Isoperimetric inequalities and Dirichlet functions of Riemann surfaces

José Manuel Rodríguez García

Universidad Carlos III de Madrid
Escuela Politécnica Superior
C/ Butarque, 15
28.911 Leganés (Madrid)
1. Introduction.

In this paper we study the relationship between linear isoperimetric inequalities and the existence of harmonic functions with finite Dirichlet integral on Riemann surfaces.

By $S$ we denote a Riemann surface (whose universal covering space is the unit disk $\Delta$) endowed with its Poincaré metric, i.e. the metric obtained by projecting the Poincaré metric of the unit disk: $ds = 2(1 - |z|^2)^{-1} |dz|$. With this metric, $S$ is a complete Riemannian manifold with constant curvature $-1$. The only Riemann surfaces which are left out are the sphere, the plane, the punctured plane and the tori.

We shall say that a Riemann surface $S$ satisfies a “linear isoperimetric inequality” (LII) if there exists a finite constant $h(S)$ so that for every relatively compact open set $G$ with smooth boundary we have

$$A(G) \leq h(S) L(\partial G).$$

Here and from now on, $A$, $L$, $\partial$ and $B$ refer to Poincaré area, length, distance and open ball of $S$.

There are connections between (LII) and some conformal invariants on Riemann surfaces: the bottom of the spectrum of the Laplace-Beltrami operator, $b(S)$, and the exponent of convergence $\delta(S)$:

**Theorem A** ([Ch], [B, p.228], [FR]). A Riemann surface $S$ satisfies a linear isoperimetric inequality if and only if $b(S) > 0$. In fact,

$$\frac{1}{4} \leq b(S) h(S)^2 \quad \text{and} \quad b(S) h(S) < \frac{3}{2}.$$  

The next result is a well known theorem of Elstrodt-Patterson-Sullivan:

**Theorem B** [S, p.333]. A Riemann surface $S$ satisfies a linear isoperimetric inequality if and only if $\delta(S) < 1$. In fact,

$$b(S) = \begin{cases} 
\frac{1}{4}, & \text{if } 0 \leq \delta(S) \leq \frac{1}{2}, \\
\delta(S) (1 - \delta(S)), & \text{if } \frac{1}{2} \leq \delta(S) \leq 1.
\end{cases}$$

A theorem of Myrberg [T, p.522] states that if $\delta(S) < 1$ (if $S$ satisfies a LII) then $S$ has a Green’s function ($S \notin O_G$ in the language of classification theory). If $S$ is a plane domain (in fact, if $S$ is a surface of almost finite genus [SN, p.193]), $S$ has Green’s function if and only if $S$ has a non-constant harmonic function with finite Dirichlet integral [SN, p.194] ($S \notin O_{HD}$ in the language of classification theory).
One would like to understand the relationship between the classes $O_{HD}$ and $\mathcal{B}$ (the Riemann surfaces which do not satisfy a LII). As we have said above, in the case of surfaces of almost finite genus, $O_G = O_{HD} \subset \mathcal{B}$. The inclusion is strict, as it is shown by the example $S_0 = \Delta \setminus \bigcup_{k=1}^{\infty} \{2^{-k}\} \cup \{0\}$: $S_0 \notin O_G$ because it is a plane domain whose boundary has positive logarithmic capacity [T, p.81]; $S_0 \in \mathcal{B}$ because $\bigcup_{k=1}^{\infty} \{2^{-k}\} \cup \{0\}$ is a discrete set with an accumulation point in $\Delta$ [FR, Theorem 4].

The inclusion $O_{HD} \subset \mathcal{B}$ is true, in general, with an extra hypothesis:

**Theorem 1.** Let $S$ be a Riemann surface which satisfies a linear isoperimetric inequality. If there exists in $S$ a set of disjoint simple closed curves $\{\gamma_j\}_{j=1}^{m}$, such that $S \setminus \bigcup_j \gamma_j$ contains $n$ connected components of infinite area $S_1, \ldots, S_n$, then

$$\dim HD(S) \geq n.$$  

This inequality is the best possible.

Here $HD(S)$ denotes the (real) linear space of harmonic functions in $S$ with finite Dirichlet integral.

