FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

DOCTORAL THESIS

Daniel Cameron Campbell

Properties of mappings of finite distortion

Department of Mathematical Analysis

Supervisor of the doctoral thesis: doc. RNDr. Stanislav Hencl, Ph.D.
Study programme: Mathematics
Study branch: Mathematical Analysis

Prague 2017

I declare that I carried out this doctoral thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Sb., the Copyright Act, as amended, in particular the fact that the Charles University has the right to conclude a license agreement on the use of this work as a school work pursuant to Section 60 subsection 1 of the Copyright Act.

In .... date .... signature of the author

Title: Properties of mappings of finite distortion

Author: Daniel Cameron Campbell

Department: Department of Mathematical Analysis

Supervisor: doc. RNDr. Stanislav Hencl, Ph.D., Department of Mathematical Analysis

Abstract: In the following thesis we will be mostly concerned with questions related to the regularity of solutions to non-linear elasticity models in the calculus of variations. An important step in this is question is the approximation of Sobolev homeomorphisms by diffeomorphisms. We refine an approximation result from [10] which, for a given planar Sobolev (or Sobolev-Orlicz) homeomorphism, constructs a diffeomorphism arbitrarily close to the original map in uniform convergence and in terms of the Sobolev-Orlicz norm. Further we show, in dimension 4 or higher, that such an approximation result cannot hold in Sobolev spaces $W^{1,p}$ where p is too small by constructing a sense-preserving homeomorphism with Jacobian negative on a set of positive measure.

The class of mappings referred to as mappings of finite distortion have been proposed as possible models for deformations of bodies in non-linear elasticity. In this context a key property is their continuity. We show, by counter-example, the surprising sharpness of the modulus of continuity with respect to the integrability of the distortion function. Also we prove an optimal regularity result for the inverse of a bi-Lipschitz Sobolev map in $W^{k,p}$ and composition of Lipschitz maps in $W^{k,p}$ comparable with the classical inverse mapping theorem. As a consequence we retrieve a Sobolev equivalent of the implicit function theorem.

Keywords: non-linear elasticity, regularity, approximation, homeomorphism

I would like to thank my supervisor, advisor, teacher, mentor and friend Standa Hencl for his good-natured care and help. His support during the course of my doctoral studies has been invaluable. I truly regard him as an inspiring mathematician with whom it has been a privilege to work. I also owe many thanks to Honza Malý for his excellent tuition and for the insight which he has helped me gain. Besides the astute mathematical comments and questions he has helped me with I am also very grateful for the humour, which he has brought to my time as a doctoral student. There are several other mathematicians I am particularly grateful to for their help and hospitality. I would like to thank professor Jani Onninen for his kind invitation and the time he has dedicated to our discussions and joint research. Also I am very appreciative to Tadeusz Iwaniec for his inspiring comments and the time he has spent with me in discussion. I am also obliged to professor Carlo Sbordone and professor Luigi Grecco for their invitations and cooperation. My thanks also belong to my dear colleague and friend Emanuela Radici for her help and interesting questions.

Most of all I would like to thank my family for their support. My parents for the kindness of their rearing and my education. My daughter Julia for interesting comments and suggestions, which have helped me in completing this thesis. My son Timothy and little Emily for their support, encouragement and the interesting graphics they supply me with. I would like to express my gratitude to my parents-in-law amongst my other very sympathetic close and loved ones. Most of all I wish to thank my understanding, patient and loving wife Pavla for her unbounded help and support in all aspects of my life including academic.

Contents

Outline

We start this thesis with an introduction explaining how each of the included articles fits into the wider context of known results and ongoing work in the field. The included articles are

- A note on mappings of finite distortion: Examples for the sharp modulus of continuity; Daniel Campbell and Stanislav Hencl, Ann. Acad. Sci. Fenn. Math., 36 (2011), 531—536,
- The weak inverse mapping theorem; Daniel Campbell, F. Konopecký and Stanislav Hencl, Z. Anal. Anwendungen 34 no. 3 (2015), 321-342,
- Diffeomorphic approximation of Planar Sobolev Homeomorphisms in Orlicz-Sobolev spaces; Daniel Campbell, Preprint (2015),
- Approximation of $W^{1,p}$ Sobolev homeomorphism by diffeomorphisms and the signs of the Jacobian; Daniel Campbell, Stanislav Hencl and Ville Tengvall Preprint (2016).

1. Introduction

In the following introduction we will discuss a variational model for elastic bodies in non-linear elasticity and hyper-elastic materials. Our particular point of interest in this are classes of mappings, in which the solution may lie and their properties. We start with two domains in $\mathbb{R}^n$ which will represent an elastic body, Ω the body in rest state and Δ the body under deformation. Our aim is to find the map which describes the deformation of Ω onto Δ. We may require that our map has prescribed boundary values from $\partial\Omega$ onto $\partial\Delta$ or simply that the image of $\overline{\Omega}$ is $\overline{\Delta}$ with the boundary being mapped onto the boundary. We may then ask in which class of mappings should we search for our solution. Our expectations of the model are that an elastic deformation does not break the body or allow the formation of cracks or cavities. Moreover we impose the concept of the non-interpenetration of matter. This leads us to expect the solution to be a homeomorphism. There are a number of models whose solutions are known to be homeomorphisms. Since our model will work with with spacial derivatives, we will require that a candidate for the deformation belongs to some Sobolev space making the class of Sobolev homeomorphisms of particular interest.

Our criteria for the deformation u will be to minimise an energy functional of the type

$$I(u) = \int_{\Omega} F(\nabla u)dx,$$

where one considers a homogenous body (thus F does not depend on x) in free space (thus alleviating the dependance on the functional values of u). Observation and intuition suggest that whenever we deform the exterior of a body in a linear manner that the deformation is also linear inside the body. In terms of the calculus of variations this property of the functional I is known as quasi-convexity. On the other hand it is also natural to expect that

$$F(A) \to \infty \text{ as } |A| \to \infty \text{ and } F(A) \to \infty \text{ as } \det A \to 0^+$$

with $F(A) = \infty$ whenever $\det A \leq 0$. The first property deals with stretching and the second with squeezing of the body. It is the second of these properties that prevents F from having p-growth which would allow us to guarantee the existence of a minimiser in the weak sequential closure of $W^{1,p}$ homeomorphisms by the result of [1]. On the other hand we do not have a complete answer from the result of [2] either as for this we need polyconvexity and we expect only quasiconvexity, which is a weaker concept.

In a number of simplified cases however the above results give us the existence of a solution. Under given conditions it can be deduced from the properties of the functional that the minimiser is a homeomorphism. This however still leaves us with the important question of the regularity of these minimisers. Since an important tool in many regularity proofs is to test the solution against itself in the weak formulation of the related equation (or the variational formulation itself) it is natural to try to apply this technique. Unfortunately the test may not make sense unless we have enough a priori regularity of the solution. Therefore we wish to test our solution against smooth maps approaching our solution and then pass to the limit. Unless we test with a map admissible to our problem (i.e. a homeomorphism) then the test may fail to have any relevance to our original task. Thus we arrive at the well-known Ball-Evans problem of approximating a Sobolev homeomorphism by diffeomorphisms in $W^{1,p} \cap L^\infty$.

The Ball-Evans problem is an important first step towards our final goal. It satisfies our requirements for the test in the simplest case we may take the p-harmonic energy

$$I(u) = \int_\Omega |\nabla u|^p dx.$$

In many cases we will have much more complicated energy functionals and want the approximation to be close to the original map not only in terms of the $W^{1,p}$ norm but also in terms of this energy. However, for other functionals, we will be interested in different distances, perhaps including also the inverse, for example if

$$I(u) = \int_\Omega |\nabla u|^p dx + \int_\Delta |\nabla u^{-1}|^q dy$$

then we would naturally approximate u in $W^{1,p}$ and u^{-1} in $W^{1,q}$ simultaneously. There are other, more complicated, energies for which we would be interested in approximating other values simultaneously. Although the planar Ball-Evan's problem has been solved, with the exception of the $W^{1,1} - W^{1,1}$ problem (see [19]), the question of simultaneous approximation of the inverse, or approximation in other more general energies, is open.

Approximating homeomorphisms with diffeomorphisms in uniform convergence is a challenging question in itself and was thoroughly studied in the 20$^{\text{th}}$ century. The interested reader can find a review of the over 60-year history of this interesting problem in the introduction of [10]. In our case we require something somewhat stronger (also convergence in terms of energy) but have the additional advantage that our map is in the Sobolev space.

As of yet the question is still rather open. Known positive results are all planar. The first result was in 2009 due to Mora-Corral [15] in the case where the homeomorphism is smooth outside a point. This was followed by the breakthrough of Iwaniec, Kovalev and Onninen [12; 13] in 2011 where it was proven that any planar homeomorphism in $W^{1,p}$, $1 < p < \infty$, can be approximated in $W^{1,p} \cap L^\infty$. The approach relies on so-called p-harmonic replacement on refining 'grids'. In 2011 Mora-Corral and Pratelli, [16], proved that a similar important question of approximating by piecewise affine homeomorphisms is in fact equivalent with approximating by diffeomorphisms. Using this result the remaining open case of approximating $W^{1,1}$ homeomorphisms in $W^{1,1}$ was solved in 2014 by

Hencl and Pratelli in [10]. Their approach made use of a rather different technique combining analytical properties of planar homeomorphisms like Lebesgue points and a.e. differentiability with an extension theorem which uses geodesic filling. Following this it was proved by Radici [20] that the core extension theorem extends also to the case $W^{1,p}$ enabling one to approximate a homeomorphism of finite distortion also in $W^{1,p}$ using the technique developed by Hencl and Pratelli. The author further extends the technique in the included article in Chapter 4 to approximate general (for example with $J_f = 0$ a.e.) Δ_2 Orlicz-Sobolev homeomorphisms in their respective Orlicz-Sobolev space using the standard Luxembourg norm. In a result by Pratelli [19] it was shown that, using (an adaption of) the same approach, one can approximate a homeomorphism in $W^{1,1}$ and its inverse in $W^{1,1}$ simultaneously. A further refinement by Pratelli and Radici is the result in preparation that allows one to approximate also planar BV homeomorphisms in area strict convergence along with the inverse.

One is naturally led to ask the question, whether the above techniques can be generalised to higher dimension. It is known that even for boundary data on a sphere uniformly approaching identity that the harmonic replacement may fail to be injective in dimension higher than two. As of yet there is no known generalisation of the Hencl-Pratelli geodesic filling theorem to higher dimension. Thus the most critical question of dimension three is still widely open and requires some new tools.

Further Iwaniec and Onninen [14] prove an approximation result for the class of $W^{1,p}, 1 < p < \infty$, planar monotone mappings which, for $p \geq 2$, are exactly the weak sequential closure of homeomorphisms in $W^{1,p}$. This allows them to apply the direct method and they further show that their model allows for collapsing of matter, while maintaining a weaker invertibility condition. On the other hand one would still be interested in a characterisation of the closure of $W^{1,p}$ homeomorphisms for $p < 2$ and also in BV. As of yet this question is open.

A partial answer to the Ball-Evans approximation question is also known in higher dimension. Assume that we have a homeomorphism u in $W^{1,p}$ whose Jacobian changes sign; by which we mean that

$$\mathcal{L}^n(\{J_u < 0\}) > 0 \text{ and } \mathcal{L}^n(\{J_u > 0\}) > 0.$$

We claim that such a homeomorphism cannot be approximated by diffeomorphisms. Indeed assume that one has an approximating sequence of diffeomorphisms which are (without loss of generality) sense-preserving, as is u. Then it is well known that the Jacobian of these mappings cannot change sign, in fact they are strictly positive. Now, from the convergence in $W^{1,p}$, we have the convergence of the derivatives (at least for a subsequence) point-wise almost everywhere, especially almost everywhere in $\{J_u < 0\}$. But the limit of positive functions is non-negative and J_u is negative on this set (of positive measure). Thus we arrive simply at a contradiction. Such a mapping provides a satisfactory answer to the Ball-Evans problem as stated. On the other hand if F heavily penalises small and non-positive values of the determinant one may restrict ones interest to homeomorphisms of positive Jacobian. We conjecture that there exists a $W^{1,p}$ mapping equal to $(x_1, x_2, \ldots, x_n) \to (-x_1, x_2, \ldots, x_n)$ on the boundary of the unit cube but with positive Jacobian almost everywhere. Any diffeomorphism approximating this homeomorphism would have negative Jacobian everywhere because the topological degree of the image is -1. As of yet no such map has been constructed and this question remains open.

Moreover the question of the sign of the Jacobian of a homeomorphism, which we attribute to Hajłasz, is one of interest in itself. The reason for this is that the Jacobian carries important analytical information about a mapping and is often used for proving topological results, for example continuity, openness and discreteness of mappings of finite distortion. Under sufficient regularity the sign of the Jacobian closely corresponds to the degree of a map, not only in the smooth case but also in the grand Sobolev space $W^{1,n)}$. Further, amongst many other applications, the Jacobian plays an essential role in change of variable formulas deriving from the area-coarea formula.

It was previously known (see [9]) that the sign of the Jacobian of a homeomorphism $u \in W^{1,p}$ with $p > [n/2]$ (here $[n/2]$ is the lower integer part of the number $n/2$) cannot change sign. It has been conjectured by Hajłasz that this result can be improved to include also the border-line case. This question became much more interesting recently when, surprisingly, it was proven that $p = [n/2]$ is in fact the border-line case for the stability of the sign of the Jacobian of a homeomorphism.

The first construction of a Sobolev homeomorphism in [11] with Jacobian changing sign is due to Hencl and Vejnar and belongs to $W^{1,1}$. By careful refinement and extension of the ideas used there the author together with Hencl and Tengvall constructed, for any $1 \leq p < [n/2]$ a homeomorphism in $W^{1,p}$ whose Jacobian changes sign and the construction can be found in the article reprinted in Chapter 5. This paper also identified a gap in the argument of Hencl and Vejnar and fixes it.

The class of mappings of finite distortion have also been proposed as model deformations of elastic bodies. In fact it is known that any sufficiently regular bi-Sobolev homeomorphism is a mapping of finite distortion and a $W^{1,n}$ mapping of finite distortion with $K \in L^p$, $p > n - 1$ and homeomorphic boundary values is a homeomorphism (see [8]). We include an example proving that an estimate of the modulus of continuity dependent on the integrability of the distortion is sharp. This essentially tells us, under given information on distortion of a map, how much it possibly stretches the material. This result is included in Chapter 2.

Further we include in Chapter 3 a result, which generalises the classical and well known inverse mapping theorem and implicit function theorems to the context of bi-Lipschitz mappings in $W^{k,p}$. We prove that the inverse of a bi-Lipschitz mapping in $W^{k,p}$ is also in $W^{k,p}$ for all $k \in \mathbb{N}$ and $p > 1$. We further prove that the composition of a Lipschitz $W^{k,p}$ Sobolev mapping with a bi-Lipschitz $W^{k,p}$ mapping is a $W^{k,p}$ mapping and the product of two Lipschitz $W^{k,p}$ mappings are also $W^{k,p}$ maps. A Sobolev version of the implicit function theorem is also proved here.

2. A note on mappings of finite distortion: Examples for the sharp modulus of continuity

Daniel Campbell and Stanislav Hencl

Ann. Acad. Sci. Fenn. Math., 36 (2011), 531—536.

Annales Academiæ Scientiarum Fennicæ
Mathematica
Volumen 36, 2011, 1–6

A NOTE ON MAPPINGS OF FINITE DISTORTION: EXAMPLES FOR THE SHARP MODULUS OF CONTINUITY

Daniel Campbell and Stanislav Hencl

Charles University, Department of Mathematical Analysis
Sokolovská 83, 186 00 Prague 8, Czech Republic; danuk@centrum.cz

Charles University, Department of Mathematical Analysis
Sokolovská 83, 186 00 Prague 8, Czech Republic; hencl@karlin.mff.cuni.cz

Abstract. We construct a mapping with exponentially integrable distortion which attains a modulus of continuity by Onninen and Zhong, showing that it is sharp.

1. Introduction

Let $\Omega \subset \mathbf{R}^n$, $n \geq 2$, be a connected and open set. A Sobolev mapping $f \in W^{1,1}_{\mathrm{loc}}(\Omega, \mathbf{R}^n)$ is said to have finite distortion if the Jacobian $J_f(x)$, i.e., the determinant of the matrix of derivatives $Df(x)$ is locally integrable and there is a measurable function $K(x) \geq 1$ finite almost everywhere such that

$$|Df(x)|^n \leq K(x) J_f(x) \quad \text{a.e. } x \in \Omega.$$

Here we have used the operator norm of the differential matrix with respect to the Euclidean distance.

If we, moreover, require that $K(x) \in L^\infty(\Omega)$, we arrive at mappings of bounded distortion also called quasiregular mappings. In [7] Reshetnyak proved among many other things that quasiregular mappings are Hölder continuous with the exponent $1/K$, where K is the L^∞ norm of the distortion. It has been shown recently that mappings of finite distortion with exponentially integrable distortion

$$\exp(\lambda K(x)) \in L^1(\Omega)$$

share many nice properties of mappings of bounded distortion. We would like to point the reader's attention to the monographs [1] and [3] for the motivation, applications and the history of the subject.

Our aim is to study the modulus of continuity of the mappings of finite distortion with $\exp(\lambda K) \in L^1(\Omega)$. Let us first recall the history of such estimates. First, it was shown by Iwaniec, Koskela and Onninen [2] that mappings in this class are continuous and satisfy

$$|f(x) - f(y)| \leq \frac{C}{\log^{1/n} \log(e^e + 1/|x-y|)}.$$

This was later improved by Koskela and Onninen [5] to

$$|f(x) - f(y)| \leq \frac{C}{\log^{\lambda/n - \varepsilon}(1/|x-y|)}$$

doi:10.5186/aasfm.2011.3633
2010 Mathematics Subject Classification: Primary 30C65, 46E35.
Key words: Mapping of finite distortion, modulus of continuity.

and finally using very delicate arguments it has been shown by Onninen and Zhong [6] that

$$|f(x)-f(y)|\le \frac{C}{\log^{\lambda/n}(1/|x-y|)}. \tag{1.1}$$

Extremal mappings for continuity of mappings of finite distortion are usually radial maps and therefore the natural candidate for the extremal map is

$$f_0(x)=\frac{x}{|x|}\frac{1}{\log^{\lambda/n}(1/|x|)}.$$

Standard computations (see (2.2) below) give us

$$K(x)=\frac{n}{\lambda}\log\frac{1}{|x|}$$

and hence

$$\int_{B(0,\frac12)}\exp(\lambda K(x))\,dx=\int_{B(0,\frac12)}\frac{1}{|x|^n}\,dx=\infty.$$

This elementary computation suggests that there is some room for improvement in the estimate (1.1) and maybe we can add some supplementary term like $\log\log 1/|x-y|$ to some negative power to our estimate. We show that, surprisingly, this is not the case and the modulus of continuity (1.1) is already sharp.

Theorem 1.1. *Given $\lambda>0$, there is a mapping of finite distortion $f\colon B(0,\frac12)\to \mathbf{R}^n$ such that*

$$\int_{B(0,\frac12)}\exp(\lambda K(x))\,dx<\infty$$

and

$$|f(x)-f(0)|\ge\frac{C}{\log^{\lambda/n}(1/|x|)}\quad\text{for all } x\in B(0,\tfrac12).$$

There have also been studies on mappings of subexponentially integrable distortion (see e.g. [4]). One requires that

$$\int_B \exp\big(\mathscr{A}(K(x))\big)\,dx<\infty \tag{1.2}$$

for some Orlicz function $\mathscr{A}$ and the above mentioned example corresponds to the case $\mathscr{A}(t)=\lambda t$. We call an infinitely differentiable and strictly increasing function $\mathscr{A}\colon[0,\infty)\to[0,\infty)$ with $\mathscr{A}(0)=0$ and $\lim_{t\to\infty}\mathscr{A}(t)=\infty$ an Orlicz function. As usual we impose additional condition

$$\int_1^\infty\frac{\mathscr{A}'(s)}{s}=\infty. \tag{1.3}$$

It is easy to see that the critical functions for this condition are

$$\begin{aligned}&\mathscr{A}_1(t)=\lambda t,\quad \mathscr{A}_2(t)=\lambda\frac{t}{\log(e+t)},\\ &\mathscr{A}_3(t)=\lambda\frac{t}{\log(e+t)\log(e+\log(e+t))}\quad\text{and so on.}\end{aligned}\tag{1.4}$$

We will also require that

$$
\begin{aligned}
&\text{(i)} \quad \exists t_0 > 0 \ \forall t > t_0 \ \ \mathscr{A}^{-1}(nt) < t^{\frac{3}{2}},\\
&\text{(ii)} \quad \mathscr{A}'(t) \text{ is non-increasing},\\
&\text{(iii)} \quad b'(t) \text{ is non-increasing for } b(t) := \frac{t}{\mathscr{A}(t)},\\
&\text{(iv)} \quad b(0) := \lim_{t\to 0+} b(t) \text{ is finite and positive.}
\end{aligned}
\tag{1.5}
$$

Let us note that the critical functions from (1.4) satisfy these conditions and therefore these assumptions are not substantially restrictive. It has been shown in [4] that a mapping f is continuous under the assumptions (1.2) and (1.3) and that the assumption (1.3) is sharp.

It was proved in [6] that under the assumptions (1.2) and (1.3) we have

$$|f(x) - f(y)| \le C \exp\Big(- \int_{|x-y|}^{R} \frac{dt}{t\mathscr{A}^{-1}(n \log C/t)} \Big)$$

for $|x-y|$ sufficiently small and $B(x, 80R) \subset \Omega$. Note further that if we put $\mathscr{A}_1(t) = \lambda t$ we arrive at the modulus given in (1.1). Our result shows the sharpness of this estimate.

Theorem 1.2. *Suppose that an Orlicz function $\mathscr{A}$ satisfies* (1.3) *and* (1.5)*. Then there is a ball $B := B(0, r)$ and a mapping of finite distortion $f\colon B \to \mathbf{R}^n$ such that*

$$\int_B \exp\big(\mathscr{A}(K(x))\big)\, dx < \infty$$

and

$$|f(x) - f(0)| \ge C \exp\Big(- \int_{|x|}^{1/2} \frac{dt}{t\mathscr{A}^{-1}(n \log 1/t)} \Big) \quad \textit{for all } x \in B. \tag{1.6}$$

2. Proofs of the theorems

To prove Theorem 1.1 we simply set

$$f(x) = \frac{x}{|x|} \frac{(\log 1/|x|)^{\frac{a}{\log 1/|x|}}}{\log^{\lambda/n}(1/|x|)},$$

where $a > 0$. The additional term clearly satisfies

$$\lim_{|x|\to 0} (\log 1/|x|)^{\frac{a}{\log 1/|x|}} = 1$$

and thus the modulus of continuity of our f is exactly as required in (1.1). On the other hand, the additional term slightly affects the distortion and the standard computation (see the general case below for details) will give us

$$K(x) \sim \frac{n}{\lambda} \log \frac{1}{|x|} - \frac{n^2 a}{\lambda^2} \log\log \frac{1}{|x|},$$

and hence

$$\int_{B(0,\frac{1}{2})} \exp(\lambda K(x))\, dx < \infty$$

for sufficiently large a.

To prove Theorem 1.2, let us put $B := B(0, \min\{\exp(-t_0), e^{-e}\})$ and choose $\alpha > b(0)^{-1}n^{-2}$. Without loss of generality we can assume that t_0 is big enough such that

$$t^{\frac{3}{2}} < \frac{1}{\alpha(\alpha+1)}\frac{t^2}{\log t} \quad \text{for all } t > t_0. \tag{2.1}$$

We define the function f as,

$$f(x) := \frac{x}{|x|}\exp\Big(-\int_{|x|}^{\frac{1}{2}} \frac{1}{t\mathscr{A}^{-1}(n\log\frac{1}{t})}dt\Big)(\log|x|^{-1})^{\frac{\alpha+2}{\log|x|^{-1}}} \quad \text{for } x \neq 0.$$

Note that

$$\lim_{|x|\to 0}(\log|x|^{-1})^{\frac{\alpha+2}{\log|x|^{-1}}} = \lim_{|x|\to 0}\exp\Big(\frac{(\alpha+2)\log\log|x|^{-1}}{\log|x|^{-1}}\Big) = 1,$$

which easily gives that f satisfies the condition given in (1.6).

Let $\rho\colon (0,\infty) \to (0,\infty)$ be a strictly monotone, differentiable function and let us consider the mapping

$$f(x) = \frac{x}{|x|}\rho(|x|), \quad x \neq 0.$$

It can be verified by an elementary computation (see e.g. [1, Chapter 2.6.]) that

$$\begin{aligned} |Df(x)| &= \max\Big\{\frac{\rho(|x|)}{|x|}, |\rho'(|x|)|\Big\}, \quad \text{and thus} \\ K(x) &= \max\Big\{\frac{\rho(|x|)}{|x|\ |\rho'(|x|)|}, \frac{|x|\ |\rho'(|x|)|}{\rho(|x|)}\Big\}. \end{aligned} \tag{2.2}$$

It follows that for our mapping we obtain

$$|Df(x)| = \frac{|f(x)|}{|x|}\max\Big\{1, \Big(\frac{1}{\mathscr{A}^{-1}(n\log|x|^{-1})} + (\alpha+2)\frac{\log\log|x|^{-1}-1}{\log^2|x|^{-1}}\Big)\Big\}.$$

Clearly,

$$\lim_{x\to 0}\Big(\frac{1}{\mathscr{A}^{-1}(n\log|x|^{-1})} + (\alpha+2)\frac{\log\log|x|^{-1}-1}{\log^2|x|^{-1}}\Big) = 0,$$

and therefore the greater element is the first one. From (1.5) (i) and (2.1) we obtain

$$\mathscr{A}^{-1}(nt) < t^{\frac{3}{2}} < \frac{1}{\alpha(\alpha+1)}\frac{t^2}{\log t} \quad \text{for all } t > t_0.$$

This, however, implies that

$$\alpha(\alpha+1)\frac{\mathscr{A}^{-1}(nt)\log t}{t^2} < 1 \quad \text{for all } t > t_0.$$

Now, by multiplying both sides by $\frac{\mathscr{A}^{-1}(nt)\log t}{t^2}$ and by substituting $t = \log|x^{-1}|$ we get that

$$\mathscr{A}^{-1}(n\log|x|^{-1})\frac{\log\log|x|^{-1}}{\log^2|x|^{-1}} > \alpha(\alpha+1)\Big(\mathscr{A}^{-1}(n\log|x|^{-1})\frac{\log\log|x|^{-1}}{\log^2|x|^{-1}}\Big)^2$$

for all $x \in B$. Using this fact and because $\log\log|x|^{-1} > 1$ for all $x \in B$, we deduce that

$$\begin{aligned}
K(x) &= \frac{1}{\left(\frac{1}{\mathscr{A}^{-1}(n\log|x|^{-1})} + (\alpha+2)\frac{\log\log|x|^{-1}-1}{\log^2|x|^{-1}}\right)} \\
&\le \frac{\mathscr{A}^{-1}(n\log|x|^{-1})}{1+(\alpha+1)\mathscr{A}^{-1}(n\log|x|^{-1})\frac{\log\log|x|^{-1}}{\log^2|x|^{-1}}} \\
&\le \mathscr{A}^{-1}(n\log|x|^{-1})\Big(1-\alpha\mathscr{A}^{-1}(n\log|x|^{-1})\frac{\log\log|x|^{-1}}{\log^2|x|^{-1}}\Big) =: \tilde{K}(x).
\end{aligned}$$

Note that

$$\mathscr{A}^{-1}(n\log|x|^{-1}) - \tilde{K}(x) = \alpha n^2\Big(\frac{\mathscr{A}^{-1}(n\log|x|^{-1})}{n\log|x|^{-1}}\Big)^2 \log\log|x|^{-1}. \tag{2.3}$$

By (1.5) (iii) we obtain that

$$b(s) - b(0) = b'(\xi)s \ge b'(s)s$$

and therefore

$$\mathscr{A}'(s)\Big(\frac{s}{\mathscr{A}(s)}\Big)^2 = \frac{b(s)-sb'(s)}{b^2(s)}b^2(s) \ge b(0). \tag{2.4}$$

From (1.5) (ii) we know that $\mathscr{A}'(t)$ is a non-increasing function and hence

$$\mathscr{A}(a-d) = \mathscr{A}(a) - \mathscr{A}'(\xi)d \le \mathscr{A}(a) - \mathscr{A}'(a)d \tag{2.5}$$

for some $\xi \in (a-d, a)$. We now use (2.5) putting

$$a := \mathscr{A}^{-1}\Big(n\log\frac{1}{|x|}\Big), \quad d := \mathscr{A}^{-1}\Big(n\log\frac{1}{|x|}\Big) - \tilde{K}(x)$$

using (2.3) and then (2.4) (where we put $s := \mathscr{A}^{-1}(n\log|x|^{-1})$) to get that

$$\begin{aligned}
&\mathscr{A}(K(x)) \le \mathscr{A}(\tilde{K}(x)) \\
&\le \mathscr{A}\big(\mathscr{A}^{-1}(n\log|x|^{-1})\big) - \mathscr{A}'\big(\mathscr{A}^{-1}(n\log|x|^{-1})\big)\Big[\mathscr{A}^{-1}\big(n\log|x|^{-1}\big) - \tilde{K}(x)\Big] \\
&\le n\log|x|^{-1} - \alpha n^2\mathscr{A}'\big(\mathscr{A}^{-1}(n\log|x|^{-1})\big)\Big(\frac{\mathscr{A}^{-1}(n\log|x|^{-1})}{n\log|x|^{-1}}\Big)^2\log\log|x|^{-1} \\
&\le n\log|x|^{-1} - b(0)\alpha n^2\log\log|x|^{-1}.
\end{aligned}$$

But this, for $\alpha > b(0)^{-1}n^{-2}$, yields

$$\begin{aligned}
\int_B \exp(\mathscr{A}(K(x)))\,dx &\le \int_B \exp\Big(n\log\frac{1}{|x|} - b(0)\alpha n^2\log\log\frac{1}{|x|}\Big)\,dx \\
&\le \int_B \frac{1}{|x|^n\log^{b(0)\alpha n^2}\frac{1}{|x|}}\,dx < \infty.
\end{aligned}$$

□

References

[1] Astala, K., T. Iwaniec, and G. Martin: Elliptic partial differential equations and quasiconformal mappings in the plane. - Princeton Math. Ser. 48, Princeton Univ. Press, Princeton, NJ, 2009.

[2] Iwaniec, T., P. Koskela, and J. Onninen: Mappings of finite distortion: Monotonicity and continuity. - Invent. Math. 144, 2001, 507–531.

[3] Iwaniec, T., and G. Martin: Geometric function theory and nonlinear analysis. - Oxford Math. Monogr., Clarendon Press, Oxford, 2001.

[4] Kauhanen, J., P. Koskela, J. Malý, J. Onninen, and X. Zhong: Mappings of finite distortion: Sharp Orlicz-conditions. - Rev. Mat. Iberoamericana 19, 857–872, 2003.

[5] Koskela, P., and J. Onninen: Mappings of finite distortion: the sharp modulus of continuity. - Trans. Amer. Math. Soc. 355:5, 2003, 1905–1920.

[6] Onninen, J., and X. Zhong: A note on mappings of finite distortion: the sharp modulus of continuity. - Michigan Math. J. 53:2, 2005, 329–335.

[7] Reshetnyak, Yu. G.: Space mappings with bounded distortion. - Transl. Math. Monogr. 73, Amer. Math. Soc., 1989.

Received 28 May 2010

3. The weak inverse mapping theorem

Daniel Campbell, F. Konopecký and Stanislav Hencl

Z. Anal. Anwendungen 34 no. 3 (2015), 321-342.

Zeitschrift für Analysis und ihre Anwendungen
Journal for Analysis and its Applications
Volume 32 (2013), 301–323
DOI: 10.4171/ZAA/3445

The weak inverse mapping theorem

Daniel Campbell, Stanislav Hencl and František Konopecký

Abstract. We prove that if a bilipschitz mapping f is in $W^{m,p}_{\mathrm{loc}}(\mathbb{R}^n, \mathbb{R}^n)$ then the inverse f^{-1} is also a $W^{m,p}_{\mathrm{loc}}$ class mapping. Further we prove that the class of bilipschitz mappings belonging to $W^{m,p}_{\mathrm{loc}}(\mathbb{R}^n, \mathbb{R}^n)$ is closed with respect to composition and multiplication without any restrictions on $m, p \geq 1$. These results can be easily extended to smooth n-dimensional Riemannian manifolds and further we prove a form of the implicit function theorem for Sobolev mappings.

Keywords. Bilipschitz mappings, Inverse mapping theorem

Mathematics Subject Classification (2010). Primary 46E35, secondary 26B10

1. Introduction

It is well known that for a mapping of $\mathbb{R}^n$ to $\mathbb{R}^n$, which is of class $\mathcal{C}^m$ and has positive Jacobian at some point x we can find neighbourhoods of x and $f(x)$ such that the restriction of f is a $\mathcal{C}^m$ class diffeomorphism of the two neighbourhoods. Building on [7] and [9], we ask if this result can be extended to classes of Sobolev mappings.

Since the seminal paper of Arnold [2] a variety of techniques have been applied to hydrodynamics and other partial differential equations using certain properties of the spaces of Sobolev diffeomorphisms on smooth Riemannian manifolds. We refer the reader to [9] for more detailed motivation and applications.

In [9] the authors considered the regularity of the inverse of a Sobolev diffeomorphism and the composition of a Sobolev mapping in $W^{s+r,2}(\mathbb{R}^n, \mathbb{R}^n)$ with a mapping pertaining to a certain class of $W^{s,2}_{\mathrm{loc}}$ Sobolev $\mathcal{C}^1$ diffeomorphisms, $s \in \mathbb{R}$. Our result is a generalization of the above, in that we remove the condition $s > n/2+1$ completely and consider $p \in [1, \infty]$ arbitrary. We also relax

D. Campbell: Department of Mathematics, University of Hradec Králové, Rokitanského 62, 500 03 Hradec Králové, Czech Republic; daniel.campbell@uhk.cz
S. Hencl, F. Konopecký: Department of Mathematical Analysis, Charles University, Sokolovská 83, 186 00 Prague 8, Czech Republic; hencl@karlin.mff.cuni.cz; frakon@gmail.com

the condition that f needs to be a diffeomorphism and require only bilipschitz regularity. Our result is also an extension of that proved in [7, Theorem 1.3], where it was assumed that $m = 2$. Our main result is as follows.

Theorem 1.1. *Let $\Omega, \Omega' \subset \mathbb{R}^n$ be open. Let $m \in \mathbb{N}$, $p \in [1, \infty]$ and $f : \Omega \to \Omega' = f(\Omega)$ be such that*

$$f \in W^{m,p}_{\mathrm{loc}}(\Omega, \mathbb{R}^n) \cap \mathrm{Bilip}_{\mathrm{loc}}(\Omega, \mathbb{R}^n).$$

Then

$$f^{-1} \in W^{m,p}_{\mathrm{loc}}(\Omega', \mathbb{R}^n).$$

Further, if $D^m f \in \mathrm{BV}_{\mathrm{loc}}(\Omega, \mathbb{R}^{n^{m+1}})$ then $D^m f^{-1} \in \mathrm{BV}_{\mathrm{loc}}(\Omega', \mathbb{R}^{n^{m+1}})$.

The counterexamples given in [7] (see also Remark 3.5) show that the bilipschitz condition is vital to guarantee that the inverse is Sobolev or BV, i.e. the result may fail if f is not Lipschitz and it may fail if f^{-1} is not Lipschitz. For the definition of bilipschitz mappings, $\mathrm{Bilip}_{\mathrm{loc}}(\Omega, \mathbb{R}^n)$, see the preliminaries. It is not difficult to check that the Sobolev imbedding theorem gives that if $r \in \mathbb{N}$ is such that $r < m - \frac{n}{p}$ then f and f^{-1} are $\mathcal{C}^r$ mappings.

In [9] the authors pose the question whether the composition of two diffeomorphisms in $W^{s,2}(M, M)$, $f \circ g$, is a $W^{s,2}(M, M)$ map, $s \in \mathbb{N}, s > \frac{n}{2}$. We prove this is true for $W^{m,p}_{\mathrm{loc}}(\mathbb{R}^n, \mathbb{R}^n)$ mappings if the interior function is bilipschitz and the exterior Lipschitz. Our result places no constraint on m, p other than $m \in \mathbb{N}$ and $p \in [1, \infty]$. Note that previous results about the regularity of the composition (see e.g. [3] and [12]) usually assume that all lower order derivatives of f are bounded which is not necessary if we moreover assume that g is bilipschitz. The results from [9] are extended also for fractional order Sobolev spaces. We have not pursued this direction but we don't see any obstacles in doing so. Our result is as follows.

Theorem 1.2. *Let $\Omega, \Omega' \subset \mathbb{R}^n$ be open. Let $m \in \mathbb{N}$, $p \in [1, \infty]$ and $g : \Omega \to \Omega' = f(\Omega)$ be such that*

$$g \in W^{m,p}_{\mathrm{loc}}(\Omega, \mathbb{R}^n) \cap \mathrm{Bilip}_{\mathrm{loc}}(\Omega, \mathbb{R}^n), \;\; f \in W^{m,p}_{\mathrm{loc}}(\Omega', \mathbb{R}^n) \cap W^{1,\infty}(\Omega', \mathbb{R}^n).$$

Then

$$f \circ g \in W^{m,p}_{\mathrm{loc}}(\Omega, \mathbb{R}^n).$$

Again this result may fail if g^{-1} is not Lipschitz and it may fail if f is not Lipschitz even for bilipschitz g - see Remark 3.4. We also prove the following result about the product of two functions in the class considered above.

Theorem 1.3. *Let $\Omega, \subset \mathbb{R}^n$ be open. Let $m \in \mathbb{N}$, $p \in [1, \infty]$ and*

$$f, g \in W^{m,p}_{\mathrm{loc}}(\Omega) \cap \mathrm{Lip}_{\mathrm{loc}}(\Omega).$$

Then

$$fg \in W^{m,p}_{\mathrm{loc}}(\Omega) \cap \mathrm{Lip}_{\mathrm{loc}}(\Omega).$$

It is not difficult to check that the Sobolev imbedding theorem gives that if $r \in \mathbb{N}$ is such that $r < m - \frac{n}{p}$ then $g \circ f$ and fg are $\mathcal{C}^r$ mappings.

Let us give the brief idea of our proof of the main result Theorem 1.1. By twice differentiation $D^2(f \circ f^{-1}) = D^2(id) = 0$ we obtain the identity

$$D^2 f^{-1}(y) = -Df^{-1}(y) D^2 f(f^{-1}(y)) Df^{-1}(y) Df^{-1}(y),$$

see Preliminaries for the interpretation of the higher order derivatives. Using the Leibnitz rule and the chain rule we derive this identity further and we express $D^k f^{-1}$ as a product of lower order derivatives of f and f^{-1}. We estimate the integrability of the lower order terms using induction and the famous Gagliardo-Nirenberg interpolation inequality. Simple use of the Hölder's inequality gives our final claim. Let us note that the simple use of the Sobolev embedding theorem would not be sufficient to prove the claim without some extra assumption on the lower order derivatives. Fortunately the Gagliardo-Nirenberg interpolation inequality gives us a better integrability of the lower order terms which gives us exactly the desired result.

It is not difficult to show that similar results hold on smooth n-dimensional Riemannian manifolds. In Sections 4 we show that it is enough to apply our Euclidean result for the composition with reference maps. Finally in Section 5 we prove a variant of the Implicit Mapping Theorem for Sobolev mappings using our Inverse Mapping Theorem 1.1.

2. Preliminaries

2.1. Results on Sobolev Functions. We start by defining locally bilipschitz mappings.

Definition 2.1. Let $\Omega \subset \mathbb{R}^n$ be open and $u : \Omega \to \mathbb{R}^d$. The space $\mathrm{Bilip}_{\mathrm{loc}}(\Omega, \mathbb{R}^d)$ is the class of mappings $u : \Omega \to \mathbb{R}^d$ such that for all $x_0 \in \Omega$ there exists some $\delta > 0$ and $C_1, C_2 > 0$, such that for all $x, x' \in \Omega \cap B(x_0, \delta)$ it holds that

$$C_1 |x - x'| < |u(x) - u(x')| < C_2 |x - x'|.$$

For the following Theorem see [1, Theorem 3.16 and Corollary 3.19]:

Theorem 2.2. *Let $\Omega, \Omega' \subset \mathbb{R}^n$ be open. Let $u : \Omega \to \mathbb{R}^d$. Suppose that $F : \Omega \to \Omega' = F(\Omega)$ is Lipschitz and a homeomorphism.*

1. *If $u \in BV_{\mathrm{loc}}(\Omega, \mathbb{R}^d)$ then $u \circ F^{-1} \in BV_{\mathrm{loc}}(\Omega', \mathbb{R}^d)$.*
2. *If $u \in W^{1,1}_{\mathrm{loc}}(\Omega, \mathbb{R}^d)$ and F^{-1} is Lipschitz, then $u \circ F^{-1} \in W^{1,1}_{\mathrm{loc}}(\Omega', \mathbb{R}^d)$ and*

$$Du \circ F^{-1}(y) = Du(F^{-1}(y)) DF^{-1}(y) \quad \textit{for almost all } y \in \Omega'.$$

We will refer to the following lemma as the product rule or the Leibnitz rule.

Lemma 2.3. *Let $\Omega \subset \mathbb{R}^n$ be open and $f, g \in W^{1,1}_{\mathrm{loc}}(\Omega, \mathbb{R}^d)$. Suppose that*

$$fDg, gDf \in L^1_{\mathrm{loc}}(\Omega, \mathbb{R}^d).$$

Then

$$D(f(x)g(x)) = g(x)Df(x) + f(x)Dg(x) \quad \text{for almost all } x \in \Omega.$$

Similarly we can prove the formula for differentiation of three or more terms under the assumption that all summands are integrable.

Proof. By considering the component functions we may assume that $d = 1$. We approximate the functions by truncation. Our truncated functions belong to $W^{1,1}_{\mathrm{loc}}(\Omega) \cap L^\infty(\Omega)$ and here we can use the standard convolution approximations to get our result for the truncated functions. The function $C(|gDf| + |fDg|)$ is integrable and dominates the derivative of our truncation approximations. This however gives us the desired equality almost everywhere. $\square$

The following theorem [7, Theorem 1.3] will be a start for our induction process.

Theorem 2.4. *Let $\Omega, \Omega' \subset \mathbb{R}^n$ be open, $p \geq 1$ and suppose that $f : \Omega \to \Omega'$ is a bilipschitz mapping. If $Df \in W^{1,p}_{\mathrm{loc}}(\Omega, \mathbb{R}^{n^2})$, then $Df^{-1} \in W^{1,p}_{\mathrm{loc}}(\Omega', \mathbb{R}^{n^2})$.*

The Sobolev Embedding Theorem is well known.

Theorem 2.5. *Let $\Omega \subset \mathbb{R}^n$ be open and have a Lipschitz boundary. Further let $p \in [1, n)$ and $f \in W^{1,p}(\Omega)$. Then $f \in L^{\frac{np}{n-p}}(\Omega)$ and moreover*

$$\|f\|_{L^{\frac{np}{n-p}}(\Omega)} \leq c\|f\|_{W^{1,p}(\Omega)}.$$

Further if $f \in W^{k,p}(\Omega)$ and $kp < n$ then for every $i \in \{0, , \dots, k-1\}$ we have $f \in W^{i,p_i}(\Omega)$ where $p_i = \frac{np}{n-(k-i)p}$. If $f \in W^{1,p}(\Omega)$ for some $p > n$ then $f \in L^\infty(\Omega)$. This also holds for mappings with values in $\mathbb{R}^d$.

The following theorem is referred to as the Gagliardo-Nirenberg interpolation inequality and is a result of [13].

Theorem 2.6. *Let $u : \mathbb{R}^n \to \mathbb{R}$, $q, r \in [1, \infty]$ and $\hat{k} \in \mathbb{N}$. Further let $j \in \{1, \dots, \hat{k}\}$, $\hat{p} \in [1, \infty)$ and $\alpha \in [\frac{j}{\hat{k}}, 1]$ be such that*

$$\frac{1}{\hat{p}} = \frac{j}{n} + \alpha\Big(\frac{1}{r} - \frac{\hat{k}}{n}\Big) + \frac{1 - \alpha}{q}.$$

Then $u \in L^q(\mathbb{R}^n)$ and $|D^{\hat{k}}u| \in L^r(\mathbb{R}^n)$ implies that $|D^j u| \in L^{\hat{p}}(\mathbb{R}^n)$ and this embedding is continuous.

Let us now prove that this result can be extended as follows.

Theorem 2.7. *Let $u : \mathbb{R}^n \to \mathbb{R}$ and $q \in [1, \infty]$ and $\hat{k} \in \mathbb{N}$. Further let $j \in \{1, \ldots, \hat{k} - 1\}$, $\hat{p} \in [1, \infty)$ and $\alpha \in [\frac{j}{\hat{k}}, 1]$ be such that*

$$\frac{1}{\hat{p}} = \frac{j}{n} + \alpha\Big(1 - \frac{\hat{k}}{n}\Big) + \frac{1-\alpha}{q}.$$

Then $u \in L^q(\mathbb{R}^n)$ and $D^{\hat{k}-1}u \in \mathrm{BV}(\mathbb{R}^n)$ implies that $|D^j u| \in L^{\hat{p}}(\mathbb{R}^n)$.

Sketch of the proof. For f_k, the convolution approximations of f, it holds that $f_k \in W^{m,1}(\mathbb{R}^n)$ and we may apply Theorem 2.6. Since $||f_k||_q \leq ||f||_q$ and $||D^m f_k||_1 \leq ||D^m f||_1$ (where $||D^m f||_1$ signifies the total variation) we have that $D^j f_k$ is a bounded sequence and by moving if necessary to a subsequence we find for every $(a_1, \ldots, a_j) \in \{1, \ldots, n\}^j$ a $w_{a_j,\ldots,a_1} \in L^{\hat{p}_j}(\mathbb{R}^n)$ such that for any $\varphi \in \mathcal{C}_c^\infty(\mathbb{R}^n)$ we have

$$\int_{\mathbb{R}^n} D_{a_j,\ldots,a_1} f_k \varphi \to \int_{\mathbb{R}^n} w_{a_j,\ldots,a_1}\varphi$$

and clearly

$$\int_{\mathbb{R}^n} D_{a_j,\ldots,a_1} f_k \varphi = (-1)^j \int_{\mathbb{R}^n} f_k D_{a_j,\ldots,a_1}\varphi \to (-1)^j \int_{\mathbb{R}^n} f D_{a_j,\ldots,a_1}\varphi.$$

This however implies that $D_{a_j,\ldots,a_1} f = w_{a_j,\ldots,a_1} \in L^{\hat{p}_j}(\mathbb{R}^n)$. $\square$

We will use the so called ACL classification of Sobolev mappings.

Theorem 2.8. *Let $\Omega \subset \mathbb{R}^n$ be open and let $f \in W^{1,p}(\Omega)$ then there exists a representative $\hat{f}$ of f such that $\hat{f}$ is absolutely continuous on almost all lines parallel to each of the coordinate axes. Further the classical partial derivatives of $\hat{f}$ equal the weak derivatives of f almost everywhere.*

It is not difficult to use the previous theorem recurrently to see that if $f \in W^{k,p}(\Omega)$ then there exists a representative of f such that for all $1 \leq j \leq k-1$, the derivative $D^j f$ is absolutely continuous on almost all lines parallel to the coordinate axis. Further the classical partial derivatives of our representative equal the weak derivatives of f almost everywhere.

2.2. Representation of Higher Order Derivatives. During the course of this work we will need to represent and work with higher order derivatives. We refer the reader to [11] for more information about multi-linear algebra which could be used to represent our higher-order derivatives.

To shorten our notation D_i, $i \in \{1, 2, \dots, n\}$, will denote the weak derivative in the direction of the i-th canonic basis vector. Given a finite sequence $a_1, a_2, \dots, a_m \in \{1, 2, \dots, n\}$ we define the symbol

$$D_{a_m,\dots,a_2,a_1} f(x) = D_{a_m}\Big(\dots \big(D_{a_2}\big(D_{a_1} f(x)\big)\big) \dots \Big).$$

We will use the following notation for the components of a mapping $f : \mathbb{R}^n \to \mathbb{R}^n$, $f(x) = (f^1(x), f^2(x), \dots, f^n(x))$. This notation will be useful as in the chain rule we will always sum over indices, one of which is a superscript and the other a subscript. This corresponds to common practice in tensor notation.

Given some $f \in W^{m,1}_{\text{loc}}(\Omega, \mathbb{R}^n)$ we can define the symbol $D^m f$ as the mapping from $\{1, 2, \dots, n\}^m \to L^1(\Omega, \mathbb{R}^n)$ such that technically speaking for $a_i \in \{1, 2, \dots, n\}$ we have

$$D^m f(a_1, \dots, a_m) = D_{a_m,\dots,a_1} f = D_{a_m}(\dots D_{a_2}(D_{a_1} f)) \dots),$$

but we will write

$$D^m f = \Big(D_{a_m,\dots,a_1} f\Big)_{a_1,\dots,a_m \in \{1,2,\dots,n\}}.$$

It therefore follows that for $f \in W^{m,1}_{\text{loc}}(\Omega, \mathbb{R}^n)$ we can identify $D^m f$ with an element of the set $L^1_{\text{loc}}(\Omega, \mathbb{R}^{n^{m+1}})$. It is an easy result of the definition of the weak derivative that the order of partial derivatives is interchangeable, i.e.

$$D_{a_m,\dots,a_1} f = D_{a_{\pi(m)},\dots,a_{\pi(1)}} f$$

for any $\pi \in S_m$ the symmetric group. It suffices to take the definition of the weak derivatives, where in the integral we can change the order of derivatives on the test function as it is smooth. Thus our derivatives $D^j f$ are symmetric.

We shall now expound the iterated chain rule, which is a result of repeatedly using the chain rule and the products rule, for a pair of smooth mappings f and g. We will later prove that the same holds for Sobolev mappings given g is bilipschitz.

Let f and g be smooth. Clearly

$$D(f \circ g)(x) = Df(g(x))Dg(x)$$

where the multiplication above is standard matrix multiplication. Now apply D again and use the products rule. We get

$$D_{j,i}(f \circ g)(x) = \sum_{k,l=1}^{n} D_{k,l} f(g(x))\big(D_i g(x)\big)^k \big(D_j g(x)\big)^l + \sum_{l=1}^{n} D_l f(x)\big(D_{j,i} g(x)\big)^l.$$

This can be symbolically rewritten as

$$D^2(f \circ g)(x) = D^2 f(g(x))Dg(x)Dg(x) + Df(x)D^2 g(x), \tag{1}$$

where we expand the concept of matrix multiplication as follows: in the first term we sum over $\{1,2,\dots,n\}^2$, all second partials of f and the components of the two Dg terms and over $\{1,2,\dots,n\}$, the first partials of f and the components of D^2g in the second term.

Notice that if we make the following definition of $\mathcal{K}_m$ with $m \in \mathbb{N}$

$$\mathcal{K}_m = \Big\{(k_0,\dots,k_m) \in \{0,1,\dots,m\}^{m+1} : (k_0 \geq 1) \ \& \\ (k_i = 0 \Leftrightarrow i > k_0) \ \& \ \big(\sum_{i=1}^{m} k_i = m\big)\Big\}$$

we may write the above as follows

$$D^2(f \circ g)(x) = \sum_{k \in \mathcal{K}_2} D^{k_0} f(g(x)) \prod_{i=1}^{k_0} D^{k_i} g(x).$$

We will now prove the following equality for all $m \in \mathbb{N}$ by induction

$$D^m\big(f \circ g\big)(x) = \sum_{k \in \mathcal{K}_m, \chi \in X_k} D^{k_0} f(g(x)) \prod_{i=1}^{k_0} D^{k_i} g(x), \tag{2}$$

where we define X_k below. Here k_0 corresponds to the order of derivative of f and other k_i correspond to the derivatives of g. Note that the definition of $\mathcal{K}_m$ gives us $\sum_{i=1}^m k_i = m$ since in each step we differentiate some term with g once more and also $k_i = 0, i > k_0$ since after deriving f k_0 times we have at most k_0 terms with g. In fact, for given numbers $\{k_i\}_{i=1}^m$ there is a number of permissible permutations (orderings in which the derivatives may be applied), which we will denote as X_k for some $k \in \mathcal{K}_m$. This corresponds to the fact that each term $D^{k_0} f(g(x)) \prod_{i=1}^{k_0} D^{k_i} g(x)$ is multiplied by some fixed natural number $\#X_k$.

It is no problem for us to work with the identity (2) and with its interpretation. In our proof we will deduce the integrability of each term in the product $D^{k_0} f(g(x)) \prod_{i=1}^{k_0} D^{k_i} g(x)$ and then we will apply the Hölder inequality. Therefore we do not need to care which component is multiplied with which component of the next term. We have finitely many terms and each of them can be estimated in the same way.

For our component-wise notation we will need to take a finite sequence in $\{1,2,\dots,n\}$ and apply each corresponding derivative to one of our factors. We define

$$X_k = \Big\{\chi : \{1,\dots,m\} \to \{1,\dots,k_0\};\ \#\{j : \chi(j) = i\} = k_i \\ \min\{j \leq m : \chi(j) = i+1\} > \min\{j \leq m : \chi(j) = i\} : 1 \leq i \leq m-1\Big\}.$$

We carry on to define b^i_j by denoting

$$b^i_j = a_l,$$

where l is the j-th index which is mapped to i by χ i.e. $\chi(l) = i$ and $\#\{\chi(o) = i : o < l\} = j - 1$. Assume that the following holds for all $m < \hat{m}$

$$D_{a_m,\dots,a_1}(f \circ g)(x) = \sum_{k \in \mathcal{K}_m} \sum_{\chi \in X_k} \sum_{b^0_1,\dots,b^0_{k_0}=1}^{n} D_{b^0_{k_0},\dots,b^0_1} f(g(x)) \prod_{i=1}^{k_0} (D_{b^i_{k_i},\dots,b^i_1} g(x))^{b^0_i}. \tag{3}$$

Let us now differentiate the equation above, where $m = \hat{m}-1$, by applying $D_{a_{\hat{m}}}$. We apply the Leibnitz rule (product rule) repeatedly and when differentiating $D_{b^0_{k_0},\dots,b^0_1} f(g(x))$ we also use the chain rule and get

$$\begin{aligned} &D_{a_{\hat{m}},\dots,a_1}(f \circ g)(x) = \\ &\sum_{\hat{j}=1}^{n} \sum_{k \in \mathcal{K}_{\hat{m}}} \sum_{\chi \in X_k} \sum_{b^0_1,\dots,b^0_{k_0}=1}^{n} D_{\hat{j},b^0_{k_0},\dots,b^0_1} f(g(x)) \prod_{i=1}^{k_0} (D_{b^i_{k_i},\dots,b^i_1} g(x))^{b^0_i} \cdot (D_{a_{\hat{m}}} g(x))^{\hat{j}} + \\ &\sum_{\hat{i}=1}^{k_0} \sum_{k \in \mathcal{K}_{\hat{m}}} \sum_{\chi \in X_k} \sum_{b^0_1,\dots,b^0_{k_0}=1}^{n} D_{b^0_{k_0},\dots,b^0_1} f(g(x)) \prod_{i \in \{1,\dots k_0\} \setminus \{\hat{i}\}} (D_{b^i_{k_i},\dots,b^i_1} g(x))^{b^0_i} \cdot (D_{a_{\hat{m}} b^{\hat{i}}_{k_{\hat{i}}},\dots,b^{\hat{i}}_1} g(x))^{b^0_{\hat{i}}}. \end{aligned} \tag{4}$$

We can express this equation using the same notation as in (3). To prove this consider two terms in the above sums, firstly where $k_0 = s$, secondly where $k_0 = s + 1$ and add the first line of (4) for $k_0 = s$ and the second line of (4) for $k_0 = s + 1$. Putting $M = \{k \in \mathcal{K}_{\hat{m}} : k_0 = s + 1\}$ we get

$$\sum_{k \in M} \sum_{\chi \in X_k} \sum_{b^0_1,\dots,b^0_{k_0}=1}^{n} D_{b^0_{k_0},\dots,b^0_1} f(g(x)) \prod_{i=1}^{k_0} (D_{b^i_{k_i},\dots,b^i_1} g(x))^{b^0_i}.$$

Summing this over s gives our desired result.

It will be very useful to be able to express the above in some more concise way, therefore we will introduce the following convention. We will shorten

$$D_{a_m,\dots,a_1}(f \circ g)(x) = \sum_{k \in \mathcal{K}_m} \sum_{\chi \in X_k} \sum_{b^0_1,\dots,b^0_{k_0}=1}^{n} D_{b^0_{k_0},\dots,b^0_1} f(g(x)) \prod_{i=1}^{k_0} (D_{b^i_{k_i},\dots,b^i_1} g(x))^{b^0_i} \tag{5}$$

by realizing that the order of the derivative of f (i.e. k_0) is the same as the number of Dg factors and we sum over the derivative indices of f and component indices of Dg-type terms as above, writing only

$$D^m f \circ g(x) = \sum_{k \in \mathcal{K}_m, \chi \in X_k} D^{k_0} f(g(x)) \prod_{i=1}^{k_0} D^{k_i} g(x) \tag{6}$$

where the previous equalities hold at all points x where $f \circ g$ is defined. As previously mentioned we will extend this result for bilipschitz Sobolev mappings in sections 3 and 4.

3. Regularity of the Inverse

Recall that if we have some mapping $h : \mathbb{R}^n \to \mathbb{R}^l$, we denote the components of h as $h(x) = (h^1(x), h^2(x), \dots, h^l(x))$.

Lemma 3.1. *Let $\Omega, \Omega' \subset \mathbb{R}^n$ be open. Let $f : \Omega \to \Omega' = f(\Omega)$ be a bilipschitz mapping such that $f \in W^{2,1}_{\mathrm{loc}}(\Omega, \mathbb{R}^n)$. Then $f^{-1} \in W^{2,1}_{\mathrm{loc}}(\Omega', \mathbb{R}^n)$ and*

$$\begin{aligned}\Big(D_{j,i} f^{-1}(y)\Big)^k &= \\ &= -\sum_{l=1}^{n} \sum_{l_1,l_2=1}^{n} \Big(D_l f^{-1}(y)\Big)^k \Big(D_{l_2,l_1} f(f^{-1}(y))\Big)^l \Big(D_i f^{-1}(y)\Big)^{l_1} \Big(D_j f^{-1}(y)\Big)^{l_2},\end{aligned} \tag{7}$$

which may also be written as

$$D^2 f^{-1}(y) = -Df^{-1}(y) D^2 f(f^{-1}(y)) Df^{-1}(y) Df^{-1}(y) \tag{8}$$

for almost all $y \in \Omega'$.

Remark 3.2. In the above lemma, the left-most factor on the right hand side of (8) is the matrix $Df^{-1}(y)$. The reason for this is that it then corresponds to standard composition (multiplication) of matrices. Notice however that in (7) the order of the factors is irrelevant given we sum the correct upper and lower indices.

Proof of Lemma 3.1. First let us note that $f^{-1} \in W^{2,1}_{\mathrm{loc}}(\Omega', \mathbb{R}^n)$ is a direct result of Theorem 2.4. Clearly we have almost everywhere that

$$0 = D(I) = D^2(id) = D^2(f \circ f^{-1}) = D\big(D(f \circ f^{-1})\big)$$

where I is the $(n \times n)$ identity matrix and id the identity mapping on $\mathbb{R}^n$. Now we can use Theorem 2.2 on $f \circ f^{-1}$ as f is bilipschitz. Therefore we get

$$0 = D_j \Bigg(\sum_{l_1=1}^{n} D_{l_1} f(f^{-1}(y)) \big(D_i f^{-1}(y)\big)^{l_1} \Bigg)$$

for all $i, j \in \{1, 2, \dots, n\}$ and for almost all $y \in \Omega'$. As $Df(f^{-1}), Df^{-1} \in L^\infty(\Omega')$ and f, f^{-1} are of Sobolev type $W^{2,1}_{\mathrm{loc}}$ it is easy to see that we will have no trouble in applying the product rule, Lemma 2.3. Deriving again, using the product rule and then Theorem 2.2 on $Df(f^{-1})$, we get,

$$0=\sum_{l_1,l_2=1}^{n} D_{l_2,l_1}f(f^{-1}(y))\big(D_i f^{-1}(y)\big)^{l_1}\big(D_j f^{-1}(y)\big)^{l_2}+ \\ +\sum_{l=1}^{n} D_l f(f^{-1}(y))\Big(D_{j,i}f^{-1}(y)\Big)^{l}.$$

Notice that thanks to the Lipschitz qualities of f and f^{-1} we have $Df^{-1}(y)=\big(Df(f^{-1}(y))\big)^{-1}$ almost everywhere in terms of the inverse of matrices. We apply $Df^{-1}(y)$ (matrix multiplication from the left) to get

$$0=\sum_{l=1}^{n}\sum_{l_1,l_2=1}^{n}\Big(D_l f^{-1}(y)\Big)^{k}\Big(D_{l_2,l_1}f(f^{-1}(y))\Big)^{l}\Big(D_i f^{-1}(y)\Big)^{l_1}\Big(D_j f^{-1}(y)\Big)^{l_2}+ \\ +\Big(D_{j,i}f^{-1}(y)\Big)^{k}$$

for all components $k\in\{1,2,\dots,n\}$ and for almost all $y\in\Omega'$. □

Lemma 3.3. *Let $\Omega,\Omega'\subset\mathbb{R}^n$ be open. Let $f:\Omega\to\Omega'=f(\Omega)$ be a bilipschitz mapping such that $f\in W^{k,p}_{\mathrm{loc}}(\Omega,\mathbb{R}^n)$. Then*

$$\big|D\big(D^{k-1}f(f^{-1})\big)\big|\in L^p_{\mathrm{loc}}(\Omega')$$

and for all k-tuples $a_1,a_2,\dots a_k\in\{1,2,\dots,n\}$ we have

$$D_{a_k}(D_{a_{k-1},\dots,a_1}f(f^{-1}(y)))=\sum_{l=1}^{n} D_{l,a_{k-1},\dots,a_1}f(f^{-1}(y))\big(D_{a_k}f^{-1}(y)\big)^{l}.$$

Proof. Clearly $D_{a_{k-1},\dots,a_1}f\in W^{1,p}_{\mathrm{loc}}(\Omega,\mathbb{R}^n)$ and therefore belongs also to the Sobolev space $W^{1,1}_{\mathrm{loc}}(\Omega,\mathbb{R}^n)$. Since f^{-1} is bilipschitz, Theorem 2.2 implies that the weak derivative of $D_{a_{k-1},\dots,a_1}f(f^{-1}(\cdot))$ exists and since as f^{-1} is a bilipschitz change of variables we have $D_{a_{k-1},\dots,a_1}f(f^{-1}(\cdot))\in W^{1,p}_{\mathrm{loc}}(\Omega',\mathbb{R}^n)$. From Theorem 2.2 we get

$$D_{a_k}(D_{a_{k-1},\dots,a_1}f(f^{-1}(y)))=\sum_{l=1}^{n} D_l D_{a_{k-1},\dots,a_1}f(f^{-1}(y))(D_{a_k}f^{-1}(y))^{l} \tag{9}$$

for almost all $y\in\Omega'$. Since $|Df^{-1}|\in L^\infty(\Omega')$ we get

$$\big|D\big(D^{k-1}f(f^{-1}(\cdot))\big)\big|\in L^p_{\mathrm{loc}}(\Omega').$$

By applying (9) to all $(a_1,\dots a_{k-1})\in\{1,2,\dots,n\}^{k-1}$ we get that by our convention, described between (5) and (6), that we may write

$$D(D^{k-1}f(f^{-1}(y)))=D^k f(f^{-1}(y))Df^{-1}(y)\quad\text{for almost all } y\in\Omega'.$$

□

Proof of Theorem 1.1. Consider the case $m = 1$. Our claim is trivial as f is bilipschitz. For the case $m = 2$ Theorem 2.4 gives our result. Our proof will continue by induction. We continue to prove the case $m = 3$ explicitly to aid the comprehension of the reader.

Firstly, we know by Lemma 3.1 that (7) holds. We differentiate (7) according to Theorem 2.2, Lemma 3.3 and Lemma 2.3, initially formally, and will afterward verify the assumptions. Thus,

$$\begin{aligned}
&D_{a_3,a_2,a_1} f^{-1}(y) = \\
&= -D_{a_3}\Bigg(\sum_{l_0,l_1,l_2=1}^{n} D_{l_0} f^{-1}(y) \big(D_{l_2,l_1} f(f^{-1}(y))\big)^{l_0} \big(D_{a_1} f^{-1}(y)\big)^{l_1} \big(D_{a_2} f^{-1}(y)\big)^{l_2} \Bigg) \\
&= \sum_{l_0,l_1,l_2=1}^{n} D_{a_3,l_0} f^{-1}(y) \big(D_{l_2,l_1} f(f^{-1}(y))\big)^{l_0} \big(D_{a_1} f^{-1}(y)\big)^{l_1} \big(D_{a_2} f^{-1}(y)\big)^{l_2} \\
&+ \sum_{l_0,l_1,l_2,l_3=1}^{n} D_{l_0} f^{-1}(y) \big(D_{l_3,l_2,l_1} f(f^{-1}(y))\big)^{l_0} \cdot \\
&\qquad\qquad \cdot \big(D_{a_1} f^{-1}(y)\big)^{l_1} \big(D_{a_2} f^{-1}(y)\big)^{l_2} \big(D_{a_3} f^{-1}(y)\big)^{l_3} \\
&+ \sum_{l_0,l_1,l_2=1}^{n} D_{l_0} f^{-1}(y) \big(D_{l_2,l_1} f(f^{-1}(y))\big)^{l_0} \big(D_{a_3,a_1} f^{-1}(y)\big)^{l_1} \big(D_{a_2} f^{-1}(y)\big)^{l_2} \\
&+ \sum_{l_0,l_1,l_2=1}^{n} D_{l_0} f^{-1}(y) \big(D_{l_2,l_1} f(f^{-1}(y))\big)^{l_0} \big(D_{a_1} f^{-1}(y)\big)^{l_1} \big(D_{a_3,a_2} f^{-1}(y)\big)^{l_2},
\end{aligned}$$

which can be summarized, according to our convention, by omitting indices as follows

$$\begin{aligned}
D^3 f^{-1}(y) =& D^2 f^{-1}(y) D^2 f(f^{-1}(y)) D f^{-1}(y) D f^{-1}(y) \\
&+ D f^{-1}(y) D^3 f(f^{-1}(y)) D f^{-1}(y) D f^{-1}(y) D f^{-1}(y) \\
&+ D f^{-1}(y) D^2 f(f^{-1}(y)) D^2 f^{-1}(y) D f^{-1}(y) \\
&+ D f^{-1}(y) D^2 f(f^{-1}(y)) D f^{-1}(y) D^2 f^{-1}(y).
\end{aligned}$$

We want to prove that the norms of the objects on the right hand side are $L^p_{\mathrm{loc}}(\Omega')$ functions. The second term is trivial as the point-wise norms $|Df^{-1}(y)|$ are uniformly bounded almost everywhere (remember $f^{-1} \in W^{1,\infty}(\Omega', \mathbb{R}^n)$) and by our hypothesis $|D^3 f| \in L^p_{\mathrm{loc}}(\Omega)$ and the f^{-1}-bilipschitz change of variables does not effect this.

We know by Theorem 2.4 that if $|D^2 f| \in L^q_{\mathrm{loc}}(\Omega)$ then $|D^2 f^{-1}| \in L^q_{\mathrm{loc}}(\Omega')$. The first and the last two terms are essentially the same. We have two bounded factors of $|Df^{-1}|$ and two factors with the same integrability as $|D^2 f|$. We now apply Theorem 2.6 by choosing $i \in \{1, 2, \dots, n\}$ and taking $u = D_i f$, $r = p$, $q = \infty$, $\hat{k} = 2$, $j = 1$, $\alpha = \frac{1}{2}$ and get

$$|D^2 f| \in L^{2p}_{\text{loc}}(\Omega).$$

Hereby we get that each of the three terms in question are in $L^p_{\text{loc}}(\Omega')$ because clearly $\frac{1}{1/(2p)+1/(2p)} = p$. This both proves that we can use the chain and product rule (see Lemma 2.3) and thus $D^3 f^{-1} \in L^p_{\text{loc}}(\Omega')$, which was our claim.

We now continue to the induction step, where we consider $m \geq 4$ and assume that $|D^k f| \in L^q_{\text{loc}}(\Omega)$ implies $|D^k f^{-1}| \in L^q_{\text{loc}}(\Omega')$ for $1 \leq k \leq m-1$ and any $q \in [1, \infty]$. We take the equation (7) and differentiate it $(m-2)$-times, again initially formally and verifying the hypothesis later. The differences here compared to the preliminaries $D^m(f \circ g)$ derivatives are as follows. Instead of equation (1) we start with equation (8) and we proceed similarly with the help of Theorem 2.2. We have an extra Df^{-1} factor, whose derivatives are summed with the components of the Df factor, and therefore we adjust the set $\mathcal{K}_m$ to the set $\mathcal{K}'_m$ as follows

$$\begin{aligned}\mathcal{K}'_m = \Big\{ &(k_0, \dots, k_{m+1}) \in \{0, 1, \dots, m\}^{m+2} : \\ &(k_0 \geq 2) \ \&\ (k_i = 0 \Leftrightarrow i > k_0 + 1) \ \&\ \big(\sum_{i=1}^{m+1} k_i = m+1\big) \Big\}.\end{aligned}$$

We take the extra Df^{-1} factor to correspond to the index 1. The set of orderings Y_k will also differ slightly from X_k as follows,

$$\begin{aligned} Y_k \ni \chi : \{1, \dots, m\} &\to \{1, \dots, k_0 + 1\} \\ \min\{j \leq m : \chi(j) = i+1\} &> \min\{j \leq m : \chi(j) = i\} && \text{for all } 2 \leq i \leq k_0 \\ \#\{j : \chi(j) = i\} &= k_i && \text{for all } 2 \leq i \leq k_0 + 1 \\ \#\{j : \chi(j) = 1\} &= k_1 - 1 && . \end{aligned}$$

Deriving according to Lemma 2.3 and Lemma 3.3 analogously as we did in (5) and (6) we get by Theorem 2.2

$$D^m f^{-1}(y) = \sum_{k \in \mathcal{K}'_m, \chi \in Y_k} D^{k_0} f(f^{-1}(y)) \prod_{i=1}^{k_0+1} D^{k_i} f^{-1}(y) \quad \text{for almost all } y \in \Omega'. \tag{10}$$

The calculations are almost identical to those leading up to (4) and so we omit the details.

Clearly $k_i \leq m-1$ for all $i \geq 1$ and thereby our induction hypothesis tells us that

$$|D^{k_i} f| \in L^q_{\text{loc}}(\Omega) \Rightarrow |D^{k_i} f^{-1}| \in L^q_{\text{loc}}(\Omega') \tag{11}$$

for all $1 \leq i \leq k_0 + 1$ and any $q \in [1, \infty]$. We take the norms of the derivatives from (10) and estimate their integrability. We have $(k_0 + 1)$ factors, which are in the spaces $L^{p_i}_{\text{loc}}(\Omega'), i = 0, 1, \dots, k_0$. We use (11), boundedness of the first order derivatives and Theorem 2.6 (with $u = D_o f$ for $o \in \{1, 2, \dots, n\}$, $r = p$, $j = k_i - 1$, $\hat{k} = m - 1$ and $\alpha = \frac{k_i - 1}{m-1}$) to get that the product in (10) is in the Lebesgue space $L^q_{\text{loc}}(\Omega')$ where

$$\begin{aligned} \frac{1}{q} &= \sum_{i=0}^{k_0+1} \frac{1}{p_i} \\ \frac{1}{p_i} &= \frac{k_i - 1}{n} + \frac{k_i - 1}{m - 1}\left(\frac{1}{p} - \frac{m-1}{n}\right) = \frac{k_i - 1}{p(m-1)} \\ \frac{1}{q} &= \frac{k_0 - 1}{p(m-1)} + \frac{\sum_{i=1}^{k_0+1}(k_i - 1)}{p(m-1)} \\ &= \frac{k_0 - 1}{p(m-1)} + \frac{m - k_0}{p(m-1)} = \frac{1}{p}. \end{aligned} \tag{12}$$

Hence $q = p$. This implies that the norm of the expression in (10) is in $L^p_{\text{loc}}(\Omega')$ and therefore our use of the chain rule and the products rule is correct (see Lemma 2.3 and Theorem 2.2). More significantly, it also implies that $|D^m f^{-1}| \in L^p_{\text{loc}}(\Omega')$, which was the first part of our claim.

Now let us return to BV-regularity of the inverse and assume further that $D^m f \in \text{BV}_{\text{loc}}(\Omega, \mathbb{R}^{n^{m+1}})$. Hence $|D^m f| \in L^1_{\text{loc}}(\Omega)$ and from the previous result we know that $f \in W^{m,1}_{\text{loc}}(\Omega, \mathbb{R}^n)$ implies $|D^m f^{-1}| \in L^1_{\text{loc}}(\Omega)$. Moreover, we have

$$D^m f^{-1}(y) = \sum_{k \in \mathcal{K}'_m, \chi \in Y_k} D^{k_0} f(f^{-1}(y)) \prod_{i=1}^{k_0+1} D^{k_i} f^{-1}(y) \quad \text{for almost all } y \in \Omega' \tag{13}$$

and we need to show that the right hand side is a BV function, i.e. its derivative exists and it is a measure. All the terms with $k_0 < m$ can be dealt with as in the previous part of the proof with the help of the Gagliardo-Nirenberg inequality for BV functions Theorem 2.7 and we obtain that the part of the sum with $k_0 < m$ belongs even to $W^{1,1}_{\text{loc}}(\Omega', \mathbb{R}^{n^{m+1}})$. It remains to consider the term

$$D^m f(f^{-1}(y)) \prod_{i=1}^{m+1} D^{k_i} f^{-1}(y) = D^m f(f^{-1}(y)) \left(D f^{-1}(y)\right)^{m+1} .$$

Take any $\tilde{\Omega} \subset\subset \Omega$ and corresponding $f(\tilde{\Omega}) = \tilde{\Omega}' \subset\subset \Omega'$. Let us define

$$a_l(y) := D^m f_l(f^{-1}(y)) \left(D f^{-1}(y)\right)^{m+1} ,$$

where f_l denotes the convolution approximations of f, i.e. $D^m(f_l) \to D^m(f)$ in $L^1(\tilde{\Omega})$ and hence also $D^m(f_l(f^{-1}(y))) \to D^m(f(f^{-1}(y)))$ in $L^1(\tilde{\Omega}', \mathbb{R}^{n^{m+1}})$ as f

is bilipschitz. Moreover, $D^{m+1}f_l$ (and hence also $D^{m+1}f_l(f^{-1})$) form a bounded sequence in $L^1(\tilde{\Omega}', \mathbb{R}^{n^{m+1}})$ and therefore it is not difficult to show using the chain rule that

$$Da_l \text{ is a bounded sequence in } L^1(\tilde{\Omega}', \mathbb{R}^{n^{m+1}}).$$

We select a subsequence still denoted as Da_l which converges w^* to some Radon measure μ. For all $\varphi \in \mathcal{D}(\tilde{\Omega}')$ we get by the definition of the weak derivative

$$\begin{aligned}\int_{\tilde{\Omega}'} Da_l(y)\varphi(y)\, dy &= -\int_{\tilde{\Omega}'} a_l(y) D\varphi(y)\, dy = \\ &= -\int_{\tilde{\Omega}'} D^m f_l(f^{-1}(y)) \left(Df^{-1}(y)\right)^{m+1} \ D\varphi(y)\, dy\ .\end{aligned}$$

The left-hand side converges to $\int \varphi\, d\mu$ and hence in the limit we have

$$\int_{\tilde{\Omega}'} \varphi(y)\, d\mu(y) = -\int_{\tilde{\Omega}'} D^m f(f^{-1}(y)) \left(Df^{-1}(y)\right)^{m+1} D\varphi(y)\, dy$$

since $D^m(f_l(f^{-1}(y))) \to D^m(f(f^{-1}(y)))$ in $L^1(\tilde{\Omega}', \mathbb{R}^{n^{m+1}})$. Clearly our derivative μ is defined uniquely on any open set $G \subset\subset \Omega'$, as for any two $\tilde{\Omega}_1, \tilde{\Omega}_2 \supset G$ and any $\varphi \in \mathcal{D}(G)$ we have $\int_G \varphi d\mu_1 = \int_G \varphi d\mu_2$. This shows that the remaining term (13) belongs to $\mathrm{BV}_{\mathrm{loc}}(\Omega', \mathbb{R}^{n^{m+1}})$ and finishes the proof. □

Remark 3.4. To prove the necessity of our assumptions in Theorem 1.2 let us consider the composition of two mappings $f \circ g$, with $f, g : \mathbb{R}^n \to \mathbb{R}^n$.To see that the inverse of the interior must be a Lipschitz mapping consider g to be the projection of $\mathbb{R}^n$ to $\mathbb{R}e_1$ and some f which is not measurable on $\mathbb{R}e_1$.

The next example shows that if f is not Lipschitz the composition may fail to have the original degree of integrability even for bilipschitz g. Consider $n = 9$, $p = \frac{3}{2}$, $\varepsilon \in (0, \frac{1}{3})$,

$$g(x) = x + x|x|^{1+\varepsilon} \sin(|x|^{-1}) \text{ and } f(x) = x|x|^{\varepsilon-3}$$

with $\varepsilon > 0$. Main part of the functions is of the form $\frac{x}{|x|}\varphi(|x|)$ and hence we can compute the derivative in a standard way (see e.g. [8, Lemma 2.1]). We know,

$$|D^j f(x)| \approx \frac{f(x)}{|x|^j} \leq |x|^{\varepsilon-2-j} \tag{14}$$

giving us that $f \in W^{4,p}(\Omega, \mathbb{R}^n)$. Further we may estimate the j-th derivative by

$$|D_{1,\dots,1} g(x)| \geq C\frac{|x|^{2+\varepsilon}}{|x|^{2j}} \tag{15}$$

for some C independent of x for a set with positive density at the origin (i.e. on $S := \{x : \ |\sin(|x|^{-1})| > \frac{1}{4} \text{ and } |\cos(|x|^{-1})| > \frac{1}{4}\}$). From (14) and (15) we get that $f, g \in W^{4,p}(B(0,1))$. We may calculate that

$$|D^4(f \circ g)(x)| \approx |Df(g(x))| \cdot |D^4 g(x)| \approx \frac{1}{|x|^{9-2\varepsilon}} > \frac{1}{|x|^8}$$

on a set with positive density at the origin. Hereby we see that $f \circ g \in W^{4,1}(B(0,1))$ but $D^4 f \circ g \notin L^p(B(0,1))$ as $8p > 9$.

Remark 3.5. The optimal assumptions for the first order regularity of the inverse $f^{-1} \in W^{1,q}$ usually contain some sort of the assumption about the integrability of the distortion function (see e.g. [4], [8] or [10, Chapter 6]).

One can ask if the bilipschitz assumption can be replaced by some condition on the distortion. It is evident that only very restricted results could hold. For example the radial stretching

$$f(x) = \frac{x}{|x|}|x|^\alpha$$

with non-zero $\alpha \in \mathbb{R}$ has bounded distortion, as does its inverse. Chose $n, m \in \mathbb{N}$ and put $\alpha = m + 1$. Although $f \in W^{m,\infty}(\Omega, \mathbb{R}^n)$ it can be calculated that $|D^m f^{-1}|^\beta \in L^1_{\text{loc}}(\Omega')$ only for

$$\beta < \frac{n}{m - (m+1)^{-1}}.$$

In the previous example we did not have f^{-1} Lipschitz. Nevertheless the counterexample in [7] is a homeomorphism with f^{-1} Lipschitz and the distortion function satisfies $K \in L^{cn}(\Omega)$ so even given these assumptions we can barely expect any a priori results without f bilipschitz.

4. The Algebra of Bilipschitz Sobolev Mappings and its Application to Smooth Riemannian Manifolds

First let us show that Bilipschitz Sobolev mappings form an algebra, i.e they are closed under composition, multiplication and inverse. We start with the following lemma.

Lemma 4.1. *Let $\Omega, \Omega' \subset \mathbb{R}^n$ be open. Let $g : \Omega' \to \Omega$ be such that $g \in W^{k,p}_{\text{loc}}(\Omega', \mathbb{R}^n) \cap \text{Bilip}_{\text{loc}}(\Omega', \mathbb{R}^n)$ and let $f \in W^{k,p}_{\text{loc}}(\Omega, \mathbb{R}^n)$. Then*

$$\big|D\big(D^{k-1} f(g)\big)\big| \in L^p_{\text{loc}}(\Omega')$$

and

$$D_{a_k}(D_{a_{k-1},\dots,a_1} f(g(x))) = \sum_{l=1}^{n} D_{l,a_{k-1},\dots,a_1} f(g(x))\big(D_{a_k} g(x)\big)^l$$

or more simply

$$D(D^{k-1}f(g(x))) = D^k f(g(x))Dg(x).$$

Proof. The proof is similar to that of Lemma 3.3. Instead of the assumption that f^{-1} is bilipschitz, we use that g is bilipschitz and the rest of the reasoning is the same. $\square$

Proof of Theorem 1.2. For $m = 1$ our result follows directly from Theorem 2.2 and the fact that g is bilipschitz. Now let $m = 2$ and use Lemma 4.1 to get that

$$D^2 f \circ g = D^2 f(g) DgDg + Df(g)D^2 g.$$

But $|Df(g)|$ and $|Dg|$ are both bounded almost everywhere and $|D^2 f(g)|$ and $|D^2 g|$ are both in $L^p(\Omega')$, therefore our claim holds.

