# Anisotropic Piecewise Linear Approximation

Andrianarivo Fabien Rabarison

Department of Mathematics and Statistics

University of Strathclyde

Glasgow, UK

August 2012

This thesis is submitted to the University of Strathclyde for the degree of Doctor of Philosophy in the Faculty of Science.

# Acknowledgements

I would like to give my special thanks to my supervisor, Dr. Oleg Davydov, for his invaluable input, advice and patience throughout the duration of my studies.

I would also like to thank my family and friends for their unequivocal love and support, especially at critical times.

Many thanks to the University of Strathclyde and the Scottish Foundation Council (SFC) which provided me with the necessary financial support for this work.

# Abstract

The subject of this thesis includes the design of new partitioning methods for the approximation of a function $f$ on a domain $\Omega \subset \mathbb{R}^d$, $d \geq 2$, by piecewise linear functions, and the derivation of errors estimations in $L_p$-norm and $W_p^1$-seminorm. In the two-dimensional setting, we develop a construction of a sequence of anisotropic triangulations, where the approximation provided by the piecewise linear interpolant for a given $f \in C^2(\Omega)$ with a positive definite Hessian, is asymptotically optimal in $L_p$-norm and in the same time optimal in $W_p^1$-seminorm with respect to the number of degrees of freedom. As a preparation for this result, we review various local error bounds for the interpolation by linear polynomials on a triangle, and derive a number of new estimates of this type. In addition, for functions of $d \geq 2$ variables, we propose a new approximation method, where several overlaying partitions of $\Omega$ are designed such that the sum of piecewise constant or piecewise linear polynomials over these partitions provides a better approximation order than the one obtainable by using a single partition.

# Contents

# List of Figures

# Chapter 1

# Introduction

We are interested in the approximation of a function $f \in C^2(\Omega)$ by using piecewise polynomials on a fixed bounded domain $\Omega \subset \mathbb{R}^d$, $d \geq 2$. The approximant $\bar{f}$ is a piecewise polynomial with respect to a partition $\Delta$ of the domain. As changing a partition means changing also the approximant, it is only natural to look for a partition where the approximant matches closely the target function. To achieve this, the domain needs to be appropriately divided into *cells* where local approximations are performed. We leave aside the *uniform* method where all cells of the partition have fixed shapes. Instead, we use *adaptive* methods where the domain is partitioned by using some properties of the function.

To begin with, we first investigate the two-dimensional case $d = 2$ (a cell is then a triangle) where the target function has a strictly positive definite Hessian on $\Omega$, and the approximant is required to be continuous. Obviously, we shall need local error analysis in order to obtain a global error estimation. Various local estimations are available from [1, 9, 13, 22, 23] (see Section 2.2). In Chapter 2, we review these standard estimates and the concept of an optimal triangle [2, 3, 29, 32] based on an intermediate quadratic polynomial approximation of a given function. We derive a number of new local error bounds, in particular with respect to the Sobolev seminorm $|\cdot|_{W^1_p}$, making preparations for the design of triangulations and estimates of Chapter 3.

Our partitioning method in Chapter 3 is inspired from [2, 3, 29, 30] in that the domain $\Omega$ is covered by two classes of triangles, the *regular* and *irregular* triangles. Regular triangles are designed by using the spectrum of the Hessian $H_f$ at some pre-selected points of $\Omega$, and the area covered by the irregular triangles is negligible. Our first aim consists in developing a triangulation $\Delta_N$ of at most $N$ triangles which allows us to estimate both the $L_p$-norm and $W_p^1$-seminorm of the approximation error. Such a triangulation requires a much more delicate procedure than the ones in [2, 3, 29] where only $L_p$-norm bounds have been obtained. The problem of designing optimal triangulation for the derivatives has been addressed in [30]. In contrast to [30], our triangulation is not only optimal in $W_p^1$-seminorm but also asymptotically optimal in $L_p$-norm which makes it more difficult to achieve and requires a new approach. Our triangulation differs in the following respects:

*i.* all regular triangles are isosceles;

*ii.* irregular triangles are obtained by drawing diagonals of a certain region obtained from connecting vertices from a regular region to a neighboring one;

*iii.* local $W_p^1$-seminorm error on all triangles is independent of their maximum angles.

The approximant $\bar{f}$ is the continuous piecewise linear polynomial which interpolates $f$ at the vertices of each triangle. With $1 \leq p < \infty$, our estimation for $\|f - \bar{f}\|_{L_p(\Omega)}$ is *optimal* in the sense that it cannot be improved on the so-called *admissible* triangulations, see (1.12), thereby achieving the same asymptotic estimation as in [2, 3, 29]. However, in addition to this estimation, careful checking of maximum interior angles enables us to estimate the $W_p^1$-seminorm $|f - \bar{f}|_{W_p^1(\Omega)}$, see (1.13), which is one of the key results of this thesis.

In addition, in Chapter 4 for any $d \geq 2$, we consider discontinuous piecewise linear approximants. Extending the work in [16], we approximate the target function $f$ by using a finite number of overlaying partitions $(\Delta_k)_{1 \leq k \leq n}$. This is a completely new method. Each partition is anisotropic and is obtained by splitting $\Omega$ using either *fixed* or *non-fixed* directions, the latter being related to the properties of the function. Each partition $\Delta_k$ contributes to the design of the approximant $\bar{f}$ which is a sum of piecewise polynomials. The set of overlaying partitions is denoted by $\mathcal{P}$, with $|\mathcal{P}|$ denoting the number of cells in all partitions. Thus, the total number of degrees of freedom of the sum of piecewise linear is $N = (d+1)|\mathcal{P}|$. Although the approximant is generally discontinuous, there are two main advantages from using these methods:

*a.* the improvement of the approximation orders in both $L_p$-norm and $W_p^1$-seminorm;

*b.* the simplicity of splitting partitions compared to designing a single anisotropic partition suitable for interpolation by a continuous piecewise linear function.

