1/4$. Thus the last corollary shows that if the area of M is finite then it is possible to give a better lower bound. To get a feel for how sharp this is for $A > 0$ let $\mathcal{D}(A)$ be the geodesic disk of area A in the complete simply connected surface with $K \equiv -1$. Then by the form of the Faber-Krahn inequality given in the book [3, p. 87] of Chavel if M is a domain in a simply connected surface with $K \leq -1$ then $\lambda_1(M) \geq \lambda_1(\mathcal{D}(A(M)))$ with equality iff M is isometric to $\mathcal{D}(A(M))$. Thus $\lambda_1(M) \geq \lambda_1(\mathcal{D}(A(M)))$ is the sharp lower bound for the first eigenvalue of domains in M. However the function $A \mapsto \lambda_1(\mathcal{D}(A))$ is not easily computable so the estimate of the last theorem is still interesting. As $\lim_{A\to\infty} \lambda_1(\mathcal{D}(A)) = 1/4$ this estimate is good for domains that contain a large geodesic disk. At the other extreme using that we can compute the eigenvalues of a Euclidean disk in terms of zeros of Bessel functions and that a small geodesic disk in surface has its eigenvalues asymptotic to those of a of a disk of the same area in the Euclidean plane, we find for $A(M) \searrow 0$

$$\lambda_1(M) \geq \lambda_1(\mathcal{D}(A(M))) \sim \frac{(5.7831859\ldots)\pi}{A(M)}$$

so for domains of very small area the lower bound given by the corollary is off by a factor of close to six, but still of the right order of growth.

4. Remarks and an Open Problem

Let H^n be the n dimensional hyperbolic space. Then is would be interesting to find an analytic inequality "equivalent" to the isoperimetric inequality in H^n. Let ω_n be the surface area of the unit sphere S^{n-1} in $\mathbf{R}^n$. Let $V(r)$ be the volume of a geodesic ball of radius r in H^n and let $A(r)$ be the surface area measure of a geodesic sphere of radius r. As the geodesic balls in H^n solve the isoperimetric problem for H^n, the isoperimetric inequality in H^n is given by the relationship between $A(r)$ and $V(r)$. They are given by

$$A(r) = \omega_n \sinh^{n-1}(r), \qquad V(r) = \omega_n \int_0^r \sinh^{n-1}(t)\, dt.$$

When $n = 3$, $A(r) = 4\pi \sinh^2(r)$ and $V(r) = 2\pi(\cosh(r)\sinh(r) - r)$. But $\cosh(r)$ and $\sinh(r)$ are rational functions in e^r and e^r is transcendental over the field of rational functions in r. Thus in this case there is no algebraic relationship between $A(r)$ and $V(r)$. A similar argument shows there is no algebraic relationship between $A(r)$ and $V(r)$ whenever n is odd. If n is even then both $V(r)$ and $A(r)$ are rational functions in e^r and thus there is a polynomial relation between $V(r)$ and $A(r)$, but for $n \geq 4$ this polynomial is rather complicated as can be seen by computing it for $n = 4$. Thus it seems that the results here do not have a straightforward generalization to higher dimensions.

Problem. *Find a Sobolev type inequality for functions of bounded variation on the n dimensional hyperbolic H^n space that is equivalent to the sharp isoperimetric inequality in H^n.*

Acknowledgments: The idea of introducing the parameter λ and taking the derivative used in Case 2 of the proof of Theorem 1.1 is the result of several very enjoyable conversations with Harold Shaperio. Eberhard Teufel read a version of the manuscript and made several useful comments. I also benefited from conversations and/or correspondence with Lars Andersson, Mike Gage, and Bernd Süssmann.

References

[1] T. F. Banchoff and W. F. Pohl, *A generalization of the isoperimetric inequality*, Jour. Diff. Geo. **6** (1971), 175–213.

[2] J. D. Burago and V. A. Zalgaller, *Geometric inequalities*, Grundlehren, vol. 285, Springer, Berlin, 1980.

[3] I. Chavel, *Eigenvalues in Riemannian geometry*, Pure and Applied Mathematics, vol. 115, Academic Press, New York, 1984.

[4] K. Enomoto, *A generalization of the isoperimetric inequality on S^2 and flat tori in S^3*, Proc. Amer. Math. Soc. **120** (1994), no. 2,, 553–558.

[5] H. Federer and W. H. Fleming, *Noraml and integral currents*, Ann. Math. **72** (1960), 458–520.

[6] W. Fleming and R. Rishel, *An integral formula for the total gradient variation*, Arch. Math. **11** (1960), 218–222.

[7] L. Gysin, *The isoperimetric inequality for nonsimple closed curves*, Proc. Amer. Math. Soc **118** (1993), no. 1, 197–203.

[8] H. P. McKean, *An upper bound for the spectrum of Δ on a manifold of negative curvature*, Jour. Diff. Geo. **4** (1970), 359–366.

[9] R. Osserman, *The isoperimetric inequality*, Bull. Amer. Math. Soc. **84** (1978), no. 6, 1182–1238.

[10] B. Süssmann, *Curve shorting and the Banchoff-Pohl inequality on surfaces of nonpositive curvature*, Preprint (1996).

[11] E. Teufel, *A generalization of the isoperimetric inequality in the hyperbolic plane*, Arch. Math. **57** (1991), no. 5, 508–513.

[12] ———, *Isoperimetric inequalities for closed curves in spaces of constant curvature*, Results Math. **22** (1992), 622–630.

[13] ———, *On integral geometry in Riemannian spaces*, Abh. Math. Sem. Univ. Hamburg **63** (1993), no. 2, 17–27.

[14] J. L. Weiner, *A generalization of the isoperimetric inequality on the* 2*-sphere*, Indiana Univ. Math. Jour. **24** (1974), 243–248.

[15] ———, *Isoperimetric inequalities for immersed closed spherical curves*, Proc. Amer. Math. Soc. **120** (1994), no. 2, 501–506.

[16] S. T. Yau, *Isoperimetric constants and the first eigenvalue of a compact manifold*, Ann. Sci. Ec. Norm. Super., Paris **8** (1975), 487–507.

[17] W. P. Ziemer, *Weakly differentiable functions*, Graduate Texts in Mathematics, vol. 120, Springer-Verlag, New York, 1989.