

Harmonic immersions of the Sierpinski gasket into the hyperbolic plane.

Ugo Bessi*

Abstract

Many fractals G admit a harmonic immersion into \mathbf{R}^n , i.e. an immersion which minimises a natural energy under fixed boundary conditions; we look for harmonic immersions of the Sierpinski gasket into the hyperbolic plane. We show that, given any three points $\tilde{A}, \tilde{B}, \tilde{C}$ in the hyperbolic plane there is a harmonic map bringing the three points A, B, C of the boundary of the gasket to $\tilde{A}, \tilde{B}, \tilde{C}$ respectively. Moreover, if the points $\tilde{A}, \tilde{B}, \tilde{C}$ are sufficiently close in the hyperbolic distance, then the harmonic map is unique and depends differentiably on $\tilde{A}, \tilde{B}, \tilde{C}$. Lastly, we show that, if the harmonic map ϕ is injective, then it brings geodesics of the gasket G into geodesics of $\phi(G)$.

Introduction

Many fractals $G \subset \mathbf{R}^n$ are defined in the following way ([15], [21]): we are given finitely many invertible, affine contractions

$$F_i: \mathbf{R}^n \rightarrow \mathbf{R}^n, \quad i \in \{1, \dots, p\}$$

and G is the unique non-empty, compact set of \mathbf{R}^n which satisfies

$$G = \bigcup_{i=1}^p F_i(G). \quad (1)$$

For several of these fractals, one can define a bilinear form on G which is formally similar to Dirichlet's energy on \mathbf{R}^n . More precisely, it is possible to define, in a natural way, a Borel probability measure κ on G (a.k.a. Kusuoka's measure, [18]) and a Borel field E_x of symmetric, semidefinite positive matrices from the dual space of \mathbf{R}^n to itself which induce the bilinear form

$$\begin{aligned} \mathcal{E}^d: C^1(\mathbf{R}^n, \mathbf{R}) \times C^1(\mathbf{R}^n, \mathbf{R}) &\rightarrow \mathbf{R} \\ \mathcal{E}^d(u, v) &:= \int_G (du(x)E_x, dv(x)) d\kappa(x). \end{aligned} \quad (2)$$

* Dipartimento di Matematica, Università Roma Tre, Largo S. Leonardo Murialdo, 00146 Roma, Italy.
email: bessi@matrm3.mat.uniroma3.it Work partially supported by the PRIN2023 grant "Stability in Hamiltonian Systems and beyond"

The inner product in the formula above is the natural one in the dual of \mathbf{R}^n and d is the standard differential.

Following [14] and [16], we consider E_x as a Riemannian tensor on the "cotangent space" of G and κ as a Riemannian volume on G . We have written $du(x)E_x$ because we adopt the convention that row vectors are in the cotangent space and column vectors are in the tangent one.

Having \mathcal{E}^d , we say that a function u is harmonic if

$$\mathcal{E}^d(u, \phi) = 0$$

for all $\phi \in C^1(\mathbf{R}^n, \mathbf{R})$ which vanish on the "boundary" of the fractal.

It is standard (see for instance [8]) that, when the coordinate functions are harmonic, the form \mathcal{E}^d is closable in $L^2(G, \kappa)$ and thus it extends to a Dirichlet form defined on $\mathcal{D}(\mathcal{E}^d) \subset L^2(G, \kappa)$ with $C^1(\mathbf{R}^n, \mathbf{R}) \subset \mathcal{D}(\mathcal{E}^d)$. Some famous cases of this situation are the harmonic Sierpinski gasket ([14], [16]), the level- n Sierpinski gasket ([10], [25]) and the stretched Sierpinski gasket ([2], [7]).

Saying that the coordinate functions are harmonic is equivalent to say that, under suitable boundary conditions, the identity map minimises the energy of the immersion into \mathbf{R}^n . The energy $\tilde{\mathcal{E}}$ of the immersion $\phi: G \subset \mathbf{R}^n \rightarrow \mathbf{R}^n$ is defined in the standard way: if

$$\phi(x_1, \dots, x_n) = (u_1(x_1, \dots, x_n), \dots, u_n(x_1, \dots, x_n))$$

then, for the form \mathcal{E}^d of (2),

$$\tilde{\mathcal{E}}(\phi) = \mathcal{E}^d(u_1, u_1) + \dots + \mathcal{E}^d(u_n, u_n).$$

Harmonic immersions of a Riemannian manifold into another one have been studied extensively since [11]; a natural question is whether there are energy-minimising immersions of fractals into manifolds different from \mathbf{R}^n .

In this paper we study a particular case and embed the harmonic Sierpinski gasket G into the hyperbolic plane

$$H_+ = \{x + iy : y > 0\} \tag{3}$$

endowed with the Poincaré metric. We briefly recall that, if v and w are in the tangent space $T_{(x,y)}H_+$, then their hyperbolic inner product is given by

$$(v, w)_{Hyp} = \frac{(v, w)}{y^2}$$

where (\cdot, \cdot) denotes the Euclidean inner product.

Now we can follow [11] and define the energy of an immersion

$$\phi = u + iv: G \rightarrow H_+ \tag{4}$$

with $u, v \in \mathcal{D}(\mathcal{E})$ as

$$\tilde{\mathcal{E}}(\phi) = \int_G \frac{1}{v^2} [(duE_x, du) + (dvE_x, dv)] d\kappa. \tag{5}$$

We recall from [14] that the function in $\mathcal{D}(\mathcal{E}^d)$ are continuous, which allows us to define the boundary conditions in the following way.

Definitions. Let $\tilde{A}, \tilde{B}, \tilde{C}$ be three assigned points in H_+ ; let $\{A, B, C\} \subset \mathbf{R}^2$ be the vertices of the equilateral triangle of figure 1 below; they are the "boundary" of the harmonic gasket. We define \mathcal{A} as the set of the maps $\phi = u + iv$ as in (4) such that

- 1) $u, v \in \mathcal{D}(\mathcal{E}^d)$.
- 2) $\phi(A) = \tilde{A}, \phi(B) = \tilde{B}, \phi(C) = \tilde{C}$.

For the energy $\tilde{\mathcal{E}}$ of (5) we define

$$\alpha = \inf_{\phi \in \mathcal{A}} \tilde{\mathcal{E}}(\phi). \quad (6)$$

We say that the map $\phi \in \mathcal{A}$ is harmonic if $\tilde{\mathcal{E}}(\phi) = \alpha$.

We shall prove the following theorem.

- Theorem 1.**
- 1) Given any three points $\tilde{A}, \tilde{B}, \tilde{C} \in H_+$, the number α of (6) is finite and there is at least one harmonic map, i.e. a map $\phi \in \mathcal{A}$ with $\tilde{\mathcal{E}}(\phi) = \alpha$.
 - 2) There is $r_0 > 0$ such that, if $\tilde{A}, \tilde{B}, \tilde{C}$ are contained in a ball of hyperbolic radius smaller than r_0 , then the minimal map ϕ is unique and depends differentiably on $\tilde{A}, \tilde{B}, \tilde{C}$.
 - 3) If the immersion ϕ is injective, then it is totally geodesic: in other words, curves in G which locally are geodesics are brought into curves in $\phi(G)$ which locally are geodesics.

A few comments. First, when the target space is \mathbf{R}^n , the maps F_i of (1) are closely related to the harmonic extension maps on the pre-fractals ([21], [23]); they are affine because the harmonic extension is linear. This is no more the case in our setting, which is non-linear, but we can modify (1) in the following way. We call $\phi(\tilde{A}, \tilde{B}, \tilde{C}): G \rightarrow H_+$ the harmonic map with boundary conditions $\tilde{A}, \tilde{B}, \tilde{C}$; then, for the points a, b, c of figure 1 below,

$$\phi(\tilde{A}, \tilde{B}, \tilde{C})(G) = \phi(\tilde{A}, \phi(\tilde{A}, \tilde{B}, \tilde{C})(c), \phi(\tilde{A}, \tilde{B}, \tilde{C})(b))(G) \cup$$

$$\phi(\phi(\tilde{A}, \tilde{B}, \tilde{C})(c), \tilde{B}, \phi(\tilde{A}, \tilde{B}, \tilde{C})(a))(G) \cup \phi(\phi(\tilde{A}, \tilde{B}, \tilde{C})(b), \phi(\tilde{A}, \tilde{B}, \tilde{C})(a), \tilde{C})(G).$$

We won't give a proof of this fact, which says that the image of G under ϕ is the union of the images of the three harmonic maps which bring the triple (A, B, C) into the triples

$$(\tilde{A}, \phi(\tilde{A}, \tilde{B}, \tilde{C})(c), \phi(\tilde{A}, \tilde{B}, \tilde{C})(b)), (\phi(\tilde{A}, \tilde{B}, \tilde{C})(c), \tilde{B}, \phi(\tilde{A}, \tilde{B}, \tilde{C})(a)), (\phi(\tilde{A}, \tilde{B}, \tilde{C})(b), \phi(\tilde{A}, \tilde{B}, \tilde{C})(a), \tilde{C}).$$

We also note that we do not know whether a harmonic map ϕ is injective, even if the points $\tilde{A}, \tilde{B}, \tilde{C}$ do not lie on the same geodesic of H_+ , and that's the reason for the additional hypothesis in point 3) of theorem 1. When instead of a fractal G we have a manifold a stronger fact holds, namely that being harmonic is equivalent to being totally geodesic.

Since the energy $\tilde{\mathcal{E}}$ is non linear, the compactness needed for the existence of the minimum will follow from the Ascoli-Arzelà theorem; to apply it we shall need some delicate results of [1], [5] and [17], which imply that the domain of \mathcal{E}^d embeds in a space of Hölder functions.

The paper is organised as follows: in section 1 we recall the construction of the harmonic gasket, in section 2 we recall some properties of the energy \mathcal{E} of (2). In section 3 we recall from [1], [5], and [17] some facts about function spaces defined on G , in section 4 we prove point 1) of theorem 1, in section 5 we prove point 2) and in section 6 we prove point 3).

§1

The harmonic gasket and its energy

The harmonic Sierpinski gasket. We briefly recall the definition and some properties of the Harmonic Sierpinski Gasket in \mathbf{R}^2 ; we refer the reader to [14] and [16] for first hand information. We set

$$T_1 = \begin{pmatrix} \frac{3}{5} & 0 \\ 0 & \frac{1}{5} \end{pmatrix}, \quad T_2 = \begin{pmatrix} \frac{3}{10} & \frac{\sqrt{3}}{10} \\ \frac{\sqrt{3}}{10} & \frac{1}{2} \end{pmatrix}, \quad T_3 = \begin{pmatrix} \frac{3}{10} & -\frac{\sqrt{3}}{10} \\ -\frac{\sqrt{3}}{10} & \frac{1}{2} \end{pmatrix},$$

$$A = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 \\ \frac{1}{\sqrt{3}} \end{pmatrix}, \quad C = \begin{pmatrix} 1 \\ -\frac{1}{\sqrt{3}} \end{pmatrix}$$

and

$$F_1(x) = T_1(x), \quad F_2(x) = B + T_2(x - B), \quad F_3(x) = C + T_3(x - C).$$

Referring to figure 1 below, F_1 brings the equilateral triangle ABC into Abc ; F_2 brings ABC into Bac and F_3 brings ABC into Cba .

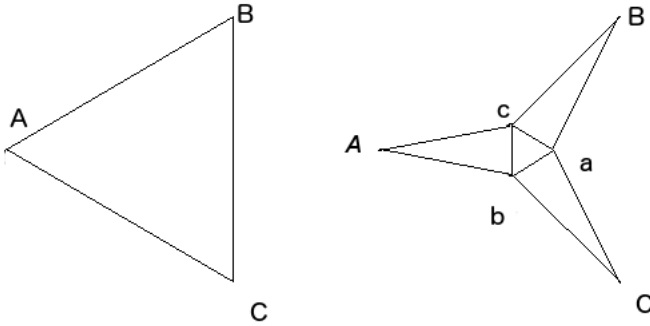


Figure 1

The bilinear form. We recall a standard fact from [14] (for alternative approaches using dynamical systems, see [12], [19], [20] or [6].) Let $\Lambda^1(\mathbf{R}^2)$ denote the dual space of \mathbf{R}^2 ; we endow it with the inner

product induced by the standard inner product of \mathbf{R}^2 , i.e.

$$(dx, dx) = 1 = (dy, dy) \quad \text{and} \quad (dx, dy) = 0.$$

In coordinates, for us the vectors of \mathbf{R}^2 will be column vectors and those of $\Lambda^1(\mathbf{R}^2)$ will be row vectors.

We denote by M the space of symmetric matrices from $\Lambda^1(\mathbf{R}^2)$ to itself. Then we can define, in a natural way, a Borel probability measure κ on G (which is called Kusuoka's measure, [18]) and a Borel function $E: G \rightarrow M$; it is known ([14]) that E_x is a one-dimensional projection for κ -a.e. $x \in G$.

These objects induce a bilinear form

$$\begin{aligned} \mathcal{E}: C(G, \Lambda^1(\mathbf{R}^2)) \times C(G, \Lambda^1(\mathbf{R}^2)) &\rightarrow \mathbf{R} \\ \mathcal{E}(u, v) &= \int_G (u(x)E_x, v(x))d\kappa(x). \end{aligned} \tag{1.1}$$

Note again that the symmetric matrix E_x appears on the right because $u(x)$ is a row vector.

It turns out ([14], but see also the next section) that the bilinear form of (2), i.e.

$$\begin{aligned} \mathcal{E}^d: C^1(\mathbf{R}^2, \mathbf{R}) \times C^1(\mathbf{R}^2, \mathbf{R}) &\rightarrow \mathbf{R} \\ \mathcal{E}^d(f, g) &:= \mathcal{E}(df, dg) \end{aligned} \tag{1.2}$$

is closable in $L^2(G, \kappa)$; its closure is a local Dirichlet form on G .

The Lebesgue space of one-forms. We call \mathcal{L} be the space of the Borel functions $u: G \rightarrow \Lambda^1(\mathbf{R}^2)$ such that $u(x) \in \text{Rank}(E_x)$ for κ -a.e. $x \in G$ and we define

$$L^2(G, \Lambda^1(G), \kappa) := \{u \in \mathcal{L} : \mathcal{E}(u, u) < +\infty\}.$$

We forego the easy proof that $L^2(G, \Lambda^1(G), \kappa)$ is a Hilbert space for the natural inner product

$$(u, v)_{L^2} := \mathcal{E}(u, v) = \int_G (u(x)E_x, v(x))d\kappa(x). \tag{1.3}$$

We shall call $L^\infty(G, \Lambda^1(G), \kappa)$ the space of the elements of \mathcal{L} which are bounded κ -a.e., with the norm of the essential supremum.