The gain in $b$ is self-explanatory, whereas the one in $a$ can be explained by an example: In the case of approximation by sums of piecewise linear polynomials, the approximation order $O(N^{-6/(2d+1)})$ is attained in the $L_p$-norm estimation, improving to $O(N^{-6/5})$ the approximation order $O(N^{-1})$ ($d = 2$) for continuous piecewise linear approximation on a single triangulation with number of degrees of freedom $N$. In addition, we attain the order $O(N^{-3/(2d+1)})$ in the $W_p^1$-seminorm estimation, which, for $d = 2$ gives $O(N^{-3/5})$ comparing to $O(N^{-1/2})$ achieved on a single triangulation.

The main results in the thesis are divided into three chapters. The second chapter on local linear estimations in $\mathbb{R}^2$ addresses the analysis of error estimations on triangles. This chapter is essential in order to understand what kind of estimations are needed after the triangulation in Chapter 3 is constructed. The third chapter discusses the construction of anisotropic triangulations, and also provides the asymptotic estimations in both $L_p$-norm and $W_p^1$-seminorm. The fourth chapter is devoted to the approximation of functions by using sums of piecewise polynomials. The general overview of these chapters is elaborated in the sections below.

## 1.1 Local linear estimations

Recall in the first setting that we consider a triangulation in two-dimensions, and the local approximation consists in interpolating the function at the triangles' vertices. Given a triangle $T$, the approximant $I_T f$ is a linear polynomial whose coefficients are determined by simple linear systems. To estimate the errors in $L_p$-norm and $W_p^1$-seminorm, one can use standard estimations such as the one found in [13],

$$|f - I_T f|_{W_p^k(T)} \leq C \frac{h_T^2}{\rho_T^k} |f|_{W_p^2(T)}, \quad k = 0, 1, \tag{1.1}$$

where $C$ is an absolute constant, $h_T$ is the diameter of $T$, and $\rho_T$ is the radius of the largest inscribed circle in $T$. Except in (1.1), we shall henceforth use the notation $\rho_T$ for the smallest height of a triangle $T$. As we shall see in Section 2.2, alternative estimations can be used when the so-called aspect ratio $\frac{h_T}{\rho_T}$ is high. For instance, many triangles in the triangulation constructed in Chapter 3 may have large aspect ratios which makes the above estimation unusable for $k = 1$.

We first start with considerable studies of the local approximation of a homogeneous quadratic polynomials $\pi$. The idea consists in using a homogeneous quadratic polynomial as an intermediate term (see (1.2) below) in order to approximate a given function. For a given triangle $T$, we use the measure of non-degeneracy $\rho_\pi(T)$ introduced in [31], and obtain local error estimations for the derivatives, in the longest edge and in the smallest height directions. Considering a triangle $T$, the polynomial $\pi$ is chosen to be the homogeneous quadratic polynomial $\pi_z \in \mathbb{H}_2$ whose coefficients are the entries of the Hessian matrix $H_f(z)$ of $f \in C^2(\Omega)$ at the point $z$, for some $z$ belonging to a neighborhood of $T$. Assume that $f$ behaves like $\pi_z$ in the neighborhood of $z$. Then, instead of using the right hand side $Ch_T^2|f|_{W_p^2(T)}$ of (1.1) ($k = 0$) we use the triangular inequality,

$$\|f - I_T f\|_{L_p(T)} \leq \|(f - \pi_z) - I_T(f - \pi_z)\|_{L_p(T)} + \|\pi_z - I_T \pi_z\|_{L_p(T)}, \tag{1.2}$$

and estimate separately the two terms on the right hand side of (1.2). Such a method is known as the quadratic model [18, 17, 19, 33], justified by the simple reason that the function behaves locally as a quadratic polynomial given by its Taylor expansion. First, if the point $z$ belongs to $T$, we show that

$$\|(f-\pi_z) - I_T(f-\pi_z)\|_{L_p(T)} \leq 6h_T^2|T|^{\frac{1}{p}} \max_{|z'-z''|\leq h_T} \|\pi_{z'} - \pi_{z''}\|, \tag{1.3}$$

where $\|\pi\|$ denotes the maximum (in absolute value) of the coefficients of $\pi \in \mathbb{H}_2$. For the second term, we introduce the *shape* function [29] defined by

$$K_p(\pi) := \inf_{|T|=1} \|\pi - I_T\pi\|_{L_p(T)}, \quad \pi \in \mathbb{H}_2. \tag{1.4}$$

An explicit expression for the shape function can be obtained (see Section 2.4). A triangle $T$ satisfying (1.4) is called an *optimal* triangle for $\pi$. The shape function plays a crucial role in the design of the so-called *regular* triangles described in Section 3.2.2. Briefly speaking, a regular triangle $T$ is a scaled and shifted version of some optimal triangle for $\pi_t$ for some $t \in \Omega$. Any other triangle of the triangulation is called an *irregular* triangle. A regular triangle is a triangle which is stretched in the directions of the eigenvalues of $H_f(t)$, with stretching constants proportional to powers of the condition number of $H_f(t)$. For such a triangle, given $z \in T$, the estimations in (1.3) and (1.4) ensure the existence of a constant $C_{t,z,T}$ (see Proposition 2.5.4) so that

$$\|f - I_Tf\|_{L_p(T)} \leq \Big(K_p(\pi_z) + C_{t,z,T}\Big)|T|^{1+\frac{1}{p}}, \tag{1.5}$$

and $C_{t,z,T} \to 0$ as $|z-t| \to 0$ and $h_T \to 0$. For the irregular triangles $T$, a coarse $L_p$-norm error bound in terms of $h_T$ is sufficient for our purposes.