Assume $|D^j f \circ g| \in L^1_{\text{loc}}(\Omega')$, for $j \leq m-1$. By Lemma 4.1, Theorem 2.2 and Lemma 2.3 repeatedly we get,

$$D^m f \circ g(y) = \sum_{k \in \mathcal{K}_m, \chi \in X_k} D^{k_0} f(g(y)) \prod_{i=1}^{k_0} D^{k_i} g(y)$$

for almost all $y \in \Omega'$, given that the expression on the right is in $L^1_{\text{loc}}(\Omega')$. This can be shown by the calculations corresponding to those in (12) to get that, $D^m(f \circ g) \in L^p(\Omega')$. $\square$

Proof of Theorem 1.3. Let $1 \leq j \leq m-1$. We calculate using Theorem 2.6 similarly as we did in (12) that $|D^j f|, |D^j g| \in L^{q_j}_{\text{loc}}(\Omega)$, where

$$\frac{1}{q_j} = \frac{j-1}{p(m-1)}.$$

Therefore using the Hölder inequality we get that the product $|D^j f D^{m-j} g| \in L^q_{\text{loc}}(\Omega)$ where

$$\frac{1}{q} = \frac{j-1}{p(m-1)} + \frac{m-j-1}{p(m-1)} < \frac{j-1}{p(m-1)} + \frac{m-1-(j-1)}{p(m-1)} = \frac{1}{p}.$$

Consider the two remaining cases $gD^m f$ and $fD^m g$. These are both clearly in $L^p_{\text{loc}}(\Omega)$. Therefore we can derive fg m-times by the product rule and get that $D^m(fg) \in L^p_{\text{loc}}(\Omega)$. $\square$

It is possible to show that similar result holds also on $\mathcal{C}^\infty$ compact n-dimensional Riemannian manifolds.

Theorem 4.2. *Let M and N be $\mathcal{C}^\infty$ compact n-dimensional connected Riemannian manifolds. Let $m \in \mathbb{N}$, $p \in [1,\infty]$ and*

$$\begin{aligned} f &\in W^{m,p}(M,N) \cap \mathrm{Bilip}_{\mathrm{loc}}(M,N),\\ u &\in W^{m,p}(M,N) \cap \mathrm{Lip}(M,N),\\ \varphi &\in W^{m,p}(M,M) \cap \mathrm{Bilip}_{\mathrm{loc}}(M,M) \text{ and}\\ g,h &\in W^{m,p}(M,\mathbb{R}) \cap \mathrm{Lip}(M,\mathbb{R}). \end{aligned}$$

Then

$$\begin{aligned} f^{-1} &\in W^{m,p}(N,M),\\ u\circ\varphi &\in W^{m,p}(M,N) \text{ and}\\ gh &\in W^{m,p}(M,\mathbb{R}). \end{aligned}$$

It is necessary we clarify the meaning of $W^{k,p}(M,N)$ (see e.g. [6]). To begin with we explain that $W^{1,\infty}(M,N) = \mathrm{Lip}(M,N)$ and $W^{1,\infty}(M,\mathbb{R}) = \mathrm{Lip}(M,\mathbb{R})$. Let ρ be the induced metric of the compact Riemannian manifold M. Firstly, notice that all $u \in \mathrm{Lip}(M,\mathbb{R})$ satisfy the following Poincaré type inequality on every ball $B(z,r) \subset M$ and for every $x \in B$,

$$|u(x) - u_B| \le \fint_B |u(x) - u(y)| \le \mathrm{Lip}_u \fint_B \rho(x,y) \le cr\,\mathrm{Lip}_u .$$

Notice however that conversely for any mapping u satisfying $|u(x) - u_B| \le cr$ where the constant c depends on u but not on $B(z,r) \subset M$, is in the class $\mathrm{Lip}(M,\mathbb{R})$. Take any z such that $x,y \in B(z, 3\rho(x,y))$ (given $\rho(x,y) < \infty$) and calculate

$$|u(x) - u(y)| \le |u(x) - u_B| + |u(y) - u_B| \le c\rho(x,y).$$

Further it holds, if we have $f : M \to N$, ϕ a map on N and χ a map on M, that

$$\phi \circ f \circ \chi^{-1} \in W^{1,\infty}(U,\mathbb{R}^n) \Leftrightarrow f \in \mathrm{Lip}(M,N),$$

where U is the open set where χ^{-1} is defined. This is because maps are in fact bilipschitz mappings between the manifold and $\mathbb{R}^n$.

Since we have no problems with continuity or in the spaces $W^{1,p}$(M, N) we may use the classical definition below. We will assume without loss of generality that for our $\mathcal{C}^\infty$ compact n-dimensional Riemannian manifolds we have the finite reference atlases $\{\chi_1, \chi_2, \dots, \chi_k\}, \chi_i : M \to \mathbb{R}^n$, and $\{\phi_1, \phi_2, \dots, \phi_k\}, \phi_i : N \to \mathbb{R}^n$. Using a division of unity we can conclude that $f \in W^{k,p}(M,N) \cap W^{1,\infty}(M,N)$ if and only if

$$\phi_j \circ f \circ \chi_i^{-1} \in W^{k,p}(U_i,\mathbb{R}^n) \cap W^{1,\infty}(U_i,\mathbb{R}^n)$$

for all $1 \le i,j \le k$, where U_i is the open set where χ_i^{-1} is defined.

Proof of Theorem 4.2. Without loss of generality $f(M) = N$. Taking our reference atlases $\{\chi_1, \chi_2, \dots, \chi_k\}$ on M, and $\{\phi_1, \phi_2, \dots, \phi_k\}$ on N and by using Theorem 1.1 we obtain that

$$\chi_i \circ f^{-1} \circ \phi_j^{-1} \in W^{k,p}_{\mathrm{loc}}(V_j, \mathbb{R}^n) \cap W^{1,\infty}_{\mathrm{loc}}(V_j, \mathbb{R}^n),$$

where V_j is the open set where ϕ_i^{-1} is defined. Since our reference atlases are finite we easily get that

$$f^{-1} \in W^{k,p}(N, M) \cap W^{1,\infty}(N, M).$$

Following a similar argument as above and calculating

$$\phi_j \circ u \circ \ \varphi \circ \chi_i^{-1}(x) = \phi_j \circ u \circ \chi_l^{-1} \circ \chi_l \circ \varphi \circ \chi_i^{-1}(x)$$

for such x that the expression on the right is defined. Both $\phi_j \circ u \circ \chi_l^{-1}$ and $\chi_l \circ \varphi \circ \chi_i^{-1}$ are $W^{k,p}$-maps where defined. Therefore their composition is also a $W^{k,p}_{\mathrm{loc}}$-map where defined according to Theorem 1.2. The compactness of M again means that the integrability of $u \circ \varphi$ is global. Apply a similar argument to g and h using Theorem 1.3 to get the last result. $\square$

5. The Implicit Function Theorem

In this section we prove a theorem analogous to the implicit mapping theorem. Before stating it, let us define some sets. Let $\Omega \subset \mathbb{R}^d \times \mathbb{R}^n$ be open and $u : \Omega \to \mathbb{R}^n$. Then we define

$$\Omega^x_z = \{x \in \mathbb{R}^d : \exists y \in \ \mathbb{R}^n : (x, y) \in \Omega, u(x, y) = z\} \text{ and}$$
$$\Omega^y = \{y \in \mathbb{R}^n : \exists x \in \ \mathbb{R}^d : (x, y) \in \Omega\}.$$

We will consider mappings which are bilipschitz 'in the second variable'. We include our definition here.

Definition 5.1. Let $\Omega \subset \mathbb{R}^d \times \mathbb{R}^n$ and $u : \Omega \to \mathbb{R}^m$. The space $\mathrm{Bilip}^2_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$ is the class of mappings $u : \Omega \to \mathbb{R}^m$ such that for all $(x_0, y_0) \in \Omega$, $x_0 \in \mathbb{R}^d, y_0 \in \mathbb{R}^n$, there exists some $\delta > 0$ and $C_1, C_2 > 0$ such that for all $(x, y), (x, y') \in \Omega \cap B((x_0, y_0), \delta)$ it holds that

$$C_1|y - y'| < |u(x, y) - u(x, y')| < C_2|y - y'|.$$

Theorem 5.2. *Let $k, n, d \in \mathbb{N}, p \in [1, \infty]$, let $\Omega \subset \mathbb{R}^d \times \mathbb{R}^n$ be open and*

$$u \in W^{k,p}_{\mathrm{loc}}(\Omega, \mathbb{R}^n) \cap \mathrm{Lip}_{\mathrm{loc}}(\Omega, \mathbb{R}^n) \cap \mathrm{Bilip}^2_{\mathrm{loc}}(\Omega, \mathbb{R}^n).$$

Then for all $z \in u(\Omega)$, Ω_z^x is open in $\mathbb{R}^d$ and for all $x \in \mathbb{R}^d, y, z \in \mathbb{R}^n$ such that $u(x,y) = z$ there exists a neighbourhood $U_x \subset \Omega_z^x$ of x and $V_y \subset \Omega_y$ of y and exactly one mapping $f_z : U_x \to V_y$ such that

$$u(x',y') = z \Leftrightarrow f_z(x') = y' \quad \text{for all } (x',y') \in U_x, \times V_y$$

Further given such a triplet x, y, z we have that $f_z \in W^{1,\infty}(U_x, \mathbb{R}^n)$ and for almost all $z \in u(U_x, V_y)$ we have

$$f_z \in W^{k,p}(U_x, \mathbb{R}^n).$$

Remark 5.3. We cannot expect that our hypothesis will guarantee $W^{k,p}$ regularity for every value of z. This can be seen by considering the following function,

$$u(x,y) = y + (x^2+y^2)^\alpha \sin\left(\frac{1}{x^2+y^2}\right)$$

with $\alpha \in (2, \frac{5}{2})$ and $(x,y) \in \mathbb{R}^2 \setminus \{(0,0)\}$ extended continuously at the origin. Since $u(x,y) - y$ is a radial mapping it is easy to calculate that the norm of the derivative for $j \geq 2$ is

$$|D^j u(x,y)| \leq C \frac{|(x,y)|^{2\alpha}}{|(x,y)|^{3j}}$$

for $|(x,y)| > 0$ and that reverse inequality with different C holds on a a set with positive density at the origin. Notice that u fulfills the hypothesis of Theorem 5.2 with $k = 2$ and any $p < \frac{1}{3-\alpha}$. The derivative of u has a singularity only at the origin, which lies on the graph of the corresponding implicit function f_0.

We want to prove that $\int_0^1 |f_0''(s)| = \infty$ and thus $f_0 \notin W^{2,1}$. Set

$$s_m = \frac{1}{\sqrt{\pi m}}$$

and note that $u(s_m, 0) = 0$ which implies that $f_0(s_m) = 0$. We have that

$$\int_0^1 |f_0''(s)| = c + \lim_{j\to\infty} \sum_{m=1}^{j} \int_{s_{m+1}}^{s_m} |f_0''(s)| \geq \lim_{j\to\infty} \sum_{m=1}^{j} |f_0'(s_m+1) - f_0'(s_m)|.$$

The classical implicit function theorem gives that $f|_{(0,1)} \in \mathcal{C}^\infty(0,1)$ and allows us to calculate

$$f_0'(s_m) = -\frac{D_1 u(s_m, 0)}{D_2 u(s_m, 0)} = (-1)^m 2 s_m^{2\alpha-4+1},$$

where we used $\sin(1/(s_m^2+0^2)) = 1$ and $\cos(1/(s_m^2+0^2)) = (-1)^m$. Now $\alpha < 5/2$ implies

$$\lim_{j\to\infty} \sum_{m=1}^{j} |f_0'(s_m+1) - f_0'(s_m)| \geq C \lim_{j\to\infty} \sum_{m=1}^{j} \left(\frac{1}{m\pi}\right)^{\alpha-3/2} + \left(\frac{1}{(m+1)\pi}\right)^{\alpha-3/2} = \infty$$

and therefore $f \notin W^{2,1}((0,1))$. In fact we have $f \notin W^{2,1}((-\delta,\delta))$ for any $\delta > 0$.

We start by proving the following lemma.

Lemma 5.4. *Let $k, n, d \in \mathbb{N}, p \in [1, \infty]$ and*

$$u \in W^{k,p}_{\text{loc}}(\mathbb{R}^d \times \mathbb{R}^n, \mathbb{R}^n) \cap \text{Lip}_{\text{loc}}(\mathbb{R}^d \times \mathbb{R}^n, \mathbb{R}^n)$$

Further let $r_1, r_2, L_1 > 0$, $x_0 \in \mathbb{R}^d, y_0 \in \mathbb{R}^n$ be such that

$$|u(x, y) - u(x, y')| > L_1|y - y'| \quad \text{for all } y, y' \in B_{\mathbb{R}^n}(y_0, r_2) \tag{16}$$

and any $x \in B_{\mathbb{R}^d}(x_0, r_1)$. Let $z \in u(x_0, B_{\mathbb{R}^n}(y_0, \frac{L_1 r_2}{2L_2}))$ then there exists $\delta > 0$ and a mapping $f_z : B_{\mathbb{R}^d}(x_0, \delta) \to B_{\mathbb{R}^n}(y_0, r_2)$ such that

$$u(x, y) = z \Leftrightarrow f_z(x) = y \quad \text{for all } x \in B_{\mathbb{R}^d}(x_0, \delta).$$

Further $f_z \in W^{1,\infty}(B_{\mathbb{R}^d}(x_0, \delta), \mathbb{R}^n)$ for all $z \in u\Big(x_0, B_{\mathbb{R}^n}\big(y_0, \frac{L_1 r_2}{2L_2}\big)\Big)$ and

$$f_z \in W^{k,p}(B_{\mathbb{R}^d}(x_0, \delta), \mathbb{R}^n) \ \text{ for almost all } z \in u\Big(x_0, B_{\mathbb{R}^n}\big(y_0, \frac{L_1 r_2}{2L_2}\big)\Big).$$

Proof. We have $u \in \text{Lip}_{\text{loc}}(\mathbb{R}^d \times \mathbb{R}^n, \mathbb{R}^n)$ and therefore there exists some $L_2 > 0$ such that

$$|u(x, y) - u(x', y')| < L_2(|x - x'| + |y - y'|) \tag{17}$$

for all $x, x' \in B_{\mathbb{R}^n}(x_0, r_1)$, and all $y, y' \in B_{\mathbb{R}^n}(y_0, r_2)$. Put

$$\delta = \frac{r_2 L_1}{2L_2}.$$

We define $h : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}^d \times \mathbb{R}^n$ as follows

$$h(x, y) = (x, u(x, y)).$$

Evidently $h \in W^{k,p}_{\text{loc}}(\mathbb{R}^d \times \mathbb{R}^n, \mathbb{R}^{d+n})$. We want to prove that

$$h \in \text{Bilip}(B_{\mathbb{R}^d}(x_0, r_1) \times B_{\mathbb{R}^n}(y_0, r_2), \mathbb{R}^{d+n}).$$

It is evident that h is Lipschitz as its component mappings are Lipschitz. Consider $x, x' \in B_{\mathbb{R}^n}(x_0, r_1)$, and $y, y' \in B_{\mathbb{R}^n}(y_0, r_2)$, such that

$$\frac{|y - y'|}{|x - x'|} \leq \frac{2L_2}{L_1}.$$

We have

$$|h(x, y) - h(x', y')| \geq |x - x'| \geq c(|x - x'| + |y - y'|)$$

for some $c > 0$. Now conversely take

$$\frac{|y-y'|}{|x-x'|} \geq \frac{2L_2}{L_1}$$

and by (16) and (17) we get

$$|h(x,y) - h(x',y')| \geq L_1|y-y'| - L_2|x-x'| \geq \frac{L_1|y-y'|}{2} \geq c(|x-x'| + |y-y'|).$$

for some $c > 0$. Now we may denote $\Omega' = h(B_{\mathbb{R}^d}(x_0, r_1), B_{\mathbb{R}^n}(y_0, r_2))$. Clearly Ω' is open. By Theorem 1.1 we get the regularity

$$h^{-1} \in W^{k,p}_{\mathrm{loc}}(\Omega', \mathbb{R}^{d+n}) \cap \mathrm{Bilip}(\Omega', \mathbb{R}^{d+n}).$$

Our goal is to define $f_z(x) = y$. Let us firstly show that if such a y exists then it is unique. The inequality (16) guarantees that if $u(x,y) = u(x,y')$ then $y = y'$. Therefore if $h(x,y) = h(x,y')$ then $y = y'$, which implies that for any given $x \in B_{\mathbb{R}^d}(x_0, r_1)$ and $z \in u(x, B_{\mathbb{R}^n}(y_0, r_2))$ there exists at most one y such that $u(x,y) = z$.

It now suffices to prove that for all $z \in \mathbb{R}^n$ such that $z = u(x_0, \hat{y})$ for some $\hat{y} \in B_{\mathbb{R}^n}(y_0, \frac{r_2 L_1}{2L_2})$ we have: for all $x \in B_{\mathbb{R}^d}(x_0, \delta)$ there exists a $y \in B_{\mathbb{R}^n}(y_0, r_2)$ such that $z = u(x,y)$. Remember that $\delta L_2 = r_2 L_1/2$ and put $z_0 = u(x_0, y_0)$. Using the fact that u is L_2-Lipschitz in the y variable and then the definition of δ we get

$$u\Big(x_0, B_{\mathbb{R}^n}\Big(y_0, \frac{r_2 L_1}{2L_2}\Big)\Big) \subset B_{\mathbb{R}^n}\Big(z_0, \frac{L_1 r_2}{2}\Big) = B_{\mathbb{R}^n}(z_0, L_1 r_2 - \delta L_2).$$

Now since u is L_2-Lipschitz in the x variable and using (16), we get that for all $x \in B_{\mathbb{R}^d}(x_0, \delta)$ that

$$B_{\mathbb{R}^n}(z_0, L_1 r_2 - \delta L_2) \subset B_{\mathbb{R}^n}((x, y_0), L_1 r_2) \subset u\big(x, B_{\mathbb{R}^n}(y_0, r_2)\big).$$

Hence we can really find $y \in B_{\mathbb{R}^n}(y_0, r_2)$ such that $u(x,y) = u(x_0, y_0)$. This gives us the existence of a mapping f_z defined for all $x \in B_{\mathbb{R}^d}(x_0, \delta)$. We show that f_z is Lipschitz (with the constant $\frac{L_2}{L_1}$). Consider two pairs (x,y) and (x',y') such that $u(x,y) = u(x',y') = z$. By (16) and (17) we have

$$\begin{aligned} |u(x,y) - u(x',y)| &< L_2|x-x'| \\ |u(x,y) - u(x',y)| = |u(x',y') - u(x',y)| &> L_1|y-y'|. \end{aligned}$$

Thus

$$|f_z(x) - f_z(x')| = |y-y'| \leq \frac{L_2|x-x'|}{L_1}.$$

It is now left to prove that for almost all $z \in u(x_0, B_{\mathbb{R}^n}(y_0, \frac{r_2 L_1}{2L_2}))$ we have $f_z \in W^{k,p}_{\mathrm{loc}}(B_{\mathbb{R}^d}(x_0,\delta), \mathbb{R}^n)$. Here it suffices to use Theorem 2.8 and the ensuing comment on the mapping h^{-1} and realize that the ACL condition implies that

$$h^{-1}(\cdot, z) \in W^{k,p}(B_{\mathbb{R}^d}(x_0,\delta), \mathbb{R}^{d+n}) \quad \text{for almost all } z \in u(x_0, B_{\mathbb{R}^n}(y_0, \frac{r_2 L_1}{2L_2})).$$

Take any such a point z and use the following notation for the coordinate mappings in the given dimensions d and n, $h^{-1} = (h_1^{-1}, h_2^{-1})$, then clearly $h_2^{-1}(\cdot, z) \in W^{k,p}(B_{\mathbb{R}^d}(x_0,\delta), \mathbb{R}^n)$. But clearly for all $x \in B_{\mathbb{R}^d}(x_0,\delta)$ it holds that

$$h_2^{-1}(x, z) = f_z(x).$$

Thus we have $f_z \in W^{k,p}_{\mathrm{loc}}(B_{\mathbb{R}^d}(x_0,\delta), \mathbb{R}^n)$ for almost all $z \in u(x_0, B_{\mathbb{R}^n}(y_0, \frac{r_2 L_1}{2L_2}))$. $\square$

Proof of Theorem 5.2. We have $u \in \mathrm{Bilip}^2_{\mathrm{loc}}(\Omega, \mathbb{R}^n)$ and therefore for any fixed $(x, y) \in \Omega$ we find $r_1, r_2 > 0$ for which we may apply Lemma 5.4 (note that our proof does not require u defined outside of $B(x_0, r_1) \times B(y_0, r_2)$). This means that for any fixed $z \in u(\Omega)$ that Ω_z^x is open in $\mathbb{R}^d$. It also implies the local existence of a Lipschitz $f_z : U_x \to V_y$ for all x, y, z and that for almost all z we have $f_z \in W^{k,p}(U_x, \mathbb{R}^n)$. $\square$

Acknowledgement. The authors would like to thank Robert Černý for carefully reading the manuscript and for his pointed comments. The first two authors were supported by the ERC CZ grant LL1203 of the Czech Ministry of Education.

References

[1] Ambrosio, L., Fusco, N. and Pallara, D., *Functions of bounded variation and free discontinuity problems.* Oxford Mathematical Monographs, The Clarendon Press, Oxford University press, New York 2000.

[2] Arnold, V., Sur la géometrie differentielle des groupes de Lie de dimension infinie et ses applications á l'hydrodynamique des fluids parfaits. *Ann. Inst. Fourier* 16 (1966), 319 – 361.

[3] Brezis, H. and Mironescu, P., Gagliardo-Nirenberg, composition and products in fractional Sobolev spaces. *J. Evol. Equ.* 1 (2001), 387 – 404.

[4] Csörnyei, M., Hencl, S. and Malý, J., Homeomorphisms in the Sobolev space $W^{1,n-1}$. *J. Reine Angew. Math* 644 (2010), 221 – 235.

[5] Hajlasz, P., Iwaniec, T., Malý, J. and Onninen, J., *Weakly differentiable mappings between manifolds.* Mem. Amer. Math. Soc. 899 (2008), No. 899.

[6] Hajlasz, P., Sobolev mappings between manifolds and metric spaces. *Sobolev Spaces in Mathematics, Int. Math. Ser., Springer, New York* I (2009), 185 – 222.

[7] Hencl, S., Bilipschitz mappings with derivative of bounded variation. *Publ. Math.* 52 (2008), 91 – 99.

[8] Hencl, S. and Koskela, P., *Lectures on Mappings of finite distortion.* Lecture notes in Mathematics 2096, Springer 2014.

[9] Inci, H., Kappeler, T. and Topalov, P., *On the regularity of the composition of diffeomorphisms.* Mem. Amer. Math. Soc. (2013), No. 226.

[10] Iwaniec, T. and Martin, G., *Geometric function theory and nonlinear analysis.* Oxford Mathematical Monographs, Clarendon Press, Oxford 2001.

[11] Lee, J. M., *Manifolds and Differential Geometry.* Graduate Studies in Mathematics, 107. American Mathematical Society 2009.

[12] Mazya, V. and Shaposhnikova, T., An elementary proof of the Brezis and Mironescu theorem on the composition operator in fractional Sobolev spaces. *J. Evol. Equ.* 2 (2002), 113 – 125.

[13] Nirenberg, L., On elliptic partial differential equations. *Ann. Scuola Norm. Sup. Pisa* 3 (1959), 115 – 162.

Received R$^m April 1, 2012; revised \mathbb{R}^m January 31, 2013$

4. Diffeomorphic approximation of Planar Sobolev Homeomorphisms in Orlicz-Sobolev spaces

Daniel Campbell

Preprint

DIFFEOMORPHIC APPROXIMATION OF PLANAR SOBOLEV HOMEOMORPHISMS IN ORLICZ-SOBOLEV SPACES.

DANIEL CAMPBELL

Abstract. Let $\Omega \subseteq \mathbb{R}^2$ be a domain, let Φ be a Δ_2 Young function and let $f \in W^{1,\Phi}(\Omega, \mathbb{R}^2)$ be a homeomorphism between Ω and $f(\Omega)$. Then there exists a sequence of diffeomorphisms f_k converging to f in the Sobolev-Orlicz space $W^{1,\Phi}(\Omega, \mathbb{R}^2)$. Further for an injective continuous map $\varphi \in W^{1,\Phi}(\partial(-1,1)^2, \mathbb{R}^2)$ we find a diffeomorphism in $W^{1,\Phi}((-1,1)^2, \mathbb{R}^2)$ that equals φ on the boundary.

1. Introduction

The problem of approximating homeomorphisms $f : \mathbb{R}^n \supseteq \Omega \longrightarrow f(\Omega) \subseteq \mathbb{R}^n$ with either diffeomorphisms or piecewise-affine homeomorphisms has proven to be both very challenging and of great interest in a variety of contexts. Although there are a number of elementary tools available for constructing smooth approximations of a mapping, for example convolution approximation or Lipschitz extension, it is not generally true that these processes maintain the injectivity of a mapping, and there are a number of applications where the injectivity is crucial.

We recommend [12] to the reader for an account of some of the the most fundamental results related to this problem. Initially, it was uniform convergence that was of interest in connection with geometric topology but after the L^∞-approximation problem had been completely solved, the question of approximating homeomorphisms revived again in the altogether different context of non-linear elasticity initiated by Ball [3]. The variational model takes an elastic body in a reference configuration, which we call the domain Ω, and given boundary values on $\partial\Omega$, which determine the shape of the deformed body after deformation. The model deformation of the body Ω is the homeomorphism f that satisfies the boundary data, while minimising an energy functional of the form

$$I(f) = \int_\Omega W(Df)\, dx\,,$$

where $W : \mathbb{R}^{n\times n} \to \mathbb{R}$ is the stored-energy functional satisfying

$$W(A) \to +\infty \quad \text{as } \det A \to 0 \qquad W(A) = +\infty \quad \text{if } \det A \leq 0\,. \tag{1.1}$$

There are two points of interest here, the characteristics of W and the fact that we require f to be a homeomorphism. As pointed out by Ball in [4, 5] we must require that f is a homeomorphism because, as matter is impenetrable, our mapping must be one-to-one, secondly continuity corresponds to the material not breaking during the deformation. For one, the conditions in (1.1) prevent excessive compressions of the elastic body and secondly guarantee that the orientation of the material remains

2000 *Mathematics Subject Classification.* 46E35.

The author was supported by the ERC CZ grant LL1203 of the Czech Ministry of Education.

unchanged. This means that, if f is an admissible deformation with finite energy, then one has that

$$\det Df > 0 \quad \text{a.e. in } \Omega.$$

As stated by Ball (for example in [2]) it is often natural to expect the stored energy function W to be quasiconvex. Unfortunately, there are no known results, which given (1.1) and quasiconvexity guarantee the existence of a solution. Either one is forced to drop the condition (1.1) and impose p-growth conditions on W (see [19, 1]), or assume that W is *polyconvex* and that some coercivity conditions are satisfied (see [2, 20]). Moreover even where the existence of $W^{1,p}$ minimisers is known, very little is known about their regularity.

In [4, 5] Ball, shows that understanding the regularity of minimisers of a quasiconvex W satisfying (1.1) would be greatly aided if one was able to find a minimising sequence of piecewise-affine homeomorphisms or diffeomorphisms. To this end it would be useful to be able to approximate a homeomorphism $u \in W^{1,p}(\Omega; \mathbb{R}^n)$, $p \in [1, +\infty)$ in $W^{1,p}$ by piecewise-affine or smooth homeomorphisms.

One very significant reason why this would be desirable, is that regularity is typically often proven by testing the weak equation or the variation formulation by the solution itself; but unless one has some a priori regularity of the solution, such a test may not make sense. In order to solve this problem it would be possible to test the equation with a smooth test mapping which is close to the given homeomorphism instead. Here we see the necessity for the approximations to be homeomorphisms whose image is the same as that of the approximated map, otherwise this sequence would have nothing in common with our original problem. Besides non-linear elasticity, an approximation result of homeomorphisms with diffeomorphisms would be a very useful tool in and of itself as it would allow a number of proofs to be significantly simplified and lead to some stronger results.

We will now review some of the techniques used to approximate homeomorphisms with piecewise-affine homeomorphisms or diffeomorphisms. Let us start by mentioning the result of Mora-Corral and Pratelli [18], which shows that (on the plane) it is not important whether we approximate using piecewise-affine homeomorphisms or diffeomorphisms in the $W^{1,p}$ case as they are equivalent. In fact, this result can be used to approximate in Δ_2-Orlicz-Sobolev spaces also.

The first positive results were achieved by Mora-Corral [17] in 2009 on homeomorphisms smooth outside a point and by Bellido and Mora-Corral [6] in 2011 on approximation in Hölder continuous maps. Shortly following this came the famous and very significant Sobolev approximation result by care of Iwaniec, Kovalev and Onninen [14], [15]. In their papers, published in 2011 and 2012, they found diffeomorphic approximations to any homeomorphism $f \in W^{1,p}(\Omega, \mathbb{R}^2)$, for any $1 < p < \infty$ in the $W^{1,p}$ norm. This celebrated result was a breakthrough in terms of the approximation problem and stimulated much interest in the subject. The only open case left open was for $p = 1$. This however has been solved very recently by Hencl and Pratelli in [12]. The problem of approximating homeomorphisms with diffeomorphisms cannot be considerred entirely closed even in the planar case. We would like to know how to approximate both a map and its inverse simultaneously in $W^{1,p}$. Building on the techniques pioneered in [12], Pratelli has answered this question for $p = 1$ in the paper [21]. The cases $p > 1$ (especially $p = 2$) which are even more important in terms of their application are still open.

At present there are two known approaches to finding diffeomorphisms close to a homeomorphism. In [14] the authors equate $\mathbb{R}^2$ with $\mathbb{C}$ and use p-harmonic functions to extend a mapping with predefined boundary values. On the other hand [12] divides Ω into squares based on Lebesgue points of the derivative, some of which can be approximated simply by triangulation, some of which must be approximated using an extension theorem and some where one uses a sophisticated combination of both of these techniques. Finally Hencl and Pratelli apply the result from [18] to smooth their approximations. This leads to an essential question; can the approach used in [12] be modified to apply in the general case, $p \in [1, \infty)$? In order to answer this question we state our main result.

By a Young function we mean a convex $\Phi : [0, \infty] \to [0, \infty]$ such that $\Phi(0) = 0$ and $\lim_{t\to\infty} \Phi(t) = \infty$. The Luxembourg norm is defined as

$$\|f\|_{W^{1,\Phi}(\Omega,\mathbb{R}^n)} = \inf \Big\{ \lambda > 0 : \int_\Omega \Phi\Big(\frac{|Df|}{\lambda}\Big) + \Phi\Big(\frac{|f|}{\lambda}\Big) < 1 \Big\}.$$

and is a generalization of the $W^{1,p}$ norm in the case $\Phi(t) = t^p$.

Theorem 1.1. *Let $\Omega \subseteq \mathbb{R}^2$ and let ε_k be a decreasing sequence of positive numbers tending to zero. Let Φ be a Young function satisfying:*

(Δ_2) *There exists a constant $C_2 > 0$ such that for all $t \in (0, \infty)$,*
$$\Phi(2t) \le C_2\Phi(t).$$

For any homeomorphism $f \in W^{1,\Phi}(\Omega, \mathbb{R}^2)$ there exists a sequence of diffeomorphisms $f_k \in W^{1,\Phi}(\Omega, \mathbb{R}^2)$ such that

$$\|f - f_k\|_{L^\infty(\Omega,\mathbb{R}^2)} < \varepsilon_k \text{ and } \|f - f_k\|_{W^{1,\Phi}(\Omega,\mathbb{R}^2)} \le \varepsilon_k,$$

where $\|\cdot\|_{W^{1,\Phi}(\Omega,\mathbb{R}^2)}$ is the Luxembourg norm. Moreover for each k we have $f_k(\Omega) = f(\Omega)$ and if f can be continuously extended onto the boundary, then f_k coincides with f on $\partial\Omega$.

We can replace our approximating diffeomorphisms with piecewise-affine homeomorphisms and, in the case that $\partial\Omega$ is a polygon and f is piecewise-linear on $\partial\Omega$, then we can construct f_k such that $f_k = f$ on $\partial\Omega$ and the triangulation of f_k is finite.

Since [14] uses p-harmonic functions it is not at all clear that their result could be extended to Orlicz-Sobolev classes as there is no variant of the Radó-Kneser-Choquet in that setting. Nevertheless, it would be desirable to be able to approximate in Orlicz-Sobolev to get sharp results. Since the extensions in [14] use solutions to the p-Laplacian the approximation generated is fundamentally different for each p. In the techniques we apply here, which were pioneered in [12] we are able to approximate f using basically the same technique for any $p \in [1, \infty)$ or in fact a Δ_2-Young function. Another point of interest is that some techniques used here may prove to inspire approximation in higher dimensions as it does not need to identify $\mathbb{R}^2$ and $\mathbb{C}$. In these senses our result extends the result in [14].

The higher dimensional case remains a very interesting and challenging question. It has been proven very recently by Hencl and Vejnar that an approximation theorem like Theorem 1.1 could not hold for $n \ge 4$ for all homeomorphisms in $W^{1,1}(\Omega, \mathbb{R}^n)$ by their example in [13]. It is still possible that a positive result may be found, however, for homeomorphisms in $W^{1,p}$ given that p is large enough.

We will now state a secondary result, which we will prove in this paper. Given a Sobolev and one-to-one mapping on $\partial\Omega$ we would like to be able to construct a homeomorphism which has exactly those boundary values. Further it would be desirable that we have a bound on the norm, or at least the modular of our extension. Recall that homeomorphic extension theorems have proven to be key in proving the claims in both [14] and [12]. We prove the following extension theorem.

Theorem 1.2. *Let $\Omega' \subseteq \mathbb{R}^2$ be a Lipschitz domain. Let Φ be a Δ_2-Young function and $\varphi \in W^{1,\Phi}(\partial(-1,1)^2, \mathbb{R}^2)$ be a homeomorphism onto $\partial\Omega'$. Then there exists a diffeomorphism $f \in W^{1,\Phi}((-1,1)^2, \mathbb{R}^2)$ such that $f = \varphi$ on $\partial(-1,1)^2$ and*

$$\fint_{(-1,1)^2} \Phi(|Df|) \leq C \fint_{\partial(-1,1)^2} \Phi(|D_\tau \varphi|). \tag{1.2}$$

Here $D_\tau\varphi$ is the derivative of φ on the one-dimensional object $\partial(-1,1)$ as we consider φ to be Sobolev on the boundary of the square. For the definition see the preliminaries.

Notice that one could also get an extension theorem using a similar approach as in [14] by using the Radó-Kneser-Choquet theorem. One would apply Ψ a bi-Lipschitz change of variables so that $\Psi(\Omega')$ is a square, extend the boundary values and then return the bi-Lipschitz change of variables. This would give one a bi-Lipschitz homeomorphism, from which one could construct a diffeomorphism. Nevertheless the modular of the diffeomorphism could be highly dependent on the geometry of Ω'. If one imagines a star shaped domain, one could change the bi-Lipschitz constant of our change of variables enormously by changing the values of φ only very slightly. Hence the constant C in (1.2) is potentially heavily dependant on the precise shape of Ω'. Our result however enjoys the property that the C in (1.2) is an absolute constant and the average modular of the extension depends truly only on the average modular of φ.

1.1. **Brief description of the proof of Theorem 1.1.** In this subsection we outline the basic plan of our proof of Theorem 1.1. We will follow the approach introduced by Hencl and Pratelli. We find piecewise-affine approximations in $W^{1,\Phi}$ and then use [18] to smooth them into diffeomorphisms. It must be noted that our triangulation of Ω may be infinite, nevertheless, whenever we have a compactly embedded domain in Ω, the triangulation on this subset is finite. In this sense we say that our triangulation is locally finite in Ω. We find a finite triangulation whenever $\partial\Omega$ is a polygon on which f is piecewise-linear. If either of these two conditions are not satisfied there is no finite triangulation of f.

The most substantial part of our proof is in Theorem 4.1. Here we use the fact that $\mathcal{L}^2(\Omega) < \infty$ and find an approximation f_1 such that $\Phi(|Df_1 - Df|)$ is arbitrarily small. To do this we will divide Ω into a locally finite grid of very small squares. We will call a square "good" given that the derivative satisfies a Lebesgue-point-type estimate in terms of L^1 and in terms of L^Φ and given we can approximate f by the affine mapping $(f(x_0) + Df(x_0)(x - x_0))$ to a given degree of accuracy. If we have a good square where the derivative is neither too small nor too big and the Jacobian is not too small then we can approximate by simple triangulation. By this we mean we separate the square diagonally into two triangles and the values of f at the corners define an affine mapping on each triangle. These affine mappings will be very close to $f(x_0)+Df(x_0)(x-x_0)$ and therefore very close to f in $W^{1,\Phi}$. Therefore our approximation is close to f in $W^{1,\Phi}$ on such goldilocks squares. Since almost all points are Lebesgue points for the derivative we will use this simple approximation technique on the vast majority of Ω where the Jacobian is non-zero.

In the process of the above it may be necessary to slightly move the corners of the squares and get a somewhat deformed grid so that the following estimates hold

$$\fint_{\partial Q}|D_\tau f|\leq K\fint_{2Q}|Df| \text{ and } \fint_{\partial Q}\Phi(|D_\tau f|)\leq K\fint_{2Q}\Phi(|Df|)\,,$$

where K is a big, but fixed, constant. Although they are now polygons we will still refer to elements of the grid as squares.

We still want to define an approximation in bad squares, (where the Jacobian is very small, or the derivative is very large or small) and good squares with zero Jacobian. To do this we define f_1, an approximation of f on the boundaries of our distorted squares, by piecewise-linear approximation, which is fine enough to guarantee that this approximation is one-to-one on the grid and close to the original mapping f. Automatically the triangle inequality and convexity of Φ will give us that

$$\int_{\partial Q}|D_\tau f_1|\leq\int_{\partial Q}|D_\tau f|, \quad \text{and} \quad \int_{\partial Q}\Phi(|D_\tau f_1|)\leq\int_{\partial Q}\Phi(|D_\tau f|).$$

On these bad squares we will use the theorem proven in Section 2, which allows us to extend a mapping with one-to-one piecewise-linear boundary values φ to a finite piecewise-affine homeomorphism h inside the square, while maintaining the following control on its size

$$\fint_{Q}\Phi(|Dh|)\leq C\fint_{\partial Q}\Phi(|D_\tau\varphi|)\,.$$

One can find a rigorous proof of this extension theorem, introduced by Hencl and Prattelli in [12] for $W^{1,1}$, and by Radici in [22] in the $W^{1,p}$ case. Therefore we will only trace certain details in the process of proving this result for $W^{1,\Phi}$.

The construction is done by first choosing a pair of opposing points on the square where average integrals of the derivative around these points on the boundary is not too big. We then move these points to opposite corners. Next we choose lines going diagonally across the square and define h on those lines as the constant-speed parametrization of the shortest curve joining the images of the chosen points inside the image of the square. As long as we place our lines close enough together and at the right points we will easily be able to construct a piecewise-affine mapping in the strip between these two lines because the shortest curves are themselves piecewise-linear. We slightly alter our curves in the image to separate them from each other as there may be places where the shortest curves coincide. This ensures that our construction gives a homeomorphism. As proven in [12] this works in $W^{1,1}$ but in fact the Lipschitz constant of h on each strip can be bounded by the Lipschitz constant of φ the intersection of the strip and the boundary. This means we can prove

$$\fint_{Q}\Phi(|Dh|)\leq C\fint_{\partial Q}\Phi(|D_\tau\varphi|)\leq C\fint_{\partial Q}\Phi(|D_\tau f|)\leq CK\fint_{2Q}\Phi(|Df|).$$

Our choice of parameters can ensure that the union of all bad squares has tiny measure and therefore the integral

$$\int_{\bigcup_{\mathcal{B}} Q} \Phi(|f - f_1|) + \Phi(|Df - Df_1|) < C\varepsilon$$

thanks to the absolute continuity of the integral. Thus we have taken care of the bad squares.

We deal with good squares with zero Jacobian, of which there may be very many (possibly full measure, see [10]), in Section 3. In the corresponding step in [12] the authors used bi-Lipschitz change of variables without an a priori bound on the bi-Lipschitz constant, therefore we cannot apply their approach in our context. Our completely new approach avoids the need for changes of variables as we will now describe.

We have a square Q centred at x_0 and piecewise-linear boundary values and we want to define an extension g close to f on the square. We take W a large central part of the square Q and squash it so that it follows part of the image of the boundary of Q with constant speed. The image of the square W, $g(W)$, is a snake-shaped object inside the snake $g(Q)$. On W we will calculate that our mapping is close to $f(x_0) + Df(x_0)(x - x_0)$ in $W^{1,1}$. Since we maintain a bound on the Lipschitz constant of our approximation we will be able to use an interpolative inequality to estimate the modular of $Dg - Df(x_0)$ in L^Φ. We divide the very narrow remaining annulus, $Q \setminus W$ into squares and treat them as bad squares, using the result from Section 2. Since the overall square is good, we know that g is close to f in the centre. On the annulus we know the functional values differ very little from the affine map $f(x_0) + Df(x_0)(x - x_0)$, which means that Dg on each small square is nicely bounded. Therefore, as the annulus has small measure, we get the integrals of $\Phi(|Df|)$ and $\Phi(|Dg|)$ over the annulus are small.

In Theorem 4.2 we show how to find a finite triangulation of f if f is piecewise-linear on $\partial\Omega$, which is a polygon. The strategy of the proof is to separate a very thin tube around the boundary into bi-Lipschitz image of squares. Then we can define f_1 appropriately on the boundary of bi-Lipschitz images of squares which lie around $\partial\Omega$ and use Theorem 2.1 to extend these values, which gives us a homeomorphism on the tube around the boundary. Then, because we are working entirely in a tube of tiny measure we will find that the integral of $\Phi(|Df|)$ and $\Phi(|Df_1|)$ around the boundary will be very small. As each section around the boundary will be the bi-Lipschitz image of a square of fixed size we will cover the tube around the boundary with a finite number of such sets. This makes the total triangulation finite because the triangulation of a compactly embedded subdomain in Ω is finite. This approach is different to the approach in [12].

In Theorem 4.3 we consider general Ω, which we exhaust with a monotone sequence of compactly embedded domains in Ω. On each domain we apply the above approximation techniques with stricter and stricter bounds on the modular of the error. This means that $f_k \to f$ in $W^{1,\Phi}(\Omega, \mathbb{R}^2)$ and in fact that the error vanishes on the boundary of Ω. By this we mean that we can approximate $f - f_k$f by a $g_m \in \mathcal{C}_c^\infty(\Omega)$ whose $W^{1,\Phi}(\Omega, \mathbb{R}^2)$ norm is less than $2^{-m}\varepsilon_k$ for $m \in \mathbb{N}$.

1.2. **Preliminaries.** In this subsection we shortly list the basic notation that will be used throughout the paper. The set $Q(c,r) = \{(x,y) \in \mathbb{R}^2 : |x-c_1| \leq r, |y-c_2| \leq r\}$ will denote the square centred at c with side length $2r$. The set $\mathcal{S}_{r_0} = \{(x,y) \in \mathbb{R}^2 : |x| + |y| < r_0\}$ will be the "rotated square". Similarly, $\mathcal{B}(c,r)$ is the ball centred at c with radius r. For the ease of notation, for $t > 0$ we will denote $tQ(c,r) = Q(c,tr)$, and $t\mathcal{B}(c,r) = \mathcal{B}(c,tr)$; more in general, for a generic set $\mathcal{Q}$ with a notional centre $c \in \mathcal{Q}$ and $t > 0$, we write $t\mathcal{Q} = \{x \in \mathbb{R}^2 : \ (x-c)/t \in \mathcal{Q}\}$.

The points in the preimage Ω will be always denoted by capital letters, such as A, B and so on, while points in the image $f(\Omega)$ will be always denoted by bold capital letters, such as $\boldsymbol{A}$, $\boldsymbol{B}$ and similar. To shorten the notation and help the reader, whenever we use the same letter A for a point in the domain and $\boldsymbol{A}$ (in bold) for a point in the target, this always means that $\boldsymbol{A}$ is the image of A under the mapping that we are considering in that moment.

We will use the following notation for segments $[AB]$ (resp., $[\boldsymbol{AB}]$), which is the set $\operatorname{co}\{A,B\}$ (resp. $\operatorname{co}\{\boldsymbol{A},\boldsymbol{B}\}$). The length of this segment is $\mathcal{H}^1(AB) = |A-B|$ and more generally $\mathcal{H}^1(\gamma)$ is the length of a curve γ. Given three non-aligned points A, B, C we denote the triangle $[AB] \cup [BC] \cup [CA]$ as ABC and set inclosed by this triangle is $\operatorname{co}\{A,B,C\}$. If we have three adjacent vertices in a polygon we call the angle between $[AB]$ and $[BC]$ measured over the interior of the polygon $\measuredangle ABC$. We use $\angle ABC$ if we need to refer to an angle itself rather than the value of an angle. Given two vectors u, v we call $\measuredangle uv$ the angle in $(0,\pi)$ between them.

We will denote the modulus of the horizontal and vertical derivatives of any mapping $f = (f^1, f^2) : \mathbb{R}^2 \to \mathbb{R}^2$ as

$$|D_1 f| = \sqrt{\left(\frac{\partial f^1}{\partial x}\right)^2 + \left(\frac{\partial f^2}{\partial x}\right)^2}, \qquad |D_2 f| = \sqrt{\left(\frac{\partial f^1}{\partial y}\right)^2 + \left(\frac{\partial f^2}{\partial y}\right)^2}.$$

Analogously, the derivatives of the components f^1 and f^2 are written as

$$D_1 f^1 = \frac{\partial f^1}{\partial x}, \quad D_2 f^1 = \frac{\partial f^1}{\partial y}, \qquad D_1 f^2 = \frac{\partial f^2}{\partial x}, \quad D_2 f^2 = \frac{\partial f^2}{\partial y}.$$

We refer to $\Phi : [0,\infty] \to [0,\infty]$ as a Young function if Φ is convex, $\Phi(0) = 0$ and $\lim_{t\to\infty} \Phi(t) = \infty$. We will say that a Young function Φ satisfies Δ_2 if there exists a constant C_2 such that for all $t \in (0,\infty)$, we have

$$\Phi(2t) \leq C_2 \Phi(t).$$

Given a function f on Ω we refer to the integral

$$\int_\Omega \Phi(|f|)$$

as the modular of f on Ω.

Let Ω be an open set. Given a Δ_2-Young function we will define $W^{1,\Phi}(\Omega,\mathbb{R}^n)$ as the set of $f \in W^{1,1}(\Omega,\mathbb{R}^n)$ such that

$$\int_\Omega \Phi(|Df|) < \infty.$$

We define the so-called Luxembourg norm $\|f\|_{W^{1,\Phi}}$ for all $f \in W^{1,\Phi}(\Omega, \mathbb{R}^n)$ as

$$\|f\|_{W^{1,\Phi}(\Omega,\mathbb{R}^n)} = \inf\Big\{\lambda > 0 : \int_\Omega \Phi\Big(\frac{|Df|}{\lambda}\Big) + \Phi\Big(\frac{|f|}{\lambda}\Big) < 1\Big\}. \tag{1.3}$$

It is well known that thus defined $W^{1,\Phi}$ is a Banach space.

Sometimes we will work on 1-dimensional objects in $\mathbb{R}^2$, which can be parametrised from $[0,1]$ by a Lipschitz mapping φ, for example segments, and various polygons but also the boundary of a Lipschitz domain. We may assume that our φ is one-to-one and $|\varphi'|$ is constant almost everywhere. For almost all $t \in (0,1)$ there exists a vector $\frac{\varphi'(t)}{|\varphi'(t)|}$ which we call the tangential vector at the point $\varphi(t)$ and denote this vector as $\tau = \tau(\varphi(t))$. If a mapping f is defined on $\varphi([0,1])$ and $f \circ \varphi$ is absolutely continuous, then we call

$$D_\tau f = \frac{(f\circ\varphi)'}{|\varphi'|}.$$

For different choices of parametrization φ_1, φ_2 we may get $\tau_1 = -\tau_2$ and so $D_{\tau_1} f = -D_{\tau_2} f$. Nevertheless all of our calculations are independent of a fixed choice of φ and our results hold independently from the choice of parametrization. Specifically notice that $|D_{\tau_1} f| = |D_{\tau_2} f|$. In this respect let us have a Lipschitz, one-to-one $\varphi : [0,1] \to \mathbb{R}^2$ and $F = \varphi([0,1])$. We define $W^{1,1}(F, \mathbb{R}^n)$ as the mappings $f : F \to \mathbb{R}^n$ such that $f\circ\varphi$ is absolutely continuous and $\int_F |D_\tau f| < \infty$ or equivalently $f \circ \varphi \in W^{1,1}((0,1), \mathbb{R}^n)$. Given a Δ_2-Young function Φ we will have $f \in W^{1,\Phi}(F, \mathbb{R}^n)$ if $f \in W^{1,1}(F, \mathbb{R}^n)$ and $\Phi(|D_\tau f|)$ is $\mathcal{H}^1$ integrable. We can norm this space using the standard Luxembourg norm.

Denote $e_1 = (1,0)$ and $e_2 = (0,1)$. If we have a mapping $\mathcal{A}$ defined on some open set of $\mathbb{R}^2$ by the formula $\mathcal{A}(x) = A + (B_{1,1}X_1 + B_{1,2}X_2)e_1 + (B_{2,1}X_1 + B_{2,2}X_2)e_2$ for some choice of constant parameters $A, B_{i,j}$ then we will refer to $\mathcal{A}$ as being an affine mapping. If we have a mapping f defined on a segment and f is the restriction of an affine map onto the segment then we refer to f as being linear.

We take advantage of standard denotation of average integrals using the symbol $⨍$. Since we integrate with respect to different measures, we emphasise the fact that we divide the integral by the measure of the set we integrated over, *where we measure the set with the same measure used in the integral.*

The symbol π_1 will be the orthogonal projection of $\mathbb{R}^2$ onto the x-axis and π_2 will be the orthogonal projection of $\mathbb{R}^2$ onto the y-axis. That is to say $X = (X_1, X_2)$ $\pi_i(X) = X_i$ for $i = 1, 2$, where we equate the images with real numbers. In Section 5, however, we use projections $\pi_{\boldsymbol{X}}$, whose images we consider to be in $\mathbb{R}^2$.

Let $X \subset \mathbb{R}^n$. We call $f : X \to \mathbb{R}^n$ an L-bi-Lipschitz mapping for $L > 0$, if for all $x, y \in X$ we have

$$\frac{|x-y|}{L} \le |f(x) - f(y)| \le L|x-y|.$$

2. Extension from the boundary of the square

In this section our aim is to prove the following extension theorem, which will allow us to construct homeomorphisms from boundary values and gives us a useful control on their modulars.

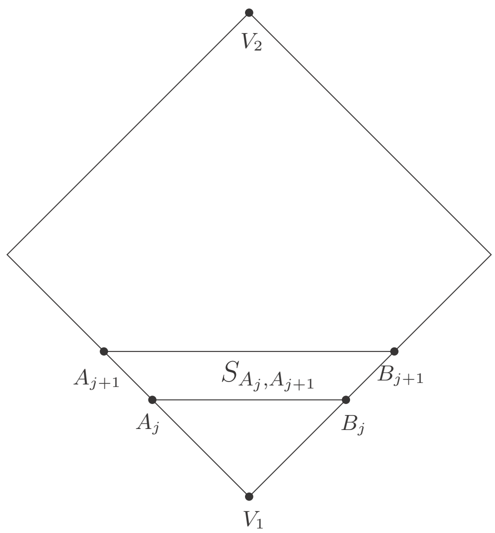


FIGURE 1. The square $\mathcal{S}_{r_0}$ and the strip S.

Theorem 2.1. *Let Φ be a Δ_2-Young function. Then there exists a constant $C > 0$ depending only on Φ such that for any $r_0 > 0$ and any piecewise-linear and one-to-one function $\varphi : \partial\mathcal{S}_{r_0} \to \mathbb{R}^2$ we can find a finitely piecewise-affine homeomorphism $h : \mathcal{S}_{r_0} \to \mathbb{R}^2$ such that*

$$\fint_{\mathcal{S}_{r_0}} \Phi(|Dh|)\, d\mathcal{L}^2 \leq C \fint_{\partial\mathcal{S}_{r_0}} \Phi(|D_\tau\varphi|)\, d\mathcal{H}^1 \,. \tag{2.1}$$

Proof. *Step 1: Choosing good corners.*

We may assume that for $V_1 = (0, -r_0)$ and $V_2 = (0, r_0)$ the derivative of φ does not accumulate too much around these points, i.e. we assume

$$\fint_{\mathcal{B}(V_i,r)\cap\partial\mathcal{S}_{r_0}} \Phi(|D_\tau\varphi|)\, d\mathcal{H}^1 \leq 6 \fint_{\partial\mathcal{S}_{r_0}} \Phi(|D_\tau\varphi|)\, d\mathcal{H}^1 \text{ for all } r \in (0, r_0),\ i \in \{1, 2\}\,. \tag{2.2}$$

The reason we may assume this is as follows. Put

$$\mathcal{A} := \left\{ P \in \partial\mathcal{S}_{r_0} : \exists\, r \in (0, r_0) : \fint_{\mathcal{B}(P,r)\cap\partial\mathcal{S}_{r_0}} \Phi(|D_\tau\varphi|) > 6 \fint_{\partial\mathcal{S}_0} \Phi(|D_\tau\varphi|) \right\}. \tag{2.3}$$

and find a disjoint system of balls $\mathcal{B}_j = \mathcal{B}(x_j, r_j)$ for which the integral estimate in (2.3) holds and $\bigcup_j 3\mathcal{B}_j \supset \mathcal{A}$. Then, as the balls are disjoint, we have,

$$\mathcal{H}^1(\mathcal{A}) \leq 6 \sum_j r_j \leq \sum_j \frac{6r_0}{6} \frac{\int_{\mathcal{B}_j\cap\partial\mathcal{S}_{r_0}} \Phi(|D_\tau\varphi|)}{\int_{\partial\mathcal{S}_{r_0}} \Phi(|D_\tau\varphi)} \leq r_0.$$

Now,

$$\mathcal{H}^1(\partial\mathcal{S}_{r_0}) = 4\sqrt{2}r_0 > 2r_0 \geq 2\mathcal{H}^1(\mathcal{A}),$$

which means that there are two points opposite one another neither of which belong to $\mathcal{A}$. Using a bi-Lipschitz transformation we may relocate these points onto V_1 and V_2 respectively.

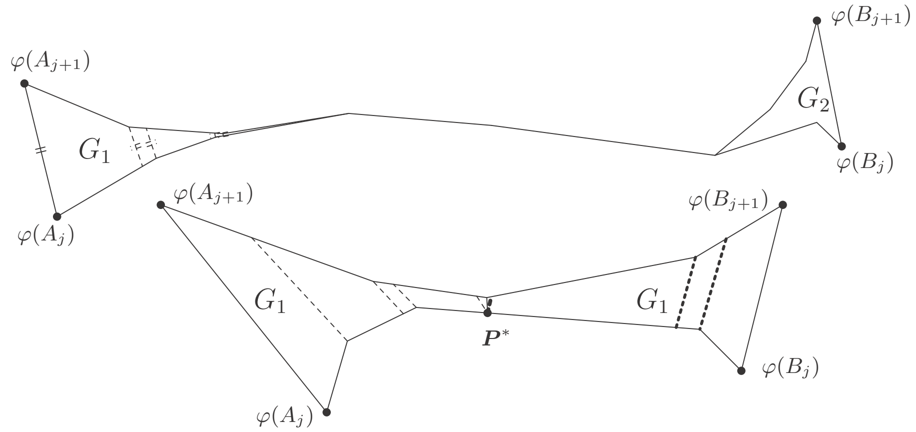


FIGURE 2. Two types of strips.

Step 2: Shortest curves between images of opposing points and 'vertical' segments. We will proceed to take advantage of some observations on paths joining opposing points on the square. Firstly let $x \in [-r_0, 0), y \in (-r_0, r_0)$ and $A = (x, y) \in \partial \mathcal{S}_{r_0}$, then we say that $B \in \partial \mathcal{S}_{r_0}$ opposes A if $B = (-x, y)$ as in Figure 1.

The details of the proof are very similar to those in [12] and [22]. Therefore we will present our argument briefly, tracing the steps of the proof in [12] and emphasising the differences between our proof and those mentioned to aid the reader follow our calculations. The following claim is [12, Theorem 2.1, Step 2, 3].

Lemma 2.2. *Let $A, B \in \partial \mathcal{S}_{r_0}$ be opposite and let Ω be the bounded component of $\mathbb{R}^2 \setminus \varphi(\partial \mathcal{S}_{r_0})$. Then there exists a uniquely determined shortest path $\tilde{\gamma} : [0, 1] \to \overline{\Omega}$, with $\tilde{\gamma}(0) = \phi(A)$ and $\tilde{\gamma}(1) = \phi(B)$. Further $\tilde{\gamma}$ is finitely piecewise-linear on $(0, 1)$ and $|\tilde{\gamma}'|$ is constant almost everywhere.*

In the following, given a point $(x_0, y_0) = A \in \partial \mathcal{S}_{r_0}, x_0 < 0$ and B its opposing point, we will define the mapping γ_A as the constant-speed parametrisation from the segment AB of the shortest path connecting $\varphi(A)$ and $\varphi(B)$ in $\overline{\Omega}$, that is

$$\gamma_A(x, y_0) = \tilde{\gamma}\left(\frac{x + r_0 - |y_0|}{2(r_0 - |y_0|)}\right),$$

with $\tilde{\gamma}$ taken from Lemma 2.2

We will also make use of segments in the image, referred to in [12] as 'vertical'. We shall now explain what these are. Given two distinct points $A_j, A_{j+1} \in \partial \mathcal{S}_{r_0}$ such that φ is linear on the segments $[A_j A_{j+1}], [B_j B_{j+1}] \subset \partial \mathcal{S}_{r_0}$, we will call the strip (depicted in Figure 2)

$$S = S_{A_j, A_{j+1}} = \text{co}\{A_j, B_j, A_{j+1}, B_{j+1}\}.$$

Writing $\gamma_j = \gamma_{A_j}$ for short, denote the set,

$$H = \varphi([A_j A_{j+1}] \cup [B_j B_{j+1}]) \cup \gamma_j([A_j B_j]) \cup \gamma_{j+1}([A_{j+1} B_{j+1}]),$$

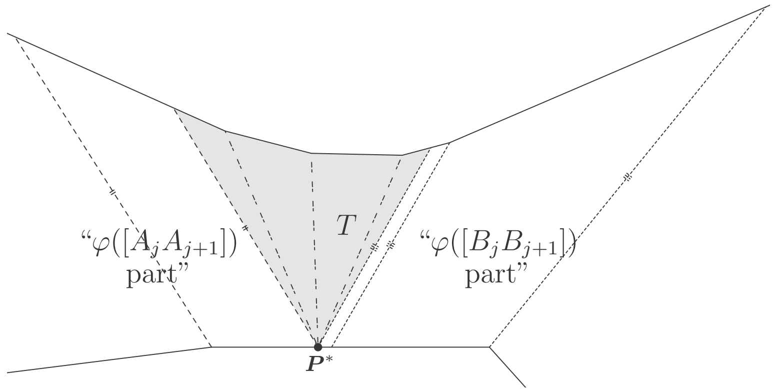


FIGURE 3. 'Vertical' segments near the centre of F_S.

then $\mathbb{R}^2 \setminus H$ has at least one and at most two bounded components as depicted in Figure 2. See [12] for all details and a rigorous argument. We denote the bounded components as G_1, G_2 (where G_2 is possibly empty). We will call the image of a strip S, $F_S = H \cup G_1 \cup G_2$. It holds that F_S is closed for all strips.

Given a strip S and the corresponding F_S we wish to define when a segment in F_S is 'vertical'. Firstly any point $(x, y) \in \gamma_j([A_jB_j]) \cap \gamma_{j+1}([A_{j+1}B_{j+1}])$ is a 'vertical' segment. Secondly, if $\mathbb{R}^2 \setminus H$ has two non-empty bounded components G_1 and G_2 with $[A_jA_{j+1}] \subset \partial G_1$ and $[B_jB_{j+1}] \subset \partial G_2$ then a 'vertical' segment in G_1 is a segment parallel with $\varphi([A_jA_{j+1}])$ and a 'vertical' segment in G_2 is a segment parallel with $\varphi([B_jB_{j+1}])$ such that the distinct endpoints of the segments lie in H. This is depicted in Figure 2

Finally let $G_1 \neq G_2 = \emptyset$. Either $\varphi([A_jA_{j+1}])$ and $\varphi([B_jB_{j+1}])$ are not parallel and we will discuss further or they are parallel and all 'vertical' segments are those segments parallel with $\varphi([A_jA_{j+1}])$ having distinct endpoints in ∂F_S.

In the case when $\varphi([A_jA_{j+1}])$ and $\varphi([B_jB_{j+1}])$ are not parallel (as depicted in Figure 3) then we clearly have that the angle between them is less than π because otherwise $G_1 = \emptyset$. Therefore there must be some point $\boldsymbol{P}^*$ such that segments starting at $\boldsymbol{P}^*$, parallel to $\varphi([A_jA_{j+1}])$ and those parallel to $\varphi([B_jB_{j+1}])$ meet H at precisely two points, one of these being $\boldsymbol{P}^*$ and the other point being on the opposite side of F_S. Since there are in fact many such points $\boldsymbol{P}^*$ (the preimage of such points in φ is a segment) we take the point at the centre of this segment. Any segment starting at $\boldsymbol{P}^*$ lying between the two segments parallel to $\varphi([A_jA_{j+1}])$ and $\varphi([B_jB_{j+1}])$ is called 'vertical' (i.e. segments in the polygon T from Figure 3). Now we can consider the rest of G_1, i.e. $G_1 \setminus T$, which has two components, the "$\varphi([A_jA_{j+1}])$ part" and the "$\varphi([B_jB_{j+1}])$ part". We treat these parts as we treated G_1 and G_2 before, i.e. segments parallel to $\varphi([A_jA_{j+1}])$ resp. $\varphi([B_jB_{j+1}])$ are called 'vertical'.

Now we can select those 'vertical' segments, which we will use. Let S be a given strip. We define points $\boldsymbol{P}_i \in \gamma_j([A_jB_j])$ and $\boldsymbol{R}_i \in \gamma_{j+1}([A_{j+1}B_{j+1}])$ such that they are either a vertex of γ_j or can be connected with a vertex of the other set by a 'vertical' segment (in the case where $G_2 = \emptyset$ we use only those segments which connect $\boldsymbol{P}^*$ with another vertex in T). Now on any strip we have corresponding $P_i = \gamma_j^{-1}(\boldsymbol{P}_i)$

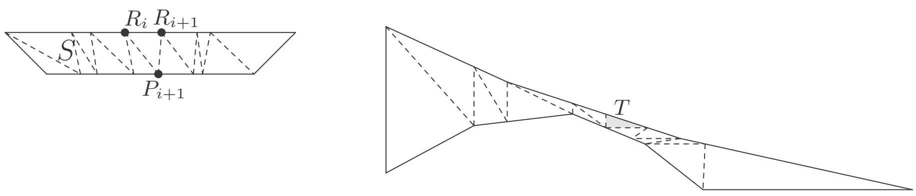


FIGURE 4. Triangular mesh induced in S by triangular mesh created in F_S.

and $R_i = \gamma_{j+1}^{-1}(\boldsymbol{R}_i)$ which may not be distinct for every i, but there are only finitely many of them for each j.

The following claim can be found in [12, Theorem 2.1, Step 4,5,6].

Lemma 2.3. *Let $S = S_{A_j,A_{j+1}}$ be the strip determined by A_j, A_{j+1} with opposing points B_j, B_{j+1}. Then all 'vertical' segments $[\boldsymbol{PR}]$ lie in F_S, the set referred to as the image of the strip S, and*

$$\mathcal{H}^1([\boldsymbol{PR}]) \leq \max\{\mathcal{H}^1\big(\varphi([A_jA_{j+1}])\big), \mathcal{H}^1\big(\varphi([B_jB_{j+1}])\big)\}.$$

Step 3: Creating a triangular mesh in the image and transferring it to the pre-image. Now for our given φ defined on $\partial\mathcal{S}_{r_0}$ we will define a mesh of horizontal lines in $\mathcal{S}_{r_0}$ which we will use to create our extension. The mapping φ is piecewise-linear and therefore its tangential, $D_\tau\varphi$, exists everywhere except the corners of the squares and a finite set. Index this set as follows $\tilde{A}_0 = V_1, \tilde{A}_1, \tilde{A}_2, \dots \tilde{A}_i = (-r_0, 0), \dots \tilde{A}_{\tilde{k}} = V_2$ where for all j the first coordinate of the points $\tilde{A}_j$ is non-positive, the second coordinate is increasing in j and $D_\tau\varphi$ does not exist either at A_j or at its opposing B_j. We now refine our mesh adding a finite number of points in between $\tilde{A}_1$ and $\tilde{A}_{\tilde{k}-1}$ to get a set $A_0, A_1 \dots, A_k$, with $A_0 = \tilde{A}_0$, $A_1 = \tilde{A}_1$, $A_{k-1} = \tilde{A}_{\tilde{k}-1}$ and $A_k = \tilde{A}_{\tilde{k}}$ such that

$$\begin{gathered}\mathcal{H}^1([A_jA_{j+1}]) \leq \mathcal{H}^1([A_0A_j])\\ \mathcal{H}^1([A_jA_{j+1}]) \leq \mathcal{H}^1([A_{j+1}A_k])\\ \max\{\mathcal{H}^1(\varphi([A_jA_{j+1}])), \mathcal{H}^1(\varphi([B_jB_{j+1}]))\} \leq \min\{\mathcal{H}^1(\gamma_j([A_jB_j])), \mathcal{H}^1(\gamma_{j+1}([A_{j+1}B_{j+1}]))\}\end{gathered} \tag{2.4}$$

for $j = 1, 2, \dots, k-2$, where γ_j is the shortest path defined on $[A_jB_j]$ connecting $\varphi(A_j)$ and $\varphi(B_j)$ in the sense of Lemma 2.2.

Having selected our horizontal mesh we separate each of the strips S_j where $S_j = \text{co}\{A_j, A_{j+1}, B_j, B_{j+1}\}$, $j = 1, 2, \dots, k-2$ as follows. On $[A_jB_j]$ define P_i and corresponding R_i on $[A_{j+1}B_{j+1}]$ such that P_i and R_i are connected by 'vertical' segments and outside these points the functions γ_j and γ_{j+1} are both derivable.

We define an initial mapping $\tilde{h}$ on $\mathcal{S}_{r_0}$ which will not be injective but will satisfy our bound (2.1) and then later we will slightly refine it to make it injective whilst maintaining (2.1).

Define $\tilde{h} = \varphi$ on $\partial\mathcal{S}_{r_0}$, $\tilde{h} = \gamma_j$ on $[A_jB_j]$ for $j = 1, 2, \dots k-1$. Now fix j and the strip S_j and consider the linear mappings, which maps $[P_iR_i]$ onto $[\boldsymbol{P}_i\boldsymbol{R}_i]$ and the linear mappings, which map the segments $[P_iR_{i+1}]$ onto $[\boldsymbol{P}_i\boldsymbol{R}_{i+1}]$. This gives us a triangular mesh in S_j and linear maps on each side of the triangles. Define $\tilde h$ on the triangular mesh in S_j as these linear mappings, then $\tilde h$ is continuous and on each triangle $P_iP_{i+1}R_{i+1}$, resp. $P_{i+1}R_iR_{i+1}$ and there exists exactly one affine map which extends $\tilde h$ onto the convex hull of the triangle. This is depicted in Figure 4.

We still must define $\tilde h$ on the triangles $\operatorname{co}\{A_0, A_1, B_1\}$ and $\operatorname{co}\{A_k, A_{k-1}, B_{k-1}\}$. The mapping γ_1 is either differentiable on $[A_1B_1]$ or there is exactly one point (this point is $P_1 \in [A_1B_1]$) where its derivative does not exist. If no such P_1 exists then φ and γ_1 determine a unique affine map by prescribing continuous linear boundary values. Otherwise we take the segment $[A_0P_1]$, which divides our triangle into two triangles. On each of these triangles φ and γ_1 determine a unique affine map by prescribing continuous linear boundary values.

Step 4: Estimating the derivatives of $\tilde h$.

We will now need to estimate $|D_1\tilde h|$ and $|D_2\tilde h|$. Firstly take the triangle $A_0A_1B_1$ (the argument on $A_kA_{k-1}B_{k-1}$ is the same) and call it $\mathcal{T}$. We have that

$$\mathcal{H}^1(\varphi([A_1B_1])) \leq \mathcal{H}^1(\varphi([A_0A_1])) + \mathcal{H}^1(\varphi([A_0B_1])) \text{ and } \mathcal{H}^1([A_1B_1]) = \sqrt{2}\mathcal{H}^1([A_0A_1]).$$

Since we consider linear mappings whose derivative is constant we have $|D_1\tilde h(x,y)| = |D_\tau\gamma_1|$. In the following estimates we can take any $X \in [A_0A_1] \setminus \{A_0, A_1\}$ and any $Y \in [A_0B_1] \setminus \{A_0, B_1\}$, we use $\mathcal{H}^1([A_1B_1]) = \sqrt{2}\mathcal{H}^1([A_0A_1])$ and the fact that γ_1 is the shortest curve to get,

$$\begin{aligned}
|D_1\tilde h(x,y)| = |D_\tau\gamma_1| &= \frac{\mathcal{H}^1(\gamma_1([A_1B_1]))}{\mathcal{H}^1([A_1B_1])}\\
&= \sqrt{2}\frac{\mathcal{H}^1(\gamma_1([A_1B_1]))}{\mathcal{H}^1([A_0A_1])}\\
&\leq \sqrt{2}\frac{\mathcal{H}^1(\varphi([A_0A_1])) + \mathcal{H}^1(\varphi([A_0B_1]))}{\mathcal{H}^1([A_0A_1])}\\
&= \sqrt{2}\big(|D_\tau\varphi(X)| + |D_\tau\varphi(Y)|\big)\\
&\leq 2\sqrt{2}\fint_{[A_0A_1]\cup[A_0B_1]} |D_\tau\varphi(t)|\,d\mathcal{H}^1(t).
\end{aligned}$$

If we call $C_1 = \frac{1}{2}(A_1 + B_1)$ the midpoint of A_1 and B_1, then we can estimate

$$|D_2\tilde h| = \sqrt{2}\fint_{[A_0A_1]} D_\tau\varphi + \fint_{[A_1C_1]} D_1 h \leq \sqrt{2}\frac{\mathcal{H}^1\big(\varphi([A_0A_1])\big) + \mathcal{H}^1\big(\varphi([A_0B_1])\big)}{\mathcal{H}^1([A_0A_1])} + |D_1\tilde h|.$$

This gives altogether, for almost all $(x,y) \in \operatorname{co}\mathcal{T}$, that

$$|D\tilde h(x,y)| \leq C|D_1\tilde h(x,y)| + C|D_2\tilde h(x,y)| \leq C\fint_{[A_0A_1]\cup[A_0B_1]} |D_\tau\varphi(t)|\,d\mathcal{H}^1(t)\ , \quad (2.5)$$

where C is an absolutely fixed geometrical constant. Now we apply Φ, take advantage of the Δ_2 quality of Φ, use the Jensen inequality and then (2.2) to get for almost all

$(x,y) \in \operatorname{co}\mathcal{T}$ that,

$$\begin{aligned}
\Phi\big(|D\tilde{h}(x,y)|\big) &\leq \Phi\Big(C \fint_{[A_0A_1]\cup[A_0B_1]} |D_\tau\varphi(t)| d\mathcal{H}^1(t)\Big)\\
&\leq \tilde{C}\Phi\Big(\fint_{[A_0A_1]\cup[A_0B_1]} |D_\tau\varphi(t)| d\mathcal{H}^1(t)\Big)\\
&\leq \tilde{C}\fint_{[A_0A_1]\cup[A_0B_1]} \Phi\big(|D_\tau\varphi(t)|\big) d\mathcal{H}^1(t)\\
&\leq 6\tilde{C}\fint_{\partial\mathcal{S}_{r_0}} \Phi\big(|D_\tau\varphi(t)|\big) d\mathcal{H}^1(t)\ ,
\end{aligned} \tag{2.6}$$

where $\tilde{C} = C_2^k$, given that $C \leq 2^k$, has been derived from the Δ_2 condition. So there exists a constant C depending only on Φ such that the integral over our triangle $\operatorname{co}\mathcal{T}$

$$\int_{\operatorname{co}\mathcal{T}} \Phi\big(|D\tilde{h}(x,y)|\big) \leq C|\operatorname{co}\mathcal{T}| \fint_{\partial\mathcal{S}_{r_0}} \Phi\big(|D_\tau\varphi(t)|\big) d\mathcal{H}^1(t).$$

We now move our attention to a strip S_j, which for the sake of simplicity is bellow the x-axis (the argument on the half of the square above the axis is symmetric). Firstly we will estimate $D_1\tilde{h}$ and then $D_2\tilde{h}$ later. Almost all points in S_j belong to the interior of a triangle $P_iP_{i+1}R_{i+1}$, or $P_{i+1}R_iR_{i+1}$. Since $\tilde{h}$ is affine on each triangle, the derivative is constant inside each triangle. Recall that $|D_\tau\gamma_j|$ is constant on $[A_jB_j]$. Since one side of each triangle is horizontal and lies on either $[A_jB_j]$ or $[A_{j+1}B_{j+1}]$ then

$$|D_1\tilde{h}(x,y)| = |D_\tau\gamma_j| = \frac{\mathcal{H}^1(\gamma([A_jB_j]))}{\mathcal{H}^1([A_jB_j])}$$

for any point inside a triangle of type $P_iP_{i+1}R_{i+1}$ and

$$|D_1\tilde{h}(x,y)| = |D_\tau\gamma_{j+1}| = \frac{\mathcal{H}^1(\gamma([A_{j+1}B_{j+1}]))}{\mathcal{H}^1([A_{j+1}B_{j+1}])}$$

for any point inside a triangle of type $P_{i+1}R_iR_{i+1}$. For almost all points $(x,y) \in S_j$ we have

$$|D_1\tilde{h}(x,y)| \leq \max\Big\{\frac{\mathcal{H}^1(\gamma_j([A_jB_j]))}{\mathcal{H}^1([A_jB_j])}, \frac{\mathcal{H}^1(\gamma_{j+1}([A_{j+1}B_{j+1}]))}{\mathcal{H}^1([A_{j+1}B_{j+1}])}\Big\}. \tag{2.7}$$

Notice however that $\sqrt{2}\mathcal{H}^1([A_jB_j]) \geq \mathcal{H}^1([A_jA_0]\cup[A_0B_j])$ and since γ_j is the shortest curve it is not longer than the image of φ on $[A_jA_0]\cup[A_0B_j]$. Use this and (2.2) to get

$$\begin{aligned}
\frac{\mathcal{H}^1(\gamma_j([A_jB_j]))}{\mathcal{H}^1([A_jB_j])} &\leq \sqrt{2}\frac{\mathcal{H}^1(\varphi([A_jA_0]\cup[A_0B_j]))}{\mathcal{H}^1([A_jA_0]\cup[A_0B_j])}\\
&= \sqrt{2}\fint_{[A_jA_0]\cup[A_0B_j]} |D_\tau\varphi(t)| d\mathcal{H}^1(t)\\
&\leq C\fint_{\partial\mathcal{S}_{r_0}} |D_\tau\varphi(t)| d\mathcal{H}^1(t)\ .
\end{aligned} \tag{2.8}$$

Combining (2.8) applied to j and $j+1$ with (2.7) gives

$$|D_1\tilde{h}(x,y)| \leq C\fint_{\partial\mathcal{S}_{r_0}} |D_\tau\varphi(t)| d\mathcal{H}^1(t) \tag{2.9}$$

almost everywhere in S_j. Now use the Jensen inequality to get

$$\begin{aligned}\Phi\big(|D_1\tilde{h}(x,y)|\big) &\leq \Phi\Big(C \fint_{\partial\mathcal{S}_{r_0}} |D_\tau\varphi(t)|d\mathcal{H}^1(t)\Big)\\ &\leq \tilde{C}\Phi\Big(\fint_{\partial\mathcal{S}_{r_0}} |D_\tau\varphi(t)|d\mathcal{H}^1(t)\Big)\\ &\leq \tilde{C}\fint_{\partial\mathcal{S}_{r_0}} \Phi\big(|D_\tau\varphi(t)|\big)d\mathcal{H}^1(t)\end{aligned} \tag{2.10}$$

for almost all $(x,y) \in S_j$.

We must now estimate $|D_2\tilde{h}(x,y)|$. Intuitively, since the longest 'vertical' segment in F_S the image of the strip S is at most $\delta = \max\{\mathcal{H}^1(\varphi([A_jA_{j+1}])), \mathcal{H}^1(\varphi([B_jB_{j+1}]))\}$ (see (2.4)) and the strip S itself has a width of say ε and the preimage of 'vertical' segments in $\tilde{h}$ are roughly vertical, we should be able to bound $|D_2\tilde{h}(x,y)|$ by $C|D_1\tilde{h}(x,y)| + C\frac{\delta}{\varepsilon}$. The details can be found in [12][Theorem 2.1, Step9] and the following Lemma is the penultimate estimate of that step.

Lemma 2.4. *Let S_j be a strip in $\mathcal{S}_{r_0}$ and the piecewise-affine $\tilde{h}$ defined by γ_j from Lemma 2.2 and φ on a mesh satisfying* (2.4) *and $\boldsymbol{P}_i, \boldsymbol{R}_i$ as defined above. Then*

$$\begin{aligned}|D_2\tilde{h}(x,y)| \leq C\Bigg(&\max\left\{\frac{\mathcal{H}^1\big(\gamma_j([A_jB_j])\big)}{\mathcal{H}^1([A_jB_j])}, \frac{\mathcal{H}^1\big(\gamma_{j+1}([A_{j+1}B_{j+1}])\big)}{\mathcal{H}^1([A_{j+1}B_{j+1}])}\right\}\\ &+\max\left\{\frac{\mathcal{H}^1\big(\varphi([A_jA_{j+1}])\big)}{\mathcal{H}^1([A_jA_{j+1}])}, \frac{\mathcal{H}^1\big(\varphi([B_jB_{j+1}])\big)}{\mathcal{H}^1([B_jB_{j+1}])}\right\}\Bigg).\end{aligned}$$

We have already estimated the first term in the previous to get

$$\max\left\{\frac{\mathcal{H}^1\big(\gamma_j([A_jB_j])\big)}{\mathcal{H}^1([A_jB_j])}, \frac{\mathcal{H}^1\big(\gamma_{j+1}([A_{j+1}B_{j+1}])\big)}{\mathcal{H}^1([A_{j+1}B_{j+1}])}\right\} \leq C\fint_{\partial\mathcal{S}_{r_0}} |D_\tau\varphi(t)|d\mathcal{H}^1(t).$$

The second term can be estimated as follows

$$\max\left\{\frac{\mathcal{H}^1\big(\varphi([A_jA_{j+1}])\big)}{\mathcal{H}^1([A_jA_{j+1}])}, \frac{\mathcal{H}^1\big(\varphi([B_jB_{j+1}])\big)}{\mathcal{H}^1([B_jB_{j+1}])}\right\} \leq 2\fint_{[A_jA_{j+1}]\cup[B_jB_{j+1}]} |D_\tau\varphi(t)|d\mathcal{H}^1(t).$$

Together that means that

$$|D_2\tilde{h}(x,y)| \leq \fint_{[A_jA_{j+1}]\cup[B_jB_{j+1}]} |D_\tau\varphi(t)|d\mathcal{H}^1(t) + C\fint_{\partial\mathcal{S}_{r_0}} |D_\tau\varphi(t)|d\mathcal{H}^1(t). \tag{2.11}$$

Now we can apply Φ, use the fact that Φ satisfies Δ_2 and use the Jensen inequality to get

$$\begin{aligned}\Phi(|D_2\tilde{h}(x,y)|) &\leq \Phi\Big(C\fint_{\partial\mathcal{S}_{r_0}} |D_\tau\varphi(t)|d\mathcal{H}^1(t) + 2\fint_{[A_jA_{j+1}]\cup[B_jB_{j+1}]} |D_\tau\varphi(t)|d\mathcal{H}^1(t)\Big)\\ &\leq \tilde{C}\Phi\Big(\fint_{\partial\mathcal{S}_{r_0}} |D_\tau\varphi(t)|d\mathcal{H}^1(t)\Big) + \tilde{C}\Phi\Big(\fint_{[A_jA_{j+1}]\cup[B_jB_{j+1}]} |D_\tau\varphi(t)|d\mathcal{H}^1(t)\Big)\\ &\leq \tilde{C}\fint_{\partial\mathcal{S}_{r_0}} \Phi\big(|D_\tau\varphi(t)|\big)d\mathcal{H}^1(t) + \tilde{C}\fint_{[A_jA_{j+1}]\cup[B_jB_{j+1}]} \Phi\big(|D_\tau\varphi(t)|\big)d\mathcal{H}^1(t)\end{aligned} \tag{2.12}$$

for almost all $(x,y) \in S_j$. So combining (2.10) and (2.12) using the Δ_2 condition, we get a C such that

$$\Phi\big(|D\tilde{h}(x,y)|\big) \leq C \fint_{\partial\mathcal{S}_{r_0}} \Phi\big(|D_\tau\varphi(t)|\big) d\mathcal{H}^1(t) + C \fint_{[A_jA_{j+1}]\cup[B_jB_{j+1}]} \Phi\big(|D_\tau\varphi(t)|\big) d\mathcal{H}^1(t).$$

Now integrate over S_j to get

$$\begin{aligned}\int_{S_j} \Phi\big(|D\tilde{h}(x,y)|\big) d\mathcal{L}^2(x,y) \leq C\mathcal{L}^2(S_j) \fint_{\partial\mathcal{S}_{r_0}} \Phi\big(|D_\tau\varphi(t)|\big) d\mathcal{H}^1(t)\\ + Cr_0 \int_{[A_jA_{j+1}]\cup[B_jB_{j+1}]} \Phi\big(|D_\tau\varphi(t)|\big) d\mathcal{H}^1(t).\end{aligned}$$

Summing over all strips

$$\begin{aligned}\int_{\mathcal{S}_{r_0}} \Phi\big(|D\tilde{h}(x,y)|\big) d\mathcal{L}^2(x,y) \leq & Cr_0^2 \fint_{\partial\mathcal{S}_{r_0}} \Phi\big(|D_\tau\varphi(t)|\big) d\mathcal{H}^1(t)\\ & + Cr_0 \int_{\partial\mathcal{S}_{r_0}} \Phi\big(|D_\tau\varphi(t)|\big) d\mathcal{H}^1(t).\end{aligned}$$

Dividing by r_0^2 gives our required (2.1)-type estimate for our initial mapping $\tilde{h}$. It now remains to alter $\tilde{h}$ so that it is injective, while maintaining all other properties.