The inclusion $O_{HD} \subset \mathcal{B}$ is not true in the general case, even with the extra customary hypothesis of bounded geometry [K], which in our context means that the injectivity radius $\iota(S)$ is positive. $\iota(S)$ is defined as

$$\iota(S) = \inf \{\iota(p) : p \in S\},$$

where $\iota(p)$ is the injectivity radius of the geodesic exponential map centered at $p$.

**Theorem 2.** There exists a Riemann surface $R \in O_{HD}$, with $\iota(R) > 0$, which satisfies a linear isoperimetric inequality.

ACKNOWLEDGEMENTS. I would like to thank J. L. Fernández for many useful conversations about these results, and J. Llorente for his careful reading of the manuscript and for some helpful suggestions.

2. Proof of Theorem 1.

Without loss of generality, we can assume that $\{\gamma_j\}_{j=1}^{m}$ are simple closed geodesics. If this is not the case, we can substitute each curve by the geodesic in its same free homotopy class.

Let $S_k$ be a component of infinite area of $S \setminus \bigcup_j \gamma_j$, and let $S_k^*$ be the Schottky double of $S_k$ (see [AS, p.26] for the definition).

CLAIM. $S_k^*$ satisfies a linear isoperimetric inequality.
If the claim is true, the theorem of Myrberg [T, p.522] states that \( S_k^* \) has Green’s function. This implies that the Royden’s harmonic boundary of \( S_k^* \) is not empty [SN, p.166].

\( S_k^* \) is symmetrical with respect to \( \partial S_k \), a compact set which separates \( S_k^* \) in two connected components. Then the Royden’s harmonic boundary of \( S_k^* \) is also symmetrical with respect to \( \partial S_k \), and contains at least two points, one of them corresponding to \( S_k \) (one of them is in the closure of \( S_k \) in the Royden’s compactification of \( S_k^* \)).

This is true for \( k = 1, \ldots, n \). Therefore, the Royden’s harmonic boundary of \( S \) contains at least \( n \) points [SN, p.191]. This is equivalent [SN, p.166] to

\[
\dim HD(S) \geq n.
\]

This inequality is the best possible:

Let \( R \) be the Riemann surface given by Theorem 2 (\( R \) will be constructed without any mention to Theorem 1) and consider \( R_n \), a \( n \)-covering of \( R \) based in a closed simple geodesic \( \gamma \subset R \). \( R_n \) satisfies the hypothesis of this theorem and also \( \dim HD(R_n) = n \) (the Royden’s harmonic boundary of \( R_n \) consists of \( n \) points, because the Royden’s harmonic boundary of \( R \) consists of one point (\( R \in O_{HD} \setminus O_G \)) [SN, p.166]).

To finish the proof of Theorem 1, we only need to prove the Claim:

By a geodesic domain in a Riemann surface we mean a connected domain \( G \) with finite area, such that \( \partial G \) consists of finitely many closed simple geodesics. \( G \) does not have to be relatively compact since it may “surround” finitely many punctures.

The following lemma will be very useful:

**Lemma** [FR, p.168]. A Riemann surface satisfies a LII if and only if it satisfies LII for geodesic domains. Moreover, if \( h \) and \( h_g \) are, respectively, the usual and geodesic isoperimetric constants, then

\[
h_g \leq h \leq 2 + h_g.
\]

Therefore, we must verify LII only for geodesic domains of \( S_k^* \). By the symmetry of \( S_k^* \) and the LII of \( S \), we just need to check this for geodesic domains which are symmetrical with respect to \( \partial S_k \). Then, we must verify

\[
A(G) \leq c L(\partial_0 G)
\]

for geodesic domains \( G \) of \( S_k \), such that \( \partial G \cap \partial S_k \neq \emptyset \), where \( \partial_0 G \) means

\[
\partial_0 G \equiv \partial G \setminus \partial S_k.
\]
Consider the open sets $C_t = \{ p \in S_k : d(p, \partial S_k) < t \}$ for positive $t$. Let $G_t$ be the geodesic domain "corresponding" to $C_t$ (each puncture or boundary curve of $G_t$ is freely homotopic to a boundary curve of $C_t$). If $G_t$ is empty for all positive $t$, then $S_k$ is a doubly connected domain (a funnel), $S_k^*$ is an annulus, and the claim is true with constant $1$. Then, we can assume that $G_t$ is connected and not empty for $t \geq t_0$. $G_t$ is non decreasing in $t$, and if $t_1 < t_2$ are such that $A(G_{t_1}) < A(G_{t_2})$, the constant curvature $-1$ and the Gauss-Bonnet theorem give $A(G_{t_1}) + 2\pi \leq A(G_{t_2})$.