The gasket is totally geodesic. In [16] and [14] it is proven that any two points $x, y \in G$ are joined by a shortest curve with image in G ; we define $d(x, y)$ as the length of this curve. It is standard that d is a distance on G ; in [14] it is shown that the two topologies on G , the one induced by the immersion in \mathbf{R}^2 and the one induced by d , coincide. From now on, we shall work on the measured metric space (G, d, κ) .

§2

Differential forms on the gasket

Differential and codifferential. First of all, we recall an integration by parts formula from [24].

We consider the following three vectors of \mathbf{R}^2 , which play the rôle of the exterior normal to the fractal:

$$\nu_A = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \quad \nu_B = \begin{pmatrix} \frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix}, \quad \nu_C = \begin{pmatrix} \frac{1}{2} \\ -\frac{\sqrt{3}}{2} \end{pmatrix}.$$

If $u \in C^2(\mathbf{R}^2, \mathbf{R})$ we define

$$\Delta u = \text{Tr}(E_x D^2 u)$$

where Tr denotes the trace, $D^2 u$ is the matrix of second order derivatives and E_x is as in (1.1). Let $u \in C^2(\mathbf{R}^2, \mathbf{R})$ and $v \in C^1(\mathbf{R}^2, \mathbf{R})$; denoting by $\partial_\nu u$ the directional derivative, we recall from [24] (an alternative proof is in [7]) that

$$\mathcal{E}(du, dv) = - \int_G \Delta u \cdot v d\kappa + \sum_{j \in \{A, B, C\}} v(j) \partial_{\nu_j} u(j). \quad (2.1)$$

This yields ([8]) a formula for the codifferential of C^1 one-forms; namely, if $u = adx + bdy$ with $a, b \in C^1(\mathbf{R}^2, \mathbf{R})$ and E_x is as in (1.1), we define

$$\bar{d}_G u = -[(E_x da, dx) + (E_x db, dy)]. \quad (2.2)$$

Then, for all $\phi \in C^1(\mathbf{R}^2, \mathbf{R})$ we have

$$\mathcal{E}(u, d\phi) = \int_G (\bar{d}_G u) \cdot \phi d\kappa + \sum_{j \in \{A, B, C\}} \phi(j) u(j) (\nu_j) \quad (2.3)$$

where $u(j)(\nu_j)$ means that we apply the one-form $u(j) \in \Lambda^1(\mathbf{R}^2)$ to the vector $\nu_j \in \mathbf{R}^2$.

Now we can define differential and codifferential in the spirit of [9].

Definitions. We say that the function $\bar{d}_G u \in L^2(G, \mathbf{R}, \kappa)$ is the weak codifferential of $u \in L^2(G, \Lambda^1(G), \kappa)$ if, for all $\phi \in C^1(\mathbf{R}^2, \mathbf{R})$, the following integration by parts formula holds.

$$\mathcal{E}(u, d\phi) = \int_G (\bar{d}_G u) \cdot \phi d\kappa. \quad (2.4)$$

Since $C^1(\mathbf{R}^2, \mathbf{R})$ is dense in $L^2(G, \mathbf{R}, \kappa)$, the formula above implies easily that the weak codifferential, if it exists, is unique; we call $\mathcal{D}(\bar{d}_G) \subset L^2(G, \Lambda^1(G), \kappa)$ the domain of the codifferential.

For the vectors ν_A, ν_B, ν_C defined at the beginning of this section we set

$$C_{abs}^1(\mathbf{R}^2, \Lambda^1(\mathbf{R}^2)) := \{u \in C^1(\mathbf{R}^2, \Lambda^1(\mathbf{R}^2)) : u(j)(\nu_j) = 0 \text{ for } j \in \{A, B, C\}\}.$$

Now formula (2.3) implies that $C_{abs}^1(\mathbf{R}^2, \Lambda^1(\mathbf{R}^2)) \subset \mathcal{D}(\bar{d}_G)$; moreover, if $u \in C_{abs}^1(\mathbf{R}^2, \Lambda^1(\mathbf{R}^2))$, then its weak codifferential is given by (2.2). The notation for C_{abs}^1 is taken from the "absolute" boundary conditions of [22], which generalise Neumann's to k -forms.

We just saw that $C_{abs}^1(\mathbf{R}^2, \Lambda^1(\mathbf{R}^2)) \subset \mathcal{D}(\bar{d}_G)$; since clearly

$$\{E_x u(x) : u \in C_{abs}^1(\mathbf{R}^2, \mathbf{R})\}$$

is dense in $L^2(G, \Lambda^1(G), \kappa)$, we get that $\mathcal{D}(\bar{d}_G)$ is dense in $L^2(G, \Lambda^1(\mathbf{R}^2), \kappa)$.

We say that the one-form $d_w v \in L^2(G, \Lambda^1(G), \kappa)$ is the weak differential of $v \in L^2(G, \mathbf{R}, \kappa)$ if, for all $u \in \mathcal{D}(\bar{d}_G)$, the following integration by parts formula holds.

$$\mathcal{E}(u, d_w v) = \int_G (\bar{d}_G u) \cdot v d\kappa. \quad (2.5)$$

Since we saw above that $\mathcal{D}(\bar{d}_G)$ is dense in $L^2(G, \Lambda^1(G), \kappa)$, we get as above that the weak differential is unique.

We call $\mathcal{D}(d_w) \subset L^2(G, \mathbf{R}, \kappa)$ the domain of the weak differential; the definition of the codifferential immediately implies that $C^1(\mathbf{R}^2, \mathbf{R}) \subset \mathcal{D}(d_w)$ and that the weak differential $d_w u$ satisfies

$$d_w u = E_x du(x)$$

when $u \in C^1(\mathbf{R}^2, \mathbf{R})$. From now on, we shall often write du instead of $d_w u$.

We recall the proof that the weak differential d is closed in $L^2(G, \mathbf{R}, \kappa)$.

Lemma 2.1. *Let $\{v_n\}_{n \geq 1} \subset \mathcal{D}(d)$ be such that*

- 1) $v_n \rightarrow v$ in $L^2(G, \mathbf{R}, \kappa)$ and
- 2) $dv_n \rightarrow V$ in $L^2(G, \Lambda^1(G), \kappa)$.

Then, $v \in \mathcal{D}(d)$ and $dv = V$.

Proof. Since $v_n \in \mathcal{D}(d)$, for all $u \in \mathcal{D}(\bar{d}_G)$ we have

$$\mathcal{E}(dv_n, u) = \int_G v_n (\bar{d}_G u) d\kappa.$$

Recall that the inner product in $L^2(G, \Lambda^1(G), \kappa)$ is that of (1.3); now points 1) and 2) above imply respectively the first and second limit below.

$$\int_G v_n (\bar{d}_G u) d\kappa \rightarrow \int_G v (\bar{d}_G u) d\kappa \quad \text{and} \quad \mathcal{E}(dv_n, u) \rightarrow \mathcal{E}(V, u).$$

From the last two formulas we get that, for all $u \in \mathcal{D}(\bar{d}_G)$,

$$\mathcal{E}(V, u) = \int_G v (\bar{d}_G u) d\kappa$$

which implies the thesis by (2.5).

\\

Now we consider the subspace

$$\{(u, du) : u \in \mathcal{D}(d)\} \subset L^2(G, \mathbf{R}, \kappa) \times L^2(G, \Lambda^1(G), \kappa)$$

which we endow with the natural inner product

$$(u, v)_{\mathcal{D}(\mathcal{E}^d)} := (u, v)_{L^2(G, \mathbf{R}, \kappa)} + \mathcal{E}(du, dv). \quad (2.6)$$

The last lemma implies that this subspace is closed for the topology induced by (2.6), which is equivalent to say that the bilinear form

$$\mathcal{E}^d: \mathcal{D}(d) \times \mathcal{D}(d) \rightarrow \mathbf{R}$$

$$\mathcal{E}^d(u, v) = \int_G (du \cdot E_x, dv) d\kappa$$

is closed in $L^2(G, \kappa)$, i.e. it is a Dirichlet form. Since it is proven in [9] that the operator d is local, we get that the Dirichlet form \mathcal{E}^d is local. Defining \mathcal{E}^d as in the formula above obviously implies that $\mathcal{D}(\mathcal{E}^d) = \mathcal{D}(d)$ and we shall often use this fact in the following. In lemma 2.2 below we are going to show that $C^1(\mathbf{R}^2, \mathbf{R})$ is dense in $\mathcal{D}(d)$ for the topology of (2.6); in other words, the closure of $\{(u, du) : u \in C^1(\mathbf{R}^2, \mathbf{R})\}$ is the graph of \mathcal{E}^d on $\mathcal{D}(d)$.

The Poincaré inequality. For starters, we recall a standard notation: if $u \in L^1(G, \kappa)$ and $F \subset G$ is a Borel set with $\kappa(F) > 0$, we set

$$u_F = \frac{1}{\kappa(F)} \int_F u d\kappa. \quad (2.7)$$

From now on, $B(x, r)$ will denote the open ball of radius r and centre $x \in G$ with respect to the geodesic distance d on G . Proposition 4.26 of [14] says that there are $C > 0$ and $\lambda \geq 1$ such that for all $u \in \mathcal{D}(d)$, $x \in G$ and $r > 0$ we have

$$\int_{B(x, r)} |u - u_{B(x, r)}|^2 d\kappa \leq Cr^2 \int_{B(x, \lambda r)} (E_x du, du) d\kappa. \quad (2.8)$$

We recall from theorem 2.6 of [14] that on G there is a distance $\mathcal{R}^{\frac{1}{2}}$ (which is called the resistance distance) such that, if $u \in \mathcal{D}(d)$, the first equality below holds; for the second one, C is the $\mathcal{R}^{\frac{1}{2}}$ -diameter of G , which is finite since G is compact and $\mathcal{R}^{\frac{1}{2}}$ induces the same topology as the geodesic distance d .

$$|u(x) - u(y)|^2 \leq \mathcal{R}(x, y) \mathcal{E}(du, du) \leq C^2 \mathcal{E}(du, du). \quad (2.9)$$

A consequence of the first inequality above is that the functions $u \in \mathcal{D}(\mathcal{E}^d)$ are continuous; from the second inequality we get there is $C > 0$ such that, for all $u \in \mathcal{D}(\mathcal{E}^d)$,

$$\|u\|_{L^\infty} \leq C \left[\min(|u(A)|, |u(B)|, |u(C)|) + \sqrt{\mathcal{E}^d(u, u)} \right].$$

This implies that, if $\alpha, \beta, \gamma \in \mathbf{R}$, then the convex set of the $u \in \mathcal{D}(\mathcal{E}^d)$ such that $u(A) = \alpha$, $u(B) = \beta$ and $u(C) = \gamma$ is closed for the topology of (2.6). The formula above also implies that there is $C_1 > 0$, depending only on α, β and γ , such that, for all such u ,

$$\int_G |u|^2 d\kappa \leq C_1 [1 + \mathcal{E}(du, du)]. \quad (2.10)$$

A particular case is the subspace H_0 of the $\phi \in \mathcal{D}(\mathcal{E}^d)$ such that $\phi(A) = \phi(B) = \phi(C) = 0$; we saw above that it is closed and the inequality before (2.10) implies that the norm $\|\cdot\|_{H_0}$ on H_0 defined by

$$\|\phi\|_{H_0}^2 := \mathcal{E}(d\phi, d\phi)$$

is equivalent to

$$\|\phi\|_{\mathcal{D}(\mathcal{E}^d)}^2 = \|\phi\|_{L^2}^2 + \mathcal{E}(d\phi, d\phi).$$

We shall use often the norm $\|\cdot\|_{H_0}$, which is the one induced by the inner product

$$(u, v)_{H_0} := \mathcal{E}(du, dv). \quad (2.11)$$

Since we saw above that H_0 is closed in the Hilbert space $\mathcal{D}(\mathcal{E}^d)$, we get that H_0 is Hilbert for the inner product of (2.11).

Other properties of the differential. We prove that test functions are dense.

Lemma 2.2. *The space $C^1(\mathbf{R}^2, \mathbf{R})$ is dense in $\mathcal{D}(d) = \mathcal{D}(\mathcal{E}^d)$ with the topology of (2.6).*

Proof. When we defined the operator d we saw that $C^1(\mathbf{R}^2, \mathbf{R}) \subset \mathcal{D}(d)$; now we show that the orthogonal space of $C^1(\mathbf{R}^2, \mathbf{R})$ in $\mathcal{D}(\mathcal{E}^d)$ with the natural inner product of (2.6) is reduced to zero.

Let $u \in \mathcal{D}(d)$ be such that

$$\int_G u \cdot \phi d\kappa + \int_G (E_x du, d\phi) d\kappa = 0 \quad \text{for all } \phi \in C^1(\mathbf{R}^2, \mathbf{R}). \quad (2.12)$$

First of all we assert that, with the notation of (2.7), $u_G = 0$. Let us suppose by contradiction that this is not the case; then, we can add a constant to ϕ in such a way that the first integral above is zero; since $d(\phi + c) = d\phi$ we have that

$$\int_G (E_x du, d\phi) d\kappa = 0 \quad \text{for all } \phi \in C^1(\mathbf{R}^2, \mathbf{R}). \quad (2.13)$$

This means that u is harmonic; by [14], this implies that u is affine. Since in the integration by parts formula (2.13) there are no boundary terms, formula (2.1) implies that $\partial_{\nu_A} u(A) = \partial_{\nu_B} u(B) = \partial_{\nu_C} u(C) = 0$; since u is affine, we get that u is constant; now (2.12) implies that

$$\int_G u \cdot \phi d\kappa = 0$$

for all $\phi \in C^1(\mathbf{R}^2, \mathbf{R})$, i.e. that $u \equiv 0$, which contradicts $u_G \neq 0$.