Step 5: *Defining h to obtain a homeomorphism.*

Here we repeat the argument from [12, Theorem 2.1, Step10]. It still remains to slightly alter our initial mapping $\tilde{h}$ in order for it to be injective, which it is not if any of the curves meet the image of the boundary or if they meet each other at some point. We can use a construction similar to the one we introduce in Lemma 3.1 to move those endpoints of segments of two lines which coincide at these endpoints. Seeing it suffices to move the endpoints by some arbitrarily small amount, the lengths of the curves involved are only longer by some insignificantly small factor and so the estimates above also hold for h with a coefficient at most larger only by a factor of C_2. □

We may also observe that the following modular estimates hold for our extension theorem.

Corollary 2.5. *Let $r_0 > 0$ and $\varphi : \partial\mathcal{S}_{r_0} \to \mathbb{R}^2$ be a piecewise-linear and one-to-one function. For all $\lambda > 0$ the following estimate holds for the finitely piecewise-affine homeomorphism $h : \mathcal{S}_{r_0} \to \mathbb{R}^2$ found in Theorem 2.1,*

$$\fint_{\mathcal{S}_{r_0}} \Phi(\lambda|Dh|)\, d\mathcal{L}^2 \leq C \fint_{\partial\mathcal{S}_{r_0}} \Phi(\lambda|D_\tau\varphi|)\, d\mathcal{H}^1\,. \tag{2.13}$$

Proof. It suffices to multiply the estimates (2.5), (2.9) and (2.11) by λ on both sides and then use these new estimates in (2.6), (2.10) and (2.12). □

3. Extension in the degenerate case $|Df(c)| \neq 0$ but $J_f(c) = 0$

Our first lemma will be a simple observation that when one creates a polygon inside another by following the boundary of the first, the perimeter is not increased significantly (see Figure 5).

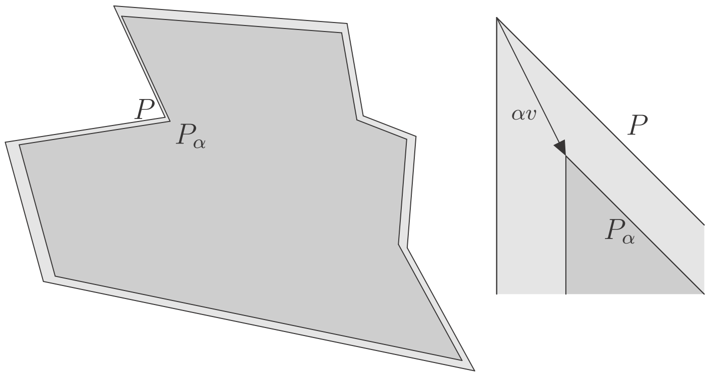


FIGURE 5. A simple observation that following the boundary inside a polygon does not increase the perimeter much.

Lemma 3.1. *Let $\mathcal{D} = \{\boldsymbol{D}_1, \boldsymbol{D}_2, \dots, \boldsymbol{D}_m = \boldsymbol{D}_0\} \subset \mathbb{R}^2$ be a set of m distinct points such that*

$$[\boldsymbol{D}_{j-1}\boldsymbol{D}_j] \cap [\boldsymbol{D}_{k-1}\boldsymbol{D}_k] \subset \mathcal{D} \text{ if } j \neq k,$$

i.e. no two segments intersect each other, except for adjacent segments which share endpoints. We will call the polygon $\bigcup_{j=1}^m [\boldsymbol{D}_{j-1}\boldsymbol{D}_j] = P$. The mapping $\varphi : \partial Q \to P$ will denote the piecewise-linear constant-speed parametrization of P from ∂Q, where Q is a square of side length $r \in (0, 1]$. Let $\varepsilon > 0$ then there exists a polygon (which we will denote P_α) lying entirely inside P, whose length differs from $\mathcal{H}^1(P)$ by at most ε.

Proof. We denote the inside of P as $\mathcal{S}$, that is to say $\mathcal{S}$ is the bounded component of $\mathbb{R}^2 \setminus P$. Clearly φ is one-to-one. Therefore denote $D_i = \varphi^{-1}(\boldsymbol{D}_i)$ for all $\boldsymbol{D}_i \in \mathcal{D}$ and $\mathbb{D} = \{D_1, D_2, \dots, D_m\}$, where we assume that the vertices of Q already belong to $\mathbb{D}$. For our proof we will require the use of a 'normal' vector at $(x, y) \in \partial Q$. For all points $(x, y) \in \partial Q \setminus \mathbb{D}$ we define a 'normal' vector as $v_{(x,y)}$, the unit vector perpendicular to $D_\tau \varphi(x, y)$ such that for some small $\eta_0 > 0$ we have $\varphi(x, y) + \eta v_{(x,y)} \in \mathcal{S}$ for all $\eta \in (0, \eta_0)$.

Then for $D_i \in \mathbb{D}$ we take the average of the two adjacent vectors, i.e.

$$v_i = \lim_{s \to 0^+} \frac{\int_{B(D_i,s)\cap\partial Q} v_u d\mathcal{H}^1(u)}{|\int_{\mathcal{B}(D_i,s)\cap\partial Q} v_u d\mathcal{H}^1(u)|} \ . \tag{3.1}$$

We shall choose a small number α_0 as follows. Using the notation $\boldsymbol{D}_{m+1} = \boldsymbol{D}_1$ we calculate

$$\alpha_0 = \frac{1}{4} \min\{\operatorname{dist}\big(\boldsymbol{D}_i, P \setminus \big([\boldsymbol{D}_{i-1}\boldsymbol{D}_i] \cup [\boldsymbol{D}_i\boldsymbol{D}_{i+1}]\big),\ i = 1, \dots, m\} \tag{3.2}$$

unless this number is larger than 1, in which case $\alpha_0 = 1$. Notice that if we choose any $\alpha \in (0, \alpha_0)$ for α_0 the segments $[(\boldsymbol{D}_i + \alpha v_i)(\boldsymbol{D}_{i+1} + \alpha v_{i+1})]$ are pairwise disjoint (except for the endpoints of adjacent segments) and they lie entirely in $\mathcal{S}$. Thus we

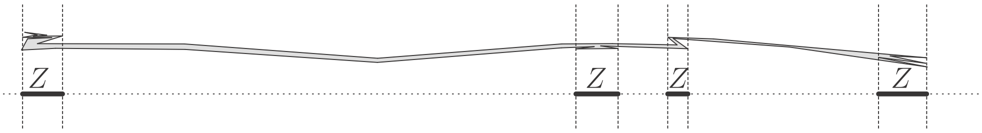


FIGURE 6. The set Z where the boundary “doubles up”.

construct and get a polygon

$$P_\alpha = \bigcup_{i=1}^{m} [(\boldsymbol{D}_i + \alpha v_i)(\boldsymbol{D}_{i+1} + \alpha v_{i+1})],$$

whose length is close to that of the original, i.e. differs from it by no more than $2\alpha m$, by simple use of the triangle inequality. It now suffices to choose $\alpha < \varepsilon(2m)^{-1}$ and our claim is proven. $\square$

In the following we will consider some geometrical properties of the following mappings and sets. We will have $r_0 > 0$ and Q will be the image of a 2-bi-Lipschitz mapping which is equal to an affine mapping on the triangles $\text{co}\{(0,0),(0,r_0),(r_0,0)\}$ and $\text{co}\{(r_0,r_0),(0,r_0),(r_0,0)\}$. Further we will have a $\varphi \in W^{1,1}(\partial Q, \mathbb{R}^2)$, which is piecewise-linear and one-to-one. Further we will assume that there exists an $M > 0$ and an $\varepsilon \ll \min\{r_0, Mr_0\}$ such that

$$\int_{\partial Q} \left| D_\tau \varphi(t) - \begin{pmatrix} M, 0 \\ 0, 0 \end{pmatrix} \tau \right| \, d\mathcal{H}^1(t) < \varepsilon. \tag{3.3}$$

We will denote $\mathcal{S}$ as the closure of the bounded component of $\mathbb{R}^2 \setminus \varphi(\partial Q)$.

The following lemma proves that the set of vertical lines intersecting $\varphi(\partial Q)$ at more than two points is very small. Compare with Figure 6.

Lemma 3.2. *Let $\varphi, Q, M, \varepsilon$ be as given in* (3.3)*. Use π_1 to denote the projection of $\mathbb{R}^2$ onto the x-axis. Call the set where a vertical line meets $\varphi(\partial Q)$ in more than two points*

$$Z = \{(\boldsymbol{Z}_1, \boldsymbol{Z}_2) \in \varphi(\partial Q);\ \mathcal{H}^0\big(\pi_1^{-1}(\boldsymbol{Z}_1) \cap \varphi(\partial Q)\big) > 2\}.$$

Then we have

$$\mathcal{H}^1(Z) < \frac{3\varepsilon}{2}.$$

Calling $P = \varphi(\partial Q)$ and taking P_α constructed in Lemma 3.1 we can find an α small enough that $\mathcal{H}^1(Z_{P_\alpha}) < 2\varepsilon$, where we define Z_{P_α} in the same way as Z, only replacing $\varphi(\partial Q)$ with P_α.

Proof. We clearly have that the set $\{x \in \mathbb{R};\ \mathcal{H}^0(\{x\} \times \mathbb{R} \cap \varphi(\partial Q)) = 1\}$ has 1 dimension measure 0, in fact it has at most two points. Indeed the set of lines intersecting $\varphi(\partial Q)$ in an odd number of points has 1-dimensional measure 0. Therefore almost all vertical lines intersecting $\varphi(\partial Q)$ do so, at 2 points or more. Notice that (3.3) easily gives that

$$\mathcal{H}^1(\varphi(\partial Q)) < 2M\mathcal{H}^1(\pi_1(Q)) + \varepsilon \tag{3.4}$$

and

$$\mathcal{H}^1\big(\pi_1(\varphi(\partial Q))\big) > M\mathcal{H}^1(\pi_1(Q)) - \varepsilon. \tag{3.5}$$

Hereby we get that

$$\mathcal{H}^1(\varphi(\partial Q)) - 2\mathcal{H}^1\big(\pi_1(\varphi(\partial Q))\big) < 3\varepsilon. \tag{3.6}$$

With respect to the fact that almost all lines intersect the polygon at an even number of points almost every vertical line intersecting Z must also intersect $\varphi(\partial Q)$ at a minimum of 4 points. This allows us to make the estimate

$$\mathcal{H}^1(Z) < \frac{3\varepsilon}{2}. \tag{3.7}$$

From Lemma 3.1 we know that $\mathcal{H}^1(P_\alpha) \leq \mathcal{H}^1(\varphi(\partial Q)) + 2m\alpha$, which we can combine with (3.4). It is evident by our construction that that we have

$$0 < \mathcal{H}^1\big(\pi_1(\varphi(\partial Q))\big) - \mathcal{H}^1\big(\pi_1(P_\alpha)\big) \leq 2\alpha,$$

which we can combine with (3.5). Therefore we get the following equation corresponding to (3.6) for P_α

$$\mathcal{H}^1(Z_{P_\alpha}) < \frac{3\varepsilon}{2} + 2(m+2)\alpha.$$

Repeating the estimate in (3.7), with α sufficiently small, we get

$$\mathcal{H}^1(Z_{P_\alpha}) < 2\varepsilon.$$

$\square$

Definition 3.3. *Let $r_0 > 0$ and Q be the image of a 2-bi-Lipschitz mapping which is equal to an affine mapping on* $\operatorname{co}\{(0,0),(0,r_0),(r_0,0)\}$ *and* $\operatorname{co}\{(r_0,r_0),(0,r_0),(r_0,0)\}$. *We will call a curve $\tilde{\gamma}$ in ∂Q direct if for every parametrization $\tilde{\gamma}^* = (\tilde{\gamma}_1^*, \tilde{\gamma}_2^*) : [0,1] \to \mathbb{R}^2$, $\tilde{\gamma}^*([0,1]) = \tilde{\gamma}$, we have that either $(\tilde{\gamma}_1^*)' \geq 0$ or $(\tilde{\gamma}_1^*)' \leq 0$ everywhere on* $(0,1)$.

Direct curves are such that the first coordinate function of their parametrizations are monotone. The significance of direct curves is that we will be able to easily estimate the length of the image of a direct curve using (3.3) as follows

$$\begin{aligned} H^1(\varphi(\tilde{\gamma})) &= \int_{\tilde{\gamma}} |D_\tau \varphi| d\mathcal{H}^1 \\ &\leq \int_{\tilde{\gamma}} \Big| \begin{pmatrix} M, 0 \\ 0, 0 \end{pmatrix} \tau \Big| d\mathcal{H}^1 + \varepsilon \\ &\leq M\mathcal{H}^1(\pi_1(\tilde{\gamma})) + \varepsilon. \end{aligned} \tag{3.8}$$

If we have two points $A = (A_1, A_2), B = (B_1, B_2) \in \partial Q$, ($\boldsymbol{A} = (\boldsymbol{A}_1, \boldsymbol{A}_2) = \varphi(A), \boldsymbol{B} = (\boldsymbol{B}_1, \boldsymbol{B}_2) = \varphi(B)$) we can simply take any curve $\tilde{\gamma}$ (we do not need that $\tilde{\gamma}$ is direct) in ∂Q having endpoints at A and B and estimate

$$\boldsymbol{A}_1 - \boldsymbol{B}_1 = \pi_1\Big(\int_{\tilde{\gamma}} D_\tau \varphi\Big) < M(A_1 - B_1) + \varepsilon,$$

and similarly

$$\boldsymbol{A}_1 - \boldsymbol{B}_1 = \pi_1\Big(\int_{\tilde{\gamma}} D_\tau \varphi\Big) > M(A_1 - B_1) - \varepsilon.$$

From this we get

$$\frac{\boldsymbol{A}_1 - \boldsymbol{B}_1}{M} - \frac{\varepsilon}{M} < A_1 - B_1 < \frac{\boldsymbol{A}_1 - \boldsymbol{B}_1}{M} + \frac{\varepsilon}{M}. \tag{3.9}$$

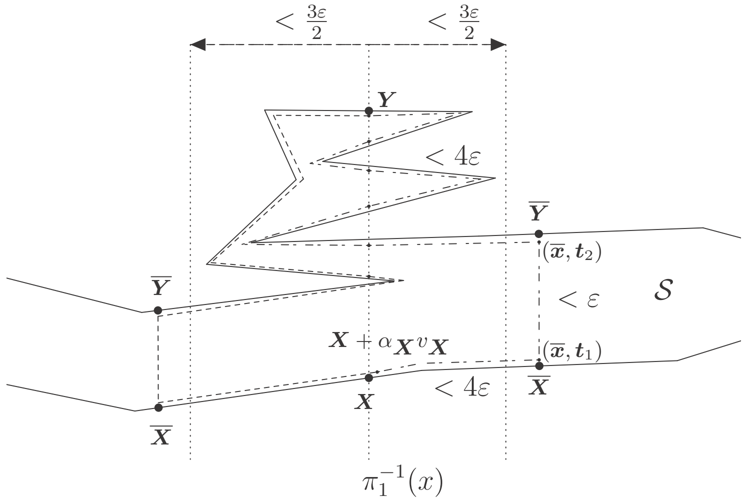


FIGURE 7. Following the boundary until we find a good point, where we can "bridge".

Lemma 3.4. *Let $\varphi, Q, M, \varepsilon$ be as given in* (3.3) *and $\mathcal{S}$ be the closure of the bounded component of $\mathbb{R}^2 \setminus \varphi(\partial Q)$. Then for any two distinct points $\boldsymbol{X} = (\boldsymbol{X}_1, \boldsymbol{X}_2), \boldsymbol{Y} = (\boldsymbol{Y}_1, \boldsymbol{Y}_2) \in \varphi(\partial Q)$ satisfying $\boldsymbol{X}_1 = \boldsymbol{Y}_1$ there exists a a one-to-one piecewise-linear curve of length no more than 10ε, which lies entirely inside $\mathcal{S}^\circ$ apart from its end-points, which are $\boldsymbol{X}$ and $\boldsymbol{Y}$.*

Proof. There are a number of cases which we will consider. The first is that the segment $[\boldsymbol{X}\boldsymbol{Y}]$ lies in $\mathcal{S}^\circ$ apart from its endpoints. In this case we can simply use this curve. We know that the difference in y-coordinate is small thanks to (3.3), that is

$$H^1([\boldsymbol{X}\boldsymbol{Y}]) = |\boldsymbol{X}_2 - \boldsymbol{Y}_2| < \varepsilon. \tag{3.10}$$

The second case is if there exists a path in $\varphi(\partial Q)$ connecting $\boldsymbol{X}$ and $\boldsymbol{Y}$, whose length is no more than 9ε. If this is true then we need only follow a path close to the boundary and we will get a path of length at most 10ε. To be explicit, consider the path consisting of those points in P_α corresponding to the points in our path on $\varphi(\partial Q)$. We use the segments generated by $\boldsymbol{X} + tv_{\boldsymbol{X}}$ and $\boldsymbol{Y} + tv_{\boldsymbol{Y}}$ to connect $\boldsymbol{X}$ and $\boldsymbol{Y}$ with this path. Since we choose

$$\alpha \leq \frac{\varepsilon}{2m+4}$$

we know that our path is not more than ε longer than the original path in $\varphi(\partial Q)$ and therefore less than 10ε.

If neither of the above cases hold then our plan is to do the following. We will find an appropriate pair of points close to $\boldsymbol{X}$ and $\boldsymbol{Y}$, which we will be able to connect with a vertical segment. We then connect this segment to our original points. Initially we will use a curve on $\varphi(\partial Q)$, which we afterwards move slightly inside $\mathcal{S}$ so that only the endpoints of the curve lie in $\varphi(\partial Q)$. Our strategy is illustrated in Figure 7.

As the second case does not hold, we know that there is no curve on $\varphi(\partial Q)$ connecting our two points $\boldsymbol{X}$, and $\boldsymbol{Y}$ with length shorter than 9ε. We will use this to prove that $\boldsymbol{X}, \boldsymbol{Y}$ are not too close to the left or right extreme of points of $\varphi(\partial Q)$. The same will hold for their preimages X, Y. This will allow us to show that the curves we use are direct and we will be able to use the estimates in (3.8).

To this end denote $l = \min \pi_1(Q)$, $L \in \partial Q$ is a point such that $\pi_1(L) = l$ and $\boldsymbol{L} = \varphi(L)$. Take the left end of ∂Q and call it

$$E = \Big\{A \in \partial Q;\ \pi_1(A) \leq l + \frac{3\varepsilon}{M}\Big\}.$$

Using (3.3) we can calculate

$$\mathcal{H}^1(\varphi(E)) < 2M\mathcal{H}^1(\pi_1(E)) + \varepsilon = 7\varepsilon.$$

It is hereby clear that at least one of the pair X, Y is not in E because otherwise we could find a curve in $\varphi(\partial Q)$ with endpoints $\boldsymbol{X}$ and $\boldsymbol{Y}$ with length less than 7ε but we know this is not true. Assume that X is in E. Since $\boldsymbol{X}_1 = \boldsymbol{Y}_1$ we can use (3.3) to show that there exists a curve in $\varphi(\partial Q)$ connecting $\boldsymbol{Y}$ with a point in $\varphi(E)$ of length less than 2ε. But this means that there is a curve connecting $\boldsymbol{X}$ and $\boldsymbol{Y}$ shorter than 9ε and we have excluded this possibility. Now consider those points $\boldsymbol{A} = (\boldsymbol{A}_1, \boldsymbol{A}_2) = \varphi(A) \in \varphi(\partial Q)$ such that

$$|\boldsymbol{A}_1 - \boldsymbol{X}_1| < 2\varepsilon.$$

We get by (3.9) that

$$|A_1 - X_1| < \frac{3\varepsilon}{M}.$$

Now since $|A_1 - X_1| < \frac{3\varepsilon}{M} < |L_1 - X_1|$ we get that exactly one of the two possibilities hold

$a)$ All curves $\tilde{\gamma} \subset \partial Q$ with endpoints A and X are direct,

$b)$ All curves $\tilde{\gamma} \subset \partial Q$ with endpoints A and Y are direct.

Obviously the same argument holds at the right-hand end of Q. Also for each $\boldsymbol{x}$ such that $|\boldsymbol{X}_1 - \boldsymbol{x}| < 2\varepsilon$ we can find $\boldsymbol{y}_1$, and $\boldsymbol{y}_2$ such that $(\boldsymbol{x}, \boldsymbol{y}_1), (\boldsymbol{x}, \boldsymbol{y}_2) \in \varphi(\partial Q)$ and $A = (\boldsymbol{x}, \boldsymbol{y}_1)$ corresponds to case $a)$ and $A = (\boldsymbol{x}, \boldsymbol{y}_2)$ corresponds to case $b)$. See also Figure 8.

Lemma 3.2 implies that we can find an $\boldsymbol{x}$ with $|\boldsymbol{x} - \boldsymbol{X}_1| < 2\varepsilon$ such that the vertical line $\{(\overline{\boldsymbol{x}}, \boldsymbol{t});\ \boldsymbol{t} \in \mathbb{R}\}$, intersects $\varphi(\partial Q)$ at only two points which are $\overline{\boldsymbol{X}} = \varphi(\overline{X})$ and $\overline{\boldsymbol{Y}} = \varphi(\overline{Y})$. Further from the above we know that the curves on ∂Q connecting X with $\overline{X}$ and Y with $\overline{Y}$ are direct. With respect to (3.9) we get that

$$|X_1 - \overline{X}_1| < \frac{3\varepsilon}{M}.$$

This means we can estimate the length of the image $\gamma = \varphi(\tilde{\gamma})$ of the direct curve $\tilde{\gamma}$ having endpoints X and $\overline{X}$ using (3.8) as

$$\mathcal{H}^1(\gamma) < 4\varepsilon.$$

As in (3.10), the length of the segment $[\overline{\boldsymbol{X}\boldsymbol{Y}}]$ is less than ε. Thus we can connect $\boldsymbol{X}$ with $\boldsymbol{Y}$ with two curves in $\varphi(\partial Q)$ and a segment having a total length less than 9ε.

It now suffices alter the curve slightly so that it does not touch the boundary of $\mathcal{S}$ anywhere else than $\boldsymbol{X}$ and $\boldsymbol{Y}$. We consider the curve that consists of the segment $[\boldsymbol{X}, (\boldsymbol{X} + \alpha_{\boldsymbol{X}} v_{\boldsymbol{X}})]$ then follows P_α until it meets $(\overline{\boldsymbol{x}}, \boldsymbol{t})$, goes vertically along $(\overline{\boldsymbol{x}}, \boldsymbol{t})$ until it meets P_α again at the other end, follows P_α until it gets to a point $\boldsymbol{Y} + \alpha_{\boldsymbol{Y}} v_{\boldsymbol{Y}}$

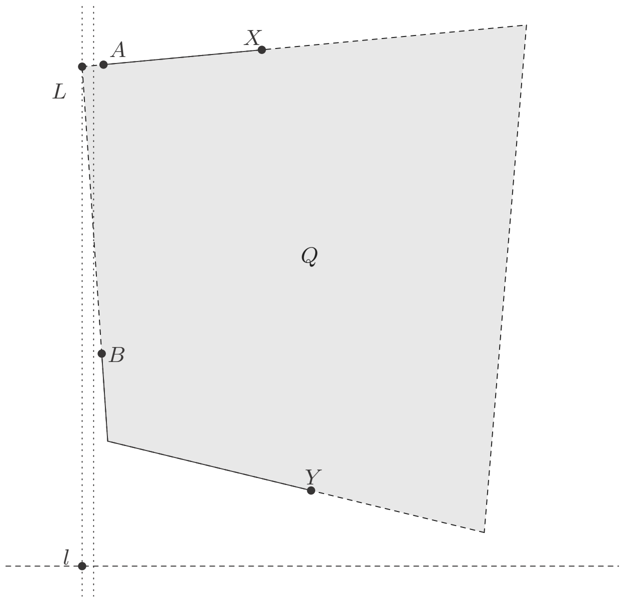


FIGURE 8. Cutting off the ends of the square ensures paths between A and X, respectively B and Y will be direct.

and then follows the final segment into $\boldsymbol{Y}$. Note that in general $\alpha_{\boldsymbol{X}}, \alpha_{\boldsymbol{Y}} \le \alpha$ with equality if $\boldsymbol{X}$ or $\boldsymbol{Y}$ are vertices of $\varphi(\partial Q)$. In doing this we increase the length of our curve by no more than ε given that α is smaller than $\varepsilon(4m+4)^{-1}$, where m is the number of vertices of $\varphi(\partial Q)$. Thus the final curve is shorter than 10ε. □

Lemma 3.5. *Let $\varphi, Q, M, \varepsilon$ be as given in* (3.3) *and $\mathcal{S}$ be the closure of the bounded component of $\mathbb{R}^2 \setminus \varphi(\partial Q)$. Then for any two points in $\boldsymbol{X} = \varphi(X), \boldsymbol{Y} = \varphi(Y) \in \varphi(\partial Q)$ we can find a piecewise-linear curve with endpoints $\boldsymbol{X}$ and $\boldsymbol{Y}$, lying inside $\mathcal{S}^\circ$ (apart from its endpoints), whose length is no more than $|\boldsymbol{X}_1 - \boldsymbol{Y}_1| + 15\varepsilon$.*

Proof. The strategy of our proof is to find a point $\tilde{\boldsymbol{Y}} = \varphi(\tilde{Y})$ such that we can use Lemma 3.4 to connect $\boldsymbol{Y}$ and $\tilde{\boldsymbol{Y}}$ and there exists a direct curve between X and $\tilde{Y}$. The length of the image of this direct curve can be estimated using (3.8). Then we can use move the curve inside $\mathcal{S}$ making it only ε longer. Finally we notice that if at any point the curve crosses itself we can cut out the loop which was formed thus shortening the curve and making it a one-to-one path.

As we saw in the proof of Lemma 3.4,

$$\mathcal{H}^1(\varphi(E)) < 9\varepsilon,$$

where $E = \{A \in \partial Q;\ A_1 \le l + 3\varepsilon M^{-1}\}$. So if $Y \in E$ we can find a point $\tilde{Y} \in E$ and a direct $\tilde{\gamma}$ path in ∂Q with endpoints X and $\tilde{Y}$. then we can estimate the length of the image of this direct path as

$$\mathcal{H}^1(\varphi(\tilde{\gamma})) < M|X_1 - \tilde{Y}_1| + \varepsilon < |\boldsymbol{X}_1 - \boldsymbol{Y}_1| + 5\varepsilon.$$

We can connect $\boldsymbol{Y}$ and $\tilde{\boldsymbol{Y}}$ with a curve of length less than 9ε and so we have connected these points in $\varphi(\partial Q)$ with a curve of length less than 14ε. Moving this curve inside $\mathcal{S}$ as we did in the proof of Lemma 3.4 will increase the length by at most ε.

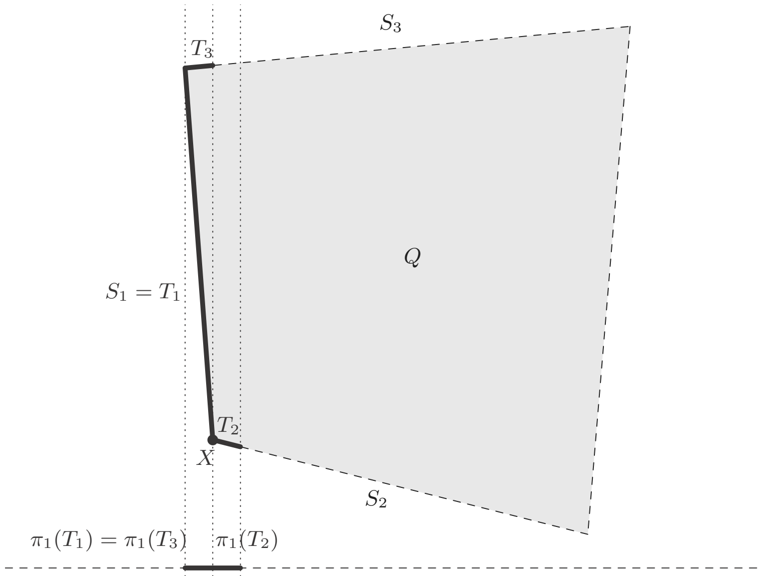


FIGURE 9. The preimage of one of two adjoining segments in ∂Q having the same size projections cannot be too big.

If neither X nor $Y = (Y_1, Y_2)$ are close to the edge of ∂Q then we know that the set $\mathcal{F} = \{(Y_1, t) \in \varphi(\partial Q)\}$ has at least two elements. At least one of these, $\tilde{Y}$ can be connected with X by a direct curve $\tilde{\gamma}$. Since

$$|X_1 - \tilde{Y}_1| < |X_1 - Y_1| + \frac{\varepsilon}{M},$$

we have

$$\mathcal{H}^1(\varphi(\tilde{\gamma})) < M|X_1 - Y_1| + 2\varepsilon < |\boldsymbol{X}_1 - \boldsymbol{Y}_1| + 3\varepsilon.$$

So we can apply Lemma 3.4 to connect $\tilde{\boldsymbol{Y}} = \varphi(\tilde{Y})$ with $\boldsymbol{Y}$ and the length of our curve will be less than $|\boldsymbol{X}_1 - \boldsymbol{Y}_1| + 13\varepsilon$. We move this curve inside $\mathcal{S}$ and get a new curve no longer than $|\boldsymbol{X}_1 - \boldsymbol{Y}_1| + 14\varepsilon$.

It is possible that this curve crosses itself but we need (the parametrizations of) our curve to be one-to-one. To this end take a parametrization ϕ of our curve from $[0, 1]$. If ϕ is not one-to-one we have points $t_1 < t_2$ such that $\phi(t_1) = \phi(t_2)$. Take the curve defined by $\phi([0, t_1] \cup [t_2, 1])$. Clearly this curve is shorter than the original. If necessary we can repeat this process until ϕ is one-to-one. This does not increase the length of our curve. $\square$

The following lemma shows that the 2-bi-Lipschitz piecewise-affine image of a square cannot have two adjoining sides which are both nearly vertical and so we can bound the length of one of a pair of segments given we know that its projection onto the x-axis of the segments is small. See also Figure 9. Recall that by τ we denote the tangential vector to a one-dimensional object, as explained in the preliminaries.

Lemma 3.6. *Let $r_0 > 0$ and Q will be the image of a 2-bi-Lipschitz mapping which is equal to an affine mapping on the triangles* $\text{co}\{(0,0),(0,r_0),(r_0,0)\}$ *and* $\text{co}\{(r_0,r_0),(0,r_0),(r_0,0)\}$. *Let S_1, S_2 be two sides of Q, which meet at the corner*

X of Q and let T_1 and T_2 be segments in S_1 and S_2 respectively, both containing X, such that $\mathcal{H}^1(\pi_1(T_1)) = \mathcal{H}^1(\pi_1(T_2))$. Then it holds that

$$\min\{\mathcal{H}^1(T_1), \mathcal{H}^1(T_2)\} \leq 20r_0 \min\{|\pi_1(\tau_{S_1})|, |\pi_1(\tau_{S_2})|\}. \tag{3.11}$$

Proof. Without loss of generality we may assume that S_1 is nearly vertical because if $\min_i |\pi_1(\tau_{S_i})| \geq \frac{1}{10}$ then the right hand side of (3.11) is greater than or equal to $2r_0$. The left hand side however is trivially bounded by the same and the claim holds. Therefore assume that $|\pi_1(\tau_{S_1})| \leq \frac{1}{10}$.

By our assumptions on Q, simple direct computation gives that the angle between S_1 and S_2 is bounded from below by $\arctan \frac{1}{4}$ and therefore, for any two sides with a common vertex it holds that

$$|\pi_1(\tau_{S_2})| \geq \frac{1}{10}.$$

We know that $\mathcal{H}^1(S_1) \leq 2r_0$ and therefore the length of its projection onto the x-axis is at most $2r_0|\pi_1(\tau_{S_1})|$. Now the two segments have the same length of projection onto the x-axis and $|\pi_1(\tau_{S_2})|$ is bounded from below by $\frac{1}{10}$ therefore

$$\mathcal{H}^1(T_2) \leq 20r_0|\pi_1(\tau_{S_1})|,$$

thus our claim has been proved. □

Theorem 3.7. *Let $M > 0$, $0 < \delta < \frac{1}{100}$, $r_0 \in (0,1)$ and Q will be the image of a 2-bi-Lipschitz mapping which is equal to an affine mapping on $\operatorname{co}\{(0,0),(0,r_0),(r_0,0)\}$ and $\operatorname{co}\{(r_0,r_0),(0,r_0),(r_0,0)\}$. Denote $\sigma = \min\{M, M^{-1}, \Phi(M), \Phi(M)^{-1}\}$. Let $\varphi : \partial Q \to \mathbb{R}^2$ be a piecewise-linear and one-to-one mapping with*

$$\int_{\partial Q} \left| D_\tau \varphi(t) - \begin{pmatrix} M, 0 \\ 0, 0 \end{pmatrix} \tau \right| d\mathcal{H}^1(t) < \delta\sigma r_0, \tag{3.12}$$

and

$$\int_{\partial Q} \Phi\Big(\Big| D_\tau \varphi(t) - \begin{pmatrix} M, 0 \\ 0, 0 \end{pmatrix} \tau \Big| \Big) \, d\mathcal{H}^1(t) < \delta\sigma r_0 \ , \tag{3.13}$$

where τ is the unit tangential vector to ∂Q. Then there is a piecewise-linear homeomorphism $g : Q \to \mathbb{R}^2$ such that $g = \varphi$ on ∂Q and

$$\int_Q \Phi\Big(\Big| Dg(x) - \begin{pmatrix} M, 0 \\ 0, 0 \end{pmatrix} \Big| \Big) \ dx \leq C\delta r_0^2 \ . \tag{3.14}$$

Proof. Firstly let us briefly outline the idea of the proof. We will take some large central part of Q and call it W. On W we define g almost as a curve in $\mathcal{S}$ so that g is close to Mxe_1 in L^1 on W. We then use the fact that $|Dg| < 2M$almost everywhere in $W$to get (3.14). We split the annulus $Q \setminus W$ into very small square-like objects where we use bi-Lipschitz mappings and Theorem 2.1. A combination of Lemma 3.5 and the fact that the annulus has very small measure guarantee that overall the L^Φ norm here is tiny.

Step 1: Basic setup.

Let the snake $\mathcal{S}$ denote the closure of the bounded component of $\mathbb{R}^2 \setminus \varphi(\partial Q)$. We need to denote the corners of Q somehow. If all corners have different x-coordinates we call A_1 the vertex of Q having the least x-coordinate, A_2 is the vertex having second least x-coordinate, A_3 the third least and A_4 is the vertex having the greatest x-coordinate. If we need to assign the corners A_i and A_{i+1} but we have two vertices

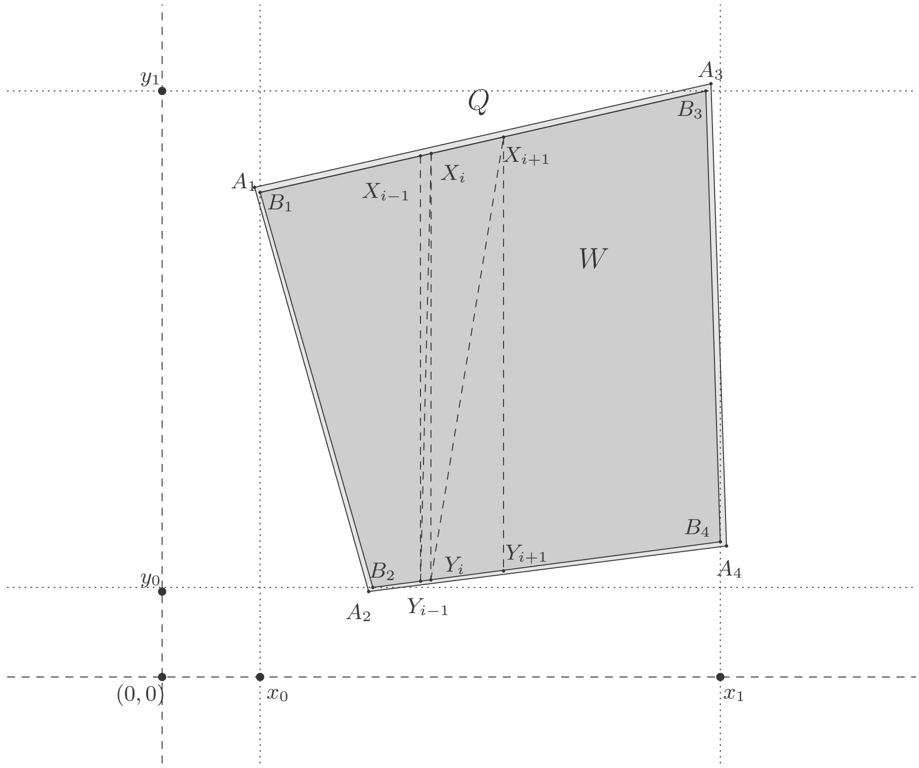


FIGURE 10. Decomposition of Q into its small annulus and the large central part W, which we divide into triangles having vertical sides. The values x_0, x_1, y_0, y_1 being the extremities of W.

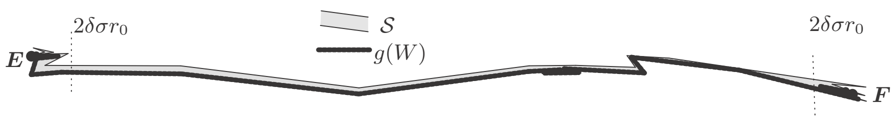


FIGURE 11. We define a curve that follows the boundary of $\mathcal{S}$, starting near its left end and ending near its right.

with the same x-coordinate then A_i is the corner with the greatest y-coordinate of the two and the other is A_{i+1}. See Figure 10.

Call A the central point of Q defined as $A = (A_1 + A_2 + A_3 + A_4)/4$. We will define our mapping in one way on a large central part of Q and then use the annulus to fill in the remaining part of $\mathcal{S}$. Let us define

$$B_i = A_i + \delta\sigma(A - A_i).$$

We proceed to define g on $W = \mathrm{co}\{B_1, B_2, B_3, B_4\}$.

Step 2: Choosing a good path along the boundary of $\mathcal{S}$.

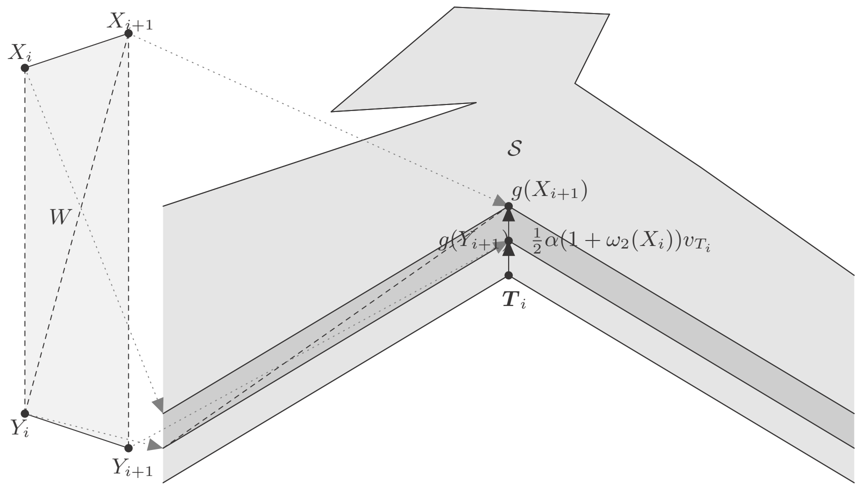


FIGURE 12. Defining $g(X_i)$ and $g(Y_i)$ as points close to $\boldsymbol{T}_i$ just inside $\mathcal{S}$ and controlling the distance from $\boldsymbol{T}_i$ using their y-coordinate by the function ω_2.

We will use a direct curve $\tilde{\gamma} \subset \partial Q$ such that $\pi_1(\tilde{\gamma}) = \pi_1(\partial Q)$. Depending on the exact shape of Q one of the following will suffice, $[A_1A_4]$, $[A_1A_2] \cup [A_2A_4]$, $[A_1A_2] \cup [A_2A_3] \cup [A_3A_4]$, $[A_1A_3]$ or $[A_2A_4]$. We call $\tilde{E} = (\tilde{E}_1, \tilde{E}_2)$ and $\tilde{F} = (\tilde{F}_1, \tilde{F}_2)$ the endpoints of $\tilde{\gamma}$, where $\tilde{E}_1 < \tilde{F}_1$ and $\tilde{\boldsymbol{E}} = \varphi(\tilde{E})$, $\tilde{\boldsymbol{F}} = \varphi(\tilde{F})$. Also we call $\gamma = \varphi(\tilde{\gamma})$.

We refer to a vertex of $\varphi(\partial Q)$ as the image of any point, at which $D_\tau\varphi$ does not exist. By K we denote the number of vertices of $\varphi(\partial Q)$ in the curve γ. Naming the constant-speed parametrization of γ as $\gamma^* : [0,1] \to \mathbb{R}^2$ we find $0 = t_1 < t_2 < \cdots < t_K = 1$ such that $\boldsymbol{T}_i = \gamma^*(t_i)$ are the vertices of γ and $\tilde{\boldsymbol{E}} = \boldsymbol{T}_1, \tilde{\boldsymbol{F}} = \boldsymbol{T}_K$.

Similarly as in (3.8) we can estimate the length of γ as

$$M|\tilde{F}_1 - \tilde{E}_1| - \delta\sigma r_0 < \mathcal{H}^1(\gamma) < M|\tilde{F}_1 - \tilde{E}_1| + \delta\sigma r_0.$$

Step 3: Definition of g on W.

Now we are ready to define g on W. The image of W is depicted in Figure 11. We will pick pairs of points on ∂W where the points of each pair have the same x-coordinate. The segments connecting these points will be perfectly vertical and D_2g on any triangle having this segment as one of its sides will be the same as the derivative of g along the vertical segment itself. By ensuring that g has very small derivative on all of these segments we will know that D_2g is tiny everywhere on W. We depict this strategy in Figure 12. In order to achieve the above we will use the numbers

$$x_0 = \min\{x;\ (x,y) \in W, y \in \mathbb{R}\} \text{ and } x_1 = \max\{x;\ (x,y) \in W, y \in \mathbb{R}\},$$
$$y_0 = \min\{y;\ (x,y) \in W, x \in \mathbb{R}\} \text{ and } y_1 = \max\{y;\ (x,y) \in W, x \in \mathbb{R}\},$$

depicted in Figure 10, to define

$$\omega_1(x,y) = \frac{x - x_0}{x_1 - x_0} \qquad \omega_2(x,y) = \frac{y - y_0}{y_1 - y_0} \qquad (x,y) \in W.$$

Note that any two points $X, Y \in W$ such that $\omega_1(X) = \omega_1(Y)$ have the same x-coordinate. We choose points X_i and Y_i as depicted in Figure 10, they are those points on ∂W such that

$$\begin{aligned} a) &\qquad \omega_2(X_i) > \omega_2(Y_i) \text{ for } 2 \leq i \leq K-1 \\ b) &\quad \omega_1(X_i) = \omega_1(Y_i) = t_i \text{ for } 1 \leq i \leq K. \end{aligned} \tag{3.15}$$

We still need to define X_1, Y_1, X_K and Y_K. There are two cases, if $\omega_1(B_1) = \omega_1(B_2)$ then

$$X_1 = B_1, Y_1 = B_2, \text{ otherwise } X_1 = Y_1 = B_1.$$

Similarly $X_K = B_3$ and $Y_K = B_4$ if B_3 and B_4 have the same x-coordinate and if they do not, we put $X_K = Y_K = B_4$.

Firstly we define $g(X_i)$ and $g(Y_i)$ and then define g on triangles made from the points X_i and Y_i to get a homeomorphism on W. Let us define the number λ as

$$\lambda = \min\{|X_{i+1} - X_i| : i = 1, \dots K-1\} \cup \{|Y_{i+1}Y_i| : i = 1, \dots K-1\} \cup \{1\}.$$

Call $P = \varphi(\partial Q)$ and let α_0 be the minimum of the number defined in (3.2) and $\frac{1}{10}$. Choose

$$\alpha \in (0, \delta\sigma r_0 \alpha_0 \lambda^2 K^{-1}). \tag{3.16}$$

The following definition is crucial for our construction. It determines that on W we follow the boundary of the image with constant speed which guarantees that our approximation is $2M$ Lipschitz everywhere in W. Recall the definition of 'normal' vectors from Lemma 3.1, (3.1) and for $i = 1, 2, \dots, K$ define

$$g(X_i) = \varphi(T_i) + \frac{\alpha}{2}(1 + \omega_2(X_i))v_{T_i} \text{ and } g(Y_i) = \varphi(T_i) + \frac{\alpha}{2}(1 + \omega_2(Y_i))v_{T_i}. \tag{3.17}$$

Now for $i = 1, 2, \dots, K-1$ define g on the segments $[X_iX_{i+1}]$, $[Y_iY_{i+1}]$, $[X_iY_i]$ and $[X_{i+1}Y_i]$ and on $[X_KY_K]$ as linear. In this way we subdivide W into triangles formed by the segments $[X_iX_{i+1}]$, $[X_iY_i]$ and $[X_{i+1}Y_i]$ and those triangles consisting of $[X_{i+1}Y_{i+1}]$, $[Y_iY_{i+1}]$ and $[X_{i+1}Y_i]$. We have defined g as linear on each segment and g is continuous on each triangle $X_iX_{i+1}Y_i$ and $X_{i+1}Y_iY_{i+1}$. There is therefore a uniquely determined piecewise-affine mapping on each triangle and we define g on the convex hull of each triangle as this mapping. Hereby we have a finite piecewise-affine mapping defined on W. It is not difficult to notice that each affine map is regular and in fact g is a homeomorphism on W. This is easy because Lemma 3.1 ensures that our curves P_α are pair-wise disjoint from each other, which means that the triangular images of our triangles in W do not touch each other (except at the images of their common points or sides in W).

Step 4: Estimates for Dg on W.

Now we need to estimate $Dg = (D_1g, D_2g)$. Firstly we estimate D_2g. Every triangle has a vertical side, which we can denote as $[X_iY_i]$ and $D_2g(x,y)$ for any point inside this triangle is equal to $D_\tau g$ on $[X_iY_i]$. By direct calculation (recall (3.17)) we have

$$\begin{aligned} D_2g(x,y) = D_{\tau_{[X_iY_i]}}g &= \frac{g(X_i) - g(Y_i)}{|X_i - Y_i|} = \frac{g(X_i) - g(Y_i)}{(y_1 - y_0)(\omega_2(X_i) - \omega_2(Y_i))} \\ &= \frac{\alpha}{2} \frac{(\omega_2(X_i) - \omega_2(Y_i))v_{T_i}}{(\omega_2(X_i) - \omega_2(Y_i))(y_1 - y_0)} = \frac{\alpha v_{T_i}}{2(y_1 - y_0)}. \end{aligned}$$

Recalling (3.16) we notice that our choice of α ensures that

$$|D_2 g(x,y)| < \delta\sigma K^{-1}\lambda^2\alpha_0.$$

Now we estimate $D_1 g$. Each triangle G has a non-vertical side lying on ∂W, $[X_i X_{i+1}]$ or $[Y_i Y_{i+1}]$ (without loss of generality assume $[X_i X_{i+1}]$). We calculate the derivative $D_\tau g$ on this segment and then clearly the triangle having this segment as its side fulfils

$$D_\tau g = D_1 g(x,y)\tau_1 + D_2 g(x,y)\tau_2,$$

where $\tau = (\tau_1, \tau_2)$ is the tangential unit vector to the segment $[X_i X_{i+1}] \subset \partial W$ as defined in the preliminaries. We are free to assume that $\tau_1 \geq 0$ by our choice of parametrization. We start by showing that roughly speaking $|D_\tau g| \approx \mathcal{H}^1(\gamma)\tau_1(x_1 - x_0)^{-1}$. In Step 2 we chose the points X_i so that $\omega_1(X_i) = t_i$. It is important to note that for $X \in [X_i X_{i+1}]$ we have

$$D_\tau \omega_1(X) = \frac{\pi_1(X+\tau) - \pi_1(X)}{x_1 - x_0} = \frac{\tau_1}{x_1 - x_0},$$

since $[X_i X_{i+1}]$ is a segment lying on ∂W, τ is constant there, which means that ω_1 is linear on each side of ∂W and so $D_\tau \omega_1$ is constant. From this we calculate

$$\omega_1(X_{i+1}) - \omega_1(X_i) = \frac{|X_{i+1} - X_i|\tau_1}{x_1 - x_0}.$$

Now we notice that since $|(\gamma^*)'| = \mathcal{H}^1(\gamma)$ is constant on all (t_i, t_{i+1}) and use (3.15) *(b)* to get

$$\frac{|\varphi(T_{i+1}) - \varphi(T_i)|}{\mathcal{H}^1(\gamma)} = \frac{|\gamma^*(t_{i+1}) - \gamma^*(t_i)|}{\mathcal{H}^1(\gamma)} = \frac{|(\gamma^*)'|(t_{i+1} - t_i)}{\mathcal{H}^1(\gamma)} = \omega_1(X_{i+1}) - \omega_1(X_i).$$

Together that is

$$\frac{|\varphi(T_{i+1}) - \varphi(T_i)|}{|X_{i+1} - X_i|} = \mathcal{H}^1(\gamma)\frac{\omega_1(X_{i+1}) - \omega_1(X_i)}{|X_{i+1} - X_i|} = \frac{\mathcal{H}^1(\gamma)\tau_1}{x_1 - x_0}.$$

So recalling our definition of g in (3.17), we can calculate $D_\tau g$ on $[X_i X_{i+1}]$ as

$$\begin{aligned} D_\tau g &= \frac{g(X_{i+1}) - g(X_i)}{|X_{i+1} - X_i|} \\ &= \frac{\varphi(T_{i+1}) - \varphi(T_i)}{|X_{i+1} - X_i|} + \frac{\alpha\big(1+\omega_2(X_{i+1})\big)v_{T_{i+1}} - \alpha\big(1+\omega_2(X_i)\big)v_{T_i}}{2|X_{i+1} - X_i|} \\ &= \frac{\mathcal{H}^1(\gamma)\tau_1}{(x_1 - x_0)}\frac{\varphi(T_{i+1}) - \varphi(T_i)}{|\varphi(T_{i+1}) - \varphi(T_i)|} + \frac{\alpha\big(1+\omega_2(X_{i+1})\big)v_{T_{i+1}} - \alpha\big(1+\omega_2(X_i)\big)v_{T_i}}{2|X_{i+1} - X_i|}. \end{aligned}$$

Now $1 + \omega_2(X_j) \leq 2$, $|X_{i+1} - X_i| \geq \lambda$ and v_{T_j} are all unit vectors. Therefore

$$\left| D_\tau g - \frac{\mathcal{H}^1(\gamma)\tau_1}{(x_1 - x_0)}\frac{\varphi(T_{i+1}) - \varphi(T_i)}{|\varphi(T_{i+1}) - \varphi(T_i)|}\right| \leq 2\alpha\lambda^{-1} < 2\delta\sigma r_0\alpha_0\lambda K^{-1}. \tag{3.18}$$

It is clear that given a triangle G, an affine mapping $\mathcal{A}$ on G and if we know $D_2\mathcal{A}$ and the value of $D_\tau\mathcal{A}$ on a non-vertical edge of G we can calculate $D_1\mathcal{A}$. We have

$$D_\tau\mathcal{A} = D\mathcal{A}\tau = D_1\mathcal{A}\tau_1 + D_2\mathcal{A}\tau_2.$$

Our choice of λ allows us to effectively estimate D_1g using $D_\tau g$ on nearly vertical lines. Recall that we have already calculated that $|D_2g\tau_2| \leq \alpha r_0^{-1}$. Therefore, since Lemma 3.6 gives $\lambda \leq 20r_0\tau_1$ and applying (3.18), we know that

$$\begin{aligned}
&\left|D_1g - \frac{\mathcal{H}^1(\gamma)}{(x_1-x_0)}\frac{\varphi(T_{i+1})-\varphi(T_i)}{|\varphi(T_{i+1})-\varphi(T_i)|}\right|\\
&\quad= \left|\frac{D_\tau g - D_2g\tau_2}{\tau_1} - \frac{\mathcal{H}^1(\gamma)}{(x_1-x_0)}\frac{\varphi(T_{i+1})-\varphi(T_i)}{|\varphi(T_{i+1})-\varphi(T_i)|}\right|\\
&\quad< \frac{2\delta\sigma r_0\alpha_0\lambda K^{-1} + \delta\sigma\alpha_0\lambda^2K^{-1}}{\tau_1}\\
&\quad< 40\delta\sigma r_0^2\alpha_0K^{-1} + 20\delta\sigma r_0\alpha_0\lambda K^{-1}.
\end{aligned}$$

We have $\mathcal{H}^1(\gamma) \approx M(x_1-x_0)$ and D_2g tiny so we can use this to conclude that g is Lipschitz on W, thus

$$|Dg| \leq 2M \text{ and so } \left|Dg - \begin{pmatrix} M,0\\ 0,0\end{pmatrix}\right| \leq 3M, \tag{3.19}$$

almost everywhere.

It holds that

$$\begin{aligned}
\int_G\left|Dg - \begin{pmatrix} M,0\\ 0,0\end{pmatrix}\right| &\leq \int_G\left|D_1g - \begin{pmatrix} M\\ 0\end{pmatrix}\right| + \int_G|D_2g|\\
&\leq \int_G|\pi_1(D_1g)-M| + \int_G|\pi_2(D_1g)| + |G|\|D_2g\|_\infty.
\end{aligned}$$

We have already calculated that $\|D_2g\|_\infty < \delta\sigma\alpha_0\lambda^2K^{-1}$. Notice that for any triangle G whose non-vertical side is $[X_iX_{i+1}]$ or $[Y_iY_{i+1}]$ we have $|G| \leq 4r_0|\pi_1(X_{i+1}-X_i)| = 4r_0|X_{i+1}-X_i|\tau_1$ and on G we know that all our derivatives are constant. So we can calculate

$$\begin{aligned}
\int_G\left|D_1g - \begin{pmatrix} M\\ 0\end{pmatrix}\right| &\leq \int_G\left|\frac{D_\tau g}{\tau_1} - \begin{pmatrix} M\\ 0\end{pmatrix}\right| + \int_G\left|\frac{D_\tau g}{\tau_1} - D_1g\right|\\
&\leq \int_G\left|\frac{D_\tau g}{\tau_1} - \begin{pmatrix} M\\ 0\end{pmatrix}\right| + |G|\|D_2g\|_\infty\\
&\leq 4r_0|X_{i+1}-X_i|\tau_1\left|\frac{D_\tau g}{\tau_1} - \begin{pmatrix} M\\ 0\end{pmatrix}\right| + |G|\|D_2g\|_\infty\\
&\leq 4r_0\int_{[X_iX_{i+1}]}\left|D_\tau g - \begin{pmatrix} M,0\\ 0,0\end{pmatrix}\tau\right|d\mathcal{H}^1 + |G|\delta\sigma\alpha_0K^{-1}.
\end{aligned}$$

Now we sum over all triangles in W to get

$$\int_W\left|D_1g - \begin{pmatrix} M\\ 0\end{pmatrix}\right| \leq 4r_0\int_{\partial W}\left|D_\tau g - \begin{pmatrix} M,0\\ 0,0\end{pmatrix}\tau\right|d\mathcal{H}^1 + 4r_0^2\delta\sigma\alpha_0K^{-1}. \tag{3.20}$$

Define

$$\partial W = \mathcal{X}\cup\mathcal{Y}, \text{ where } \mathcal{X} = \cup[X_iX_{i+1}] \text{ and } \mathcal{Y} = \cup[Y_iY_{i+1}]. \tag{3.21}$$

We can estimate both of the integrals $\int_{\mathcal{X}}|D_\tau g(t) - M\tau_1e_1|d\mathcal{H}^1(t)$ and $\int_{\mathcal{Y}}|D_\tau g(t) - M\tau_1e_1|d\mathcal{H}^1(t)$ in the same way as follows. We have

$$\int_{\mathcal{X}}\left|D_\tau g - \begin{pmatrix} M,0\\ 0,0\end{pmatrix}\tau\right|d\mathcal{H}^1 \leq \int_{\mathcal{X}}|\pi_1(D_\tau g) - M\tau_1| + \int_{\mathcal{X}}|\pi_2(D_\tau g)|.$$

The second integral can be estimated using (3.12) and the triangle inequality as

$$\int_{\mathcal{X}} |\pi_2(D_\tau g)| d\mathcal{H}^1 \leq \delta\sigma r_0 + 2K\alpha \leq 2\delta\sigma r_0.$$

Further we can calculate

$$\int_{\mathcal{X}} |\pi_1(D_\tau g) - M\tau_1| = \int_{\mathcal{X}} \big(\pi_1(D_\tau g) - M\tau_1\big)^+ - \int_{\mathcal{X}} \big(\pi(D_\tau g) - M\tau_1\big)^-.$$

Now it is clear that

$$\begin{aligned}\int_{\mathcal{X}} \big(\pi_1(D_\tau g) - M\tau_1\big)^+ &\leq \Big|\mathcal{H}^1\big(g(\mathcal{X})\big) - M(x_1 - x_0)\Big| \\ &\leq |\mathcal{H}^1(\gamma) + 2\alpha K - M(x_1 - x_0)| \\ &\leq 3\delta\sigma r_0.\end{aligned}$$

We also have by using (3.12) and (3.17) that

$$\begin{aligned}\int_{\mathcal{X}} \big(\pi_1(D_\tau g) - M\tau_1\big)^+ + \int_{\mathcal{X}} \big(\pi_1(D_\tau g) - M\tau_1\big)^- &\geq \int_{\tilde{\gamma}} |D_\tau \varphi| - 2\delta\sigma r_0 \\ &> M(x_1 - x_0) - 3\delta\sigma r_0.\end{aligned}$$

Therefore we get that

$$\int_{\mathcal{X}} \big(\pi_1(D_\tau g) - M\tau_1\big)^- > -6\delta\sigma r_0.$$

Combining the above renders

$$\int_{\mathcal{X}} |\pi_1(D_\tau g) - M\tau_1| < 9\delta\sigma r_0.$$

Obviously the same holds for the integral over $\mathcal{Y}$. This and the fact that

$$\int_{\partial W} |\pi_2(D_\tau g)| \leq 4\delta\sigma r_0$$

together give us a fixed geometrical constant C such that

$$\int_{\partial W} \left| D_\tau g - \begin{pmatrix} M, 0 \\ 0, 0 \end{pmatrix} \tau \right| < C\delta\sigma r_0. \tag{3.22}$$

Now combining this with the fact that $\|D_2 g\|_\infty \leq \delta\sigma$ and $|W| \leq 4r_0^2$ we get with respect to (3.20) that,

$$\int_{W} \left| Dg - \begin{pmatrix} M, 0 \\ 0, 0 \end{pmatrix} \right| \leq C\delta\sigma r_0^2, \tag{3.23}$$

where C is a fixed geometrical constant.

We will now get a (3.14)-type estimate on W by combining (3.23) with (3.19) and the fact that Φ satisfies the Δ_2 condition. Firstly notice by (3.19) that almost everywhere

$$\Phi\left(\left| Dg - \begin{pmatrix} M, 0 \\ 0, 0 \end{pmatrix} \right|\right) \leq C_2^2 \Phi(M)$$

with C_2 the Δ_2 constant of Φ. Now, as Φ is convex it is clear that

$$\Phi(t) \leq C_2^2 t \frac{\Phi(M)}{M}$$

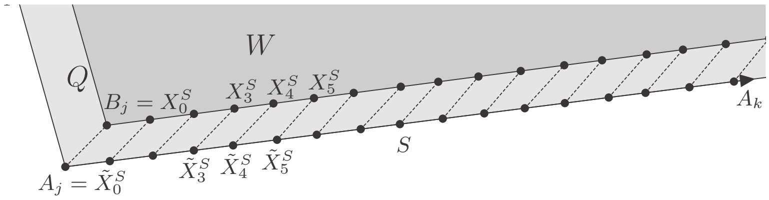


FIGURE 13. Separating the part of the annulus on the side $S = [A_j A_k]$ into similar 2-bi-Lipschitz images of a square.

for all $t \leq 4M$. Combing this with (3.23) we get

$$\begin{aligned}\int_W \Phi\Big(\Big|Dg - \begin{pmatrix} M,0 \\ 0,0 \end{pmatrix}\Big|\Big) &\leq C_2^2 \frac{\Phi(M)}{M} \int_W \Big|Dg - \begin{pmatrix} M,0 \\ 0,0 \end{pmatrix}\Big| \\ &< C\frac{\Phi(M)}{M}\delta\sigma r_0^2 \leq C\delta r_0^2,\end{aligned} \tag{3.24}$$

where in the final inequality we have used the definition of σ and the fact that $\frac{\Phi(M)\sigma}{M}$ is a bounded function of M. The constant C depends only on Φ.

Step 5: Division of the annulus $Q \setminus W$.

We will now turn our attention to the annulus $Q \backslash W$. We want to apply Theorem 2.1 on the annulus, therefore we will split it into small areas, each of which are uniformly bi-Lipschitz equivalent with a square as we show in Figure 13. To this end choose $m = [\delta^{-1}\sigma^{-1}] \in \mathbb{N}$. Each side S of ∂Q corresponds to some $S = [A_j A_k]$. Thus we define

$$\tilde{X}_i^S = \frac{iA_k + (m-i)A_j}{m} \qquad i = 0, 1 \dots m, \tag{3.25}$$

which yields $|\tilde{X}_i^S - \tilde{X}_{i+1}^S| \approx \delta\sigma r_0$.

For any point $\tilde{X}$ in ∂Q we have a corresponding point X in ∂W such that

$$X = \tilde{X} + \delta\sigma(A - \tilde{X}) \text{ that is to say } \tilde{X} = \frac{X - \delta\sigma A}{1 - \delta\sigma}.$$

Either way we can easily estimate

$$|X - \tilde{X}| \leq 4\delta\sigma r_0.$$

We define $g = \varphi$ on ∂Q and we want to estimate the distance $|g(\tilde{X}_i^S) - g(X_i^S)|$. Since we have (3.12) and (3.22) we will be able to apply the same type estimates as we used in (3.8)-(3.9). At the start of step 2 we defined $\tilde{E}$ as either A_1 or A_2 and $\tilde{F}$ as either A_3 or A_4. If $\tilde{E} = A_i$ then $E = B_i$ and the same for F so that we have

$$E = \tilde{E} + \delta\sigma(A - \tilde{E}), \quad F = \tilde{F} + \delta\sigma(A - \tilde{F}).$$

Also $\boldsymbol{E} = g(E)$, $\boldsymbol{F} = g(F)$. First of all we know that

$$|\boldsymbol{E} - \tilde{\boldsymbol{E}}| < \delta\sigma r_0. \tag{3.26}$$

Now, given a pair of points, $X \in \partial W$, $\tilde{X} \in \partial Q$ we want to estimate $|\boldsymbol{X}_1 - \tilde{\boldsymbol{X}}_1|$ so that we can apply Lemma 3.5. The length of the image of a vertical segment in ∂Q

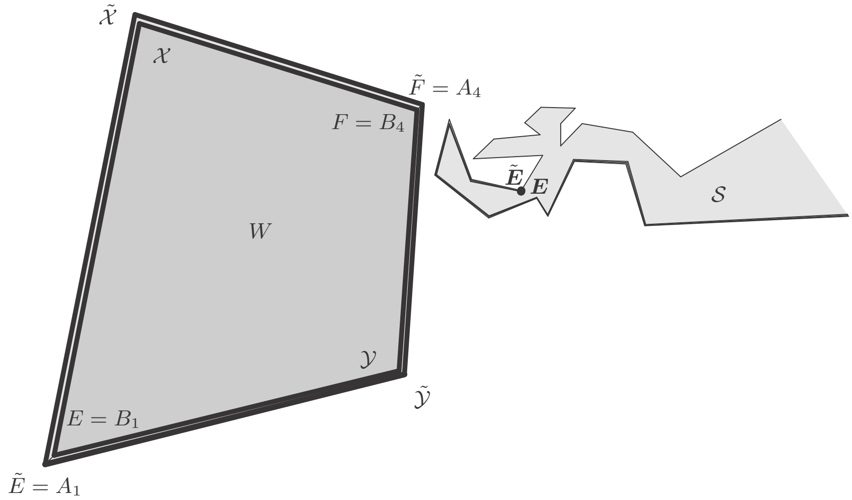


FIGURE 14. Dividing $Q \setminus W$ into two components bordered by $\mathcal{X}, \tilde{\mathcal{X}}, [E\tilde{E}], [F\tilde{F}]$ and $\mathcal{Y}, \tilde{\mathcal{Y}}, [E\tilde{E}], [F\tilde{F}]$ and similarly in the image by joining $\boldsymbol{E}$ and $\tilde{\boldsymbol{E}}$ with a segment in $\mathcal{S} \setminus g(W)$.

or ∂W must be less than $C\delta\sigma r_0$ by (3.12) and (3.22). This fact and (3.26) mean that if X and $\tilde{X}$ lie on vertical segments then

$$|\boldsymbol{X}_1 - \tilde{\boldsymbol{X}}_1| < C\delta\sigma r_0.$$

Points $\tilde{X}$ and X, which do not lie on vertical segments can be connected with $\tilde{E}$ and E with a direct curve in ∂Q, resp. ∂W. For such a pair $X = (X_1, X_2) \in \partial W$ and any $\tilde{X} = (\tilde{X}_1, \tilde{X}_2) \in \partial Q$ with $\tilde{\gamma}_X$ a direct curve in ∂W with endpoints E and X and $\tilde{\gamma}_{\tilde{X}}$ a direct curve in ∂Q with endpoints $\tilde{E}$ and $\tilde{X}$ we can calculate using (3.12), (3.22) and (3.26) to get

$$\begin{aligned} |g(\tilde{X}) - g(X)| &= \Big|\boldsymbol{E} - \tilde{\boldsymbol{E}} + \int_{\tilde{\gamma}_X} D_\tau g - \int_{\tilde{\gamma}_{\tilde{X}}} D_\tau g\Big| \\ &\leq \delta\sigma r_0 + \Big|M(X_1 - E_1) - M(\tilde{X}_1 - \tilde{E}_1) + (1 + C)\delta\sigma r_0\Big| \\ &\leq (2 + C)\delta\sigma r_0 + M|X_1 - \tilde{X}_1| < C\delta\sigma r_0, \end{aligned} \tag{3.27}$$

for an appropriate fixed C.

We proceed by separating the annulus into 2 parts as we depict in Figure 14. Recall how in (3.21) we separated ∂W into $\mathcal{X}$ and $\mathcal{Y}$. Now we make a similar division of ∂Q, call

$$\tilde{\mathcal{X}} = \Big\{\frac{X - \delta\sigma A}{1 - \delta\sigma};\ X \in \mathcal{X}\Big\} \text{ and } \tilde{\mathcal{Y}} = \Big\{\frac{Y - \delta\sigma A}{1 - \delta\sigma};\ Y \in \mathcal{Y}\Big\}.$$

Our choice of α guarantees that $[\tilde{\boldsymbol{E}}\boldsymbol{E}]$ and $[\tilde{\boldsymbol{F}}\boldsymbol{F}]$ lie in $\mathcal{S}^\circ \setminus g(W)$ apart from their endpoints. Now there are two bounded components of

$$\mathbb{R}^2 \setminus (g(\partial Q) \cup g(W) \cup [\boldsymbol{E}\tilde{\boldsymbol{E}}] \cup [\boldsymbol{F}\tilde{\boldsymbol{F}}]).$$

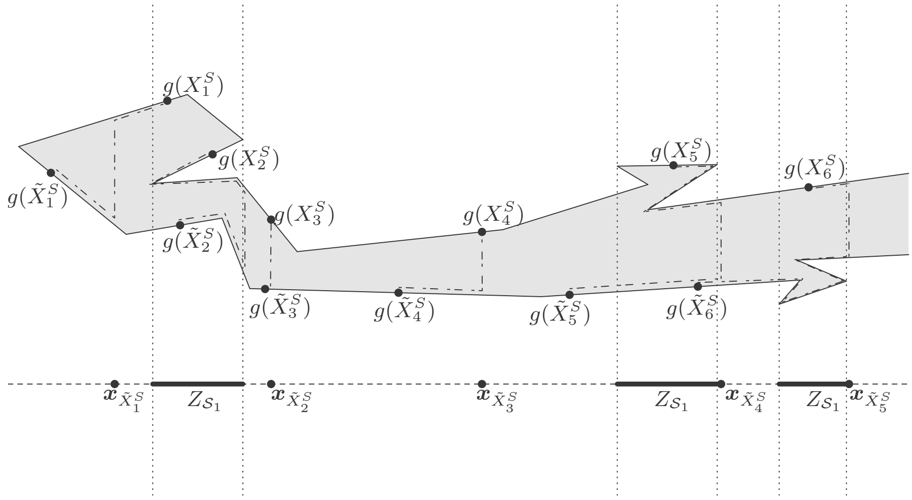


FIGURE 15. Our strategy for connecting points in $g(\mathcal{X})$ or $g(\mathcal{Y})$ with their corresponding points on the boundary of $\mathcal{S}_1$ or $\mathcal{S}_2$.

We take the completion of these components, which we will call $\mathcal{S}_1$ and $\mathcal{S}_2$ depending on their boundaries

$$\partial\mathcal{S}_1 = g(\mathcal{X}) \cup g(\tilde{\mathcal{X}}) \cup [\boldsymbol{E}\tilde{\boldsymbol{E}}] \cup [\boldsymbol{F}\tilde{\boldsymbol{F}}] \text{ and } \partial\mathcal{S}_2 = g(\mathcal{Y}) \cup g(\tilde{\mathcal{Y}}) \cup [\boldsymbol{E}\tilde{\boldsymbol{E}}] \cup [\boldsymbol{F}\tilde{\boldsymbol{F}}].$$

Obviously we define g linearly on $[E\tilde{E}]$ and $[F\tilde{F}]$ so that $g(E) = \boldsymbol{E}$ and $g(\tilde{E}) = \tilde{\boldsymbol{E}}$ and similarly for F and $\tilde{F}$. We already know by the definition of g that $D_\tau g$ on $[E\tilde{E}]$ is bounded. Therefore calling $U = \mathcal{X}\cup\tilde{\mathcal{X}}\cup[E\tilde{E}]\cup[F\tilde{F}]$ and $L = \mathcal{Y}\cup\tilde{\mathcal{Y}}\cup[E\tilde{E}]\cup[F\tilde{F}]$, with respect to (3.12) and (3.22), we get

$$\int_U \left| D_\tau g - \begin{pmatrix} M, 0 \\ 0, 0 \end{pmatrix} \tau \right| d\mathcal{H}^1 < C\delta\sigma r_0 \text{ and } \int_L \left| D_\tau g - \begin{pmatrix} M, 0 \\ 0, 0 \end{pmatrix} \tau \right| d\mathcal{H}^1 < C\delta\sigma r_0. \quad (3.28)$$

We want to apply Lemma 3.5 where $\varphi = g$, $\partial Q = U$ respectively L, $\varepsilon = C\delta\sigma r_0$ and $M = M$. Although the geometry of U and L does not correspond to the geometry of ∂Q this will not cause a problem. The result depends only on the geometry of the image, not the pre-image and the geometry of $\mathcal{S}_1$, the image of U in g corresponds to that in the lemma. To be specific take for example U. We move the bottom half of U, i.e. $\mathcal{X}\cup[E\tilde{E}]\cup[F\tilde{F}]$ vertically down onto $\tilde{\mathcal{Y}}$. Now we have a new map defined on ∂Q and since we move points in our pre-image vertically only, none of our estimates have altered at all. We apply the lemma and get a curve in $\mathcal{S}_1$ which is the image for our altered map. But the images of U in g and ∂Q in our altered map are identical. Therefore we can use the curve we found in the image for our altered map, also when we work with U.

This being clear, we know by Lemma 3.5 and (3.27), that if the side S of ∂Q is a subset of U then we can find a curve in $\mathcal{S}_1$ with endpoints $\boldsymbol{X}_i^S$ and $\tilde{\boldsymbol{X}}_i^S$, whose length is less than $C\delta r_0$, where the points $\boldsymbol{X}_i^S = g(X_i^S)$ and $\tilde{\boldsymbol{X}}_i^S = g(\tilde{X}_i^S)$ were chosen in (3.25). Obviously the same is true for sides of ∂Q in L and curves in $\mathcal{S}_2$ as well.

Step 6: Definition of g on the grid in the annulus $Q \setminus W$.

Now our aim is to slightly refine the curves we constructed in Lemma 3.5 to make sure that they are pairwise disjoint, without increasing their length significantly. We demonstrate our construction on $\mathcal{S}_1$ and then apply it to $\mathcal{S}_2$ as all the arguments are the same. We have depicted our strategy in Figure 15

Take the constant C in (3.28) and for every $\tilde{X} \in \tilde{\mathcal{X}}$ such that $X_1 < \tilde{F}_1 - 3C\delta\sigma r_0$ (recall from Step 2 that $\tilde{F}$ is a point at the right end of ∂Q) we find a unique x-coordinate in the image $\boldsymbol{x}_X \in \pi_1(\mathcal{S}_1) \setminus Z_{\mathcal{S}_1}$ with Z_{S_1} from Lemma 3.2. We choose $\boldsymbol{x}_X$ so that it is an increasing function of X_1. If $[A_1A_2]$ is vertical segment in $\tilde{\mathcal{X}}$ then we choose $\boldsymbol{x}_X$ as an increasing function of X_2 on $[A_1A_2]$. We write $\boldsymbol{X} = (\boldsymbol{X}_1, \boldsymbol{X}_2) = g(X)$. Since $\mathcal{H}^1(Z_{\mathcal{S}_1}) < C\delta\sigma r_0$ we can choose $\boldsymbol{x}_X$ such that $|\boldsymbol{x}_X - \boldsymbol{X}_1| < C\delta\sigma r_0$. Let α_0^1 be the number defined in (3.2) in Lemma 3.1, where $P = \mathcal{S}_1$. If m_1 denotes the number of vertices in $\mathcal{S}_1$, then for all $X \in \mathcal{X}$ we choose an

$$\alpha_X \in \left(\frac{C\alpha_0^1 \delta\sigma r_0}{4(4m_1+8)}, \frac{C\alpha_0^1 \delta\sigma r_0}{2(4m_1+8)} \right)$$

such that α_X is a decreasing function of X_1, (respectively an increasing function of X_2 on $[B_1B_2]$ if this segment is vertical).

If we have a side S of ∂Q and $S \subset U$, we want to find curves in $\mathcal{S}_1$ with endpoints $\tilde{X}_i^S$ and X_i^S for $i = 1, \dots m$, which we will call γ_i^S. We recall the estimate (3.27), which we achieved thanks to (3.12) and (3.22), as it is critical for this step. We use the construction in Lemma 3.5 (with $\varepsilon = C\delta\sigma r_0$ by (3.28)) but choose some of the parameters specifically. Part of our construction was following the boundary using a curve in P_α (defined in Lemma 3.1) for α sufficiently small. Now we use a curve in $P_{\alpha_{X_i^S}}$ to follow the boundary. Also we choose $\overline{\boldsymbol{x}} = \boldsymbol{x}_X$ in our application of Lemma 3.4. There exists an absolute constant C such that the length of all γ_i^S is bounded by $C\delta\sigma r_0$. We define g on $[\tilde{X}_i^S X_i^S]$ as the constant-speed parametrization of the curve γ_i^S. Remember we have already defined g on $[E\tilde{E}]$ and $[F\tilde{F}]$ which may correspond to $[\tilde{X}_0^S \tilde{X}_0^S]$ or $[\tilde{X}_m^S \tilde{X}_m^S]$. Since the length of γ_i^S is bounded by $C\delta\sigma r_0$ we get an absolute constant K such that

$$|D_\tau g| \leq K \ \mathcal{H}^1\text{-almost everywhere on } [X_i^S, \tilde{X}_i^S]. \tag{3.29}$$

Step 7: Estimates on $\int_{Q\setminus W} \Phi(|Dg|)$.

It is very easy to notice that $R = \mathrm{co}\{X_i^S, \tilde{X}_i^S, X_{i+1}^S, \tilde{X}_{i+1}^S\}$ is 8-bi-Lipschitz equivalent with a square of diameter $\delta\sigma r_0$. Upon calculating the L^Φ modular of the derivative on ∂R we can use the bi-Lipschitz piecewise-affine change of variables and then apply Theorem 2.1, take the mapping from the theorem and reverse the bi-Lipschitz change of variables and we will have defined g.

We know from (3.29) that the derivative of g is bounded on $[\tilde{X}_i^S X_i^S]$ by a constant K and bounded by $2M$ on ∂W, as was shown in (3.19). We estimate,

$$\begin{aligned}\fint_{\partial R}\Phi\big(|D_\tau g|\big)d\mathcal{H}^1 &\leq \frac{C}{\delta\sigma r_0}\Big(\int_{[X_i^S\tilde{X}_i^S]}\Phi\big(|D_\tau g|\big)d\mathcal{H}^1+\int_{[X_{i+1}^S\tilde{X}_{i+1}^S]}\Phi\big(|D_\tau g|\big)d\mathcal{H}^1\\&\qquad+\int_{[X_i^S X_{i+1}^S]}\Phi\big(|D_\tau g|\big)d\mathcal{H}^1+\int_{[\tilde{X}_i^S\tilde{X}_{i+1}^S]}\Phi\big(|D_\tau g(t)|\big)d\mathcal{H}^1(t)\Big)\\&\leq C\Phi(K)+C\Phi(2M)+C\frac{\int_{[\tilde{X}_i^S,\tilde{X}_{i+1}^S]}\Phi\big(|D_\tau g|\big)d\mathcal{H}^1}{\delta\sigma r_0}.\end{aligned}$$

Now using Theorem 2.1 we get an extension of g on R such that

$$\int_R\Phi\big(|Dg|\big)\leq C\delta^2\sigma^2r_0^2\Phi(K)+C\delta^2\sigma^2r_0^2\Phi(2M)+C\delta\sigma r_0\int_{[\tilde{X}_i^S,\tilde{X}_{i+1}^S]}\Phi\big(|D_\tau g(t)|\big)d\mathcal{H}^1(t).$$

We sum this over all quadrilaterals in $Q\setminus W$ of which there are less than $5\delta^{-1}\sigma^{-1}$, then we use the Δ_2 quality of Φ, the definition of σ and (3.13) to get

$$\begin{aligned}\int_{Q\setminus W}\Phi\big(|Dg|\big)&\leq C\frac{\delta^2\sigma^2r_0^2\big(\Phi(2M)+\Phi(K)\big)}{\delta\sigma}+C\delta\sigma r_0\int_{\partial Q}\Phi\big(|D_\tau g(t)|\big)d\mathcal{H}^1(t)\\&\leq C\delta r_0^2+C\delta\sigma r_0\int_{\partial Q}\Phi(M)\\&\leq C\delta r_0^2.\end{aligned}\tag{3.30}$$

This constant C depends only on Φ.

The combination of (3.24) and (3.30) together prove our claim. □

4. Proof of Theorem 1.1

4.1. **A preliminary approximation result on bounded domains.** The following theorem is the mainstay of the proof of Theorem 1.1.