This implies that there exists a positive number $T$ such that $G_t = G_T$ for all $t \geq T$, or $A(G_t) \to \infty$ as $t \to \infty$.

The first possibility is easy: there are only a finite number of geodesic domains. Without loss of generality, we can assume that $A(G_t) \to \infty$ as $t \to \infty$.

**Case 1.** $A(G) \geq 2h(S)\ell$, with $\ell = \sum_{j=1}^m L(\gamma_j)$. In this case,

$$2h(S)\ell \leq A(G) \leq h(S)L(\partial G) \leq h(S)(L(\partial_0 G) + \ell)$$

and

$$\ell \leq L(\partial_0 G).$$

Therefore,

$$A(G) \leq 2h(S)L(\partial_0 G).$$

**Case 2.** $A(G) < 2h(S)\ell$.

Let $\Omega$ be a geodesic domain in $S_k$ such that

$$\partial S_k \subset \partial \Omega \quad \text{and} \quad A(\Omega) \geq 2h(S)\ell.$$

We can choose $\Omega$, for example, as the first geodesic domain $G_t$ satisfying $A(G_t) \geq 2h(S)\ell$.

We define

$$a = \min \{ L(\gamma) : \gamma \text{ closed simple geodesic, } \gamma \subset \overline{\Omega} \},$$

$$b = \max \{ L(\gamma) : \gamma \text{ closed simple geodesic, } \gamma \subset \partial_0 \Omega \}.$$

Since $A(\Omega) > A(G)$ and $\Omega \cap G \neq \emptyset$, one of the two next possibilities holds:

**Case 2.1.** There exists a closed simple geodesic $\gamma \subset \overline{\Omega} \cap \partial_0 G$. Then

$$L(\partial_0 G) \geq L(\gamma) \geq a.$$

**Case 2.2.** There exists a closed simple geodesic $\eta$ in $\partial_0 G$, which meets some $\gamma \subset \partial_0 \Omega$. 

Then, the Collar Lemma [R] says that \( L(\eta) \geq 4d_0 \), where \( d_0 \) (the width of the greater collar of \( \gamma \)) satisfies

\[
\cosh d_0 \geq \coth \frac{L(\gamma)}{2} \geq \coth \frac{b}{2},
\]

and

\[
d_0 \geq D \equiv \text{arc \cosh} \left( \coth \frac{b}{2} \right).
\]

Randol [R] states the Collar Lemma if the surface is compact, but the same proof, without any change, works for a general Riemann surface.

Therefore,

\[
L(\partial_0 G) \geq L(\eta) \geq 4D.
\]

In both cases (2.1 and 2.2) \( L(\partial_0 G) \geq \min \{a, 4D\} \equiv c_0 \). Then

\[
A(G) \leq h(S)(L(\partial_0 G) + \ell) \leq h(S) \left( L(\partial_0 G) + \ell \frac{L(\partial_0 G)}{c_0} \right)
\]

and

\[
A(G) \leq h(S) \left( 1 + \frac{\ell}{c_0} \right) L(\partial_0 G).
\]

Obviously, \( \ell \geq a \geq c_0 \) and \( 1 + \ell/c_0 \geq 2 \). Therefore, in any case,

\[
A(G) \leq h(S) \left( 1 + \frac{\ell}{c_0} \right) L(\partial_0 G).
\]

Consequently,

\[
h(S_k) \leq 2 + h_g(S_k) \leq 2 + h(S) \left( 1 + \frac{\ell}{c_0} \right),
\]

and the proof of Theorem 1 is now complete.

3. Proof of Theorem 2.

The desired Riemann surface \( \mathcal{R} \) will be obtained with the help of a graph \( G \). We will construct this graph in three steps.