We estimate also the $W_p^1$-seminorm of the error. In the literature, some conditions have to be met by the triangle in order to estimate the derivatives of the error, namely either the *minimum angle* condition which is referred to as the *Zlámal's condition* [13] or the *maximum angle* condition [1, 4, 23, 24], and they often appear in the error bounds. Note that the minimum angle condition implies the maximum angle condition [8], and is not applicable to anisotropic triangles. In [1], it is shown that for $f \in W_p^2(T)$,

$$|f - I_T f|_{L_p(T)} \lesssim h_T |D_{\boldsymbol{\sigma}_{h_T}} f|_{W_p^1(T)} + \rho_T |D_{\boldsymbol{\sigma}_{\rho_T}} f|_{W_p^1(T)}, \tag{1.6}$$

with constant depending on the maximum interior angle $\gamma(T)$, and where $\boldsymbol{\sigma}_{h_T}$ and $\boldsymbol{\sigma}_{\rho_T}$ are unit vectors associated with the triangle $T$, both being described in Figure 2.3. We are interested in the case of positive definite Hessian $H_f$ where $K_p(\pi_z)$ is attained at some isosceles triangle (see Chapter 3) with maximum angle satisfying $\gamma(T) < \frac{\pi}{2}$.

The estimation in (1.6) provides the first step in estimating the derivatives of the error (see Proposition 2.5.7) on regular triangles which are designed to be isosceles in Chapter 3. We cannot do the same for irregular triangles $T$ having an arbitrary shape with no control of their maximum interior angle $\gamma(T)$. Note also that we cannot use (1.1) for $k = 1$ since it may cause an overestimation if the aspect ratio $\frac{h_T}{\rho_T}$ is unbounded, for example when $T$ is strongly anisotropic. One of the two methods discussed in Section 2.6 consists in using an invertible affine map $\varphi$ with its condition number $\mathrm{cond}(\varphi)$ bounded and such that the maximum interior angle $\gamma(\varphi^{-1}(T))$ is well-distant from the flat angle. In this case, we show (see Lemma 2.6.1) that

$$|f - I_T f|_{W_p^1(T)} \lesssim \mathrm{cond}(\varphi)^2 h_T |f|_{W_p^2(T)}, \tag{1.7}$$

where $\mathrm{cond}(\varphi)$ denotes the condition number of the invertible matrix associated with $\varphi$. The above result is an alternative to (1.6), allowing us to estimate the derivatives of the approximation error on irregular regions. More reviews on local estimations can be found in [1, 24]. The second method consists in using the quadratic polynomial $\pi_z$ for some $z \in T$. If the triangle $T$ has its measure $\rho_{\pi_z}(T)$ bounded, then (see Proposition 2.6.2)

$$|f - I_T f|_{W_p^1(T)} \leq C\Big(\frac{h_T}{\rho_T}\omega(h_T) + \rho_{\pi_z}(T)\sqrt{|\det \pi_z|}\Big) h_T |T|^{\frac{1}{p}}, \tag{1.8}$$

where $\omega$ is the modulus of continuity of the function $z \mapsto \pi_z$ (see (2.75)). By an example, we show that the above estimation can be useful instead of ensuring that $\gamma(T)$ is bounded.

## 1.2 Triangulation and asymptotic estimates

The function $f$ is approximated on a square domain $\Omega$ which we triangulate according to the properties of $f$, namely by using the eigenvalues and eigenvectors of the Hessian $H_f$ at some pre-selected points. We assume that $f \in C^2(\Omega)$ with positive definite Hessian $H_f$. There are two principal goals for Chapter 3, the first one is the construction of an *optimal* triangulation where the interior angles of the triangles or their certain images are controlled, and the second one the derivation of asymptotic error in $L_p$-norm and $W_p^1$-seminorm by using the results in Chapter 2.

Given a triangulation $\Delta_N$ of $\Omega$ consisting of at most $N$ triangles, the approximant $f_N$ for $f$ is given by

$$f_N = \sum_{T \in \Delta_N} \ell_T \chi_T, \;\; \ell_T \in \Pi_2, \tag{1.9}$$

where $\Pi_2$ denotes the space of linear polynomials, whereas $\chi_T$ is the characteristic function on the triangle $T$. Each linear polynomial $\ell_T = I_T f$ interpolates $f$ at the vertices of the triangle $T \in \Delta_N$. The global error is obtained by combining the errors on each triangle,

$$\|f - f_N\|_{L_p} := \left( \sum_{T \in \Delta_N} \|f - I_T f\|_{L_p(T)}^p \right)^{\frac{1}{p}}. \tag{1.10}$$

With $f - f_N$ viewed as a distribution, we define its Sobolev seminorm by

$$|f - f_N|_{W_p^1(\Omega)} := \left( \sum_{T \in \Delta_N} |f - I_T f|_{W_p^1(T)}^p \right)^{\frac{1}{p}}. \tag{1.11}$$

It is crucial that $\Omega$ is carefully partitioned in order to obtain accurate estimations in $L_p$-norm (1.10) and in $W_p^1$-seminorm (1.11). The domain $\Omega$ is initially divided into $m^2$ sub-squares $S_i$, $i = 1, \dots, m^2$ of side length $r > 0$. As shown in Section 3.2.2 and Section 3.2.3, the construction of $\Delta_N$ is characterized by two main steps, the establishment of *regular* regions and the triangulation of *irregular* regions. Regular regions are obtained by grouping triangles that fit into the initially prescribed sub-squares $S_i$, $i = 1, \dots, m^2$, whereas irregular regions are the left over subspaces of $\Omega$. Similar work can be found in [2, 3, 29] and [30].