Theorem 4.1. *Let $\Omega\subseteq\mathbb{R}^2$ be a bounded domain and let ε_k be a decreasing sequence of positive numbers tending to zero. Let Φ be a Δ_2-Young function. Then for any homeomorphism $f\in W^{1,\Phi}(\Omega,\mathbb{R}^2)$ there exists a sequence of diffeomorphisms $f_k\in W^{1,\Phi}(\Omega,\mathbb{R}^2)$ such that*

$$\|f-f_k\|_{W^{1,\Phi}(\Omega,\mathbb{R}^2)}\leq\varepsilon_k,$$

where $\|\cdot\|_{W^{1,\Phi}(\Omega,\mathbb{R}^2)}$ is the Luxembourg norm. Moreover if f can be continuously extended onto the boundary, then f_k coincides with f on $\partial\Omega$ and f_k converge to f uniformly.

Proof. *Step 0: Smoothing piecewise-affine maps.*

We will find (countably) piecewise-affine homeomorphisms f_k such that

$$\|Df-Df_k\|_{W^{1,\Phi}(\Omega,\mathbb{R}^2)}<\varepsilon_k.$$

The construction of f_k gives us a countable (possibly finite) covering of Ω with triangles $\mathcal{T}_i=\mathrm{co}\{X_i,Y_i,Z_i\}, i\in\mathbb{N}$ such that f_k is affine on each triangle. We choose an essentially disjoint covering of Ω with polygons W_j. We choose the vertices of each polygon W_j so that the boundary of W_j intersects the boundary of our triangles only at the midpoint of a side. Let $L_{j,k}$ denote the Lipschitz constant of f_k on W_j. We apply

[18, Theorem A] on f_k choosing $\Omega = W_j$, $p = 1$, $q = 1$, $\varepsilon = 14L_{j,k}\varepsilon_k/2^j\Phi(14L_{j,k}\varepsilon_k^{-1})$ and get diffeomorphisms $\tilde{f}_{j,k}$ on W_j.

Set $\tilde{f}_k(x) = \tilde{f}_{j,k}(x)$ for $x \in \overline{W_j}$. We claim that we can carry out the above smoothing in such a way that $\tilde{f}_k$ is a diffeomorphism on Ω. In the process of applying the above theorem we take care that if the side $[X_iY_i]$ of the triangle $\mathcal{T}_i$ intersects two of our polygons W_{j_1} and W_{j_2} then we choose the same parameter for smoothing along that side of the triangle in both polygons. This ensures that $\tilde{f}_k$ is indeed a diffeomorphism because the boundary of the polygons do not approach the vertices of the triangles, $\tilde{f}_k$ has been set to be a diffeomorphism on sides which intersect the boundary of a polygon and away from the boundary of the triangles we have $\tilde{f}_k = f_k$ which is an affine map in the set in question.

We have $|D\tilde{f}_k| \leq 13L_{j,k}$ almost everywhere in W_j and therefore $|D\tilde{f}_k - Df_k| \leq 14L_{j,k}$ almost everywhere in W_j. Since Φ is convex and $\Phi(0) = 0$, we know that

$$\Phi(t) \leq t\frac{\Phi(14L_{j,k}\varepsilon_k^{-1})}{14L_{j,k}\varepsilon_k^{-1}}$$

for $t \leq 14L_{j,k}\varepsilon_k^{-1}$. Therefore

$$\int_{W_j}\Phi\Big(\frac{|D\tilde{f}_k - Df_k|}{\varepsilon_k}\Big) \leq \frac{\Phi(14L_{j,k}\varepsilon_k^{-1})}{14L_{j,k}}\int_{W_j}|D\tilde{f}_k - Df_k| < 2^{-j}.$$

Since we may require that $\|f - f_k\| < \tilde{\varepsilon}$ for any $\tilde{\varepsilon} > 0$ we will have no difficulty in estimating $\|\tilde{f}_k - f_k\|_{L^\Phi}$. Therefore we have

$$\|D\tilde{f}_k - Df_k\|_{W^{1,\Phi}(\Omega,\mathbb{R}^2)} < \varepsilon_k \text{ and } \|Df - D\tilde{f}_k\|_{W^{1,\Phi}(\Omega,\mathbb{R}^2)} < 2\varepsilon_k.$$

Step 1: Basic setup.

We will choose some $\varepsilon > 0$ and show how to find an approximation f_1 such that

$$\int_\Omega \Phi(|f - f_1|) + \Phi(|Df - Df_1|) < \varepsilon$$

given that $\mathcal{L}^2(\Omega) < \infty$. This will then imply that for any $\varepsilon_k > 0$ we can find f_k such that $\|f - f_k\|_{W^{1,\Phi}(\Omega,\mathbb{R}^2)} < \varepsilon_k$. Therefore let us have $\varepsilon > 0$ arbitrary.

We will determine and restrict some set, where the behaviour of f is bad. By the absolute continuity of the Lebesgue integral we can find $0 < \eta < \varepsilon$ such that

$$\int_A \Phi(|Df|) < \varepsilon \text{ for any set } A \text{ such that } |A| < 16\eta. \tag{4.1}$$

Without loss of generality we can assume that $J_f \geq 0$ a.e. (see e.g. [11, Theorem 5.22]). By fixing M large enough and a small enough $\delta > 0$ (also require $\delta < \varepsilon$) we will get a set F with $|F| < \frac{\eta}{4}$, where

$$\begin{aligned} F &= F_1 \cup F_2 \cup F_3\\ F_1 &= \big\{x \in \Omega :\ |Df(x)| > M\big\}\\ F_2 &= \big\{x \in \Omega :\ |Df(x)| < M^{-1}\big\}\\ F_3 &= \big\{x \in \Omega :\ J_f(x) \neq 0 \text{ but } \exists v_1, v_2 \in \mathbb{R}^2 \text{ with } |v_1| = |v_2| = 1,\\ &\qquad |v_1 \cdot v_2| < \frac{1}{2} \text{ and } |Df(c)v_1 - Df(c)v_2| < 8\delta\big\}. \end{aligned} \tag{4.2}$$

We put $\sigma = \min\{1, M^{-1}, \big(\Phi(M)\big)^{-1}\}$.

We will subdivide our domain Ω using a Whitney-type covering. On the majority of our squares f will behave well and the centre of such squares will lie in a so-called good set G. To this purpose we recall a well-known result (see [9] or [11, Lemma A.28]) which says that a planar homeomorphism in the Sobolev space $W^{1,1}_{\rm loc}(\Omega, \mathbb{R}^2)$ is differentiable almost everywhere. It follows that almost every point of Ω is a point of differentiability for f and a Lebesgue point of Df and $\Phi(|Df|)$. Simply by the convexity of Φ we get

$$\Phi(b-a) = \int_0^{b-a} \Phi'(s) ds \leq \int_a^b \Phi'(s) ds = \Phi(b) - \Phi(a).$$

Therefore at all Lebesgue points X of $\Phi(|Df|)$ we know that

$$\lim_{h\to 0+} \fint_{Q(X,h)} \Phi(|Df(X) - Df(Y)|)\, dY = 0.$$

Therefore if we choose a h_0 which satisfies

$$h_0 < \frac{\varepsilon}{M} \tag{4.3}$$

and is small enough, we get the set

$$\begin{aligned} G := \Big\{X \in \Omega :\ & \text{for every } 0 < h \leq \min\{h_0, \operatorname{dist}(X, \partial\Omega)\} \text{ we have} \\ & \fint_{Q(X,4h)} \Phi(|Df(X) - Df(Y)|)\, dY < \delta\sigma, \\ & \fint_{Q(X,4h)} |Df(X) - Df(Y)|\, dY < \delta\sigma \text{ and} \\ & |f(Y) - f(X) - Df(X)(Y-X)| < \delta h \text{ for all } y \in Q(X,4h)\Big\}, \end{aligned} \tag{4.4}$$

which satisfies $|\Omega \setminus G| < \frac{\eta}{4}$.

Step 2: *A good covering of* Ω.

It follows from the definition of G and F that we can find a Whitney-style covering of Ω with squares that have pairwise disjoint interior and $\operatorname{diam}(Q) = 2^{-j_Q} h_0$ with $j_Q \in \mathbb{N}$. We call the set of squares $\mathcal{C}_0$ and $\Omega = \bigcup_{Q\in\mathcal{C}_0} Q$, which also has

$$\sum_{Q(c,r)\in I} |Q(c,r)| > |\Omega| - \eta, \text{ where } I = \{Q(c,r) \in \mathcal{C}_0 :\ c \in G \text{ and } c \notin F\}. \tag{4.5}$$

We refer to two squares in our grid as neighbours is they have a common vertex. We require that the squares in our covering satisfy the following condition: If Q_1 and Q_2 are neighbours and Q_2 and Q_3 are neighbours then the ratio of the radii of any pair of these three squares is at most 2. That is $\operatorname{diam} Q_i \leq 2 \operatorname{diam} Q_j \leq 4 \operatorname{diam} Q_i$ for $i, j = 1, 2, 3$. This fact means that if

$$Q \in \mathcal{C}_0 \text{ then } \operatorname{card}\{Q' \in \mathcal{C}_0 : 2Q \cap Q' \neq \emptyset\} \leq \iota' \tag{4.6}$$

and ι' is universal for all Q irrespective of our choice of $\mathcal{C}_0$. We will alter our grid in the next step of our construction. This may mean that we have a different (possibly larger) value ι such that (4.6) is satisfied on our final grid but the value ι is also an absolute constant.

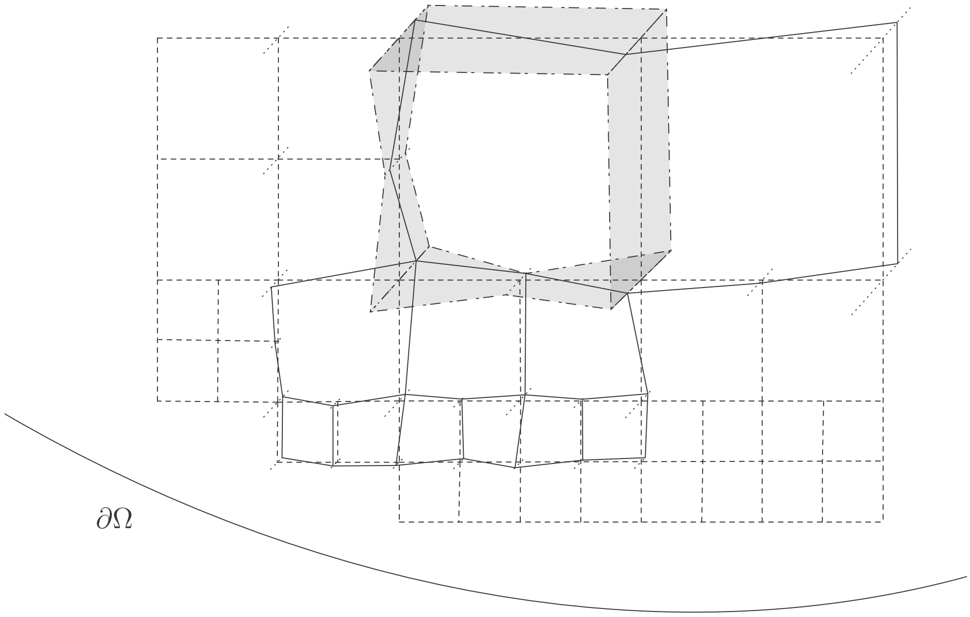


FIGURE 16. Modifying the initial grid in Ω to achieve nice boundary values on the squares.

We classify the squares in our covering as good, null and bad as follows

$$\begin{aligned}
\mathcal{G}_0 &= \{Q(c,r) \in \mathcal{C}_0 : \ c \in G \setminus F\},\\
\mathcal{N}_0 &= \{Q(c,r) \in \mathcal{C}_0 : \ c \in G \setminus (F_1 \cup F_2), \ J_f(c) = 0\},\\
\mathcal{B}_0 &= \mathcal{C}_0 \setminus (\mathcal{G}_0 \cup \mathcal{N}_0).
\end{aligned}$$

The boundaries of our squares create a grid in Ω and we would like the following estimates to hold

$$⨍_{\partial Q(c,r)} \Phi(|D_\tau f|) \, d\mathcal{H}^1 \leq C ⨍_{Q(c,r)} \Phi(|Df|) \tag{4.7}$$

and

$$⨍_{\partial Q(c,r)} \Phi(|D_\tau f - Df(c)\tau|) \, d\mathcal{H}^1 < C\delta\sigma \text{ if } c \in G, \tag{4.8}$$

(for the relevant definitions see the preliminaries). Nevertheless, the estimates (4.7) and (4.8) may not hold on squares $Q \in \mathcal{C}_0$ and therefore we need to slightly shift the corners of these squares to have better boundary values, which we depict in Figure 16. In this way we obtain a covering of Ω by quadrangles that are close to squares.

We start by considering adjoining squares which have the same radius r and later we explain how to proceed in the general case. For each vertex $V = [v_1, v_2]$ in the grid (corner point of some square) we define a segment of length $\sqrt{2}r/4$ through this point

$$S_V := \{[x,y] : x \in [v_1 - \tfrac{r}{8}, v_1 + \tfrac{r}{8}], \ y - v_2 = x - v_1\}.$$

We claim that for every square $\mathrm{co}\{V_1, V_2, V_3, V_4\} = Q \in \mathcal{C}$, there exists a quadrilateral $\mathcal{Q} = \mathrm{co}\{\tilde{V}_1, \tilde{V}_2, \tilde{V}_3, \tilde{V}_4\}$ with $\tilde{V}_i \in S_{V_i}$ such that f is absolutely continuous on each side

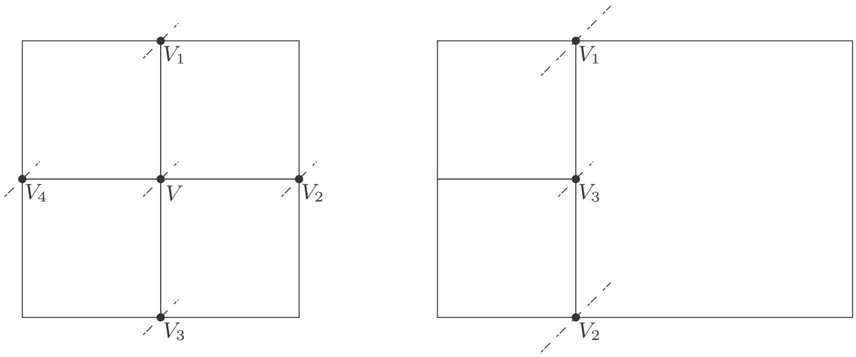


FIGURE 17. Neighbouring corners of the grid.

of $\partial\mathcal{Q}$ and further

$$\begin{aligned} \fint_{\partial\mathcal{Q}} \Phi(|D_\tau f|) &\le C \fint_{2Q} \Phi(|Df|), \\ \fint_{\partial\mathcal{Q}} \Phi(|D_\tau f - Df(c)\tau|) &< C\delta\sigma \quad \text{ if } c \in G. \end{aligned} \tag{4.9}$$

We prove this claim as follows. The mapping f is absolutely continuous on almost all lines parallel to a fixed direction because $f \in W^{1,1}_{\mathrm{loc}}$. It follows that for almost all $(X_1, X_2) \in S_{V_1} \times S_{V_2}$ (with respect to two dimensional measure) that f is absolutely continuous on the segment $[X_1X_2]$. A simple computation shows that

$$\begin{aligned} &\int_{S_{V_1}} \int_{S_{V_2}} \Big(\int_{[X_1X_2]} \Phi(|D_\tau f(t)|)\, d\mathcal{H}^1(t) \Big)\, d\mathcal{H}^1(X_2)\, d\mathcal{H}^1(X_1) \le C \operatorname{diam}(Q) \int_{2Q} \Phi(|Df|), \\ &\int_{S_{V_1}} \int_{S_{V_2}} \Big(\int_{[X_1X_2]} \Phi(|D_\tau f(t) - Df(c)\tau|)\, d\mathcal{H}^1(t) \Big)\, d\mathcal{H}^1(X_2)\, d\mathcal{H}^1(X_1) \\ &\qquad\qquad\qquad \le C \operatorname{diam}(Q) \int_{2Q} \Phi(|Df - Df(c)|), \end{aligned}$$

as $\mathcal{H}^1(S_{V_1}) \approx \operatorname{diam} Q$ and $\operatorname{co}(S_{V_1} \cup S_{V_2}) \subset 2Q$. Similar estimates hold for any two neighbouring vertices.

In fact, we basically have

$$\fint_{S_{V_1} \times S_{V_2}} \Big(\fint_{[X_1X_2]} \Phi(|D_\tau f(t)|)\, d\mathcal{H}^1(t) \Big)\, d\mathcal{L}^2 \le C \fint_{2Q} \Phi(|Df|). \tag{4.10}$$

Let a be a non-negative function on a finite-measure space A and $\lambda \ge 1$. It is not difficult to observe that

$$\mu\Big(\Big\{ a > \lambda \fint_A a \Big\} \Big) < \frac{\mu(A)}{\lambda} \; . \tag{4.11}$$

It follows that we can find a constant C_1 such that: For each vertex V with neighbouring vertices V_1, V_2, V_3 and V_4 depicted in Figure 17 we can find a set $\mathcal{S}_V \subset S_V$ with $|\mathcal{S}_V| \ge \frac{4}{5}|S_V|$ and for every $X \in \mathcal{S}_V$ we know that

$$\Big| \Big\{ Y \in S_{V_i} : \ \operatorname{diam} Q \int_{[XY]} \Phi(|D_\tau f|) \le C_1 \int_{2Q} \Phi(|Df|) \Big\} \Big| \ge \frac{4}{5} |S_{V_i}|, \tag{4.12}$$

for each $i = 1, 2, 3, 4$. In the last equation we know that $\int_{XY} \Phi(|D_\tau f|)$ exists, as f is absolutely continuous on the segment $[XY]$. Further, in case of $c \in G$ we use (4.4) and we can require that

$$\begin{aligned}\Big|\Big\{Y \in S_{V_i} :\ \operatorname{diam} Q \int_{[XY]} \Phi(|D_\tau f|) \leq C_1 \int_{2Q} \Phi(|Df|) \text{ and}\\ \fint_{[XY]} \Phi(|D_\tau f - Df(c)\tau|) \leq C_1 \delta\sigma\Big\}\Big| > \frac{4}{5}|S_{V_i}|.\end{aligned} \tag{4.13}$$

The second case depicted in Figure 17, where we have a common side of a larger square with corners V_1 and V_2 and two smaller squares, whose common corner divides this side in half at V_3 is practically identical. Here instead of (4.10) we have

$$\fint_{S_{V_1}\times S_{V_2}\times S_{V_3}} \Big(\fint_{[X_1X_2]} \Phi(|D_\tau f(t)|)\, d\mathcal{H}^1(t)\Big)\, d\mathcal{L}^3 \leq C \fint_{2Q} \Phi(|Df|),$$

where, of course, $\mathcal{H}^1(S_{V_3})$ is half the size of the other corners.

We will also need our approximation to be good with respect to the L^1 norm. In order to achieve this we repeat the above process so that the following estimates hold

$$\begin{aligned}\fint_{\partial\mathcal{Q}} |D_\tau f| \leq C \fint_{2\mathcal{Q}} |Df|,\\ \fint_{\partial\mathcal{Q}} |D_\tau f - Df(c)\tau| < C\delta\sigma \quad \text{ if } c \in G.\end{aligned} \tag{4.14}$$

which can be achieved at the cost of increasing the absolute constant C_1.

For each vertex V we have found three or four sets $\mathcal{S}_V$ corresponding to the three or four vertices connected to our vertex V. Thanks to (4.12) and (4.13) we have that the intersection of these sets has positive measure and therefore is non-empty. We will now replace each vertex with a new one. It suffices to order our vertices according to some countable index set and sequentially choose the new vertices. If we have already chosen a neighbouring vertex we must take care when choosing our new vertex so that a (4.14)-type estimate holds on that segment. As we always choose our points from the sets $\mathcal{S}_V$, we will know that we can find appropriate points for neighbouring vertices. Hereby, for every $Q(c)$ in our original grid, we have a $\mathcal{Q}(c)$ and we will have satisfied the estimate (4.9) on the boundary of all $\mathcal{Q}$. Further we set

$$\begin{aligned}\mathcal{C} = \{\mathcal{Q} : Q \in \mathcal{C}_0\},\ \mathcal{G} = \{\mathcal{Q} : Q \in \mathcal{G}_0\},\\ \mathcal{B} = \{\mathcal{Q} : Q \in \mathcal{B}_0\},\ \mathcal{N} = \{\mathcal{Q} : Q \in \mathcal{N}_0\}.\end{aligned}$$

As each $\mathcal{Q}(c) \in \mathcal{C}$ is close to $Q(c, r) \in \mathcal{C}_0$ it is easy to check that

$$\begin{aligned}&(i)\ \fint_{2\mathcal{Q}(c)} \Phi(|Df(c) - Df(Y)|)\, dY < C\delta\sigma \text{ for all } \mathcal{Q} \in \mathcal{G} \cup \mathcal{N}\\ &(ii)\ |f(Y) - f(c) - Df(c)(Y - c)| < \delta r \text{ for all } Y \in 2\mathcal{Q} \text{ and } \mathcal{Q} \in \mathcal{G} \cup \mathcal{N}.\end{aligned} \tag{4.15}$$

We would like to estimate the area contained in bad squares. Since the diameter of the distorted square is bounded by 2 times the diameter of the original square, we can use (4.5) and the fact that $\Omega = \bigcup_{\mathcal{Q}\in\mathcal{C}}$ to deduce that

$$\sum_{\mathcal{Q}\in\mathcal{B}} |\mathcal{Q}| < 4\eta. \tag{4.16}$$

<u>Step 3:</u> *Defining a piecewise-linear approximation of f on the grid $\mathcal{R}$.*

We need to find a piecewise-linear mapping φ on the grid $\mathcal{R} := \bigcup_{Q\in\mathcal{C}} \partial Q$ which is one-to-one and for all $Q(c,r) \in \mathcal{C}$ satisfies

$$\begin{aligned} \int_{\partial Q} \Phi(|D_\tau\varphi(t) - D_\tau f(t)|)\, d\mathcal{H}^1(t) < C\delta\sigma r \\ \|f-\varphi\|_{L^\infty(\partial Q)} < C\delta r. \end{aligned} \tag{4.17}$$

Recall that f is absolutely continuous on ∂Q for all $Q \in \mathcal{C}$ and hence the integral is well defined.

First we will do this for good $\mathcal{Q} \in \mathcal{G}$. We start by denoting the vertices of $\mathcal{Q}$ as K_i, $i = 1, 2, \dots, o, o+1$ with $K_1 = K_{o+1}$, where $o \leq 8$ and $[K_i, K_{i+1}] \subset \partial\mathcal{Q}$. Since $\mathcal{Q} \in \mathcal{G}$, we can simply use φ that is linear on each side $[K_iK_{i+1}]$ and connects points $f(K_i)$ and $f(K_{i+1})$. Now if we denote $\mathcal{A}(X) := f(c) + Df(c)(X-c)$ as the affine function that well approximates f on Q. We can use the definition of the good set (4.4) (i.e. $|f(X) - \mathcal{A}(X)| \leq \delta r$) to obtain

$$\begin{aligned} \Big|D_\tau\varphi - Df(c)\tau\Big| &= \Big|\frac{f(K_i)-f(K_{i+1})}{|K_i-K_{i+1}|} - Df(c)\frac{(K_i-K_{i+1})}{|K_i-K_{i+1}|}\Big| \\ &\leq \frac{|f(K_i)-\mathcal{A}(K_i)| + |f(K_{i+1})-\mathcal{A}(K_{i+1})|}{|K_i-K_{i+1}|} \leq C\delta. \end{aligned} \tag{4.18}$$

Together with (4.9) (note that as Φ is a young function, we know that there exists some C such that $\Phi(s) < Cs$ for small s), this implies the first inequality in (4.17) for $Q \in \mathcal{G}$. For the second inequality consider $t \in [K_iK_{i+1}]$, use (4.4) for t and the fact that difference of two linear functions is biggest at the end of a segment to obtain

$$\begin{aligned} |f(t)-\varphi(t)| &\leq \delta r + |\varphi(t) - \mathcal{A}(t)| \\ &\leq \delta r + \max\{|f(K_i)-\mathcal{A}(K_i)|, |f(K_{i+1})-\mathcal{A}(K_{i+1})|\} \leq 2\delta r. \end{aligned} \tag{4.19}$$

We need to show that this approximation φ is one-to-one on ∂Q and well-oriented. We know that $J_f(c) > 0$ (as $J_f \geq 0$ a.e.) and as Q is close to the square that the scalar product satisfies $|\frac{(K_1-K_2)}{|K_1-K_2|} \cdot \frac{(K_3-K_2)}{|K_3-K_2|}| < \frac{1}{2}$. Using (4.2) (recall $c \notin F$ by the definition of $\mathcal{G}_0$) we obtain that

$$|\mathcal{A}(K_1) - \mathcal{A}(K_3)| = |Df(c)(K_1-K_2) - Df(c)(K_3-K_2)| > 4\delta r\ .$$

Together with (4.4) this implies that

$$|f(K_1) - f(K_3)| \geq |\mathcal{A}(K_1) - \mathcal{A}(K_3)| - 2\delta r > \delta r$$

and it is also easy to see that the orientation of vectors $f(K_1) - f(K_2)$ and $f(K_3) - f(K_2)$ is the same as orientation of vectors $\mathcal{A}(K_1) - \mathcal{A}(K_2)$ and $\mathcal{A}(K_3) - \mathcal{A}(K_2)$ and hence the same as orientation of $K_1 - K_2$ and $K_3 - K_2$, i.e. all three couples are either clockwise or all anticlockwise. This shows that φ is one-to-one on ∂Q and well-oriented. We now have everything we need to define our approximation on the good $\mathcal{Q} \in \mathcal{G}$ but we will do this after we have defined φ on the rest of the grid $\mathcal{R}$.

Now we need to construct φ on parts of $\mathcal{R}$ which are not borders of squares in $\mathcal{G}$ but lie adjacent to them. We need φ to be piecewise-linear and one-to-one close to the vertices of $Q \in \mathcal{G}$. Our setup is pictured in Figure 18. Let $Q_1 \in \mathcal{C} \setminus \mathcal{G}$ have a common boundary with $Q \in \mathcal{G}$ and let $[X_1X_2]$ be a segment in the boundary of Q_1 with $X_1 \in \overline{Q}$ but $X_2 \notin \overline{Q}$ and set $X_3 = \frac{1}{2}X_1 + \frac{1}{2}X_2$. As $Q \in \mathcal{G}$ is good and f is close to the affine mapping $\mathcal{A}$ on $2\mathcal{Q}$ (see (4.15)) it is easy to see that we can define φ linearly

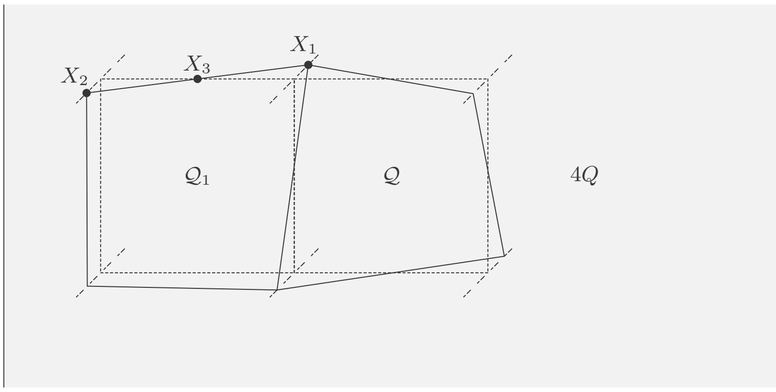


FIGURE 18. Lay of points X_1, X_2 and X_3 on $\mathcal{Q}$ and $\mathcal{Q}_1$ inside $4Q$.

on the line segment $[X_1X_3]$ with $\varphi(X_1) = f(X_1)$ and $\varphi(X_3) = f(X_3)$. Analogously as before we have (4.17) on the segment $[X_1X_3]$ and the function φ is again one-to-one on all boundaries of $Q \in \mathcal{G}$ and all half-segments touching Q. Now we need to define φ on other parts of our skeleton $\mathcal{R}$. It is clear that the image in f of the part of the grid $\mathcal{R}$, where we have not yet defined φ has positive distance from the image of $\partial\mathcal{Q}$ in φ (or f) for all $\mathcal{Q} \in \mathcal{G}$. Therefore it will be easy to approximate f on this part of the grid using segments (as shown in Figure 19) while guaranteeing that φ is one-to-one on $\mathcal{R}$.

Let us take a vertex $V \in \overline{\mathcal{Q}}$ of a square $\mathcal{Q} \in \mathcal{C} \setminus \mathcal{G}$. Assume that our vertex is inside Ω and therefore it is the vertex of three or four squares. We may as well assume that V is the vertex of four squares $\mathcal{Q}_1, \mathcal{Q}_2, \mathcal{Q}_3, \mathcal{Q}_4$ (the reasoning is analogous if V is the vertex of three squares). We will assume that $Q_1, Q_2, Q_3, Q_4 \in \mathcal{C}$ are clockwise oriented, which corresponds to $V_i \in \mathcal{Q}_i$, $i = 1, 2, 3, 4$ in Figure 19. For every vertex $V \in \Omega$ of $\mathcal{Q} \in \mathcal{C} \setminus \mathcal{G}$ we choose a radius r_V small enough such that

$$\begin{aligned} &(i)\ B_V := \mathcal{B}(f(V), r_V) \text{ are pairwise disjoint,} \\ &(ii)\ r_V \leq \delta, \\ &(iii) \int_{f^{-1}(B_V)\cap\partial\mathcal{Q}} \Phi(|D_\tau f|) \leq \delta\sigma \int_{\partial\mathcal{Q}} \Phi(|D_\tau f|) \text{ for all } \mathcal{Q} \in \mathcal{C}. \end{aligned} \tag{4.20}$$

Then we can find points

$$\begin{aligned} &\boldsymbol{P}_1 \in \partial B_V \cap f(\partial Q_1 \cap \partial Q_2),\ \boldsymbol{P}_2 \in \partial B_V \cap f(\partial Q_2 \cap \partial Q_3), \\ &\boldsymbol{P}_3 \in \partial B_V \cap f(\partial Q_3 \cap \partial Q_4),\ \boldsymbol{P}_4 \in \partial B_V \cap f(\partial Q_4 \cap \partial Q_1), \end{aligned} \tag{4.21}$$

such that $P_i = f^{-1}(\boldsymbol{P}_i)$ is the most distant point from V lying in $\partial Q_i \cap \partial Q_{i+1}$ and $\boldsymbol{P}_i$ satisfies (4.21) (here $Q_5 = Q_1$). It is easy to see that $\boldsymbol{P}_1$, $\boldsymbol{P}_2$, $\boldsymbol{P}_3$ and $\boldsymbol{P}_4$ are also clockwise oriented. We define φ on the line segment $[VP_i]$ as the linear map determined by $\varphi(V) = f(V)$ and $\varphi(P_i) = f(P_i) = \boldsymbol{P}_i$. This means that on $[VP_i]$ we have

$$\fint_{[VP_i]} D_\tau\varphi = \fint_{[VP_i]} D_\tau f$$

$$\fint_{[VP_i]} |D_\tau\varphi| = \left|\fint_{[VP_i]} D_\tau\varphi\right| = \left|\fint_{[VP_i]} D_\tau f\right| \leq \fint_{VP_i} |D_\tau f|$$

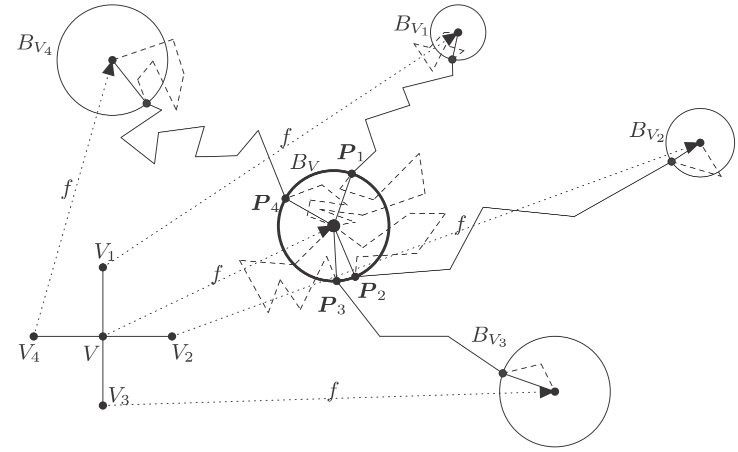


FIGURE 19. Defining f on sides $[VV_i]$ which do not belong to a square in $\mathcal{G}$.

and since $D_\tau\varphi$ is constant on the segment $[VP_i]$, we get

$$\fint_{[VP_i]} \Phi(|D_\tau\varphi|) = \Phi\Big(\fint_{[VP_i]} |D_\tau\varphi|\Big) \leq \Phi\Big(\fint_{[VP_i]} |D_\tau f|\Big) \leq \fint_{[VP_i]} \Phi(|D_\tau f|).$$

Given $Q \in \mathcal{C}$ with corners V^1, V^2, V^3, V^4 we have already defined φ close to these vertices and by (4.20) (ii) and (iii) we have

$$\begin{aligned}\int_{f^{-1}(B_v)\cap\mathcal{R}} \Phi(|D_\tau f - D_\tau\varphi|) &\leq C\int_{f^{-1}(B_v)\cap\mathcal{R}} \Phi(|D_\tau f|) + C\int_{f^{-1}(B_v)\cap\mathcal{R}} \Phi(|D_\tau\varphi|)\\ &\leq C\int_{f^{-1}(B_v)\cap\mathcal{R}} \Phi(|D_\tau f|) < C\delta\sigma\int_{\partial Q} \Phi(|D_\tau f|).\end{aligned}$$

As f is absolutely continuous on ∂Q we can simply take a fine net of points in the rest of $\mathcal{R}$ and define φ as piecewise-linear on segments and $\varphi(X) = f(X)$ for every X endpoint of such a segment. As the points $\boldsymbol{P}_i$ around each vertex are clockwise oriented it is not difficult to see that we can achieve both (4.17) and have φ one-to-one simply by making our division fine enough. Also we make our net fine enough to garantee that the image of the grid lies inside $f(\Omega)$.

Step 4: Defining a piecewise-affine approximation of f on Ω.

Now we are in a position to define our piecewise-affine mapping f_1. We define the approximation f_1 as the function φ on the grid $\mathcal{R}$. We may start with the bad 'squares'. Note that Theorem 2.1 can be applied not only to a square but also to polygon $\mathcal{Q}$ which is a 2-bi-Lipschitz piecewise-affine image of a square as can be easily seen by bi-Lipschitz change of variables. If we are given such a $\mathcal{Q}(r)$ and a φ defined on $\partial\mathcal{Q}$, then we use the piecewise-affine 2-bi-Lipschitz change of variables to define a φ_1 on S_r. We then apply Theorem 2.1 to φ_1 and reverse the 2-bi-Lipschitz change of variables gives us a finitely piecewise-affine homeomorphism h on $\mathcal{Q}$ satisfying

$$\fint_{\mathcal{Q}} \Phi(|Dh(Y)|)\, dY \leq C\fint_{\partial\mathcal{Q}} \Phi(|D_\tau\varphi(t)|)\, d\mathcal{H}^1(t).$$

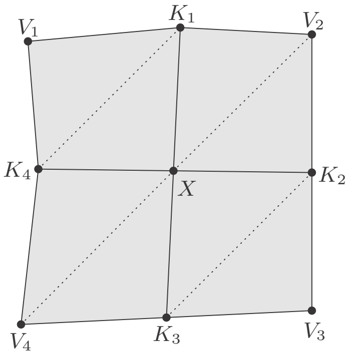


FIGURE 20. How we divide $\mathcal{Q}$ into quadrilaterals, which are each the 2-bi-Lipschitz piecewise-affine images of a pair of triangles.

So given that φ satisfies (4.17) on a $\mathcal{Q}$ where we also have (4.9) we get

$$\begin{aligned}\int_{\mathcal{Q}} &\Phi\big(|Df(Y) - Dh(Y)|\big)\, dY\\ &\leq C\int_{\mathcal{Q}} \Phi\big(|Df(Y)|\big)\, dY + C \operatorname{diam} Q \int_{\partial\mathcal{Q}} \Phi\big(|D_\tau\varphi(t)|\big)\, d\mathcal{H}^1(t) \\ &\leq C\int_{2\mathcal{Q}} \Phi\big(|Df(Y)|\big)\, dY\end{aligned} \tag{4.22}$$

and C depends only on Φ.

On null polygons $\mathcal{Q} = \mathcal{Q}(c) \in \mathcal{N}$ we would like to use Theorem 3.7. Firstly let us assume that $\mathcal{Q}$ is a quadrilateral and we will explain the difference afterwards. Without loss of generality (up to an isometric rotation in the image or pre-image, which cannot effect the quality of our estimates) we can assume that $Df(c) = \begin{pmatrix} d, 0\\ 0, 0\end{pmatrix}$ for some $M^{-1} < d < M$. By (4.9) and (4.17) we know that

$$\int_{\partial\mathcal{Q}} \Phi(|D_\tau\varphi(t) - Df(c)\tau|)\, d\mathcal{H}^1(t) < C\delta\sigma \operatorname{diam} \mathcal{Q}. \tag{4.23}$$

But applying Theorem 3.7 we get g, a finitely piecewise-affine homeomorphism on $\mathcal{Q}$ satisfying

$$\int_{Q} \Phi\Big(\big|Dg(x) - Df(c)\big|\Big)\, dx \leq C\delta|\mathcal{Q}|. \tag{4.24}$$

Now assume that $\mathcal{Q}$ is not a quadrangle itself, notice that we can divide it into quarters, which will be quadrangles, which are 2-bi-Lipschitz affine images of a pair of triangles as described in the theorem. The situation is depicted in Figure 20. Firstly applying the necessary rotations so that the situation corresponds to that in Section 3. Call the vertices of $\mathcal{Q}$ which were added near the middle of a side of the original square Q, K_1, K_2, K_3, K_4 ordered clockwise around ∂Q so that K_1 and K_3 oppose each other. We want to define φ on $[K_1K_3]$ and $[K_2K_4]$. Firstly we need to define φ at the point X where these segments intersect.

The mapping φ is close to an affine map $\mathcal{A}(Y) = A + de_1Y_1$. Find a point $\boldsymbol{X} = (\boldsymbol{X}_1, \boldsymbol{X}_2)$ in bounded component of $\mathbb{R}^2 \setminus \varphi(\partial\mathcal{Q})$ (the set we called the snake $\mathcal{S}^\circ$) such that $\boldsymbol{X}_1 = \pi_1(\mathcal{A}(X))$. We now define $\varphi(X) = \boldsymbol{X}$. Now we connect each vertex K_i with X using curves described in Lemma 3.5. But making sure that each pair of curves is disjoint apart from the point $\boldsymbol{X}$. Define φ as the constant-speed parametrization of the relevant curve. The estimate (4.23) is still satisfied perhaps with a slightly larger constant.

On the remaining $\mathcal{Q} = \mathcal{Q}(c) \in \mathcal{G}$ we extend our φ as follows. If $\mathcal{Q}$ is a quadrangle we have a linear mapping on each side of $\mathcal{Q}$. So if we divide $\mathcal{Q}$ diagonally by connecting two of its vertices then the values of φ determine a unique affine map on each of the triangles. Note that if $\mathcal{Q}$ is not a quadrilateral we can divide it as above into quarters and it suffices to find a point $\tilde{X} \in \frac{1}{4}\mathcal{Q}(c)$ such that (4.9) holds on $[K_i\tilde{X}]$. We define $\varphi(\tilde{X}) = f(\tilde{X})$ and φ as linear on $[K_i\tilde{X}]$ and we can repeat the calculations from (4.18).

Given a quadrilateral $\mathcal{Q}$ we take two adjacent sides S_1 and S_2 and here define a mapping $k_{\mathcal{Q}}$ on $\mathrm{co}(S_1 \cup S_2)$ as the affine mapping that coincides with φ on $S_1 \cup S_2$. There are two such triangles in $\mathcal{Q}$. We get $k_{\mathcal{Q}} : \mathcal{Q} \to \mathbb{R}^2$ such that $k_Q = \varphi$ on ∂Q. By (4.18) and Δ_2 we know that

$$\Phi\big(|Dk_{\mathcal{Q}} - Df(c))|\big) \leq C\Phi(\delta),$$

where, since $\mathcal{Q}$ is nearly a square in the sense that adjacent sides of $\mathcal{Q}$ are not nearly parallel, we can control the size of $|Dk_{\mathcal{Q}} - Df(c)|$ in $\mathcal{Q}$ by $C|D_\tau k_{\mathcal{Q}} - Df(c)\tau|$ on $\partial\mathcal{Q}$ for some fixed geometrical constant C. It follows that

$$\int_{\mathcal{Q}} \Phi\big(|Dk_{\mathcal{Q}} - Df(c))|\big) \leq C\Phi(\delta)|Q|.$$

Combining this with (4.4) we get

$$\begin{aligned}\int_{\mathcal{Q}} \Phi\big(|Df(Y) - Dk_Q(Y)|\big)\, dY \leq C \int_{\mathcal{Q}} \Phi\big(|Df(c) - Df(Y)|\big)\, dY +\\ + C\int_{\mathcal{Q}} \Phi\big(|Df(c) - Dk_Q(Y)|\big)\, dY \leq C\delta|\mathcal{Q}|.\end{aligned} \tag{4.25}$$

Step 5: Estimates on the derivative and functional values.
Altogether we obtain a mapping such that

$$f_1(x) = \begin{cases} k_Q(x) & \text{for } x \in Q,\ Q \in \mathcal{G}, \\ h_Q(x) & \text{for } x \in Q,\ Q \in \mathcal{B}, \\ g_Q(x) & \text{for } x \in Q,\ Q \in \mathcal{N}. \end{cases}$$

It is easy to check that f_1 is a piecewise-affine homeomorphism. On the bad squares we can estimate our error as follows. We use the finite-overlap quality of our grid given in (4.6), the estimate (4.1), (4.16) and (4.22) to show that

$$\int_{\bigcup_{\mathcal{B}} \mathcal{Q}} \Phi\big(|Df_1|\big) \leq C\iota \sum_{\mathcal{B}} \int_{2\mathcal{Q}} \Phi\big(|Df|\big) < C\iota\varepsilon.$$

Therefore also recalling that C_2 is the Δ_2-constant of Φ we get

$$\int_{\bigcup_{\mathcal{B}} \mathcal{Q}} \Phi\big(|Df_1 - Df|\big) < C_2 \int_{\bigcup_{\mathcal{B}} \mathcal{Q}} \Phi\big(|Df_1|\big) + C_2 \int_{\bigcup_{\mathcal{B}} \mathcal{Q}} \Phi\big(|Df_1|\big) < C_2(C\iota + 1)\varepsilon.$$

For other squares we can simply make use of (4.24) and (4.25) to estimate the distance of our approximation from the original function. We get

$$\begin{aligned}\int_\Omega \Phi\big(|Df(x) - Df_1(x)|\big)\,dx \leq &\sum_{\mathcal{Q}\in\mathcal{B}} \int_{\mathcal{Q}} \Phi\big(|Df(x) - Dh_Q(x)|\big)\,dx\\ &+\sum_{\mathcal{Q}\in\mathcal{N}} \int_{\mathcal{Q}} \Phi\big(|Df(x) - Dg_Q(c)|\big)\,dx\\ &+\sum_{\mathcal{Q}\in\mathcal{G}} \int_{\mathcal{Q}} \Phi\big(|Df(x) - Dk_Q(x)|\big)\,dx\\ \leq\, &C(1+|\Omega|)\varepsilon.\end{aligned}$$

Summarising the above we write

$$\int_\Omega \Phi\Big(|Df(x) - Df_1(x)|\Big)\,dx \leq C|\Omega|\varepsilon. \tag{4.26}$$

Now we will estimate the distance of the functional values. On good squares $\mathcal{Q} \in \mathcal{G}$, it immediately follows from (4.4) and (4.19) that

$$|f_1(x) - f(x)| \leq C\delta \operatorname{diam} \mathcal{Q}.$$

Considering the proof of Theorem 3.7 we know for $\mathcal{Q} \in \mathcal{N}$ that our approximation f_1 is close (error of less than $C\delta\sigma \operatorname{diam} \mathcal{Q}$) to the affine mapping given by $f(x_0) + Df(x_0)(x - x_0)$, which is close to f on $\mathcal{Q}$.

The last case we have to consider is $\mathcal{Q} \in \mathcal{B}$. We know that at some point $Y \in \partial\mathcal{Q}$ we have $f_1(Y) = f(Y)$. Both f_1 and f are monotone and so we can estimate their oscillation using their oscillation on $\partial\mathcal{Q}$. Also $f_1(\mathcal{Q}) \subset \operatorname{co} f(\mathcal{Q})$ and therefore $\operatorname{diam} f_1(\mathcal{Q}) \leq \operatorname{diam} f(\mathcal{Q})$. We have

$$\operatorname{diam} f_1(\mathcal{Q}) \leq \operatorname{diam} f(\partial\mathcal{Q}) = \int_{\partial\mathcal{Q}} |Df| < C \operatorname{diam} \mathcal{Q} \fint_{2\mathcal{Q}} |Df|.$$

It is not immediate, however, that this tends to zero as $\operatorname{diam} \mathcal{Q}$ tends to zero. Therefore we will make the following alteration. If we have $\mathcal{Q} \in \mathcal{B}$ then we subdivide it with a grid of quadrangles which are all close to squares of radius $\operatorname{diam} \mathcal{Q}/2l$ for some $l \in \mathbb{N}$. This will suffice to ensure that the error on $\mathcal{Q}$ is as small as we want. The mapping f is uniformly continuous on $\mathcal{Q}$ and the Lipschitz constant of f_1 on $\mathcal{Q}$ is bounded by the Lipschitz constant of the approximation of f on the grid. So we see that for all grids (all values of l) that f_1 and f are equally uniformly continuous. Therefore by choosing l large enough we can ensure that the error on $\mathcal{Q}$ is as small as we want. We must however take care when making this subdivision that we maintain our estimates on the derivative as well.

This is very similar to Step 2 so our argument will be considerably briefer. We may assume that $\mathcal{Q}$ is infact a square as otherwise we may divide it into quarters and apply bi-Lipschtz piecewise-affine mappings to achieve four squares. We divides the square $\mathcal{Q}$ evenly into l^2 squares. Then we repeat the shifting argument from Step 2 with the difference that if a vertex lies on a side of $\mathcal{Q}$ then we position it somewhere on a a segment in $\partial\mathcal{Q}$ not on a segment parallel with $(1,1)$, see Figure 21. It is now clear that on a square $\tilde{Q}$ compactly contained in $\mathcal{Q}$ we have

$$\fint_{\partial\tilde{Q}} \Phi(|D_\tau f_1|) \leq C \fint_{2\tilde{Q}} \Phi(|Df|).$$

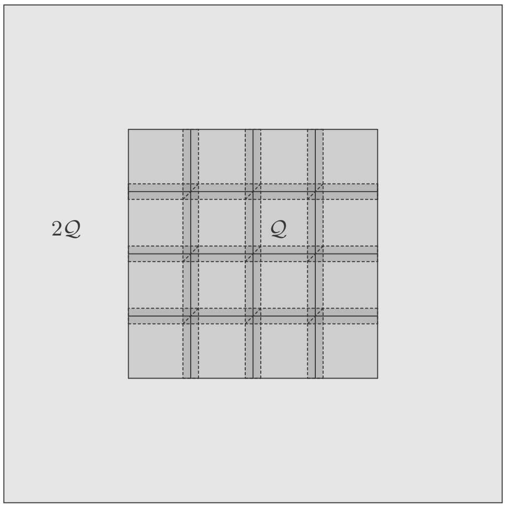


FIGURE 21. Subdividing a bad square so that the integral over the grid inside the bad square can be bounded using the integral over the shaded area around the grid.

Now we repeat our approximation-on-the-grid argument from Step 3 and use Theorem 2.1 to redefine f_1 in the squares of the subdivision of $\mathcal{Q}$. This process may require us to redefine f_1 on $\partial\mathcal{Q}$ with a finer approximation and smaller r_V, this is not a problem however. Summing over the squares $\tilde{Q}$ compactly contained in $\mathcal{Q}$ gives

$$\sum_{\tilde{Q}\Subset\mathcal{Q}} \int_{\tilde{Q}} \Phi(|Df_1|) \leq \sum_{\tilde{Q}\Subset\mathcal{Q}} C\iota \operatorname{diam}\tilde{Q} \int_{\partial\tilde{Q}} \Phi(|D_\tau f_1|) \leq C\iota \int_{\mathcal{Q}} \Phi(Df).$$

It remains to consider those squares around the edge of $\mathcal{Q}$. Here the situation is not much more complicated, we have

$$\begin{aligned}\sum_{\tilde{Q}\cap\partial\mathcal{Q}\neq\emptyset} \int_{\tilde{Q}} \Phi(|Df_1|) &< C\iota\frac{\operatorname{diam}\tilde{Q}}{l} \int_{\partial\mathcal{Q}} \Phi(|D_\tau f_1|) + C\iota\frac{\operatorname{diam}\tilde{Q}}{l} \int_{\bigcup(\partial\tilde{Q}\setminus\partial\mathcal{Q})} \Phi(|D_\tau f_1|)\\ &< C\iota \int_{2\mathcal{Q}} \Phi(|Df_1|) + C\iota \int_{\mathcal{Q}} \Phi(|Df_1|).\end{aligned}$$

In conclusion we have increased the modular of the derivative on $\mathcal{Q}$ by at most some constant factor C and we have achieved $|f(X) - f_1(X)| < \varepsilon$ for $X \in \mathcal{Q} \in \mathcal{B}$ for any given $\varepsilon > 0$ and C is geometric, independant of $\mathcal{Q}$ and ε.

Let us take the sequence ε_k given in the theorem. The constant C_2 is the Δ_2 constant of Φ. We find p_k such that $2^{-p_k} < \varepsilon_k$, then we find approximations f_k as above such that

$$\int_\Omega \Phi(|f - f_k|) + \Phi(|Df - Df_k|) \leq C_2^{-p_k} \text{ and } \|f_k - f\|_{L^\infty(\Omega)} < \varepsilon_k.$$

By the Δ_2 condition we have that

$$\int_\Omega \Phi\Big(\frac{|f-f_k|}{2^{-p_k}}\Big)+\Phi\Big(\frac{|Df-Df_k|}{2^{-p_k}}\Big)\leq C_2^{p_k}C_2^{-p_k}=1.$$

Therefore $\|f-f_k\|_{W^{1,\Phi}(\Omega,\mathbb{R}^2)}\leq \varepsilon_k$.

Step 6: Behaviour on the boundary.

Let us assume that f can be continuously extended onto $\partial\Omega$. Then we have an extension which we will also call f which is continuous on $\overline{\Omega}$ and is therefore uniformly continuous on Ω. Since f and f_k are both uniformly continuous and in every $\mathcal{Q}\in\mathcal{C}$ there are points X where $f_k(X)=f(X)$, we know that for $Y\in\mathcal{Q}$ the error $|f_k(Y)-f(Y)|$ must be less than 2ε given that $\operatorname{diam}\mathcal{Q}<\delta$ for suitable ε and δ. We know that the diameters of our squares tend uniformly to zero with respect to their distance from $\partial\Omega$. This and the fact that $f_k(\mathcal{Q})\subset f(\Omega)$ easily imply that we can extend each f_k with the same boundary values as f on $\partial\Omega$. Further, in this case it is clear that $f_k(\Omega)=f(\Omega)$. □

4.2. **Finite triangulation if f is piecewise-linear on a polygon.** Now we will use Theorem 4.1 on a compactly embedded subdomain in Ω. Then we improve the boundary behaviour given that $\partial\Omega$ is a polygon and f is piecewise-linear on $\partial\Omega$. The strategy of the proof below is to separate a very thin tube around the boundary into bi-Lipschitz image of squares. In order to have a uniform bound on the bi-Lipschitz constant we must eliminate any sharp angles in the boundary. Then we can define f_1 appropriately on the boundary of bi-Lipschitz images of squares which lie around $\partial\Omega$ and use Theorem 2.1 to extend these values which give us a homeomorphism on the tube around the boundary. Then, because we are working entirely in a tube of tiny measure, we will find that the integral of $\Phi(|Df|)$ and $\Phi(|Df_1|)$ around the boundary will be very small. As each section in the boundary will be the bi-Lipschitz image of a square of fixed size we will cover the tube around the boundary with a finite number of such sets. A finite triangulation around $\partial\Omega$ means that the total triangulation is finite, because the triangulation of a compactly embedded subdomain in Ω is finite.

Theorem 4.2. *Let Φ be a Δ_2-Young function and let ε_k be a decreasing sequence of positive numbers tending to zero. Let the domain $\Omega\subset\mathbb{R}^2$ be such that $\partial\Omega$ is a polygon in the sense of Lemma 3.1. Let $f\in W^{1,\Phi}(\Omega,\mathbb{R}^2)$ be a homeomorphism from $\overline{\Omega}$ into $\mathbb{R}^2$ and f is piecewise-linear on $\partial\Omega$. Then there exist globally finite piecewise-affine homeomorphisms f_k such that*

$$\|f-f_k\|_{W^{1,\Phi}(\Omega,\mathbb{R}^2)}<\varepsilon_k,\ \ \|f-f_k\|_{L^\infty(\Omega,\mathbb{R}^2)}<\varepsilon_k$$

and $f_k=f$ everywhere on $\partial\Omega$.

Proof. *Step 0: Uniform convergence* The mapping f is uniformly continuous. Thanks to Theorem 4.1 we can aproximate f in L^∞ as acurately as we like in any $G\Subset\Omega$. Also, calling $\lambda_X=\inf\{|X-Y|;Y\in\partial\Omega\}$ if $\sup\{\lambda_X,X\in G\}<\delta$ then we can easily arrange that $|f-f_k|<\varepsilon$. So we will have no problem in ensuring that $|f-f_k|<\varepsilon_k$ by choosing in our following construction α small enough.

Step 1: Eliminating sharp angles in $\partial\Omega$.

It is necessary we recall the simple Lemma 3.1, where we constructed polygons close to an original. We construct the polygons P_α from Lemma 3.1, where $P=\partial\Omega$

and call Ω_α the bounded component of $\mathbb{R}^2 \setminus P_\alpha$. Let $\eta > 0$. It is a simple observation that

$$\mathcal{L}^2(\overline{\Omega} \setminus \Omega_\alpha) \leq C\mathcal{H}^1(\partial\Omega)\alpha.$$

So we can easily find an $\alpha' > 0$ such that

$$\mathcal{L}^2(\overline{\Omega} \setminus \Omega_{\alpha'}) \leq \eta$$

for any given η.

We may assume that $\partial\Omega$ does not contain any acute angles. By this we mean that if we have three subsequent vertices of the polygon D_i, D_{i+1}, D_{i+2} then $\frac{\pi}{2} \leq \measuredangle D_i D_{i+1} D_{i+2} \leq \frac{3\pi}{2}$ (for the definition of $\measuredangle D_i D_{i+1} D_{i+2}$ see the preliminaries). The reason we can make this assumption is that if it does not hold we can use a piecewise-affine bi-Lipschitz change of variables Ψ to change the polygon Ω so that it holds. Also we may do this in such a way that for any compactly embedded $G \Subset \Omega$ we may assume that Ψ is equal to identity in G and the bi-Lipschitz constant of Ψ depends on Ω but not G. Thanks to this, the bi-Lipschitz change of variables will not cause problems because thanks to the Δ_2 quality of Φ we can calculate that

$$\lim_{\mathcal{L}^2(\Omega\setminus G)\to 0} \int_{\Omega\setminus G} \Phi(\|D\Psi\|_\infty |Df|) \leq \lim_{\mathcal{L}^2(\Omega\setminus G)\to 0} C\int_{\Omega\setminus G} \Phi(|Df|) = 0$$

and therefore we can find G large enough that the above integral is as small as we like. Also note that we know that f is L-Lipschitz on $\partial\Omega$ for some L.

Step 2: Choosing a finite triangulation of f in Ω_α.

Set $P = \partial\Omega$ and find $\alpha_0(\partial\Omega)$ from (3.2). We find α such that

$$\alpha \leq \alpha_0(\partial\Omega),\ \alpha \leq \alpha',\ \text{and } \mathcal{L}^2(\Omega \setminus \tilde{\Omega}_\alpha) < \eta.$$

We know that the angles in $\partial\Omega$ are between $\frac{\pi}{2}$ and $\frac{3\pi}{2}$. This means, by the construction in Lemma 3.1, that if we take any point in $\partial\tilde{\Omega}_\alpha$ its distance from $\partial\tilde{\Omega}$ is in the interval $[\frac{\alpha}{\sqrt{2}}, \alpha]$. Now we can use Steps 1-5 of the proof of Theorem 4.1 to get a locally finite triangulation and approximation of f in Ω. In the process we construct a set of squares $\mathcal{C}_0$ and the shifted 'squares' $\mathcal{C}$. Call $\mathcal{C}_0^*$ the set of all $Q \in \mathcal{C}_0$ such that $Q \cap \tilde{\Omega}_\alpha \neq \emptyset$ or a corresponding $\mathcal{Q} \in \mathcal{C}$ has $\mathcal{Q} \cap \Omega_\alpha \neq \emptyset$. The set $\mathcal{C}^*$ is the set of distorted squares $\mathcal{Q}$ corresponding to $Q \in \mathcal{C}_0^*$ (see also Figure 22). We call $U = \bigcup_{\mathcal{Q}\in\mathcal{C}^*} \mathcal{Q}$, and $U_0 = \bigcup_{Q\in\mathcal{C}_0^*} Q$. Notice that $U, U_0 \supset \Omega_\alpha$. We will use f_k given by Theorem 4.1 on U. Note that as $U \subset\subset \Omega$, our triangulation is finite.

Now we want to construct a fine grid of squares around the boundary of U. By construction the side lengths of adjacent squares $Q \in \mathcal{C}_0^*$ vary by a factor of at most 2, therefore all of these squares are contained in $\Omega_{\frac{\alpha}{4}}$. The side length of $Q \in \mathcal{C}_0^*$ is $2^{-j_Q} h_0$ for some $j_Q \in \mathbb{N}$. We choose k such that $k > j_Q$ for all $Q \in \mathcal{C}_0^*$ and the number $r = 2^{-k} h_0 \leq \frac{\alpha}{32}$. Now we define the maximal set of essentially disjoint squares Q with side length r in $\Omega_{\frac{\alpha}{8}} \setminus U_0$ and call it $\mathcal{C}_0^\partial$. This is depicted in Figure 22.

By shifting the vertices of the $Q \in \mathcal{C}_0^\partial$ we find the set $\mathcal{C}^\partial$ of $\mathcal{Q}$ satisfying

$$\fint_{\partial\mathcal{Q}} \Phi(|D_\tau f|) \leq C \fint_{2\mathcal{Q}} \Phi(|Df|). \tag{4.27}$$

We define $f_1 = f$ on $\partial\Omega$. Now we define f_1 on $\partial\mathcal{Q}$, for $\mathcal{Q} \in \mathcal{C}^\partial$, and on ∂Q_i, for $i \in I$, as we did in Step 4 of Theorem 4.1 and depicted in Figure 19 making sure that our divisions of each side of $\mathcal{Q}$ are so fine that their images are mutually disjoint.

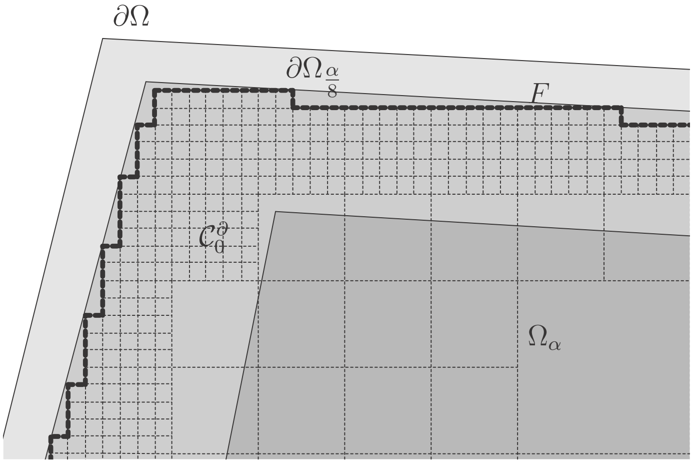


FIGURE 22. The preliminary squares $\mathcal{C}_0^*$ and $\mathcal{C}_0^\partial$ from which we create the fine, uniform grid around the edge of Ω.

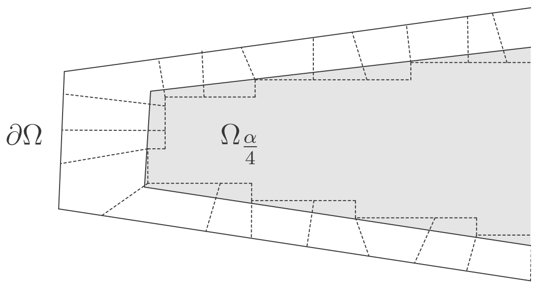


FIGURE 23. Choosing pairs of vertices around the edge of $\tilde{\Omega}$ to enclose D_k, where we can apply Theorem 2.1.

Now we can simply apply Theorem 2.1 on each $\mathcal{Q}$ in $\mathcal{C}^\partial$ using our standard change of variables argument. We get

$$\sum_{\mathcal{Q}\in\mathcal{C}^\partial}\int_{\mathcal{Q}}\Phi(|Df_1|) \leq Cr\sum_{\mathcal{Q}\in\mathcal{C}^\partial}\int_{\partial\mathcal{Q}}\Phi(|D_\tau f_1|) \leq C\sum_{\mathcal{Q}\in\mathcal{C}^\partial}\int_{2\mathcal{Q}}\Phi(|Df|) \tag{4.28}$$

<u>Step 3:</u> *Defining f_1 on $\tilde{\Omega}\setminus(\bigcup_{\mathcal{C}^\partial}\mathcal{Q}\cup U)$.*

We have defined f_1 on $G=\bigcup_{\mathcal{C}^\partial}\mathcal{Q}\cup U$ and we are left with some very small area around the boundary $\Omega\setminus G$. Let $\mathcal{F}=\partial G$ and $\mathcal{F}$ corresponds to $f(F)$ for the F depicted in Figure 22 after we have moved the vertices of the squares in $\mathcal{C}_0^\partial$. We want to separate the remaining part of Ω, where we have not yet defined f_1 into parts which are all K-bi-Lipschitz images of a square with K universally fixed, which is depicted in Figure 23. We will separate the set using pairs of points $X_k \in \mathcal{F}, Y_k \in \partial\Omega$. In this regard consider a finite set $\{X_1, \dots, X_{k_0}\} \subset \mathcal{F}$ such that $\mathcal{F} \setminus \{X_1, \dots, X_{k_0}\}$ is the disjoint union of the curves $\gamma_1, \gamma_2 \dots \gamma_{k_0}$ with the endpoints of γ_k being X_k and X_{k+1} (where $(X_{k_0+1} = X_1)$. Similarly $\partial\Omega = \{Y_1, Y_2, \dots, Y_{k_0}\} \cup \bigcup_k \phi_k$. We will have $[X_k Y_k] \setminus \{X_k, Y_k\} \subset \Omega \setminus G$ for all k and we will refer to the bounded component of $\mathbb{R}^2 \setminus ([X_k Y_k] \cup [X_{k+1} Y_{k+1}] \cup \gamma_k \cup \phi_k)$ as D_k.

We want to choose pairs of points $X_k \in \mathcal{F}$, and $Y_k \in \partial\Omega$ such that $[X_k Y_k] \setminus \{X_k, Y_k\} \subset \Omega \setminus G$ and so that we have both the inequalities,

$$\begin{aligned} \fint_{[X_k Y_k]} \Phi(|D_\tau f|) d\mathcal{H}^1 &\leq C \fint_{D_k} \Phi(|Df|) \\ \fint_{[X_{k+1} Y_{k+1}]} \Phi(|D_\tau f|) d\mathcal{H}^1 &\leq C \fint_{D_k} \Phi(|Df|). \end{aligned} \tag{4.29}$$

Further we will require

$$\frac{\alpha}{32} \leq \mathcal{H}^1(\gamma_k) \leq \frac{\alpha}{4} \quad \text{ and } \quad \frac{\alpha}{32} \leq \mathcal{H}^1(\phi_k) \leq \frac{\alpha}{4}. \tag{4.30}$$

There are clearly many sets $\{X_1, \dots, X_{k_0}\}$ and $\{Y_1, \dots, Y_{k_0}\}$ from which we can choose (e.g. choose Y_k as any point in $\partial\tilde{\Omega}$ and connect it using a segment whose angle to the to the 'normal' vector at Y_k is less than $\frac{1}{8}\pi$ with a point X_k in $\mathcal{F}$). Therefore we will not have any problems finding pairs that satisfy (4.29) and (4.30) will not be a problem.

Now we notice that for some large but fixed K we can pick X_k and Y_k such that D_k is the K-bi-Lipschitz image of a square of length $\frac{\alpha}{8}$. To this end we make the following geometric considerations. Since $\frac{1}{2}\alpha < \alpha_0(\partial\tilde{\Omega})$ we know that ϕ_k contains at most one vertex of $\partial\tilde{\Omega}$. Also we know that the angles in $\partial\Omega$ are between $\frac{1}{2}\pi$ and $\frac{3}{2}\pi$. Each $\mathcal{Q} \in \mathcal{C}^\partial$ is the 2-bi-Lipschitz piecewise-affine image of a square of length $\frac{1}{32}\alpha$ which means that we can locally straighten out $\mathcal{F}$ with a piecewise-affine bi-Lipschitz mapping. The distance from any point in $\mathcal{F}$ to $\tilde{\Omega}$ is approximately the same everywhere in $\mathcal{F}$, i.e. it lies in the interval $[\frac{\alpha}{8}, \frac{3\alpha}{16})$, which is approximately the same as $H^1(\gamma_k)$, resp. $\mathcal{H}^1(\phi_k)$. Thanks to these geometric considerations, we know that for some large but universally fixed K we can choose X_k and Y_k such that all D_k are the K-bi-Lipschitz image of a square of length $\frac{\alpha}{8}$. All we need to do is to choose our X_k on vertices of $\mathcal{Q} \in \mathcal{C}^\partial$ or near the middle of sides of $\mathcal{Q}$. Now on each D_k we apply the K-bi-Lipschitz change of variables, Theorem 2.1 and reverse the change of variables to define f_1 on D_k.

Now take any D_k and apply Theorem 2.1 to get

$$\int_{D_k} \Phi(|Df_1|) \leq C\alpha \int_{\partial D_k} \Phi(|D_\tau f_1|).$$

Remember that our choices of vertices in $\mathcal{C}^\partial$ we get (4.27), which bounds $\int_{\gamma_k} |D_\tau f|$, where γ_k is the part of F which is between X_k and X_{k+1}. Combine this with (4.29) and $|D_\tau f| = |D_\tau f_1| \leq L$ to get that

$$\begin{aligned}\int_{D_k} \Phi(|Df_1|) &\leq C\alpha \int_{\partial D_k} \Phi(|D_\tau f_1|) \leq C\alpha \int_{\partial D_k} \Phi(|D_\tau f|)\\ &\leq C \int_{\partial D_k} \big(\Phi(|D_\tau f|) + \Phi(L)\big) + C \sum_{\{\mathcal{Q}\in\mathcal{C}^\partial : \overline{\mathcal{Q}}\cap\overline{D_k}\neq\emptyset\}} \int_{2\mathcal{Q}} \Phi(|Df|).\end{aligned}$$

For each $2\mathcal{Q}$ in our sum it holds that $2\mathcal{Q} \subset U$.

Now it suffices to choose η small enough so that despite the change of variables Ψ we applied in Step 1 of our proof we will have that $\int_A \Phi(|Df \circ \Psi|) \leq \varepsilon$ for any A such that $\mathcal{L}^2(A) < \eta$ and choose α so that $\alpha\|D\Psi\|^2_{L^\infty}\mathcal{H}^1(\partial\Omega) < \eta$. Then we get

$$\int_\Omega \Phi(|Df - Df_1|) = \int_U \Phi(|Df - Df_1|) + \int_{\Omega\setminus U} \Phi(|Df|) < C\iota\varepsilon.$$

If $2^{-p} < \varepsilon_k$ and we have $C\iota\varepsilon < C_2^p$, then this estimate implies that we can find a finitely piecewise-affine f_k such that

$$\|f - f_1\|_{W^{1,\Phi}(\Omega,\mathbb{R}^2)} < \varepsilon_k.$$

$\square$

4.3. Approximation on general Ω, which ensures that the error vanishes on the boundary.

Theorem 4.3. *Let $\Omega \subseteq \mathbb{R}^2$ be a domain and let ε_k be a decreasing sequence of positive numbers tending to zero and let Φ be a Δ_2-Young function. For any homeomorphism $f \in W^{1,\Phi}(\Omega,\mathbb{R}^2)$ there exists a sequence of diffeomorphisms $f_k \in W^{1,\Phi}(\Omega,\mathbb{R}^2)$ such that*

$$\|f - f_k\|_{W^{1,\Phi}(\Omega,\mathbb{R}^2)} \leq \varepsilon_k \text{ and } \|f - f_k\|_{L^\infty(\Omega,\mathbb{R}^2)} < \varepsilon_k.$$

Moreover $f - f_k \in W_0^{1,\Phi}(\Omega,\mathbb{R}^2)$ for all $k \in \mathbb{N}$.

Proof. Initially we will consider only the Luxembourg semi-norm, i.e. we will estimate $\|Df_k - Df\|_{L^\Phi(\Omega,\mathbb{R}^2)}$ and later explain how to modify our proof to simultaneously estimate functional values.

Step 1: Dividing Ω into subdomains.

We will start by choosing a sequence of bounded, embedded subdomains which cover Ω, i.e. we find

$$\emptyset = U_0 \subsetneq U_1 \Subset U_2 \Subset \cdots \Subset U_m \Subset \cdots \Subset \Omega \text{ and } \bigcup_{m=1}^\infty U_m = \Omega.$$

We can do this in such a way that for each $m \in \mathbb{N}$, ∂U_m is a polygon with sides parallel to $(1,0)$ or $(0,1)$ (for example U_m is a partial Whitney covering of Ω). We denote κ_m as the length of the shortest side of ∂U_m.

Step 2: Creating an approximation of f on $U_m \setminus U_{m-1}$ and joining them near the boundary of the sets U_m.

It follows from Theorem 4.1 that on every $U_m \setminus U_{m-1}$ we can find locally finite piecewise-affine homeomorphisms $\tilde{f}_{k,m}$ such that

$$\int_{U_m\setminus U_{m+1}} \Phi\bigg(2^{m-j}\frac{|Df - D\tilde{f}_{k,m}|}{\varepsilon_k}\bigg) < 2^{-j-3} \text{ for all } j = 0, 1, \ldots, m. \tag{4.31}$$

In the process of constructing these approximations we use preliminary grids in $U_m \setminus U_{m-1}$ which we call $\mathcal{C}_0^m$ to create an altered grid in $U_m \setminus U_{m-1}$ which we call $\mathcal{C}^m$ as per Step 2 of the proof of Theorem 4.1. We call the radius of a 'square' in $\mathcal{C}^m$ the radius of the corresponding square in $\mathcal{C}_0^m$. We will use the approximations $\tilde{f}_{k,m}$ to construct a single mapping f_k defined on Ω. In order to do this we need to modify our approximations near where they meet so that f_k is locally finitely piecewise-affine in Ω. To do this we will define tiny 'tubes' around ∂U_m which we will call G_m. Some parts of the construction are very similar to the proof of Theorem 4.2 so we will be slightly briefer about some of the details. For

$$\rho < \rho_m^0 = \min\left\{\frac{\kappa_m}{16}, \frac{\operatorname{dist}(U_m, U_{m+1})}{8}, \frac{\operatorname{dist}(U_m, U_{m-1})}{8}\right\}$$

we define

$$\begin{aligned} G_m(\rho) &= \bigcup_{x\in\partial U_m} \mathcal{B}\big(x,\rho\big),\\ G_m^1(\rho) &= \bigcup_{x\in\partial U_m} \mathcal{B}\Big(x,\frac{\rho}{4}\Big),\\ G_m^2(\rho) &= \bigcup_{x\in\partial U_m} \mathcal{B}\Big(x,\frac{\rho}{8}\Big). \end{aligned}$$

It is a simple observation that $\mathcal{L}^2(G_m) < \infty$ and tends to 0 as $\rho \to 0$. It suffices us to choose any $\rho_m \in (0, \rho_m^0)$ such that

$$\begin{aligned} \int_{U_m\cap G_m(\rho_m)} \Phi\Big(2^{m-j}\frac{|Df|}{\varepsilon_k}\Big) &< \frac{2^{-j-3}}{C_4(m)} \text{ for all } j = 0, 1, 2\dots m \text{ and}\\ \int_{U_{m+1}\cap G_m(\rho_m)} \Phi\Big(2^{m+1-j}\frac{|Df|}{\varepsilon_k}\Big) &< \frac{2^{-j-3}}{C_4(m)} \text{ for all } j = 0, 1, 2\dots m+1, \end{aligned}$$

where $C_4(m)$ is a function dependant only on m, which we will specify in the following. We do this for all $m \in \mathbb{N}$.

Our construction allows us to require that all squares in our preliminary grids $\mathcal{C}_0^m$ and $\mathcal{C}_0^{m+1}$ which intersect $G_m(\rho)$ have diameter less than $\frac{\rho}{64}$. Put

$$\begin{aligned} F_m^+(\rho) &= \{x \in U_m; x \in \mathcal{Q} \text{ for some } \mathcal{Q} \in \mathcal{C}^m \text{ and } \mathcal{Q} \subset G_m^1(\rho)\}\\ F_m^-(\rho) &= \{x \in U_m; x \in \mathcal{Q} \text{ for some } \mathcal{Q} \in \mathcal{C}^m \text{ and } \mathcal{Q} \subset G_{m-1}^1(\rho)\}. \end{aligned}$$

Now we create a finite triangulation of $F_m^+(\rho)$ and $F_m^-(\rho)$. We define the ρ-dependant variable h_m^+ as the radius of the smallest 'square' $\mathcal{Q} \in \mathcal{C}^m$ which intersects $G_m(\rho) \setminus F_m^+(\rho)$. Then we can define

$$G_m^3(\rho) = \bigcup_{x\in\partial U_m} \mathcal{B}(x, 4h_m^+).$$

We know that $4h_m^+ \le \frac{\rho_m}{16}$ so $G_m^3(\rho) \subset G_m^2(\rho)$. We can now alter the preliminary grid in $F_m^+(\rho)$ so that all squares in it which intersect $G_m^2(\rho)$ have radius h_m^+ and the new preliminary grid is still a Whitney type covering satisfying (4.6). The reason we can achieve this is as follows. We start with the original Whitney-type covering of U_m made of squares such that if a square Q intersects $G_m^1(\rho)$, it has diameter less than

$\frac{\rho}{64}$. This means that for any square Q intersecting ∂G^1_m we have

$$\mathrm{dist}(Q, G^2_m(\rho)) \geq \frac{\rho}{8} - \frac{\rho}{64} = \frac{7\rho}{64}.$$

Now consider the sum of the geometric series given by summing the diameters of the squares we will add to the grid. We add a pair of neighbouring squares of a given radius and then a pair of squares, both having half the radius of the first pair and so on. The sum is 4 times the diameter we started with. In this case the initial diameter is $\rho/64$ and because $4\rho/64 < 7\rho/64$ we will be able to achieve arbitrarily small squares without intersecting G^2_m. We add squares where the radius of the squares in every other generation is half that of the squares in the two previous generations until we have achieved the radius h^+_m. By this we alter $\mathcal{C}^m_0$ and get $\mathcal{C}^m_1$. We add to $\mathcal{C}^m_1$ all the squares with radius h^+_m which do not intersect $G^3_m(\rho)$.

It is necessary to alter the boundary of the new squares in $\mathcal{C}^m_1$ (which correspond to 'squares' in $F^+_m \setminus G^3_m(\rho)$) as per Step 2 from Theorem 4.1. We need, however, to conduct our steps somewhat more carefully. In Theorem 4.1 we controlled the modular of the derivative but now we will need to control $m+1$ functions. We want

$$\int_{\partial\mathcal{Q}} \Phi\Big(2^{m-j}\frac{|D_\tau f|}{\varepsilon_k}\Big) < \hat{C}(m)\int_{2\mathcal{Q}} \Phi\Big(2^{m-j}\frac{|Df|}{\varepsilon_k}\Big) \text{ for } j = 0, 1 \dots m.$$

It suffices to consider the equation (4.11) to see that we can find a set $S^{i,j}_V(\lambda) \subset S_V$ with $|S^{i,j}_V(\lambda)| \geq (1-\lambda^{-1})|S_V|$ and for all $X \in S^{i,j}_V(\lambda)$ there exists a set in S_{V_i} which satisfies the following,

$$\Big|\Big\{Y \in S_{V_i} : \operatorname{diam} Q \int_{[XY]} \Phi\Big(2^{m-j}\frac{|D_\tau f|}{\varepsilon_k}\Big) \leq 2C\lambda \int_{2Q} \Phi\Big(2^{m-j}\frac{|Df|}{\varepsilon_k}\Big)\Big\}\Big| \geq \big(1 - \tfrac{1}{\lambda}\big)|S_{V_i}|$$

for all $j = 0, 1 \dots m$. So, if we put $\hat{C}(m) = 10C(m+1)$ we get a set $\mathcal{S}_V \subset S_V$ with $|\mathcal{S}_V| \geq \frac{|S_V|}{5}$ and for all $j = 0, 1, \dots, m$ and $i = 1, 2, 3, 4$ we have

$$\Big|\Big\{Y \in S_{V_i} : \operatorname{diam} Q \int_{[XY]} \Phi\Big(2^{m-j}\frac{|D_\tau f|}{\varepsilon_k}\Big) \leq \hat{C}(m) \int_{2Q} \Phi\Big(2^{m-j}\frac{|Df|}{\varepsilon_k}\Big)\Big\}\Big| \geq \frac{5m+4}{5m+5}|S_{V_i}|.$$

Now we are in a similar situation as we were in Step 2 of Theorem 4.1. Therefore we can we can repeat our shifting argument verbatim for all functions simultaneously. This gives us a new refined grid $\mathcal{C}^m_2$ which we construct from $\mathcal{C}^m_1$. We define a piecewise-linear approximation of f as in Step 3 of Theorem 4.1, making it fine enough to ensure that it is one-to-one on the union of the boundaries of all our squares in $\mathcal{C}^2_m$ (we do this for squares in $F^+_m(\rho)$ because we have already done it for other squares). We repeat this process in $F^-_m(\rho)$ considering the the newly added squares (resp. 'squares') as squares in $\mathcal{C}^m_1$ (resp. $\mathcal{C}^m_2$).

We still have to deal with the remaining portion of U_m in $G_m(\rho)$ and $G_{m-1}(\rho)$, which is

$$E^+_m(\rho) = \big\{x \in U_m \cap G_m \setminus \big(\bigcup_{\mathcal{Q}\in\mathcal{C}^m_2} \mathcal{Q}\big)\big\}$$

$$E^-_m(\rho) = \big\{x \in U_m \cap G_{m-1} \setminus \big(\bigcup_{\mathcal{Q}\in\mathcal{C}^m_2} \mathcal{Q}\big)\big\}.$$

The geometry of $E^+_m(\rho)$ is a tube of roughly constant width and has no sharp angles (all angles in ∂U_m are close to π). This means it is locally bi-Lipschitz equivalent to

$(0,4h_0^*)\times(0,t)$ for some $t\gg h_0^*$ and the bi-Lipschitz constant is bounded universally and is independent of ρ and m. We call this bound on the bi-Lipschitz constant K_1. It is now easy to separate $(0,4h_0^*)\times(0,t)$ into K_2-bi-Lipschitz equivalents of squares with radius h_0^*, where K_2 is a purely geometrical constant. Moreover, as we continue to prove, we can do this in such a way that when we return both the bi-Lipschitz changes of variables we get sets H_m^i (which we refer to as quasi-squares) such that

$$\fint_{\partial H_m^i}\Phi\Big(2^{m-j}\frac{|D_\tau f|}{\varepsilon_k}\Big)<C\hat{C}(m)\fint_{\mathcal{B}(X_m^i,2\operatorname{diam}H_m^i)}\Phi\Big(2^{m-j}\frac{|Df|}{\varepsilon_k}\Big)\text{ for }j=0,1,\dots,m,\tag{4.32}$$

where the points X_m^i are the 'centres' of H_m^i i.e. $X_m^i=\fint_{H_m^i}X\,dX$. As long as the segments we choose to separate $(0,4h_0^*)\times(0,t)$ into square-like objects are not too close or far away from each other (for example in the interval $(2h_m^*,8h_m^*)$) and nearly vertical (angle to the vertical less than $\pi/8$) we can fix a K_2 such that they are K_2-bi-Lipschitz equivalent with squares. Similarly to Step 2 in the proof of Theorem 4.1 and Step 3 of Theorem 4.2, for a suitable value of C the set of admissible segments is large enough to find segments that satisfy (4.32).