In the set of vertices of any connected graph we can define a natural distance:

\[
d(p,q) = \inf \{ \text{length of the paths from } p \text{ to } q \}.
\]

This will be “the distance” in all graphs of this section.

First, let \( T \) be the infinite complete binary tree with root \( r_0 \). Secondly, let \( V_n \) be the subset of \( 2^n \) vertices of \( T \) at distance \( n \) of \( r_0 \). We can construct new graphs \( G_n \ (n \geq 1) \) with vertices \( V_n \). In \( G_1 \) there is one edge between the two vertices of \( V_1 \). The edges of \( G_n \ (n \geq 2) \) are chosen as follows: \( 2^n - 1 \) vertices of \( V_n \) are connected
by a complete binary tree with \(2^{n-1}\) leaves and with root \(r_n\) (in any way); we add another edge between \(r_n\) and the last vertex \(v_n\) of \(V_n\). In this way, the degree of the vertices of \(G_n\) is one (if the vertex is a leave) or three (if the vertex is not a leave). The leaves are at distance \(n-1\) of \(r_n\), except for \(v_n\) which is at distance 1. Hence, the diameter of \(G_n\) is \(2n-2\), if \(n \geq 2\).

Finally, we are ready to construct the graph \(G\). The vertices of \(G\) are the vertices of \(T\). The edges of \(G\) are the union of the edges of \(T\) and the edges of \(G_n\), for all \(n \geq 1\). The root \(r_0\) of \(G\) has degree two. The other vertices of \(G\) have degree four or six.

To build up our Riemann surface \(\mathcal{R}\), modelled upon the graph \(G\), we will need the so called L{"o}bell Y-pieces, which are a standard tool for constructing Riemann surfaces. A clear description of these Y-pieces and their use is given in [C, Chapter X.3].

A L{"o}bell Y-piece is a three-holed sphere, endowed with a metric of constant negative curvature \(-1\), so that the boundary curves are geodesics. We also require that the lengths of the boundary curves are the same, say \(2\alpha\), and the distance between any two of these boundary curves is \(\beta\), say. Then \(\alpha\) and \(\beta\) are related by

\[
\sinh \left( \frac{\alpha}{2} \right) \sinh \left( \frac{\beta}{2} \right) = \frac{1}{2}.
\]

This is the unique restriction on \(\alpha\) and \(\beta\). See [C, p.248] for details.

Fix \(\alpha\) and \(\beta\) satisfying the above relation.

A X-piece (\(\ast\)-piece) is a four-holed (six-holed) sphere, endowed with a metric of curvature \(-1\), so that the boundary curves are geodesics of the same length \(2\alpha\). We can construct these pieces, for example, joining two (four) Y-pieces, by identifying corresponding boundary curves.

If we now put together these pieces following the combinatorial design of \(G\), with the X-pieces (\(\ast\)-pieces) in the place of the vertices of degree four (six), we obtain a complete surface of constant negative curvature \(-1\).

The only non-standard vertex is \(r_0\), which has degree two. There is not problem if we forget \(r_0\) and consider that the two vertices of \(V_1\), are connected by a double edge. Since we have used only two distinct pieces to build up \(\mathcal{R}\), it is trivial to see that \(\iota(\mathcal{R}) > 0\).

First of all, we will prove that \(\mathcal{R} \in O_{HD}\). Let \(u\) be a harmonic function in \(\mathcal{R}\) with finite Dirichlet integral. Without loss of generality we can assume that \(u\) is a bounded function [AS, p.203] [SN, p.178]. We want to verify that \(u\) is constant.

If \(u\) has limit at infinity, there is a point \(p\) in \(\mathcal{R}\) such that \(u(p)\) is the maximum or the minimum of \(u\) in \(\mathcal{R}\). The maximum principle implies that \(u\) is constant.
If $u$ is non-constant and has not limit at infinity, we can assume that $u$ is positive and
\[
\limsup_{z \to \infty} u(z) > 4 \quad \text{and} \quad \liminf_{z \to \infty} u(z) < 1.
\]
The maximum (minimum) principle implies that each connected component of the set \( \{ u > 4 \} (\{ u < 1 \} ) \) is not a relatively compact set of \( \mathcal{R} \).