The main novelty of our method is a new approach to partitioning the irregular regions. For each $i = 1, \dots, m^2$, the *regular region* $R_i$ is defined by two systems of parallel segments $\mathcal{L}_i$ and $\bar{\mathcal{L}}_i$ with directional vectors $\mathbf{e}_i$ and $\bar{\mathbf{e}}_i$ derived from the eigenvectors of the Hessian of $f$ at the center of $S_i$. With the specific maneuvers described in Section 3.2.3, each segment $\ell_0$ belonging to $\mathcal{L}_i \cup \bar{\mathcal{L}}_i$ is extended into the neighboring regular regions of $R_i$, the extension being done in a direction of either $\mathbf{e}_i$ or $\bar{\mathbf{e}}_i$. The procedure of segment extensions creates the *irregular* and *boundary* regions which are polygons of at most six edges. The polygons are then divided into at most four triangles by drawing diagonals which are described in Section 3.2.4. Various interesting properties of the resulting triangulation are given in Section 3.3.

For the approximation of $f$ by using $\Delta_N$, with the Hessian $H_f$ being positive definite, we show in Theorem 3.4.8 of Section 3.4.6 that the resulting $L_p$-norm and $W_p^1$-seminorm of the approximation error satisfy the asymptotic estimates (1.12) and (1.13) below, with $\frac{1}{q} := 1 + \frac{1}{p}$,

$$\limsup_{N\to\infty} N\|f - f_N\|_{L_p(\Omega)} \le \left( \int_\Omega K_p(\pi_z)^q \mathrm{d}z \right)^{\frac{1}{q}}, \tag{1.12}$$

$$\limsup_{N\to\infty} N^{\frac{1}{2}}|f - f_N|_{W_p^1(\Omega)} \le C_p |f|^{\frac{1}{2}}_{W_p^2(\Omega)} \left( \int_\Omega K_p(\pi_z)^q \mathrm{d}z \right)^{\frac{1}{2q}}, \tag{1.13}$$

where $C_p$ is a constant depending on $p$ only. The estimation (1.12) is optimal in the sense that it cannot be improved amongst all *admissible*[1] triangulations as in [2, 3, 29]. However, at the same time $f_N$ satisfies (1.13), which cannot be

[1] $\Delta_N$ is *admissible* if $\sup_{T\in\Delta_N} \operatorname{diam}(T) \le CN^{-1/2}$, with $C$ independent of $N$.

guaranteed on the partitions suggested in [2, 3, 29] because the Delaunay triangulations of the irregular regions employed in these papers may contain triangles of arbitrary shapes leading to uncontrolled errors in $W_p^1$-seminorm estimations. By a different method, a $W_p^1$-seminorm error bound is obtained in [30] where the triangulation is designed to be asymptotically optimal for the derivatives but not for the function, and no error bound for $\|f - f_N\|_{L_p(\Omega)}$ is given.

## 1.3 Sums of piecewise polynomials

For the multi-dimensional setting, a partition $\Delta$ of the square domain $\Omega \subset \mathbb{R}^d$, $d \geq 2$, consists of convex sub-domains called *cells*. To approximate a function $f \in C(\Omega)$ in Chapter 4, we use several anisotropic partitions instead of one, though the resulting approximant $\bar{f}$ is discontinuous. The design of the partitions in $\mathcal{P}$ may or may not depend on the properties of the function.

Consider a function $f \in C(\Omega)$. Given a system $\mathcal{P} = \{\Delta^{(1)}, \dots, \Delta^{(n)}\}$ of several overlaying partitions of $\Omega$, where each cell of a partition is convex, we consider the space of sums of piecewise polynomials

$$S_k(\mathcal{P}) = \Big\{ \sum_{\nu=1}^{n} \sum_{\omega \in \Delta^{(\nu)}} q_{\nu,\omega} \chi_\omega \, : \, q_{\nu,\omega} \in \Pi_k^d \Big\}, \tag{1.14}$$

where $\Pi_k^d$, $k \geq 1$, denotes the space of polynomials of total degree $< k$ in $d$ variables, and where $\chi_\omega$ denotes the characteristic function of the cell $\omega$. A function in $S_k(\mathcal{P})$ is the sum of $n$ piecewise polynomials respectively belonging to $\Pi_k^d$. The corresponding best approximation error is measured in $L_p$-norm

$$E_k(f, \mathcal{P})_p := \inf_{s \in S_k(\mathcal{P})} \|f - s\|_{L_p(\Omega)}, \qquad 1 \leq p \leq \infty.$$

The cardinality $|\mathcal{P}| = \sum_{\nu=1}^{n} |\Delta^{(\nu)}|$ is the sum of the cardinalities of each partition.

The best approximation error on a cell $\omega \in \Delta$ is defined by

$$E_k(f, \Delta)_p := \inf_{s \in S_k(\Delta)} \|f - s\|_{L_p(\Omega)}. \qquad 1 \le p \le \infty. \tag{1.15}$$

As the most used inequality in Chapter 4 for our local analysis on $\omega$ (see Section 4.1), the Bramble-Hilbert lemma for convex domains (see [21]) states that there is a polynomial $q \in \Pi_k^d$, $k \ge 0$, such that

$$|f - q|_{W_p^r(\omega)} \le \rho_{d,k} \operatorname{diam}^{k-r}(\omega) |f|_{W_p^k(\omega)}, \quad r = 0, \dots, k, \tag{1.16}$$

where $\rho_{d,k}$ denotes a constant depending only on $d$ and $k$.

Extending the work in [16] on approximations by constants, we design several partitions of $\Omega$ to achieve the estimation below for $f \in W_p^2(\Omega)$,

$$E_1(f, \mathcal{P})_p \le C_d |\mathcal{P}|^{-2/(d+1)} (|f|_{W_p^1(\Omega)} + |f|_{W_p^2(\Omega)}), \tag{1.17}$$

where $C_d$ is a constant depending only on $d$, improving the saturation order $E_k(f, \Delta)_p = O(|\Delta|^{-k/d})$ which is obtainable on an isotropic single partition. Our extension to approximation by sums of linear polynomials yields that, for $f \in W_p^3(\Omega)$, there is a function $s \in S_2(\mathcal{P})$ so that

$$\|f - s\|_{L_p(\Omega)} \le C_1 |\mathcal{P}|^{-6/(2d+1)} (|f|_{W_p^2(\Omega)} + |f|_{W_p^3(\Omega)}), \tag{1.18}$$

$$|f - s|_{W_p^1(\Omega)} \le C_2 |\mathcal{P}|^{-3/(2d+1)} (|f|_{W_p^2(\Omega)} + |f|_{W_p^3(\Omega)}), \tag{1.19}$$

where $C_1, C_2$ are absolute constants.