Now we can find a fine piecewise-linear approximation of f on ∂H_m^i such that our approximation is one-to-one on the boundary of squares in the grid. We repeat this process in E_m^-. The difference between finding a piecewise-linear approximation on H_m^i and a typical bad square is that it is not possible to require that the piecewise-linear approximation of f on ∂U_m stays entirely inside $f(U_m)$ but we do not need this. It suffices us that the values of the approximation on ∂U_m lie in $f\big(E_m^+(\rho)\cup E_{m+1}^-(\rho)\big)$ and that the approximation is so fine that it is one-to-one on the grid as a whole. We add the above quasi-squares to $\mathcal{C}_2^m$ to get our final grid, which we call $\mathcal{C}_3^m$. Notice that H_m^i is a finite family of sets.

Now we can define an approximation f_k^ρ on Ω. We have $f_k^\rho=\tilde{f}_{k,m}$ on $U_m\setminus(F_m^+(\rho)\cup F_m^-(\rho))$. We define f_k^ρ on $\mathcal{Q}$ for all $\mathcal{Q}\in\mathcal{C}_2^m$ using a K_3-bi-Lipschitz change of variables and Theorem 2.1. For $H_m^i\in\mathcal{C}_3^m\setminus\mathcal{C}_2^m$ we use two bi-Lipschitz changes of variables and then Theorem 2.1 to define f_k^ρ. For all ρ the mapping f_k^ρ is a locally finite piecewise affine homeomorphism on Ω. We take care when approximating f on the grid inside $F_m^+(\rho)$ and $F_m^-(\rho)$ so that f_k is one-to-one on the grid and also

$$|f(X)-f_k(X)|\leq\frac{\operatorname{dist}(f(\partial U_m),f(\partial U_{m+1}))}{4}\tag{4.33}$$

for all $X\in\partial U_m$.

<u>Step 3:</u> *Modular estimates and choice of ρ_m.*

The 'squares' $\mathcal{Q}\subset F_m^+(\rho)\setminus E_m^+(\rho)$ are K_3-bi-Lipschitz equivalent with squares. We use the bi-Lipschitz change of variables Ψ and then Theorem 2.1 and return the change of variables to define f_k on $\mathcal{Q}$. We can content ourselves with the most elementary estimates in our change of varaibles arguments. Find a $p\in\mathbb{N}$ so that $2^p>\max\{K_1K_2,K_3\}$. Using the above we can make the following estimate on the modular of f_k,

$$\begin{aligned}
\fint_{\mathcal{Q}} \Phi\Big(2^{m-j}\frac{|Df_k|}{\varepsilon_k}\Big) &< K_3^2 \fint_{Q(0,\operatorname{diam}\mathcal{Q}/2)} \Phi\Big(2^{m-j}K_3\frac{|D_\tau f_k\circ\Psi|}{\varepsilon_k}\Big)\\
&< CK_3^2 \fint_{\partial Q(0,\operatorname{diam}\mathcal{Q}/2)} \Phi\Big(2^{m-j}K_3\frac{|D_\tau f_k\circ\Psi|}{\varepsilon_k}\Big)\\
&< CK_3^3 \fint_{\partial\mathcal{Q}} \Phi\Big(2^{2p+m-j}\frac{|D_\tau f_k|}{\varepsilon_k}\Big)\\
&< C_2^{2p}2^{3p}C \fint_{\partial\mathcal{Q}} \Phi\Big(2^{m-j}\frac{|D_\tau f|}{\varepsilon_k}\Big)\\
&< C_2^{2p}2^{3p}C\hat{C}(m) \fint_{2\mathcal{Q}} \Phi\Big(2^{m-j}\frac{|Df|}{\varepsilon_k}\Big) \text{ for all } j=0,1\dots m.
\end{aligned}$$

Similarly putting $C_3(m) = C_2^{2p}2^{3p}C\hat{C}(m)$ we can estimate for any quasi-square $H_m^i \subset E_m^+$

$$\fint_{H_m^i} \Phi\Big(2^{m-j}\frac{|Df_k|}{\varepsilon_k}\Big) < C_3(m)\fint_{\mathcal{B}(X_m^i,2\operatorname{diam}H_m^i)} \Phi\Big(2^{m-j}\frac{|Df|}{\varepsilon_k}\Big) \text{ for all } j=0,1\dots m.$$

Up to exchange $\hat{C}(m+1)$ for $\hat{C}(m)$ these same estimates hold in $F_{m+1}^-(\rho)$ and $E_{m+1}^-(\rho)$.

For simplicity call the set $Z_m(\rho) = (F_m^+(\rho)\cup F_{m+1}^-(\rho))\setminus(E_m^+(\rho)\cup E_{m+1}^-(\rho))$. We would like to sum the above integrals over all $\mathcal{Q}\subset Z_m(\rho)$. We can find an ι_2 which satisfies the (4.6)-type condition for $\mathcal{C}_3^m$ (generally we may have $\iota_2 > \iota$ since we include also quasi-squares into our grid). Then we can estimate

$$\begin{aligned}
\sum_{H_m^i\subset E_m^+\cup E_{m+1}^-} \int_{\mathcal{B}(x_m^i,2\operatorname{diam}H_m^i)} \Phi\Big(2^{m-j}\frac{|Df|}{\varepsilon_k}\Big) + \sum_{\mathcal{Q}\subset Z_m(\rho)}\int_{2\mathcal{Q}}\Phi\Big(2^{m-j}\frac{|Df|}{\varepsilon_k}\Big)\\
< \iota_2\int_{G_m(\rho)}\Phi\Big(2^{m-j}\frac{|Df|}{\varepsilon_k}\Big).
\end{aligned}\tag{4.34}$$

And the integral on the right all tends to 0 as ρ tends to 0 for each $j=0,1\dots m$. Put $C_4(m) = \iota_2(1+C_2)C_2^{2p}2^{3p}C\hat{C}(m)$. If we fix m we have that $\hat{C}(m)$ is a constant. Therefore for each m we can find ρ_m such that

$$C_4(m)\int_{G_m(\rho_m)}\Phi\Big(2^{m-j}\frac{|Df|}{\varepsilon_k}\Big) < 2^{-j-3} \text{ for all } j=0,1\dots m, \tag{4.35}$$

recalling also we require $\rho_m < \rho_m^0$. Combining (4.34) and (4.35) we get that

$$\begin{aligned}
&\int_{F_m^+}\Phi\Big(2^{m-j}\frac{|Df|+|Df_k|}{\varepsilon_k}\Big) < 2^{-j-3} \text{ and}\\
&\int_{F_m^-}\Phi\Big(2^{m-j}\frac{|Df|+|Df_k|}{\varepsilon_k}\Big) < 2^{-j-3}.
\end{aligned}\tag{4.36}$$

Step 4: Estimates on functional values and uniform convergence.

We have f defined on ∂U_m, therefore by Theorem 4.1 we can construct $\tilde{f}_{k,m}$ such that $\|f-\tilde{f}_{k,m}\|_{L^\infty(U_m\setminus U_{m-1})}$ are as small as we like. On the tube we can repeat the refining argument from Step 5 of the proof of Theorem 4.1, which may require us to refine the piecewise-linear approximation on the grid but this poses no obstacle. Therefore we will have therefore no problem in ensuring both $\|f_k - f\|_{L^\infty(\Omega)} < \varepsilon_k$ and

$$\int_{U_m\setminus U_{m-1}} \Phi\Big(2^{m-j}\frac{\|f_k - f\|_{L^\infty(U_m\setminus U_{m-1})}}{\varepsilon_k}\Big) < 2^{-j-1} \tag{4.37}$$

for all $m \in \mathbb{N}$ and $j = 0, 1, \dots m$.

Step 5: Show that $f - f_k \in W_0^{1,\Phi}(\Omega, \mathbb{R}^2)$.

We want to show that $\|f - f_k\|_{W_0^{1,\Phi}(\Omega\setminus U_m,\mathbb{R}^2)} < 2^{-m}\varepsilon_k$. Choose $m_0 \in \mathbb{N}$ and $m > m_0$ notice that (4.31) implies that

$$\int_{U_m\setminus(U_{m+1}\cup F_m^+(\rho_m)\cup F_m^-(\rho_{m-1}))} \Phi\Big(2^{m_0}\frac{|Df - D\tilde{f}_{k,m}|}{\varepsilon_k}\Big) < \frac{2^{m_0-m}}{8}.$$

When we combine this with (4.36) we get

$$\int_{U_m\setminus U_{m+1}} \Phi\Big(2^{m_0}\frac{|Df - D\tilde{f}_{k,m}|}{\varepsilon_k}\Big) < 2^{m_0-m-1}.$$

Therefore summing we get

$$\sum_{m=m_0+1}^{\infty}\int_{U_m\setminus U_{m-1}} \Phi\Big(2^{m_0}\frac{|Df - Df_k|}{\varepsilon_k}\Big) < \frac{1}{2}$$

and

$$\sum_{m=m_0+1}^{\infty}\int_{U_m\setminus U_{m-1}} \Phi\Big(2^{m_0}\frac{|Df - Df_k|}{\varepsilon_k}\Big) + \Phi\Big(2^{m_0}\frac{\|f_k - f\|_{L^\infty(U_m\setminus U_{m-1})}}{\varepsilon_k}\Big) < 1$$

which is what we wanted. Therefore if we define $g_{k,m} = (f - f_k)\cdot\chi_{\Omega\setminus U_{m_0}}$ then

$$\|g_{k,m} - (f - f_k)\|_{W^{1,\Phi}(\Omega,\mathbb{R}^2)} < 2^{-m_0}\varepsilon_k.$$

Now we find some $j \in \mathbb{N}$ such that the convolution approximation $(g_{k,m})_j$ is within $2^{-m_0}\varepsilon_k$ of $g_{k,m}$. Also we may assume that $j^{-1} \ll \operatorname{dist}(\partial U_m, \partial\Omega)$ and therefore $(g_{k,m})_j \in \mathcal{C}_0^\infty(\Omega,\mathbb{R}^2)$. We have

$$\|(g_{k,m})_j - (f - f_k)\|_{W^{1,\Phi}(\Omega,\mathbb{R}^2)} < 2^{1-m_0}\varepsilon_k.$$

We repeat the smoothening argument in Step 0 of the proof of Theorem 4.1 for f_k. Now it suffices to send m_0 to ∞ and then k to ∞.

Step 5: Show that $f_k(\Omega) = f(\Omega)$.

We know from Theorem 4.1 that $\tilde{f}_{k,m}(\Omega) = f(\Omega)$. We altered our approximation in the tubes near ∂U_m and therefore it is not clear that $f_k(U_m) \subset f(U_m)$. Nevertheless it is obvious that $f_k(U_m) \subset f(U_{m+1})$ because we only alter the values very slightly. But similarly considering (4.33) we have

$$f(U_m) \subset f(U_m) + \mathcal{B}(0, d) \subset f_k(U_{m+1})$$

where

$$d = \frac{\operatorname{dist}(f(\partial U_m), f(\partial U_{m+1}))}{4}.$$

Together this gives that $f_k(\Omega) = f(\Omega)$.

□

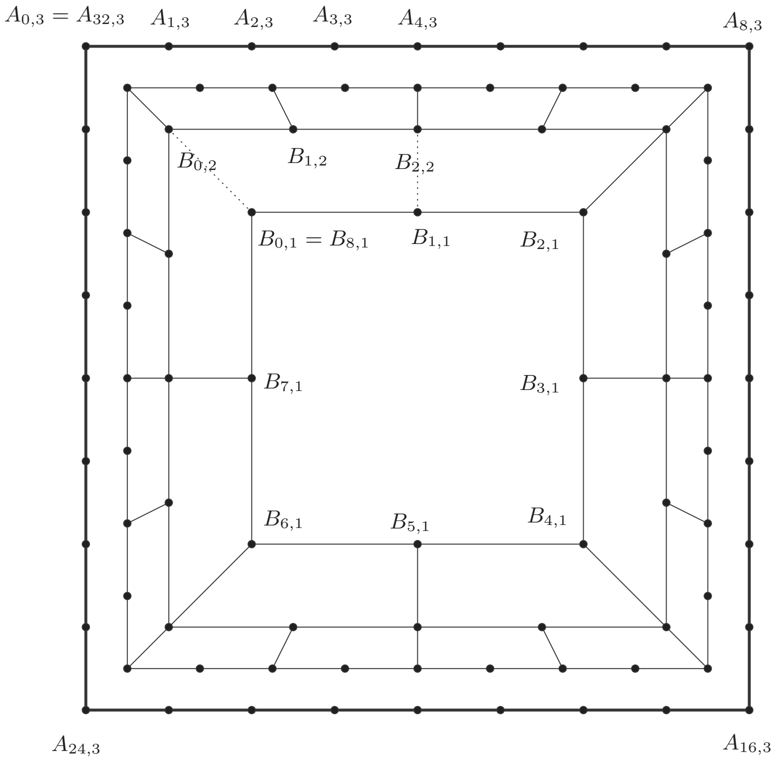


FIGURE 24. The square and the points $A_{i,j}$ and $B_{i,j}$.

Remark 4.4. *The above proof can be simplified if* Φ *satisfies the so called* Δ' *condition (for example* $W^{1,p}$*) in the sense that it suffices to estimate the largest terms, i.e. the terms where* $j = 0$.

4.4. **Proof of Theorem 1.1.** Theorem 1.1 is a direct result of Theorem 4.1, Theorem 4.3 and Theorem 4.2. □

5. PROOF OF THE EXTENSION THEOREM, THEOREM 1.2

We may use the same smoothing argument as in Step 0 of the proof of Theorem 4.1. Therefore it suffices to find a piecewise-affine homeomorphism, with the given properties.

Step 1: Defining some useful mappings and a special set.

We will use a curve $h : [0, 8] \to \mathbb{R}^2$ to parametrize the square. We do this as follows.

$$h = \begin{cases} (t-1, 1) & t \in [0, 2] \\ (1, 3-t) & t \in [2, 4] \\ (5-t, -1) & t \in [4, 6] \\ (-1, t-7) & t \in [6, 8]. \end{cases}$$

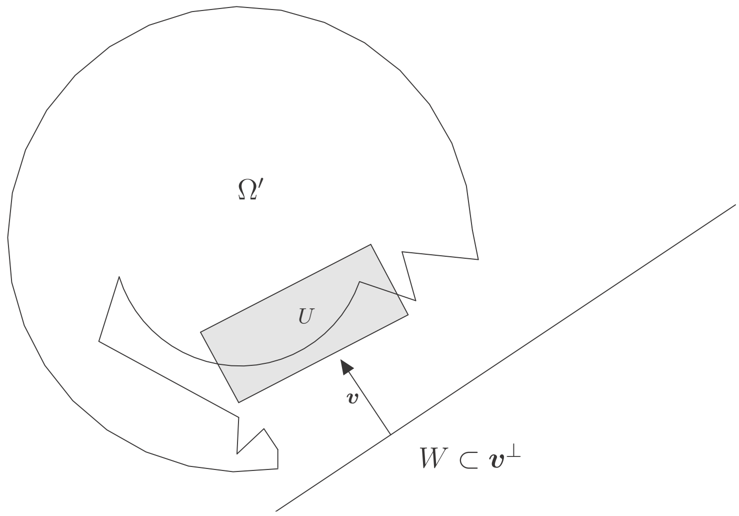


FIGURE 25. A Lipschitz Domain.

See the details of the setup in Figure 24. We further define h on $\mathbb{R}$ as the 8-periodic extension of h. Let us consider the points

$$A_{k,j} = h\left(\frac{8k}{2^{j+2}}\right) \quad k = 0, 1, \ldots, 2^{j+2},\ j \in \mathbb{N}.$$

Then we define

$$\mathcal{A} = \bigcup_{k,j} \varphi\big(A_{k,j}\big).$$

It is obvious but important that $\mathcal{A}$ is countable. Using the following mappings,

$$h_j(t) = (1 - 2^{-j})h(t),$$

we define $B_{k,j}$ as follows,

$$B_{k,j} = h_j\left(\frac{8k}{2^{j+2}}\right) \quad k = 0, 1, \ldots, 2^{j+2},\ j \in \mathbb{N}.$$

Step 2: Choosing n-vectors in the image.

For a given point on the boundary of the image we will need to use vectors directed into the image. Let us recall an equivalent definition of a Lipschitz domain as pictured in Figure 25.

Definition 5.1. *The domain $G \subset \mathbb{R}^n$ is a Lipschitz domain if for every $\boldsymbol{X} \in \partial G$ there exists a neighbourhood U of $\boldsymbol{X}$, a vector $\boldsymbol{v}_{\boldsymbol{X}} \in \mathbb{R}^n$ and $\pi_{\boldsymbol{X}}$ the projection of $\mathbb{R}^n$ onto $\boldsymbol{v}_{\boldsymbol{X}}^{\perp}$ with the following properties:*

a) Let $W_{\boldsymbol{X}} = \pi_{\boldsymbol{X}}(U)$ be the projection of U onto $\boldsymbol{v}_{\boldsymbol{X}}^{\perp}$ then there exists a Lipschitz function $L_{\boldsymbol{X}}$ such that the mapping $f : W_{\boldsymbol{X}} \to \mathbb{R}^n$ defined as

$$f(w) = w + \boldsymbol{v}_{\boldsymbol{X}} L_{\boldsymbol{X}}(w),$$

is a one-to-one mapping onto $\partial G \cap U$.

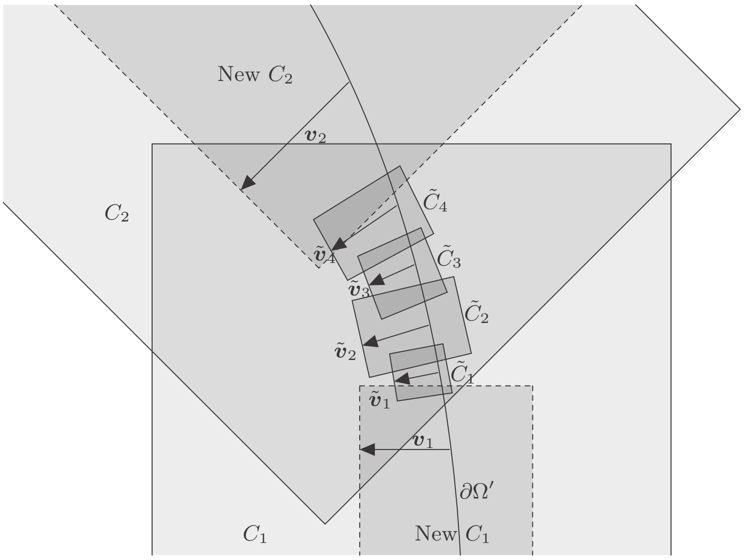


FIGURE 26. Reducing the angle between neighbouring n-vectors to less than a pre-prescribed θ.

b) If $L^*(w) > L_{\boldsymbol{X}}(w) > L_*(w)$ *and* $w + \boldsymbol{v}_{\boldsymbol{X}} L^*(w), w + \boldsymbol{v}_{\boldsymbol{X}} L_*(w) \in U$ *then*

$$w + \boldsymbol{v}_{\boldsymbol{X}} L^*(w) \in G \text{ and } w + \boldsymbol{v}_{\boldsymbol{X}} L_*(w) \in \mathbb{R}^n \setminus \overline{G}.$$

Given a point $\boldsymbol{X} \in \partial G$ *we will refer to the unit vector* $\boldsymbol{v}_{\boldsymbol{X}}$ *as the n-vector of* $\boldsymbol{X}$*.*

Note that the n-vector of a point may not be uniquely determined and will depend on our chosen covering. Part of our construction is choosing specific n-vectors for each point on the boundary of our Lipschitz domain Ω'. Note that for a given open set U we can actually use $\boldsymbol{v}_{\boldsymbol{X}} = \boldsymbol{v}_{\boldsymbol{Y}}$ for all $\boldsymbol{X}, \boldsymbol{Y} \in U \cap \partial\Omega'$. So in fact we have an n-vector $\boldsymbol{v}_U$ for each sufficiently small open set intersecting the boundary.

We claim that, since Ω' is a pre-compact Lipschitz domain, if we are given a small angle θ we can find a finite cover $\mathcal{B}$ of $\partial\Omega'$ with rectangles C_i satisfying *a), b)* from Definition 5.1, with the following properties,

$$\begin{aligned}
&i) \text{ If } C_1, C_2 \in \mathcal{B} \text{ and } \partial\Omega' \cap C_1 \cap C_2 = \emptyset \text{ then } C_1 \cap C_2 = \emptyset.\\
&ii) \text{ If } C_1, C_2, C_3 \in \mathcal{B} \text{ are distinct rectangles then } C_1 \cap C_2 \cap C_3 = \emptyset.\\
&iii) \text{ The set } C_i \setminus \partial\Omega' \text{ has precisely 2 components.}\\
&iv) \text{ If } C_1, C_2 \in \mathcal{B} \text{ and } C_1 \cap C_2 \neq \emptyset \text{ then } \measuredangle \boldsymbol{v}_{C_1} \boldsymbol{v}_{C_2} \leq \theta.
\end{aligned} \tag{5.1}$$

Also we have a universal bound M on the Lipschitz constants of L_C which is independent of θ for $\theta \in (0, \theta_0)$.

Our first step is to take any covering $\{G_i; i \in I\}$ satisfying Definition 5.1. We then cover $\partial\Omega'$ with rectangles inside G_i whose sides are either parallel or perpendicular to the given n-vector $\boldsymbol{v}_{G_i}$. Since $\partial\Omega'$ is compact we will always be able to do this with a finite number of rectangles. Properties *i), iii)* are achieved simply by choosing rectangles small enough. Similarly we easily get *ii)* since we are in dimension 2. We show how we achieve property *iv)* in Figure 26. Let us now expound.

Assume we have $\mathcal{B}$, a covering of $\partial\Omega'$ with rectangles as described above satisfying *i), ii), iii)*. For every $C \in \mathcal{B}$ we have an n-vector $\boldsymbol{v}_C$ and a Lipschitz function L_C. Since we have a finite covering we may define

$$M = \max\{\operatorname{Lip} L_C : C \in \mathcal{B}\} \cup \{1\}$$

to get that all L_C are M-Lipschitz. Further notice that thanks to (5.1) *ii)* we have that for any $\boldsymbol{Y} \in \partial\Omega'$ there are at most two $C_1, C_2 \in \mathcal{B}$ containing $\boldsymbol{Y}$. Property *iii)* guarantees that there exists an interval $(a, b) \subset \mathbb{R}$ such that

$$\varphi\big(h((a, b))\big) = \partial\Omega' \cap C_1 \cap C_2 =: F.$$

Notice that $\beta := \measuredangle \boldsymbol{v}_{C_1}\boldsymbol{v}_{C_2} < \pi$ because L_{C_1} and L_{C_2} are both Lipschitz functions. Calculate

$$n = \left[\frac{\beta}{\theta}\right] + 1.$$

Now choose the points

$$t_i = a + \frac{i(b-a)}{n+1}, \quad i = 1, 2, \dots, n.$$

We find rectangles $\tilde{C}_i \ni \varphi(h(t_i))$ which are contained in $C_1 \cap C_2$ and $\{\tilde{C}_i; i = 1, 2, \dots, n\}$ satisfy conditions (5.1) *i), ii), iii)* with respect to each other. We now reduce the size of C_1 and C_2 so that $\{C_1, C_2, \tilde{C}_i; i = 1, 2 \dots, n\}$ satisfy *i), ii), iii)* with respect to each other. We proceed to choose n-vectors $\tilde{\boldsymbol{v}}_i$ for our new finite covering of F, $\{C_1, C_2, \tilde{C}_i; i = 1, 2 \dots, n\}$. We choose $\tilde{\boldsymbol{v}}_i$ as unitised vectors from the convex hull of the vectors $\boldsymbol{v}_{C_1}$ and $\boldsymbol{v}_{C_2}$ such that

$$\measuredangle \tilde{\boldsymbol{v}}_i \tilde{\boldsymbol{v}}_{i+1} = \frac{\beta}{n+1} \leq \theta$$

and since F can be expressed as the image of an M-Lipschitz mapping with respect to both $\boldsymbol{v}_{C_1}$ and $\boldsymbol{v}_{C_2}$ our choice of $\tilde{\boldsymbol{v}}_i$ guarantees that our L_i are also M-Lipschitz.

So we may assume that we have a finite covering $\mathcal{B} = \{C_1, C_2, \dots, C_m\}$ of $\partial\Omega'$ satisfying Definition 5.1 and (5.1) *i) - iv)*, where all mappings considered are M-Lipschitz and

$$\theta = \arcsin \frac{1}{4M}.$$

Notice that our choice of θ guarantees that

$$M \sin\theta \leq \frac{1}{4}, \ \ \sin(\theta) \leq \frac{1}{4}, \ \ \cos(\theta) > \frac{3}{4}. \tag{5.2}$$

We may assume that the rectangles are ordered around $\partial\Omega'$ as shown in Figure 27. That is, we may assume that

$$C_i \cap C_{i+1} \neq \emptyset \ i = 1, 2, \dots, m,$$

where for simplicity we have called $C_{m+1} = C_1$. Choose $s_1, s_2, \dots, s_m \in [0, 8]$, ($s_{m+1} = s_1 + 8$) such that we have

$$\boldsymbol{Y}_i = \varphi \circ h(s_i) \in (C_i \cap C_{i+1}) \setminus \mathcal{A},$$

recalling that $\mathcal{A}$ is countable and so easy to avoid. We define $H = \{\boldsymbol{Y}_1, \boldsymbol{Y}_2, \dots, \boldsymbol{Y}_m\}$ and sometimes refer to points in H as being critical points because the n-vector

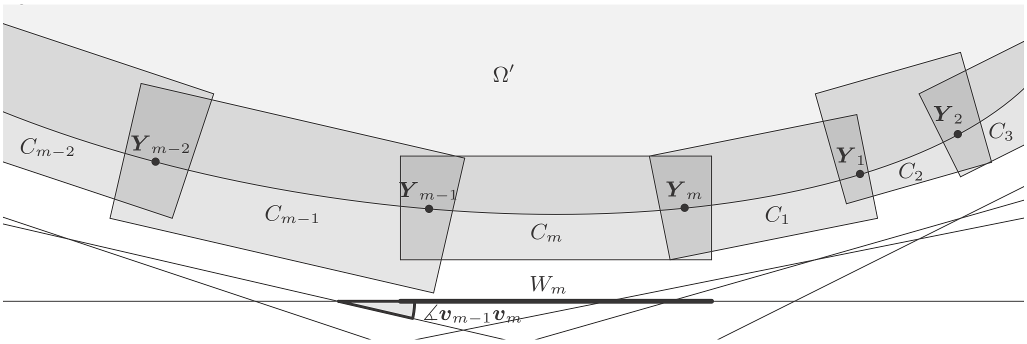


FIGURE 27. An example of a nice boundary covering and the set H.

changes at those points. Now for each $\boldsymbol{X} \in \partial\Omega' \setminus H$ we can choose a single n-vector for $\boldsymbol{X}$ and that is

$$\boldsymbol{v}_{\boldsymbol{X}} = \boldsymbol{v}_i \text{ for } \boldsymbol{X} \in \varphi \circ h\big((s_i, s_{i+1})\big).$$

Notice that in fact our covering $\mathcal{B}$ also covers the set $\{x + t\boldsymbol{v}_x; x \in \partial\Omega', t \in (-\varepsilon, \varepsilon)\}$, for some small $\varepsilon \le 1$.

Step 3: Creating γ_j, paths close to the boundary.

Our main concern now is, given a segment $[B_{k,j}B_{k+1,j}]$, how do we define f on this segment. In order to do this we will need the small numbers α_j. The maximum number of points of H in such a segment of a given j is,

$$K_j = \max_{0 \le i \le 2^{j+2}-1} \mathcal{H}^0\big(H \cap \varphi([A_{k,j}A_{k+1,j}])\big).$$

Make the following definitions

$$\alpha_1 = \frac{\varepsilon}{2K_1} \min_{0 \le k \le 7} \mathcal{H}^1(\varphi([A_{k,1}A_{k+1,1}]))$$

and for $j \ge 2$

$$\alpha_j = \min\Big\{\frac{\varepsilon}{2K_j} \min_{0 \le k \le 2^{j+2}-1} \mathcal{H}^1(\varphi([A_{k,j}A_{k+1,j}])), \alpha_{j-1}\Big\}. \tag{5.3}$$

This gives us a non-increasing sequence α_j.

Notice that the intervals $(2^{-3j-1}, 2^{-3j+1})$ are pairwise disjoint. They remain pairwise disjoint if we multiply by some non-increasing sequence, i.e. the intervals,

$$J_j = (2^{-3j-1}\alpha_j, 2^{-3j+1}\alpha_j), \tag{5.4}$$

are also disjoint. This will be useful because we will want to define f on $h_j([0,8])$ as a piecewise-linear function and we will require that the image of each $h_j([0,8])$ is disjoint from the others. If we have that

$$\operatorname{dist}(f \circ h_j(t), \partial\Omega') \in J_j \text{ for all } t \in [0,8] \text{ and all } j \in \mathbb{N},$$

then we will have

$$f \circ h_{j_1}([0,8]) \cap f \circ h_{j_2}([0,8]) = \emptyset \quad \text{if} \quad j_1 \neq j_2.$$

In the j-th iteration of our construction we want to define a curve inside Ω' lying close to the boundary with a similar shape to the boundary and then define f on $h_j([0,8])$ as an appropriate parametrization of this curve. There are two steps we

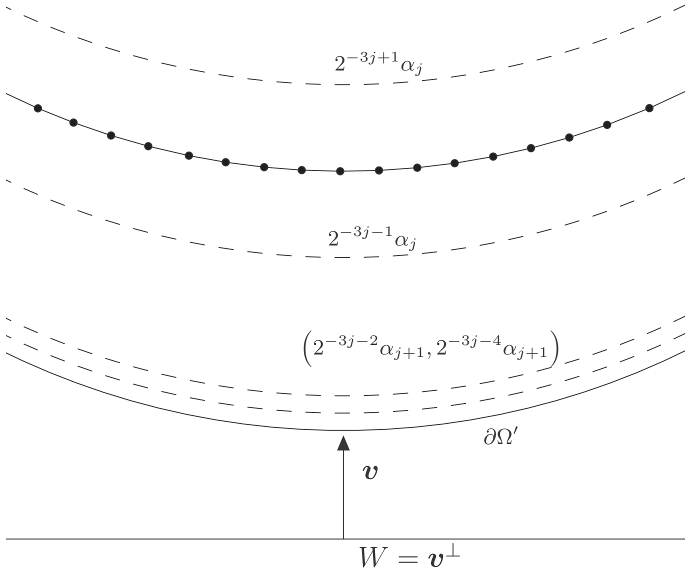


FIGURE 28. Approximations of the boundary, whose distance from it lie in J_j.

conduct in order to do so. Firstly we take a section of the boundary lying between two points in H approximate it using segments and then move it inside Ω' as is depicted in Figure 28. The second step is to connect such curves in a reasonable way near their ends to make sure that, together, they form a (piecewise-linear) closed path.

In the first of the two steps described above we have two neighbouring critical points $\boldsymbol{Y}_i, \boldsymbol{Y}_{i+1} \in H$ and $F = \varphi \circ h((s_i, s_{i+1}))$, the set of all points on $\partial\Omega'$ between $\boldsymbol{Y}_i$ and $\boldsymbol{Y}_{i+1}$. At the end of step 2 we found an $\varepsilon > 0$ such that $F + \big((-\varepsilon, \varepsilon)\boldsymbol{v}_i\big)$ is an open set which satisfies the role of U in Definition 5.1 and $W_i = \pi_i(F)$ the projection of F onto $\boldsymbol{v}_i^{\perp}$. Then W_i is a segment in $\boldsymbol{v}_i^{\perp}$, and let its length be called d. Choose a number

$$N = N_{j,F} = \left\lceil \frac{2dM8^j}{\alpha_j} \right\rceil + 2.$$

The number N is the number of points we allocate on W_i dividing it into $N-1$ equal parts each of length

$$l = \frac{d}{N-1} \leq \frac{\alpha_j}{2M8^j}.$$

Call these points t_k, $k = 1, \dots, N$. We know that the points $\boldsymbol{T}_k = t_k + L_i(t_k)\boldsymbol{v}_i$ lie on $\partial\Omega'$ so we define a curve which will approximate the boundary by connecting the points $\boldsymbol{T}_k$ with segments. These segments can obviously be expressed using a new M-Lipschitz function $\tilde{L}_i$ so that the approximation can be expressed as

$$w + \tilde{L}_i(w)\boldsymbol{v}_i \qquad w \in W_i.$$

We know that $\tilde{L}_i(t_k) = L_i(t_k)$ and that these two mappings are both M-Lipschitz, making their difference $2M$-Lipschitz. This means that

$$\|\cdot + L(\cdot)v - (\cdot + \tilde{L}(\cdot)v)\|_\infty \leq 2M\frac{l}{2} \leq \frac{\alpha_j}{2\cdot 8^j}.$$

Now we move this approximation inside Ω'. Define $\hat{L}_i(w) = \tilde{L}_i(w) + \frac{\alpha_j}{8^j}$ and put

$$\gamma_{i,j} = \{w + \hat{L}_i(w)\boldsymbol{v}_i : \ w \in W_i\} = \{w + \tilde{L}_i(w)\boldsymbol{v}_i + \frac{\alpha_j}{8^j}\boldsymbol{v}_i : \ w \in W_i\}.$$

It is easy to note that

$$\gamma_{i,j} \cap F = \emptyset.$$

It is well known that the shortest path connecting two points is a segment, so considering the way we defined $\hat{L}_i$ it is easy to see that

$$\mathcal{H}^1(\gamma_{i,j}) \leq \mathcal{H}^1(F). \tag{5.5}$$

In the second step we may assume that we have the curves $\gamma_{i,j}$ and $\gamma_{i+1,j}$ and we need to connect them near to $\boldsymbol{Y}_i \in H$. Start by defining a cone,

$$\kappa = \{\boldsymbol{Y}_i + \lambda_1\boldsymbol{v}_i + \lambda_2\boldsymbol{v}_{i+1} : \lambda_1, \lambda_2 > 0\}$$

and define $\tilde{\gamma}_{i,j} = \gamma_{i,j} \setminus \kappa$. We claim that there is precisely one point that lies in both $\gamma_{i,j}$ and the boundary of the cone. This is in fact quite simple. It is not possible for $\gamma_{i,j}$ to have two points on $\partial\kappa$ because the boundary of the cone is parallel to one of the n-vectors and $\gamma_{i,j}$ is the graph of a Lipschitz function with respect to both $\boldsymbol{v}_i$ and $\boldsymbol{v}_{i+1}$ on an open set containing $\boldsymbol{Y}_i$. All endpoints considered are clearly contained in $C_i \cap C_{i+1}$ thanks to our choice of ε.

This gives us two uniquely determined endpoints $\boldsymbol{E}_i$, $\boldsymbol{E}_{i+1}$ of our two curves $\tilde{\gamma}_{i,j}$ and $\tilde{\gamma}_{i+1,j}$, which we simply connect with a segment that lies in $\overline{\kappa}$. We have conducted two operations, the first being $\setminus\kappa$ which did not make our curve longer and the second is connecting our two endpoints. It is necessary to estimate how long such a segment may be. It is definitely shorter than the length of the curve which goes around the boundary of κ. As shown in Figure 29, there are two mutually exclusive possibilities

$$a)\ \boldsymbol{Y}_i \in \boldsymbol{E}_i + \mathbb{R}\boldsymbol{v}_i$$
$$b)\ \boldsymbol{Y}_i \in \boldsymbol{E}_i + \mathbb{R}\boldsymbol{v}_{i+1}.$$

In the first case we know that $|L_i - \hat{L}_i| \leq \frac{\alpha_j}{2\cdot 8^j}$. This means that

$$|\boldsymbol{E}_i - \boldsymbol{Y}_i| < \frac{2\alpha_j}{8^j}$$

and the total length around $\partial\kappa$ from $\boldsymbol{E}_i$ to $\boldsymbol{E}_{i+1}$ is at most $\frac{4\alpha_j}{8^j}$.

The second case we will get a similar estimate because the angle between our two n-vectors is very small. In order to show this we have to make some geometric considerations. Define $w_{i+1} = \pi_i(\boldsymbol{E}_{i+1})$ and $w_i = \pi_i(\boldsymbol{E}_i) \in W_i$. We intend to calculate the distance between w_i and w_{i+1} in W_i.

Consider the triangle $T = \boldsymbol{E}_i, \boldsymbol{Y}_i, \boldsymbol{P}$ pictured in Figure 30, where we have denoted $\zeta = \measuredangle\boldsymbol{v}_i\boldsymbol{v}_{i+1} \leq \theta$. Simple direct computation gives

$$\frac{|w_i - w_{i+1}|}{\tan\zeta} = |\boldsymbol{E}_i - \boldsymbol{P}| = r + L_i(w_i) - L(w_{i+1}) \leq r + M|w_{i+1} - w_i|. \tag{5.6}$$

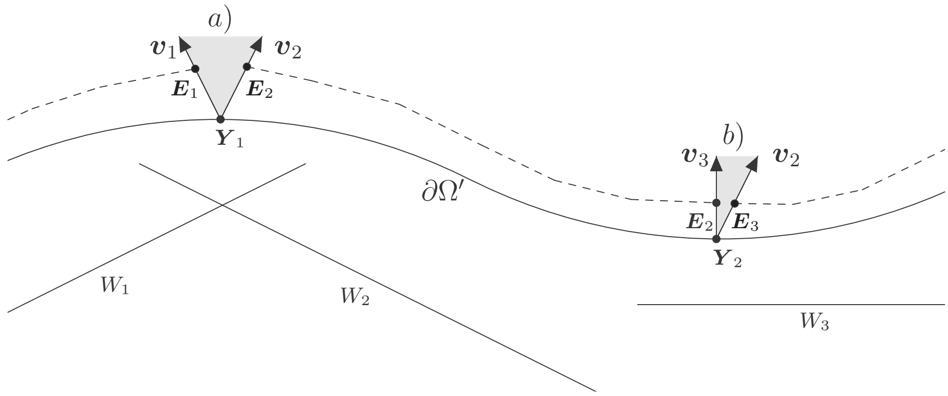


FIGURE 29. Two possibilities at a critical point $\boldsymbol{Y}_i$.

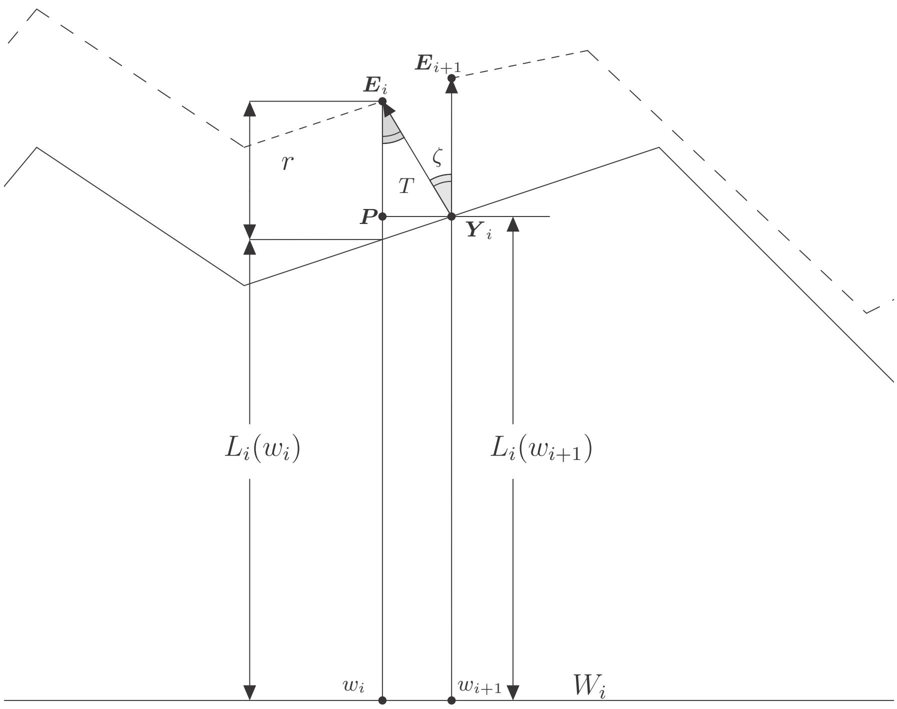


FIGURE 30. Geometry of a b)-type critical point.

Since (5.2) implies that $\tan\zeta < \frac{1}{3M}$ it follows that

$$\frac{1}{\tan\zeta} - M > \frac{2}{3\tan\zeta}.$$

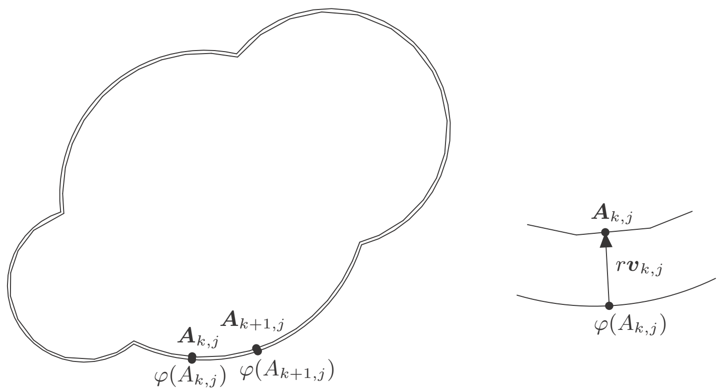


FIGURE 31. Dividing γ_j into sections that correspond to $\varphi([A_{k,j}A_{k+1,j}])$.

Applying the above to (5.6) and using the fact that L_1 is M-Lipschitz we get

$$\Big(\frac{1}{\tan\zeta} - M\Big)|w_i - w_{i+1}| \leq r$$

$$|w_i - w_{i+1}| \leq \frac{r}{\tan^{-1}\zeta - M} < \frac{3r\tan\zeta}{2}.$$

Now using the fact that $r \in J_j$ we can get

$$|\boldsymbol{E}_i - \boldsymbol{Y}_i| = \frac{|w_i - w_{i+1}|}{\sin\zeta} \leq \frac{3r}{2\cos\zeta} < \frac{4\alpha_j}{8^j}.$$

Finally we a similar calculation to estimate the distance $|\boldsymbol{E}_{i+1} - \boldsymbol{Y}_i|$.

We can conclude the above as follows. Firstly we have curves $\tilde{\gamma}_{i,j}$ which are pairwise disjoint with respect to $i = 1, 2, \dots n$ and $j \in \mathbb{N}$. The endpoints of $\tilde{\gamma}_{i,j}$ are such that,

$$|\boldsymbol{E}_i - \boldsymbol{E}_{i+1}| < \frac{8\alpha_j}{8^j}. \tag{5.7}$$

Secondly, if we connect these endpoints with segments, then the newly added segments do not intersect $\tilde{\gamma}_{i,j}$ anywhere other than $\boldsymbol{E}_i$ and $\boldsymbol{E}_{i+1}$ thanks to the operation $\backslash\kappa$. All the newly added segments are pairwise disjoint with respect to j because of (5.4). Using (5.5) and (5.7) we know that the length the total curve (which we now call γ_j) is

$$\mathcal{H}^1(\gamma_j) \leq \mathcal{H}^1(\partial\Omega') + \frac{8m\alpha_j}{8^j}. \tag{5.8}$$

Now for each given $j \in \mathbb{N}$ we will split up γ_j into those parts, which correspond to the parts of the boundary $\varphi([A_{k,j}, A_{k+1,j}])$. Remember that for all $\boldsymbol{Y}_i \in H$ we have $s_i \in [0, 8]$ such that $\varphi \circ h(s_i) = \boldsymbol{Y}_i$. Also $A_{k,j} = h(2^{1-j}k)$. We know that $H \cap \mathcal{A} = \emptyset$ and therefore for every $k = 0, 1, \dots, 2^{j+2} - 1$ and for every $j \in \mathbb{N}$ we make the definition

$$\boldsymbol{v}_{k,j} = \boldsymbol{v}_{\varphi(A_{k,j})}.$$

As depicted in Figure 31 there exists exactly one $\boldsymbol{A}_{k,j} \in \gamma_j$ such that $\varphi(A_{k,j}) + r\boldsymbol{v}_{k,j} = \boldsymbol{A}_{k,j}$ with $r \in J_j$. We now find a parametrization g_j of the curve γ_j such that $g_j : [0, 8] \to \mathbb{R}^2$ has

$$g_j(k2^{1-j}) = \boldsymbol{A}_{j,k}$$

and $|g_j'|$ is constant (almost everywhere) on $(k2^{1-j}, (k+1)2^{1-j})$. We now wish to estimate the length of the parts of γ_j between $\boldsymbol{A}_{k,j}$ and $\boldsymbol{A}_{k+1,j}$. The only difference between this and (5.8) is that we replace the number m with the number of critical points in our interval. Thereby, and recalling our definition of α_j in (5.3), we get,

$$\begin{aligned}\mathcal{H}^1\Big(g_j\big([k2^{1-j}, (k+1)2^{1-j}]\big)\Big) &\leq \mathcal{H}^1(\varphi([A_{k,j}A_{k+1,j}])) + \frac{8K_j\alpha_j}{8^j}\\ &\leq \mathcal{H}^1(\varphi([A_{k,j}A_{k+1,j}])) + \frac{8\mathcal{H}^1\big(\varphi([A_{k,j}A_{k+1,j}])\big)}{8^j}.\end{aligned} \tag{5.9}$$

Step 4: Defining f on Q.

Now we will define f on $[B_{k,j}B_{k+1,j}]$ as the constant-speed parametrization of the curve $g_j\big([k2^{1-j}, (k+1)2^{1-j}]\big)$ such that $f(B_{k,j}) = \boldsymbol{A}_{k,j}$. By (5.9) we have,

$$|D_\tau f(X)| \leq 2 ⨍_{[A_{k,j}A_{k+1,j}]} |D_\tau \varphi| d\mathcal{H}^1 , \tag{5.10}$$

for $\mathcal{H}^1$-almost every $X \in [B_{k,j}B_{k+1,j}]$. Notice that

$$\boldsymbol{A}_{k,j} - \boldsymbol{A}_{2k,j+1} = \lambda_{k,j}\boldsymbol{v}_{k,j} \text{ for some appropriate } \lambda_{k,j} > 0.$$

So $\boldsymbol{A}_{k,j}$ and $\boldsymbol{A}_{2k,j+1}$ can be connected with a segment parallel to $\boldsymbol{v}_{k,j} = \boldsymbol{v}_{2k,j+1}$. Recalling (5.3) we have

$$\mathcal{H}^1([\boldsymbol{A}_{k,j}\boldsymbol{A}_{2k,j+1}]) \leq 2^{-3j+1}\alpha_j < \mathcal{H}^1(\varphi([A_{k,j}A_{k+1,j}])). \tag{5.11}$$

We define f on $[B_{k,j}B_{2k,j+1}]$ as the constant speed parametrization of the segment $[\boldsymbol{A}_{k,j}\boldsymbol{A}_{2k,j+1}]$.

We can now define f in the quadrilateral $Q_{k,j} = \mathrm{co}\{B_{k,j}, B_{k+1,j}, B_{2k,j+1}, B_{2k+2,j+1}\}$. All $Q_{k,j}$ are 4-bi-Lipschitz equivalent with a square so we can apply the piecewise-affine bi-Lipschitz change of variables, use Theorem 2.1 and then reverse the change of variables to get a piecewise-affine mapping on $Q_{k,j}$. We need to calculate the integral of $\Phi(|D_\tau f|)$ over the boundary $\partial Q_{k,j}$. Using (5.10) we can see that

$$\int_{[B_{k,j}B_{k+1,j}]} |D_\tau f| \leq C \int_{[A_{k,j}A_{k+1,j}]} |D_\tau \varphi| \text{ for all } j \in \mathbb{N},\ k = 0, 1, 2, \dots 2^{j+2} - 1.$$

As $|D_\tau f|$ is constant on $[B_{k,j}B_{k+1,j}]$ and $[B_{k,j}B_{2k,j+1}]$, the integral over three sides of $\partial Q_{k,j}$ can be estimated very easily. The first side is estimated as

$$\begin{aligned}⨍_{[B_{i,j}B_{i+1,j}]} \Phi(|D_\tau f|) &= \Phi\Big(⨍_{[B_{i,j}B_{i+1,j}]} |D_\tau f|\Big)\\ &\leq C\Phi\Big(⨍_{[A_{i,j}A_{i+1,j}]} |D_\tau \varphi|\Big)\\ &\leq C ⨍_{[A_{i,j}A_{i+1,j}]} \Phi(|D_\tau \varphi|).\end{aligned}$$

We can do the same on the two sides of $Q_{k,j}$, namely $[B_{k,j}B_{2k,j+1}]$ and $[B_{k+1,j}B_{2k+2,j+1}]$ thanks to (5.11). The size of $D_\tau f$ is constant on our segments and without loss of generality assume that $|D_\tau f(X)| \geq |D_\tau f(Y)|$ for $\mathcal{H}^1$ almost all $X \in [B_{2k,j+1}B_{2k+1,j+1}]$

and almost all $Y \in [B_{2k+1,j+1}B_{2k+2,j+1}]$. We can calculate

$$\begin{aligned}
\fint_{[B_{2k,j+1}B_{2k+2,j+1}]} \Phi(|D_\tau f|) &\leq \fint_{[B_{2k,j+1}B_{2k+1,j+1}]} \Phi(|D_\tau f|)\\
&= \Phi\Big(\fint_{[B_{2k,j+1}B_{2k+1,j+1}]} |D_\tau f|\Big)\\
&\leq C\Phi\Big(\fint_{[A_{2k,j+1}A_{2k+1,j+1}]} |D_\tau \varphi|\Big)\\
&\leq C\Phi\Big(2\fint_{[A_{2k,j+1}A_{2k+2,j}]} |D_\tau \varphi|\Big)\\
&\leq C\fint_{[A_{2k,j+1}A_{2k+2,j+1}]} \Phi(|D_\tau \varphi|) = C\fint_{[A_{k,j}A_{k+1,j}]} \Phi(|D_\tau \varphi|).
\end{aligned}$$

Altogether we get

$$\fint_{\partial Q_{k,j}} \Phi(|D_\tau f|) \leq C\fint_{[A_{k,j}A_{k+1,j}]} \Phi(|D_\tau \varphi|)$$

with C a fixed constant depending only on Φ. Now using Theorem 2.1 we can extend f inside $Q_{k,j}$ with the estimate

$$\fint_{Q_{k,j}} \Phi(|Df|) \leq C\fint_{[A_{k,j}A_{k+1,j}]} \Phi(|D_\tau \varphi|).$$

Notice that $|Q_{k,j}| \approx 2^{-2j}$ and that $\mathcal{H}^1([A_{k,j}A_{k+1,j}]) \approx 2^{-j}$ so we get

$$\int_{Q_{k,j}} \Phi(|Df|) \leq C2^{-j}\int_{[A_{k,j}A_{k+1,j}]} \Phi(|D_\tau \varphi|).$$

Sum this over k to get

$$\int_{1-2^{-j-1}\leq|x|_\infty\leq 1-2^{-j}} \Phi(|Df|) \leq C2^{-j}\int_{\partial(-1,1)^2} \Phi(|D_\tau \varphi|).$$

Summing over j gives us our result. □

Acknowledgement. The author would like to thank Stanislav Hencl for carefully reading the manuscript and for his pointed comments. The author would also like to express gratitude to Aldo Pratelli and Emanuela Radici for their helpful remarks. The author was supported by the ERC CZ grant LL1203 of the Czech Ministry of Education.

References

[1] E. Acerbi and N. Fusco, Semicontinuity problems in the calculus of variations, *Arch. Rat. Mech. Anal.* **86** (1984), 125–145.

[2] J.M. Ball, Convexity conditions and existence theorems in nonlinear elasticity, *Arch. Rat. Mech. Anal.* **63** (1977), 337–403.

[3] J.M. Ball, Discontinuous equilibrium solutions and cavitation in nonlinear elasticity, *Philos. Trans. R. Soc. Lond. A* **306** n.1496 (1982), 557–611.

[4] J.M. Ball, *Singularities and computation of minimizers for variational problems*, Foundations of computational mathematics (Oxford, 1999), 1–20, London Math. Soc. Lecture Note Ser., 284, Cambridge Univ. Press, Cambridge, 2001.

[5] J.M. Ball, *Progress and puzzles in Nonlinear Elasticity*, Proceedings of course on Poly-, Quasi- and Rank-One Convexity in Applied Mechanics, CISM, Udine, to appear.

[6] J.C. Bellido and C. Mora-Corral, Approximation of Hölder continuous homeomorphisms by piecewise affine homeomorphisms, *Houston J. Math.* **37** no.2 (2011), 449–500.
[7] S. Daneri and A. Pratelli, *Smooth approximation of bi-Lipschitz orientation-preserving homeomorphisms*, to appear in Ann. Inst. H. Poincaré Anal. Non Linéaire.
[8] S. Daneri and A. Pratelli, *A planar bi-Lipschitz extension theorem*, to appear in Adv. Calc. Var.
[9] F. W. Gehring and O. Lehto, On the total differentiability of functions of a complex variable, *Ann. Acad. Sci. Fenn. Ser. A I* **272** (1959), 1–9.
[10] S. Hencl, Sobolev homeomorphism with zero Jacobian almost everywhere, *J. Math. Pures Appl.* **95** (2011), 444–458.
[11] S. Hencl and P. Koskela, *Lectures on Mappings of finite distortion*, Lecture Notes in Mathematics 2096, Springer, 2014, 176pp.
[12] S. Hencl and A. Pratelli, *Diffeomorphic approximation of $W^{1,1}$ planar Sobolev homeomorphisms*, Preprint.
[13] S. Hencl, B. Vejnar, *Sobolev homeomorphism that cannot be approximated by diffeomorphisms in $W^{1,1}$*., Arch. Rational Mech. Anal., to appear.
[14] T. Iwaniec, L.V. Kovalev and J. Onninen, Diffeomorphic Approximation of Sobolev Homeomorphisms, *Arch. Rational Mech. Anal.* **201** no. 3 (2011), 1047–1067.
[15] T. Iwaniec, L.V. Kovalev and J. Onninen, Hopf Differentials and Smoothing Sobolev Homeomorphisms , *International Mathematics Research Notices* **14** (2012), 3256–3277.
[16] J. Luukkainen, Lipschitz and quasiconformal approximation of homeomorphism pairs, *Topology Appl.* **109** (2001), 1–40.
[17] C. Mora-Corral, Approximation by piecewise affine homeomorphisms of Sobolev homeomorphisms that are smooth outside a point, *Houston J. Math.* **35** (2009), 515–539.
[18] C. Mora-Corral and A. Pratelli, Approximation of Piecewise Affine Homeomorphisms by Diffeomorphisms, *J Geom. Anal* **24** (2014), 1398–1424.
[19] C.B. Morrey, Quasi-convexity and the semicontinuity of multiple integrals, *Pacific J. Math.* **2** (1952), 25–53.
[20] S. Müller, Q. Tang and B. S. Yan, On a new class of elastic deformations not allowing for cavitation, *Ann. Inst. Henri Poincaré* **11** (1994), 217–243.
[21] A. Pratelli, *On the bi-Sobolev planar homeomorphisms and their approximation*, Preprint.
[22] E. Radici, *A planar Sobolev extension theorem for piecewise linear homeomorphisms*, Preprint.

D. Campbell: Department of Mathematics, University of Hradec Králové, Rokitanského 62, 500 03 Hradec Králové, Czech Republic

E-mail address: daniel.campbell@uhk.cz

5. Approximation of $W^{1,p}$ Sobolev homeomorphism by diffeomorphisms and the signs of the Jacobian

Daniel Campbell, Stanislav Hencl and Ville Tengvall

Preprint

APPROXIMATION OF $W^{1,p}$ SOBOLEV HOMEOMORPHISM BY DIFFEOMORPHISMS AND THE SIGNS OF THE JACOBIAN

DANIEL CAMPBELL, STANISLAV HENCL, AND VILLE TENGVALL

Dedicated to Professor Jan Malý on his 60th birthday

Abstract. Let $\Omega \subset \mathbb{R}^n$, $n \geq 4$, be a domain and $1 \leq p < [n/2]$, where $[a]$ stands for the integer part of a. We construct a homeomorphism $f \in W^{1,p}((-1,1)^n, \mathbb{R}^n)$ such that $J_f = \det Df > 0$ on a set of positive measure and $J_f < 0$ on a set of positive measure. It follows that there are no diffeomorphisms (or piecewise affine homeomorphisms) f_k such that $f_k \to f$ in $W^{1,p}$.

1. Introduction

The problem of approximating homeomorphisms $f : \mathbb{R}^n \supseteq \Omega \longrightarrow f(\Omega) \subseteq \mathbb{R}^n$ with either diffeomorphisms or piecewise-affine homeomorphisms has proven to be both very challenging and of great interest in a variety of contexts. As far as we know, in the simplest non-trivial setting (i.e. $n = 2$, approximations in the L^∞-norm) the problem was solved by Radó [38]. Due to its fundamental importance in geometric topology, the problem of finding piecewise affine homeomorphic approximations in the L^∞-norm and dimensions $n > 2$ was deeply investigated in the '50s and '60s. In particular, it was solved by Moise [33] and Bing [8] in the case $n = 3$ (see also the survey book [34]), while for contractible spaces of dimension $n \geq 5$ the result follows from theorems of Connell [13], Bing [9], Kirby [29] and Kirby, Siebenmann and Wall [30] (for a proof see, e.g., Rushing [40, Theorem 4.11.1.] or Luukkainen [31]). Finally, twenty years later, while studying the class of quasi-conformal manifolds, Donaldson and Sullivan [16] proved that the result is false in dimension 4.

After the L^∞-approximation problem had been completely solved, the question of approximating homeomorphisms revived again in the altogether different context for variational models in nonlinear elasticity. Let us briefly explain this. Let $\Omega \subset \mathbb{R}^n$ be a domain which models a body made out of homogeneous elastic material, and suppose that a mapping $f : \Omega \to \mathbb{R}^n$ is modeling the deformation of this body with prescribed boundary values. If we want to study the properties of the deformation in the setting of nonlinear elasticity theory of Antman, Ball and Ciarlet, see e.g. [2, 3, 4, 12], we are led to study the existence and regularity properties of minimizers of the energy functionals of the form

$$I(f) = \int_\Omega W(Df)\,\mathrm{d}x,$$

1991 *Mathematics Subject Classification.* 46E35.

All authors were supported by the ERC CZ grant LL1203 of the Czech Ministry of Education. V. Tengvall was also supported by the Vilho, Yrjö and Kalle Väisälä foundation and the Academy of Finland Project 277923.

where $W : \mathbb{R}^{n\times n} \to \mathbb{R}$ is so-called stored-energy functional, and Df is the differential matrix of a deformation f. In order for this model to be physically relevant we have to require this model to satisfy the following conditions:

(W1) $W(A) \to +\infty$ as $\det A \to 0$, which prevents too high compression of the elastic body.

(W2) $W(A) = +\infty$ if $\det A \leq 0$, which guarantees that the orientation is preserved.

In particular, it follows that if f is an admissible deformation with finite energy, then we have

$$J_f(x) := \det Df(x) > 0 \quad \text{for a.e. } x \in \Omega.$$

Using other assumptions one can prove that the mapping with finite energy is continuous and one-to-one, which corresponds to the non-impenetrability of the matter. Therefore the natural candidate for a minimizer is in fact a homeomorphism. Hence, when we study this model it is natural to restrict our attention only on Sobolev homeomorphisms where the Jacobian does not change sign.

As pointed out by Ball in [5, 6] (who ascribes the question to Evans [18]), an important issue toward understanding the regularity of the minimizers in this setting would be to show the existence of minimizing sequences given by piecewise affine homeomorphisms or by diffeomorphisms. In particular, a first step would be to prove that any homeomorphism $u \in W^{1,p}(\Omega;\mathbb{R}^n)$, $p \in [1,+\infty)$, can be approximated in $W^{1,p}$ by piecewise affine ones or smooth ones. One very significant reason why this would be desirable, is that regularity is typically often proven by testing the weak equation or the variation formulation by the solution itself; but unless one has some a priori regularity of the solution, such a test may not make sense. In order to solve this problem it would be possible to test the equation with a smooth test mapping which is close to the given homeomorphism instead. Here we see the necessity for the approximations to be homeomorphisms whose image is the same as that of the approximated map, otherwise this sequence would have nothing in common with our original problem. Besides non-linear elasticity, an approximation result of homeomorphisms with diffeomorphisms would be a very useful tool in and of itself as it would allow a number of proofs to be significantly simplified and lead to some stronger results. Let us note that finding diffeomorphisms near a given homeomorphism is not an easy task, as the usual approximation techniques like mollification or Lipschitz extension using the maximal operator destroy, in general, injectivity.

Let us describe the results in this direction. The first positive results were achieved by Mora-Corral [35] on planar homeomorphisms smooth outside a point and by Bellido and Mora-Corral [7] on approximation in Hölder continuous maps. Let us also note that the problem of approximation by smooth or piecewise affine planar homeomorphisms are in fact equivalent by the result of Mora-Corral and Pratelli [36]. The celebrated breakthrough result in the area which stimulated much interest in the subject was given by Iwaniec, Kovalev and Onninen in [27], [28], where they found diffeomorphic approximations to any homeomorphism $f \in W^{1,p}(\Omega,\mathbb{R}^2)$, for any $1 < p < \infty$ in the $W^{1,p}$ norm. The remaining missing case $p = 1$ in the plane has been solved by Hencl and Pratelli in [25] by a different method. This method was extended by Campbell [10] to give a different proof of the $W^{1,p}$, $p > 1$, case and to prove the result also for Orlicz-Sobolev spaces. The problem of approximating homeomorphisms with diffeomorphisms cannot be considered entirely closed even in the planar case. Another problem mentioned in [27] is to approximate both a map and its inverse simultaneously in $W^{1,p}$. The first results in this area was given by Daneri and Pratelli in [15] for all $1 \leq p < \infty$ under the additional assumption that the mapping is bi-Lipschitz. Recently Pratelli [37] has answered this question for $p = 1$ (without any additional assumptions) using the technique of [25]. The cases $p > 1$ (especially $p = 2$) which are even more important in terms of their application are still open.

And even more interesting open problem is the approximation of Sobolev homeomorphism in dimension $n = 3$ as there are no results in this direction so far. The only breakthrough result in higher dimension is the result of Hencl and Vejnar in [26] that there is a homeomorphism in $W^{1,1}$ for $n \geq 4$ which cannot be approximated by diffeomorphisms. The main result of this paper is the following extension, which shows that the problem is not in the special choice of nonreflexive space $W^{1,1}$.

Theorem 1.1. *Let $n \geq 4$ and $1 \leq p < [n/2]$. Then there exists a homeomorphism $f \in W^{1,p}((-1,1)^n, \mathbb{R}^n)$ such that there are no diffeomorphisms (or piecewise affine homeomorphisms) $f_k : (-1,1)^n \to \mathbb{R}^n$ such that $f_k \to f$ in $W^{1,p}_{\mathrm{loc}}((-1,1)^n, \mathbb{R}^n)$.*

Here $[n/2]$ denotes the integer part of $n/2$, i.e. $1 \leq p < 2$ for $n = 4, 5$, $1 \leq p < 3$ for $n = 6, 7$ and so on. This result is deeply connected with the sign of the Jacobian of a homeomorphism. As we mentioned before in models of nonlinear elasticity one usually assumes that $J_f > 0$ a.e. (or at least $J_f \geq 0$ a.e.). It is therefore natural to ask if this condition is automatically satisfied (up to a reflection) in the reasonable class of mappings. This problem was promoted by Hajlasz, see e.g. Goldstein and Hajlasz [21]. As each homeomorphism on a domain is either sense-preserving or sense-reversing (see Preliminaries) we can equivalently ask if the topological (sense-preserving) and analytical ($J_f \geq 0$) notion of orientation are the same.

Another reason to study nonnegativity of the Jacobian comes from the well-known area formula which is one of the most fundamental tools in the area. For a Sobolev homeomorphism $f : \Omega \to \mathbb{R}^n$ for which the Lusin's condition (N) (i.e. sets of null measure are always mapped to sets of null measure) holds we have

$$\int_{\Omega} \eta(f(x))\, |J_f(x)| \,\mathrm{d}x = \int_{f(\Omega)} \eta(y) \,\mathrm{d}y \tag{1.1}$$

for every nonnegative Borel function $\eta : f(\Omega) \to [0, \infty]$ (see Federer [19]). If we knew that $J_f \geq 0$ a.e. we could write the formula (1.1) without absolute values.

It is relatively easy to show that every topologically sense-preserving Sobolev homeomorphism which is differentiable almost everywhere has nonnegative Jacobian almost everywhere, see [32, Lemma 2.14]. Therefore every sense-preserving planar homeomorphism in $W^{1,1}_{\mathrm{loc}}(\Omega, \mathbb{R}^2)$, and more generally every sense-preserving homeomorphism in $W^{1,p}_{\mathrm{loc}}(\Omega, \mathbb{R}^n)$ with $p > n - 1$, satisfies $J_f \geq 0$ a.e. (see [23, Corollary 2.25 and Theorem 5.22.]). However, when we study homeomorphisms in $W^{1,p}_{\mathrm{loc}}(\Omega, \mathbb{R}^n)$ with $n \geq 3$ and $1 \leq p \leq n - 1$ it might happen that the mapping is nowhere differentiable even under some additional assumptions, see e.g. [14]. Thus the previous argument which heavily uses differentiability of the mapping cannot be used anymore when $f \in W^{1,p}$, $p \in [1, n - 1]$.

In [24] Hencl and Malý were able to overcome the difficulties caused by the lack of differentiability by giving the first nontrivial positive answer to the question about the nonnegativity of the Jacobian of Sobolev homeomorphisms. More precisely, they showed that every sense-preserving Sobolev homeomorphism $f \in W^{1,p}_{\mathrm{loc}}(\Omega, \mathbb{R}^n)$ with $p > [n/2]$ has nonnegative Jacobian at almost every point. The proof was based on the approximative differentiability of Sobolev mappings and on the topological invariance of the linking number under homeomorphisms. The restriction $p > [n/2]$ in their proof comes from the linking number argument where one has to require the mapping to behave geometrically nicely on both "links". Here we show that somewhat surprisingly the strange exponent $[n/2]$ is indeed the borderline exponent for this question.

Theorem 1.2. *Let $n \geq 4$ and $1 \leq p < [n/2]$. Then there is a homeomorphism $f \in W^{1,p}((-1,1)^n, \mathbb{R}^n)$ such that $J_f > 0$ on a set of positive measure and $J_f < 0$ on a set of positive measure.*

This result for $p = 1$ was shown by Hencl and Vejnar in [26] and as in their paper Theorem 1.1 now follows easily. Indeed, assume on the contrary that f from the statement can be approximated by diffeomorphisms (or piecewise affine homeomorphisms) $\{f_k\}_{k=1}^{\infty}$, then the pointwise limit of a subsequence (which we denote the same) satisfies

$$Df_k(x) \to Df(x) \quad \text{and} \quad J_{f_k}(x) \to J_f(x)$$

for almost every $x \in (-1,1)^n$. As f_k are locally Lipschitz we know that $J_{f_k} \geq 0$ a.e. in $(-1,1)^n$ or $J_{f_k} \leq 0$ a.e. in $(-1,1)^n$, see e.g. [24] and [23, Theorem 5.22]. The pointwise limit of nonnegative (or nonpositive) functions J_{f_k} cannot change sign which gives us contradiction.

Let us also recall that the Jacobian of a $W^{1,p}$, $1 \leq p < n$, Sobolev homeomorphism may behave strangely as it may vanish a.e. (see [22], [11] and [17]). As mentioned before the Jacobian of a homeomorphism cannot change sign if $p > [n/2]$ by [24] and therefore the method of sign-changing Jacobian for providing a counterexample in Theorem 1.1 cannot be improved to $p > [n/2]$. On the other hand, there might be a different way of producing a counterexample to the Ball-Evans approximation problem or there might be even a positive result in $\mathbb{R}^n$, $n \geq 4$, for $W^{1,p}$ if p is large enough (but definitely we must have $p \geq [n/2]$). Also the question whether the Jacobian can have both positive and negative Jacobian in a sets of positive measure in the borderline case $p = [n/2]$ remains open.

Now we outline the rough idea of our construction. We fix a Cantor type set $\mathcal{C}_A \subset (-1,1)$ of positive measure and we set

$$\begin{aligned} \mathcal{K}_A := &\big(\mathcal{C}_A \times \mathcal{C}_A \times \mathcal{C}_A \times [-1,1]\big) \cup \big(\mathcal{C}_A \times \mathcal{C}_A \times [-1,1] \times \mathcal{C}_A\big) \cup \\ &\cup \big(\mathcal{C}_A \times [-1,1] \times \mathcal{C}_A \times \mathcal{C}_A\big) \cup \big([-1,1] \times \mathcal{C}_A \times \mathcal{C}_A \times \mathcal{C}_A\big). \end{aligned} \tag{1.2}$$

We also fix a Cantor type set $\mathcal{C}_B \subset (-1,1)$ of zero measure (in fact its Hausdorff dimension δ is small) and define the set $\mathcal{K}_B$ similarly as above. Our first mapping $S_q : \mathbb{R}^n \to \mathbb{R}^n$ squeezes $\mathcal{K}_A$ onto $\mathcal{K}_B$ homeomorphically in a natural way. Then we find a bi-Lipschitz sense-preserving homeomorphism F such that

$$F(x_1, x_2, x_3, x_4) = (x_1, x_2, x_3, -x_4) \text{ for every } x \in \mathcal{K}_B. \tag{1.3}$$

Indeed, we can find a direction in $\mathbb{R}^4$ such that the projection of $\mathcal{K}_B$ to the corresponding hyperplane is one-to-one. The rough reason for that is that the set of directions where the projection is not one-to-one has Hausdorff dimension at most

$$\dim \mathcal{K}_B + \dim \mathcal{K}_B = 2 + 6\delta \tag{1.4}$$

(starting+ending point of the direction) and this is smaller than 3-the dimension of all directions. This projection of $\mathcal{K}_B$ can be extended to the homeomorphism $g : \mathbb{R}^4 \to \mathbb{R}^4$ which is bi-Lipschitz. By the turnover of the 3-dimensional hyperplane with respect to x_1 direction (which can be done by a sense-preserving homeomorphism of $\mathbb{R}^4$) and the composition with g^{-1} we obtain our mapping F. In view of the turnover of the hyperplane we obtain the key property (1.3). At last we find a mapping $S_t : \mathbb{R}^n \to \mathbb{R}^n$ which stretches $\mathcal{C}_B \times \mathcal{C}_B \times \mathcal{C}_B \times \mathcal{C}_B$ back to $\mathcal{C}_A \times \mathcal{C}_A \times \mathcal{C}_A \times \mathcal{C}_A$ such that lines in $\mathcal{K}_B$ through the Cantor set are not prolonged too much and that S_t is locally Lipschitz outside of $\mathcal{C}_B \times \mathcal{C}_B \times \mathcal{C}_B \times \mathcal{C}_B$.

We verify that $f = S_t \circ F \circ S_q$ belongs to $W^{1,p}$ by using the ACL property. It is thus crucial for us that lines parallel to coordinate axes that intersect $\mathcal{C}_A \times \mathcal{C}_A \times \mathcal{C}_A \times \mathcal{C}_A$ are mapped to lines by S_q, then to the same lines (with possibly reverse orientation in x_4-direction) by F (see (1.3)) and to something of reasonable length by S_t. To control the derivative on the lines parallel to coordinate axes that do not intersect $\mathcal{C}_A \times \mathcal{C}_A \times \mathcal{C}_A \times \mathcal{C}_A$ we use explicit form of mappings S_q and S_t and it is essential for us that F is Lipschitz everywhere and that S_t is locally Lipschitz far away from $\mathcal{C}_B \times \mathcal{C}_B \times \mathcal{C}_B \times \mathcal{C}_B$.

Let us compare this result to the methods in [26]. In [26] the authors only showed that the length of the images of line segments are finite (which is enough for $\int |Df| < \infty$) but here we need to write explicit formulas for the mappings and to differentiate them, which requires much more details, precision and a delicate case study. More importantly there are three new main essential ingredients here. First there is a gap in the argument of [26] in the construction of the last mapping. During our detailed estimates we have found this gap and we have repaired it by giving a different last mapping S_t such that lines in $\mathcal{K}_B$ through the Cantor set are not prolonged too much. Secondly in [26] it was enough to find any bi-Lipschitz extension of the projection to construct a mapping F. Here we need to know that line segments close to $\mathcal{K}_B$ but far away from $\mathcal{C}_B \times \mathcal{C}_B \times \mathcal{C}_B \times \mathcal{C}_B$ are mapped to line segments (see Section 3) so that the partial derivatives corresponding to different directions do not mix (and the big derivative in one direction is not multiplied by a big derivative in other direction). This requires a novel construction of the mapping F in Section 3. The third main ingredient is the extension to higher dimension as in $W^{1,1}$ it was enough to extend simply as $\tilde{f}(x) = (f(x_1, x_2, x_3, x_4), x_5, \ldots, x_n)$. Here it requires much more work and it is essential for us to consider not only line segments through $\mathcal{C}_A \times \mathcal{C}_A \times \ldots \times \mathcal{C}_A$ as in (1.2) but $[n/2] - 1$ dimensional planes through the Cantor set (i.e. 2 dimensional planes for $n = 6, 7$ and so on). Then F reflects not only line segments through $\mathcal{C}_A \times \mathcal{C}_A \times \ldots \times \mathcal{C}_A$ (see (1.3)) but it reflects $[n/2] - 1$ dimensional planes through the Cantor set as the analogy of (1.4) is now

$$\dim \mathcal{K}_B + \dim \mathcal{K}_B = 2([n/2] - 1) + 2(n - [n/2] + 1)\delta < n - 1.$$

This allows us to control the derivative only on lines that do not belong to this $[n/2]-1$ dimensional planes and the measure of this set is very small close to the Cantor set.

2. Preliminaries

2.1. **Notation.** A point $x \in \mathbb{R}^n$ in coordinates is denoted as $(x_1, x_2, \dots, x_n)$. We denote by $|x| := \sqrt{\sum_{i=1}^n x_i}$ the Euclidean norm of a point $x \in \mathbb{R}^n$, and $\|x\| := \sup_i |x_i|$ will denote the supremum norm of x. We also define the distance of two sets $A, B \subset \mathbb{R}^n$ as

$$\operatorname{dist}(A, B) := \inf\{|x - y| : x \in A \text{ and } y \in B\}.$$

We will denote by

$$Q(a, r) := (a_1 - r, a_1 + r) \times \cdots \times (a_n - r, a_n + r)$$

the open cube centered at $a \in \mathbb{R}^n$ with sidelength $2r > 0$. The interior of a set $A \subset \mathbb{R}^n$ is sometimes denoted also by A°.

We will denote by $C := C(p_1, \dots, p_k)$ a positive constant which depends only on the given parameters $p_1, \dots p_k$. The constant C might change from line to line. Furthermore, for given functions f and g we denote $f \lesssim g$ if there exists a positive constant $C > 0$ such that $f(x) \leq Cg(x)$ for all points x. If both conditions $f \lesssim g$ and $g \lesssim f$ are satisfied we denote $f \sim g$.

2.2. **Sobolev spaces and the ACL condition.** Let $\Omega \subset \mathbb{R}^n$ be an open set. We say that $f : \Omega \to \mathbb{R}^m$ belongs to the Sobolev space $W^{1,p}(\Omega, \mathbb{R}^m)$, $1 \leq p < \infty$, if f is p-integrable and if the coordinate functions of f have p-integrable distributional derivatives. We say that f belongs to the space $W^{1,p}_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$ if $f \in W^{1,p}(\Omega', \mathbb{R}^m)$ for every subdomain $\Omega' \subset\subset \Omega$.

Let $i \in \{1, 2, \dots, n\}$ and denote by π_i the projection on the given hyperplane $H_i = \{x \in \mathbb{R}^m : x_i = 0\}$ perpendicular to the x_i-axis. We say that a mapping $f \in L^1_{\mathrm{loc}}(\Omega, \mathbb{R}^m)$ is absolutely continuous on lines (abbr. $f \in \mathrm{ACL}(\Omega, \mathbb{R}^m)$) if the following ACL conditions holds:

(ACL) For every cube $Q(a, r) = (a_1 - r, a_1 + r) \times \cdots \times (a_n - r, a_n + r) \subset\subset \Omega$ and for every $i \in \{1, 2, \dots, n\}$ the coordinate functions of the mapping

$$f^i(t; x) := f(x_1, \dots, x_{i-1}, x_i + t, x_{i+1}, \dots, x_n)$$

are absolutely continuous on $(a_i - r, a_i + r)$ for $\mathcal{L}^{n-1}$-almost every $x \in \pi_i(Q(a, r))$.

The following characterization of Sobolev spaces is classical and can be found e.g. in [1, Section 3.11] and [23, Theorem A.15].

Proposition 2.1. *Let $1 \leq p < \infty$, $\Omega \subset \mathbb{R}^n$ be an open set and $f \in L^p(\Omega, \mathbb{R}^m)$. Then $f \in W^{1,p}_{\mathrm{loc}}(\Omega, \mathbb{R}^n)$ if and only if there is a representative of f which is a $\mathrm{ACL}(\Omega, \mathbb{R}^m)$ mapping with locally L^p-integrable partial derivatives on Ω.*

2.3. **Topological degree.** For a given smooth map f from $\Omega \subset \mathbb{R}^n$ into $\mathbb{R}^n$ we can define the *topological degree* as

$$\deg(f, \Omega, y_0) = \sum_{\{x \in \Omega : f(x) = y_0\}} \operatorname{sgn}(J_f(x))$$

if $J_f(x) \neq 0$ for each $x \in f^{-1}(y_0)$. This definition can be extended to arbitrary continuous mappings and each point, see e.g. [20].

A continuous mapping $f : \Omega \to \mathbb{R}^n$ is called *sense-preserving* if

$$\deg(f, \Omega', y_0) > 0$$

for all subdomains $\Omega' \subset\subset \Omega$ and for all $y_0 \in f(\Omega') \setminus f(\partial\Omega')$. Similarly we call f *sense-reversing* if $\deg(f, \Omega', y_0) < 0$ for all Ω' and $y_0 \in f(\Omega') \setminus f(\partial\Omega')$. Let us recall that each homeomorphism on a domain is either sense-preserving or sense-reversing, see e.g. [39, II.2.4., Theorem 3].

2.4. **Hausdorff dimension.** Let $\alpha > 0$. We define α-dimensional Hausdorff measure of a set $E \subset \mathbb{R}^n$ by

$$\mathcal{H}^\alpha(E) = \lim_{\varepsilon \to 0^+} \mathcal{H}^\alpha_\varepsilon(E),$$

where for a given $\varepsilon > 0$ we define

$$\mathcal{H}^\alpha_\varepsilon(E) = \inf\left\{ \sum_{i=1}^{\infty} (\operatorname{diam} A_i)^\alpha : E \subset \bigcup_{i=1}^{\infty} A_i,\ \operatorname{diam} A_i < \varepsilon \right\}.$$

We define the Hausdorff dimension of a set E as

$$\dim_{\mathcal{H}}(E) = \sup\{\alpha > 0 : \mathcal{H}^\alpha(E) = \infty\} = \inf\{\alpha > 0 : \mathcal{H}^\alpha(E) = 0\}.$$

We point out that Lipschitz mappings do not raise the Hausdorff dimension of a set and furthermore if $E = \bigcup_{i=1}^\infty E_i$ then

$$\dim_{\mathcal{H}}(E) = \sup_i \dim_{\mathcal{H}}(E_i).$$

2.5. **Construction of the Cantor set C_A and the set $\mathcal{K}_A$.** Denote by $\mathbb{V}$ the set of 2^4 vertices of the cube $[-1, 1]^4$. The sets

$$\mathbb{V}^k = \mathbb{V} \times \cdots \times \mathbb{V}, \quad k \in \mathbb{N},$$

will serve as the set of indices for our construction of Cantor sets.

We will define next the Cantor set C_A with positive measure for our construction. For this fix $\alpha > 0$. Let us define the sequence $\{a_k\}_{k=0}^\infty$ by setting

$$a_k = \frac{1}{2}\left(1 + \frac{1}{(k+1)^\alpha}\right).$$

Set $z_0 = 0$ and let us define

$$r_k = 2^{-k} a_k.$$

It follows that $Q(z_0, r_0) = (-1, 1)^4$ and further we proceed by induction. For $\boldsymbol{v}(k) = (v_1, \ldots, v_k) \in \mathbb{V}^k$ we denote $\boldsymbol{w}(k) = (v_1, \ldots, v_{k-1})$ and we define

$$z_{\boldsymbol{v}(k)} = z_{\boldsymbol{w}(k)} + \frac{1}{2} r_{k-1} v_k = z_0 + \frac{1}{2} \sum_{j=1}^{k} r_{j-1} v_j,$$

$$Q'_{\boldsymbol{v}(k)} = Q\big(z_{\boldsymbol{v}(k)}, 2^{-k} a_{k-1}\big) \text{ and } Q_{\boldsymbol{v}(k)} = Q\big(z_{\boldsymbol{v}(k)}, 2^{-k} a_k\big).$$

Formally we should write $\boldsymbol{w}(\boldsymbol{v}(k))$ instead of $\boldsymbol{w}(k)$ but for the simplification of the notation we will avoid this. Sometimes we may even denote $\boldsymbol{v}$ and $\boldsymbol{w}$ instead of $\boldsymbol{v}(k)$ and $\boldsymbol{w}(k)$.