This implies that, for each \( n \geq n_0 \), there exist points \( p_n, q_n \) in the pieces of \( \mathcal{R} \) corresponding to \( G_n \), such that
\[
u(p_n) > 4 \quad \text{and} \quad u(q_n) < 1.
\]

Since \( u \) is a positive harmonic function, the Harnack’s Theorem says that there exists a positive number \( \varepsilon < \iota(\mathcal{R}) \), independent of \( n \), such that,
\[
u(z) \geq 3, \quad \text{for all } z \in B(p_n, \varepsilon),
\]
\[
u(z) \leq 2, \quad \text{for all } z \in B(q_n, \varepsilon).
\]

Let the manifold with boundary \( \mathcal{R}_n \) be the union of the pieces in \( \mathcal{R} \) corresponding to the vertices \( V_n \) of \( G_n \). We need a geodesic \( \gamma_n \) between \( p_n \) and \( q_n \), completely contained in \( \mathcal{R}_n \), which minimizes distance inside \( \mathcal{R}_n \). To prove the existence of such geodesic, consider the Riemann surface \( \Omega \)
\[
\Omega \equiv \{ z \in \mathbb{C} : 1 < |z| < \nu^2 \},
\]
where the constant \( \nu \) is chosen so that the geodesic \( \{ |z| = \nu \} \) has length \( 2\alpha \), the length of each boundary curve of \( \mathcal{R}_n \). If we join a copy of \( \Omega_0 \equiv \{ 1 < |z| \leq \nu \} \) in each boundary curve of \( \mathcal{R}_n \), we obtain a new Riemann surface \( \mathcal{R}_n^0 \). Since \( \mathcal{R}_n^0 \) is complete, there is a geodesic \( \gamma_n \) between \( p_n \) and \( q_n \) such that the length \( L_n \) of \( \gamma_n \) is equal to the distance between \( p_n \) and \( q_n \). \( \gamma_n \) is “completely contained in \( \mathcal{R}_n \)” because if \( \gamma \) enters in some copy of \( \Omega_0 \), it lies there forever.

Consider now the Fermi coordinates \( (r, t) \) [C, p.247], where \( r \in [0, L_n] \) describes the curve \( \gamma_n \), and \( t \in [-\varepsilon, \varepsilon] \) describes the orthogonal geodesics to \( \gamma_n \).

Observe that if we choose
\[
\varepsilon < \frac{1}{2} \arccosh \sqrt{\cosh \iota(\mathcal{R})} < \iota(\mathcal{R}),
\]
\((0, L_n) \times (-\varepsilon, \varepsilon) \) corresponds injectively to a region \( \Lambda_n \subset \mathcal{R} \). Let us denote by \( \pi \) this correspondence.

Assume that there exists two points \((r_1, t_1) \neq (r_2, t_2)\) in \((0, L_n) \times (-\varepsilon, \varepsilon)\), corresponding to the same \( p \in \Lambda_n \). By the definition of \( \iota(\mathcal{R}) \), it is not possible that \((r_2, t_2) \in B((r_1, 0), \iota(\mathcal{R}))\). This implies \(|r_1 - r_2| > 2\varepsilon\), because if \( d \equiv d((r_2, t_2), (r_1, 0))\), hyperbolic trigonometry [F, p.92] gives
\[
cosh(r_2 - r_1) \cosh t_2 = \cosh d \geq \cosh \iota(\mathcal{R}) > \cosh^2(2\varepsilon),
\]
and we have

\[ \cosh(r_2 - r_1) \cosh \varepsilon > \cosh(r_2 - r_1) \cosh t_2 > \cosh(2\varepsilon) \cosh \varepsilon. \]

Then

\[ |r_2 - r_1| > 2\varepsilon, \]

and

\[
d(\pi(r_1, 0), \pi(r_2, 0)) \leq d(\pi(r_1, 0), \pi(r_1, t_1)) + d(\pi(r_2, t_2), \pi(r_2, 0)) = t_1 + t_2 < 2\varepsilon.
\]

But \( |r_1 - r_2| > 2\varepsilon \) and \( d(\pi(r_1, 0), \pi(r_2, 0)) < 2\varepsilon \) contradict that \( \gamma_n \) minimizes length between \( p_n \) and \( q_n \).