Thus, if u is orthogonal to C^1 we must have that $u_G = 0$ and now it suffices to show that the functions C^1 with zero average are dense among the functions $u \in \mathcal{D}(\mathcal{E}^d)$ with $u_G = 0$. Formula (2.8) implies that $\mathcal{E}^d(\cdot, \cdot)$ is an equivalent inner product in the subspace of the functions u with $u_G = 0$ and we are reduced to show that, if u with $u_G = 0$ satisfies (2.13) for all $\phi \in C^1(\mathbf{R}^2, \mathbf{R})$ with $\phi_G = 0$, then $u \equiv 0$. Since again $d(\phi + c) = d\phi$, we get that u satisfies (2.13) for all $\phi \in C^1(\mathbf{R}^2, \mathbf{R})$; this implies as above that u is constant; since $u_G = 0$, we get that $u \equiv 0$, as we wanted.

\\

Lemma 2.3. *1) Let $u \in \mathcal{D}(d)$ and $v \in \mathcal{D}(\bar{d}_G)$; let us suppose in addition that $u \in L^\infty(G, \mathbf{R}, \kappa)$ and $v \in L^\infty(G, \Lambda^1(G), \kappa)$. Then, $u \cdot v \in \mathcal{D}(\bar{d}_G)$ and*

$$\bar{d}_G(u \cdot v) = -(E_x du, v) + u \bar{d}_G v. \quad (2.14)$$

The signs in the formula above are due to the fact that \bar{d}_G is minus the divergence.

2) Let $\psi \in C^1(\mathbf{R}, \mathbf{R})$ have bounded derivative and let $u \in \mathcal{D}(d)$. Then, $\psi \circ u \in \mathcal{D}(d)$ and

$$d(\psi \circ u) = \psi' \circ u \cdot du. \quad (2.15)$$

3) Let $u \in \mathcal{D}(d)$, let $c \in \mathbf{R}$ and let us define

$$v(x) = \max(c, u(x)).$$

Then, $v \in \mathcal{D}(d) = \mathcal{D}(\mathcal{E}^d)$ and for κ -a.e. $x \in G$ we have

$$dv(x) = \begin{cases} 0 & \text{if } u(x) \leq c \\ du(x) & \text{if } u(x) > c. \end{cases}$$

Proof. We begin with point 1); note that, by our hypotheses, the right hand side of (2.14) is in $L^2(G, \mathbf{R}, \kappa)$. For starters, we suppose that $u \in C^1(\mathbf{R}^2, \mathbf{R})$.

Let $\phi \in C^1(\mathbf{R}^2, \mathbf{R})$; the first equality below follows from the definition of \mathcal{E} in (1.1); the second one is obvious and the third one comes from the Leibnitz rule for C^1 functions; the last one is the integration by parts formula (2.4), which we can apply because $v \in \mathcal{D}(\bar{d}_G)$ by assumption.

$$\begin{aligned} \mathcal{E}(uv, d\phi) &= \int_G (E_x(uv), d\phi) d\kappa = \int_G (E_x v, u d\phi) d\kappa = \\ &= \int_G (E_x v, d(u\phi)) d\kappa - \int_G \phi (E_x v, du) d\kappa = \\ &= \int_G (\bar{d}_G v) u \phi d\kappa - \int_G \phi (E_x v, du) d\kappa. \end{aligned}$$

This proves (2.14) when $u \in C^1(\mathbf{R}^2, \mathbf{R})$; we recall the standard density argument for the general case. By lemma 2.2, $C^1(\mathbf{R}^2, \mathbf{R})$ is dense in $\mathcal{D}(d)$ with the topology of (2.6); thus, we can find a sequence $\{u_n\} \subset C^1(\mathbf{R}^2, \mathbf{R})$ such that

$$u_n \rightarrow u, \quad \text{and} \quad du_n \rightarrow du \quad \text{in } L^2. \quad (2.16)$$

Since $v, \phi, d\phi, E_x \in L^\infty$, we can take limits in the left and right hand sides of

$$\mathcal{E}(u_n v, d\phi) = \int_G (\bar{d}_G v) u_n \phi d\kappa - \int_G \phi (E_x v, du_n) d\kappa$$

getting (2.14) in the general case.

For the proof of (2.15) we consider the sequence $\{u_n\}$ of (2.16); up to taking a subsequence, which we denote with the same index, we can suppose that the convergences of (2.16) are κ -a.e. and dominated. Since ψ' is bounded, this implies that

$$\psi \circ u_n \rightarrow \psi \circ u \quad \text{and} \quad d(\psi \circ u_n) \rightarrow d(\psi \circ u) \quad (2.17)$$

κ -a.e. and dominated and thus in $L^2(G, \kappa)$. Since $\psi \circ u_n \in C^1(\mathbf{R}^2, \mathbf{R})$, the definition of the codifferential implies that for all $\phi \in \mathcal{D}(\bar{d}_G)$ we have the first equality below, while the second one comes from the chain rule for C^1 functions.

$$\int_G \psi \circ u_n \cdot \bar{d}_G \phi d\kappa = \int_G (E_x d(\psi \circ u_n), \phi) d\kappa = \int_G \psi' \circ u_n (E_x du_n, \phi) d\kappa.$$

Using (2.17) and the fact that $\psi' \circ u_n$ is bounded, we can take limits in the formula above and get

$$\int_G \psi \circ u \cdot \bar{d}_G \phi d\kappa = \int_G \psi' \circ u (E_x du, \phi) d\kappa$$

which implies (2.15) by the definition of d in (2.5).

As for point 3), it just says that the operator d is local, a fact proven in [9]. Let us recall a different proof, equally standard. We find a sequence $\{\psi_n\} \subset C^1(\mathbf{R}, \mathbf{R})$ with bounded derivatives such that $\psi_n(s) = c$ when $s \leq c$ and $\psi_n(s) \rightarrow s$ when $s > c$; we can also require that $\psi'_n(s) = 1$ when $s \geq c + \frac{1}{n}$. We apply point 2) to the composition $\psi_n \circ u$ and then we take limits.

\\

The codifferential in the relative sense and the Dirichlet problem. We define the codifferential in the relative sense, i.e we ask that (2.4) holds, but only for test functions which vanish on the boundary of the fractal.

Definitions. Let $v \in L^2(G, \Lambda^1(\mathbf{R}^2), \kappa)$; we say that $\bar{d}_G v \in L^2(G, \mathbf{R}, \kappa)$ is the codifferential in the relative sense if formula (2.4) holds for all $\phi \in C^1(\mathbf{R}^2, \mathbf{R})$ such that $\phi(A) = \phi(B) = \phi(C) = 0$. Formula (2.3) implies easily that, if $u \in C^1(\mathbf{R}^2, \Lambda^1(\mathbf{R}^2))$, then it has a codifferential in the relative sense which is given by (2.2).

Let $f \in L^1(G, \mathbf{R}, \kappa)$, let $g \in L^2(G, \Lambda^1(G), \kappa)$ and let H_0 be the space defined before (2.11); we say that $u \in H_0$ is a weak solution of the equation

$$-\Delta u = f + \bar{d}_G g \tag{2.18}$$

if for all $\phi \in C^1(\mathbf{R}^2, \mathbf{R})$ with $\phi(A) = \phi(B) = \phi(C) = 0$ we have that, for the weak differential operator d of (2.5),

$$\mathcal{E}(du, d\phi) = \int_G f \cdot \phi d\kappa + \int_G (d\phi, g) d\kappa. \tag{2.19}$$

We have the following.

Lemma 2.4. *Let $f \in L^1(G, \mathbf{R}, \kappa)$ and let $g \in L^2(G, \Lambda^1(G), \kappa)$; then, equation (2.18) has a unique weak solution $u \in H_0$. Moreover, there is $C > 0$, independent of f and g , such that*

$$\|u\|_{H_0} \leq C (\|f\|_{L^1} + \|g\|_{L^2}).$$

Proof. Recall that the space H_0 is Hilbert for the inner product $(\cdot, \cdot)_{H_0}$ of (2.11). The inequality after (2.9) implies that, if $u \in H_0$,

$$\|u\|_{L^\infty} \leq C \|u\|_{H_0}.$$

Together with Hölder's inequality this implies that the linear functional

$$L: H_0 \rightarrow \mathbf{R}$$

$$L: \phi \rightarrow \int_G f \cdot \phi d\kappa + \int_G (d\phi, g) d\kappa$$

is continuous for $\|\cdot\|_{H_0}$; actually, by the last two formulas a bound on the operator norm of L is $C(\|f\|_{L^1} + \|g\|_{L^2})$. Now the thesis follows from the standard Lax-Milgram argument: the functional L is represented by a unique $u \in H_0$, which satisfies (2.19) for all $\phi \in C^1(\mathbf{R}^2, \mathbf{R})$ with $\phi(A) = \phi(B) = \phi(C) = 0$.

\\

§3

The function spaces

As in the last section, we consider the harmonic gasket G with the geodesic distance d and Kusuoka's measure κ . By (2.9), the functions of $\mathcal{D}(\mathcal{E}^d)$ are $\frac{1}{2}$ -Hölder with respect to the distance $\mathcal{R}^{\frac{1}{2}}$; since this distance does not coincide with the geodesic distance d , we recall some results from [1] and [5].

Volume doubling. We recall from [16] and theorem 4.25 of [14] that (G, d, κ) has the volume doubling property: there is $C > 0$ such that, for all $x \in G$ and $r > 0$, we have

$$0 < \kappa(B(x, 2r)) \leq C\kappa(B(x, r)).$$

A consequence is that there are $C, Q > 0$ such that, for all $0 < r \leq R$ we have

$$\frac{\kappa(B(x, R))}{\kappa(B(x, r))} \leq \left(\frac{R}{r}\right)^Q.$$

The number Q , which in many formulas takes the rôle of the dimension, is computed in [14], where it is shown that

$$Q = \log_5 15 \in (1, 2). \tag{3.1}$$

Cheeger's energy. We denote by $\text{Lip}(G, \mathbf{R})$ the space of Lipschitz functions on G ; if $u: G \rightarrow \mathbf{R}$ is Lipschitz, we define the local Lipschitz constant as

$$\text{Lip}(u, x) := \limsup_{r \rightarrow 0} \sup_{y \in B(x, r)} \frac{|u(x) - u(y)|}{r}.$$

We are going to relax this function in $L^2(G, \mathbf{R}, \kappa)$. Let $u \in L^2(G, \mathbf{R}, \kappa)$ and let us suppose that there is $h \in L^2(G, \mathbf{R}, \kappa)$ and a sequence $\{u_n\} \subset \text{Lip}(G, \mathbf{R})$ such that the following two points hold.

- 1) $u_n \rightarrow u$ in $L^2(G, \kappa)$ and
- 2) $\text{Lip}(u_n, \cdot) \rightarrow h$ in $L^2(G, \kappa)$.

If the set of such functions h is nonempty, then ([4], [13]) it contains a unique element $|Du|_*$ which is minimal in the following sense: for all h as in points 1) and 2) above, we have

$$\| |Du|_* \|_{L^2(G, \mathbf{R}, \kappa)} \leq \|h\|_{L^2(G, \mathbf{R}, \kappa)} \quad \text{and} \quad |Du|_*(x) \leq h(x) \quad \text{for } \kappa\text{-a.e } x \in G.$$

We define Cheeger's energy as

$$Ch(u) = \int_G |Du|_*^2 d\kappa$$

if $|Du|_*$ exists, and we set $Ch(u) = +\infty$ otherwise.

Clearly, the minimality of $|Du|_*$ implies that, if $u \in \text{Lip}(G, \mathbf{R})$, then

$$|Du|_*(x) \leq \text{Lip}(u, x) \quad \text{for } \kappa\text{-a.e } x \in G.$$

A result of [17] says that, on the harmonic gasket with the geodesic distance and Kusuoka's measure, $\mathcal{D}(Ch) = \mathcal{D}(\mathcal{E}^d)$ and that for all $u \in \mathcal{D}(\mathcal{E}^d) = \mathcal{D}(d)$ we have that

$$|Du|_*^2 = (du(x)E_x, du(x)) \quad \text{for } \kappa\text{-a.e } x \in G. \quad (3.2)$$

The last formula and the Poincaré inequality (2.8) imply that, if $u \in \mathcal{D}(\mathcal{E}^d) = \mathcal{D}(d)$,

$$\int_{B(x,r)} |u - u_{B(x,r)}|^2 d\kappa \leq Cr^2 \int_{B(x,\lambda r)} |Du|_*^2 d\kappa(x). \quad (3.3)$$

The Korevaar-Schoen energy. We recall some definitions and results from [1] and [5].

First of all, if $u \in L^2(G, \mathbf{R}, \kappa)$, we define

$$E_{2,1}(u, r) := \int_G d\kappa(x) \frac{1}{\kappa(B(x, r))} \int_{B(x,r)} \frac{|f(y) - f(x)|^2}{r^2} d\kappa(y)$$

and

$$\|u\|_{\mathcal{B}^{1,2}}^2 := \|u\|_{L^2(G, \kappa)}^2 + \limsup_{r \rightarrow 0} E_{2,1}(u, r). \quad (3.4)$$

This is known as the Korevaar-Schoen norm of u ; the space \mathcal{KS} of the functions u such that $\|u\|_{\mathcal{B}^{1,2}} < +\infty$ is a Banach space for the norm $\|\cdot\|_{\mathcal{B}^{1,2}}$ of (3.4).

We saw at the beginning of this section that (G, d, κ) has the volume doubling property; since the Poincaré inequality of (3.3) holds, we are in the hypotheses of theorem 3.1 of [1]; by point (v) of this theorem, $\|\cdot\|_{\mathcal{B}^{1,2}}$ is equivalent to the (apparently stronger) norm $\|\cdot\|_{\bullet}$ defined by

$$\|u\|_{\bullet}^2 := \|u\|_{L^2(G, \kappa)}^2 + \sup_{r>0} E_{2,1}(u, r). \quad (3.5)$$

For the number Q of (3.1) we set

$$\alpha := 1 - \frac{Q}{2} \in \left(0, \frac{1}{2}\right)$$

and we recall the standard notation

$$|u|_{0,\alpha} := \sup \left\{ \frac{|u(x) - u(y)|}{d(x, y)^\alpha} : x \neq y \in G \right\}$$

for the Hölder seminorm, and

$$\|u\|_{0,\alpha} := \|u\|_{\text{sup}} + |u|_{0,\alpha}$$

for the Hölder norm.