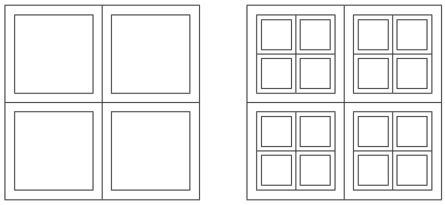


FIGURE 1. Two-dimensional projection of the cubes $Q_{\boldsymbol{v}(k)}$ and $Q'_{\boldsymbol{v}(k)}$ for $k = 1, 2$.

Then for the measure of the k-th frame $\mathbb{A}_{\boldsymbol{v}(k)} := Q'_{\boldsymbol{v}(k)} \setminus Q_{\boldsymbol{v}(k)}$, $k \in \mathbb{N}$, we have

$$\mathcal{L}^4(\mathbb{A}_{\boldsymbol{v}(k)}) = 2^{-4k+4}(a_{k-1}^4 - a_k^4) = 2^{-4k}\left[\left(1+\frac{1}{k^\alpha}\right)^4 - \left(1+\frac{1}{(k+1)^\alpha}\right)^4\right]. \tag{2.1}$$

The number of the cubes in $\{Q_{\boldsymbol{v}(k)} : \boldsymbol{v}(k) \in \mathbb{V}^k\}$ is 2^{4k}. It is not difficult to find out that the resulting Cantor set

$$\bigcap_{k=1}^{\infty} \bigcup_{\boldsymbol{v}(k)\in\mathbb{V}^k} Q_{\boldsymbol{v}(k)} =: C_A[\{a_k\}_{k=0}^{\infty}] = \mathcal{C}_A \times \mathcal{C}_A \times \mathcal{C}_A \times \mathcal{C}_A$$

is a product of 4 Cantor sets $\mathcal{C}_A$ in $\mathbb{R}$. Moreover, the measure of the set C_A can be calculated as

$$\mathcal{L}^4(C_A) = \lim_{k\to\infty} 2^{4k}(2a_k 2^{-k})^4 = \lim_{k\to\infty}\left(1+\frac{1}{(k+1)^\alpha}\right)^4 = 1. \tag{2.2}$$

Furthermore, we may write the 1-dimensional Cantor set $\mathcal{C}_A$ as

$$\mathcal{C}_A = \bigcap_{k=1}^{\infty}\bigcup_{i=1}^{2^k} I_{i,k}$$

where $I_{i,k}$ are closed intervals of length $2^{-k}\big(1+\frac{1}{(k+1)^\alpha}\big)$, $I_{i,k} \cap I_{j,k} = \emptyset$ for $i \neq j$, and $I_{2i-1,k} \cup I_{2i,k} \subset I_{i,k-1}$. Throughout this paper we will also denote

$$U_k := \bigcup_{i=1}^{2^k} I_{i,k}, \quad \mathcal{M}_k := U_k \times U_k \times U_k \times U_k, \quad P_k := U_k \times U_k \times U_k,$$

and in view (2.2) it is easy to see that

$$\mathcal{H}^1\big(U_k \setminus \mathcal{C}_A\big) \leq 2^k 2^{-k}\Big(1+\frac{1}{(k+1)^\alpha}\Big) - 1 \leq \frac{C}{k^\alpha}. \tag{2.3}$$

Further we denote

$$\begin{aligned} \mathcal{A}_k :=& \big(U_k \times U_k \times U_k \times \mathbb{R}\big) \cup \big(U_k \times U_k \times \mathbb{R} \times U_k\big) \\ &\cup \big(U_k \times \mathbb{R} \times U_k \times U_k\big) \cup \big(\mathbb{R} \times U_k \times U_k \times U_k\big). \end{aligned}$$

It is easy to see that

$$C_A = \mathcal{C}_A \times \mathcal{C}_A \times \mathcal{C}_A \times \mathcal{C}_A = \bigcap_{k=1}^{\infty} \mathcal{M}_k.$$

Furthermore, we also denote

$$\begin{aligned}\mathcal{K}_A :=&(\mathcal{C}_A \times \mathcal{C}_A \times \mathcal{C}_A \times [-1,1]) \cup (\mathcal{C}_A \times \mathcal{C}_A \times [-1,1] \times \mathcal{C}_A)\\ &\cup (\mathcal{C}_A \times [-1,1] \times \mathcal{C}_A \times \mathcal{C}_A) \cup ([-1,1] \times \mathcal{C}_A \times \mathcal{C}_A \times \mathcal{C}_A),\end{aligned}$$

and then we have

$$\mathcal{K}_A = [-1,1]^4 \cap \bigcap_{k=1}^{\infty} \mathcal{A}_k.$$

It is easy to see that $\mathcal{L}^4(\mathcal{K}_A) > 0$. Analogously to (2.3) we can estimate

$$\mathcal{H}^3(P_k \setminus P_{k+1}) \leq \Big(2^k 2^{-k}\big(1 + \frac{1}{(k+1)^\alpha}\big)\Big)^3 - \Big(2^{k+1} 2^{-(k+1)}\big(1 + \frac{1}{(k+2)^\alpha}\big)\Big)^3 \leq \frac{C}{k^{\alpha+1}}. \tag{2.4}$$

2.6. **Construction of the Cantor set C_B and the set $\mathcal{K}_B$.** Next, we will define the Cantor set C_B of zero measure for our construction. The definition of the index set $\mathbb{V}^k$ remains the same as in the subsection 2.5.

To define C_B we fix $0 < \delta < 1/7$. Let us define the sequence $\{b_k\}_{k=0}^{\infty}$ by setting

$$b_k = 2^{-k\beta},$$

where $\beta = \frac{1-\delta}{\delta}$. Analogously to the previous section we set $\hat{z}_0 = 0$ and define

$$\hat{r}_k = 2^{-k} b_k.$$

Then it follows that $Q(\hat{z}_0, \hat{r}_0) = (-1,1)^4$ and further we proceed by induction. For $\boldsymbol{v}(k) = (v_1, \dots, v_k) \in \mathbb{V}^k$ we denote $\boldsymbol{w}(k) = (v_1, \dots, v_{k-1})$ and we define

$$\hat{z}_{\boldsymbol{v}(k)} = \hat{z}_{\boldsymbol{w}(k)} + \frac{1}{2}\hat{r}_{k-1} v_k = \hat{z}_0 + \frac{1}{2}\sum_{j=1}^{k} \hat{r}_{j-1} v_j,$$

$$\hat{Q}'_{\boldsymbol{v}(k)} = Q\big(\hat{z}_{\boldsymbol{v}(k)}, 2^{-k} b_{k-1}\big) \ \text{ and } \ \hat{Q}_{\boldsymbol{v}(k)} = Q\big(\hat{z}_{\boldsymbol{v}(k)}, 2^{-k} b_k\big).$$

Index $\boldsymbol{w}(k) = (v_1, \dots, v_{k-1})$ is called as the *parent* of the index $\boldsymbol{v}(k) = (v_1, \dots, v_k)$. For the measure of the k-th frame $\mathbb{B}_{\boldsymbol{v}(k)} := \hat{Q}'_{\boldsymbol{v}(k)} \setminus \hat{Q}_{\boldsymbol{v}(k)}$, $k \in \mathbb{N}$, we have

$$\mathcal{L}^4(\mathbb{B}_{\boldsymbol{v}(k)}) = 2^{-4k+4}(b_{k-1}^4 - b_k^4) = 2^{-4k+4-4\beta k}(2^{4\beta} - 1). \tag{2.5}$$

Analogously to the previous section, it is not difficult to find out that the resulting Cantor set

$$\bigcap_{k=1}^{\infty} \bigcup_{\boldsymbol{v}(k) \in \mathbb{V}^k} \hat{Q}_{\boldsymbol{v}(k)} =: C_B[\{b_k\}_{k=0}^{\infty}] = \mathcal{C}_B \times \mathcal{C}_B \times \mathcal{C}_B \times \mathcal{C}_B$$

is a product of n Cantor sets $\mathcal{C}_B$ in $\mathbb{R}$. Moreover, the measure of the set C_B can be calculated as

$$\mathcal{L}^4(C_B) = \lim_{k\to\infty} 2^{4k}(2 b_k 2^{-k})^n = \lim_{k\to\infty} 2^{4-4\beta k} = 0. \tag{2.6}$$

Furthermore, we may write the 1-dimensional Cantor set $\mathcal{C}_B$ as

$$\mathcal{C}_B = \bigcap_{k=1}^{\infty} \bigcup_{i=1}^{2^k} \hat{I}_{i,k}$$

where $\hat{I}_{i,k}$ are closed intervals of length $2b_k 2^{-k}$, $\hat{I}_{i,k} \cap \hat{I}_{j,k} = \emptyset$ for $i \neq j$, and $\hat{I}_{2i-1,k} \cup \hat{I}_{2i,k} \subset \hat{I}_{i,k-1}$. Throughout this paper we denote

$$\hat{U}_k := \bigcup_{i=1}^{2^k} \hat{I}_{i,k}\,, \quad \hat{\mathcal{M}}_k := \hat{U}_k \times \hat{U}_k \times \hat{U}_k \times \hat{U}_k\,, \quad \hat{P}_k = \hat{U}_k \times \hat{U}_k \times \hat{U}_k. \tag{2.7}$$

Furthermore, we also denote

$$\begin{aligned} \hat{\mathcal{A}}_k :=& \big(\hat{U}_k \times \hat{U}_k \times \hat{U}_k \times \mathbb{R}\big) \cup \big(\hat{U}_k \times \hat{U}_k \times \mathbb{R} \times \hat{U}_k\big) \\ & \cup \big(\hat{U}_k \times \mathbb{R} \times \hat{U}_k \times \hat{U}_k\big) \cup \big(\mathbb{R} \times \hat{U}_k \times \hat{U}_k \times \hat{U}_k\big), \end{aligned} \tag{2.8}$$

and

$$\begin{aligned} \mathcal{K}_B :=& (\mathcal{C}_B \times \mathcal{C}_B \times \mathcal{C}_B \times [-1,1]) \cup (\mathcal{C}_B \times \mathcal{C}_B \times [-1,1] \times \mathcal{C}_B) \\ & \cup (\mathcal{C}_B \times [-1,1] \times \mathcal{C}_B \times \mathcal{C}_B) \cup ([-1,1] \times \mathcal{C}_B \times \mathcal{C}_B \times \mathcal{C}_B). \end{aligned}$$

It is easy to see that $\mathcal{L}^4(\mathcal{K}_B) = 0$. Furthermore, we may find out that $\dim_{\mathcal{H}} \mathcal{C}_B = \delta$ as in the k-th step of construction we have 2^k intervals of length $2b_k 2^{-k} = 2 \cdot 2^{-k-k\beta} = 2 \cdot 2^{-\frac{k}{\delta}}$. Therefore, as $0 < \delta < 1/7$, we conclude that

$$\dim_{\mathcal{H}} \mathcal{K}_B \leq 1 + 3\delta < \frac{3}{2}.$$

2.7. **The mapping S_q.** Suppose that C_A and C_B are the Cantor sets in subsections 2.5 and 2.6. Let $q : \mathbb{R} \to \mathbb{R}$ be the natural piecewise linear homeomorphism which takes each interval in the set $U_k \backslash U_{k+1}$, $k \in \mathbb{N}$, onto corresponding interval in $\hat{U}_k \backslash \hat{U}_{k+1}$ linearly. Then it is easy to see that q is an odd function, i.e. $q(-s) = -q(s)$ for every $s \in \mathbb{R}$. We define the homeomorphism $S_q : (-1,1)^n \to (-1,1)^n$ by setting

$$S_q(x_1, \dots, x_n) = (q(x_1), \dots, q(x_n)).$$

It is easy to see that S_q maps $\mathcal{K}_A$ onto $\mathcal{K}_B$. Moreover, we may notice that S_q is a Lipschitz mapping which takes each line segment parallel to x_i-axis to a line segment parallel to x_i-axis for every $i = 1, 2, 3, 4$. Furthermore, we also have that:

(1) For each $x \in (-1,1)^4$ such that $x_i \in U_k \setminus U_{k+1}$, $i = 1, 2, 3, 4$, we have

$$|D_i S_q(x)| = \frac{b_k - b_{k+1}}{a_k - a_{k+1}} \leq C k^{\alpha+1} 2^{-\beta k}, \tag{2.9}$$

where the constant $C = C(\alpha, \beta) > 0$ depends only on parameters α and β.

(2) For each $x \in (-1,1)^4$ such that $x_i \in \mathcal{C}_A$, $i = 1, 2, 3, 4$, we have

$$|D_i S_q(x)| = 0.$$

Here and in what follows $D_i g$ denotes the derivative of a mapping g along the x_i-direction for $i \in \{1, 2, 3, 4\}$.

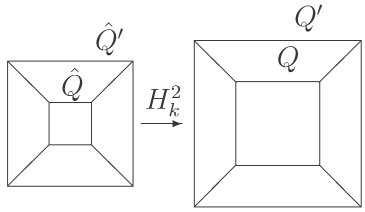


FIGURE 2. The transformation of $\hat{Q}' \setminus \hat{Q}^\circ$ onto $Q' \setminus Q^\circ$ in two dimensions.

2.8. **Frames to frames mapping of $(n-1)$-dimensional Cantor sets.** Suppose that $n \geq 3$. Analogously to the constructions of C_A and C_B we can define the $(n-1)$-dimensional Cantor type sets

$$\underbrace{\mathcal{C}_B \times \cdots \times \mathcal{C}_B}_{n-1 \text{ times}} \text{ and } \underbrace{\mathcal{C}_A \times \cdots \times \mathcal{C}_A}_{n-1 \text{ times}}.$$

We will need to find a mapping which maps the first set onto the second and the corresponding frames around it to corresponding frames around the second set. Instead of the index set $\mathbb{V}^k$ we use now the set $\mathbb{W}^k$ where $\mathbb{W}$ denotes the vertices of the cube $[-1,1]^{n-1}$. Analogously to previous notation we denote $\boldsymbol{w} \in \mathbb{W}^k$ instead of $\boldsymbol{v} \in \mathbb{V}^k$ and we work with cubes

$$Q'_{\boldsymbol{w}(k)}, Q_{\boldsymbol{w}(k)}, \hat{Q}'_{\boldsymbol{w}(k)} \text{ and } \hat{Q}_{\boldsymbol{w}(k)}$$

defined analogously to subsections 2.3 and 2.4 but now in $n-1$ dimensions.

We will find a sequence of homeomorphisms $H_k^{n-1} : (-1,1)^{n-1} \to (-1,1)^{n-1}$. We set $H_0^{n-1}(x) = x$ and we proceed by induction. We will give a mapping F_1 which stretches each cube $\hat{Q}_{\boldsymbol{w}}$, $\boldsymbol{w} \in \mathbb{W}^1$, homogeneously so that $H_1^{n-1}(\hat{Q}_{\boldsymbol{w}})$ equals $Q_{\boldsymbol{w}}$. On the annulus $\hat{Q}'_{\boldsymbol{w}} \setminus \hat{Q}_{\boldsymbol{w}}$, H_1^{n-1} is defined to be an appropriate radial map with respect to $\hat{z}_{\boldsymbol{w}}$ and $z_{\boldsymbol{w}}$ in the image in order to make H_1^{n-1} a homeomorphism. The general step is the following: If $k > 1$, H_k^{n-1} is defined as H_{k-1}^{n-1} outside the union of all cubes $\hat{Q}'_{\boldsymbol{w}}$, $\boldsymbol{w} \in \mathbb{W}^k$. Further, H_k^{n-1} remains equal to H_{k-1}^{n-1} at the centers of cubes $\hat{Q}_{\boldsymbol{w}}$, $\boldsymbol{w} \in \mathbb{W}^k$. Then H_k^{n-1} stretches each cube $\hat{Q}_{\boldsymbol{w}}$, $\boldsymbol{w} \in \mathbb{W}^k$, homogeneously so that $H_k^{n-1}(\hat{Q}_{\boldsymbol{w}})$ equals $Q_{\boldsymbol{w}}$. On the annulus $\hat{Q}'_{\boldsymbol{w}} \setminus \hat{Q}_{\boldsymbol{w}}$, H_k^{n-1} is defined to be an appropriate radial map with respect to $\hat{z}_{\boldsymbol{w}}$ in preimage and $z_{\boldsymbol{w}}$ in image to make H_k^{n-1} a homeomorphism (see Fig. 2). Notice that the Jacobian determinant $J_{H_k^{n-1}}(x)$ will be strictly positive almost everywhere in $(-1,1)^{n-1}$.

In the following definition of H_k^{n-1} we use the notation $\|x\|$ for the supremum norm of $x \in \mathbb{R}^{n-1}$. The mappings H_k^{n-1}, $k \in \mathbb{N}$, are formally defined as

(2.10)
$$H_k^{n-1}(x) = \begin{cases} H_{k-1}^{n-1}(x) & \text{for } x \notin \bigcup_{\boldsymbol{w} \in \mathbb{W}^k} \hat{Q}'_{\boldsymbol{w}} \\ H_{k-1}^{n-1}(\hat{z}_{\boldsymbol{w}}) + (\alpha_k \|x - \hat{z}_{\boldsymbol{w}}\| + \beta_k) \frac{x - \hat{z}_{\boldsymbol{w}}}{\|x - \hat{z}_{\boldsymbol{w}}\|} & \text{for } x \in \hat{Q}'_{\boldsymbol{w}} \setminus \hat{Q}_{\boldsymbol{w}},\ \boldsymbol{w} \in \mathbb{W}^k \\ H_{k-1}^{n-1}(\hat{z}_{\boldsymbol{w}}) + \frac{r_k}{\tilde{r}_k}(x - \hat{z}_{\boldsymbol{w}}) & \text{for } x \in \hat{Q}_{\boldsymbol{w}},\ \boldsymbol{w} \in \mathbb{W}^k \end{cases}$$

where the constants α_k and β_k are given by

$$\alpha_k \hat{r}_k + \beta_k = r_k \text{ and } \alpha_k \tfrac{\hat{r}_{k-1}}{2} + \beta_k = \tfrac{r_{k-1}}{2}. \tag{2.11}$$

It is not difficult to find out that each H_k^{n-1} is a homeomorphism and maps

$$\bigcup_{\boldsymbol{w}\in\mathbb{W}^k} \hat{Q}_{\boldsymbol{w}} \text{ onto } \bigcup_{\boldsymbol{w}\in\mathbb{W}^k} Q_{\boldsymbol{w}}.$$

The limit $H^{n-1}(x) = \lim_{k\to\infty} H_k^{n-1}(x)$ is clearly one-to-one and continuous and therefore a homeomorphism. Moreover, it is easy to see that H^{n-1} is differentiable almost everywhere (as $\mathcal{L}^{n-1}(\mathcal{C}_B^{n-1}) = 0$) and maps $\mathcal{C}_B^{n-1}$ onto $\mathcal{C}_A^{n-1}$.

Fix $j \in \mathbb{N}$. We claim that the mapping H^{n-1} is Lipschitz on $(\hat{U}_j)^{n-1} \setminus (\hat{U}_{j+1})^{n-1}$ where the sets $\hat{U}_j$ are defined analogously to subsection 2.6 (the Lipschitz constant of course depends on the fixed j). This is in fact easy to see as the mapping is given by simple formula (2.10) on each $\hat{Q}'_{\boldsymbol{w}} \setminus \hat{Q}_{\boldsymbol{w}}$ for every $\boldsymbol{w} \in \mathbb{W}^j$. Analogously to [23, Lemma 2.1 and proof of Theorem 4.10] we can estimate

$$|DH_j^{n-1}(x)| = |DH^{n-1}(x)| \sim \max\Big\{\frac{r_j}{\hat{r}_j}, \alpha_j\Big\} \leq C\max\{2^{\beta j}, 2^{\beta j} j^{-(\alpha+1)}\} \leq C2^{\beta j} \tag{2.12}$$

for every $x \in \hat{Q}'_{\boldsymbol{w}} \setminus \hat{Q}_{\boldsymbol{w}}$ and $\boldsymbol{w} \in \mathbb{W}^j$. This is because

$$\begin{aligned} |D_l H_j^{n-1}(x)| &\leq \frac{r_j}{\hat{r}_j} \leq C2^{\beta j} \text{ for } l \neq i \text{ and} \\ |D_i H_j^{n-1}(x)| &\leq \alpha_j \leq C2^{\beta j} j^{-(\alpha+1)} \end{aligned} \tag{2.13}$$

if x_i is the direction which realizes the supremum norm distance from the center of the cube $\hat{z}_{\boldsymbol{w}}$. From (2.10) it is also easy to see that

$$|DH_j^{n-1}(x)| \sim \frac{r_j}{\hat{r}_j} \leq C2^{\beta j} \text{ for } x \in \hat{Q}_{\boldsymbol{w}} \text{ and } \boldsymbol{w} \in \mathbb{W}^j. \tag{2.14}$$

In our construction we will need to know that for each $\alpha \in (0,1)$ and $k \in \mathbb{N}$ the mapping

$$\alpha H_{3k-3}^{n-1}(x) + (1-\alpha) H_{3k}^{n-1}(x) \tag{2.15}$$

is a homeomorphisms. Outside of $\bigcup_{\boldsymbol{w}\in\mathbb{W}^{3k-3}} \hat{Q}_{\boldsymbol{w}}$ both mapping are equal and hence the mapping is a homeomorphism there. Let us fix $\hat{Q}_{\boldsymbol{w}}$ for some $\boldsymbol{w} \in \mathbb{W}^{3k-3}$. We know by (2.10) that H_{3k}^{n-1} is a frame to frame mapping on $\hat{Q}_{\boldsymbol{w}}$ which maps corresponding squares with sizes $\tilde{r}_{3k-2}$ (resp. $\tilde{r}_{3k-1}$ and $\tilde{r}_{3k}$) to squares with sizes r_{3k-2} (resp. r_{3k-1} and r_{3k}). We also know by (2.10) that H_{3k-3}^{n-1} is a linear mapping

$$\frac{r_{3k-3}}{\tilde{r}_{3k-3}}(x - \tilde{z}_{\boldsymbol{w}}) \text{ on } \hat{Q}_{\boldsymbol{w}}$$

but this can be also viewed as a frame to frame mapping on $\hat{Q}_{\boldsymbol{w}}$ which maps corresponding squares with sizes $\tilde{r}_{3k-2}$ (resp. $\tilde{r}_{3k-1}$ and $\tilde{r}_{3k}$) to squares with sizes $\frac{r_{3k-3}}{\tilde{r}_{3k-3}}\tilde{r}_{3k-2}$ (resp. $\frac{r_{3k-3}}{\tilde{r}_{3k-3}}\tilde{r}_{3k-1}$ and $\frac{r_{3k-3}}{\tilde{r}_{3k-3}}\tilde{r}_{3k}$). Thus it is not difficult to see that the mapping (2.15)

on $\hat{Q}_{\boldsymbol{w}}$ is a frame to frame mapping which maps corresponding squares with sizes $\tilde{r}_{3k-2}$ (resp. $\tilde{r}_{3k-1}$ and $\tilde{r}_{3k}$) to squares with sizes
$\alpha r_{3k-2}+(1-\alpha)\frac{r_{3k-3}}{\tilde{r}_{3k-3}}\tilde{r}_{3k-2}$ (resp. $\alpha r_{3k-1}+(1-\alpha)\frac{r_{3k-3}}{\tilde{r}_{3k-3}}\tilde{r}_{3k-1}$ and $\alpha r_{3k}+(1-\alpha)\frac{r_{3k-3}}{\tilde{r}_{3k-3}}\tilde{r}_{3k}$).
Analogously to the fact that each H_k defined by (2.10) is a homeomorphism we can conclude that the mapping (2.15) given by formula analogous to (2.10) is also a homeomorphism.

3. A sense-preserving bi-Lipschitz mapping F equal to a reflection in the last variable on $\mathcal{K}_B$

This section is dedicated to constructing a bi-Lipschitz mapping which equals the reflection in the last variable on $\mathcal{K}_B$. Especially, this means that the mapping will map lines in $\mathcal{K}_B$ to lines in $\mathcal{K}_B$. In fact even more than this the mapping will map certain line segments close to $\mathcal{K}_B$ to line segments (recall that $\mathcal{K}_B$ and $\hat{U}_k$ are defined in subsection 2.5). Also see Fig. 3.

Theorem 3.1. *If $\beta > 0$ is sufficiently large in the definition of the Cantor set C_B in subsection 2.6 then there exists a mapping $F : (-1,1)^4 \to (-1,1)^4$, which is a sense-preserving bi-Lipschitz extension of the map*

$$F(x_1, x_2, x_3, x_4) = (x_1, x_2, x_3, -x_4) \quad x \in \mathcal{K}_B, \tag{3.1}$$

and a constant $N_F \in \mathbb{N}$ such that for each $j, k \in \mathbb{N}$ satisfying $N_F < j \le k$ the image of the intersection of a line parallel to e_i with the set

$$A_{i,j-N_F-1,k+N_F} := \{x \in \mathbb{R}^4 : x_i \in [-1,1] \setminus \hat{U}_{j-N_F-1}, x_l \in \hat{U}_{k+N_F}, l \neq i\}$$

in the map F is a line segment parallel to e_i which lies in the set

$$A_{i,j-1,k} = \{x \in \mathbb{R}^4 : x_i \in [-1,1] \setminus \hat{U}_{j-1}, x_l \in \hat{U}_k, l \neq i\}.$$

Moreover, the derivative along this segment satisfies

$$D_iF(x) = \begin{cases} e_i & \text{if } i = 1,2,3 \\ -e_i & \text{if } i = 4 \end{cases}$$

for every $x \in A_{i,j-N_F-1,k+N_F}$.

The concept of the following type of mapping is key to our proof. We will show the obvious fact that they are bi-Lipschitz maps.

Definition 3.2. *Let $n \in \mathbb{N}$, $n \geq 2$, and let $v \in \mathbb{R}^n$ be a vector such that $v_n \neq 0$. Denote $X := \mathbb{R}^{n-1} \times \{0\}$. Let $g : X \to \mathbb{R}$ be a Lipschitz function and define a projection P_v of $\mathbb{R}^n$ onto X as follows*

$$P_v(x) = x - \frac{x_n}{v_n}v. \tag{3.2}$$

Then we define the spaghetti strand map $F_{g,v}$ as follows

$$F_{g,v}(x) = x + vg(P_v(x)).$$

Lemma 3.3. *Spaghetti strand maps from Definition 3.2 are bi-Lipschitz maps.*

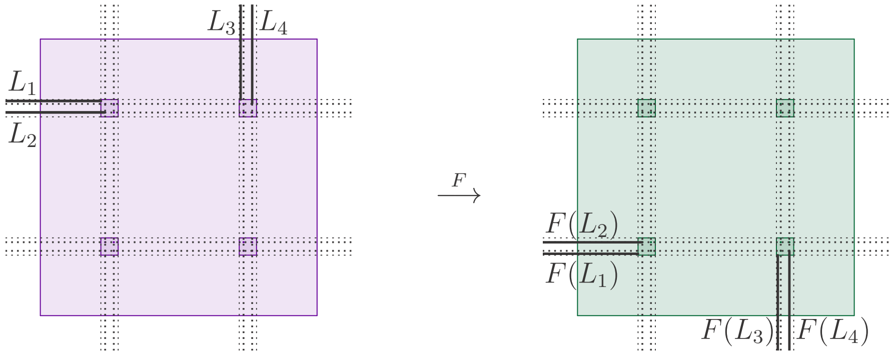


FIGURE 3. A sense preserving bi-Lipschitz map that reflects in e_4 and maps certain lines to lines

Proof. It is easy to see that every spaghetti strand map is Lipschitz as a composition of Lipschitz maps. Moreover $P_v(\alpha v) = 0$ for each $\alpha \in \mathbb{R}$ which implies that

$$x + vg(P_v(x)) - vg\big(P_v(x + vg(P_v(x)))\big) = x + vg(P_v(x)) - vg(P_v(x)) = x$$

and hence the inverse of a spaghetti strand map is the spaghetti strand map corresponding to $-g$. This inverse is also Lipschitz and therefore we see that these maps are bi-Lipschitz. □

Firstly, let us outline our strategy for the rest of the section. We construct F from the composition of two spaghetti strand maps. Firstly we must choose a vector v and prove that the projection P_v is one-to-one on the set $\mathcal{K}_B$ and further there exists a Lipschitz function g so that $F_{g,v}(x) = P_v(x)$ for all $x \in \mathcal{K}_B$. This step is contained in Lemma 3.4. If we take $u = (-v_1, -v_2, \dots, -v_{n-1}, v_n)$ then we can define $F = F_{g,u} \circ F_{g,v}$ and it is not difficult to deduce that (3.1) holds (this is done in (3.37) below).

Lemma 3.4. *Let $v = \left(\frac{1}{16}, \frac{1}{8}, \frac{1}{4}, 1\right)$, $u = \left(-\frac{1}{16}, -\frac{1}{8}, -\frac{1}{4}, 1\right)$. Then there is $\beta \geq 6$ and a corresponding set $\mathcal{K}_B$ given by the subsection 2.6 such that P_v is one-to-one on $\mathcal{K}_B$, and the function g defined on $P_v(\mathcal{K}_B)$ as $g(P_v(x)) = -x_4$ can be extended onto X as a Lipschitz function. Furthermore, it is possible to find a Lipschitz extension of the function g which guarantees that*

$$D_i\big(F_{g,u} \circ F_{g,v}\big)(x) = \begin{cases} e_i & \text{if } i = 1,2,3 \\ -e_i & \text{if } i = 4 \end{cases} \tag{3.3}$$

whenever $k \in \mathbb{N}$, $x_i \in [-1,1] \setminus \hat{U}_k$ and $x_j \in \hat{U}_{k+2}$ for all $j \neq i$.

Proof. Let us start by defining some notation we will use throughout the proof. We will denote $\tilde{v} := \left(\frac{1}{16}, \frac{1}{8}, \frac{1}{4}\right)$. Furthermore, if $\hat{Q}_{\boldsymbol{v}(k)} := Q(\hat{z}_{\boldsymbol{v}(k)}, \hat{r}_k)$, $\boldsymbol{v}(k) \in \mathbb{V}^k$, are the cubes used in the definition of the Cantor set C_B in subsection 2.6, then we define

$$\hat{G}^i_{\boldsymbol{v}(k)} := \hat{Q}_{\boldsymbol{v}(k)} + \mathbb{R}e_i. \tag{3.4}$$

These sets are called *k-bars*.

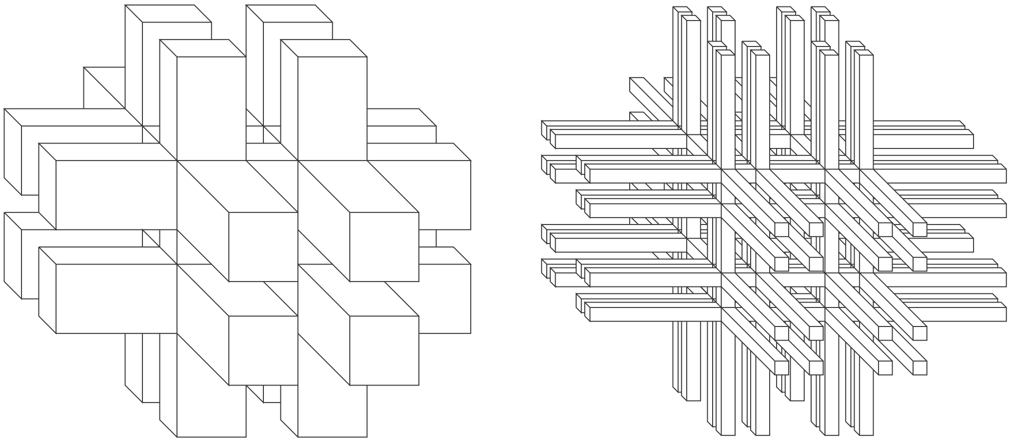


FIGURE 4. All 1-bars and 2-bars in three dimensions.

By the construction of the Cantor set we have $Q(\hat{z}_{\boldsymbol{v}(k-2)}, \hat{r}_{k-2}) \cap Q(\hat{z}_{\hat{\boldsymbol{v}}(k-2)}, \hat{r}_{k-2}) = \emptyset$, whenever $\boldsymbol{v}(k) \neq \hat{\boldsymbol{v}}(k)$. Therefore we have the equality for the so-called "sliced" bar

$$\hat{S}^i_{\boldsymbol{v}(k)} := \hat{G}^i_{\boldsymbol{v}(k)} \setminus \Big(\bigcup_{\boldsymbol{w} \in \mathbb{V}^{k-2}_{\boldsymbol{v}(k)}(i)} Q\big(\hat{z}_{\boldsymbol{w}}, \hat{r}_{k-2}\big)\Big) = \hat{G}^i_{\boldsymbol{v}(k)} \setminus \Big(\bigcup_{\boldsymbol{w} \in \mathbb{V}^{k-2}} Q\big(\hat{z}_{\boldsymbol{w}}, \hat{r}_{k-2}\big)\Big), \tag{3.5}$$

where

$$\mathbb{V}^{k-2}_{\boldsymbol{v}(k)}(i) := \big\{\boldsymbol{w} \in \mathbb{V}^{k-2} : (\hat{z}_{\boldsymbol{v}(k)} + \mathbb{R}e_i) \cap Q(\hat{z}_{\boldsymbol{w}}, \hat{r}_{k-2}) \neq \emptyset\big\}.$$

It is easy to see that there is $\beta_1 > 0$ such that for $\beta \geq \beta_1$ (in the definition of C_B) we can replace the index set $\mathbb{V}^{k-2}_{\boldsymbol{v}(k)}(i)$ by much nicer set $\mathbb{V}^{k-2}$ in the definition of $\hat{S}^i_{\boldsymbol{v}(k)}$.

More precisely, a sliced k-bar $\hat{S}^i_{\boldsymbol{v}(k)}$ can be considered as a k-bar where we have removed all the cubes around the Cantor set from the $(k-2)$-nd generation of the construction.

In similar fashion we also define

$$\hat{\mathcal{S}}^i_{\boldsymbol{v}(k)} := \Big(Q^3\big(P_v(\hat{z}_{\boldsymbol{v}(k)}), q\hat{r}_k\big) + P_v(\mathbb{R}e_i)\Big) \setminus \Big(\bigcup_{\boldsymbol{w} \in \mathbb{V}^{k-1}_{\boldsymbol{v}(k)}(i)} Q^3\big(P_v(\hat{z}_{\boldsymbol{w}}), \tfrac{7}{8}\hat{r}_{k-1}\big)\Big) \subset X, \tag{3.6}$$

where $Q^3(z, r)$ denotes the 3 dimensional cube in $X := \mathbb{R}^3 \times \{0\}$ with radius $r > 0$ and centered at $z \in X$ and $q \geq \frac{5}{4}$ is a constant we will determine later. We also denote "sliced k-pipes" as follows

$$\hat{H}^i_{\boldsymbol{v}(k)} := \partial_X \Big(Q^3\big(P_v(\hat{z}_{\boldsymbol{v}(k)}), q\hat{r}_k\big) + P_v(\mathbb{R}e_i)\Big) \setminus \Big(\bigcup_{\boldsymbol{w} \in \mathbb{V}^{k-1}_{\boldsymbol{v}(k)}(i)} Q^3\big(P_v(\hat{z}_{\boldsymbol{w}}), \tfrac{7}{8}\hat{r}_{k-1}\big)\Big) \subset X,$$

where $\partial_X A$ denotes the relative boundary of a set A in X. We will later see that also in the definition of the sets $\hat{\mathcal{S}}^i_{\boldsymbol{v}(k)}$ and $\hat{H}^i_{\boldsymbol{v}(k)}$ the index sets $\mathbb{V}^{k-1}_{\boldsymbol{v}(k)}(i)$ can be replaced by $\mathbb{V}^{k-1}$ when $\beta > 0$ in the definition of C_B is just large enough.

Now let us briefly outline the rest of the proof. We prove that our choice of a vector v gives that P_v is one-to-one on $\mathcal{K}_B$. Then we prove that each $\hat{S}^i_{\boldsymbol{v}(k)}$ is projected into

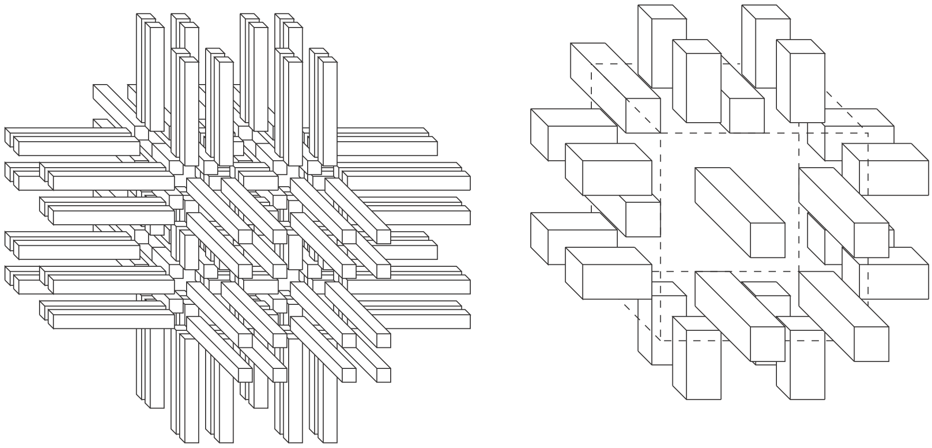


FIGURE 5. In the picture on the left we have all sliced 2-bars by 1-st generation cubes in three dimensions. In the later we slice k generational bars with $k-2$ generation cubes. A choice of β guarantees that in comparison the bars are as thin as required in comparison to the cube. In the picture on the right we have zoomed in one of the removed cubes (drawn with dashed line) from the picture on the left.

$\hat{\mathcal{S}}^i_{\boldsymbol{v}(k)}$ which, for fixed k, are pairwise disjoint. This allows us to define a Lipschitz function g on $\mathbb{R}^3 \times \{0\}$ such that $F_{g,v} = P_v$ on $\mathcal{K}_B$. A careful extension of g onto $\mathbb{R}^3 \times \{0\}$ guarantees (3.3). We divide the proof into several steps.

Step 1: The projection is one-to-one on C_B. Our first step is simple, we want to show that the projection is one-to-one on the set $C_B = \mathcal{C}_B \times \mathcal{C}_B \times \mathcal{C}_B \times \mathcal{C}_B$. Consider the first stage of our Cantor construction, i.e. we have the cube $\hat{Q}_0 = Q(0,1)$ and the set of cubes $\hat{Q}_{\boldsymbol{v}(1)} := Q(\hat{z}_{\boldsymbol{v}(1)}, \hat{r}_1)$, $\boldsymbol{v}(1) \in \mathbb{V}$. We will show that the images of these 2^4 cubes in P_v are pairwise disjoint. Then we can use the same calculations to show that the projections of the next generation of cubes in our construction are also pairwise disjoint because the construction is self-similar. We can repeat this argument inductively to get that P_v is one-to-one on C_B. Therefore it suffices to show that the images of $\hat{Q}_{\boldsymbol{v}(1)}$ are pairwise disjoint. Although this step is slightly redundant it aids the understanding of the reader and so we include it here.

We will deal with two separate cases. The first case is where we are considering the projections of a pair of boxes $\hat{Q}_{\boldsymbol{v}(1)}$ and $\hat{Q}_{\hat{\boldsymbol{v}}(1)}$, whose centers have the same 4-th coordinate. The second case is where $(\hat{z}_{\boldsymbol{v}(1)})_4 \neq (\hat{z}_{\hat{\boldsymbol{v}}(1)})_4$. For any of the first generation cubes $\hat{Q}_{\boldsymbol{v}(1)}$ we can calculate its image in P_v using (3.2) and $v = \big(\frac{1}{16}, \frac{1}{8}, \frac{1}{4}, 1\big)$ as

$$P_v(\hat{Q}_{\boldsymbol{v}(1)}) = Q^3\big(P_v(\hat{z}_{\boldsymbol{v}(1)}), \hat{r}_1\big) + (-\hat{r}_1, \hat{r}_1)\big(\tfrac{1}{16}, \tfrac{1}{8}, \tfrac{1}{4}\big) \subset Q^3\big(P_v(\hat{z}_{\boldsymbol{v}(1)}), \hat{r}_1(1+\tfrac{1}{4})\big), \tag{3.7}$$

where

$$P_v(\hat{z}_{\boldsymbol{v}(1)}) = \hat{z}_{\boldsymbol{v}(1)} - (\hat{z}_{\boldsymbol{v}(1)})_4(\tfrac{1}{16}, \tfrac{1}{8}, \tfrac{1}{4}, 1) = \big((\hat{z}_{\boldsymbol{v}(1)})_1, (\hat{z}_{\boldsymbol{v}(1)})_2, (\hat{z}_{\boldsymbol{v}(1)})_3, 0\big) \mp \tfrac{1}{2}(\tfrac{1}{16}, \tfrac{1}{8}, \tfrac{1}{4}, 0). \tag{3.8}$$

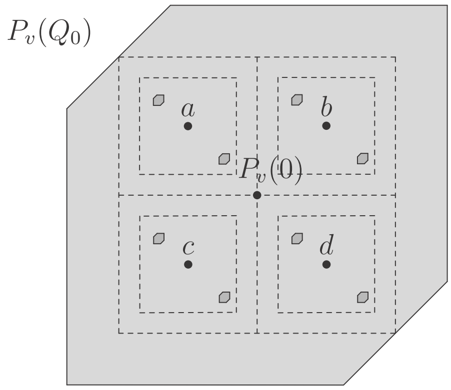


FIGURE 6. An illustration of the image of two generations of cubes in the projection P_v from three dimensions to the plane. For printing reasons we have now increased significantly $\hat{r}_1$ and changed somewhat v. The shaded regions (one big and eight smaller ones) describe the images of the cubes in P_v. The black dot in the middle describes the point $P_v(0) = 0$ the center of the large cube. The other four black dots a, b, c and d describe the centers of the small dashed cubes of radius $\frac{1}{4} + \frac{5}{4}\hat{r}_1$ which always contain the image of a pair of cubes symmetrical about the hyperplane, see Case 2B. In one case we consider a pair of cubes symmetrical about the hyperplane and the images of their centers are separated by $\tilde{v}$. In the other cases the images of cubes are disjoint because they lie in different dotted cubes, which are disjoint.

Case 1 (Step 1): In the first case we have a distinct pair of centers $\hat{z}_{\boldsymbol{v}(1)}$ and $\hat{z}_{\hat{\boldsymbol{v}}(1)}$ such that $(\hat{z}_{\boldsymbol{v}(1)})_4 = (\hat{z}_{\hat{\boldsymbol{v}}(1)})_4$. Since the pair is distinct we can find at least one $i \in \{1, 2, 3\}$ such that

$$|(\hat{z}_{\boldsymbol{v}(1)})_i - (\hat{z}_{\hat{\boldsymbol{v}}(1)})_i| = 1.$$

This means that $|\hat{z}_{\boldsymbol{v}(1)} - \hat{z}_{\hat{\boldsymbol{v}}(1)}| \geq 1$. But since $(\hat{z}_{\boldsymbol{v}(1)})_4 = (\hat{z}_{\hat{\boldsymbol{v}}(1)})_4$ we have

$$P_v(\hat{z}_{\boldsymbol{v}(1)}) - \hat{z}_{\boldsymbol{v}(1)} = P_v(\hat{z}_{\hat{\boldsymbol{v}}(1)}) - \hat{z}_{\hat{\boldsymbol{v}}(1)},$$

and therefore

$$|P_v(\hat{z}_{\boldsymbol{v}(1)}) - P_v(\hat{z}_{\hat{\boldsymbol{v}}(1)})| = |\hat{z}_{\boldsymbol{v}(1)} - \hat{z}_{\hat{\boldsymbol{v}}(1)}| \geq 1.$$

This together with (3.7) and the fact that $2\hat{r}_1(1 + \frac{1}{4}) < 1$ (recall that $\hat{r}_1 = 2^{-1}2^{-\beta}$ with $\beta \geq 6$) implies that

$$P_v(\hat{Q}_{\boldsymbol{v}(1)}) \cap P_v(\hat{Q}_{\hat{\boldsymbol{v}}(1)}) = \emptyset.$$

Case 2A (Step 1): Suppose now that $(\hat{z}_{\boldsymbol{v}(1)})_4 \neq (\hat{z}_{\hat{\boldsymbol{v}}(1)})_4$. We shall consider first a pair of boxes, whose centers $\hat{z}_{\boldsymbol{v}(1)}$ and $\hat{z}_{\hat{\boldsymbol{v}}(1)}$ are on a line parallel to e_4. To see that the images of these boxes are disjoint we observe that

$$P_v(\hat{z}_{\boldsymbol{v}(1)}) - P_v(\hat{z}_{\hat{\boldsymbol{v}}(1)}) = P_v(\hat{z}_{\boldsymbol{v}(1)} - \hat{z}_{\hat{\boldsymbol{v}}(1)}) = P_v(\pm e_4) = \mp\left(\tfrac{1}{16}, \tfrac{1}{8}, \tfrac{1}{4}\right). \tag{3.9}$$

Furthermore, as

$$P_v(\hat{Q}_{\boldsymbol{v}(1)}) \subset Q(P_v(\hat{z}_{\boldsymbol{v}(1)}), \tfrac{5}{4}\hat{r}_1), \tag{3.10}$$

and since $2\hat{r}_1\frac{5}{4} < \frac{1}{16} < \|P_v(\hat{z}_{\boldsymbol{v}(1)}) - P_v(\hat{z}_{\hat{\boldsymbol{v}}(1)})\|$ then the projection of $\hat{Q}_{\boldsymbol{v}(1)}$ and $\hat{Q}_{\hat{\boldsymbol{v}}(1)}$ must be disjoint. Here $\|x\| := \sup_i |x_i|$ denotes the supremum norm.

Case 2B (Step 1): We still need to consider the pairs of cubes with centers that vary from each other in the 4-th variable and in another variable. In other words, let us suppose that it holds for $\hat{z}_{\boldsymbol{v}(1)}$ and $\hat{z}_{\hat{\boldsymbol{v}}(1)}$ that

$$(\hat{z}_{\boldsymbol{v}(1)})_4 \neq (\hat{z}_{\hat{\boldsymbol{v}}(1)})_4 \quad \text{and} \quad (\hat{z}_{\boldsymbol{v}(1)})_i \neq (\hat{z}_{\hat{\boldsymbol{v}}(1)})_i \text{ for some } i \in \{1,2,3\},$$

and let us denote

$$a := ((\hat{z}_{\boldsymbol{v}(1)})_1, (\hat{z}_{\boldsymbol{v}(1)})_2, (\hat{z}_{\boldsymbol{v}(1)})_3, 0) \quad \text{and} \quad b = ((\hat{z}_{\hat{\boldsymbol{v}}(1)})_1, ((\hat{z}_{\hat{\boldsymbol{v}}(1)})_2, ((\hat{z}_{\hat{\boldsymbol{v}}(1)})_3, 0).$$

By applying (3.7) and (3.8) we get

$$P_v(Q_{\boldsymbol{v}(1)}) \subset Q\big(a, \tfrac{1}{4} + \tfrac{5}{4}\hat{r}_1\big) \quad \text{and} \quad P_v(Q_{\hat{\boldsymbol{v}}(1)}) \subset Q\big(b, \tfrac{1}{4} + \tfrac{5}{4}\hat{r}_1\big),$$

where $\frac{1}{4} + \frac{5}{4}\hat{r}_1 < \frac{9}{64} < \frac{1}{2}$. Thus, it follows from the fact $|a-b| \geq 1$ that

$$\begin{aligned} \operatorname{dist}(P_v(\hat{Q}_{\boldsymbol{v}(1)}), P_v(\hat{Q}_{\hat{\boldsymbol{v}}(1)})) &\geq \operatorname{dist}\big(Q\big(a, \tfrac{1}{4} + \tfrac{5}{4}\hat{r}_1\big), Q\big(b, \tfrac{1}{4} + \tfrac{5}{4}\hat{r}_1\big)\big) \\ &\geq |a-b| - 2\big(\tfrac{1}{4} + \tfrac{5}{4}\hat{r}_1\big) > 0, \end{aligned}$$

which gives us that the sets $P_v(\hat{Q}_{\boldsymbol{v}(1)})$ and $P_v(\hat{Q}_{\hat{\boldsymbol{v}}(1)})$ are disjoint. This implies that the remaining pairs of cubes to consider (i.e. the pairs of cubes with centers that vary from each other in the 4-th variable and in another variable) are also disjoint.

It follows now from Cases 1, 2A and 2B that images of the first generation cubes $\hat{Q}_{\boldsymbol{v}(1)}$ in the projection P_v are pairwise disjoint. The self similarity argument mentioned above implies that P_v is one-to-one on C_B. The reason why the self similarity argument works here is because the ratio $\hat{r}_{k-1}/\hat{r}_k = 2^{\beta+1}$ is not depending on k. Geometrically this means that if we rescale a cube $\hat{Q}_{\boldsymbol{v}(k-1)}$ and the smaller cubes $\hat{Q}_{\boldsymbol{v}(k)}$ which lies inside this cube by factor $2^{k-1}2^{\beta(k-1)}$ we see that there will be as much space to project the cubes of the k-th step as there was in the first step (see Fig. 7).

Step 2: The projection is one-to-one on $\mathcal{K}_B$. We will start this step by showing that if a and b are any two vertices of $Q(0, \frac{1}{2})$ and $e_i, e_j \in \mathbb{R}^4$ are two (possibly identical) canonical basis vectors of $\mathbb{R}^4$, then

$$P_v(a + \mathbb{R}e_i) \cap P_v(b + \mathbb{R}e_j) = P_v\big((a + \mathbb{R}e_i) \cap (b + \mathbb{R}e_j)\big). \tag{3.11}$$

This gives us that if ℓ and $\hat{\ell}$ are two distinct lines parallel to coordinate axes through some vertices a and b of $Q(0, \frac{1}{2})$ then their projections $P_v(\ell)$ and $P_v(\hat{\ell})$ meet at most at one point which is the image $P_v(z)$ of the intersection point z of ℓ and $\hat{\ell}$. We use this to show that images of sliced k-bars are disjoint and finally by this observation and by Step 1 we will conclude that P_v is one-to-one on $\mathcal{K}_B$. It is good to remark that the argument bellow does not work if the dimension of the space is three or smaller.

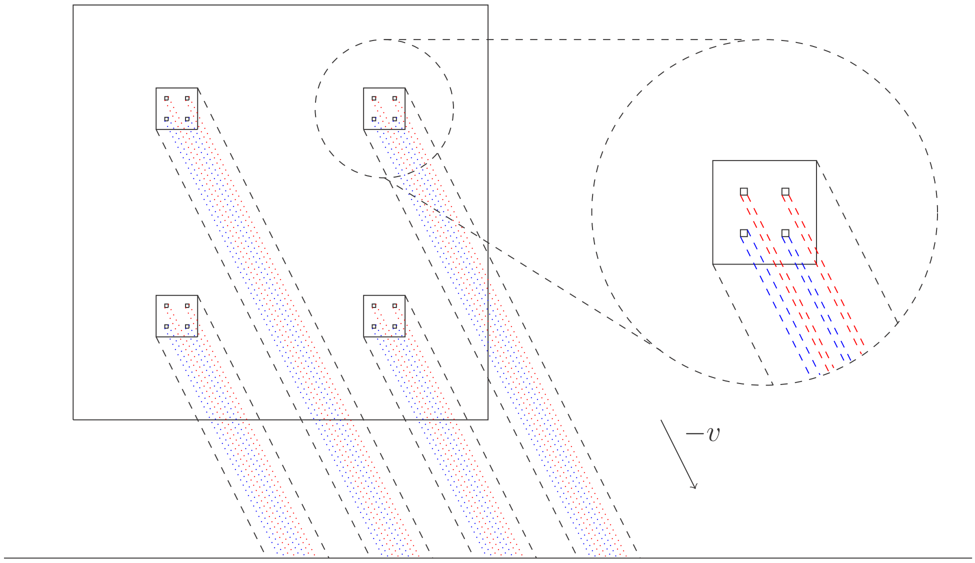


FIGURE 7. An illustration of the idea behind the self similarity argument.

Step 2A: Proving the equation (3.11). To prove (3.11) it suffices to show that

$$P_v(a+\mathbb{R}e_i)\cap P_v(b+\mathbb{R}e_j)\subset P_v\big((a+\mathbb{R}e_i)\cap(b+\mathbb{R}e_j)\big) \tag{3.12}$$

as the opposite inclusion is obvious. To prove (3.12) we recall the following elementary dimension formula for the linear map $P_v:\mathbb{R}^4\to X$:

$$\dim\big(\ker P_v\big)+\dim\big(\operatorname{im} P_v\big)=4,$$

where $\ker P_v$ stands for the kernel of the linear map P_v, and $\operatorname{im} P_v$ equals the image $P_v(\mathbb{R}^4)$. It is easy to see that $\dim\big(\operatorname{im} P_v\big)=3$ from the definition of P_v and from the observation that

$$P_v(e_l)=\begin{cases} e_l & \text{if } l=1,2,3\\ e_4-v & \text{if } l=4.\end{cases} \tag{3.13}$$

Thus, when $v=\big(\frac{1}{16},\frac{1}{8},\frac{1}{4},1\big)$ we conclude that

$$\ker P_v=\langle v\rangle,$$

where $\langle v\rangle$ stands for the linear span of the vector v. This follows from the fact that $\dim\ker P_v=1$ and $v\in\ker P_v$.

Next, we may assume that $P_v(a+\mathbb{R}e_i)\cap P_v(b+\mathbb{R}e_j)\neq\emptyset$ as otherwise the inclusion in (3.12) is obvious. Then there exists $t\in\mathbb{R}$ and $s\in\mathbb{R}$ such that

$$P_v(a+te_i-b-se_j)=0,$$

or equivalently, there exists $t,s\in\mathbb{R}$ and $r\in\mathbb{R}$ such that

$$(a-b)+te_i-se_j=rv. \tag{3.14}$$

To prove (3.11) we need to show that the equation (3.14) can have only trivial solutions (i.e. solutions for which $r=0$). In other words, we need to show that if the intersection $\big((a-b)+\mathbb{R}e_i \oplus \mathbb{R}e_j\big) \cap \langle v \rangle$ is nonempty, then

$$\big((a-b)+\mathbb{R}e_i \oplus \mathbb{R}e_j\big) \cap \langle v \rangle = \{0\}.$$

Because $(a+b)+\mathbb{R}e_i \oplus \mathbb{R}e_j$ is an affine vector space which is parallel to coordinate axes, and

$$\dim\big((a+b)+\mathbb{R}e_i \oplus \mathbb{R}e_j\big) \leq 2,$$

it is easy to see that for the vector $v = \big(\frac{1}{16}, \frac{1}{8}, \frac{1}{4}, 1\big)$ the intersection $\big((a-b)+\mathbb{R}e_i \oplus \mathbb{R}e_j\big) \cap \langle v \rangle$ can contain at most one point z. Then there are two possible cases we need to consider:

Case 1: Suppose first that $(a+\mathbb{R}e_i) \cap (b+\mathbb{R}e_j) \neq \emptyset$. In this case it follows that $z=0$ and the claim will follow because all the solutions to (3.14) are then trivial.

Case 2: Let us next assume that $(a+\mathbb{R}e_i) \cap (b+\mathbb{R}e_j) = \emptyset$. Then, because a and b were assumed to be vertices of $Q(0, \frac{1}{2})$ it follows that there is an index $i_1 \notin \{i,j\}$ such that

$$(a-b)_{i_1} \in \{1,-1\}.$$

Moreover, because $\dim\langle e_i, e_j, e_{i_1}\rangle < 4$ it will follow that there is also an index $i_2 \notin \{i, j, i_1\}$ such that

$$(a-b)_{i_2} \in \{1,0,-1\}.$$

However, this is a contradiction with the fact that the equation (3.14) was assumed to have a solution. Indeed, otherwise it would follow that there is $r \in \mathbb{R}$ such that

$$|rv_{i_1}| = 1 \quad \text{and} \quad |rv_{i_2}| \in \{1,0\}$$

which is not the case when $v = \big(\frac{1}{16}, \frac{1}{8}, \frac{1}{4}, 1\big)$.

Step 2B: Proving that the sets $\hat{\mathcal{S}}^i_{\boldsymbol{v}(k)}$ are disjoint. Recall now that

$$\hat{\mathcal{S}}^i_{\boldsymbol{v}(k)} := \Big(Q^3\big(P_v(\hat{z}_{\boldsymbol{v}(k)}), q\hat{r}_k\big) + P_v(\mathbb{R}e_i)\Big) \setminus \Big(\bigcup_{\boldsymbol{w}\in\mathbb{V}^{k-1}_{\boldsymbol{v}(k)}(i)} Q^3\big(P_v(\hat{z}_{\boldsymbol{w}}), \tfrac{7}{8}\hat{r}_{k-1}\big)\Big),$$

where $\mathbb{V}^{k-1}_{\boldsymbol{v}(k)}(i) := \big\{\boldsymbol{w} \in \mathbb{V}^{k-1} : (\hat{z}_{\boldsymbol{v}(k)} + \mathbb{R}e_i) \cap Q(\hat{z}_{\boldsymbol{w}}, \hat{r}_{k-1}) \neq \emptyset\big\}$. We claim that if we choose $\beta > 0$ sufficiently large in the definition of the Cantor set C_B then:

(1) We may replace the index set $\mathbb{V}^{k-1}_{\boldsymbol{v}(k)}(i)$ in the definition of $\hat{\mathcal{S}}^i_{\boldsymbol{v}(k)}$ by the index set $\mathbb{V}^{k-1}$. This will be only a technical detail which helps us to work with sets $\hat{\mathcal{S}}^i_{\boldsymbol{v}(k)}$ more easily.

(2) The sets $\hat{\mathcal{S}}^i_{\boldsymbol{v}(k)}$ are pairwise disjoint for each fixed $k \in \mathbb{N}$ (recall that $\hat{r}_k = 2^{-k}2^{-\beta k}$).

Proof of (1): We need to show that for every fixed q there exists $\beta_2 := \beta_2(q) > 0$ such that if we choose $\beta \geq \beta_2$ in the definition of the Cantor set C_B then for each fixed $\boldsymbol{v}(k) \in \mathbb{V}^k$ we have

$$\Big(Q^3\big(P_v(\hat{z}_{\boldsymbol{v}(k)}), q\hat{r}_k\big) + P_v(\mathbb{R}e_i)\Big) \cap Q^3\big(P_v(\hat{z}_{\boldsymbol{w}}), \tfrac{7}{8}\hat{r}_{k-1}\big) = \emptyset,$$

whenever $\boldsymbol{w} \in \mathbb{V}^{k-1} \setminus \mathbb{V}^{k-1}_{\boldsymbol{v}(k)}(i)$. It suffices to prove this for $k = 2$ because after this the general case follows from the self similarity of the construction.

First, we may find $\beta_2^1(q) > 0$ such that if $\beta \geq \beta_2^1$ in the definition of C_B then we have

$$Q^3(P_v(\hat{z}_{\boldsymbol{v}(2)}), q\hat{r}_2) + P_v(\mathbb{R}e_i) \subset\subset Q^3\big(P_v(\hat{z}_{\boldsymbol{w}}), \tfrac{7}{8}\hat{r}_1\big) + P_v(\mathbb{R}e_i),$$

whenever $\boldsymbol{w} \in \mathbb{V}$ is the parent of a given index $\boldsymbol{v}(2) \in \mathbb{V}^2$.

Next, by applying (3.11) and continuity of P_v we may find $\beta_2^2 > 0$ such that if $\beta \geq \beta_2^2$ in the definition of the Cantor set C_B then

$$\Big(Q^3\big(P_v(\hat{z}_{\boldsymbol{w}}), \tfrac{7}{8}\hat{r}_1\big) + P_v(\mathbb{R}e_i)\Big) \cap \Big(Q^3\big(P_v(\hat{z}_{\hat{\boldsymbol{w}}}), \tfrac{7}{8}\hat{r}_1\big) + P_v(\mathbb{R}e_j)\Big) = \emptyset,$$

whenever $\boldsymbol{w}, \hat{\boldsymbol{w}} \in \mathbb{V}$ are indices for which the intersection of the lines $l_{\boldsymbol{w}} = \hat{z}_{\boldsymbol{w}} + \mathbb{R}e_i$ and $l_{\hat{\boldsymbol{w}}} = \hat{z}_{\hat{\boldsymbol{w}}} + \mathbb{R}e_j$ is empty, i.e. if the intersection of lines is empty then the intersection of small neighborhoods is also empty.

Suppose now that $\beta \geq \beta_2 := \max\{\beta_2^1, \beta_2^2\}$. Let us fix $\boldsymbol{v}(2) \in \mathbb{V}^2$ and suppose that $\boldsymbol{w} \in \mathbb{V}$ is the parent of $\boldsymbol{v}(2)$. Let us also assume that $\hat{\boldsymbol{w}} \in \mathbb{V} \setminus \mathbb{V}^1_{\boldsymbol{v}(2)}(i)$. Then it follows that the lines $l_{\boldsymbol{w}} = \hat{z}_{\boldsymbol{w}} + \mathbb{R}e_i$ and $l_{\hat{\boldsymbol{w}}} = \hat{z}_{\hat{\boldsymbol{w}}} + \mathbb{R}e_i$ do not intersect each other, and therefore

$$\begin{aligned}\Big(Q^3\big(P_v(\hat{z}_{\boldsymbol{v}(2)}),&q\hat{r}_2\big) + P_v(\mathbb{R}e_i)\Big) \cap Q^3\big(P_v(\hat{z}_{\hat{\boldsymbol{w}}}), \tfrac{7}{8}\hat{r}_1\big)\\ &\subset \Big(Q^3\big(P_v(\hat{z}_{\boldsymbol{w}}), \tfrac{7}{8}\hat{r}_1\big) + P_v(\mathbb{R}e_i)\Big) \cap \Big(Q^3\big(P_v(\hat{z}_{\hat{\boldsymbol{w}}}), \tfrac{7}{8}\hat{r}_1\big) + P_v(\mathbb{R}e_i)\Big) = \emptyset,\end{aligned}$$

and (1) follows and we may write from now on

$$\hat{\mathcal{S}}^i_{\boldsymbol{v}(k)} = \Big(Q^3\big(P_v(\hat{z}_{\boldsymbol{v}(k)}), q\hat{r}_k\big) + P_v(\mathbb{R}e_i)\Big) \setminus \Big(\bigcup_{\boldsymbol{w}\in\mathbb{V}^{k-1}} Q^3\big(P_v(\hat{z}_{\boldsymbol{w}}), \tfrac{7}{8}\hat{r}_{k-1}\big)\Big).$$

Proof of (2): Again, by the self similarity of the construction it is enough to prove (2) in the case $k = 1$. Let us first assume that z is one of the vertices of the cube $Q(0, \frac{1}{2})$. Then, recalling that the center of the cube $\hat{Q}_0$ is $\tilde{z}_0 = 0$, we have that

$$\|P_v(z) - P_v(\hat{z}_0)\| = \|P_v(z) - P_v(0)\| = \|z - z_4 v\| < \tfrac{7}{8}. \tag{3.15}$$

This gives us that $P_v(z) \in Q^3(P_v(\hat{z}_0), \frac{7}{8}\hat{r}_0)$ for each vertex z of the cube $Q(0, \frac{1}{2})$. Suppose next that $\hat{z}_{\boldsymbol{v}(1)}$ and $\hat{z}_{\hat{\boldsymbol{v}}(1)}$ are two (possibly identical) vertices of $Q(0, \frac{1}{2})$, and consider two (nonidentical) lines $M_1 = \hat{z}_{\boldsymbol{v}(1)} + \mathbb{R}e_i$ and $M_2 = \hat{z}_{\hat{\boldsymbol{v}}(1)} + \mathbb{R}e_j$ through the points $\hat{z}_{\boldsymbol{v}(1)}$ and $\hat{z}_{\hat{\boldsymbol{v}}(1)}$. Then by applying (3.15) to points $\hat{z}_{\boldsymbol{v}(1)}$ and $\hat{z}_{\hat{\boldsymbol{v}}(1)}$, and using (3.11) we get

$$\big(P_v(M_1)\setminus Q^3\big(P_v(\hat{z}_0), \tfrac{7}{8}\hat{r}_0\big)\big) \cap \big(P_v(M_2)\setminus Q^3\big(P_v(\hat{z}_0), \tfrac{7}{8}\hat{r}_0\big)\big) = \emptyset.$$

Therefore, by linearity and Lipschitz continuity of P_v and by the fact that the lines $P_v(M_1)$ and $P_v(M_2)$ intersect at most at one point it is easy to see that there exists

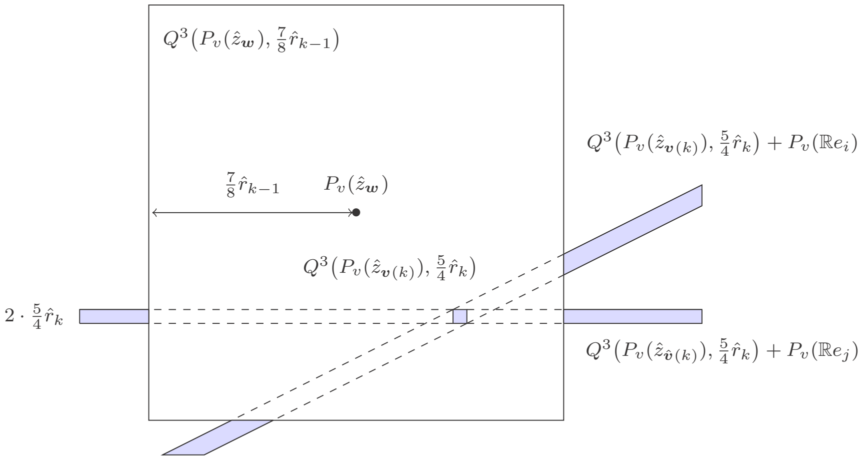


FIGURE 8. When we choose $\beta > 0$ large enough the ratio $\hat{r}_{k-1}/\hat{r}_k = 2^{1+\beta}$ will be very large. This gives enough space for the lines $P_v(\hat{z}_{\boldsymbol{v}(k)} + \mathbb{R}e_i)$ and $P_v(\hat{z}_{\hat{\boldsymbol{v}}(k)} + \mathbb{R}e_j)$ to recede from each other before they reach the boundary of the big cube. Especially, it follows from this and the linearity of the mapping P_v that the intersection of the sets $Q^3\big(P_v(\hat{z}_{\boldsymbol{v}(k)}), \frac{5}{4}\hat{r}_k\big) + P_v(\mathbb{R}e_i)$ and $Q^3\big(P_v(\hat{z}_{\hat{\boldsymbol{v}}(k)}), \frac{5}{4}\hat{r}_k\big) + P_v(\mathbb{R}e_j)$ is empty outside the cubes $Q^3\big(P_v(\hat{z}_{\boldsymbol{w}}), \frac{7}{8}\hat{r}_{k-1}\big)$, $\boldsymbol{w} \in \mathbb{V}^{k-1}$.

$\beta_3 > 0$ such that if we choose $\beta \geq \beta_3$ in the definition of the Cantor set C_B then we get

$$\begin{aligned}\hat{\mathcal{S}}^i_{\boldsymbol{v}(1)} \cap \hat{\mathcal{S}}^j_{\hat{\boldsymbol{v}}(1)} = &\Big(\big(Q^3\big((P_v(\hat{z}_{\boldsymbol{v}(1)}), q\hat{r}_1\big) + P_v(\mathbb{R}e_i)\big) \cap \big(Q^3\big((P_v(\hat{z}_{\hat{\boldsymbol{v}}(1)}), q\hat{r}_1\big) + P_v(\mathbb{R}e_j)\big)\Big)\\ &\setminus Q^3\big(P_v(\hat{z}_0), \tfrac{7}{8}\hat{r}_0\big) = \emptyset,\end{aligned}$$

see also Fig. 8. By working through all the combinations $\boldsymbol{v}(1), \hat{\boldsymbol{v}}(1) \in \mathbb{V}$ we may also assume that $\beta_3 > 0$ is independent on the pair $(\boldsymbol{v}(1), \hat{\boldsymbol{v}}(1)) \in \mathbb{V} \times \mathbb{V}$. This gives us that the sets $\hat{\mathcal{S}}^i_{\boldsymbol{v}(k)}$ are pairwise disjoint for $k = 1$.

To see that $\hat{\mathcal{S}}^i_{\boldsymbol{v}(k)} \cap \hat{\mathcal{S}}^j_{\hat{\boldsymbol{v}}(k)} = \emptyset$ for $k \geq 2$ one may apply self similarity of the construction together with the previous argument where $k = 1$. Self similarity argument applies to this situation as the ratio $\hat{r}_{k-1}/\hat{r}_k = 2^{1+\beta}$ stays the same for every $k \in \mathbb{N}$ (see also Fig. 8).

Step 2C: Proving the inclusion $P_v(\hat{S}^i_{\boldsymbol{v}(k)}) \subset \hat{\mathcal{S}}^i_{\boldsymbol{v}(k)}$. Let us next recall the definition of k-bars

$$\hat{G}^i_{\boldsymbol{v}(k)} := \hat{Q}_{\boldsymbol{v}(k)} + \mathbb{R}e_i.$$

We also recall (see the paragraph after (3.5)) that there exists $\beta_1 > 0$ such that if we choose $\beta \geq \beta_1$ in the definition of the Cantor set C_B, then we may define the corresponding sliced k-bars for k-bars as

$$\hat{S}^i_{\boldsymbol{v}(k)} := \hat{G}^i_{\boldsymbol{v}(k)} \setminus \Big(\bigcup_{\boldsymbol{w} \in \mathbb{V}^{k-2}} Q(\hat{z}_{\boldsymbol{w}}, \hat{r}_{k-2})\Big).$$

We want to show that $P_v(\hat{S}^i_{\boldsymbol{v}(k)}) \subset \hat{\boldsymbol{\mathcal{S}}}^i_{\boldsymbol{v}(k)}$. For this we first observe that for every $x, y \in \mathbb{R}^4$ such that $\|x-y\| < \hat{r}_k$, where $\|.\|$ denotes the maximum norm, we have by (3.2)

$$\begin{aligned}\|P_v(x) - P_v(y)\| = \big(&x_1 - y_1 - (x_4 - y_4)v_1, x_2 - y_2 - (x_4 - y_4)v_2,\\ &x_3 - y_3 - (x_4 - y_4)v_3, 0\big) \le \tfrac{5}{4}\|x-y\| < \tfrac{5}{4}\hat{r}_k.\end{aligned}$$

Thus, it follows that

$$P_v(\hat{S}^i_{\boldsymbol{v}(k)}) \subset Q^3\big(P_v(\hat{z}_{\boldsymbol{v}(k)}), \tfrac{5}{4}\hat{r}_k\big) + P_v(\mathbb{R}e_i), \tag{3.16}$$

and hence the requirement that $q \ge \frac{5}{4}$ in the definition (3.6). Therefore it suffices now to show that for every $x \in \hat{S}^i_{\boldsymbol{v}(k)}$ we have

$$\|P_v(x) - P_v(\hat{z}_{\boldsymbol{w}})\| > \tfrac{7}{8}\hat{r}_{k-1}, \tag{3.17}$$

whenever $\boldsymbol{w} \in \mathbb{V}^{k-1}$. Actually, by assuming that $\beta \ge \beta_2$ we need to verify (3.17) only for every $\boldsymbol{w} \in \mathbb{V}^{k-1}_{\boldsymbol{v}(k)}(i)$. For this, let us assume that $x \in \hat{S}^i_{\boldsymbol{v}(k)}$ and $\boldsymbol{w} \in \mathbb{V}^{k-1}_{\boldsymbol{v}(k)}(i)$. Then we have to consider two different cases:

(i) Suppose first that $i \ne 4$. If we denote $y := \hat{z}_{\boldsymbol{w}}$ we get $|x_4 - y_4| < \hat{r}_{k-1}$ and $|x_i - y_i| \ge \hat{r}_{k-2}$. Thus, it follows that

$$\|P_v(x) - P_v(y)\| = \|x - y - (x_4 - y_4)v\| > \hat{r}_{k-2} - \tfrac{1}{4}\hat{r}_{k-1} > \tfrac{7}{8}\hat{r}_{k-1},$$

simply because we know that $\beta > 2$ and (3.17) follows.

(ii) Next we assume that $i = 4$. If we write $y := \hat{z}_{\boldsymbol{v}(k)} - x$ we get that

$$|y_4| \ge \hat{r}_{k-2} \text{ and } |y_i| < \hat{r}_{k-1} \text{ for all } i = 1, 2, 3.$$

These estimates give us

$$\|P_v(\hat{z}_{\boldsymbol{v}(k)}) - P_v(x)\| = \|P_v(y)\| = \|y - y_4 v\| \ge |\tfrac{1}{4}y_4 - y_3| \ge \tfrac{1}{4}\hat{r}_{k-2} - \hat{r}_{k-1} > \tfrac{7}{8}\hat{r}_{k-1},$$

which implies (3.17) by having $\beta > 3$.

Therefore, by combining (3.16) and (3.17) together we conclude that $P_v(\hat{S}^i_{\boldsymbol{v}(k)}) \subset \hat{\boldsymbol{\mathcal{S}}}^i_{\boldsymbol{v}(k)}$ as we wanted.

Step 2D: Conclusion of Step 2. In Step 1 we have already showed that P_v is one-to-one on C_B. Thus, it suffices to show that P_v is one-to-one also on $\mathcal{K}_B \setminus C_B$ and then we can easily see, for example by the linearity of P_v, that P_v is in fact one-to-one on $\mathcal{K}_B$.

To see that P_v is one-to-one on $\mathcal{K}_B \setminus \mathcal{C}_B$ suppose that ℓ and $\hat{\ell}$ are two distinct lines in $\mathcal{K}_B$. It is easy to see that P_v is one-to-one along these lines and thus it suffices to show that

$$P_v\big(\ell \setminus C_B\big) \cap P_v\big(\hat{\ell} \setminus C_B\big) = \emptyset.$$

For this we observe that the intersection of ℓ and $\hat{\ell}$ is either an empty set or one point which lies in the set C_B. Therefore we may find an index $N \in \mathbb{N}$, and sequences $\{\hat{S}^i_{\boldsymbol{v}(k)}\}_{k=N}^{\infty}$ and $\{\hat{S}^j_{\hat{\boldsymbol{v}}(k)}\}_{k=N}^{\infty}$ of sliced k-bars such that $\{\ell \cap \hat{S}^i_{\boldsymbol{v}(k)}\}_{k=N}^{\infty}$ and $\{\hat{\ell} \cap \hat{S}^j_{\hat{\boldsymbol{v}}(k)}\}_{k=N}^{\infty}$ are two sequences of sets, and it holds that

$$\lim_{k\to\infty} \ell \cap \hat{S}^i_{\boldsymbol{v}(k)} = \ell \setminus C_B, \quad \lim_{k\to\infty} \hat{\ell} \cap \hat{S}^j_{\hat{\boldsymbol{v}}(k)} = \hat{\ell} \setminus C_B, \text{ and } \quad \hat{S}^i_{\boldsymbol{v}(k)} \cap \hat{S}^j_{\hat{\boldsymbol{v}}(k)} = \emptyset$$

for every $k \geq N$. Furthermore, by step 2B and step 2C we have $P_v(\hat{S}^i_{\boldsymbol{v}(k)}) \subset \hat{\mathcal{S}}^i_{\boldsymbol{v}(k)}$ and $P_v(\hat{S}^j_{\hat{\boldsymbol{v}}(k)}) \subset \hat{\mathcal{S}}^j_{\hat{\boldsymbol{v}}(k)}$ where $\hat{\mathcal{S}}^i_{\boldsymbol{v}(k)} \cap \hat{\mathcal{S}}^j_{\hat{\boldsymbol{v}}(k)} = \emptyset$, and therefore

$$P_v\big(\ell \setminus C_B\big) \cap P_v\big(\hat{\ell} \setminus C_B\big) \subset \lim_{k\to\infty} P_v\big(\hat{S}^i_{\boldsymbol{v}(k)}\big) \cap P_v\big(\hat{S}^j_{\hat{\boldsymbol{v}}(k)}\big) \subset \lim_{k\to\infty} \hat{\mathcal{S}}^i_{\boldsymbol{v}(k)} \cap \hat{\mathcal{S}}^j_{\hat{\boldsymbol{v}}(k)} = \emptyset,$$

which ends this step.

Step 3: Defining the function g on X. We have that the sets $\hat{\mathcal{S}}^i_{\boldsymbol{v}(k)}$ (recall that their definition is dependent on a positive parameter q) are disjoint if distinct. Therefore also the sets, which one could call (punctured) pipes,

(3.18)

$$\begin{aligned}\hat{H}^i_{\boldsymbol{v}(k)} := \hat{H}^i_{\hat{z}_{\boldsymbol{v}(k)}} &:= \partial_X \Big(Q^3(P_v(\hat{z}_{\boldsymbol{v}(k)}), q\hat{r}_k) + P_v(\mathbb{R}e_i)\Big) \setminus \Big(\bigcup_{\boldsymbol{w}\in\mathbb{V}^{k-1}_{\boldsymbol{v}(k)}(i)} Q^3(P_v(\hat{z}_{\boldsymbol{w}}), \tfrac{7}{8}\hat{r}_{k-1})\Big) \\ &= \partial_X \Big(Q^3(P_v(\hat{z}_{\boldsymbol{v}(k)}), q\hat{r}_k) + P_v(\mathbb{R}e_i)\Big) \setminus \Big(\bigcup_{\boldsymbol{w}\in\mathbb{V}^{k-1}} Q^3(P_v(\hat{z}_{\boldsymbol{w}}), \tfrac{7}{8}\hat{r}_{k-1})\Big),\end{aligned}$$

are pairwise disjoint sets for distinct bars. Here $\partial_X A$ denotes the relative boundary of a set A in $X = \mathbb{R}^3 \times \{0\}$.

It is worth noticing that lines in $\mathcal{K}_B$ parallel to e_i are contained in the interior of $\hat{G}^i_{\boldsymbol{v}(k)}$-type bars and therefore also the projection of the line is contained in the 3-dimensional interior of the projection of the bar. Especially the projection of a line in $\mathcal{K}_B$ never intersects a punctured pipe. When we say that the projection of a line in $\mathcal{K}_B$ is inside a pipe $\hat{H}^i_{\hat{z}_{\boldsymbol{v}(k)}}$ we mean that the line in $\mathcal{K}_B$ lies in the bar $\hat{G}^i_{\boldsymbol{v}(k)}$ from which we derived the pipe. In fact we can claim not only that the projection of lines in $\mathcal{K}_B$ do not intersect pipes, but further we know that the projection of a line in $\mathcal{K}_B$ lies inside the projection of some $(k+1)$-bar and that means that there are no lines in $\mathcal{K}_B$ whose projection intersects the set

$$L^i_{\boldsymbol{v}(k)} := \hat{\mathcal{S}}^i_{\boldsymbol{v}(k)} \setminus \big(\bigcup_{\boldsymbol{v}(k+1)\in\mathbb{V}^{k+1}} \hat{\mathcal{S}}^i_{\boldsymbol{v}(k+1)}\big). \tag{3.19}$$

With respect to this fact we will extend our Lipschitz function g in the following way. We will define g on the projection of lines in $\mathcal{K}_B$ and on punctured pipes. We will show that our definition is Lipschitz and then extend it in a Lipschitz way inside $L^i_{\boldsymbol{v}(k)}$. We will take care during the extension to guarantee that (3.3) holds, which is not difficult. Then there will be some remaining part of X where we can define g practically arbitrarily as long as we maintain the Lipschitz property.

As mentioned in our outline, we will define

$$g(P_v(x)) = -x_n \text{ for } x \in \mathcal{K}_B.$$

We now wish to show that this can be extended in a Lipschitz way onto X. Our argument will make use of pipes, but for pipes of type $\hat{H}^i_{\hat{z}_{\boldsymbol{v}(k)}}$, with $i = 1, 2, 3$ it is slightly more simple than for $i = 4$. We will deal with the simpler case first then note the difference for the case $i = 4$.

Step 3A. First we take a pipe $\hat{H}^i_{\hat{z}_{\boldsymbol{v}(k)}}$, with $i = 1, 2, 3$ and $k \geq 2$. Then we define

$$g(x) = -(\hat{z}_{\boldsymbol{v}(k)})_4 \quad \text{for all } x \in \hat{H}^i_{\hat{z}_{\boldsymbol{v}(k)}}.$$

Let us note that this definition is well-defined, if $\hat{H}^i_{\hat{z}_{\boldsymbol{v}(k)}} = \hat{H}^i_{\hat{z}_{\tilde{\boldsymbol{v}}(k)}}$ then $(\hat{z}_{\boldsymbol{v}(k)})_4 = (\hat{z}_{\tilde{\boldsymbol{v}}(k)})_4$. Now it is very easy to notice that if we have two pipes, one inside another (that is $\hat{z}_{\boldsymbol{v}(k+1)} + \mathbb{R}e_i$ intersects $\hat{Q}_{\boldsymbol{v}(k)}$), then

$$\text{dist}(\hat{H}^i_{\hat{z}_{\boldsymbol{v}(k)}}, \hat{H}^i_{\hat{z}_{\boldsymbol{v}(k+1)}}) \geq C\hat{r}_k \text{ for a suitable } C > 0 \text{ independent of } k. \tag{3.20}$$

Further considering $x \in \hat{H}^i_{\hat{z}_{\boldsymbol{v}(k)}}$ and $y \in \hat{H}^i_{\hat{z}_{\boldsymbol{v}(k+1)}}$ we have

$$|g(x) - g(y)| = |-(\hat{z}_{\boldsymbol{v}(k)})_4 + (\hat{z}_{\boldsymbol{v}(k+1)})_4| = \tfrac{1}{2}\hat{r}_k.$$

Considering two distinct pipes of the same generation, both inside $\hat{H}^i_{\hat{z}_{\boldsymbol{v}(k)}}$ we see that

$$\text{dist}(\hat{H}^i_{\hat{z}_{\hat{\boldsymbol{v}}(k+1)}}, \hat{H}^i_{\hat{z}_{\boldsymbol{v}(k+1)}}) \geq C\hat{r}_k \text{ for a suitable } C > 0 \text{ independent of } k.$$

Furthermore, for $x \in \hat{H}^i_{\hat{z}_{\hat{\boldsymbol{v}}(k+1)}}$ and $y \in \hat{H}^i_{\hat{z}_{\boldsymbol{v}(k+1)}}$ we have

$$|g(x) - g(y)| = |-(z_{\hat{\boldsymbol{v}}(k+1)})_4 + (z_{\boldsymbol{v}(k+1)})_4| \leq \hat{r}_k.$$

This proves that g, thus defined, on the pipes $\hat{H}^i_{\hat{z}_{\boldsymbol{v}(k)}}$, $i = 1, 2, 3$, is Lipschitz with respect to parallel pipes.