Therefore, \((0, L_n) \times (-\varepsilon, \varepsilon)\) corresponds injectively to a region \( \Lambda_n \subset \mathcal{R} \).

It is easy to see that for all \( t \in (-\varepsilon, \varepsilon) \), if \( \gamma^t_n \equiv \{ \pi(r, t) : 0 \leq r \leq L_n \}, \)

\[
1 \leq u(\pi(0, t)) - u(\pi(L_n, t)) = \left| \int_{\gamma^t_n} \nabla u \, ds \right|
\leq \int_{\gamma^t_n} |\nabla u| \, ds \leq \left( \int_{\gamma^t_n} |\nabla u|^2 \, ds \right)^{1/2} \left( L(\gamma^t_n) \right)^{1/2}.
\]

But the metric in Fermi coordinates is expressed by

\[ ds^2 = \cosh^2 t \, dt^2 + dr^2, \]

and so

\[
L(\gamma^t_n) = \int_0^{L_n} \cosh t \, dr = L_n \cosh t < L_n \cosh \varepsilon \leq 2Dn \cosh \varepsilon,
\]

if \( D \) is the maximum of the diameters of the \( X \)-pieces and the *-pieces, because the diameter of \( G_n \) is \( 2n - 2 \). This gives

\[
\int_{\gamma^t_n} |\nabla u|^2 \, ds \geq \frac{1}{2Dn \cosh \varepsilon},
\]

and

\[
\int_{\Lambda_n} |\nabla u|^2 \geq \frac{\varepsilon}{Dn \cosh \varepsilon}.
\]

Therefore

\[
\int_\mathcal{R} |\nabla u|^2 \geq \frac{1}{2} \sum_{n \geq n_0} \frac{\varepsilon}{Dn \cosh \varepsilon} = \infty,
\]
and so \( u \notin HD(\mathcal{R}) \).

This proves that \( \mathcal{R} \in O_{HD} \).

To prove that \( \mathcal{R} \) has a LII we need to precise the metric relationship between \( G \) and \( \mathcal{R} \). Following Kanai’s terminology [K], we say that an application \( \varphi \), not necessarily continuous, between two metric spaces

\[
\varphi : (M_1, d_1) \rightarrow (M_2, d_2)
\]

is a “rough isometry” if the following two conditions are satisfied:

(i) There are constants \( a \geq 1 \) and \( b \geq 0 \) such that

\[
a^{-1}d_1(x, y) - b \leq d_2(\varphi(x), \varphi(y)) \leq a d_1(x, y) + b,
\]

for all \( x, y \in M_1 \).

(ii) For some \( \varepsilon > 0 \), the \( \varepsilon \)-neighborhood of \( \varphi(M_1) \) covers \( M_2 \).

A metric space \( M_1 \) is said to be “roughly isometric” to a metric space \( M_2 \) if there exists a rough isometry from \( M_1 \) into \( M_2 \). Obviously being roughly isometric is an equivalence relation between metric spaces.

It is evident that the graph \( G \) and the surface \( \mathcal{R} \) are roughly isometric.

If \( F \) is a graph, for a subset \( P \) of vertices of \( F \) we define its “boundary” \( \partial P \) by

\[
\partial P \equiv \{ v \in V(F) : \ d(v, P) = 1 \}.
\]

If \( | \cdot | \) denotes the cardinal of a subset of vertices, the “linear isoperimetric constant” of \( F \) is defined by

\[
h(F) = \sup_{P} \frac{|P|}{|\partial P|},
\]

where \( P \) ranges over all the non-empty finite subsets of vertices of \( F \).

Combining two lemmas of Kanai [K, p.401] one obtains that the surface \( \mathcal{R} \) and the graph \( G \) verify a LII simultaneously. Moreover, the definition of the linear isoperimetric constant in a graph implies that \( G \) has a LII if the binary tree \( T \) has a LII, because both have the same vertices and \( G \) has more edges.

It is not difficult to prove, by induction in the number of vertices of \( P \), that

\[
|P| \leq |\partial P|,
\]

for all non-empty finite subsets of vertices \( P \) of \( T \). This complete the proof that \( \mathcal{R} \) has a LII.
References.