Now theorem 3.2 of [5] says that there is $C > 0$ such that, for all $u \in \mathcal{KS}$,

$$\|u\|_{0,\alpha} \leq C \|u\|_{\bullet}.$$

We saw before (3.5) that $\|\cdot\|_{\bullet}$ and $\|\cdot\|_{\mathcal{B}^{1,2}}$ are equivalent; thus, possibly enlarging C , for all $u \in \mathcal{KS}$,

$$\|u\|_{0,\alpha} \leq C \|u\|_{\mathcal{B}^{1,2}}. \quad (3.6)$$

In [1] it is shown that the Γ -limit of $E_{2,1}(\cdot, r)$ as $r \rightarrow 0$ is comparable to Cheeger's energy, which implies the second inequality below; for the first inequality we use theorem 3.1 of [1], which says that the Γ -limit is comparable to $\limsup_{r \rightarrow 0} E_{2,1}(u, r)$.

$$\limsup_{r \rightarrow 0} E_{2,1}(u, r) \leq C \cdot \Gamma - \lim_{r \rightarrow 0} E_{2,1}(u, r) \leq C_1 \cdot Ch(u).$$

By (3.2) this implies that $\mathcal{D}(d) \subset \mathcal{KS}$; together with (3.4), it also implies the second inequality below for all $u \in \mathcal{D}(d)$; the first inequality is (3.6) and the equality at the end comes again from (3.2).

$$\|u\|_{0,\alpha} \leq C \|u\|_{\mathcal{B}^{1,2}} \leq C \left[\|u\|_{L^2(G,\kappa)}^2 + Ch(u) \right]^{\frac{1}{2}} = C \left[\|u\|_{L^2(G,\kappa)}^2 + \mathcal{E}(du, du) \right]^{\frac{1}{2}}. \quad (3.7)$$

This formula will be essential in the next sections.

§4

Existence of the minimum

Invariance under automorphisms. We recall the classical proof that the functional of (5) is invariant under automorphisms of H_+ .

Lemma 4.1. *Let $\zeta: H_+ \rightarrow H_+$ be an automorphism and let $\tilde{\mathcal{E}}$ be as in (5). Then, for all $u + iv \in \mathcal{D}(d) \oplus i\mathcal{D}(d)$ we have that*

$$\tilde{\mathcal{E}}(\zeta \circ (u + iv)) = \tilde{\mathcal{E}}(u + iv).$$

Proof. After identifying \mathbf{C} with \mathbf{R}^2 , we write $\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} u^1 \\ u^2 \end{pmatrix}$, $\zeta = \begin{pmatrix} \zeta^1 \\ \zeta^2 \end{pmatrix}$ and we denote by $g_{i,j}$ the Riemannian tensor of H_+ . We denote by $\begin{pmatrix} x^1 \\ x^2 \end{pmatrix}$ the coordinates on G and, for the matrix field E_x of (2) we set $E(x^1, x^2) = a^{i,j}(x^1, x^2)$ and adopt the usual convention of summation on repeated indices. With this notation, the definition (5) becomes

$$\tilde{\mathcal{E}}(u^1 + iu^2) = \int_G a^{i,j} \cdot g_{i,k} \circ (u + iv) \partial_i u^l \partial_j u^k d\kappa.$$

This implies the last equality below; together with the chain rule, it implies the first one; the second equality follows from the fact that ζ , being an automorphism, preserves the Riemannian tensor of H_+ .

$$\begin{aligned}\tilde{\mathcal{E}}(\zeta \circ (u^1 + iu^2)) &= \int_G a^{i,j} \cdot g_{l,k} \circ \zeta \circ (u^1 + iu^2) \cdot \partial_r \zeta^l|_{u^1+iu^2} \partial_i u^r \partial_s \zeta^k|_{u+iv} \partial_j u^s d\kappa = \\ &= \int_G a^{i,j} \cdot g_{r,s} \circ (u + iv) \partial_i u^r \partial_j u^s d\kappa = \tilde{\mathcal{E}}(u^1 + iu^2).\end{aligned}$$

\\

The infimum is finite. We consider three points of H_+ , not necessarily distinct; however, if they are the vertices of a non-degenerate Euclidean triangle, we ask that they are labeled in the clockwise way, as in figure 1; more precisely, the automorphism bringing \tilde{A} , \tilde{B} to i , it respectively with $t > 1$ brings \tilde{C} to the right hand plane.

$$\tilde{A} = a_1 + ib_1, \quad \tilde{B} = a_2 + ib_2, \quad \tilde{C} = a_3 + ib_3. \quad (4.1)$$

We briefly prove that the number α of (6) is finite: on one side, $\alpha \geq 0$ since $\tilde{\mathcal{E}} \geq 0$. In order to show that $\alpha < +\infty$, we build $\psi_0 \in \mathcal{A}$ such that $\tilde{\mathcal{E}}(\psi_0) < +\infty$.

We consider the canonical map

$$J: \mathbf{R}^2 \rightarrow \mathbf{C}, \quad J: (x, y) \rightarrow x + iy.$$

Let $\tilde{\psi}_0: \mathbf{R}^2 \rightarrow \mathbf{R}^2$ be the unique affine map such that

$$\tilde{\psi}_0(A) = J^{-1}(\tilde{A}), \quad \tilde{\psi}_0(B) = J^{-1}(\tilde{B}), \quad \tilde{\psi}_0(C) = J^{-1}(\tilde{C}).$$

We define $\psi_0 = u_0 + iv_0$ as

$$\psi_0 = J \circ \tilde{\psi}_0.$$

This is an affine map from \mathbf{R}^2 to \mathbf{C} bringing A, B, C respectively to $\tilde{A}, \tilde{B}, \tilde{C}$; in particular, it belongs to the set \mathcal{A} defined in the introduction. Since ψ_0 is affine, du_0, dv_0 are constant one-forms, while $v_0|_G$ is bounded away from 0 and $+\infty$; now (5) implies that $\tilde{\mathcal{E}}(\psi_0) < +\infty$.

A bounded minimising sequence. By the last paragraph, there is a minimising sequence $\{\phi_n\}$ for (6); since we want $\phi_n(G)$ to lie in a hyperbolically bounded set of H_+ , independent of n , we need the next definitions.

Definitions. Let $D, E \in H_+$; we denote by

$$\eta_{D,E}: [0, 1] \rightarrow H_+$$

the unique hyperbolic geodesic with $\eta_{D,E}(0) = D$ and $\eta_{D,E}(1) = E$. We recall that $\eta_{D,E}$ is an arc of a circle with diameter on the real axis, or a segment on a vertical line.

We denote by $T(\tilde{A}, \tilde{B}, \tilde{C})$ the closed hyperbolic triangle with vertices \tilde{A} , \tilde{B} and \tilde{C} ; in other words, $T(\tilde{A}, \tilde{B}, \tilde{C})$ is the closure of the bounded connected component of

$$H_+ \setminus (\eta_{\tilde{A}, \tilde{B}}([0, 1]) \cup \eta_{\tilde{B}, \tilde{C}}([0, 1]) \cup \eta_{\tilde{C}, \tilde{A}}([0, 1])). \quad (4.2)$$

We have the following.

Lemma 4.2. *There is a minimising sequence $\{\tilde{\phi}_n\} \subset \mathcal{A}$ such that*

$$\tilde{\phi}_n(G) \subset T(\tilde{A}, \tilde{B}, \tilde{C}) \quad \text{for all } n \geq 1. \quad (4.3)$$

Proof. Since $\alpha < +\infty$, there is a minimising sequence $\{\phi_n\} \subset \mathcal{A}$; we are going to use $\{\phi_n\}$ to build the minimising sequence $\{\tilde{\phi}_n\}$ of the thesis. The idea is the following: up to an automorphism of H_+ , we can suppose that $\tilde{A} = i$, $\tilde{B} = ti$ with $t \geq 1$ and \tilde{C} is in the right hand half plane. We are going to compose ϕ_n with a map that is the identity on the right hand half plane and projects the left hand half plane onto the imaginary axis; this will put $\phi_n(G)$ in the right hand half plane, i.e on the right side of $\eta_{\tilde{A}, \tilde{B}}([0, 1])$. The composition continues to be minimising; indeed, we shall see that its energy is smaller than that of ϕ_n . Repeating this construction for all sides of the triangle, we get $\tilde{\phi}_n$.

Now to the rigorous proof.

First of all, we consider a cyclical permutations of $(\tilde{A}, \tilde{B}, \tilde{C})$

$$(\tilde{L}, \tilde{M}, \tilde{N}) \in \{(\tilde{A}, \tilde{B}, \tilde{C}), (\tilde{B}, \tilde{C}, \tilde{A}), (\tilde{C}, \tilde{A}, \tilde{B})\}$$

and we make the following assertion.

Let $(\tilde{L}, \tilde{M}, \tilde{N})$ be as in the formula above; let $\{\phi_n\}$ be minimising and let $\zeta_{\tilde{L}, \tilde{M}}$ be the automorphism of H_+ such that $\zeta_{\tilde{L}, \tilde{M}}(\tilde{L}) = i$ and $\zeta_{\tilde{L}, \tilde{M}}(\tilde{M}) = it$ with $t \geq 1$. Since $\zeta_{\tilde{L}, \tilde{M}}$ preserves orientation, our labelling of the vertices implies that $\zeta_{\tilde{L}, \tilde{M}}(\tilde{N})$ is in the closed right hand half plane. We assert that we can build another minimising sequence $\{\tilde{\phi}_n\} \subset \mathcal{A}$ such that $\zeta_{\tilde{L}, \tilde{M}} \circ \tilde{\phi}_n(G)$ is contained in the closed right hand plane.

Before proving the assertion, we show that it implies the thesis. Indeed, applying it three times to the three cyclical permutations of $(\tilde{A}, \tilde{B}, \tilde{C})$, we get a new minimising sequence $\{\tilde{\phi}_n\}$ such that the three sets

$$\zeta_{\tilde{A}, \tilde{B}} \circ \tilde{\phi}_n(G), \quad \zeta_{\tilde{B}, \tilde{C}} \circ \tilde{\phi}_n(G), \quad \zeta_{\tilde{C}, \tilde{A}} \circ \tilde{\phi}_n(G)$$

are contained in the right hand half plane; by the definitions of the maps $\zeta_{\tilde{L}, \tilde{M}}$ this is equivalent to (4.3), ending the proof.

Now we prove the assertion; to fix ideas we suppose that $(\tilde{L}, \tilde{M}, \tilde{N}) = (\tilde{A}, \tilde{B}, \tilde{C})$. By lemma 4.1, for all $n \geq 1$ we have that

$$\tilde{\mathcal{E}}(\zeta_{\tilde{A}, \tilde{B}} \circ \tilde{\phi}_n) = \tilde{\mathcal{E}}(\tilde{\phi}_n). \quad (4.4)$$

We consider the function "positive part"

$$m(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } x > 0 \end{cases}$$

and we define

$$\begin{aligned} g: H_+ &\rightarrow H_+ \\ g(x + iy) &= m(x) + iy \end{aligned} \quad (4.5)$$

which is the identity on $\text{Re}(z) > 0$ and projects $\text{Re}(z) \leq 0$ orthogonally onto the imaginary axis. We give a name to the real and imaginary parts of $g \circ \zeta_{\tilde{A}, \tilde{B}} \circ \phi_n$ and $\zeta_{\tilde{A}, \tilde{B}} \circ \phi_n$, and we define $\psi_n, \tilde{\phi}_n$:

$$\psi_n := g \circ \zeta_{\tilde{A}, \tilde{B}} \circ \phi_n = \tilde{a}_n + i b_n, \quad \zeta_{\tilde{A}, \tilde{B}} \circ \phi_n = a_n + i b_n, \quad \tilde{\phi}_n := \zeta_{\tilde{A}, \tilde{B}}^{-1} \circ \psi_n. \quad (4.6)$$

Note that the imaginary part of ψ_n coincides with that of $\zeta_{\tilde{A}, \tilde{B}} \circ \phi_n$ by the definition of g in (4.5).

The first formula of (4.6) and (4.5) imply that $\psi_n(G)$ is contained in the closed right half plane; by the last formula of (4.6) this implies part of the assertion, namely that $\zeta_{\tilde{A}, \tilde{B}} \circ \tilde{\phi}_n(G)$ is contained in the right half plane.

Thus, the proof of the lemma is complete if we show first that $\{\tilde{\phi}_n\} \subset \mathcal{A}$, second that $\{\tilde{\phi}_n\}$ is minimising.

We prove that $\{\tilde{\phi}_n\} \subset \mathcal{A}$. First of all, by lemma 2.3 and (4.6) we have that both components of $\tilde{\phi}_n$ are in $\mathcal{D}(d)$; it remains to show that

$$\tilde{\phi}_n(A) = \tilde{A}, \quad \tilde{\phi}_n(B) = \tilde{B}, \quad \tilde{\phi}_n(C) = \tilde{C}. \quad (4.7)$$

We saw at the beginning of the proof that

$$\zeta_{\tilde{A}, \tilde{B}}(\tilde{A}), \zeta_{\tilde{A}, \tilde{B}}(\tilde{B}), \zeta_{\tilde{A}, \tilde{B}}(\tilde{C}) \subset \{z \in H_+ : \text{Re}(z) \geq 0\}.$$

Since by (4.5) g is the identity on the right half plane, the first equality of (4.6) implies that ψ_n brings A, B, C to $\zeta_{\tilde{A}, \tilde{B}}(\tilde{A}), \zeta_{\tilde{A}, \tilde{B}}(\tilde{B}), \zeta_{\tilde{A}, \tilde{B}}(\tilde{C})$ respectively; by the last one of (4.6), $\tilde{\phi}_n$ brings A, B, C to $\tilde{A}, \tilde{B}, \tilde{C}$ respectively, proving (4.7).

Lastly, we show that $\{\tilde{\phi}_n\}$ is minimising. The first and last equalities below come from (4.6) and the definition of $\tilde{\mathcal{E}}$ in (5), the second equality follows from point 3) of lemma 2.3; recall that, by (4.5), $\tilde{a}_n = m \circ a_n$. The inequality follows since $m'(x)$ is either zero or one.

$$\begin{aligned} \tilde{\mathcal{E}}(\psi_n) &= \int_G \frac{1}{b_n^2} [(E_x d\tilde{a}_n, d\tilde{a}_n) + (E_x db_n, db_n)] d\kappa = \\ &\int_G \frac{1}{b_n^2} [(m' \circ a_n(x))^2 \cdot (E_x da_n, da_n) + (E_x db_n, db_n)] d\kappa \leq \\ &\int_G \frac{1}{b_n^2} [(E_x da_n, da_n) + (E_x db_n, db_n)] d\kappa = \tilde{\mathcal{E}}(\zeta_{\tilde{A}, \tilde{B}} \circ \phi_n). \end{aligned}$$

The first equality below is the definition of $\tilde{\phi}_n$ in (4.6), the second one follows from lemma 4.1 since $\zeta_{\tilde{A}, \tilde{B}}$ is an automorphism.

$$\tilde{\mathcal{E}}(\tilde{\phi}_n) = \tilde{\mathcal{E}}(\zeta_{\tilde{A}, \tilde{B}}^{-1} \circ \psi_n) = \tilde{\mathcal{E}}(\psi_n).$$

The last two formulas together with (4.4) imply that

$$\tilde{\mathcal{E}}(\tilde{\phi}_n) \leq \tilde{\mathcal{E}}(\phi_n).$$

Since $\{\phi_n\}$ is minimising, $\{\tilde{\phi}_n\}$ is minimising too.