Step 3B. Now consider a line l through the Cantor set C_B parallel to e_i, $i \in \{1, 2, 3\}$, whose projection lies inside the pipe $\hat{H}^i_{\hat{z}_{\boldsymbol{v}(k)}}$. For each such line l we define

$$g(P_v(l)) = -x_4 \quad \text{where } x \in l \cap C_B. \tag{3.21}$$

Next, we calculate that

$$\text{dist}(P_v(l), \hat{H}^i_{\hat{z}_{\boldsymbol{v}(k)}}) \geq C\hat{r}_k \text{ for a suitable } C \text{ independent of } k.$$

On the other hand we may observe that g is constant on each line $P_v(l)$ described above, and thus by taking $z \in P_v(l), y \in \hat{H}^i_{\hat{z}_{\boldsymbol{v}(k)}}$ we observe

$$|g(z) - g(y)| = |-x_4 + (\hat{z}_{\boldsymbol{v}(k+1)})_4| \leq 2\hat{r}_k.$$

But this shows that we have defined g Lipschitz on the set of pipes and projection of lines through the Cantor set for those pipes and lines parallel to e_i, $i = 1, 2, 3$.

Strictly speaking we should check that our definition of g is Lipschitz, when we compare $x \in \hat{H}^i_{\hat{z}_{\boldsymbol{v}(k)}}$ and $y \in \hat{H}^j_{\hat{z}_{\boldsymbol{v}(k)}}$ for $i, j \in \{1, 2, 3\}$ also for $i \neq j$ but the considerations and calculations from step 1 and step 2 show that the distance between these pipes is at least $C\hat{r}_k$ and $|g(x) - g(y)| \leq 2\hat{r}_k$ and so this part of the argument is easy.

Step 3C. Now we define g on $L^i_{\boldsymbol{v}(k)}$, $i = 1, 2, 3$, as follows (recall that $L^i_{\boldsymbol{v}(k)}$ are defined in (3.19)). Choose $i \in \{1, 2, 3\}$ and fix a 2-dimensional hyperplane $Y_i \subset\subset X$ perpendicular to e_i, such that Y_i intersects all of the pipes $\hat{H}^i_{\hat{z}_{\boldsymbol{v}(k)}}$, $k \geq 2$. We may write

$$Y_i = \Big\{ y_i e_i + \sum_{j \neq i} t_j e_j : t_j \in \mathbb{R} \text{ for every } j \neq i \Big\},$$

where $y_i \in \mathbb{R}$ is fixed. We define a projection π_{Y_i} from X onto Y_i by

$$(\pi_{Y_i}(x))_j = \begin{cases} x_j & j \neq i \\ y_i & j = i. \end{cases}$$

We use the McShane extension theorem on the hyperplane Y_i to extend g on those parts of the set $Y_i \cap \big(\bigcup_{k=2}^\infty \hat{\mathcal{S}}^i_{\hat{z}_{\boldsymbol{v}(k)}}\big)$ where we did not define g during the previous steps (step 3A and 3B) and then we define g at other points x in $\hat{\mathcal{S}}^i := \bigcup_{k=2}^\infty \hat{\mathcal{S}}^i_{\hat{z}_{\boldsymbol{v}(k)}}$ by simply projecting x onto Y_i and then using g. In other words

$$g(x) = g(\pi_{Y_i}(x)) \quad \text{for all } x \in \hat{\mathcal{S}}^i. \tag{3.22}$$

Thus defined the function g is constant on the intersection of lines parallel to e_i with the set $\hat{\mathcal{S}}^i$.

Step 3D. Our argument in the projection of bars parallel to e_4 is identical to the previous, up to the fact that we do not define g as constant equal to $-(\hat{z}_{\boldsymbol{v}(k)})_4$ on pipes generated by $\hat{G}^4_{\boldsymbol{v}(k)}$-type bars but by using an appropriate affine function. For this we recall that $\tilde{v} = \big(\frac{1}{16}, \frac{1}{8}, \frac{1}{4}\big)$ and we denote

$$Y_4 := \{w \in \mathbb{R}^3 : \langle w, \tilde{v} \rangle = 0\}.$$

Then we may separate $\mathbb{R}^3$ into the direct sum $\mathbb{R}\tilde{v} \oplus Y_4$. Now, suppose that $\lambda_0 \in \mathbb{R}$ and $w_0 \in Y_4$ are such that

$$P_v(\hat{z}_{\boldsymbol{v}(k)}) = w_0 + \lambda_0 \tilde{v}. \tag{3.23}$$

Then, if $\tilde{x} \in \hat{H}^4_{\boldsymbol{v}(k)} \cup P_v(\hat{z}_{\boldsymbol{v}(k)} + \mathbb{R}e_4)$ we may find $\lambda \in \mathbb{R}$ and $w \in Y_4$ such that $\tilde{x} = w + \lambda \tilde{v}$ which leads us to define

$$g(\tilde{x}) = \lambda - \lambda_0 - (\hat{z}_{\boldsymbol{v}(k)})_4 \quad \text{for every } \tilde{x} \in \hat{H}^4_{\boldsymbol{v}(k)} \cup P_v(\hat{z}_{\boldsymbol{v}(k)} + \mathbb{R}e_4). \tag{3.24}$$

We proceed to prove that by defining

$$\begin{array}{ll} g(\tilde{x}) = \lambda - \lambda_0 - (\hat{z}_{\boldsymbol{v}(k)})_4 & \text{for every } \tilde{x} \in \hat{H}^4_{\boldsymbol{v}(k)} \cup P_v(\hat{z}_{\boldsymbol{v}(k)} + \mathbb{R}e_4) \\ g(P_v(x)) = -x_4 & \text{for every } x \in \mathcal{K}_B \end{array}$$

we get a Lipschitz function on $\hat{H}^4_{\boldsymbol{v}(k)} \cup P_v(\hat{z}_{\boldsymbol{v}(k)} + \mathbb{R}e_4) \cup P_v(\mathcal{K}_B)$. A first observation is that for every $x \in \hat{z}_{\boldsymbol{v}(k)} + \mathbb{R}e_4$ we find α such that $x = \hat{z}_{\boldsymbol{v}(k)} + \alpha e_4$ and then by (3.13) and (3.23) we get

$$P_v(x) = P_v(\hat{z}_{\boldsymbol{v}(k)}) + \alpha P_v(e_4) = \big(w_0 + \lambda_0 \tilde{v} - \alpha \tilde{v}, 0\big).$$

Now we will apply (3.24) with $\tilde{x} = P_v(x) = w_0 + \lambda_0 \tilde{v} - \alpha \tilde{v}$ to get

$$g(P_v(x)) = \lambda_0 - \alpha - \lambda_0 - (\hat{z}_{\boldsymbol{v}(k)})_4 = -x_4 \text{ for all } x \in \hat{z}_{\boldsymbol{v}(k)} + \mathbb{R}e_4.$$

The rest of the argument will be a case of proving that g has similar values on $\hat{H}^4_{\boldsymbol{v}(k)}$ (up to an error of $C\hat{r}_k$) and the distance between $P_v(C_B+\mathbb{R}e_4)$ and $\hat{H}^4_{\boldsymbol{v}(k)}$ is $C\hat{r}_k$. Let us continue to expound.

Our choice of $\beta>1$ guarantees that the Cantor set C_B is at a distance of at least $\frac{1}{4}\hat{r}_k$ from the boundary of the cubes $\hat{Q}_{\hat{z}_{\boldsymbol{v}(k)}}$. Now we will take any $c\in C_B\cap\hat{Q}_{\boldsymbol{v}(k)}$ and (recall the definition of k-bars from (3.4)) we will see that

$$c+\mathbb{R}e_4+Q(0,\tfrac{1}{4}\hat{r}_k)\subset\subset\hat{G}^4_{\hat{z}_{\boldsymbol{v}(k)}}.$$

Our projection P_v is continuous onto X and therefore there is a $C>0$ such that

$$P_v(c+\mathbb{R}e_4)+Q^3(0,C\hat{r}_k)\subset P_v(\hat{G}^4_{\hat{z}_{\boldsymbol{v}(k)}})$$

implying that there is $C_1>0$ such that

$$\operatorname{dist}(\hat{H}^4_{\boldsymbol{v}(k)},P_v(c+\mathbb{R}e_4))\geq C_1\hat{r}_k \tag{3.25}$$

for all $\boldsymbol{v}(k)$, all c and all k. Exactly the same argument gives that

$$\operatorname{dist}(\hat{H}^4_{\boldsymbol{v}(k)},P_v(\hat{z}_{\boldsymbol{v}(k)}+\mathbb{R}e_4))\geq C_1\hat{r}_k. \tag{3.26}$$

Furthermore we can make the opposite estimates since for some $C>0$

$$c+\mathbb{R}e_4+Q(0,C\hat{r}_k)\supset\hat{G}^4_{\hat{z}_{\boldsymbol{v}(k)}}$$

and the continuity of our projection gives

$$\operatorname{dist}(x,P_v(c+\mathbb{R}e_4))\leq C_2\hat{r}_k \tag{3.27}$$

for any $x\in\hat{H}^4_{\boldsymbol{v}(k)}$ and similarly

$$\operatorname{dist}(x,P_v(\hat{z}_{\boldsymbol{v}(k)}+\mathbb{R}e_4))\leq C_2\hat{r}_k, \tag{3.28}$$

for any $x\in\hat{H}^4_{\boldsymbol{v}(k)}$.

Now we will be able to show that g is a Lipschitz function when restricted to $\hat{H}^4_{\boldsymbol{v}(k)}\cup P_v(\hat{z}_{\boldsymbol{v}(k)}+\mathbb{R}e_4)$. For every point $x\in\hat{H}^4_{\boldsymbol{v}(k)}$ we find a point $w\in P_v(\hat{z}_{\boldsymbol{v}(k)}+\mathbb{R}e_4)$ such that $g(x)=g(w)$ and $|x-w|$ is bounded by a constant multiple of $\hat{r}_k$. Finally, since g is linear on the line $P_v(\hat{z}_{\boldsymbol{v}(k)}+\mathbb{R}e_4)$ we will be able to prove the desired Lipschitz quality of g by (3.26), when y is close to x and w, and by $|w-y|\approx|g(w)-g(y)|$, when y is far from x and w.

Now recall that $P_v(\mathbb{R}e_4)=\mathbb{R}\tilde{v}$ (see (3.13)) and Y_4 is the linear space perpendicular to $\tilde{v}$. We take $x\in\hat{H}^4_{\boldsymbol{v}(k)}$ and claim that there exists a unique $w\in P_v(\hat{z}_{\boldsymbol{v}(k)}+\mathbb{R}e_4)\cap(x+Y_4)$, which is obvious because $P_v(\hat{z}_{\boldsymbol{v}(k)}+\mathbb{R}e_4)$ and $x+Y_4$ are a pair of perpendicular affine spaces in a 3 dimensional space and the sum of their dimensions is 3. Quite simply because $x-w\in Y_4$ and $P_v(\hat{z}_{\boldsymbol{v}(k)}+\mathbb{R}e_4)$ is perpendicular to Y_4 we see that w is the closest point to x in $P_v(\hat{z}_{\boldsymbol{v}(k)}+\mathbb{R}e_4)$. Using (3.26) and (3.28) we may estimate

$$C_1\hat{r}_k\leq|w-x|\leq C_2\hat{r}_k.$$

Also, since (3.24) gives that g is constant on the intersection of any affine plane parallel to Y_4 with the set $\hat{H}^4_{\boldsymbol{v}(k)}\cup P_v(\hat{z}_{\boldsymbol{v}(k)}+\mathbb{R}e_4)$, we have that

$$g(x)=g(w).$$

Now we may take any $x \in \hat{H}^4_{\boldsymbol{v}(k)}$, its corresponding $w \in (x + Y_4) \cap P_v(\hat{z}_{\boldsymbol{v}(k)} + \mathbb{R}e_4)$ and any $y \in P_v(\hat{z}_{\boldsymbol{v}(k)} + \mathbb{R}e_4)$ and calculate

$$|g(y) - g(x)| = |g(y) - g(w)| = |\langle (y - w), \tfrac{\tilde{v}}{|\tilde{v}|^2} \rangle| = |\tilde{v}|^{-1} |y - w|$$
$$|y - x| \geq |y - w| - |w - x| \geq |y - w| - C_2 \hat{r}_k.$$

When $|y - w| > 2C_2\hat{r}_k$ then $|y - w| > 2|x - w|$ and therefore $|x - y| > \frac{1}{2}|w - y|$ and we may estimate

$$\frac{|g(x) - g(y)|}{|x - y|} \leq \frac{2|g(w) - g(y)|}{|w - y|} \leq \frac{2|w - y|}{|\tilde{v}||w - y|}.$$

When $|y - w| \leq 2C_2\hat{r}_k$ we use (3.26)

$$\frac{|g(x) - g(y)|}{|x - y|} \leq \frac{|w - y|}{|\tilde{v}| C_1 \hat{r}_k} \leq \frac{2C_2}{C_1 |\tilde{v}|}.$$

By a very similar argument we will proceed to prove that g is C_3-Lipschitz when restricted to $P_v(\mathcal{K}_B) \cup P_v(\hat{z}_{\boldsymbol{v}(k)} + \mathbb{R}e_4)$. Take a point $c \in C_B$ and the unique $\hat{z}_{\boldsymbol{v}(k)}$ such that $c \in Q_{\boldsymbol{v}(k)}$. First we observe that

$$g(P_v(\hat{z}_{\boldsymbol{v}(k)})) = -(\hat{z}_{\boldsymbol{v}(k)})_4 \text{ and } g(P_v(c)) = -c_4$$

and

$$|c_4 - (\hat{z}_{\boldsymbol{v}(k)})_4| < C\hat{r}_k.$$

If $x \in P_v(c + \mathbb{R}e_4)$ and $y \in P_v(\hat{z}_{\boldsymbol{v}(k)} + \mathbb{R}e_4)$ then

$$|g(x) - g(y)| \leq \langle (x - y), \tfrac{\tilde{v}}{|\tilde{v}|^2} \rangle + C\hat{r}_k.$$

Further, distance estimates similar to (3.25)–(3.28) hold also for $P_v(c + \mathbb{R}e_4)$ and $P_v(\hat{z}_{\boldsymbol{v}(k)} + \mathbb{R}e_4)$, and therefore

$$|x - y| \geq C\langle (x - y), \frac{\tilde{v}}{|\tilde{v}|^2} \rangle + C\hat{r}_k.$$

Hence

$$\frac{|g(x) - g(y)|}{|x - y|} \leq \frac{\langle (x - y), \frac{\tilde{v}}{|\tilde{v}|^2} \rangle + C\hat{r}_k}{C\langle (x - y), \frac{\tilde{v}}{|\tilde{v}|^2} \rangle + C\hat{r}_k} < C_3.$$

This means, that g is C_3-Lipschitz when restricted to

$$P_v(\mathcal{K}_B) \cup P_v(\hat{z}_{\boldsymbol{v}(k)} + \mathbb{R}e_4).$$

Now we can show that the restriction of g to $P_v(\mathcal{K}_B) \cup \hat{H}^4_{\boldsymbol{v}(k)} \cup P_v(\hat{z}_{\boldsymbol{v}(k)} + \mathbb{R}e_4)$ is Lipschitz. Assume that we have $x \in \hat{H}^4_{\boldsymbol{v}(k)}$ and $y \in P_v(\mathcal{K}_B)$. If

$$|g(x) - g(y)| \leq 2C_2C_3\hat{r}_k$$

then (3.25) says that g has been defined Lipschitz. Therefore we consider the case

$$|g(x) - g(y)| > 2C_2C_3\hat{r}_k.$$

We have a $w \in (x + Y_4) \cap P_v(\hat{z}_{\boldsymbol{v}(k)} + \mathbb{R}e_4)$ and $|x - w| < C_2\hat{r}_k$. Since g is C_3-Lipschitz when restricted to the set $P_v(\mathcal{K}_B) \cup \bigcup_{\boldsymbol{v}(k) \in \mathbb{V}^k} P_v(\hat{z}_{\boldsymbol{v}(k)} + \mathbb{R}e_4)$ we have that

$$|w - y| \geq \frac{|g(w) - g(y)|}{C_3} = \frac{|g(x) - g(y)|}{C_3}$$

and therefore

$$|x-y| \geq |w-y| - |x-w| \geq \frac{|g(x)-g(y)|}{C_3} - C_2\hat{r}_k.$$

We get

$$\frac{|g(x)-g(y)|}{|x-y|} \leq \frac{C_3|g(x)-g(y)|}{|g(x)-g(y)| - C_3C_2\hat{r}_k} \leq \frac{C_3}{1-\frac{1}{2}}.$$

So we prove that g is $2C_3C_2C_1^{-1}$-Lipschitz when restricted to the set

$$P_v(\mathcal{K}_B) \cup \bigcup P_v(z_{\boldsymbol{v}(k)} + \mathbb{R}e_4) \cup \bigcup_{\boldsymbol{v}(k)\in\mathbb{V}^k} \hat{H}^4_{\hat{z}_{\boldsymbol{v}(k)}}.$$

Of course self-similarity means that $2C_3C_2C_1^{-1}$ is independent of k.

It is not hard to estimate that

$$\operatorname{dist}(\hat{H}^4_{\hat{z}_{\boldsymbol{v}(k)}}, \hat{H}^4_{\hat{z}_{\tilde{\boldsymbol{v}}(\tilde{k})}}) \approx \hat{r}_{\min\{k,\tilde{k}\}}.$$

Therefore we see that the definition (3.24) is Lipschitz on the collection of all e_4 pipes. We use the construction described before (3.22) to get a Lipschitz extension which guarantees that

$$g(x+t\tilde{v}) = g(x) + t \tag{3.29}$$

everywhere in $\hat{\mathcal{S}}^4_{\boldsymbol{v}(k)}$, this time by projecting onto Y_4.

Where not yet defined we may extend g Lipschitz arbitrarily, for example by the McShane extension theorem.

Step 3E: verifying the condition (3.3)**.** Now it is quite simple to notice that we have

$$D_iF_{g,u} \circ F_{g,v}(x) = e_i,\ i = 1,2,3 \text{ and } D_4F_{g,u} \circ F_{g,v}(x) = -e_4$$

whenever $x_i \in [-1,1] \setminus \hat{U}_k$ and $x_j \in \hat{U}_{k+2}$ for all $j \neq i$. This can be seen from the following arguments. Firstly, it follows from (3.13) that $P_v(x+te_i) = P_v(x) + te_i$ for $i = 1,2,3$. On the other hand, one can see from (3.22) that if $P_v(x)$ and $P_v(x) + te_i$ lies in $\hat{\mathcal{S}}^i_{\hat{z}_{\boldsymbol{v}(k)}}$, then $g(P_v(x)) = g(P_v(x+te_i))$. Thus, we have

$$F_{g,v}(x+te_i) = x + te_i + vg(P_v(x+te_i)) = F_{g,v}(x) + te_i$$

and the similar identity holds for $F_{g,u}$. It follows that for each $i = 1,2,3$ it holds

$$\lim_{t\to 0} \frac{F_{g,u}(F_{g,v}(x+te_i)) - F_{g,u}(F_{g,v}(x))}{t} = e_i \quad \text{whenever } P_v(x) \in \hat{\mathcal{S}}^i_{\hat{z}_{\boldsymbol{v}(k)}}. \tag{3.30}$$

The argument for D_4 is similar. We know (see (3.29)) that g has the following property on a line segment parallel to $\tilde{v}$ e.g. $\{P_v(x) + t\tilde{v}, t \in I\}$ which happens to lie in $\hat{\mathcal{S}}^4_{\hat{z}_{\boldsymbol{v}(k)}}$,

$$g(P_v(x) + t\tilde{v}) = g(P_v(x)) + t.$$

Now take a line segment in $\hat{S}^4_{\hat{z}_{\boldsymbol{v}(k)}}$ parallel to e_4. From (3.13) we know that

$$P_v(x+te_4) = P_v(x) - t\tilde{v}$$

and therefore

$$g(P_v(x+te_4)) = g(P_v(x)) - t.$$

Recalling that $v = (\tilde{v}, 1)$ we get

$$\begin{aligned} F_{g,v}(x + te_4) &= x + te_4 + vg(P_v(x + te_4)) \\ &= x + te_4 - tv + vg(P_v(x)) \\ &= F_{g,v}(x) - t\tilde{v}. \end{aligned} \tag{3.31}$$

At this point we need to finally choose q in the definition of $\hat{\mathcal{S}}^i_{\boldsymbol{v}(k)}$ (see (3.6)). For every $i \in \{1, 2, 3\}$ we have defined g in (3.21) on projection of line segments through $\hat{z}_{\boldsymbol{v}(k)}$ so that $g = -(\hat{z}_{\boldsymbol{v}(k)})_4$ and hence for every $x \in (\hat{z}_{\boldsymbol{v}(k)} + \mathbb{R}e_i) \cap \hat{S}^i_{\boldsymbol{v}(k)}$ we have

$$(F_{g,v}(x))_4 = (x + vg(P_v(x)))_4 = x_4 + v_4(-(\hat{z}_{\boldsymbol{v}(k)})_4) = x_4 - (\hat{z}_{\boldsymbol{v}(k)})_4 = 0$$

and hence

$$\big(F_{g,v}\big((\hat{z}_{\boldsymbol{v}(k)} + \mathbb{R}e_i) \cap \hat{S}^i_{\boldsymbol{v}(k)}\big)\big)_4 = 0.$$

Analogously for $i = 4$ we defined g in (3.24) so that

$$\big(F_{g,v}\big((\hat{z}_{\boldsymbol{v}(k)} + \mathbb{R}e_4) \cap \hat{S}^4_{\boldsymbol{v}(k)}\big)\big)_4 = 0$$

for all $\hat{z}_{\boldsymbol{v}(k)}$. Since for all $x \in \hat{S}^i_{\boldsymbol{v}(k)}$ we find a $y \in (\hat{z}_{\boldsymbol{v}(k)} + \mathbb{R}e_i) \cap \hat{S}^i_{\boldsymbol{v}(k)}$ such that $\|x - y\| \leq \hat{r}_k$ we see that

$$|F_{g,v}(y) - F_{g,v}(x)| < C\hat{r}_k$$

and so by Lipschitz continuity of P_u,

$$|P_u(F_{g,v}(y)) - P_u(F_{g,v}(x))| < C\hat{r}_k.$$

Therefore we find a $q \geq \frac{5}{4}$ which will now ensure that (note that there is q in the definition of $\hat{\mathcal{S}}^i_{\boldsymbol{v}(k)}$ but not in the definition of $\hat{S}^i_{\boldsymbol{v}(k)}$)

$$P_u(F_{g,v}(x)) \in \hat{\mathcal{S}}^i_{\boldsymbol{v}(k)} \text{ for every } x \in \hat{S}^i_{\boldsymbol{v}(k)}. \tag{3.32}$$

Now using (3.31), applying $F_{g,u}$ and using (3.29) with (3.32) we get

$$\begin{aligned} F_{g,u} \circ F_{g,v}(x + te_4) &= F_{g,v}(x) - t\tilde{v} + ug(P_u(F_{g,v}(x)) - t\tilde{v}) \\ &= F_{g,v}(x) - t\tilde{v} + ug(P_u(F_{g,v}(x)) - tu \\ &= F_{g,v}(x) + ug(P_u(F_{g,v}(x)) - te_4 \\ &= F_{g,u} \circ F_{g,v}(x) - te_4, \end{aligned} \tag{3.33}$$

where we used $u = (-\tilde{v}, 1)$. Now (3.33) easily gives us what we wanted to prove, i.e. $D_4 F_{g,u} \circ F_{g,v} = -e_4$.

Given this, it suffices to realize that

$$\bigcup_{\boldsymbol{v}(k) \in \mathbb{V}^k} \hat{S}^i_{\boldsymbol{v}(k)} = \{x \in \mathbb{R}^n : x_i \in [-1, 1] \setminus \hat{U}_{k-2},\ x_j \in \hat{U}_k \text{ for all } j \neq i\}$$

and that $P_v(\hat{S}^i_{\boldsymbol{v}(k)}) \subset \hat{\mathcal{S}}^i_{\boldsymbol{v}(k)}$ to see that (3.3) is satisfied. This ends the proof of the lemma. □

Lemma 3.5. *Let F be a C-bi-Lipschitz map defined on $Q(0, 1)$ that maps $\mathcal{K}_B$ onto $\mathcal{K}_B$ and C_B onto C_B. Then there exists a constant $\tilde{C} > 0$ such that for every $x \in Q(0, 1)$ we have*

$$\tilde{C}^{-1} \operatorname{dist}(x, \mathcal{K}_B) < \operatorname{dist}(F(x), \mathcal{K}_B) < \tilde{C} \operatorname{dist}(x, \mathcal{K}_B) \tag{3.34}$$

and

$$\tilde{C}^{-1}\operatorname{dist}(x, C_B) < \operatorname{dist}(F(x), C_B) < \tilde{C}\operatorname{dist}(x, C_B). \tag{3.35}$$

Proof. We prove the first inequality in (3.34) by contradiction. Assume that we have a sequence $\{x_k\}_{k=1}^{\infty}$ of points with the following property,

$$\operatorname{dist}(F(x_k), \mathcal{K}_B) < \tfrac{1}{k}\operatorname{dist}(x_k, \mathcal{K}_B).$$

Then applying F^{-1} to the points $F(x_k)$ and using the fact that F^{-1} is Lipschitz map which maps $\mathcal{K}_B$ onto $\mathcal{K}_B$ we get that

$$\operatorname{dist}(x_k, \mathcal{K}_B) = \operatorname{dist}\big(F^{-1}(F(x_k)), F^{-1}(\mathcal{K}_B)\big) \le C\operatorname{dist}(F(x_k), \mathcal{K}_B) < C\frac{1}{k}\operatorname{dist}(x_k, \mathcal{K}_B)$$

for all k, which is a contradiction. Therefore we see that there exists some constant $\tilde{C}_1 > 0$ such that

$$\tilde{C}_1^{-1}\operatorname{dist}(x, \mathcal{K}_B) \le \operatorname{dist}(F(x), \mathcal{K}_B) \quad \text{for all } x.$$

The second inequality in (3.34) is implied by the first and the fact that F^{-1} is a bi-Lipschitz mapping, which maps $\mathcal{K}_B$ onto $\mathcal{K}_B$. Thus we my find also a constant $\tilde{C}_2 > 0$ such that

$$\operatorname{dist}(F(x), \mathcal{K}_B) < \tilde{C}_2\operatorname{dist}(x, \mathcal{K}_B) \quad \text{for all } x.$$

The proof of the two inequalities in (3.35) goes similarly and thus we may find constants $\tilde{C}_3, \tilde{C}_4 > 0$ such that

$$\tilde{C}_3^{-1}\operatorname{dist}(x, C_B) < \operatorname{dist}(F(x), C_B) < \tilde{C}_4\operatorname{dist}(x, C_B) \quad \text{for all } x.$$

The claim follows now by taking $\tilde{C} = \max\{\tilde{C}_1, \tilde{C}_2, \tilde{C}_3, \tilde{C}_4\}$. □

Proof of Theorem 3.1. First we need to find a suitable Cantor set C_B. For this we need to assume that $\beta \ge \max\{6, \beta_1, \beta_2, \beta_3\}$ in the definition of C_B in subsection 2.6 where $\beta_1, \beta_2, \beta_3$ are described in the proof of Lemma 3.4. Taking this Cantor set C_B we may apply Lemma 3.4. From Lemma 3.4 we get a vector v, such that P_v is one-to-one on the set $\mathcal{K}_B$ and further the function $g(P_v(x)) = -x_4$ on $P_v(\mathcal{K}_B)$ has a Lipschitz extension on $X = \mathbb{R}^3 \times \{0\}$, which we have defined in the end of the the proof of Lemma 3.4. Define the vector $u = (-v_1, -v_2, -v_3, v_4)$ and recall that we have defined $F_{g,v} : \mathbb{R}^4 \to \mathbb{R}^4$ as

$$F_{g,v}(x) = x + vg(P_v(x)). \tag{3.36}$$

Then consider the image of a point $x \in \mathcal{K}_B$ for the map $F := F_{g,u} \circ F_{g,v}$. First we observe that

$$F_{g,v}(x) = x + vg(P_v(x)) = x - \frac{x_4}{v_4}v = P_v(x) \quad \text{for every } x \in \mathcal{K}_B.$$

Furthermore, it is easy to see that the projections P_v and P_u are identities when restricted to $X = P_v(\mathbb{R}^4) = P_u(\mathbb{R}^4)$, which gives us

$$P_u(F_{g,v}(x)) = P_u(P_v(x)) = P_v(x) \quad \text{for all } x \in \mathcal{K}_B.$$

Therefore, for each $x \in \mathcal{K}_B$ we can calculate

$$
\begin{aligned}
F_{g,u} \circ F_{g,v}(x) &= x + vg(P_v(x)) + ug(P_u(F_{g,v}(x))) \\
&= x + vg(P_v(x)) + ug(P_v(x)) \\
&= x - v\frac{x_4}{v_4} - u\frac{x_4}{v_4} \\
&= (x_1, x_2, x_3, -x_4).
\end{aligned} \tag{3.37}
$$

This means that $F_{g,u} \circ F_{g,v}$ is exactly the reflection in the last coordinate on $\mathcal{K}_B$ as in (3.1).

If we redefine g so that it is constant on a small ball B in X then we can find a point $x \in \mathbb{R}^4$, which is mapped to the center of B by $F_{g,v}$. The projection P_v is continuous and so $V = P_v^{-1}(\frac{1}{2}B)$ is open. Now we call

$$U := \{y \in V; P_u(y) \in \tfrac{1}{2}B\}$$

which is also an open neighbourhood of $P_v(x) = F_{g,v}(x)$. Then $W = F_{-g,v}(U)$ is an open neighbourhood of x mapped by $F_{g,v}$ onto U. Then $P_u(F_{g,v}(w)) \in B$ for all $w \in W$. Let us denote by λ the constant value of g on B, then we have

$$g(P_v(w)) = g(P_u(F_{g,v}(w))) = \lambda.$$

Hereby we see using (3.36) that

$$F_{g,u} \circ F_{g,v}(w) = w + vg(P_v(w)) + ug(P_u(F_{g,v}(w))) = w + \lambda v + \lambda u = w + 2\lambda e_4$$

for all $w \in W$ which is an open set containing x. Our mapping $f = F_{g,u} \circ F_{g,v}$ is a translation on V and the translation is obviously sense preserving. Now $F_{g,u} \circ F_{g,v}$ is a bi-Lipschitz map that can equal a translation everywhere on a ball and therefore must be sense-preserving. This ends the first part of the proof.

Next, it follows from Lemma 3.4 that if $N_F \in \mathbb{N}$ is arbitrary and $N_F < j \leq k$ then F maps each line segment I_i parallel to e_i which lies in

$$A_{i,j-N_F-1,k+N_F} = \{x \in \mathbb{R}^4 : x_i \in [-1,1] \setminus \hat{U}_{j-N_F-1},\, x_l \in \hat{U}_{k+N_F} \text{ for } l \neq i\}$$

to a line segments parallel to e_i as the derivative along the segment satisfies

$$D_i F(x) = \begin{cases} e_i & \text{if } i = 1,2,3 \\ -e_i & \text{if } i = 4 \end{cases}$$

for every $x \in A_{i,j-N_F-1,k+N_F}$. Therefore it suffices to show that there exists $N_F \in \mathbb{N}$ such that the image $F(I_i)$ of such a line segment I_i lies always in the set $A_{i,j-1,k}$.

Let us start by recalling from (2.7) that $\hat{U}_k = \bigcup_i \hat{I}_{i,k}$, where by choosing the center points of the intervals to be $\hat{z}_{i,k}$ we can write $\hat{I}_{i,k} = [\hat{z}_{i,k} - \hat{r}_k, \hat{z}_{i,k} + \hat{r}_k]$. Thus

$$\hat{U}_k \subset \bigcup_i [\hat{z}_{i,k} - 2\hat{r}_k, \hat{z}_{i,k} + 2\hat{r}_k].$$

This immediately gives that

$$\{y \in \mathbb{R} : \operatorname{dist}(y, \mathcal{C}_B) < \hat{r}_{k+1}\} \subset \hat{U}_k \subset \{y \in \mathbb{R} : \operatorname{dist}(y, \mathcal{C}_B) < 2\hat{r}_k\}, \tag{3.38}$$

and it follows that

$$A_{i,j-1,k} \supset \{x \in \mathbb{R}^4 : \operatorname{dist}(x_i, \mathcal{C}_B) > 2\hat{r}_{j-1},\, \operatorname{dist}(x_l, \mathcal{C}_B) < \hat{r}_{k+1} \text{ for } l \neq i\} =: \hat{A}_{i,j-1,k}.$$

Therefore, it is enough to show that there is $N_F \in \mathbb{N}$ such that $F(x) \in \hat{A}_{i,j-1,k}$ for every $x \in A_{i,j-N_F-1,k+N_F}$ whenever $N_F < j \leq k$.

Suppose that $N_F \in \mathbb{N}$ and assume that $x \in A_{i,j-N_F-1,k+N_F}$ where $N_F < j \leq k$. Then it follows from (3.38) that

$$\operatorname{dist}(x, C_B + \mathbb{R}e_i) < c_1 \hat{r}_{k+N_F+1} \quad \text{and} \quad \operatorname{dist}(x, C_B) \geq c_2^{-1} \hat{r}_{j-N_F-1}, \tag{3.39}$$

where the constants $c_1 > 0$ and $c_2 > 0$ depend only on the dimension $n = 4$ and on β. If we apply the bi-Lipschitz property of F to (3.39), and the fact that $F(C_B + \mathbb{R}e_i) = C_B + \mathbb{R}e_i$ and $F(C_B) = C_B$ we get

$$\operatorname{dist}(F(x), C_B + \mathbb{R}e_i) \leq C \operatorname{dist}(x, C_B + \mathbb{R}e_i) < Cc_1 \hat{r}_{k+N_F+1}$$

and with the help of Lemma 3.5

$$\operatorname{dist}(F(x), C_B) \geq C^{-1} \operatorname{dist}(x, C_B) \geq (Cc_2)^{-1} \hat{r}_{j-N_F-1},$$

where $C \geq 1$ stands for the bi-Lipschitz constant of F. Thus, if we choose $N_F \in \mathbb{N}$ such that $\hat{r}_{N_F} < \min\{(Cc_1)^{-1}, \frac{1}{3}(Cc_2)^{-1}\}$ and use the fact that $\hat{r}_k = 2^{-k-\beta k} = \hat{r}_1^k$ we have that

$$\operatorname{dist}(F(x), C_B + \mathbb{R}e_i) < \hat{r}_{k+1} \quad \text{and} \quad \operatorname{dist}(F(x), C_B) > 3\hat{r}_{j-1}. \tag{3.40}$$

On the other hand, if we apply the triangle inequality to the point $y = F(x)$ we get

$$3\hat{r}_{j-1} < \operatorname{dist}(y, C_B) \leq \operatorname{dist}(y_i, \mathcal{C}_B) + \operatorname{dist}(y, C_B + \mathbb{R}e_i) < \operatorname{dist}(y_i, \mathcal{C}_B) + \hat{r}_{k+1},$$

where y_i is the i-th coordinate of y. Thus, it follows that $\operatorname{dist}(y_i, \mathcal{C}_B) > 2\hat{r}_{j-1}$. Furthermore, as it follows from (3.40) that

$$\operatorname{dist}(y_l, \mathcal{C}_B) \leq \operatorname{dist}(F(x), C_B + \mathbb{R}e_i) < \hat{r}_{k+1} \quad \text{for each } l \neq i,$$

we get that $F(x) \in \hat{A}_{i,j-1,k}$ and the claim follows. □

4. The mapping S_t

The purpose of this section is to define a mapping which stretches C_B back onto C_A and has the properties listed in Lemma 4.1. We use the notation $\hat{U}_k$, $\hat{\mathcal{M}}_k$ and $\hat{\mathcal{A}}_k$ introduced in (2.7) and (2.8) and we recall that

$$C_B = \mathcal{C}_B \times \mathcal{C}_B \times \mathcal{C}_B \times \mathcal{C}_B = \bigcap_{k=1}^{\infty} \hat{\mathcal{M}}_k \,.$$

Lemma 4.1. *There exists a sense-preserving homeomorphisms $S_t : (-1,1)^4 \to (-1,1)^4$ such that:*

(i) *S_t maps C_B onto C_A and $S_t = S_q^{-1}$ on C_B.*
(ii) *Mapping S_t is locally Lipschitz on $(-1,1)^4 \setminus C_B$.*
(iii) *If L_i is a line parallel to x_i-axis with $L_i \cap (\hat{U}_k)^4 \neq \emptyset$ then*

$$|D_i S_t(x)| \leq C \frac{2^{\beta k}}{k^{\alpha+1}}$$

for every $x \in L_i \cap \big((\hat{U}_{k-1})^4 \setminus (\hat{U}_k)^4\big)$.

(iv) *If* $k \le j \le 3k+2$ *and* $x \in (\hat{U}_j \setminus \hat{U}_{j+1}) \times \big((\hat{U}_k)^3 \setminus (\hat{U}_{k+1})^3\big)$ *then*

$$|DS_t(x)| \le C2^{\beta j}.$$

The same holds for $x \in \big((\hat{U}_k)^3 \setminus (\hat{U}_{k+1})^3\big) \times (\hat{U}_j \setminus \hat{U}_{j+1})$ *and also for two other permutations of coordinates.*

(v) *If* $x \in \hat{U}_{3k+3} \times \big((\hat{U}_k)^3 \setminus (\hat{U}_{k+1})^3\big)$ *then*

$$|DS_t(x)| \le C2^{\beta(3k+3)}.$$

The same holds for $x \in \big((\hat{U}_k)^3 \setminus (\hat{U}_{k+1})^3\big) \times (\hat{U}_{3k+3})^3$ *and also for two other permutations of coordinates.*

Proof. In order to aid our construction, let us first recall and define some notation we will use. We recall that if $C_A = \mathcal{C}_A \times \mathcal{C}_A \times \mathcal{C}_A \times \mathcal{C}_A$ and $C_B = \mathcal{C}_B \times \mathcal{C}_B \times \mathcal{C}_B \times \mathcal{C}_B$ are the Cantor sets defined in subsection 2.5 and 2.6 then we may write

$$\mathcal{C}_A = \bigcap_{k=1}^{\infty} \bigcup_{i=1}^{2^k} I_{i,k} \quad \text{and} \quad \mathcal{C}_B = \bigcap_{k=1}^{\infty} \bigcup_{i=1}^{2^k} \hat{I}_{i,k}, \tag{4.1}$$

where the closed intervals $I_{i,k}$ and $\hat{I}_{i,k}$ have the lengths

$$\ell_k = \mathcal{L}^1(I_{i,k}) = 2^{-k}\Big(1 + \frac{1}{(k+1)^\alpha}\Big) \quad \text{and} \quad \hat{\ell}_k = \mathcal{L}^1(\hat{I}_{i,k}) = 2^{-k\beta-k+1}. \tag{4.2}$$

Moreover, we have $I_{i,k} \cap I_{j,k} = \emptyset$ for $i \neq j$, $I_{2i-1,k} \cup I_{2i,k} \subset I_{i,k-1}$ and $I_{i,k}$ lies more to the left than $I_{i+1,k}$ (similar properties holds also for the intervals $\hat{I}_{i,k}$).

Then there exists a natural function $t : \mathbb{R} \to \mathbb{R}$ which maps $\mathcal{C}_B$ onto $\mathcal{C}_A$. In fact t is a uniform limit of functions $t_k : \mathbb{R} \to \mathbb{R}$, $k = 0, 1, 2, \ldots$, such that

(1) $t_0(x) = x$,
(2) t_k maps each $\hat{I}_{i,k}$ onto $I_{i,k}$ linearly,
(3) t_k maps each of the three parts of $\hat{I}_{i,k-1} \setminus (\hat{I}_{2i-1,k} \cup \hat{I}_{2i,k})$ onto the corresponding parts of $I_{i,k-1} \setminus (I_{2i-1,k} \cup I_{2i,k})$ linearly, and
(4) $t_k = t_{k-1}$ outside $\bigcup_{i=1}^{2^k} \hat{I}_{i,k-1}$.

Note that then we have $t = q^{-1}$ where q is the function defined in subsection 2.7. It follows that $(t(x_1), t(x_2), t(x_3), t(x_4)) = S_q^{-1}(x)$.

The definition of the mapping S_t will make use of the standard frame-to-frame maps H^3_k, H^3 and H^4_k, H^4 described in Section 2.8. In a rough, intuitive sense we want a map that behaves very much like H^4_k on parts of the frame "far away" from $\mathcal{K}_B$, (i.e. in $\big((\hat{U}_{k-1})^4 \setminus (\hat{U}_k)^4\big) \setminus \hat{\mathcal{A}}_k$) and on hyperplanes in $\hat{\mathcal{A}}_k$ perpendicular to lines in $\mathcal{K}_B$ acts like the higher iterations of the frame-to-frame map H^3_{3k}. Our strategy is to define a map which equals (up to some isometric rotation) H^3_{3k} on each face of each cube in $(\hat{U}_k)^4 = \bigcup_{\boldsymbol{v}(k)\in\mathbb{V}^k} Q_{\boldsymbol{v}(k)}$. We extend this mapping into $(\hat{U}_k)^4 \setminus (\hat{U}_{k+1})^4$ simply as the frame to frame mapping H^4_k on $\big((\hat{U}_k)^4 \setminus (\hat{U}_{k+1})^4\big) \setminus \hat{\mathcal{A}}_k$. Inside the "tubes" of type $\big((\hat{U}_k)^4 \setminus (\hat{U}_{k+1})^4\big) \cap \hat{\mathcal{A}}_k$ we use a suitable convex combination of the maps defined on the faces to extend the map inside the frame.

We refer to the i-th canonical projection π_i as the linear map

$$\pi_i(x) = (x_1, \ldots, x_{i-1}, 0, x_{i+1}, \ldots, x_4).$$

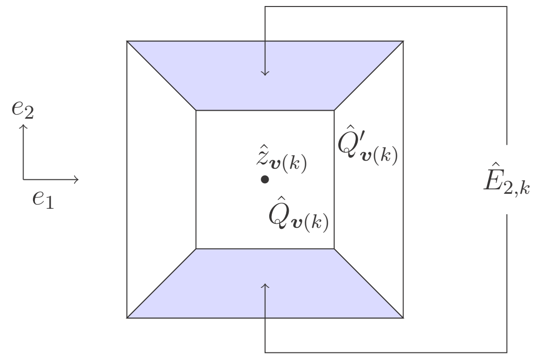


FIGURE 9. The part of the set $\hat{E}_{2,k}$ which lies inside the frame $\hat{Q}'_{\boldsymbol{v}(k)} \setminus \hat{Q}_{\boldsymbol{v}(k)}$ in two dimensions.

Take $i \in \{1,2,3,4\}$. We will also denote $\tilde{x}^i = (x_1, \dots x_{i-1}, x_{i+1}, \dots x_4) \in \mathbb{R}^3$. Using this notation we define the linear isomorphic isometry $L_i : \mathbb{R}^{i-1} \times \{0\} \times \mathbb{R}^{4-i} \to \mathbb{R}^3$ defined as $L_i(\pi_i(x)) = \tilde{x}^i$. Furthermore, we define

$$H_k^{3,i}(x) = L_i^{-1} \circ H_k^3 \circ L_i \circ \pi_i(x),$$
$$H^{3,i}(x) = L_i^{-1} \circ H^3 \circ L_i \circ \pi_i(x).$$

For a point $x \in (\hat{U}_{k-1})^4 \setminus (\hat{U}_k)^4$ we will define the functions

$$d_{i,k}(x) = \frac{\min\{|x_i - (\hat{z}_{\boldsymbol{v}(k)})_i| : \boldsymbol{v}(k) \in \mathbb{V}^k\} - \hat{r}_k}{\frac{1}{2}\hat{r}_{k-1} - \hat{r}_k}.$$

The set $\hat{I}_{i,k-1} \setminus (\hat{I}_{2i-1,k} \cup \hat{I}_{2i,k})$ of intervals, whose union is $\hat{U}_{k-1} \setminus \hat{U}_k$, can be decomposed to four closed (maximal) intervals with disjoint interiors so that each function $d_{i,k}$ is linear in x_i on each of these four subintervals. Further, if $x \in (\hat{U}_{k-1})^4$ and $x_i \in \hat{U}_{k-1} \setminus \hat{U}_k$ then we have

$$d_{i,k}(x) = 4\frac{\operatorname{dist}(x_i, \hat{U}_k)}{\hat{\ell}_{k-1} - 2\hat{\ell}_k} \tag{4.3}$$

which takes values between 0 and 1. Using these functions we can divide the frame into the parts where we are furthest from its center in the direction e_i, which are the sets (see Fig. 9.)

$$\begin{aligned}\hat{E}_{i,k} &= \big\{x \in (\hat{U}_{k-1})^4 \setminus (\hat{U}_k)^4 : \min_{\boldsymbol{v}(k)\in\mathbb{V}^k} |x_i - (\hat{z}_{\boldsymbol{v}(k)})_i| \geq \min_{\boldsymbol{v}(k)\in\mathbb{V}^k} |x_j - (\hat{z}_{\boldsymbol{v}(k)})_j| \text{ for all } j \neq i\big\}\\ &= \{x \in (\hat{U}_{k-1})^4 \setminus (\hat{U}_k)^4 : d_{i,k}(x) \geq d_{j,k}(x) \text{ for all } j \neq i\}.\end{aligned}$$

For technical reasons it is also convenient to define the corresponding sets $E_{i,k}$ in the target, i.e.,

$$E_{i,k} = \big\{y \in (U_{k-1})^4 \setminus (U_k)^4 : \min_{\boldsymbol{v}(k)\in\mathbb{V}^k} |y_i - (z_{\boldsymbol{v}(k)})_i| \geq \min_{\boldsymbol{v}(k)\in\mathbb{V}^k} |y_j - (z_{\boldsymbol{v}(k)})_j| \text{ for all } j \neq i\big\}.$$

We will next use the convex combinations of the maps $H_{3k-3}^{3,i}$ and $H_{3k}^{3,i}$ in the sets $\hat{E}_{i,k}$ together with some correction mapping to define the mapping S_t.

We cut the set $\hat{E}_{i,k}$ into hyperplane slices with hyperplanes perpendicular to e_i, $\hat{E}_{i,k}\cap\{x_i=c\}$. On these planes we apply $L_i\circ\pi_i$, which shifts it onto a corresponding hyperplane $\{x_i = 0\}$ and then rotates it onto $\mathbb{R}^3$. Then we can apply a convex combination of the 3-dimensional frame-to-frame maps

$$A_{i,k}(x) := d_{i,k}(x)H^3_{3k-3}(x) + (1-d_{i,k}(x))H^3_{3k}(x)$$

and then reverse the rotation using L_i^{-1}. Now we shift the hyperplane into the right position by adjusting the i-coordinate so that it corresponds to the i-th coordinate of the frame-to-frame map, i.e. $((H^4_k)(x))_i = t(x_i)$. In summary, we define

$$S_t(x) = \underbrace{d_{i,k}(x)H^{3,i}_{3k-3}(x) + (1-d_{i,k}(x))H^{3,i}_{3k}(x) + t(x_i)e_i}_{=A_{i,k}(x)+t(x_i)e_i} \text{ for } x\in\hat{E}_{i,k}. \tag{4.4}$$

We need to show that S_t defines a homeomorphism which satisfies all the conditions (i)-(v) in Lemma 4.1.

Step 1: Proving that S_t is a homeomorphism. First we show that (4.4) yields a homeomorphism. The first observation in this direction is that S_t maps $(\hat{U}_{k-1})^4\setminus(\hat{U}_k)^4$ onto $(U_k)^4\setminus(U_{k+1})^4$ for each $k\in\mathbb{N}$. To see this we observe that in the expression

$$S_t(x) = A_{i,k}(x) + t(x_i)e_i \text{ for every } x\in\hat{E}_{i,k}$$

the mapping $A_{i,k}$ maps each hyperplane in $\hat{E}_{i,k}$ perpendicular to e_i homeomorphically to a hyperplane in $E_{i,k}$ perpendicular to e_i. Moreover, at the end of subsection 2.8 we have shown that for each fixed $\alpha\in(0,1)$ the mapping

$$\alpha H^3_{3k-3}(x) + (1-\alpha)H^3_{3k}(x)$$

is a homeomorphism in $\mathbb{R}^3$ and thus $A_{i,k}(x)$ on the hyperplane is a homeomorphism. Furthermore, we have that $t(\hat{U}_k) = U_k$ for every k. Thus, it is quite easy to see that S_t actually maps each set $\hat{E}_{i,k}$ onto $E_{i,k}$ homeomorphically.

Next we will show that S_t defines a homeomorphism from $(\hat{U}_{k-1})^4\setminus(\hat{U}_k)^4$ onto $(U_{k-1})^4\setminus(U_k)^4$. For this it suffices to show that in the critical set

$$\hat{E}_{i,k}\cap\hat{E}_{j,k} = \bigcup_{\boldsymbol{v}(k)\in\mathbb{V}^k}\big\{x\in\hat{Q}'_{\boldsymbol{v}(k)}\setminus\hat{Q}_{\boldsymbol{v}(k)} : |x-\hat{z}_{\boldsymbol{v}(k)}| = |x_i-(\hat{z}_{\boldsymbol{v}(k)})_i| = |x_j-(\hat{z}_{\boldsymbol{v}(k)})_j|\big\}$$

the expressions in (4.4) coincide. This gives us that S_t is continuous in $(\hat{U}_{k-1})^4\setminus(\hat{U}_k)^4$. For this, let us assume that $x\in\hat{E}_{i,k}\cap\hat{E}_{j,k}$. Then one can show that

$$(H^{3,i}_{3k-3}(x))_j = (H^{3,i}_{3k}(x))_j = t(x_j) \quad\text{and}\quad (H^{3,i}_{3k-3}(x))_i = (H^{3,i}_{3k}(x))_i = t(x_i).$$

Thus, we get that

$$\begin{aligned} H^{3,i}_{3k-3}(x) &= H^{3,j}_{3k-3}(x) + (H^{3,i}_{3k-3}(x))_je_j - (H^{3,j}_{3k-3}(x))_ie_i \\ &= H^{3,j}_{3k-3}(x) + t(x_j)e_j - t(x_i)e_i. \end{aligned}$$

and

$$\begin{aligned} H^{3,i}_{3k}(x) &= H^{3,j}_{3k}(x) + (H^{3,i}_{3k}(x))_je_j - (H^{3,j}_{3k}(x))_ie_i \\ &= H^{3,j}_{3k}(x) + t(x_j)e_j - t(x_i)e_i. \end{aligned}$$

Thus, it follows that

$$\begin{aligned}A_{i,k}(x)+t(x_i)e_i &= d_{i,k}(x)H^{3,i}_{3k-3}(x)+(1-d_{i,k}(x))H^{3,i}_{3k}(x)+t(x_i)e_i\\ &= d_{j,k}(x)H^{3,i}_{3k-3}(x)+(1-d_{j,k}(x))H^{3,i}_{3k}(x)+t(x_i)e_i\\ &= d_{j,k}(x)H^{3,j}_{3k-3}(x)+(1-d_{j,k}(x))H^{3,j}_{3k}(x)+t(x_j)e_j\\ &= A_{j,k}(x)+t(x_j)e_j,\end{aligned}$$

as we wanted.

We have now shown that S_t defines a homeomorphism on each set $(\hat{U}_{k-1})^4\setminus(\hat{U}_k)^4$. Next, we will show that S_t defines a homeomorphism on $(-1,1)^4\setminus C_B$. Because S_t is a homeomorphisms on each set $(\hat{U}_{k-1})^4\setminus(\hat{U}_k)^4$ and these sets are pairwise disjoint it suffices to show that in the critical set

$$C_k:=\overline{(\hat{U}_{k-1})^4\setminus(\hat{U}_k)^4}\cap\overline{(\hat{U}_k)^4\setminus(\hat{U}_{k+1})^4}$$

the expressions in (4.4) coincide. For this, it suffices to prove that for every $i=1,2,3,4$ these expressions coincide along the lines

$$l^i_{\boldsymbol{v}(k)}=\hat{z}_{\boldsymbol{v}(k)}+\hat{x}_i+se_i,\quad s\in\mathbb{R},$$

where $\hat{x}_i=(x_1,\dots,x_{i-1},0,x_{i+1},\dots,x_4)$ with $|x_j|<\hat{r}_k$ for every $j\neq i$. However, this is clear as

$$\begin{aligned}\lim_{s\to\hat{r}_k^+}S_t(\hat{z}_{\boldsymbol{v}(k)}+\hat{x}_i+se_i) &= \lim_{s\to\hat{r}_k^+}\Big(\overbrace{\big(d_{i,k}H^{3,i}_{3k-3}+(1-d_{i,k})H^{3,i}_{3k}\big)(\hat{z}_{\boldsymbol{v}(k)}+\hat{x}_i+se_i)}^{=A_{i,k}(\hat{z}_{\boldsymbol{v}(k)}+\hat{x}_i+se_i)}+t(s)e_i\Big)\\ &= H^{3,i}_{3k}\big(\hat{z}_{\boldsymbol{v}(k)}+\hat{x}_i+\hat{r}_ke_i\big)+t(\hat{r}_k)e_i\\ &= \lim_{s\to\hat{r}_k^-}\Big(\big(d_{i,k+1}H^{3,i}_{3k}+(1-d_{i,k+1})H^{3,i}_{3k+1}\big)(\hat{z}_{\boldsymbol{v}(k)}+\hat{x}_i+se_i)+t(s)e_i\Big)\\ &= \lim_{s\to\hat{r}_k^-}S_t(\hat{z}_{\boldsymbol{v}(k)}+\hat{x}_i+se_i),\end{aligned}$$

thus we have shown that S_t defines a homeomorphism on $(-1,1)^4\setminus C_B$.

Finally, since S_t is a homeomorphism on all frames that sends frames to frames, S_t is extended homeomorphically as $S_t(x)=S_q^{-1}(x)=(t(x_1),t(x_2),t(x_3),t(x_4))$ to C_B. Especially, S_t will then take C_B onto C_A, and thus (i) follows. It is also easy to see that for a fixed k the mapping H^3_{3k} is Lipschitz and hence S_t defined by (4.4) is a locally Lipschitz mappings on on $(-1,1)^4\setminus C_B$ which implies (ii).

Step 2: Calculating the derivatives of S_t. We now calculate the derivative of the mapping S_t on $\hat{U}^4_{k-1}\setminus\hat{U}^4_k$. More precisely, we want to verify the conditions (iii)-(v). In the following calculations we will rely on (2.13) to calculate the derivative.

Step 2A: Proving the condition (iii). Suppose that L_i is a line parallel to x_i-axis with $L_i\cap(\hat{U}_k)^4\neq\emptyset$. Then it follows that $L_i\cap\big((\hat{U}_{k-1})^4\setminus(\hat{U}_k)^4\big)\subset\hat{E}_{i,k}$.

(1) Let us first assume that $x\in L_i\cap\hat{E}_{i,k}$ with $x_i\in\hat{U}_{k-1}\setminus\hat{U}_k$ and $\tilde{x}^i\in(\hat{U}_{k-1})^3\setminus(\hat{U}_{3k-3})^3$. In this case $A_{i,k}$ is a constant function in x_i-direction and therefore

$$|D_iS_t(x)|=|t'(x_i)|\leq\frac{a_{k-1}-a_k}{b_{k-1}-b_k}\leq C\frac{2^{k\beta}}{k^{\alpha+1}}.$$

(2) Let us next assume that $x \in L_i \cap \hat{E}_{i,k}$ with $x_i \in \hat{U}_{k-1} \setminus \hat{U}_k$ and $\tilde{x}^i \in (\hat{U}_{3k-3})^3$. We recall that the maps $H^{3,i}_{3k-3}$ and $H^{3,i}_{3k}$ are independent on x_i which implies

$$D_i H^{3,i}_{3k-3}(x) = 0 \text{ and } D_i H^{3,i}_{3k}(x) = 0$$

and by the construction of mappings H_{3k} we easily obtain

$$\|H^{3,i}_{3k-3} - H^{3,i}_{3k}\| \leq 2^{-3k+4}.$$

On the other hand by applying (4.3) we may conclude that

$$|D_i d_{i,k}(x)| \leq \frac{4}{\hat{\ell}_{k-1} - 2\hat{\ell}_k} \leq \frac{C}{\frac{1}{2}\hat{r}_{k-1} - \hat{r}_k}. \tag{4.5}$$

By combining these facts we get

$$\begin{aligned}
|D_i S_t(x)| &\leq |D_i A_{i,k}(x)| + |t'(x_i)| \\
&\leq |D_i d_{i,k}(x)||H^{3,i}_{3k-3}(x) - H^{3,i}_{3k}(x)| + |t'(x_i)| \\
&\leq C\frac{\|H^{3,i}_{3k-3} - H^{3,i}_{3k}\|}{\frac{1}{2}\hat{r}_{k-1} - \hat{r}_k} + \frac{a_{k-1} - a_k}{b_{k-1} - b_k} \\
&\leq C\frac{2^{\beta k}}{2^{3k-4}} + C\frac{2^{k\beta}}{k^{\alpha+1}} \leq C\frac{2^{k\beta}}{k^{\alpha+1}},
\end{aligned}$$

as we wanted. Now (1) and (2) together will give us (iii).

Step 2B: Proving the condition (iv). Let us next assume that $x \in (\hat{U}_j \setminus \hat{U}_{j+1}) \times \big((\hat{U}_k)^3 \setminus (\hat{U}_{k+1})^3\big)$ with $k \leq j \leq 3k+2$. The case $j = k$ is easy to deal with and therefore we may assume that $j > k$. In this case we have that $x \in \hat{E}_{i,k}$ for some $i \neq 1$. With the help of (4.5) we easily obtain

$$|Dd_{i,k}(x)| \leq C \max\{2^{\beta k}, 2^{\beta j}\} \text{ and } |d_{i,k}(x)| \leq 1 \quad \text{for every } x.$$

Moreover, we also have $\max\{H^{3,i}_{3k-2}(x), H^{3,i}_{3k}(x)\} \leq 1$ for every x. As $x \notin (\hat{U}_{j+1})^4$ we easily obtain $H^{3,i}_{3k}(x) = H^{3,i}_{j+1}(x)$ and thus using (2.12) that

$$|DH^{3,i}_{3k}(x)| \leq C2^{\beta j}.$$

Thus, it follows from (4.4) and (2.12) that for every $l \neq i$

$$\begin{aligned}
|D_l S_t(x)| &\leq |Dd_{i,k}(x)|\big(|H^{3,i}_{3k-2}(x)| + |H^{3,i}_{3k}(x)|\big) + |d_{i,k}(x)|\big(|D_l H^{3,i}_{3k-2}(x)| + |D_l H^{3,i}_{3k}(x)|\big) \\
&\leq C\max\{2^{\beta k}, 2^{\beta j}\} + |D_l H^{3,i}_{3k-2}(x)| + |D_l H^{3,i}_{3k}(x)| \\
&\leq C\max\{2^{\beta k}, 2^{\beta j}\} + C2^{\beta j} \leq C2^{\beta j}.
\end{aligned}$$

On the other hand, it follows from the step 2A that $|D_i S_t(x)| \leq C\frac{2^{\beta k}}{k^{\alpha+1}}$. Thus, because $j > k$ we may estimate

$$|DS_t(x)| \leq C2^{\beta j}.$$

There is no difference in the proof for $x \in \big((\hat{U}_k)^3 \setminus (\hat{U}_{k+1})^3\big) \times (\hat{U}_j \setminus \hat{U}_{j+1})$ and also for two other permutations of coordinates, and thus this ends the proof of (iv).

Step 2C: Proving the condition (v). Finally, assume that $x \in \hat{U}_{3k+3} \times \big((\hat{U}_k)^3 \setminus (\hat{U}_{k+1})^3\big)$. Again, we have that $x \in \hat{E}_{i,k}$ for some $i \neq 1$. By applying (4.4) and (2.14) we have

$$|D_l S_t(x)| \leq C \max\{2^{\beta(3k+3)}, 2^{\beta k}\} \leq C 2^{\beta(3k+3)} \quad \text{for every } l \neq i.$$

On the other hand, it follows from the step 2A that $|D_i S_t(x)| \leq C \frac{2^{\beta k}}{k^{\alpha+1}}$, and thus we conclude

$$|DS_t(x)| \leq C 2^{\beta(3k+3)}.$$

There is no difference in the proof for the other permutations of coordinates, and thus this ends the proof of (v).

□

5. Proof of Theorem 1.2 for $n = 4$

We will now define the mapping $f : (-1,1)^4 \to (-1,1)^4$ by

$$f = S_t \circ F \circ S_q\,.$$

Let us first remark that as a composition of three sense-preserving homeomorphisms S_q, F and S_t the mapping f is obviously a sense-preserving homeomorphism.

5.1. **The sign of the Jacobian:** We need to show that $J_f > 0$ on a set of positive measure and $J_f < 0$ on a set of positive measure. We know that S_q and F are Lipschitz maps, and by Lemma 4.1 (ii) that S_t is locally Lipschitz outside the set $C_B = (F \circ S_q)(C_A)$ and hence f is locally Lipschitz outside of C_A. Therefore f is a sense-preserving homeomorphism which is locally Lipschitz there and hence $J_f \geq 0$ outside of C_A (see e.g. [24]). We may also require that J_f is not identically zero on $\{x : J_f \geq 0\}$ because otherwise by [22] f would not satisfy Lusin's condition (N) on this set which cannot happen for a locally Lipschitz map, see e.g. [23, Theorem 4.2]. Hence $\mathcal{L}^4(\{x : J_f > 0\}) > 0$.

Now we show that $J_f(x) < 0$ for almost every $x \in C_A$. For this let us fix $x \in C_A$. If q and t are the functions in the definitions of homeomorphisms S_q and S_t we may observe that for every $x \in C_A$ we have

$$\begin{aligned} f(x) &= (S_t \circ F \circ S_q)(x) = (S_t \circ F)(q(x_1), q(x_2), q(x_3), q(x_4)) \\ &= S_t\big(q(x_1), q(x_2), q(x_3), -q(x_4)\big) = S_t\big(q(x_1), q(x_2), q(x_3), q(-x_4)\big) \\ &= \big(t(q(x_1)), t(q(x_2)), t(q(x_3)), t(q(-x_4))\big) = (x_1, x_2, x_3, -x_4)\,. \end{aligned} \tag{5.1}$$

Here we have used the following facts in the given order:

(i) $S_q(x) \in C_B$ for every $x \in C_A$,
(ii) $F(z) = (z_1, z_2, z_3, -z_4)$ for every $z \in C_B$,
(iii) the function q is odd, i.e. $q(-s) = -q(s)$ for every $s \in (-1,1)$,
(iv) if $x_4 \in \mathcal{C}_B$ then also $-x_4 \in \mathcal{C}_B$,
(v) $S_t(x) = (t(x_1, t(x_2), t(x_3), t(x_4))$ on C_B by Lemma 4.1 (i), and
(vi) $t = q^{-1}$.

It follows that at the points of density of C_A we know that the approximative derivate equals to the reflection in the last coordinate and hence the determinant of this matrix is -1. Once we show that f is Sobolev mapping we will know that its distributional derivative equals to approximative derivative a.e. and hence $J_f(x) = -1$ a.e. on C_A.

5.2. **ACL condition:** To verify the ACL-condition for f let us suppose that L is a line segment parallel to x_i-axis and consider the following cases:

Case 1: Suppose first that $L \cap C_A = \emptyset$. We know that both mappings S_q and F are Lipschitz maps, and by Lemma 4.1 (ii) that S_t is locally Lipschitz outside the set $C_B = (F \circ S_q)(C_A)$. Thus, the mapping $f = S_t \circ F \circ S_q$ is locally Lipschitz outside the set $\mathcal{K}_A$. It follows that f is Lipschitz and hence also absolutely continuous on the segment L.

Case 2: Suppose next that $L \cap C_A \neq \emptyset$, which means that $L \subset \mathcal{K}_A$. The line L decomposes into the part of L in C_A, and segments, which are mapped by f onto segments. On the parts of lines L intersecting C_A we use (5.1) to see that f is in fact 1-Lipschitz continuous on $L \cap C_A$. Now it remains to consider $L \setminus C_A$.

We fix $k \in \mathbb{N}$ and use the fact that

$$L \cap \big((U_k)^4 \setminus (U_{k+1})^4\big) = \bigcup_{\boldsymbol{v}(k) \in \mathbb{V}^k} L \cap (Q'_{\boldsymbol{v}(k)} \setminus Q_{\boldsymbol{v}(k)}).$$

Further, $L \cap (Q'_{\boldsymbol{v}(k)} \setminus Q_{\boldsymbol{v}(k)})$ is either empty or made up of two segments $L^1_{\boldsymbol{v}(k)}, L^2_{\boldsymbol{v}(k)}$ (recall that we assume now $L \cap C_A \neq \emptyset$). Each of these segments has length $\frac{1}{2} r_{k-1} - r_k$, which is squeezed by S_q into a segment parallel to x_i of length $\frac{1}{2}\hat{r}_{k-1} - \hat{r}_k$. We then apply the mapping F, which merely reflects the segment in the last variable (see (3.1)). Finally we apply the mapping S_t which maps each of the segments $F(S_q(L^1_{\boldsymbol{v}(k)}))$ and $F(S_q(L^2_{\boldsymbol{v}(k)}))$ onto a segment. Since $D_i f$ is constant on $L^1_{\boldsymbol{v}(k)}$ and $L^2_{\boldsymbol{v}(k)}$, we have that f maps each segment to a segment at constant speed. Therefore the restriction of f to each segment is Lipschitz. Then we can estimate the length of the image of the segment using Lemma 4.1 (iii) as follows

$$\begin{aligned}\mathcal{H}^1(f(L^1_{\boldsymbol{v}(k)})) = \mathcal{H}^1(f(L^2_{\boldsymbol{v}(k)})) &= |D_i S_t(x)| \big(\tfrac{1}{2}\hat{r}_{k-1} - \hat{r}_k\big) \\ &\le C \frac{2^{\beta k}}{k^{\alpha+1}} (2^{-k} 2^{-\beta k}) \le C\big(\tfrac{1}{2} r_{k-1} - r_k\big).\end{aligned}$$

The length of each segment has increased by no more than a factor of C. Thus we see that the restriction of f to $L \setminus C_A$ is Lipschitz continuous and hence it is Lipschitz on the whole L and therefore absolutely continuous on L.

5.3. **Sobolev regularity of the mapping:** We would like to estimate

$$\int_{(-1,1)^4} |Df(x)|^p \, \mathrm{d}x \le C \sum_{i=1}^4 \int_{(-1,1)^4} |D_i f(x)|^p \, \mathrm{d}x,$$

where $D_i f$ denotes the derivative with respect to x_i coordinate. Without loss of generality it is enough to estimate

$$\int_{(-1,1)^4} |D_1 f(x)|^p \, \mathrm{d}x = \int_{(-1,1)^3} \int_{-1}^1 |D_1 f(x_1, \tilde{x})|^p \, \mathrm{d}x_1 \, \mathrm{d}\tilde{x}, \tag{5.2}$$

where $\tilde{x} = (x_2, x_3, x_4)$ (derivatives in the other directions can be estimated analogously). For this, let us recall that the Cantor type sets C_A and C_B were constructed as the intersections of the sets

$$\bigcup_{\boldsymbol{v}(k)\in\mathbb{V}^k} Q_{\boldsymbol{v}(k)} = U_k \times U_k \times U_k \times U_k \quad \text{and} \quad \bigcup_{\boldsymbol{v}(k)\in\mathbb{V}^k} \hat{Q}_{\boldsymbol{v}(k)} = \hat{U}_k \times \hat{U}_k \times \hat{U}_k \times \hat{U}_k,$$

where $U_k = \bigcup_{i=1}^{2^k} I_{i,k}$ and $\hat{U}_k = \bigcup_{i=1}^{2^k} \hat{I}_{i,k}$. Moreover, recall also that

$$P_k := U_k \times U_k \times U_k \quad \text{and} \quad \hat{P}_k := \hat{U}_k \times \hat{U}_k \times \hat{U}_k.$$

Then the sets P_k and $\hat{P}_k$ are formed by 2^{3k} cubes.

Let us consider several possibilities. If $\tilde{x} \in \mathcal{C}_A \times \mathcal{C}_A \times \mathcal{C}_A$ then it is easy to see that f restricted to the line $[-1,1] \times \{\tilde{x}\}$ is in fact Lipschitz as it was explained at the end of subsection 5.2. It thus remains to estimate the integral (5.2) for $\tilde{x}$ in the sets $P_k \setminus P_{k+1}$. Because the mapping f is locally Lipschitz on $[-1,1]^4 \setminus C_A$ it suffices to analyze the mapping only near the set C_A, i.e. on the set $U_{k_0}^4$. For this fix now the exponent $p \in [1,2)$ and put

$$\alpha = \frac{2p}{2-p},$$

and β large enough for Theorem 3.1. Then we may find an index $k_0 \geq 4N_F + 5$, where $N_F \in \mathbb{N}$ is from Theorem 3.1, large enough so that

$$\max\{2^{-kp\beta/2} k^{(p-1)(\alpha+1)}, 2^{-p\beta}(\tfrac{k+1}{k})^{\alpha}\} < 1 \quad \text{for all } k \geq k_0. \tag{5.3}$$

Let us then fix $k \geq k_0$ and suppose that $\tilde{x} \in P_k \setminus P_{k+1}$. We will define the following divisions of the segment $L(\tilde{x}) = L := [-1,1] \times \{\tilde{x}\}$ according to $x_1 \in [-1,1]$ into the following sets

$$L_j = \{(x_1, \tilde{x}) : x_1 \in U_j \setminus U_{j+1}\} \text{ and } L_0 = \{(x_1, \tilde{x}) : x_1 \in \mathcal{C}_A\}.$$

The aim of the following calculations is to prove the estimate (5.16) below.

Case 1: Consider first those parts of the line segment L which are far away from the set C_A. More precisely, suppose that $k \geq k_0$ and

$$x \in L_j \quad \text{with } j = 1, \dots, k - 2N_F - 3.$$

First we observe that S_q maps the line segment L_j which is parallel to x_1-axis to a line segment L_j^1 which is also parallel to x_1-axis and lies inside the set $(\hat{U}_j \setminus \hat{U}_{j+1}) \times (\hat{P}_k \setminus \hat{P}_{k+1})$. Furthermore, we may estimate the derivative of S_q in the x_1-direction as

$$|D_1 S_q(x)| \leq C 2^{-\beta j} j^{\alpha+1}, \tag{5.4}$$

(see subsection 2.7).

Next we observe that

$$\begin{aligned} L_j^1 \subset ([-1,1] \setminus \hat{U}_{(j+N_F+2)-N_F-1}) \times \big((\hat{U}_{(k-N_F)+N_F})^3 \setminus (\hat{U}_{(k-N_F+1)+N_F})^3\big) \\ \subset A_{1,(j+N_F+2)-N_F-1,(k-N_F+1)+N_F}, \end{aligned}$$

where $N_F < j + N_F + 2 \leq k - N_F + 1$, and thus it follows from Theorem 3.1 that the bi-Lipschitz map F maps L_j^1 to a line segment L_j^2 parallel to x_1-axis such that

$$L_j^2 \subset A_{1,(j+N_F+2)-1,k-N_F}.$$

Moreover, because F is a bi-Lipschitz map, we have

$$|D_1 F(S_q(x))| \leq \mathrm{Lip}(F) \quad \text{for a.e. } x \in L_j, \tag{5.5}$$

where $\operatorname{Lip}(F)$ stands for the Lipschitz constant of the mapping F.

Finally, because $F(S_q(L_j)) = L_j^2$ is a line segment parallel to x_1-axis which is contained in set $[-1,1]^4 \setminus (\hat{U}_{j+N_F+1})^4$ it follows from Lemma 4.1 (iii) that

$$|D_1 S_t(F(S_q(x)))| \leq C \max_{1\leq l\leq j+N_F+1} 2^{\beta l} l^{-(\alpha+1)} \leq C 2^{\beta j} j^{-(\alpha+1)} \tag{5.6}$$

with C independent of j, k.

If we now put together the estimates (5.4), (5.5) and (5.6) the chain rule gives us

$$|D_1 f(x)| \leq C \quad \text{for a.e. } x \in L_j, \tag{5.7}$$

where C is an absolute constant.

Case 2: Let us next assume that

$$x \in L_j \quad \text{with } k - 2N_F - 3 < j \leq 3k - 3N_F - 3.$$

Again S_q maps the line segment L_j which is parallel to the x_1-axis to a line segment L_j^1 which is also parallel to x_1-axis and lies inside the set $(\hat{U}_j \setminus \hat{U}_{j+1}) \times (\hat{P}_k \setminus \hat{P}_{k+1})$. Furthermore, we have

$$|D_1 S_q(x)| \leq C 2^{-\beta j} j^{\alpha+1}, \tag{5.8}$$

with C independent of j, k.

Next, we recall again that

$$|D_1 F(S_q(x))| \leq \operatorname{Lip}(F) \quad \text{for a.e. } x \in L_j. \tag{5.9}$$

Moreover, it follows from the assumption $j \geq k - 2N_F - 2$ that

$$\begin{aligned} L_j^1 &\subset \{x \in [-1,1]^4 : \operatorname{dist}(x, \mathcal{K}_B) > \min\{\hat{r}_j - \hat{r}_{j+1}, \hat{r}_k - \hat{r}_{k+1}\}\} \\ &\subset \big\{x \in [-1,1]^4 : \operatorname{dist}(x, \mathcal{K}_B) > \min\{2^{-\beta(j+1)-(j+1)}, 2^{-\beta(k+1)-(k+1)}\}\big\} \\ &\subset \big\{x \in [-1,1]^4 : \operatorname{dist}(x, \mathcal{K}_B) > \min\{2^{-\beta(j+1)-(j+1)}, 2^{-\beta(j+2N_F+3)-(j+2N_F+3)}\}\big\} \\ &\subset \big\{x \in [-1,1]^4 : \operatorname{dist}(x, \mathcal{K}_B) > 2^{-\beta(j+2N_F+3)-(j+2N_F+3)}\big\}. \end{aligned}$$

Suppose now that $\tilde{C} > 0$ is the constant given by Lemma 3.5.

We may assume that $N_F \in \mathbb{N}$ is so large that $\tilde{C}^{-1} > 2^{3\beta+1} 2^{-\beta N_F - N_F}$. We may assume this because if Theorem 3.1 holds for a certain N_F, then it immediately holds for any $\tilde{N}_F \geq N_F$.

Then it follows from Lemma 3.5 and from the inclusion above that

$$\begin{aligned} F(L_j^1) &\Subset \big\{x \in [-1,1]^4 : \operatorname{dist}(x, \mathcal{K}_B) > \tilde{C}^{-1} 2^{-\beta(j+2N_F+3)-(j+2N_F+3)}\big\} \\ &\subset \big\{x \in [-1,1]^4 : \operatorname{dist}(x, \mathcal{K}_B) > 2^{-\beta(j+3N_F)-(j+3N_F)}\big\}. \end{aligned}$$

Thus, we have that $F(L_j^1)$ is contained in the following union of four sets

$$\begin{aligned} F(L_j^1) \subset \Bigg(&([-1,1] \setminus \hat{U}_{j+3N_F}) \times \big([-1,1]^3 \setminus (\hat{U}_{j+3N_F})^3\big) \Bigg) \cup \\ &\cdots \cup \Bigg(\big([-1,1]^3 \setminus (\hat{U}_{j+3N_F})^3\big) \times ([-1,1] \setminus \hat{U}_{j+3N_F}) \Bigg). \end{aligned}$$

Without loss of generality suppose that

$$F(L_j^1) \subset ([-1,1] \setminus \hat{U}_{j+3N_F}) \times \big([-1,1]^3 \setminus (\hat{U}_{j+3N_F})^3\big).$$

Then by Lemma 4.1 (*iv*) it follows that

$$|DS_t(F(S_q(x)))| \le C2^{\beta j} \quad \text{for every } x \in L_j. \tag{5.10}$$

If we now combine the estimates (5.8), (5.9) and (5.10) the chain rule gives us

$$|D_1 f(x)| \le Cj^{\alpha+1} \quad \text{for a.e. } x \in L_j. \tag{5.11}$$

Case 3: Let us now assume that

$$x \in L_j \quad \text{with } j > 3k - 3N_F - 3.$$

Also in this case S_q maps the line segment L_j to a line segment L_j^1 parallel to x_1-axis and inside the set $(\hat{U}_j \setminus \hat{U}_{j+1}) \times (\hat{P}_k \setminus \hat{P}_{k+1})$, and we have

$$|D_1 S_q(x)| \le C2^{-\beta j} j^{\alpha+1}. \tag{5.12}$$

Also the derivative of F can be estimated again by

$$|D_1 F(S_q(x))| \le \mathrm{Lip}(F) \quad \text{for a.e. } x \in L_j. \tag{5.13}$$

Moreover, as $x \in [-1,1]^4 \setminus (\hat{U}_{k+1})^4$ we have

$$\begin{aligned} [-1,1]^4 \setminus (\hat{U}_{k+1})^4 &\subset \{y \in [-1,1]^4 : \mathrm{dist}(y, C_B) > \hat{r}_{k+1} - \hat{r}_{k+2}\} \\ &\subset \{y \in [-1,1]^4 : \mathrm{dist}(y, C_B) > 2^{-\beta(k+2)-(k+2)}\}. \end{aligned}$$

If we then assume that $\tilde{C} > 0$ is the constant in Lemma 3.5 we may again assume that $\tilde{C}^{-1} > 2^{\beta+1}2^{-\beta N_F - N_F}$ (see case 2). Then it follows from Lemma 3.5 and from the inclusion above that

$$\begin{aligned} F(S_q(x)) &\in \{z \in [-1,1]^4 : \mathrm{dist}(z, C_B) > \tilde{C}^{-1}2^{-\beta(k+2)+(k+2)}\} \\ &\subset \{z \in [-1,1]^4 : \mathrm{dist}(z, C_B) > 2^{-\beta(k+N_F+2)-(k+N_F+2)}\} \\ &\subset [-1,1]^4 \setminus (\hat{U}_{k+N_F+2})^4. \end{aligned}$$

Thus, it follows from Lemma 4.1 (*iv*) and (*v*) that we may estimate

$$|DS_t(F(S_q(x)))| \le C2^{\beta(3k-3N_F-3)}. \tag{5.14}$$

If we now combine (5.12), (5.13) and (5.14) the chain rule gives us

$$|D_1 f(x)| \le Cj^{\alpha+1}2^{-\beta(j-3k)} \quad \text{for a.e. } x \in L_j. \tag{5.15}$$

Estimating the Sobolev norm of f: The above estimates (5.7), (5.11) and (5.15) can be summarized as follows. Suppose that $k \ge k_0$ and let $x \in L_j := (U_j \setminus U_{j+1}) \times \{\tilde{x}\}$ with $\tilde{x} \in P_k \setminus P_{k+1}$. Then

$$|D_1 f(x)| \le \begin{cases} C & \text{if } 1 \le j \le k - 2N_F - 3 \\ Cj^{\alpha+1} & \text{if } k - 2N_F - 3 < j \le 3k - 3N_F - 3 \\ Cj^{\alpha+1}2^{-\beta(j-3k)} & \text{if } j > 3k - 3N_F - 3, \end{cases} \tag{5.16}$$

where the constant C does not depend on k or j.