## 1.4 Various notations

Given a triangle $T$, its diameter and smallest height are denoted by $h_T$ and $\rho_T$. They are called the *length scales* of $T$. While confusion does not occur, we simply use $h$ and $\rho$. The area of $T$ is denoted by $|T|$. The triangle $T$ is termed *isotropic* if there exists a constant $C$ such that $\frac{h}{\rho} \leq C$. The ratio $\frac{h}{\rho}$ is called the *aspect ratio* of $T$. otherwise it is termed *anisotropic*. In general, an isotropic triangle is a triangle whose edges are all comparable in length, or a triangle whose interior angles are not too small nor too large. Anisotropic triangles on the other hand are characterized by long diameters and small heights, presenting one or two very small interior angles.

A triangulation $\Delta$ of a bounded domain $\Omega \subset \mathbb{R}^2$ is a partition of $\Omega$ into triangles where the intersection of any two of them is either an empty set, a common vertex or a common edge. It satisfies the *maximum angle* condition (see [1]) if there exists an angle $\gamma_* < \pi$ such that $\gamma(T) \leq \gamma_*$ for any triangle $T \in \Delta$, with $\gamma(T)$ denoting the maximum interior angle of $T$. It satisfies the *minimum angle* condition if there is an angle $\alpha_* > 0$ such that $\alpha(T) \geq \alpha_*$ for any triangle $T \in \Delta$, with $\alpha(T)$ denoting the minimum interior angle of $T$. We say that $\Delta$ is *isotropic* if all of its triangles are isotropic, i.e. satisfy the minimum angle condition, otherwise it is an *anisotropic* triangulation, i.e. if $\Delta$ presents some triangles which are anisotropic.

Given a matrix $A$, we denote by $\phi_A$ the associated linear map. We say that $\phi_A$ is invertible if $A$ is a non-degenerate matrix. The singular values of $\phi_A$ are those of $A$, and its condition number is defined by

$$\mathrm{cond}(\phi_A) := \mathrm{cond}(A).$$

A *partition* $\Delta$ of $\Omega$ is a set of *cells* $\omega \subset \Omega$ possessing the two properties:

i. For every $\omega, \omega' \in \Delta$, $|\omega \cap \omega'| = 0$ if $\omega \neq \omega'$;

iii. $\sum_{\omega \in \Delta} |\omega| = |\Omega|$,

where $|\omega|$ denotes the Lebesgue measure ($d$-dimensional volume) of $\omega$ (for $d = 2$, the partition is a triangulation). A partition is said to be *convex* if each cell $\omega$ is a convex domain. With a slight abuse of notation, we denote by $|D|$ the cardinality of a finite set $D$, so that $|\Delta|$ stands for the number of cells $\omega$ in $\Delta$.

The approximant $\bar{f}$ is the piecewise polynomial given by

$$\bar{f} = \sum_{\omega \in \Delta} f_\omega \chi_\omega, \tag{1.20}$$

where $f_\omega$ is a polynomial which approximates $f$ on the cell $\omega$, and $\chi_\omega$ the characteristic function of $\omega$, with $\chi_\omega(x) = 1$ if $x \in \omega$ and 0 otherwise.

Given two numbers $a, b \in \mathbb{R}$, the notation $a \lesssim b$ is used if there exists a constant $C$ independent of $a, b$ such that $a \leq Cb$. The notation $a \sim b$ is used when there are two constants $C_1$ and $C_2$ such that $C_1 b \leq a \leq C_2 b$.

We denote by $D_x f$, $D_y f$ the partial derivatives of $f$, and by $D_{\boldsymbol{\sigma}} f$ the derivative of $f$ in the direction of a unit vector $\boldsymbol{\sigma}$. We also use double indices for the second order partial derivatives: $D^2_{xx} := D^2_x$, $D^2_{xy} := D_x D_y$, $D^2_{\boldsymbol{\sigma}\boldsymbol{\sigma}} := D^2_{\boldsymbol{\sigma}}$, $D^2_{\boldsymbol{\sigma}\boldsymbol{\tau}} := D_{\boldsymbol{\sigma}} D_{\boldsymbol{\tau}}$, etc., where both $\boldsymbol{\sigma}$ and $\boldsymbol{\tau}$ are unit vectors.

Given two real sequences $(x_1, \ldots, x_n), (y_1, \ldots, y_n)$, $n \in \mathbb{N}$, the Hölder inequality states that, for $p, q \in [1, \infty)$ such that $\frac{1}{p} + \frac{1}{q} = 1$, we have

$$\sum_{k=1}^n |x_k y_k| \leq \left( \sum_{k=1}^n |x_k|^p \right)^{\frac{1}{p}} \left( \sum_{k=1}^n |y_k|^q \right)^{\frac{1}{q}}. \tag{1.21}$$

The Euclidean norm of $\nabla f$ on a bounded convex domain $\omega$ is defined by

$$\|\nabla\|_{L_p(\omega)} := \left\| \left( \sum_{k=1}^d \left| \frac{\partial f}{\partial x_k} \right|^2 \right)^{\frac{1}{2}} \right\|_{L_p(\omega)}. \tag{1.22}$$

It has been shown in [16] that, for any $1 \leq p \leq \infty$,

$$\|\nabla f\|_{L_p(\omega)} \leq |f|_{W^1_p(\omega)} \leq d^{\max\{\frac{1}{2}, 1-\frac{1}{p}\}} \|\nabla f\|_{L_p(\omega)}. \tag{1.23}$$

# Chapter 2

# Local linear approximations

In this chapter, we analyze local interpolation errors on triangles in both $L_p$-norm and $W_p^1$-seminorm. Apart from the standard estimations found in the literature, which we discuss hereafter, we also provide estimations based on the minimization of errors on unit triangles. The results in this chapter are the principal sources of estimations for Chapter 3 where we approximate a given function $f \in C^2(\Omega)$ on a constructed triangulation $\Delta_N$ of $\Omega$. Indeed, it is necessary to bound the error on each triangle $T \in \Delta_N$ by investigating their properties.