\\

A bound on the minimal. The technique of the last lemma yields the following bound, which will be useful in the next section.

Lemma 4.3. *Let $\phi \in \mathcal{A}$ be minimal; then, with the notation of (4.2),*

$$\phi(G) \subset T(\tilde{A}, \tilde{B}, \tilde{C}). \tag{4.8}$$

Proof. As we saw in the last lemma it suffices to show that

- i) if ζ is the automorphism of H_+ bringing \tilde{A} to i and \tilde{B} to ti with $t \geq 1$, or
 - ii) if ζ is the automorphism of H_+ bringing \tilde{B} to i and \tilde{C} to ti with $t \geq 1$, or
 - iii) if ζ is the automorphism of H_+ bringing \tilde{C} to i and \tilde{A} to ti with $t \geq 1$,
- then $\zeta \circ \phi(G)$ is contained in the right half plane.

We prove case i), since the other ones are analogous.

We begin to note that, since the functions in \mathcal{A} are continuous,

$$O := \{x \in G : \operatorname{Re} \zeta \circ \phi(x) < 0\} \tag{4.9}$$

is an open set of G . We suppose by contradiction that this set is non-empty.

We set $\zeta \circ \phi = u + iv$ and define $\tilde{\phi}$ as in (4.6) by $\zeta \circ \tilde{\phi} = g \circ (u + iv)$ where g is as in (4.5); formula (5) implies that

$$\begin{aligned} \tilde{\mathcal{E}}(\zeta \circ \phi) &= \int_{G \setminus O} \frac{1}{v^2} [(duE_x, du) + (dvE_x, dv)] d\kappa + \\ &\int_O \frac{1}{v^2} [(duE_x, du) + (dvE_x, dv)] d\kappa. \end{aligned}$$

By (4.5), $\zeta \circ \tilde{\phi} = \zeta \circ \phi$ on $G \setminus O$; since the first component of $\zeta \circ \tilde{\phi}$ is zero on O , we get the equality below.

$$\tilde{\mathcal{E}}(\zeta \circ \tilde{\phi}) = \int_{G \setminus O} \frac{1}{v^2} [(duE_x, du) + (dvE_x, dv)] d\kappa + \int_O \frac{1}{v^2} (dvE_x, dv) d\kappa.$$

We assert that

$$\int_O \frac{1}{v^2} (E_x du, du) d\kappa > 0.$$

Indeed, if the integral in the last formula were zero, then the Poincaré inequality (2.8) would imply that u is constant on every connected component of O . If we show that this constant is zero, we have a contradiction with (4.9). Let \mathcal{U} be a connected component of \mathcal{O} and let $u(\mathcal{U}) = \{x_0\}$ with $\operatorname{Re}(x_0) < 0$; since \mathcal{U} is the pre-image of a point it is also closed, and since G is connected we get that $\mathcal{U} = G$, contradicting the fact that $\operatorname{Re}(\tilde{C}) > 0$.

We see as in the last lemma that $\tilde{\phi} \in \mathcal{A}$; lemma 4.1 yields the first and last equalities below, the strict inequality comes from the three formulas above.

$$\tilde{\mathcal{E}}(\tilde{\phi}) = \tilde{\mathcal{E}}(\zeta \circ \tilde{\phi}) < \tilde{\mathcal{E}}(\zeta \circ \phi) = \tilde{\mathcal{E}}(\phi).$$

This contradicts the minimality of ϕ and we are done. \(\backslash\backslash\)

Convergence. Now we can prove point 1) of theorem 1.

Proposition 4.4. *There is a minimising sequence $\{\phi_n\} \subset \mathcal{A}$ which, up to subsequences, converges uniformly to a map $\phi \in \mathcal{A}$ with $\tilde{\mathcal{E}}(\phi) = \alpha$. By the definition in the introduction, the map ϕ is harmonic.*

Proof. We saw above that there is a minimising sequence $\phi_n = u_n + iv_n$; by lemma 4.2, we can suppose that $\{\phi_n\}$ satisfies (4.3). It is immediate that there is $C \geq 1$ such that $z \in T(\tilde{A}, \tilde{B}, \tilde{C})$ implies

$$\frac{1}{C} \leq \operatorname{Im} z \leq C.$$

Together with (4.3) this implies that, for all $n \geq 1$,

$$\frac{1}{C} \leq v_n(x) \leq C \quad \text{for } \kappa\text{-a.e. } x \in G. \quad (4.10)$$

By the definition of $\tilde{\mathcal{E}}$ in (5) this implies that

$$\mathcal{E}(du_n, du_n) + \mathcal{E}(dv_n, dv_n) \leq C^2 \tilde{\mathcal{E}}(\phi_n).$$

Since $\{\phi_n\}$ is minimising, $\tilde{\mathcal{E}}(\phi_n)$ tends to the finite number α ; using the formula above we get that there is $C_1 > 0$ such that, for all $n \geq 1$,

$$\mathcal{E}(du_n, du_n) + \mathcal{E}(dv_n, dv_n) \leq C_1. \quad (4.11)$$

Since $u_n \in \mathcal{A}$, we can apply the Poincaré inequality (2.10) and get that there is $C_2 > 0$ such that, for all $n \geq 1$,

$$\|u_n\|_{L^2(G, \mathbf{R})} + \|v_n\|_{L^2(G, \mathbf{R})} \leq C_2.$$

By (3.7) the last two formulas imply that there is $C_3 > 0$ such that for all $n \geq 1$ we have

$$\|u_n\|_{0, \alpha} + \|v_n\|_{0, \alpha} \leq C_3. \quad (4.12)$$

By (4.12) and the Arzelà-Ascoli theorem we can extract a subsequence (which with we denote with the same index)

$$\{\phi_n = u_n + iv_n\} \subset \mathcal{A}$$

such that, endowing $C(G, H_+)$ with the sup norm,

$$u_n \rightarrow u \quad \text{and} \quad v_n \rightarrow v \quad \text{in} \quad C(G, H_+). \quad (4.13)$$

By (4.11) and weak compactness we can further refine to have

$$du_n \rightharpoonup U \quad \text{and} \quad dv_n \rightharpoonup V \quad \text{in} \quad L^2(G, \Lambda^1(G), \kappa).$$

By lemma 2.1, the last two formulas imply that that $u, v \in \mathcal{D}(d)$ and that $U = du, V = dv$; thus, the last formula becomes

$$du_n \rightharpoonup du \quad \text{and} \quad dv_n \rightharpoonup dv \quad \text{in} \quad L^2(G, \Lambda^1(\mathbf{R}^2), \kappa). \quad (4.14)$$

We just saw that that $u, v \in \mathcal{D}(\mathcal{E}^d) = \mathcal{D}(d)$; the function $\phi = u + iv$ is our candidate minimum. Since $\phi_n \in \mathcal{A}$, (4.13) implies that $\phi(A) = \tilde{A}$, $\phi(B) = \tilde{B}$, $\phi(C) = \tilde{C}$, yielding that $\phi \in \mathcal{A}$. The proof is complete if we show that ϕ is minimal.

We note that, by (4.10), (4.11) and (4.13)

$$\lim_{n \rightarrow +\infty} \int_G \left| \frac{1}{v_n^2} - \frac{1}{v^2} \right| [(du_n E_x, du_n) + (dv_n E_x, dv_n)] d\kappa = 0. \quad (4.15)$$

Let now $v \in \mathcal{A}$ be fixed and let it satisfy $v \geq \frac{1}{C}$ as in (4.10); it is immediate that the convex functional

$$\begin{aligned} & : L^2(G, \Lambda^1(G), \kappa) \times L^2(G, \Lambda^1(G), \kappa) \rightarrow \mathbf{R} \\ & : (\psi, \eta) \rightarrow \int_G \frac{1}{v^2} [(\psi E_x, \psi) + (\eta E_x, \eta)] d\kappa \end{aligned}$$

is continuous for the strong topology of $L^2(G, \Lambda^1(G), \kappa) \times L^2(G, \Lambda^1(G), \kappa)$, and thus it is lower semicontinuous for the weak one.

The first equality below follows by the definition of $\tilde{\mathcal{E}}$, the second one by adding and subtracting; the inequality follows from (4.14), (4.15) and weak lower-semicontinuity; the equality at the end is again the definition of $\tilde{\mathcal{E}}$.

$$\begin{aligned} \lim_{n \rightarrow +\infty} \tilde{\mathcal{E}}(\phi_n) &= \lim_{n \rightarrow +\infty} \int_G \frac{1}{v_n^2} [(du_n E_x, du_n) + (dv_n E_x, dv_n)] d\kappa = \\ & \lim_{n \rightarrow +\infty} \int_G \frac{1}{v^2} [(du_n E_x, du_n) + (dv_n E_x, dv_n)] d\kappa + \\ & \lim_{n \rightarrow +\infty} \int_G \left(\frac{1}{v_n^2} - \frac{1}{v^2} \right) [(du_n E_x, du_n) + (dv_n E_x, dv_n)] d\kappa \geq \\ & \int_G \frac{1}{v^2} [(du E_x, du) + (dv E_x, dv)] d\kappa = \tilde{\mathcal{E}}(\phi). \end{aligned}$$

We saw above that $\phi \in \mathcal{A}$; sine $\{\phi_n\}$ is minimising, the last formula implies that $\tilde{\mathcal{E}}(\phi) \leq \alpha$, ending the proof.

\\

§5

Uniqueness

In this section we prove point 2) of theorem 1; we start with the following lemma.

Lemma 5.1. *There are $r_0 > 0$ and $C > 0$ such that the following holds. Let $r \in (0, r_0)$ and let*

$$\tilde{A}, \tilde{B}, \tilde{C} \in B(i, r), \tag{5.1}$$

where the ball is the Euclidean one. Let $\phi \in \mathcal{A}$ be minimal; then,

$$\phi(G) \subset B(i, r). \tag{5.2}$$

Proof. Since ϕ is minimal in \mathcal{A} , formula (4.8) holds. Recall that the ball $B(i, r)$ is also a hyperbolic ball, albeit with a different center and radius; since hyperbolic balls are geodesically convex, the geodesics connecting \tilde{A} with \tilde{B} , \tilde{B} with \tilde{C} and \tilde{C} with \tilde{A} are contained in $B(i, r)$. As a consequence, the hyperbolic triangle $T(\tilde{A}, \tilde{B}, \tilde{C})$ satisfies

$$T(\tilde{A}, \tilde{B}, \tilde{C}) \subset B(i, r).$$

Since $\phi(G) \subset T(\tilde{A}, \tilde{B}, \tilde{C})$ by (4.8), we get (5.2).

\\

Let us call $\tilde{r}(r)$ the hyperbolic radius of $B(i, r)$; if B is a ball of hyperbolic radius $\tilde{r}(r)$, there is an automorphism of H_+ bringing it into $B(i, r)$. Thus, the next proposition implies point 2) of theorem 1.

Proposition 5.2. *There is $r_0 > 0$ such that, if $r \in (0, r_0)$ and (5.1) holds, then $\tilde{\mathcal{E}}$ has a unique minimiser $\phi_{\tilde{A}, \tilde{B}, \tilde{C}}$. The function*

$$\Phi: B(i, r)^3 \rightarrow \mathcal{D}(\mathcal{E}^d) \oplus i\mathcal{D}(\mathcal{E}^d), \quad \Phi: (\tilde{A}, \tilde{B}, \tilde{C}) \rightarrow \phi_{\tilde{A}, \tilde{B}, \tilde{C}}$$

is of class C^1 .

Proof. By lemma 5.1 and proposition 4.4 it suffices to prove that there is a unique critical point ϕ of $\tilde{\mathcal{E}}$ which satisfies (5.2) and depends differentiably on $\tilde{A}, \tilde{B}, \tilde{C}$. In order to find such a critical point, in step 1 below we perform a blow-up and recast the Euler-Lagrange equation of $\tilde{\mathcal{E}}$ in the form (5.8) below; in step 2, we see that (5.8) means that $\phi_{\tilde{A}, \tilde{B}, \tilde{C}}$ is the zero of the nonlinear operator of (5.12), to which we apply the implicit function theorem in step 3.