Also note that S_q maps $\mathcal{C}_A \times \mathbb{R}^3$ onto $\mathcal{C}_B \times \mathbb{R}^3$ and using $|\mathcal{C}_A| > 0$ and $|\mathcal{C}_B| = 0$ we easily obtain $|D_1 S_q| = 0$ on $\mathcal{C}_A \times \mathbb{R}^3$. As F is just a reflection on $\mathcal{C}_B \times \mathbb{R}^3$ and S_t is locally Lipschitz on $[-1,1]^4 \setminus C_B$, we easily obtain that

$$|D_1 f| = 0 \quad \text{on } (\mathcal{C}_A \times \mathbb{R}^3) \setminus C_A.$$

Therefore, for $\tilde{x} \in P_k \setminus P_{k+1}$ we can calculate

$$\begin{aligned}\int_{(-1,1)} |D_1 f(x_1, \tilde{x})|^p \,\mathrm{d}x_1 &= \int_{(-1,1)\setminus \mathcal{C}_A} |D_1 f(x_1, \tilde{x})|^p \,\mathrm{d}x_1 \\ &= \sum_{j=1}^{\infty} \int_{U_j \setminus U_{j+1}} |Df(x_1, \tilde{x})|^p \,\mathrm{d}x_1.\end{aligned}$$

We use the fact that f is Lipschitz on $[-1,1]^4 \setminus (U_{k_0})^4$ for every fixed k_0 to see that

$$\int_{(-1,1)} |D_1 f(x_1, \tilde{x})|^p \,\mathrm{d}x_1 \leq C + \sum_{j=k_0}^{\infty} \int_{U_j \setminus U_{j+1}} |D_1 f(x_1, \tilde{x})|^p \,\mathrm{d}x_1$$

for every $\tilde{x} \in U_k^3 \setminus U_{k+1}^3$ with $k \geq k_0$.

Let us next estimate the measure of the set $\{x_1 \in [-1,1] : x_1 \in U_j \setminus U_{j+1}\}$. For every given j this set contains 2^j line segments each having length which can be approximated above by $2^{-j}\bigl(1 + \frac{1}{(j+1)^\alpha} - 1 - \frac{1}{(j+2)^\alpha}\bigr)$. Thus the measure of the set can be approximated as

$$\mathcal{L}^1(U_j \setminus U_{j+1}) \leq C 2^j 2^{-j}\Bigl(1 + \frac{1}{(j+1)^\alpha} - 1 - \frac{1}{(j+2)^\alpha}\Bigr) \leq \frac{C}{j^{\alpha+1}}.$$

Therefore, for the line segment $L = [-1,1] \times \{\tilde{x}\}$ we have using (5.16)

$$\begin{aligned}&\int_L |D_1 f|^p dx_1 = C + \sum_{j=k_0}^{\infty} \int_{L_j} |Df(x_1, \tilde{x})|^p dx_1 \\ &\leq C\Biggl(1 + \sum_{j=k_0}^{k-2N_F-3} \frac{1}{j^{\alpha+1}} + \sum_{j=k-2N_F-2}^{4k-3N_F-3} \frac{j^{p(\alpha+1)}}{j^{\alpha+1}} + \sum_{j=4k-3N_F-2}^{\infty} 2^{-p\beta(j-3k)} \frac{j^{p(\alpha+1)}}{j^{\alpha+1}}\Biggr).\end{aligned}$$

The first sum converges even if we sum to infinity, the second sum will be estimated simply by taking an estimate of the largest summand and multiplying by an estimate of the total number of summands. We will use (5.3) to estimate the final sum by a convergent geometric sum ($\sum_{l=k}^{\infty} 2^{-pl\beta/2}$). Continuing the calculation and using $k \geq 4N_F + 5$ we have

$$\begin{aligned}\int_L |D_1 f|^p \,dx_1 &\leq C + C4k(4k)^{(p-1)(\alpha+1)} + \frac{C}{1 - 2^{-kp\beta/2}} \\ &\leq C + C\frac{k^{p\alpha+p}}{k^\alpha}.\end{aligned} \tag{5.17}$$

The estimate (5.17) holds for all lines $L = [-1,1] \times \{\tilde{x}\}$ such that $\tilde{x}$ in $P_k \setminus P_{k+1}$ with $k \geq k_0$. Furthermore, since f is Lipschitz on $[-1,1]^n \setminus U_{k_0}^n$, we may estimate

$$\int_L |D_1 f|^p \,\mathrm{d}x_1 \leq C \quad \text{for all } \tilde{x} \in P_{k+1} \setminus P_k \text{ with } k < k_0,$$

which proves the validity of (5.17) for all $k \in \mathbb{N}$ (not only for $k \geq k_0$). If $\tilde{x} \in \mathcal{C}_A^3$ then we will again use the fact that

$$|D_1 f(x)| \leq C \text{ for a.e. } x \in L. \tag{5.18}$$

Now we integrate (5.17) over $\tilde{x} \in [-1,1]^3$. By (2.4) we know that

$$\mathcal{L}^3(P_k \setminus P_{k+1}) \leq \frac{C}{k^{\alpha+1}}$$

and we continue by multiplying this with (5.17) and summing over k plus (5.18) multiplied by the measure $\mathcal{L}^3(\mathcal{C}_A^3) = 1$. Since $\alpha \geq 2$ we have

$$\begin{aligned} \int_{(-1,1)^4} |D_1 f(x)|^p dx &\leq \sum_{k=1}^{\infty} C k^{-\alpha-1} + C \sum_{k=1}^{\infty} \frac{k^{p\alpha+p}}{k^{2\alpha+1}} + C \\ &\leq C + C \sum_{k=1}^{\infty} \frac{k^p}{k^{(2-p)\alpha}} = C \sum_{k=k_0}^{\infty} \frac{1}{k^p} < \infty \end{aligned} \tag{5.19}$$

by our choice of $\alpha = \frac{2p}{2-p}$ at the start of the proof. This ends the proof of Theorem 1.2 when $n = 4$. Taking our mapping f in 4 dimensions and using it to define a mapping $f^* : \mathbb{R}^5 \to \mathbb{R}^5$ as follows

$$f^*(x_1, x_2, x_3, x_4, x_5) = (f(x_1, x_2, x_3, x_4), x_5) \tag{5.20}$$

proves Theorem 1.2 when $n = 5$. □

6. The higher dimensional case $n \geq 6$

Let $\mathcal{M}(o,n)$ be the set of all linear subspaces of $\mathbb{R}^n$ of dimension o, parallel to the coordinate axes (i.e. $M \in \mathcal{M}(o,n)$ if and only if there exists a basis of M of o vectors chosen from the canonical basis). Where there is no danger of confusion we will omit n and write simply $\mathcal{M}(o)$. Previously we defined $\mathcal{K}_A$ as $\bigcup_{L \in \mathcal{M}(1,4)} C_A + L$. From now on we take $n \geq 6$ even, $m = n/2 - 1$ and define

$$\mathcal{K}_A = \bigcup_{L \in \mathcal{M}(m,n)} C_A + L \quad \text{and} \quad \mathcal{K}_B = \bigcup_{L \in \mathcal{M}(m,n)} C_B + L,$$

where $C_A = \bigcap_k U_k^n$ and $C_B = \bigcap_k \hat{U}_k^n$. Of course C_B and $\mathcal{K}_B$ depend on the parameter β. During our proof we show that if β is large enough then our mapping exists and we show how to construct the mapping for any β sufficiently large. Let us note that for n odd we can define our mapping analogously to (5.20) by using identity in the last coordinate.

We make a further explicitation to the notation used above and that is the sets $P_k = U_k^{n-1}$. It is more or less obvious how to generalize the notion from subsections 2.5 and 2.6 to the higher dimensional case, see e.g. [23, Proof of Theorem 4.9]. In this section we will show that if we fix $1 \leq p < [n/2]$ then by choosing the parameters $\alpha > 0$ and $\beta > 0$ large enough we may construct the mapping $f \in W^{1,p}$ which we have in mind in Theorem 1.2.

6.1. **Mapping F in higher dimensions:** We will introduce some sets that will aid notation for Theorem 6.1. Let $L \in \mathcal{M}(o,n)$, $1 \leq o \leq m$ then call

$$N_L = \{e_j \in \mathbb{R}^n : e_j \in L^{\perp}\}$$

and let M_L be the set of all subsets of N_L with $n-m$ elements. Let $k,l\in\mathbb{N}$ then call

$$A_{L,k,l}=\bigcup_{W\in M_L}\Big(\{x\in\mathbb{R}^n: x_i\in[-1,1]\setminus\hat{U}_k, e_i\in L\}\cap\{x\in\mathbb{R}^n: x_j\in\hat{U}_l, e_j\in W\}\Big). \tag{6.1}$$

This is the set where informally speaking we are far away from our Cantor set C_B in o directions and close in some $n-m$ directions (perpendicular to the given o directions) and in the remaining $n-o-(n-m)$ directions x_i could be arbitrary.

Theorem 6.1. *Let $m\in\mathbb{N}$ and $n=2m+2$. There exists a mapping F which is a sense-preserving bi-Lipschitz extension of the map*

$$F(x_1,x_2,\dots,x_{n-1},x_n)=(x_1,x_2,\dots,x_{n-1},-x_n)\quad x\in\mathcal{K}_B. \tag{6.2}$$

There exists an $N_F\in\mathbb{N}$ such that for each $k\in\mathbb{N}$, $k>N_F$, $1\leq o\leq m$, $L\in\mathcal{M}(o,n)$ we have

$$F\big((x+L)\cap A_{L,j-N_F,k+N_F+1}\big)\subset\big(F(x)+L\big)\cap A_{L,j,k+1} \tag{6.3}$$

for any given $x\in A_{L,j-N_F,j+N_F+1}$.

The inclusion (6.3) basically means that the image of those parts of affine spaces $x+L$ which are much closer to $\mathcal{C}_B$ in $n-m$ directions from $L^\perp$ than it is in directions from L in the map F is part of an affine space $F(x)+L$ and the distance of the affine space from $\mathcal{K}_B$ is roughly maintained.

The proof of Theorem 6.1 is similar to that of Theorem 3.1. We find vectors v and u, a Lipschitz extension g onto $X=\mathbb{R}^{n-1}\times\{0\}$ of $g(P_v(x))=-x_n$ and then $F=F_{g,u}\circ F_{g,v}$. The following lemma, corresponds to Lemma 3.4

Lemma 6.2. *Let $m\in\mathbb{N}$ and $n=2m+2$. Let $v=(2^{-n},2^{1-n},\dots,\frac{1}{4},1)$ and $u=(-2^{-n},-2^{1-n},\dots,-\frac{1}{4},1)$. Then there exists $\beta>0$ and a corresponding set $\mathcal{K}_B$ such that P_v is one-to-one on $\mathcal{K}_B$ and the function g defined on $P_v(\mathcal{K}_B)$ as $g(P_v(x))=-x_n$ can be extended onto X as a Lipschitz function. Furthermore, it is possible to find a Lipschitz extension of the function g which guarantees that*

$$D_iF_{g,u}\circ F_{g,v}(x)=\begin{cases} e_i & \text{if } i=1,2,\dots,n-1\\ -e_i & \text{if } i=n\end{cases} \tag{6.4}$$

whenever $x_i\in[-1,1]\setminus\hat{U}_k$ and we can find a set of $n-m$ indexes $\{j_1,j_2,\dots j_{n-m}\}$ such that $x_{j_1},x_{j_2},\dots,x_{j_{n-m}}\in\hat{U}_{k+2}$ and $j_l\neq i$ for every $l=1,2,\dots,n-m$.

Proof. With some small modifications the proof will mainly follow the proof of Lemma 3.4.

Step 1: The projection P_v is one-to-one on C_B. Step 1 here is the same as in the previous lemma. The reader can somewhat laboriously but easily check that $P_v(a)\neq P_v(b)$ whenever a,b are distinct vertices of $Q(0,\frac{1}{2})$. This gives us a set of 2^n distinct points and so (using Q^{n-1} to denote cubes in $\mathbb{R}^{n-1}$) there exists a $d_0>0$ such that $Q^{n-1}(P_v(a),d)\cup Q^{n-1}(b,d)=\emptyset$ whenever a and b are distinct vertices of $Q(0,\frac{1}{2})$ and $0<d\leq d_0$. By the continuity of P_v there exists a $d_1>0$ such that the sets $P_v(Q(a,d))$ are pairwise disjoint for distinct vertices a of the cube $Q(0,\frac{1}{2})$ and $0<d\leq d_1$. Thus we have proved that whenever we construct the cantor set C_B using $\beta=\log_2(d)-1$ for any $0<d\leq d_1$ we have

$$P_v(\hat{Q}_{\boldsymbol{v}(1)})\cap P_v(\hat{Q}_{\boldsymbol{v}'(1)})=\emptyset \text{ whenever } \boldsymbol{v}(1)\neq\boldsymbol{v}'(1).$$

The self-similarity argument applied in Lemma 3.4 applies here too and so we see that the image of the collection of all k-th generational cubes $\hat{Q}_{\boldsymbol{v}(k)}$ are pairwise disjoint and this holds for all k. This implies that P_v is one-to-one on C_B for $\beta > \beta_0$. In fact this is a special case of the next step for $o = 0$.

Step 2: The projection P_v is one-to-one on $\mathcal{K}_B$. We would like to prove that P_v is one-to-one on $\mathcal{K}_B$. Let $\boldsymbol{v}(k) \in \mathbb{V}^k$ and let $M \in \mathcal{M}(o,n)$, $1 \le o \le m$, then we will prove that the projection of any distinct pair of k-"bars" $\hat{S}^M_{\boldsymbol{v}(k)}$ where

$$\hat{S}^M_{\boldsymbol{v}(k)} = \Big(Q(z_{\boldsymbol{v}(k)}, \hat{r}_k) + M\Big) \setminus \Big(\bigcup_{\boldsymbol{w}\in\mathbb{V}^k} \bigcup_{L\in\mathcal{M}(o-1)} Q(\hat{z}_{\boldsymbol{w}}, \hat{r}_{k-2}) + L\Big), \tag{6.5}$$

is disjoint. Similarly to before we achieve this by projecting them into disjoint sets

$$\begin{aligned}\hat{\mathcal{S}}^M_{\boldsymbol{v}(k)} := &\Big(Q^{n-1}\big(P_v(\hat{z}_{\boldsymbol{v}(k)}), q\hat{r}_k\big) + P_v(M)\Big) \\ &\setminus \Big(\bigcup_{\boldsymbol{w}\in\mathbb{V}^k} \bigcup_{L\in\mathcal{M}(o-1)} Q^{n-1}\big(P_v(\hat{z}_{\boldsymbol{w}}), \hat{r}_{k-1}\big) + P_v(L)\Big).\end{aligned} \tag{6.6}$$

Let us note that in dimension $n = 4$ our definition was slightly different as we used $w \in \mathbb{V}^{k-1}$ (or $w \in \mathbb{V}^{k-2}$) in previous definitions. However, this is not a big change as the union of cubes over all $w \in \mathbb{V}^{k-1}$ or $w \in \mathbb{V}^k$ is similar (up to some multiple of radius) and from technical reasons this is better here.

Step 2A: The projection P_v is one-to-one on every $M \in \mathcal{M}(n-1)$. Recall the corresponding notation from Lemma 3.4,

$$\tilde{v} = (2^{-n}, 2^{1-n}, \dots, \tfrac{1}{4}).$$

The definition of P_v (3.2) immediately yields that

$$P_v(e_l) = \begin{cases} e_l & \text{if } 1 \le l \le n-1 \\ -\tilde{v} & \text{if } l = n. \end{cases} \tag{6.7}$$

We take $M \in \mathcal{M}(n-1)$ and solve $P_v(u) = 0$, $u \in M$. First we will assume that $e_n \in M^\perp$. Then (6.7) says that $P_v(u) = u$ for all $u \in M$ and the only solution to $P_v(u) = 0$ is $u = 0$ and thus P_v is one-to-one on M. Now we assume that $e_n \in M$ and we find j such that $\operatorname{span}\{e_j\} = M^\perp$. Using (6.7) we obtain

$$0 = P_v(u) = P_v\Big(\sum_{\substack{i=1\\ i\neq j}}^{n} \lambda_i e_i\Big) = \sum_{\substack{i=1\\ i\neq j}}^{n-1} \lambda_i e_i - \lambda_n \tilde{v}.$$

Thus the j-th coordinate of the last expression must be zero, which implies $\lambda_n \tilde{v}_j = 0$ and hence $\lambda_n = 0$. Thus we have reduced to the first case which has been proved already.

Step 2B: Finding β such that $\hat{\mathcal{S}}^M_{\boldsymbol{v}(k)}$ are disjoint sets. We defined sets $\hat{\mathcal{S}}^M_{\boldsymbol{v}(k)}$ in (6.6) and now we would like to show that if one chooses $\beta > \beta_2$ that these sets are pairwise disjoint. Exactly the same arguments from Lemma 3.4, Step 2B, Claim (1) can be applied here to see that the two contesting definitions from Lemma 3.4 of $\hat{\mathcal{S}}^M_{\boldsymbol{v}(k)}$ which we could generalize are equivalent.

Let A denote the set of vertices of $Q(0, \frac{1}{2})$. We define the "sliced" affine sets for $M \in \mathcal{M}(o)$ and $a \in A$

$$W^M_a = (a + M) \setminus \Big(\bigcup_{b\in A} \bigcup_{L\in\mathcal{M}(o-1)} Q(b, \tfrac{4}{5}) + L \Big). \tag{6.8}$$

The sets $W^{M_1}_{a_1}$ and $W^{M_2}_{a_2}$ are equal if and only if $M_1 = M_2$ and $a_1 - a_2 \in M_1$.

Let us make the following useful observations on P_v. Since $\|\tilde{v}\| = \frac{1}{4}$ we obtain the simple observation

$$\|P_v(x)\| \le \tilde{x} + \tfrac{1}{4}|x_n| \le \tfrac{5}{4}\|x\|, \tag{6.9}$$

recall that $\|\cdot\|$ denotes the maximum norm. Also we denote the distance with respect to this norm as $\operatorname{dist}_\infty$. Further we use the fact that P_v is one-to-one whenever restricted to any $M \in \mathcal{M}(n-1)$ and especially P_v is one-to-one on each $a + M$ for $M \in \mathcal{M}(o)$ and $a \in A$. There are a finite number of such affine spaces and P_v is one-to-one on each of them. This implies that there is a $\lambda > 0$ such that whenever we choose $M \in \mathcal{M}(o)$ and $x \in M$ that $\|P_v(x)\| \ge \lambda^{-1}\|x\|$. The fact that $\|P_v(x)\| \le \lambda\|x\|$ is shown in (6.9) if $\lambda \ge \frac{5}{4}$. Now we choose any $M \in \mathcal{M}(o)$, any $L \in \mathcal{M}(o-1)$, $a \in A$ and any $x \in W^M_a$ and conclude that

$$\lambda^{-1} \operatorname{dist}_\infty(x, W^L_a) \le \operatorname{dist}_\infty(P_v(x), P_v(W^L_a)) \le \lambda \operatorname{dist}_\infty(x, W^L_a) \tag{6.10}$$

as the distance of x to W^L_a is attained in some direction in $M \setminus L$.

Let us recall that the sets W^M_a are defined in (6.8). In the following we will be interested in pairs of distinct W^M_a. Another fact that is clear is if we have a pair of distinct $W^{M_1}_{a_1}$ and $W^{M_2}_{a_2}$ (with $M_1, M_2 \in \mathcal{M}(o)$) then either $M_1 = M_2$ and $(a_1 + M_1) \cap (a_2 + M_2) = \emptyset$ or $\dim(M_1 \cap M_2) \le o - 1$ and there exists an $L \in \mathcal{M}(\tilde{o})$, $\tilde{o} \le o - 1$, and an $a \in A$ such that

$$(a_1 + M_1) \cap (a_2 + M_2) \subset a + L \subset \bigcup_{b\in A} \bigcup_{L\in\mathcal{M}(o-1)} Q(b, \tfrac{4}{5}) + L.$$

This means that our pair $W^{M_1}_{a_1}$ and $W^{M_2}_{a_2}$ are disjoint if distinct in both cases. We can easily calculate that

$$\operatorname{dist}_\infty(W^{M_1}_{a_1}, W^{M_2}_{a_2}) \ge \frac{4}{5}$$

and so (6.10) gives that

$$\operatorname{dist}_\infty\big(P_v(W^{M_1}_{a_1}), P_v(W^{M_2}_{a_2})\big) \ge \frac{4}{5\lambda}. \tag{6.11}$$

This however immediately implies that $\{P_v(W^M_a) + Q^{n-1}(0, \frac{2}{5\lambda})\}$ is a finite family of closed pairwise disjoint sets.

Let $\delta > 0$. Assuming that $a \in A$, $M \in \mathcal{M}(o)$, $\boldsymbol{v}(1) = 2a \in \mathbb{V}$, $\hat{r}_1 < \frac{\delta}{q}$, it is simple to observe that

$$\begin{aligned} P_v(a+M) + Q(0,\delta) &= P_v(\hat{z}_{\boldsymbol{v}(1)}) + P_v(M) + Q(0,\delta) \\ &\supset Q^{n-1}(P_v(\hat{z}_{\boldsymbol{v}(1)}), q\hat{r}_1) + P_v(M). \end{aligned} \tag{6.12}$$

We will again use the fact that P_v is one-to-one on all $M \in \mathcal{M}(o)$ to see that

$$P_v\big(W_a^M\big) = P_v\big(a+M\big) \setminus P_v\big((a+M) \cap (\bigcup_{b\in A} \bigcup_{L\in\mathcal{M}(o-1)} Q(b,\tfrac{4}{5}) + L)\big)$$

and so for any $L \in \mathcal{M}(o-1)$ and for the $b \in A$ such that $\boldsymbol{w}(1) = 2b \in \mathbb{V}$

$$\begin{aligned} P_v\big((a+M) \cap (Q(b,\tfrac{4}{5}) + L)\big) &\subset P_v\big(Q(b,\tfrac{4}{5}) + L\big) \\ &= P_v\big(Q(b,\tfrac{4}{5})\big) + P_v\big(L\big) \\ &\subset Q^{n-1}(P_v(\hat{z}_w), 1) + P_v(L). \end{aligned} \tag{6.13}$$

The definition of $\hat{\mathcal{S}}^M_{\boldsymbol{v}(1)}$ (6.6) in combination with (6.12) and (6.13) show that

$$\hat{\mathcal{S}}^M_{\boldsymbol{v}(1)} \subset P_v(W_a^M) + Q^{n-1}(0,\delta), \text{ whenever } \hat{z}_{\boldsymbol{v}(1)} = a.$$

Further by applying (6.11) and assuming $\delta < \frac{1}{5\lambda}$ (and $\hat{r}_1 < \frac{\delta}{q}$) we see that the sets $P_v(W_a^M) + Q^{n-1}(0,\delta)$ are pairwise disjoint and in fact

$$\operatorname{dist}_\infty(P_v(W_{a_1}^{M_1}) + Q^{n-1}(0,\delta), P_v(W_{a_2}^{M_2}) + Q^{n-1}(0,\delta)) \geq \tfrac{1}{5\lambda},$$

whenever the pair is distinct. This implies that the sets $\hat{\mathcal{S}}^M_{\boldsymbol{v}(1)}$ satisfy

$$\operatorname{dist}(\hat{\mathcal{S}}^{M_1}_{\boldsymbol{v}(1)}, \hat{\mathcal{S}}^{M_2}_{\boldsymbol{w}(1)}) \geq \tfrac{C(n)}{5\lambda},$$

whenever distinct. Further by self similarity we get the same for all k, i.e.

$$\operatorname{dist}(\hat{\mathcal{S}}^{M_1}_{\boldsymbol{v}(k)}, \hat{\mathcal{S}}^{M_2}_{\boldsymbol{w}(k)}) \geq \tfrac{C(n)}{5\lambda}\hat{r}_{k-1}$$

whenever distinct.

Step 2C: Proving the inclusion $P_v(\hat{S}^M_{\boldsymbol{v}(k)}) \subset \hat{\mathcal{S}}^M_{\boldsymbol{v}(k)}$. We will prove the inclusion $P_v(\hat{S}^M_{\boldsymbol{v}(2)}) \subset \hat{\mathcal{S}}^M_{\boldsymbol{v}(2)}$ and for other k it will hold by self-similarity. Again we will employ (6.9) in the following to calculate that

$$\begin{aligned} P_v\big(Q(z_{\boldsymbol{v}(2)}, \hat{r}_2) + M\big) &\subset Q^{n-1}\big(P_v(\hat{z}_{\boldsymbol{v}(2)}), \tfrac{5}{4}\hat{r}_2\big) + P_v(M) \\ &\subset Q^{n-1}\big(P_v(\hat{z}_{\boldsymbol{v}(2)}), q\hat{r}_2\big) + P_v(M) \end{aligned} \tag{6.14}$$

whenever $q \geq \frac{5}{4}$. The remainder of what we need to prove is that for each $\boldsymbol{w}(2) \in \mathbb{V}^2$ and for each $L \in \mathcal{M}(o-1)$

$$Q^{n-1}\big(P_v(\hat{z}_{\boldsymbol{w}(2)}), \hat{r}_1\big) + P_v(L) \subset P_v\big(Q(\hat{z}_{\boldsymbol{w}(2)}, \hat{r}_0) + L\big)$$

which can easily be achieved by selecting $\hat{r}_1$ small enough (i.e. β large enough) as $\hat{r}_0 = 1$. This step is analogous to the proof in dimension $n = 4$ and therefore we skip the details.

Step 2D: Conclusion of Step 2. By definition

$$\mathcal{K}_B = \bigcup_{M\in\mathcal{M}(m)} C_B + M.$$

Let us consider the sets

$$K_o = \big(\bigcup_{M\in\mathcal{M}(o)} C_B + M\big) \setminus \big(\bigcup_{L\in\mathcal{M}(o-1)} C_B + L\big).$$

It is easy to see that

$$K_o = \bigcup_{k_0\geq 1}\bigcap_{k\geq k_0}\bigcup_{M\in\mathcal{M}(o)}\bigcup_{\boldsymbol{v}(k)\in\mathbb{V}^k} \hat{S}^M_{\boldsymbol{v}(k)}$$

and since $\mathcal{K}_B = \bigcup_{o=1}^m K_o$ we obtain

$$\mathcal{K}_B = \bigcup_{0\leq o\leq m}\bigcup_{k_0\geq 1}\bigcap_{k\geq k_0}\bigcap_{k\geq 1}\bigcup_{M\in\mathcal{M}(o)}\bigcup_{\boldsymbol{v}(k)\in\mathbb{V}^k} \hat{S}^M_{\boldsymbol{v}(k)}.$$

We have proven that for any fixed k the images of $\hat{S}^M_{\boldsymbol{v}(k)}$ in P_v are pairwise disjoint, whenever the pair of sets in question are distinct. Take any pair of distinct points $x, y \in \mathcal{K}_B$. If there exists $k, M_1, \boldsymbol{v}(k)$ and $M_2, \boldsymbol{w}(k)$ such that $x \in \hat{S}^{M_1}_{\boldsymbol{v}(k)} \neq \hat{S}^{M_2}_{\boldsymbol{w}(k)} \ni y$ then P_v maps x and y onto distinct points in X because as we have proven

$$P_v(\hat{S}^{M_1}_{\boldsymbol{v}(k)}) \cap P_v(\hat{S}^{M_2}_{\boldsymbol{w}(k)}) = \emptyset.$$

If for almost every k we have $x, y \in \hat{S}^M_{\boldsymbol{v}(k)}$, then $x - y \in M$ and P_v is one-to-one on M and so maps x and y to distinct points.

Step 3: Defining the function g on X. Now we expound how to perform step 3 of the proof, that is how to define g on X. In steps 3A, 3B and 3C we assume always that $e_n \notin M$ and $M \in \mathcal{M}(o)$ for some $1 \leq o \leq m$. The case $e_n \in M$ is dealt with in 3D. Step 3E then proves that g has the desired properties. We will make use of the sets

$$\hat{H}^M_{\boldsymbol{v}(k)} := \partial_X\Big(Q^{n-1}\big(P_v(\hat{z}_{\boldsymbol{v}(k)}), q\hat{r}_k\big) + P_v(M)\Big) \setminus \Big(\bigcup_{\boldsymbol{w}\in\mathbb{V}^{k-1}}\bigcup_{\substack{L\in\mathcal{M}(m-1)\\ L\subset M}} Q\big(\hat{z}_{\boldsymbol{w}}, \hat{r}_{k-1}\big) + P_v(L)\Big)$$

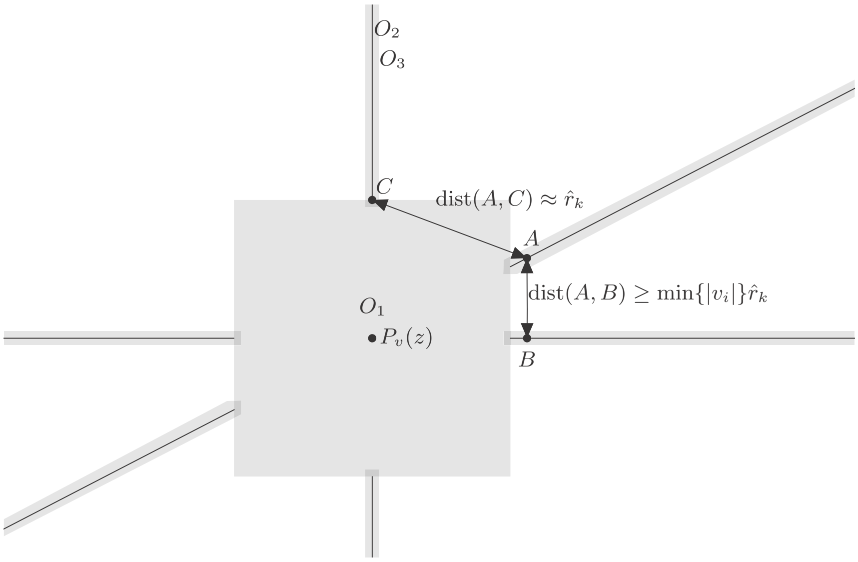
where Q^{n-1} is a cube in $\mathbb{R}^{n-1}$ (specifically in $X = \mathbb{R}^{n-1}\times\{0\}$) and $\partial_X U$ is the relative boundary of a set U with respect to X.

Step 3A. First we take a "pipe" $\hat{H}^M_{\boldsymbol{v}(k)}$ with $e_n \notin M$ and $k \geq 2$ and define

$$g(x) = -(\hat{z}_{\boldsymbol{v}(k)})_n \text{ for all } x \in \hat{H}^M_{\boldsymbol{v}(k)}.$$

Again, first we remark that if $\hat{H}^M_{\boldsymbol{v}(k)} = \hat{H}^M_{\tilde{\boldsymbol{v}}(k)}$ then $(\hat{z}_{\boldsymbol{v}(k)})_n = (\hat{z}_{\tilde{\boldsymbol{v}}(k)})_n$ because $e_n \notin M$ and therefore g is well-defined at these points. It is easy to see that if we have two pipes, both parallel to M, one inside another (that is $\hat{z}_{\boldsymbol{v}(k+1)} + M$ intersects $\hat{Q}_{\boldsymbol{v}(k)}$) then

$$\operatorname{dist}(\hat{H}^M_{\boldsymbol{v}(k)}, \hat{H}^M_{\boldsymbol{v}(k+1)}) \geq C\hat{r}_k \text{ for a suitable } C > 0 \text{ independent of } k.$$



$O_1 = \bigcup_{L' \in \mathcal{M}(o-1)} Q(P_v(z), \hat{r}_k) + L'$,

$O_2 = P_v(z+L) \setminus \bigcup_{L' \in \mathcal{M}(o-1)} Q(P_v(z), \hat{r}_k) + L'$ for some $L \in \mathcal{M}(o)$,

$O_3 = \big(P_v(x+L_1) \setminus \bigcup_{L' \in \mathcal{M}(o-1)} Q(P_v(z), \hat{r}_k) + L'\big) + Q(0, \hat{r}_{k+1})$ for some $L_1 \in \mathcal{M}(o)$.

FIGURE 10. Measuring the distances between sliced projected affine spaces reduces to the case dealt with in Lemma 3.4 where we measured the distance between sliced lines. The thickness of the 'bars' is $\hat{r}_{k+1}$ which can be made much smaller than the distance between them which is comparable to $\hat{r}_k$.

Further considering $x \in \hat{H}^M_{\boldsymbol{v}(k)}$ and $y \in \hat{H}^M_{\boldsymbol{v}(k+1)}$ we have

$$|g(x) - g(y)| = |-(\hat{z}_{\boldsymbol{v}(k)})_n + (\hat{z}_{\boldsymbol{v}(k+1)})_n| = \tfrac{1}{2}\hat{r}_k.$$

Considering two distinct pipes $\hat{H}^M_{\hat{\boldsymbol{v}}(k+1)}$ and $\hat{H}^M_{\boldsymbol{v}(k+1)}$ of the same generation, both inside $\hat{H}^M_{\boldsymbol{v}(k)}$ we see that

$$\mathrm{dist}(\hat{H}^M_{\hat{\boldsymbol{v}}(k+1)}, \hat{H}^M_{\boldsymbol{v}(k+1)}) \geq C\hat{r}_k \text{ for a suitable } C > 0 \text{ independent of } k.$$

Furthermore, for $x \in \hat{H}^M_{\boldsymbol{v}(k+1)}$ and $y \in \hat{H}^M_{\boldsymbol{v}(k+1)}$ we have

$$|g(x) - g(y)| = |-(z_{\hat{\boldsymbol{v}}(k+1)})_n + (z_{\boldsymbol{v}(k+1)})_n| \leq \hat{r}_k\,.$$

This proves that g, thus defined, on the pipes $\hat{H}^M_{\boldsymbol{v}(k)}$ with $e_n \notin M$ is Lipschitz with respect to parallel pipes, i.e. pipes given by the same subspace M.

Step 3B. Similarly, for $M \in \mathcal{M}(o)$, $1 \leq o \leq m$, and $e_n \notin M$ we define

$$g(P_v(\hat{z}_{\boldsymbol{v}(k)} + M)) = -(\hat{z}_{\boldsymbol{v}(k)})_n.$$

Also for every $x \in C_B$ we define

$$g(P_v(x + M)) = -x_n.$$

Note that by step 2 we know that these sets are pairwise disjoint whenever distinct and P_v is one-to-one on $\mathcal{K}_B$ and therefore this definition is correct. The estimates from Lemma 3.4, step 3B easily generalize to this setting showing that our definition of g is Lipschitz on the collection of all pipes, i.e., all sets of type $P_v(\hat{z}_{\boldsymbol{v}(k)} + M)$ and $P_v((C_B \cap \hat{Q}_{\boldsymbol{v}(k)}) + M)$.

Step 3C. Now we will fix $k \geq 2$, $\boldsymbol{v}(k) \in \mathbb{V}^k$, $1 \leq o \leq m$ and $M \in \mathcal{M}(o)$ (we still assume that $e_n \notin M$) and define g on

$$L^M_{\boldsymbol{v}(k)} := \hat{\mathcal{S}}^M_{\boldsymbol{v}(k)} \setminus \bigcup_{\boldsymbol{v}(k+1) \in \mathbb{V}^{k+1}} \hat{\mathcal{S}}^M_{\boldsymbol{v}(k+1)}.$$

Call $Y_M = M^\perp \cap X$ and denote π_{Y_M} the orthogonal projection onto this subspace. In general one can only claim that the projection of a pair of sets does not increase the distance between them. Here however we consider sets parallel to a given vector space M and project them onto Y_M, which is perpendicular to M. In this case the projection does not decrease the distance between the sets either. That is to say (in the following we use $P_v(M) = M$, see (6.7))

$$\operatorname{dist}(\pi_{Y_M}(\hat{H}^M_{\boldsymbol{v}(k)}), \pi_{Y_M}(\hat{H}^M_{\boldsymbol{v}(k+1)})) = \operatorname{dist}(\hat{H}^M_{\boldsymbol{v}(k)}, \hat{H}^M_{\boldsymbol{v}(k+1)}).$$

Similarly

$$\operatorname{dist}(\pi_{Y_M}(\hat{H}^M_{\boldsymbol{v}(k)}), \pi_{Y_M}(\hat{z}_{\boldsymbol{v}(k)} + M)) = \operatorname{dist}(\hat{H}^M_{\boldsymbol{v}(k)}, \hat{z}_{\boldsymbol{v}(k)} + M)$$

and

$$\operatorname{dist}(\pi_{Y_M}(\hat{H}^M_{\boldsymbol{v}(k)}), \pi_{Y_M}(x + M)) = \operatorname{dist}(\hat{H}^M_{\boldsymbol{v}(k)}, x + M)$$

for $x \in P_v(C_B \cap \hat{Q}_{\boldsymbol{v}(k)})$. We defined g as constant on sets $\hat{H}^M_{\boldsymbol{v}(k)}$, therefore we may define a function $\tilde{g}$ on Y_M as $\tilde{g}(\pi_{Y_M}(x)) = g(x)$ for any $x \in \hat{H}^M_{\boldsymbol{v}(k)}$ and this definition is correct. The above estimates on the distances of the sets projected onto Y_M shows that $\tilde{g}$ is Lipschitz with respect to the projection of those sets. Therefore we may use the McShane extension theorem to get a Lipschitz $\tilde{g}$ defined on Y_M. For $x \in L^M_{\boldsymbol{v}(k)}$ we define $g(x) = \tilde{g}(\pi_{Y_M}(x))$ and so get a function g, which is constant on the intersection of any affine space parallel to M with the set $L^M_{\boldsymbol{v}(k)}$.

Once we have defined g on all $\hat{\mathcal{S}}^M_{\boldsymbol{v}(k)}$ for all $k, \boldsymbol{v}(k)$ and M we still need to fill in certain "gaps", where we transition from $M \in \mathcal{M}(o)$ to $L \in \mathcal{M}(o-1)$. Considering Figure 10 we need to define g on sets corresponding to O_3. Specifically, for $M \in \mathcal{M}(o)$ we need to define g on

$$\begin{aligned}\hat{T}^M_{\boldsymbol{v}(k)} =& \Big(Q^{n-1}\big(P_v(z_{\boldsymbol{v}(k)}), q\hat{r}_k\big) + P_v(M)\Big) \\ & \setminus \Bigg(\bigcup_{\substack{\tilde{L} \in \mathcal{M}(o-2) \\ \tilde{L} \subset M}} \Big(Q^{n-1}\big(P_v(z_{\boldsymbol{v}(k)}), \hat{r}_{k-1}\big) + P_v(\tilde{L})\Big) \cup \hat{\mathcal{S}}^M_{\boldsymbol{v}(k)}\Bigg).\end{aligned}$$

These gaps were necessary as they made the sets $\mathcal{S}^M_{\boldsymbol{v}(k)}$ disjoint (whenever distinct) and this made the definition of g in step 3A and 3B correct. Note that it was not possible to define g as constant on entire m-dimensional subspaces (without removing $m-1$ dimensional subspaces) because they intersect other m-dimensional subspaces where g has a different value. Informally speaking, in each $o-1$ dimensional gap we project onto a corresponding perpendicular $n-1-(o-1)$ dimensional subspace on which, g is already defined in some points by step 3A and 3B. We keep those values and extend them as a Lipschitz function on the perpendicular subspace and then, by projecting along the $o-1$ dimensional subspace we define g everywhere in the gap.

If $x \in \hat{T}^M_{\boldsymbol{v}(k)}$ and $o \geq 2$ then there exists exactly one coordinate, let us say the i-th coordinate where $e_i \in M$,

$$|x_i - \big(P_v(z_{\boldsymbol{v}(k)})\big)_i| \leq \hat{r}_{k-1}$$

but for all $j \neq i$ such that $e_j \in M$ we have

$$|x_j - \big(P_v(z_{\boldsymbol{v}(k)})\big)_j| > \hat{r}_{k-1}.$$

So, set $M \in \mathcal{M}(o)$, $L \in \mathcal{M}(o-1)$ and e_i such that $\operatorname{span}\{L, e_i\} = M$. Recall that $M \subset X$ and P_v is identity on X to see that a point $x \in \hat{T}^M_{\boldsymbol{v}(k)}$ can be expressed as

$$x = P_v(z_{\boldsymbol{v}(k)}) + \sum_{e_j \in L} \lambda_j e_j + t e_i + \sum_{e_l \in M^\perp \cap X} \tilde{\lambda}_l e_l$$

where $\lambda_j > \hat{r}_{k-1}$, $\tilde{\lambda}_l < q\hat{r}_k$, $t < \hat{r}_{k-1}$. We project $\hat{T}^M_{\boldsymbol{v}(k)}$ onto $Y_L = L^\perp \cap X$. Since g is constant on affine subsets contained in $\mathcal{S}^{\hat{M}}_{\boldsymbol{v}(k)}$ parallel to L it is also constant on affine subsets of $\partial\hat{T}^M_{\boldsymbol{v}(k)} \cap \partial\hat{\mathcal{S}}^M_{\boldsymbol{v}(k)}$ parallel to L (note that these affine sets on boundaries have dimension $o-1$). Thus the following definition is correct

$$\tilde{g}(y) = g(x) \text{ whenever } y = \pi_{Y_L}(x)$$

and $x \in \partial\hat{T}^M_{\boldsymbol{v}(k)} \cap \partial\hat{\mathcal{S}}^M_{\boldsymbol{v}(k)}$. The function $\tilde{g}$ is Lipschitz and can be Lipschitz extended onto Y_L by the McShane Theorem. For $x \in \hat{T}^M_{\boldsymbol{v}(k)}$ we define

$$g(x) = \tilde{g}(\pi_{Y_L}(x)).$$

In this manner we extend g for all $M, L, \hat{z}_{\boldsymbol{v}(k)}$ and all k in the case where $e_n \notin M$.

Step 3D. Next we will define g on pipes $\hat{H}^M_{\boldsymbol{v}(k)}$ with $e_n \in M$ and $k \geq 2$. For this, let us next denote $\tilde{v} := \big(2^{-n}, 2^{1-n}, \dots, \frac{1}{4}\big)$ and define

$$Y_n := \{w \in \mathbb{R}^{n-1} : \langle w, \tilde{v}\rangle = 0\}.$$

Then we separate $\mathbb{R}^{n-1}$ into the direct sum $\mathbb{R}\tilde{v} \oplus Y_n$. Suppose now that $\lambda_0 \in \mathbb{R}$ and $w_0 \in Y_n$ are such that

$$P_v(\hat{z}_{\boldsymbol{v}(k)}) = w_0 + \lambda_0 \tilde{v}. \tag{6.15}$$

Then, if $\tilde{x} \in \hat{H}^M_{\boldsymbol{v}(k)} \cup P_v(\hat{z}_{\boldsymbol{v}(k)} + M)$ we may find $\lambda \in \mathbb{R}$ such that $\tilde{x} = w + \lambda\tilde{v}$ with $w \in Y_n$ which leads us to define

$$g(\tilde{x}) = \lambda - \lambda_0 - (\hat{z}_{\boldsymbol{v}(k)})_n \quad \text{for every } \tilde{x} \in \hat{H}^M_{\boldsymbol{v}(k)} \cup P_v(\hat{z}_{\boldsymbol{v}(k)} + M). \tag{6.16}$$

This means that g has been defined as constant on the intersections of the sets in question with hyperplanes in X parallel to $\tilde{v}^\perp$.

We claim that the definition in (6.16) and the definition $g(P_v(x)) = -x_n$ for $x \in \mathcal{K}_B$ gives us a g Lipschitz on the collection of sets $\hat{H}^M_{\boldsymbol{v}(k)}$, $P_v(\hat{z}_{\boldsymbol{v}(k)} + M)$ and $P_v(\mathcal{K}_B)$. The proof of this is just a repetition of step 3D from Lemma 3.4. Once again we extend our map by creating a Lipschitz extension on Y_M and by using $g(x) = \tilde{g}(\pi_{Y_M}(x))$.

Where not yet defined we may extend g Lipschitz arbitrarily, for example by the McShane extension theorem.

Step 3E: verifying the condition (6.4)**.** Now we define the spaghetti strand map $F_{g,v}$ again as

$$F_{g,v}(x) = x + vg(P_v(x)).$$

Analogously to the proof in dimension $n = 4$ it is possible to show that

$$D_i\big(F_{g,u} \circ F_{g,v}\big)(x) = \begin{cases} e_i & \text{if } i = 1, 2, \dots, n-1 \\ -e_i & \text{if } i = n \end{cases}$$

whenever $x_i \in [-1, 1] \setminus \hat{U}_k$ and we can find a set of $n - m$ indexes $\{j_1, j_2, \dots j_{n-m}\}$ such that $x_{j_1}, x_{j_2}, \dots, x_{j_{n-m}} \in \hat{U}_{k+2}$. There are two possibilities. Either $i \neq n$ and g is constant on $P_v(x + \mathbb{R}e_i)$ or $i = n$ and $g(P_v(x + te_n)) = c - t$. This is true because all x such that $x_i \in [-1, 1] \setminus \hat{U}_k$ and $x_{j_1}, x_{j_2}, \dots, x_{j_{n-m}} \in \hat{U}_{k+2}$ belong to some $\hat{S}^M_{\boldsymbol{v}(k)}$ which is projected into $\hat{\mathcal{S}}^M_{\boldsymbol{v}(k)}$ and we defined g on $\hat{\mathcal{S}}^M_{\boldsymbol{v}(k)}$ to have precisely these qualities. The rest of the calculations are just a repetition of step 3E of Lemma 3.4. □

Proof of Theorem 6.1. It suffices to repeat the proof from Theorem 3.1 to show that the mapping $F = F_{g,u} \circ F_{g,v}$ with g, v, u from Lemma 6.2 satisfies (6.2). Also we see that F is sense preserving for the same reason as before. The proof of the behavior of F on m-dimensional planes close to $\mathcal{K}_B$ is the same as it was for lines in the previous too. The choice of N_F follows from the same arguments as in Theorem 3.1 and an adaption of Lemma 3.5, where we replace lines with affine spaces and the proof remains the same. □

6.2. **Mapping f in higher dimensions.** We define our mapping in much the same way as in the 4-dimensional case. We set

$$f = S_t \circ F \circ S_q,$$

where $S_q = (q(x_1), q(x_2), \dots, q(x_n))$, F is the mapping from Theorem 6.1 and S_t is defined exactly as before. More precisely, we define S_t by

$$S_t(x) = d_{i,k}(x) H^{n-1,i}_{3k-3}(x) + (1 - d_{i,k}(x)) H^{n-1,i}_{3k}(x) + t(x_i)e_i \qquad \text{for } x \in \hat{E}_{i,k}, \tag{6.17}$$

where the mappings $H^{n-1,i}_{3k-3}$ and $H^{n-1,i}_{3k}$ are the obvious higher dimensional generalizations of the mappings $H^{3,i}_{3k-3}$ and $H^{3,i}_{3k}$ in Section 4, and

$$\hat{E}_{i,k} = \{x \in (\hat{U}_{k-1})^n \setminus (\hat{U}_k)^n : d_{i,k}(x) \geq d_{j,k}(x), j \neq i\}.$$

By following the arguments in Section 4 the reader may generalize Lemma 4.1 in all dimensions:

Lemma 6.3. *Suppose that $S_t : (-1, 1)^n \to (-1, 1)^n$ is defined as in* (6.17) *where $n \geq 4$. Then S_t is a sense-preserving homeomorphisms which satisfies the following conditions:*

(i) *S_t maps C_B onto C_A and $S_t = S_q^{-1}$ on C_B.*
(ii) *Mapping S_t is locally Lipschitz on $(-1, 1)^n \setminus C_B$.*

(iii) *If L_i is a line parallel to x_i-axis with $L_i \cap (\hat{U}_k)^n \neq \emptyset$ then*

$$|D_i S_t(x)| \leq C \frac{2^{\beta k}}{k^{\alpha+1}}$$

for every $x \in L_i \cap \big((\hat{U}_{k-1})^n \setminus (\hat{U}_k)^n\big)$.

(iv) *If $k \leq j \leq 3k+2$ and $z \in (\hat{U}_j \setminus \hat{U}_{j+1}) \times \big((\hat{U}_k)^{n-1} \setminus (\hat{U}_{k+1})^{n-1}\big)$ then*

$$|DS_t(x)| \leq C 2^{\beta j}.$$

The same holds for $x \in \big((\hat{U}_k)^{n-1} \setminus (\hat{U}_{k+1})^{n-1}\big) \times (\hat{U}_j \setminus \hat{U}_{j+1})$ and also for $n-2$ other permutations of coordinates.

(v) *If $x \in \hat{U}_{3k+3} \times \big((\hat{U}_k)^{n-1} \setminus (\hat{U}_{k+1})^{n-1}\big)$ then*

$$|DS_t(x)| \leq C 2^{\beta(3k+3)}.$$

The same holds for $z \in \big((\hat{U}_k)^{n-1} \setminus (\hat{U}_{k+1})^{n-1}\big) \times (\hat{U}_{3k+3})$ and also for $n-2$ other permutations of coordinates.

Sobolev regularity of f. Just as before f, as the composition of homeomorphisms, is a homeomorphism. Our aim is to prove that if $1 \leq p < [\frac{n}{2}]$ then for an aptly chosen $\alpha > 0$ in the definition of C_A the corresponding mapping f belongs to $W^{1,p}$. Therefore we are interested in calculating the integral

$$\int_{(1,1)^{n-1}} \int_{(-1,1)} |D_1 f(x_1, \tilde{x})|^p \, \mathrm{d}x_1 \mathrm{d}\tilde{x}.$$

The integrals over lines in other directions can all be estimated in the same way as the reader may easily check. Recall that n is even and we start by fixing the exponent $1 \leq p < [n/2]$ and

$$\alpha = \frac{2p}{n/2 - p}$$

and the index $k_0 \geq 4N_F + 5$, where $N_F \in \mathbb{N}$ is from Lemma 3.5, large enough so that

$$\max\{2^{-kp\beta/2} k^{(p-1)(\alpha+1)}, 2^{-p\beta}(\tfrac{k+1}{k})^{\alpha}\} < 1 \quad \text{for all } k \geq k_0. \tag{6.18}$$

The reasoning in the arguments in section 5 for the ACL condition and the use of the chain rule hold here too. Both S_q and F are Lipschitz maps. By Lemma 6.3 (ii) we see that S_t is $C(k_0)$-Lipschitz on $[-1,1]^n \setminus \hat{U}^n_{k_0+N_F}$. Therefore it follows that f is Lipschitz on $[-1,1]^n \setminus U^n_{k_0}$, and it remains to consider the set $U^n_{k_0}$.

We use the ACL property of f to make the following estimates on the derivatives of $|D_1 f|$. For convenience sake we will denote $x = (x_1, \tilde{x})$. Now we will fix $k \in \mathbb{N}$ with $k \geq k_0 \geq 4N_F + 5$ and in the further we assume that $\tilde{x} \in (U_k)^{n-1} \setminus (U_{k+1})^{n-1}$. We define the following divisions of a segment $L = [-1,1] \times \{\tilde{x}\}$:

$$L_j(\tilde{x}) = L_j = \{(x_1, \tilde{x}) : x_1 \in (U_j \setminus U_{j+1})\}.$$

In the following we use the simpler notation L_j to aid readability. The aim of the following calculations is to prove the estimate (6.31) below.

Case 1: Let us first assume that $k \geq k_0$ and

$$x \in L_j \quad \text{with } j = 1, \dots, k - 2N_F - 3.$$

Then S_q maps the line segment L_j to a line segment L_j^1 which is parallel to x_1-axis and lies inside the set $(\hat{U}_j \setminus \hat{U}_{j+1}) \times (\hat{P}_k \setminus \hat{P}_{k+1})$. Furthermore, we have

$$|D_1 S_q(x)| \leq C 2^{-\beta j} j^{\alpha+1}, \tag{6.19}$$

with C independent of j, k.

Recall that the sets $A_{L,j,k}$ are defined in (6.1). Observe that

$$L_j^1 \subset A_{\mathbb{R}e_1, (j+N_F+2)-N_F-1, (k-N_F+1)+N_F},$$

where $N_F < j + N_F + 2 \leq k - N_F + 1$, and thus it follows from Theorem 6.1 that the bi-Lipschitz map F maps L_j^1 to a line segment L_j^2 parallel to x_1-axis and such that

$$L_j^2 \subset A_{\mathbb{R}e_1, (j+N_F+2)-1, k-N_F+1}.$$

Moreover, we have

$$|DF(S_q(x))| \leq \operatorname{Lip}(F) \quad \text{for a.e. } x \in L_j. \tag{6.20}$$

Finally, because L_j^2 is a line segment parallel to the x_1-axis which is contained in $[-1,1]^n \setminus (\hat{U}_{j+N_F+1})^n$ it follows from Lemma 6.3 (iii) that

$$|D_1 S_t(F(S_q(x)))| \leq C 2^{\beta j} j^{-(\alpha+1)}, \tag{6.21}$$

with C independent of j, k.

If we now combine (6.19), (6.20) and (6.21) we get

$$|D_1 f(x)| \leq C \quad \text{for a.e. } x \in L_j, \tag{6.22}$$

with C independent of j, k.

Case 2: Let us next assume that

$$x \in L_j \quad \text{with } k - 2N_F - 3 < j \leq 3k - 3N_F - 3.$$

Again S_q maps the line segment L_j to a line segment L_j^1 which is parallel to x_1-axis and lies inside the set $(\hat{U}_j \setminus \hat{U}_{j+1}) \times (\hat{P}_k \setminus \hat{P}_{k+1})$. Furthermore, we have

$$|D_1 S_q(x)| \leq C 2^{-\beta j} j^{\alpha+1}, \tag{6.23}$$

with C independent of j, k.

Next, we recall that

$$|D_1 F(S_q(x))| \leq \operatorname{Lip}(F) \quad \text{for a.e. } x \in L_j. \tag{6.24}$$

Moreover, it follows from the assumption $j \geq k - 2N_F - 2$ that

$$\begin{aligned}
L_j^1 &\subset \{x \in [-1,1]^4 : \operatorname{dist}(x, \mathcal{K}_B) > \min\{\hat{r}_j - \hat{r}_{j+1}, \hat{r}_k - \hat{r}_{k+1}\}\} \\
&\subset \big\{x \in [-1,1]^4 : \operatorname{dist}(x, \mathcal{K}_B) > \min\{2^{-\beta(j+1)-(j+1)}, 2^{-\beta(k+1)-(k+1)}\}\big\} \\
&\subset \big\{x \in [-1,1]^4 : \operatorname{dist}(x, \mathcal{K}_B) > \min\{2^{-\beta(j+1)-(j+1)}, 2^{-\beta(j+2N_F+3)-(j+2N_F+3)}\}\big\} \\
&\subset \big\{x \in [-1,1]^4 : \operatorname{dist}(x, \mathcal{K}_B) > 2^{-\beta(j+2N_F+3)-(j+2N_F+3)}\big\}.
\end{aligned}$$

Suppose now that $\tilde{C} > 0$ is the constant given by Lemma 3.5. Then we may choose $N_F \in \mathbb{N}$ to be so large that $\tilde{C}^{-1} > 2^{3\beta+3}2^{-\beta N_F - N_F}$. Then it follows from Lemma 3.5 and from the inclusion above that

$$\begin{aligned} F(L_j^1) &\subset \big\{x \in [-1,1]^4 : \operatorname{dist}(x, \mathcal{K}_B) > \tilde{C}^{-1}2^{-\beta(j+2N_F+3)-(j+2N_F+3)}\big\} \\ &\subset \big\{x \in [-1,1]^4 : \operatorname{dist}(x, \mathcal{K}_B) > 2^{-\beta(j+3N_F)-(j+3N_F)}\big\}. \end{aligned}$$

Thus, we have that $F(L_j^1)$ is contained in the following union of n sets

$$\begin{aligned} F(L_j^1) \subset &\Big(([-1,1] \setminus \hat{U}_{j+3N_F}) \times \big([-1,1]^{n-1} \setminus (\hat{U}_{j+3N_F})^{n-1}\big)\Big) \cup \\ &\cdots \cup \Big(\big([-1,1]^{n-1} \setminus (\hat{U}_{j+3N_F})^{n-1}\big) \times ([-1,1] \setminus \hat{U}_{j+3N_F})\Big). \end{aligned}$$

Without loss of generality suppose that

$$F(L_j^1) \subset ([-1,1] \setminus \hat{U}_{j+3N_F}) \times \big([-1,1]^{n-1} \setminus (\hat{U}_{j+3N_F})^{n-1}\big).$$

Then by Lemma 6.3 (iv) it follows that

$$|DS_t(F(S_q(x)))| \leq C2^{\beta j} \quad \text{for a.e. } x \in L_j. \tag{6.25}$$

If we now combine the estimates (6.23), (6.24) and (6.25) the chain rule gives us

$$|D_1 f(x)| \leq C j^{\alpha+1} \quad \text{for a.e. } x \in L_j, \tag{6.26}$$

with C independent of j, k.

Case 3: Let us now assume that

$$x \in L_j \quad \text{with } j > 3k - 3N_F - 3.$$

Also in this case S_q maps the line segment L_j to a line segment L_j^1 parallel to x_1-axis and inside the set $(\hat{U}_j \setminus \hat{U}_{j+1}) \times (\hat{P}_k \setminus \hat{P}_{k+1})$, and we have

$$|D_1 S_q(x)| \leq C2^{-\beta j} j^{\alpha+1}, \tag{6.27}$$

with C independent of j, k. Furthermore, also the derivative of F can be estimated again by

$$|D_1 F(S_q(x))| \leq \operatorname{Lip}(F) \quad \text{for a.e. } x \in L_j. \tag{6.28}$$

Moreover, as $x \in [-1,1]^n \setminus (\hat{U}_{k+1})^n$ we have

$$\begin{aligned} [-1,1]^n \setminus (\hat{U}_{k+1})^n &\subset \{y \in [-1,1]^n : \operatorname{dist}(y, C_B) > \hat{r}_{k+1} - \hat{r}_{k+2}\} \\ &\subset \{y \in [-1,1]^n : \operatorname{dist}(y, C_B) > 2^{-\beta(k+2)-(k+2)}\}. \end{aligned}$$

If we then assume that $\tilde{C} > 0$ is the constant in Lemma 3.5 we may again assume that $\tilde{C}^{-1} > 2^{\beta+1}2^{-\beta N_F - N_F}$ (see case 2). Then it follows from Lemma 3.5 and from the inclusion above that

$$\begin{aligned} F(S_q(x)) &\in \{z \in [-1,1]^n : \operatorname{dist}(z, C_B) > \tilde{C}^{-1}2^{-\beta(k+2)+(k+2)}\} \\ &\subset \{z \in [-1,1]^n : \operatorname{dist}(z, C_B) > 2^{-\beta(k+N_F+2)-(k+N_F+2)}\} \\ &\subset [-1,1]^n \setminus (\hat{U}_{k+N_F+2})^n. \end{aligned}$$

Thus, it follows from Lemma 6.3 (iv) and (v) that we may estimate

$$|DS_t(F(S_q(x)))| \leq C2^{\beta(3k-3N_F-3)}, \tag{6.29}$$

with C independent of j, k.

If we now combine (6.27), (6.28) and (6.29) the chain rule gives us

$$|D_1 f(x)| \leq Cj^{\alpha+1}2^{-\beta(j-3k)} \quad \text{for a.e. } x \in L_j, \tag{6.30}$$

with C independent of j, k.

Estimating the Sobolev norm of f: The above estimates (6.22), (6.26) and (6.30) can be summarized as follows. Suppose that $k \geq k_0$ and let $x \in L_j := (U_j \setminus U_{j+1}) \times \{\tilde{x}\}$ with $\tilde{x} \in P_k \setminus P_{k+1}$. Then

$$|D_1 f(x)| \leq \begin{cases} C & \text{if } 1 \leq j \leq k - 2N_F - 3 \\ Cj^{\alpha+1} & \text{if } k - 2N_F - 3 < j \leq 3k - 3N_F - 3 \\ Cj^{\alpha+1}2^{-\beta(j-3k)} & \text{if } j > 3k - 3N_F - 3, \end{cases} \tag{6.31}$$

where the constant $C := C(n, \alpha, \beta, N_F, \mathrm{Lip}(F))$ does not depend on k or j.

Also note that S_q maps $\mathcal{C}_A \times \mathbb{R}^{n-1}$ onto $\mathcal{C}_B \times \mathbb{R}^{n-1}$ and using $|\mathcal{C}_A| > 0$ and $|\mathcal{C}_B| = 0$ we easily obtain $|D_1 S_q| = 0$ on $\mathcal{C}_A \times \mathbb{R}^{n-1}$. As F is just a reflection on $\mathcal{C}_B \times \mathbb{R}^{n-1}$ and S_t is locally Lipschitz on $[-1,1]^n \setminus C_B$, we easily obtain that

$$|D_1 f| = 0 \quad \text{on } (\mathcal{C}_A \times \mathbb{R}^{n-1}) \setminus C_A.$$

Therefore, for $\tilde{x} \in P_k \setminus P_{k+1}$ we can calculate

$$\begin{aligned} \int_{(-1,1)} |D_1 f(x_1, \tilde{x})|^p \, \mathrm{d}x_1 &= \int_{(-1,1)\setminus\mathcal{C}_A} |D_1 f(x_1, \tilde{x})|^p \, \mathrm{d}x_1 \\ &= \sum_{j=1}^{\infty} \int_{U_j \setminus U_{j+1}} |Df(x_1, \tilde{x})|^p \, \mathrm{d}x_1. \end{aligned}$$

We use the fact that f is C-Lipschitz on $[-1,1]^n \setminus (U_{k_0})^n$ (k_0 fixed in (6.18)) to see that

$$\int_{(-1,1)} |D_1 f(x_1, \tilde{x})|^p \, \mathrm{d}x_1 \leq C^p$$

for every $\tilde{x} \in [-1,1]^{n-1} \setminus (U_{k_0})^{n-1}$.

Therefore we may now restrict to the case that $\tilde{x} \in P_k \setminus P_{k+1}$ for $k \geq k_0$. By the same reasoning as in Section 5 we have

$$\mathcal{L}^1(U_j \setminus U_{j+1}) \leq \frac{C}{j^{\alpha+1}}.$$

Therefore, for the line segment $L = [-1,1] \times \{\tilde{x}\}$ we have using (6.31)

$$\begin{aligned} &\int_L |D_1 f|^p dx_1 = C + \sum_{j=k_0}^{\infty} \int_{L_j} |Df(x_1, \tilde{x})|^p dx_1 \\ &\leq C\Bigg(1 + \sum_{j=k_0}^{k-2N_F-3} \frac{1}{j^{\alpha+1}} + \sum_{j=k-2N_F-2}^{4k-3N_F-3} \frac{j^{p(\alpha+1)}}{j^{\alpha+1}} + \sum_{j=4k-3N_F-2}^{\infty} 2^{-p\beta(j-3k)} \frac{j^{p(\alpha+1)}}{j^{\alpha+1}}\Bigg). \end{aligned}$$

The first sum converges even if we sum to infinity, the second sum will be estimated simply by taking an estimate of the largest summand and multiplying by an estimate of the total number of summands. We will use (6.18) to estimate the final sum by a convergent geometric sum ($\sum_{l=k}^{\infty} 2^{-pl\beta/2}$). Continuing the calculation and using $k \geq 4N_F + 5$ we have

$$\begin{aligned}\int_L |D_1 f|^p \, dx_1 &\leq C + C4k(4k)^{(p-1)(\alpha+1)} + \frac{C}{1-2^{-kp\beta/2}} \\ &\leq C + C\frac{k^{p\alpha+p}}{k^{\alpha}}.\end{aligned} \tag{6.32}$$

The estimate (6.32) holds for all lines $L = [-1,1] \times \{\tilde{x}\}$ such that $\tilde{x}$ in $P_k \setminus P_{k+1}$ with $k \geq k_0$. Furthermore, since f is Lipschitz on $[-1,1]^n \setminus U^n_{k_0}$, we may estimate

$$\int_L |D_1 f|^p \, \mathrm{d}x_1 \leq C \quad \text{for all } \tilde{x} \in P_k \setminus P_{k+1} \text{ with } k < k_0, \tag{6.33}$$

which proves the validity of (6.32) for all $k \in \mathbb{N}$ (not only for $k \geq k_0$). We will use the estimate (6.32) on those lines which are not entirely contained in $\mathcal{K}_A$ and on lines which are entirely contained in $\mathcal{K}_A$ we will use again the fact that

$$|D_1 f(x)| \leq C \text{ for a.e. } x \in L. \tag{6.34}$$

Now we integrate the above estimates over $\tilde{x} \in [-1,1]^{n-1}$. Calling $\tilde{K}_A$ the set of those $\tilde{x}$ such that $L(\tilde{x})$ is contained entirely in $\mathcal{K}_A$ we claim that

$$\mathcal{L}^{n-1}\big(U_k^{n-1} \setminus (U_{k+1}^{n-1} \cup \tilde{\mathcal{K}}_A)\big) < \mathcal{L}^{n-1}\big(U_k^{n-1} \setminus \tilde{\mathcal{K}}_A\big) \leq Ck^{-m\alpha} \tag{6.35}$$

where $m := \frac{n}{2} - 1$. Once we will have established this estimate the rest of the proof follows quickly from the following calculations. We continue by multiplying (6.32) by the measure estimate (6.35) and summing over k plus (6.34) multiplied by the measure $\mathcal{L}^{n-1}([-1,1]^{n-1}) > \mathcal{L}^{n-1}(\tilde{\mathcal{K}}_B \cap [-1,1]^{n-1})$. Assuming (6.35) we have (it holds that $\alpha m > 1$)

$$\begin{aligned}\int_{(-1,1)^n} |D_1 f(x)|^p dx &\leq \sum_{k=1}^{\infty} Ck^{-m\alpha} + C\sum_{k=1}^{\infty} \frac{k^{p(\alpha+1)}}{k^{\alpha(m+1)}} + C \\ &\leq C + C\sum_{k=1}^{\infty} \frac{k^p}{k^{(m+1-p)\alpha}} = C\sum_{k=k_0}^{\infty} \frac{k^p}{k^{2p}} < \infty\end{aligned} \tag{6.36}$$

by our choice $\alpha = \frac{2p}{n/2-p}$ at the beginning of the proof.

Therefore it remains to prove (6.35). We notice that a line segment $L := [-1,1] \times \{\tilde{x}\}$ parallel to x_1-axis is contained in $\mathcal{K}_A$ if and only if $L \subset C_A + M$ for some $M \in \mathcal{M}(m)$ containing e_1. We write all such subspaces M as $\mathbb{R}e_1 + \tilde{M}$ for some $\tilde{M} \in \mathcal{M}(m-1, n-1)$. Hereby we see that

$$L \subset \mathcal{K}_A \quad \text{if and only if} \quad \tilde{x} \in \tilde{\mathcal{K}}_A := \bigcup_{\tilde{M} \in \mathcal{M}(m-1,n-1)} \Big(\mathcal{C}_A^{n-1} + \tilde{M}\Big).$$

Next, we estimate

$$\begin{aligned}\mathcal{L}^{n-1}\big(U_k^{n-1} \setminus (U_{k+1}^{n-1} \cup \tilde{\mathcal{K}}_A)\big) &< \mathcal{L}^{n-1}\big(U_k^{n-1} \setminus \tilde{\mathcal{K}}_A\big) \\ &= 2^{k(n-1)}\big(2^{-k}(1+k^{-\alpha})\big)^{n-1} - \mathcal{L}^{n-1}\big((U_k)^{n-1} \cap \tilde{\mathcal{K}}_A\big).\end{aligned} \tag{6.37}$$

and therefore it suffices to calculate $\mathcal{L}^{n-1}\big((U_k)^{n-1}\cap\tilde{\mathcal{K}}_A\big)$. We do this by decomposing the set $U_k^{n-1}\cap\tilde{\mathcal{K}}_A$ into the disjoint union of m sets $E_0^k, E_1^k, \dots, E_{m-1}^k$ and each E_i^k is the disjoint union of $\binom{n-1}{i}$ measurable rectangles. For simplicity of the notation call $G_k = U_k \setminus \mathcal{C}_A$. We denote

$$
\begin{aligned}
&E_0^k := \mathcal{C}_A^{n-1}\\
&E_1^k := (G_k\times\mathcal{C}_A^{n-2})\cup(\mathcal{C}_A\times G_k\times\mathcal{C}_A^{n-3})\cup\cdots\cup(\mathcal{C}_A^{n-3}\times G_k\times\mathcal{C}_A)\cup(\mathcal{C}_A^{n-2}\times G_k)\\
&\vdots\\
&E_{m-1}^k := (G_k^{m-1}\times\mathcal{C}_A^{m+3})\cup\cdots\cup(\mathcal{C}_A^{m+3}\times G_k^{m-1}).
\end{aligned}
$$

So each E_j^k is a union of $\binom{n-1}{j}$ sets $F_l(j,k)$, $l = 1,2,\dots,\binom{n-1}{j}$. Each $F_l(j,k)$ is a measurable rectangle and G_k appears in its product j times and the set $\mathcal{C}_A$ appears $n-1-j$ times. Each $F_l(j,k)$ is uniquely determined by the sequence of sets in its product. So if $F_l(j,k)\neq F_{l'}(j',k')$ then there is a direction such that one of the sets is projected onto $\mathcal{C}_A$ and the other is projected onto G_k and $\mathcal{C}_A\cap G_k=\emptyset$.

Now simple calculation gives

$$
\mathcal{L}^1(\tilde{\mathcal{C}}_A) = 1 \quad\text{and}\quad \mathcal{L}^1(G_k) = \frac{1}{(k+1)^\alpha}
$$

and so by definition

$$
\begin{aligned}
\mathcal{L}^{n-1}(E_j^k) &= \binom{n-1}{j}\mathcal{L}^{n-1}\big((G_k)^j\times(\tilde{\mathcal{C}}_A)^{n-1-j}\big)\\
&= \binom{n-1}{j}\big[\mathcal{L}^1(G_k)\big]^j\big[\mathcal{L}^1(\tilde{\mathcal{C}}_A)\big]^{n-1-j} = \binom{n-1}{j}\frac{1}{(k+1)^{\alpha j}}
\end{aligned}
$$

for each $j=0,1,\dots,m-1$. Therefore, we see that

(6.38)

$$
\mathcal{L}^{n-1}\big((U_k)^{n-1}\cap\tilde{\mathcal{K}}_A\big) = \mathcal{L}^{n-1}\left(\bigcup_{j=0}^{m-1}E_j^k\right) = \sum_{j=0}^{m-1}\mathcal{L}^{n-1}\big(E_j^k\big) = \sum_{j=0}^{m-1}\binom{n-1}{j}\frac{1}{(k+1)^{\alpha j}}.
$$

When we now combine (6.37) and (6.38) we get

$$
\begin{aligned}
&\mathcal{L}^{n-1}\big(U_k^{n-1}\setminus(U_{k+1}^{n-1}\cup\tilde{\mathcal{K}}_A)\big)\\
&\qquad\leq\left(1+\frac{1}{(k+1)^\alpha}\right)^{n-1} - \sum_{j=0}^{m-1}\binom{n-1}{j}\frac{1}{(k+1)^\alpha}\leq\frac{C}{k^{m\alpha}},
\end{aligned}
$$

which is exactly what we claimed in (6.35). As shown in (6.36) This ends the proof of Theorem 1.2 for all $n\geq 4$. □

Acknowledgement. The authors would like to thank Tapio Rajala for the discussion on the projections of Cantor sets into lower dimensional subspaces.

References

[1] L. Ambrosio, N. Fusco and D. Pallara, *Functions of bounded variation and free discontinuity problems*, Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2000.

[2] S.S. Antman, *Fundamental mathematical problems in the theory of nonlinear elasticity*, Numerical solution of partial differential equations, III (Proc. Third Sympos. (SYNSPADE), Univ. Maryland, College Park, Md., 1975) Academic Press, New York, 1976, 35–54.

[3] J.M. Ball, Convexity conditions and existence theorems in nonlinear elasticity, *Arch. Rational Mech. Anal.* **63**, no. 4, (1977), 337–403.

[4] J.M. Ball, The calculus of variations and material science, Current and future challenges in the applications of mathematics, (Providence, RI, 1997), *Quart. Appl. Math.* **56**, no. 4, (1998), 719–740.

[5] J.M. Ball, *Singularities and computation of minimizers for variational problems*, Foundations of computational mathematics (Oxford, 1999), 1–20, London Math. Soc. Lecture Note Ser., 284, Cambridge Univ. Press, Cambridge, 2001.

[6] J.M. Ball, *Progress and puzzles in Nonlinear Elasticity*, Proceedings of course on Poly-, Quasi- and Rank-One Convexity in Applied Mechanics. CISM Courses and Lectures. Springer, 2010.

[7] J.C. Bellido and C. Mora-Corral, Approximation of Hölder continuous homeomorphisms by piecewise affine homeomorphisms, *Houston J. Math.* **37** no.2 (2011), 449–500.

[8] R.H. Bing, Locally tame sets are tame, *Ann. of Math.* **59** (1954), 145–158.

[9] R.H. Bing, Stable homeomorphisms on E^5 can be approximated by piecewise linear ones, *Notices Amer. Math. Soc.* **10** (1963), 607–616.

[10] D. Campbell, *Diffeomorphic approximation of Planar Sobolev Homeomorphisms in Orlicz-Sobolev spaces*, Preprint.

[11] R. Černý, Homeomorphism with zero Jacobian: Sharp integrability of the derivative, *J. Math. Anal. Appl* **373** (2011), 161–174.

[12] P.G. Ciarlet, *Mathematical Elasticity. Vol. I. Three-dimensional elasticity.*, Studies in Mathematics and its Applications, 20. North-Holland Publishing Co., Amsterdam, 1987.

[13] E.H. Connell, Approximating stable homeomorphisms by piecewise linear ones, *Ann. of Math.* **78** (1963), 326–338.

[14] M. Csörnyei, S. Hencl and J. Malý, Homeomorphisms in the Sobolev space $W^{1,n-1}$, *J. Reine Angew. Math.* **644** (2010), 221–235.

[15] S. Daneri and A. Pratelli, Smooth approximation of bi-Lipschitz orientation-preserving homeomorphisms, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **31** no. 3 (2014), 567–589.

[16] S.K. Donaldson and D.P. Sullivan, Quasiconformal 4-manifolds, *Acta Math.* **163** (1989), 181–252.

[17] L. D'Onofrio, S. Hencl and R. Schiattarella, Bi-Sobolev homeomorphism with zero Jacobian almost everywhere, *Calc. Var. and PDE* **51** (2014), 139–170.

[18] L.C. Evans, Quasiconvexity and partial regularity in the calculus of variations, *Ann. of Math.* **95** no. 3 (1986), 227–252.

[19] H. Federer, Surface area. II, *Trans. Amer. Math. Soc.* **55** (1944), 438–456.

[20] I. Fonseca and W. Gangbo, *Degree Theory in Analysis and Applications*, Clarendon Press, Oxford, 1995.

[21] P. Goldstein and P. Hajlasz, *A measure and orientation preserving homeomorphism of a cube with Jacobian equal* -1 *almost everywhere*, preprint 2015.

[22] S. Hencl, Sobolev homeomorphism with zero Jacobian almost everywhere, *J. Math. Pures Appl.* **95** (2011), 444–458.

[23] S. Hencl and P. Koskela, *Lectures on mappings of finite distortion*, Lecture Notes in Mathematics 2096, Springer, Cham, 2014, 176pp.

[24] S. Hencl and J. Malý, Jacobians of Sobolev homeomorphisms, *Calc. Var. Partial Differential Equations* **38**, no. 1–2, (2010), 233–242.

[25] S. Hencl and A. Pratelli, *Diffeomorphic approximation of* $W^{1,1}$ *planar Sobolev homeomorphisms*, to appear in J. Eur. Math. Soc.

[26] S. Hencl and B. Vejnar, Sobolev homeomorphisms that cannot be approximated by diffeomorphisms in $W^{1,1}$, *Arch. Rational Mech. Anal.* **219**, no. 1, (2016), 183–202.

[27] T. Iwaniec, L.V. Kovalev and J. Onninen, Diffeomorphic Approximation of Sobolev Homeomorphisms, *Arch. Rational Mech. Anal.* **201**, no. 3, (2011), 1047–1067.

[28] T. Iwaniec, L.V. Kovalev and J. Onninen, Hopf Differentials and Smoothing Sobolev Homeomorphisms , *International Mathematics Research Notices* **14** (2012), 3256–3277.
[29] R.C. Kirby, Stable homeomorphisms and the annulus conjecture, *Ann. of Math.* **89** (1969), 575–582.
[30] R.C. Kirby, L.C. Siebenmann and C.T.C. Wall, The annulus conjecture and triangulation, *Notices Amer. Math. Soc.* **16** 432 (1969), abstract 69T-G27.
[31] J. Luukkainen, Lipschitz and quasiconformal approximation of homeomorphism pairs, *Topology Appl.* **109** (2001), 1–40.
[32] O. Martio, S. Rickman and J. Väisälä, Definitions for quasiregular mappings, *Ann. Acad. Sci. Fenn. Ser. A I* **448** (1969), 1–40.
[33] E.E. Moise, Affine structures in 3-manifolds. IV. Piecewise linear approximations of homeomorphisms, *Ann. of Math.* **55** (1952), 215–222.
[34] E.E. Moise, *Geometric topology in dimensions 2 and 3*, Graduate Texts in Mathematics 47. Springer. New York-Heidelberg 1977.
[35] C. Mora-Corral, Approximation by piecewise affine homeomorphisms of Sobolev homeomorphisms that are smooth outside a point, *Houston J. Math.* **35** (2009), 515–539.
[36] C. Mora-Corral and A. Pratelli, Approximation of Piecewise Affine Homeomorphisms by Diffeomorphisms, *J Geom. Anal* **24** (2014), 1398–1424.
[37] A. Pratelli, *On the bi-Sobolev planar homeomorphisms and their approximation*, Preprint.
[38] T. Radó, Über den Begriff Riemannschen Fläche, *Acta. Math. Szeged* **2** (1925), 101–121.
[39] T. Rado and P.V. Reichelderfer, *Continuous Transformations in Analysis*, Springer 1955.
[40] T.B. Rushing, *Topological embeddings*, Pure and Applied Mathematics 52. Academic Press, New York-London, 1973.

Department of Mathematical Analysis, Charles University, Sokolovská 83, 186 00 Prague 8, Czech Republic
E-mail address: daniel.campbell@uhk.cz; hencl@karlin.mff.cuni.cz

Department of Mathematics and Statistics, University of Jyväskylä, P.O. Box 35 (MaD), FIN-40014, Jyväskylä, Finland
E-mail address: ville.tengvall@jyu.fi

Bibliography

[1] E. Acerbi and N. Fusco, Semicontinuity problems in the calculus of variations, *Arch. Rat. Mech. Anal.* **86** (1984), 125–145.

[2] J.M. Ball, Convexity conditions and existence theorems in nonlinear elasticity, *Arch. Rat. Mech. Anal.* **63** (1977), 337–403.

[3] J.M. Ball, Discontinuous equilibrium solutions and cavitation in nonlinear elasticity, *Philos. Trans. R. Soc. Lond. A* **306** n.1496 (1982), 557–611.

[4] J.M. Ball, *Singularities and computation of minimizers for variational problems*, Foundations of computational mathematics (Oxford, 1999), 1–20, London Math. Soc. Lecture Note Ser., 284, Cambridge Univ. Press, Cambridge, 2001.

[5] J.M. Ball, *Progress and puzzles in Nonlinear Elasticity*, Proceedings of course on Poly-, Quasi- and Rank-One Convexity in Applied Mechanics, CISM, Udine, to appear.

[6] S. Daneri and A. Pratelli, *Smooth approximation of bi-Lipschitz orientation-preserving homeomorphisms*, to appear in Ann. Inst. H. Poincaré Anal. Non Linéaire.

[7] S. Hencl, Sobolev homeomorphism with zero Jacobian almost everywhere, *J. Math. Pures Appl.* **95** (2011), 444–458.

[8] S. Hencl and P. Koskela, *Lectures on mappings of finite distortion*, Lecture Notes in Mathematics 2096, Springer, Cham, 2014, 176pp.

[9] S. Hencl and J. Malý, Jacobians of Sobolev homeomorphisms, *Calc. Var. Partial Differential Equations* **38**, no. 1–2, (2010), 233–242.

[10] S. Hencl and A. Pratelli, *Diffeomorphic approximation of* $W^{1,1}$ *planar Sobolev homeomorphisms*, Preprint.

[11] S. Hencl, B. Vejnar, *Sobolev homeomorphism that cannot be approximated by diffeomorphisms in* $W^{1,1}$., Arch. Rational Mech. Anal., to appear.

[12] T. Iwaniec, L.V. Kovalev and J. Onninen, Diffeomorphic Approximation of Sobolev Homeomorphisms, *Arch. Rational Mech. Anal.* **201** no. 3 (2011), 1047–1067.

[13] T. Iwaniec, L.V. Kovalev and J. Onninen, Hopf Differentials and Smoothing Sobolev Homeomorphisms , *International Mathematics Research Notices* **14** (2012), 3256–3277.

[14] T. Iwaniec and J. Onninen, Monotone Sobolev Mappings of Planar Domains and Surfaces , *Arch. Rational Mech. Anal.* **219, 1** (2016), 159–181.

[15] C. Mora-Corral, Approximation by piecewise affine homeomorphisms of Sobolev homeomorphisms that are smooth outside a point, *Houston J. Math.* **35** (2009), 515–539.

[16] C. Mora-Corral and A. Pratelli, Approximation of Piecewise Affine Homeomorphisms by Diffeomorphisms, *J Geom. Anal* **24** (2014), 1398–1424.

[17] C.B. Morrey, Quasi-convexity and the semicontinuity of multiple integrals, *Pacific J. Math.* **2** (1952), 25–53.

[18] S. Müller, Q. Tang and B. S. Yan, On a new class of elastic deformations not allowing for cavitation, *Ann. Inst. Henri Poincaré* **11** (1994), 217–243.

[19] A. Pratelli, *On the bi-Sobolev planar homeomorphisms and their approximation*, Preprint.

[20] E. Radici, *A planar Sobolev extension theorem for piecewise linear homeomorphisms*, Preprint.