We start by studying the local estimations for homogeneous quadratic polynomial approximations which we use as intermediate steps in order to estimate the local error when approximating the function $f$. Such a method can be found in [2, 3, 29] where the function is assumed to behave locally as a quadratic polynomial. A general review of local estimations is provided below, however as we shall see later, standard estimates are difficult to apply especially in the case of $W_p^1$-seminorm estimations where the aspect ratio $\frac{h_T}{\rho_T}$ may be unbounded, or when the maximum interior angle $\gamma(T)$ is nearly the flat angle. We thus need some other approaches in order to estimate the derivatives of the approximation error.

In Section 2.1 we provide a general review of homogeneous quadratic polynomials and local approximations on triangles. Similarly, in Section 2.2 we review standard local estimations from the literature. In Section 2.3 we study the local approximations of quadratic polynomials by using a reference triangle. In particular, in Section 2.3.3, we obtain new results on the bounds for the directional derivatives of the local errors. Following the argument in [29], we introduce in Section 2.4 the concept of optimal triangles and discuss in Section 2.5 various local estimations on nearly optimal triangles designed according to the behavior of the Hessian $H_f$ at some specific points. In Section 2.6 we present two methods in order to estimate the derivatives of the error on non-optimal triangles.

## 2.1 Preliminaries and notations

In this section, we present the basic yet important steps in order to obtain the estimations of this and next chapters. We use $[x\ y]^t$ to denote a column vector where the superscript $t$ denotes the transpose operator. We denote by $[\mathbf{v}_1\ \mathbf{v}_2]$ a matrix whose column-vectors are $\mathbf{v}_1$ and $\mathbf{v}_2$.

### 2.1.1 On homogeneous quadratic polynomials

We will often denote by $\pi$ a homogeneous quadratic polynomial $\pi(x,y) = ax^2 + 2bxy + cy^2$, with $a, b, c \in \mathbb{R}$, which can be represented by the symmetric matrix $Q_\pi := \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ for which (2.1) below holds, for all $x, y \in \mathbb{R}$,

$$\pi(x,y) = [x\ y]\, Q_\pi\, [x\ y]^t = [x\ y] U_\pi \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} U_\pi^t [x\ y]^t. \tag{2.1}$$

The right hand side of (2.1) can be viewed as the Schur decomposition of $Q_\pi$, with $\lambda_1, \lambda_2$ being the eigenvalues and $U_\pi$ denoting the orthonormal matrix whose columns are the unit eigenvectors $\mathbf{v}_1, \mathbf{v}_2$ of $Q_\pi$ corresponding to $\lambda_1, \lambda_2$, respectively. When $\pi$ is convex, both eigenvalues $\lambda_1, \lambda_2$ of $Q_\pi$ are positive, whereas when $\pi$ is concave $\lambda_1, \lambda_2$ are negative. Note that (2.1) remains true when switching the positions of $\lambda_1$ and $\lambda_2$ and accordingly the positions of $\mathbf{v}_1$ and $\mathbf{v}_2$ since, by writing $U_\pi = [\mathbf{v}_1\ \mathbf{v}_2] = \begin{bmatrix}\alpha & \beta \\ \gamma & \delta\end{bmatrix}$ where $\alpha, \beta, \gamma, \delta \in \mathbb{R}$, we have

$$[\mathbf{v}_1\ \mathbf{v}_2]\begin{bmatrix}\lambda_1 & 0 \\ 0 & \lambda_2\end{bmatrix}[\mathbf{v}_1\ \mathbf{v}_2]^t = \begin{bmatrix}\lambda_1\alpha^2 + \lambda_2\beta^2 & \lambda_1\alpha\gamma + \lambda_2\beta\delta \\ \lambda_1\alpha\gamma + \lambda_2\beta\delta & \lambda_1\gamma^2 + \lambda_2\delta^2\end{bmatrix} = [\mathbf{v}_2\ \mathbf{v}_1]\begin{bmatrix}\lambda_2 & 0 \\ 0 & \lambda_1\end{bmatrix}[\mathbf{v}_2\ \mathbf{v}_1]^t.$$

Moreover, $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthogonal to each other, $\mathbf{v}_1$ can be replaced by $-\mathbf{v}_1$ and $\mathbf{v}_2$ by $-\mathbf{v}_2$. Hence, without loss of generality, we assume that

$$|\lambda_1| \leq |\lambda_2| \quad \text{and} \quad U_\pi = R_\mu := \begin{bmatrix}\cos\mu & -\sin\mu \\ \sin\mu & \cos\mu\end{bmatrix}, \tag{2.2}$$

that is, $U_\pi$ is the rotation matrix $R_\mu$ with angle $\mu = \mu(\pi)$ which is the smallest possible angle of counterclockwise rotation that transforms the coordinate unit vectors $[1\ 0]^t$ and $[0\ 1]^t$ into $\mathbf{v}_1, \mathbf{v}_2$, respectively. Note that $\mu$ is then the smallest between the non-negative angle from $[1\ 0]^t$ to $\mathbf{v}_1$ and that from $[1\ 0]^t$ to $-\mathbf{v}_1$, as shown in Figure 2.1, and necessarily $\mu \in [0, \pi)$. For example, $\mu = 0$ for the polynomial $\pi_0(x, y) = x^2 + y^2$.