Step 1: blow up. We write the function $\phi \in \mathcal{A}$ in the form $\phi = u + i(v + 1)$. By (5), the energy of ϕ is

$$\tilde{\mathcal{E}}(\phi) = \int_G \frac{1}{(1+v)^2} [(duE_x, du) + (dvE_x, dv)] d\kappa.$$

We set

$$a_r(u, v) = \frac{-2r}{(1+rv)^3} [(duE_x, du) + (dvE_x, dv)].$$

Classical results on the differentiation of functionals ([3]) show that the critical points of $\tilde{\mathcal{E}}$ satisfy

$$\int_G \left\{ \frac{1}{(1+v)^2} [(duE_x, d\phi_1) + (dvE_x, d\phi_2)] + a_1(u, v)\phi_2 \right\} d\kappa = 0$$

for all $\phi_1, \phi_2 \in \mathcal{D}(d)$ with $\phi_i(A) = \phi_i(B) = \phi_i(C) = 0$. By the definition of the codifferential in the relative sense this implies that

$$\begin{cases} \bar{d}_G \frac{du}{(1+v)^2} = 0 \\ \bar{d}_G \frac{dv}{(1+v)^2} + a_1(u, v) = 0 \\ \phi = [u + i(v + 1)] \in \mathcal{A}. \end{cases}$$

In order to make a blow-up around i we set

$$\phi_r = \frac{\phi - i}{r} + i = u_r + i(1 + rv_r), \quad \text{or equivalently} \quad \phi = u + i(v + 1) = ru_r + (1 + rv_r)i. \quad (5.3)$$

Formula (5.3) implies the first equality below, while the second one comes from the definition of $\tilde{\mathcal{E}}$ in (5).

$$\begin{aligned} \tilde{\mathcal{E}}(\phi) &= \tilde{\mathcal{E}}(ru_r + i(1 + rv_r)) = \\ &= r^2 \int_G \frac{1}{(1 + rv_r)^2} [(du_r E_x, du_r) + (dv_r E_x, dv_r)] d\kappa. \end{aligned}$$

Using again the definition of the codifferential in the relative sense, the Euler-Lagrange equation of this functional is given by the first two equations below; for the last three equations, which are the boundary conditions, we use (5.3) and the fact that $\phi(A) = \tilde{A}$, $\phi(B) = \tilde{B}$, $\phi(C) = \tilde{C}$.

$$\begin{cases} \bar{d}_G \left(\frac{du_r}{(1 + rv_r)^2} \right) = 0 \\ \bar{d}_G \left(\frac{dv_r}{(1 + rv_r)^2} \right) + a_r(u_r, v_r) = 0 \\ (u_r + iv_r)(A) = \frac{1}{r}(\tilde{A} - i) \\ (u_r + iv_r)(B) = \frac{1}{r}(\tilde{B} - i) \\ (u_r + iv_r)(C) = \frac{1}{r}(\tilde{C} - i). \end{cases}$$

Adding and subtracting, we re-write it in the following way; the last condition is implied by (5.2) and the definition of u_r, v_r in (5.3); the ball is the Euclidean one with center the origin and radius 1.

$$\left\{ \begin{array}{l} -\Delta u_r + \bar{d}_G \left[du_r \left(\frac{1}{(1+rv_r)^2} - 1 \right) \right] = 0 \\ -\Delta v_r + \bar{d}_G \left[dv_r \left(\frac{1}{(1+rv_r)^2} - 1 \right) \right] + a_r(u_r, v_r) = 0 \\ (u_r + iv_r)(A) = \frac{1}{r}(\tilde{A} - i) \\ (u_r + iv_r)(B) = \frac{1}{r}(\tilde{B} - i) \\ (u_r + iv_r)(C) = \frac{1}{r}(\tilde{C} - i) \\ (u_r + iv_r)(G) \subset B(0, 1). \end{array} \right. \quad (5.4)$$

Now we consider the affine map ψ_0 we defined after (4.1); we make a blow up as in (5.3), setting

$$\frac{1}{r}[\psi_0(z) - i] = \tilde{u}_r(z) + i\tilde{v}_r(z). \quad (5.5)$$

Since ψ_0 is affine, we have that \tilde{u}_r and \tilde{v}_r are affine too and thus by (5.1) $\tilde{u}_r + i\tilde{v}_r$ has image inside the Euclidean ball $B(0, 1)$, which implies the first inequality below; since ψ_0 is affine, the first inequality implies the second one for some $C_1 > 0$ independent of $\tilde{A}, \tilde{B}, \tilde{C} \in B(i, r)$.

$$\sup_{z \in G} |\tilde{u}_r(z) + i\tilde{v}_r(z)| \leq 1 \quad \text{and} \quad \sup_{z \in G} |d\tilde{u}_r(z) + id\tilde{v}_r(z)| \leq C_1. \quad (5.6)$$

In the first equality below, we write $u_r + iv_r$ as a perturbation of the affine map $\frac{1}{r}(\psi_0 - i)$; the second equality comes from (5.5).

$$(u_r + iv_r)(z) = \frac{1}{r}[\psi_0(z) - i] + U(z) + iV(z) = \tilde{u}_r(z) + i\tilde{v}_r(z) + U(z) + iV(z).$$

In this way we have that

$$u_r = \tilde{u}_r + U, \quad v_r = \tilde{v}_r + V,$$

where \tilde{u}_r, \tilde{v}_r are affine and depend on the points $\tilde{A}, \tilde{B}, \tilde{C}$.

Since both ϕ and ψ_0 bring A, B, C into $\tilde{A}, \tilde{B}, \tilde{C}$ respectively, (5.3) and (5.5) imply that $(u_r + iv_r)(j) = (\tilde{u}_r + i\tilde{v}_r)(j)$ for $j \in \{A, B, C\}$; by the last formula this implies that $U + iV$ vanishes on A, B and C , or that $U, V \in H_0$, where H_0 is the space we defined before (2.11). At the beginning of section 2 we saw that the differential of a constant and the Laplacian of an affine function are both zero; thus, the first two equations of (5.4) become the first two equations below, while the third one follows from (5.6) and the last one of (5.4).

$$\left\{ \begin{array}{l} -\Delta U + \bar{d}_G \left[d(\tilde{u}_r + U) \left(\frac{1}{(1+r(\tilde{v}_r + V))^2} - 1 \right) \right] = 0 \\ -\Delta V + \bar{d}_G \left[d(\tilde{v}_r + V) \left(\frac{1}{(1+r(\tilde{v}_r + V))^2} - 1 \right) \right] + a_r(\tilde{u}_r + U, \tilde{v}_r + V) = 0 \\ (U + iV)(G) \subset B(0, 2), \quad U, V \in H_0. \end{array} \right. \quad (5.7)$$

By lemma 2.4 we can apply $(-\Delta)^{-1}$ to the first two equations of (5.7), getting

$$\begin{cases} U + (-\Delta)^{-1} \bar{d}_G \left[d(\tilde{u}_r + U) \left(\frac{1}{(1+r(\tilde{v}_r + V))^2} - 1 \right) \right] = 0 \\ V + (-\Delta)^{-1} \left\{ \bar{d}_G \left[d(\tilde{v}_r + V) \left(\frac{1}{(1+r(\tilde{v}_r + V))^2} - 1 \right) \right] + a_r(\tilde{u}_r + U, \tilde{v}_r + V) \right\} = 0 \end{cases} \quad (5.8)$$

$(U + iV)(G) \subset B(0, 2), \quad U, V \in H_0.$

Note that $U + iV \equiv 0$ is a solution of (5.8) when $r = 0$; in the next two steps we shall see that this solution survives when $r > 0$ is small.

Step 2. The nonlinear operator. The solutions of (5.8) are the zeroes of a nonlinear operator; in this step we write it down explicitly and verify the hypotheses of the implicit function theorem.

We group the nonlinear terms of (5.8) in the function F defined below; it depends on the parameters $\tilde{A}, \tilde{B}, \tilde{C}$ through the affine function $\tilde{u}_r + i\tilde{v}_r$; as for the denominators, we can choose ϵ so small that $\epsilon|\tilde{v}_r + V| \leq \frac{1}{2}$ κ -a.e. for all $U + iV$ satisfying the last one of (5.8) and we take $r \in (-\epsilon, \epsilon)$.

$$\begin{aligned} F: (-\epsilon, \epsilon) \times B(0, 1)^3 \times (H_0 \oplus iH_0) &\rightarrow H_0 \oplus iH_0 \\ F(r, \tilde{A}, \tilde{B}, \tilde{C}, U + iV) &= (-\Delta)^{-1} \bar{d}_G \left[d(\tilde{u}_r + U) \left(\frac{1}{(1+r(\tilde{v}_r + V))^2} - 1 \right) \right] + \\ &i(-\Delta)^{-1} \left\{ \bar{d}_G \left[d(\tilde{v}_r + V) \left(\frac{1}{(1+r(\tilde{v}_r + V))^2} - 1 \right) \right] + a_r(\tilde{u}_r + U, \tilde{v}_r + V) \right\}. \end{aligned} \quad (5.9)$$

We prove that this operator lands in $H_0 \oplus iH_0$. Since $U \in \mathcal{D}(\mathcal{E}^d) = \mathcal{D}(d)$, we have that $dU \in L^2$, while $d\tilde{u}_r \in L^2$ since \tilde{u} is affine and we saw in section 2 that the differential of a C^1 function is the projection of the standard differential on the cotangent space.

On the other side, (5.6) and the last one of (5.8) imply that, for all $r \in (-\epsilon, \epsilon)$,

$$\left\| \left(\frac{1}{(1+r(\tilde{v}_r + V))^2} - 1 \right) \right\|_{L^\infty} \leq C_3 r.$$

By the second inequality of (5.6) this implies that

$$\begin{aligned} \left\| \left(\frac{1}{(1+r(\tilde{v}_r + V))^2} - 1 \right) d(\tilde{u}_r + U) \right\|_{L^2} &\leq C_4 r (1 + \|dU\|_{L^2}), \\ \left\| \left(\frac{1}{(1+r(\tilde{v}_r + V))^2} - 1 \right) d(\tilde{v}_r + V) \right\|_{L^2} &\leq C_4 r (1 + \|dV\|_{L^2}). \end{aligned}$$

By the definition of a_r at the beginning of step 1 we get similarly

$$\|a_r(\tilde{u}_r + U, \tilde{v}_r + V)\|_{L^1} \leq C_4 r (1 + \|dU\|_{L^2}^2 + \|dV\|_{L^2}^2)^{\frac{1}{2}}.$$

Now lemma 2.4 implies that, for some absolute constant $C_5 > 0$,

$$\begin{cases} \left\| (-\Delta)^{-1} \bar{d}_G \left[d(\tilde{u}_r + U) \left(\frac{1}{(\tilde{v}_r + rV)^2} - 1 \right) \right] \right\|_{H_0} \leq C_5 r (1 + \|dU\|_{L^2}) \\ \left\| (-\Delta)^{-1} \bar{d}_G \left[d(\tilde{v}_r + V) \left(\frac{1}{(\tilde{v}_r + rV)^2} - 1 \right) \right] \right\|_{H_0} \leq C_5 r (1 + \|dV\|_{L^2}) \\ \|(-\Delta)^{-1} a_r(\tilde{u}_r + U, \tilde{v}_r + V)\|_{H_0} \leq C_5 r (1 + \|dU\|_{L^2}^2 + \|dV\|_{L^2}^2)^{\frac{1}{2}}. \end{cases} \quad (5.10)$$

This implies that F lands in a ball of $H_0 \oplus iH_0$ which is small with r ; actually, if U, V are in a fixed ball of H_0 , then (forgetting to write some arguments of F) $F(U + iV)$ is in a ball of $H_0 \oplus iH_0$ with radius smaller than $C_6 r$.

We assert that F is differentiable in the variables (U, V) and that the differential depends continuously on all variables. Let us consider for instance the real part of F ; forgetting to write some arguments of F on the left hand side, we have

$$\operatorname{Re}F(U + iV) = (-\Delta)^{-1} \bar{d}_G \left[d(\tilde{u}_r + U) \left(\frac{1}{(1 + r(\tilde{v}_r + V))^2} - 1 \right) \right].$$

This operator is affine in U and thus we have that

$$\partial_U \operatorname{Re}F(U + iV)(h) = (-\Delta)^{-1} \bar{d}_G \left[dh \left(\frac{1}{(1 + r(\tilde{v}_r + V))^2} - 1 \right) \right]. \quad (5.11)$$

This function does not depend on U ; we show that it is continuous in $r, \tilde{A}, \tilde{B}, \tilde{C}$ and V . Indeed, the map

$$R: (r, \tilde{A}, \tilde{B}, \tilde{C}, V) \rightarrow \frac{1}{(1 + r(\tilde{v}_r + V))^2} - 1$$

is continuous from $(-\epsilon, \epsilon) \times B(i, r)^3 \times H_0$ to L^∞ by (3.7) and the definition of \tilde{v}_r . Now we consider the map

$$S: L^\infty \rightarrow \mathcal{L}(H_0, L^2), \quad S(W)(h) = Wdh.$$

It is easy to see that S is linear and continuous; by lemma 2.4 this implies that the operator

$$T: L^\infty \rightarrow \mathcal{L}(H_0, H_0)$$

$$T(W)(h) = (-\Delta)^{-1} \bar{d}_G(Wdh) = (-\Delta)^{-1} \bar{d}_G(S(W)h)$$

is linear and continuous. Since by (5.11)

$$\partial_U \operatorname{Re}F(U + iV)(h) = T \circ R(V)(h),$$

we get that the partial derivative above is continuous from $(-\epsilon, \epsilon) \times B(i, r)^3 \times H_0$ to $\mathcal{L}(H_0, H_0)$.

For the derivative with respect to V , standard results on Nemitsky operators (see for instance [3]) show that

$$\partial_V \operatorname{Re}F(U + iV)(h) = (-\Delta)^{-1} \bar{d}_G \left[d(\tilde{u}_r + U) \frac{-2rh}{(1 + r(\tilde{v}_r + V))^3} \right].$$

We see as above that this map is continuous from $(-\epsilon, \epsilon) \times B(i, r)^3 \times H_0 \oplus iH_0$ to $\mathcal{L}(H_0, H_0)$.

The same argument implies that $\operatorname{Re}F$ is C^1 in the variables $r, \tilde{A}, \tilde{B}, \tilde{C}$; we forego the analogous argument for the imaginary part.

Moreover, (5.9) implies that $D_{U,V}F(0, \tilde{A}, \tilde{B}, \tilde{C}, U + iV)$ is the zero operator. Now we define

$$G: (-\epsilon, \epsilon) \times B(0, 1)^3 \times (H_0 \oplus iH_0) \rightarrow H_0 \oplus iH_0$$

$$G(r, \tilde{A}, \tilde{B}, \tilde{C}, U + iV) := (U + iV) + F(r, \tilde{A}, \tilde{B}, \tilde{C}, U + iV).$$

The map G is C^1 in $(-\epsilon, \epsilon) \times B(0, 1)^3 \times (H_0 \oplus iH_0)$; moreover, since $D_{U,V}F(0, \tilde{A}, \tilde{B}, \tilde{C}, U + iV)$ is the zero operator, the matrix

$$\begin{pmatrix} \partial_U G(0, \cdot) & 0 \\ 0 & \partial_V G(0, \cdot) \end{pmatrix}$$

is the identity from $H_0 \oplus iH_0$ to itself and thus it is invertible.

Equation (5.8) is equivalent to

$$G(r, \tilde{A}, \tilde{B}, \tilde{C}, U + iV) = 0 \tag{5.12}$$

which has the unique solution $U \equiv V \equiv 0$ when $r = 0$.