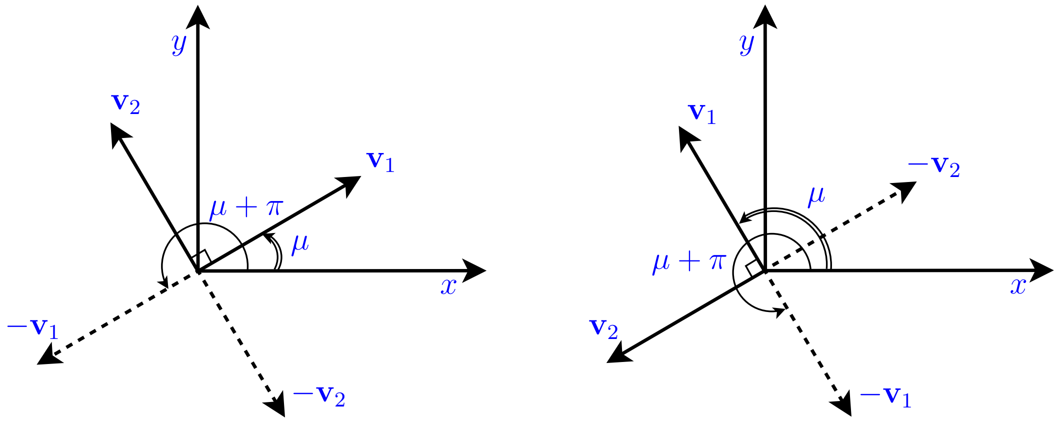


Figure 2.1: Positions of the eigenvectors of $Q_\pi$ and choice of $\mu$.

### 2.1.2 Norms on the space $\mathbb{H}_2$

The determinant and condition number of $\pi \in \mathbb{H}_2$ are defined by $\det \pi := \det Q_\pi = \lambda_1\lambda_2$ and $\operatorname{cond} \pi := \operatorname{cond} Q_\pi = |\lambda_2/\lambda_1|$, with $\lambda_1, \lambda_2$ as described in the previous section. We say that $\pi$ is *degenerate* if $\det \pi = 0$, and *non-degenerate* otherwise. The function $\pi \mapsto \|\pi\|_2 := \|Q_\pi\|_2 = |\lambda_2|$ is a norm over the linear space $\mathbb{H}_2$ of homogeneous quadratic polynomials since it is induced by the Euclidean vector norm $\|\cdot\|_2$. Also, writing $\pi(x,y) = ax^2 + 2bxy + cy^2$,

$$\|\pi\| = \max\{|a|, |2b|, |c|\} \tag{2.3}$$

is an easily proved norm on $\mathbb{H}_2$. Simple computations show that the eigenvalues of $Q_\pi$ are exactly $\frac{a+c\pm\sqrt{(a-c)^2+4b^2}}{2}$. If $a, c$ have the same sign, then clearly $\max\{|a|, |2b|, |c|\} \leq \max\{|a + c - \sqrt{(a-c)^2 + 4b^2}|, |a + c + \sqrt{(a-c)^2 + 4b^2}|\}$ whereas if they have different signs, then $\max\{|a|, |2b|, |c|\} \leq \sqrt{(a-c)^2 + 4b^2}$ holds. Thus

$$\frac{1}{2}\|\pi\| \leq \|Q_\pi\|_2 \leq \frac{3}{2}\|\pi\|, \tag{2.4}$$

holds, where the second inequality is obtained as follows: The constant in the equivalence is obtained by maximizing $g(a, 2b, c) := \frac{|a+c\pm\sqrt{(a-c)^2+4b^2}|}{2}$ for $|a|, |2b|, |c| \leq 1$. Clearly $g$ is an increasing function of each of its variables, hence the maximum is attained when $a = c = 1$ and $2b = 1$, with maximum $\frac{3}{2}$. The inequality below is also easily proved for any triangle $T$ containing the origin,

$$\|\pi\|_{L_p(T)} \leq 3h_T^2|T|^{\frac{1}{p}}\|\pi\|. \tag{2.5}$$

Given a fixed triangle $T$ and $\mathbf{e} = (x_\mathbf{e}, y_\mathbf{e})$ an edge of $T$, we denote $\pi(\mathbf{e}) := \pi(x_\mathbf{e}, y_\mathbf{e})$. Then the function

$$\pi \mapsto \|\pi\|_T := |T|^{\frac{1}{p}} \max\{|\pi(\mathbf{e})| : \mathbf{e} \text{ edge of } T\}, \tag{2.6}$$

is also a norm over the space of quadratic polynomials: To see this, we verify the axioms of a norm. For any $\alpha \in \mathbb{R}$ and any quadratic polynomials $\pi, \pi'$, we have the following.

a. $\|\pi\|_T \geq 0$ and $\|\pi\|_T = 0$ if and only if $\pi \equiv 0$: it is clear that $\|\pi\|_T \geq 0$ and if $\pi \equiv 0$, then $\|\pi\|_T = 0$. If $\pi(\mathbf{e}) = 0$ for any edge $\mathbf{e}$ of $T$, then $\pi$ vanishes at three distinct points. Since $\pi$ is quadratic, necessarily $\pi \equiv 0$;

b. it is easy to see that

$$\|\alpha\pi\|_T = |\alpha||T|^{\frac{1}{p}} \max\{|\pi(\mathbf{e})| : \mathbf{e} \text{ edge of } T\} = |\alpha|\|\pi\|_T;$$

c. we have that

$$\begin{aligned}\|\pi+\pi'\|_T &= |T|^{\frac{1}{p}} \max\{|(\pi+\pi')(\mathbf{e})| : \mathbf{e} \text{ edge of } T\} \\ &\leq |T|^{\frac{1}{p}} \max\{|\pi(\mathbf{e})| + |\pi'(\mathbf{e})| : \mathbf{e} \text{ edge of } T\} \\ &\leq |T|^{\frac{1}{p}} \Big( \max\{|\pi(\mathbf{e})| : \mathbf{e} \text{ edge of } T\} + \max\{|\pi'(\mathbf{e})| : \mathbf{e} \text{ edge of } T\}\Big) \\ &= \|\pi\|_T + \|\pi'\|_T.\end{aligned}$$

We have the result below.