Step 3. The implicit function theorem. By the last step we can apply the implicit function theorem to (5.12); we get that there is $\epsilon > 0$ with the following property. If $r \in [0, \epsilon)$ and $\tilde{A}, \tilde{B}, \tilde{C} \in B(0, 1)$, then equation (5.12) has a unique solution $U_{r, \tilde{A}, \tilde{B}, \tilde{C}} + iV_{r, \tilde{A}, \tilde{B}, \tilde{C}}$; such a solution depends differentiably on $r, \tilde{A}, \tilde{B}, \tilde{C}$ and is zero when $r = 0$; in particular, the last one of (5.8) holds. In other words, we are finding a branch of solutions which coincide with the affine $\tilde{u}_r + i\tilde{v}_r$ when $r = 0$; the dependence on all parameters is C^1 .

\\

§6

Geodesics

One of the properties of harmonic immersions is that they are totally geodesic, i.e. they bring geodesics into geodesics. Of course, this is a local property: the image of a minimising curve satisfies the geodesic equation (i.e. it is minimising on small time intervals) but it may not be the shortest path between its endpoints.

The problem is that at this point we only know that the harmonic map ϕ is Hölder, which does not even imply that the image of a geodesic of G has finite length. In this section we are going to show that the image of a geodesic on G has finite length and is a geodesic on $\phi(G)$.

In order to show that the harmonic immersions ϕ is totally geodesic, we recall from [14] and [16] some facts about geodesics on the harmonic Sierpinski gasket G of section 1.

First of all, there are six geodesics

$$\gamma_{r,s}: [0, 1] \rightarrow G, \quad r \neq s \in \{A, B, C\} \tag{6.1}$$

such that

$$\gamma_{r,s}(0) = r, \quad \gamma_{r,s}(1) = s \quad \text{for } r \neq s \in \{A, B, C\}.$$

Obviously, if we parametrise these geodesics proportional to arc-length, $\gamma_{r,s}(t) = \gamma_{s,r}(1 - t)$.

These geodesics are C^1 but not C^2 and they are the "outer boundary" of the gasket in the following sense: if U is the unbounded connected component of $\mathbf{R}^2 \setminus G$, then

$$\partial U = \gamma_{A,B}([0, 1]) \cup \gamma_{B,C}([0, 1]) \cup \gamma_{A,C}([0, 1]).$$

On smaller cells we have that the curve connecting, say, $F_{j_0 \dots j_{l-1}}(A)$ with $F_{j_0 \dots j_{l-1}}(C)$, which has minimal length and has image in the cell $F_{j_0 \dots j_{l-1}}(G)$ is $F_{j_0 \dots j_{l-1}} \circ \gamma_{A,C}$.

Now the harmonic gasket is totally geodesic: for any two points in G there is a geodesic connecting them. Moreover, this geodesic is a countable or finite chain of geodesics of the type $F_{j_0 \dots j_{l-1}} \circ \gamma_{r,s}$ with $r \neq s \in \{A, B, C\}$. The chain is finite if and only if the geodesic connects two vertices of the pre-fractal, say $\gamma(0) = F_{j_0 \dots j_{l-1}}(A)$ and $\gamma(1) = F_{i_0 \dots i_{l-1}}(B)$.

This prompts us to state the following proposition; its proof will occupy the rest of this section.

Proposition 6.1. *Let ϕ be the harmonic immersion of proposition 4.4; we suppose that ϕ is injective. Let $j_0, \dots, j_{l-1} \in \{1, 2, 3\}$ and let $r \neq s \in \{A, B, C\}$. Then, $\phi \circ F_{j_0 \dots j_{l-1}} \circ \gamma_{r,s}$ is a curve minimising length among all curves which connect $\phi \circ F_{j_0 \dots j_{l-1}}(r)$ with $\phi \circ F_{j_0 \dots j_{l-1}}(s)$ and have image in the cell $\phi \circ F_{j_0 \dots j_{l-1}}(G)$.*

Proof. For simplicity, we suppose that $l = 0$, i.e. that the map $F_{j_0 \dots j_{l-1}}$ is the identity; the general case follows applying the argument below to the cell $F_{j_0 \dots j_{l-1}}(G)$.

We must prove that $\phi \circ \gamma_{r,s}$ is a geodesic in $\phi(G)$; to fix ideas, we suppose that $r = A$, $s = B$ and thus $\gamma_{r,s} = \gamma_{A,B}$. Since ϕ is injective, \tilde{A} , \tilde{B} and \tilde{C} do not lie on the same free geodesic of H_+ : otherwise, by (4.8) $\phi(G)$ would be contained in that geodesic and then we could find, by the intermediate value theorem, $t, s \in (0, 1)$ such that $\phi \circ \gamma_{A,B}(s) = \phi \circ \gamma_{A,C}(t)$, contradicting injectivity.

Recall also that \tilde{A} , \tilde{B} , \tilde{C} are oriented in the clockwise sense. Thus, applying an automorphism to H_+ , we can also suppose that

$$\tilde{A} = i, \quad \tilde{B} = ai \quad \text{with } a > 1, \quad \text{Re}(\tilde{C}) > 0. \quad (6.2)$$

The idea is the following. We shall define a closed set $\bar{R} \subset H_+$ which is bounded on the left by $\phi \circ \gamma_{A,B}([0, 1])$; a standard connectedness argument shows that $\phi(G) \subset \bar{R}$. Next, among the curves $\bar{\eta}: [0, 1] \rightarrow H_+$ which connect \tilde{A} with \tilde{B} and have image in \bar{R} we find one with minimal energy; this is a classical "geodesic with obstacle" problem; we are going to see that the infimum is finite and that it is attained on a curve η . To end the proof, we show that $\eta([0, 1]) = \phi \circ \gamma([0, 1])$. Since η is minimal among all curves which have image in \bar{R} , a larger class than curves which have image in $\phi(G)$, we get that $\phi \circ \gamma$ is a geodesic in $\phi(G)$. Now to the rigorous proof.

Definition of the set R . Formulas (6.2) and (4.8) imply that $\phi(G)$ does not intersect the following subset of the imaginary axis:

$$U := \{it : t \in (0, 1) \cup (a, +\infty)\}.$$

We consider the closed set of H_+

$$S := \{it : t \in (0, 1] \cup [a, +\infty)\} \cup \phi \circ \gamma_{A,B}([0, 1]). \quad (6.3)$$

By (6.2) we have $\phi \circ \gamma_{A,B}(0) = \tilde{A} = i$, $\phi \circ \gamma_{A,B}(1) = \tilde{B} = ai$; thus, S is the image of a continuous curve connecting 0 with infinity; we assert that this curve has no self-intersections. First we note an obvious fact,

i.e. that U has no self-intersections. Second, $\phi \circ \gamma_{A,B}([0, 1]) \subset \phi(G)$ does not intersect the set U by (4.8). Lastly, we show that $\phi \circ \gamma_{A,B}$ has no self-intersections. We are supposing that ϕ is injective; since $\gamma_{A,B}$ is injective because it is a geodesic of G , we get that $\phi \circ \gamma_{A,B}$ is injective; these facts prove the assertion.

As a consequence, S divides $H_+ \setminus S$ into two connected components, both unbounded; we call R the one which contains $\tilde{C} + s$ for $s > 0$ large; we assert that $\tilde{C} \in R$. Indeed, the half-line $\{\tilde{C} + s : s \geq 0\}$ is connected and, by the last formula of (6.2) and (4.8), $\{\tilde{C} + s : s \geq 0\}$ does not intersect S ; as a consequence, this set lies in one of the two connected components of $H_+ \setminus S$, which must be R .

Now we note that $\phi(G) \subset \bar{R}$. Since ϕ is injective and $\phi(G) \subset T(\tilde{A}, \tilde{B}, \tilde{C})$ by (4.8), we get that $\phi(G \setminus \gamma_{\tilde{A}, \tilde{B}}([0, 1]))$ does not intersect S ; since ϕ is continuous and $G \setminus \gamma_{\tilde{A}, \tilde{B}}([0, 1])$ is connected, its image is in one of the two connected components of $H_+ \setminus S$; since we just saw that $\tilde{C} \in R$, we get that $\phi(G) \subset \bar{R}$.

The geodesic with obstacle. In the next lemma we show that there is a shortest curve in H_+ which connects \tilde{A} with \tilde{B} and has image in \bar{R} ; before stating it, we need a definition.

Definitions. We define \mathcal{R} as the set of all curves η which connect \tilde{A} with \tilde{B} and such that $\eta([0, 1]) \subset \bar{R}$. More precisely, a curve

$$\eta: [0, 1] \rightarrow H_+$$

is in \mathcal{R} if the following three points hold.

- 1) $\eta \in H^1((0, 1), H_+)$; by the properties of the Sobolev space H^1 , η has a representative in $C([0, 1], H_+)$, which justifies the following two points.
- 2) $\eta(0) = \tilde{A}$, $\eta(1) = \tilde{B}$.
- 3) $\eta([0, 1]) \subset \bar{R}$.

If $\eta(t) = \eta_1(t) + i\eta_2(t) \in H^1((0, 1), H_+)$ we define as usual its hyperbolic energy as

$$\int_0^1 \frac{1}{\eta_2(t)^2} |\dot{\eta}(t)|^2 dt.$$

For the obstacle problem, we define

$$L := \inf_{\eta \in \mathcal{R}} \int_0^1 \frac{1}{\eta_2(t)^2} |\dot{\eta}(t)|^2 dt.$$

Lemma 6.2. *The infimum above is a minimum, i.e. there is $\eta \in \mathcal{R}$ such that*

$$L = \int_0^1 \frac{1}{\eta_2(t)^2} |\dot{\eta}(t)|^2 dt. \tag{6.4}$$

Proof. We begin to prove that L is finite. We consider two free geodesics of H_+ : $\eta_{\tilde{A}, \tilde{C}}$ is the one with $\eta_{\tilde{A}, \tilde{C}}(0) = \tilde{A}$, $\eta_{\tilde{A}, \tilde{C}}(1) = \tilde{C}$ while $\eta_{\tilde{C}, \tilde{B}}$ satisfies $\eta_{\tilde{C}, \tilde{B}}(0) = \tilde{C}$, $\eta_{\tilde{C}, \tilde{B}}(1) = \tilde{B}$. They both have finite energy, implying that the absolutely continuous curve

$$\eta: [0, 1] \rightarrow H_+$$

$$\eta(t) = \begin{cases} \eta_{\bar{A}, \bar{C}}(2t) & \text{if } 0 \leq t \leq \frac{1}{2} \\ \eta_{\bar{C}, \bar{B}}(2t-1) & \text{if } \frac{1}{2} \leq t \leq 1 \end{cases}$$

has finite energy. Moreover, formula (4.8) implies that $\eta \in \mathcal{R}$. A consequence of these two facts is that L is finite. In particular, there is a minimising sequence which, by standard results, converges to a curve $\gamma = \gamma_1 + i\gamma_2 \in \mathcal{R}$ which satisfies (6.4).

\\

The geodesic with obstacle coincides with $\phi \circ \gamma_{A,B}([0,1])$. We briefly explain the idea of the proof of lemma 6.3 below. Let η minimise in (6.4) and let us suppose by contradiction that there is a maximal non-empty interval (t_1, t_2) such that $\eta(t_1, t_2) \cap \phi \circ \gamma_{A,B}([0,1]) = \emptyset$. As a consequence, $\eta(t_1, t_2)$ does not touch the obstacle and thus it is a free geodesic; up to applying an automorphism, we can suppose that $\eta(t_1, t_2)$ is a segment on the imaginary axis; since (t_1, t_2) is maximal, we have that $\eta(t_i) = \phi \circ \gamma_{A,B}(s_i)$ for $i \in \{1, 2\}$. Let us suppose, to fix ideas, that $\phi \circ \gamma_{A,B}(s_1, s_2)$ intersects the left hand half plane. We project onto the imaginary axis the part of $\phi(G)$ which stays to the right of $\phi \circ \gamma_{A,B}(s_1, s_2)$ and to the left of the imaginary axis and we show that the energy of the projection is strictly smaller than that of the minimal ϕ , obtaining a contradiction as in lemma 4.3.

Lemma 6.3. *Let $\eta \in \mathcal{R}$ minimise in (6.4); then*

$$\eta([0,1]) = \phi \circ \gamma_{A,B}([0,1]). \quad (6.5)$$

Proof. We suppose by contradiction that (6.5) does not hold; since γ and $\phi \circ \gamma_{A,B}$ are continuous,

$$\{t : \eta(t) \notin \phi \circ \gamma_{A,B}([0,1])\} \quad (6.6)$$

is an open set; we consider one of its connected components $(t_1, t_2) \subset [0,1]$:

$$\eta(t) \notin \phi \circ \gamma_{A,B}([0,1]) \quad \text{if } t \in (t_1, t_2) \quad \text{but} \quad \eta(t_1), \eta(t_2) \in \phi \circ \gamma_{A,B}([0,1]). \quad (6.7)$$

Since η is a geodesic, albeit with obstacle, it is injective and thus $\eta(t_1) \neq \eta(t_2)$. Since also $\phi \circ \gamma_{A,B}$ is injective, there are $s_1 \neq s_2$ such that $\eta(t_i) = \phi \circ \gamma_{A,B}(s_i)$ for $i \in \{1, 2\}$.

The geodesic is free. We assert that $\eta|_{(t_1, t_2)}$ is a free geodesic in H_+ ; since $\eta \in \mathcal{R}$, this is true if $\eta((t_1, t_2))$ does not intersect the set S of (6.3). We know by (6.7) that $\eta((t_1, t_2))$ does not intersect $\psi \circ \gamma_{A,B}([0,1])$; we must prove that it does not intersect $i(0,1] \cup i[a, +\infty)$. Let us suppose by contradiction that there is $t_0 \in (t_1, t_2)$ such that $\eta(t_0) \in i(0,1] \cup i[a, +\infty)$. First of all, we can exclude that $\eta(t_0) = i$ and that $\eta(t_0) = ia$: since these points also belong to $\phi \circ \gamma_{A,B}([0,1])$, in this case (t_1, t_2) would not be a connected component of (6.6). Thus, $\eta(t_0) \in i\mathbf{R}$; let $[\alpha, \beta]$ be the maximal interval which contains t_0 and such that

$$\eta(t) \in i\mathbf{R} \quad \text{for } t \in [\alpha, \beta].$$

We have that $[\alpha, \beta] \subset [t_1, t_2]$ and it is easy to see that the inclusion is proper unless $\eta([t_1, t_2])$ is contained in the imaginary axis, but in this case the assertion follows trivially because the imaginary axis is a geodesic.

We work out the case $\beta < t_2$, since the case $\alpha > t_1$ is analogous. Since $\eta \in \mathcal{R}$ and β is maximal, for $\delta > 0$ small we get that

$$\operatorname{Re}\gamma(t) = 0 \quad \text{for } t \in [\alpha, \beta] \quad \text{and} \quad \operatorname{Re}\eta(\beta + \delta) > 0.$$

Let $\tilde{\eta}$ be the free hyperbolic geodesic connecting $\eta(\alpha - \delta)$ with $\eta(\beta + \delta)$; the last formula easily implies that $\tilde{\eta}$ has image in \bar{R} . Since $\eta|_{[\alpha - \delta, \beta + \delta]}$ is clearly not a geodesic of H_+ , $\tilde{\eta}$ is shorter, contradicting the minimality of η in \mathcal{R} .

The embedding does not minimise. We reach the contradiction that the embedding ψ does not minimise the energy of (5).

Let (t_1, t_2) be as in (6.7); we apply another automorphism of H_+ in order to move $\eta(t_1), \eta(t_2)$ on the imaginary axis; in other words, there is $a > 0$ such that, for $s_1, s_2 \in [0, 1]$,

$$\eta(t_1) = i = \phi \circ \gamma_{A,B}(s_1), \quad \eta(t_2) = ia = \phi \circ \gamma_{A,B}(s_2). \quad (6.8)$$

We forego to give an explicit name to the automorphism but from now on by $\tilde{A}, \tilde{B}, \tilde{C}$ and R we shall mean their image by this map.

We saw above that $\eta|_{[t_1, t_2]}$ is a free geodesic; together with the last formula this implies that $\eta([t_1, t_2])$ is contained in the imaginary axis. We suppose that $\phi \circ \gamma_{A,B}(s_1, s_2)$ intersects the open left hand half plane; the argument for the right hand plane is analogous.

To fix ideas we suppose that $s_1 < s_2$ and we consider the loop $\tilde{\eta}$ obtained joining η with $\phi \circ \gamma_{A,B}$:

$$\tilde{\eta}(t) = \begin{cases} \eta(t) & \text{if } t \in [t_1, t_2] \\ \phi \circ \gamma_{A,B}(t_2 - t + s_2) & \text{if } t \in [t_2, t_2 + s_2 - s_1]. \end{cases}$$

Formula (6.8) implies that this curve is continuous and that it is closed, i.e. $\tilde{\eta}(t_1) = \tilde{\eta}(t_2 + s_2 - s_1)$. Moreover, $\tilde{\eta}$ has no self-intersections. Indeed, $\eta|_{[t_1, t_2]}$ has no self-intersections because it is a free geodesic of H_+ ; $\phi \circ \gamma_{A,B}|_{[s_1, s_2]}$ has no self-intersections because ϕ is injective and $\gamma_{A,B}$ is a geodesic of G . Lastly,

$$\eta(t_1, t_2) \cap \phi \circ \gamma_{A,B}(s_1, s_2) = \emptyset$$

by the definition of (t_1, t_2) in (6.7).

By Jordan's closed curve theorem, $H_+ \setminus \tilde{\eta}([t_1, t_2 + s_2 - s_1])$ has two connected components and we call $\tilde{\mathcal{O}}$ the bounded one; part of its boundary, $\eta(t_1, t_2)$, is contained in R , part of it, $\phi \circ \gamma_{A,B}(s_1, s_2)$, is contained in ∂R . Since $\tilde{\mathcal{O}}$ is a connected set which intersects R only at its boundary and contains points of R , we get that $\tilde{\mathcal{O}} \subset R$.

Since $\phi \circ \gamma_{A,B}(s_1, s_2)$ intersects the left hand plane we have that

$$\mathcal{O} := \tilde{\mathcal{O}} \cap \{\operatorname{Re}(z) < 0\} \neq \emptyset.$$

We set

$$\mathcal{U} = \phi^{-1}(\mathcal{O}).$$

We refer the reader to figure 2 below: the free geodesic $\eta|_{[t_1, t_2]}$ is the bold segment on the imaginary axis, the curve $\phi \circ \gamma_{A,B}$ is the curved line and the set \mathcal{O} is shaded.

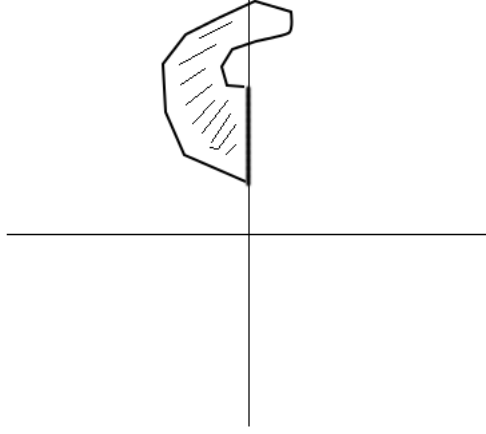


Figure 2

Now we note that \mathcal{U} is not empty; indeed, let us consider the points of G in a thin strip around $\gamma_{A,B}(s_1, s_2)$. Since ϕ is continuous, these points are sent by ϕ into a thin strip around $\phi \circ \gamma_{A,B}(s_1, s_2)$; since $\tilde{\mathcal{O}} \subset R$ it is not hard to show that this strip intersect \mathcal{O} , implying that \mathcal{U} is not empty.

For the function g of (4.5) we define $\tilde{\psi} \in \mathcal{A}$ by

$$\tilde{\phi}(x) = \begin{cases} \phi(x) & \text{if } x \notin \mathcal{U} \\ g \circ \phi(x) & \text{if } x \in \mathcal{U}. \end{cases}$$

We assert that the three points below hold.

- 1) $\tilde{\phi} \in \mathcal{D}(d) \oplus i\mathcal{D}(d)$.
- 2) $\tilde{\phi}(A) = \tilde{A}$, $\tilde{\phi}(B) = \tilde{B}$, $\tilde{\phi}(C) = \tilde{C}$.
- 3) $\tilde{\mathcal{E}}(\tilde{\phi}) < \tilde{\mathcal{E}}(\phi)$.

By the first two points $\tilde{\phi} \in \mathcal{A}$ while the third one contradicts the minimality of ϕ , ending the proof of lemma 6.4.

We forego the proof of point 2), which follows easily from the definition of $\tilde{\phi}$ and is the same as in lemma 4.3. We also skip the proof of point 3), since again it is analogous to that of lemma 4.3; we prove point 1).

Since the operator d is local, it suffices to show that, for all $x_0 \in G$, there is $r > 0$ such that $d\tilde{\phi}$ exists in $B(x_0, r)$; a more formal definition would be that, for all C^1 functions ψ supported in such a ball, $\tilde{\phi} \cdot \psi \in \mathcal{D}(d)$.

There are various cases. The first one is when $x_0 = \phi^{-1}(z_0)$ and $B(z_0, r) \subset \mathcal{O}^c$. In this case, for r' small, $\tilde{\phi} = \phi$ on $B(x_0, r') \subset \phi^{-1}(B(z_0, r))$ and we already know that ϕ is in the domain of d .

The second case is when $B(z_0, r) \subset \mathcal{O}$. We set $\phi = u + iv$ and we see that, for r' small, on $B(x_0, r') \subset \phi^{-1}(B(z_0, r))$ we have $\tilde{\phi} = iv$, and thus it is in the domain of d .

The same argument applies when $z_0 \in \partial\mathcal{O}$ but z_0 is not on the imaginary axis.

The last case is when $\phi(z_0) \in \partial\mathcal{O}$ is on the imaginary axis. In this case, $\tilde{\phi} = g \circ \phi$ in $B(x_0, r') \subset \phi^{-1}(B(z_0, r))$, and the same argument of lemma 4.2 shows that $\tilde{\phi}$ is in the domain of d .

\\

Proof of point 3) of theorem 1. We follow [16]. Let γ be a geodesic in $\phi(G)$ with $\gamma(0) = x$ and $\gamma(1) = y$. It suffices to show that $\phi(G)$ is a join of curves of the type $\phi \circ F_{j_0 \dots j_{l-1}} \circ \gamma_{i,j}$ with $i \neq j \in \{A, B, C\}$.

For a fixed l , we consider the cells $\phi \circ F_{j_0 \dots j_{l-1}}(G)$. Since G is compact and we are supposing that ϕ is injective, ϕ is a homeomorphism on its image, and thus for the interior part we have

$$[\phi \circ F_{j_0 \dots j_{l-1}}(G)]^\circ = \phi[F_{j_0 \dots j_{l-1}}(G)^\circ]. \quad (6.9)$$

We define the set $\{(t_1^\alpha, t_2^\alpha)\}_\alpha$ of maximal intervals such that

$$\gamma((t_1^\alpha, t_2^\alpha)) \subset \phi \circ F_{j_0 \dots j_{l-1}}(G)^\circ \quad (6.10)$$

for some $j_0 \dots j_{l-1}$.

Since (t_1^α, t_2^α) is maximal, (6.9) and (6.10) imply that $\gamma(t_1^\alpha), \gamma(t_2^\alpha) \in F_{j_0 \dots j_{l-1}}(\{A, B, C\})$; since γ is a geodesic, it has no self-intersections and thus $\gamma(t_1^\alpha) \neq \gamma(t_2^\alpha)$. By (6.10) and the fact that γ is a geodesic, $\gamma|_{[t_1^\alpha, t_2^\alpha]}$ is minimal among all curves with image in $\phi \circ F_{j_0 \dots j_{l-1}}(G)$. To fix ideas, we suppose that $\gamma(t_1^\alpha) = \phi \circ F_{j_0 \dots j_{l-1}}(A)$, $\gamma(t_2^\alpha) = \phi \circ F_{j_0 \dots j_{l-1}}(B)$; by proposition 6.2 and using the notation of (6.1), for $t \in [t_1^\alpha, t_2^\alpha]$, up to reparametrisation we have

$$\gamma(t) = \phi \circ F_{j_0 \dots j_{l-1}} \circ \gamma_{A,B} \left(\frac{t - t_1^\alpha}{t_2^\alpha - t_1^\alpha} \right)$$

i.e. on these intervals γ is the image of a geodesic of the harmonic gasket, as we wanted.

Recall that we fixed l ; now we note that

$$[0, 1] \setminus \bigcup_\alpha (t_1^\alpha, t_2^\alpha) = [0, \epsilon_1] \cup [1 - \epsilon_2, 1]$$

with $\epsilon_1, \epsilon_2 \rightarrow 0$ as $l \rightarrow +\infty$. Thus the thesis follows letting $l \rightarrow +\infty$.

\\

References

- [1] P. Alonso-Ruiz, F. Baudoin, Dirichlet forms on metric measure spaces as Mosco limits of Korevaar-Schoen energies, preprint 2023.
- [2] P. Alonso-Ruiz, U. Freiberg, J. Kigami, Completely symmetric resistance forms on the Stretched Sierpinski Gasket, Journal of Fractal Geometry, **5-3**, 227-277, 2018.

- [3] A. Ambrosetti, A. Malchiodi, *Nonlinear Analysis and semilinear elliptic problems*, Cambridge, 2006.
- [4] L. Ambrosio, N. Gigli, G. Savaré, *Heat flow and calculus on metric measure spaces with Ricci curvature bounded below - the compact case. Analysis and numerics of Partial Differential Equations*, 63-115, Springer, Milano, 2013.
- [5] F. Baudoin, *Korevaar-Schoen-Sobolev spaces and critical exponents in metric measure spaces*, preprint 2024.
- [6] U. Bessi, *Another point of view on Kusuoka's measure*, DCDS, **41-7**, 3251-3271, 2021.
- [7] U. Bessi, *Harmonic embeddings of the stretched Sierpinski gasket*, NoDEA, **30-6**, paper no. 80, 34pp.
- [8] U. Bessi, *Hodge theory on the harmonic gasket and other fractals*, to appear on *Nonlinear Analysis*, TMA.
- [9] G. Bouchitté, G. Buttazzo, P. Seppecher, *Energies with respect to a measure and applications to low dimensional structures*, Calc. Var. and PDE, **5**, 37-54, 1997.
- [10] S. Cao, H. Qiu, *A topological proof of the non-degeneracy of harmonic structures on Sierpinski gaskets*, *Analysis in theory and applications*, **36**, 510-516, 2020.
- [11] J. Eels, J. H. Sampson, *Harmonic mappings of Riemannian manifolds*, *American Journal of Mathematics*, **86**, 109-160, 1964.
- [12] A. Johansson, A. Öberg, M. Pollicott, *Ergodic theory of Kusuoka's measures*, *J. Fractal Geom.*, **4**, 185-214, 2017.
- [13] J. Heinonen, P. Koskela, N. Shanmugalingam, J. T. Tyson, *Sobolev spaces on metric measure spaces*, Cambridge, 2015.
- [14] N. Kajino, *Analysis and geometry of the measurable Riemannian structure on the Sierpinski gasket*, *Contemporary Mathematics*, **600**, 91-133, 2013.
- [15] J. Kigami, *Analysis on fractals*, Cambridge, 2001.
- [16] J. Kigami, *Measurable Riemannian geometry on the Sierpinski gasket: the Kusuoka measure and the Gaussian heat kernel estimates*, *Math. Ann.*, **340**, 781-804, 2008.
- [17] P. Koskela, Y. Zhou, *Geometry and Analysis of Dirichlet forms*, *Adv. Math.*, **231**, 2755-2801, 2012.
- [18] S. Kusuoka, *Dirichlet forms on fractals and products of random matrices*, *Publ. Res. Inst. Math. Sci.*, **25**, 659-680, 1989.
- [19] I. D. Morris, *Ergodic properties of matrix equilibrium state*, *Ergodic Theory and Dyn. Sys.*, **38/6**, 2295-2320, 2018.
- [20] M. Piraino, *The weak Bernoulli property for matrix equilibrium states*, *Ergodic theory Dynam. Systems*, **40**, 2219-2238, 2020.
- [21] R. S. Strichartz, *Differential equations on fractals, a tutorial*, Princeton, 2006.
- [22] M. E. Taylor, *Partial differential equations, Basic theory*, Springer, Berlin, 1996
- [23] A. Teplyaev, *Harmonic coordinates on fractals with finitely ramified cell structure*, *Canad. J. Math.*, **60**, 457-480, 2008.
- [24] A. Teplyaev, *Energy and Laplacian on the Sierpinski gasket*, *Proceedings of symposia in pure Mathematics*, **72-1**, 131-154, Providence, R. I., 2004.

[25] K. Tsougas, Non-degeneracy of the harmonic structure on Sierpinski gaskets, *J. Fractal Geom.*, **6**, 143-156, 2019.