**Lemma 2.1.1** ([31, Proposition 2.1])**.** *There exist absolute constants $c_1$ and $c_2$ such that, for any homogeneous quadratic polynomial $\pi$ and a triangle $T$,*

$$c_1\|\pi\|_T \leq \|\pi\|_{L_p(T)} \leq c_2\|\pi\|_T, \tag{2.7}$$

*where $\|\cdot\|_T$ is defined in (2.6) and $c_1, c_2$ are absolute constants.*

### 2.1.3 Reference triangle

Henceforth, the triangles that we consider are always non-degenerate, i.e non-empty and with a non-zero area. Given a triangle $T$, its edges are oriented in a counterclockwise direction. We denote by $\boldsymbol{\sigma}_{h_T}$ the unit vector on the longest edge of $T$ and $\boldsymbol{\sigma}_{\rho_T}$ the corresponding inner normal, they are as shown in Figure 2.3.

Let us fix the triangle $\hat{T}$ to have the vertices $(0,0)$, $(1,0)$, $(\frac{1}{2},1)$ as shown on the left of Figure 2.2. We shall refer $\hat{T}$ as to the *reference* triangle, its area is $\frac{1}{2}$ and its length scales are given by $\hat{h} = \sqrt{1+\frac{1}{4}} = \frac{\sqrt{5}}{2}$ and $\hat{\rho} = 2\frac{|\hat{T}|}{\hat{h}} = \frac{2\sqrt{5}}{5}$. For any arbitrary triangle $T$ there exists an invertible affine map $\psi$ such that

$$T = \psi(\hat{T}), \quad \psi(\hat{\mathbf{x}}) := M\hat{\mathbf{x}} + \mathbf{t}, \quad \hat{\mathbf{x}} \in \hat{T}, \tag{2.8}$$

where $M$ is a non-singular matrix and $\mathbf{t}$ a translation vector. Let the vectors $\hat{\boldsymbol{\sigma}} = [1\ 0]^t$, $\hat{\boldsymbol{\tau}} = [\frac{1}{2} - \frac{\delta}{h}\ \ 1]^t$ be fixed, with $\delta \in [0, h]$ depending only on $T$ as shown on the right of Figure 2.2. We can decompose $M$ by means of

$$M = R_\theta B_T, \quad B_T := \begin{bmatrix} h & \delta - \frac{h}{2} \\ 0 & \rho \end{bmatrix}, \tag{2.9}$$

where $R_\theta$ is a counterclockwise rotation matrix with angle $\theta \in [0, 2\pi)$ as shown in Figure 2.3. Then $B_T\hat{\boldsymbol{\sigma}} = [h\ 0]^t$ and $B_T\hat{\boldsymbol{\tau}} = [0\ \rho]^t$. Observe that $B_T(\hat{T})$, as shown on the right of Figure 2.2, is obtained by rotating $T$ in such a way that the longest edge becomes parallel to the $x$-axis, then shifting the rotated triangle to the first quadrant of the plane such that one of its vertices coincides with the origin. Necessarily, the longest edge of the resulting triangle lies on the $x$-axis. Moreover, since $|M\hat{\boldsymbol{\sigma}}| = |R_\theta[h\ 0]^t| = h$ and $|M\hat{\boldsymbol{\tau}}| = |R_\theta[0\ \rho]^t| = \rho$, it holds that

$$\frac{M\hat{\boldsymbol{\sigma}}}{|M\hat{\boldsymbol{\sigma}}|} = \frac{R_\theta[h\ 0]^t}{h} = \boldsymbol{\sigma}_h \quad \text{and} \quad \frac{M\hat{\boldsymbol{\tau}}}{|M\hat{\boldsymbol{\tau}}|} = \frac{R_\theta[0\ \rho]^t}{\rho} = \boldsymbol{\sigma}_\rho. \tag{2.10}$$

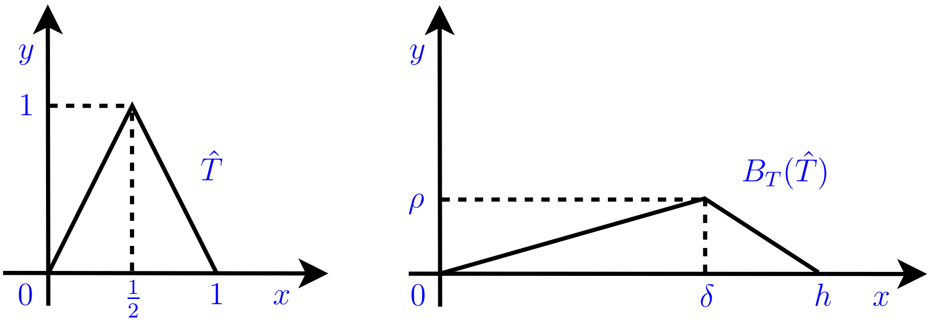


Figure 2.2: The triangle $\hat{T}$ and its image $B_T(\hat{T})$.

**Remark 2.1.1.** In Figure 2.2, the triangle $\hat{T}$ can be represented by the vectors $[1\ 0]^t$ and $[\frac{1}{2}\ 1]^t$ whose images under $B_T$ may represent the triangle $B_T(\hat{T})$, with

$$B_T[1\ 0]^t = [h\ 0]^t \quad \text{and} \quad B_T[\tfrac{1}{2}\ 1]^t = [\delta\ \rho]^t,$$

justifying the choice of $B_T$ in (2.9).

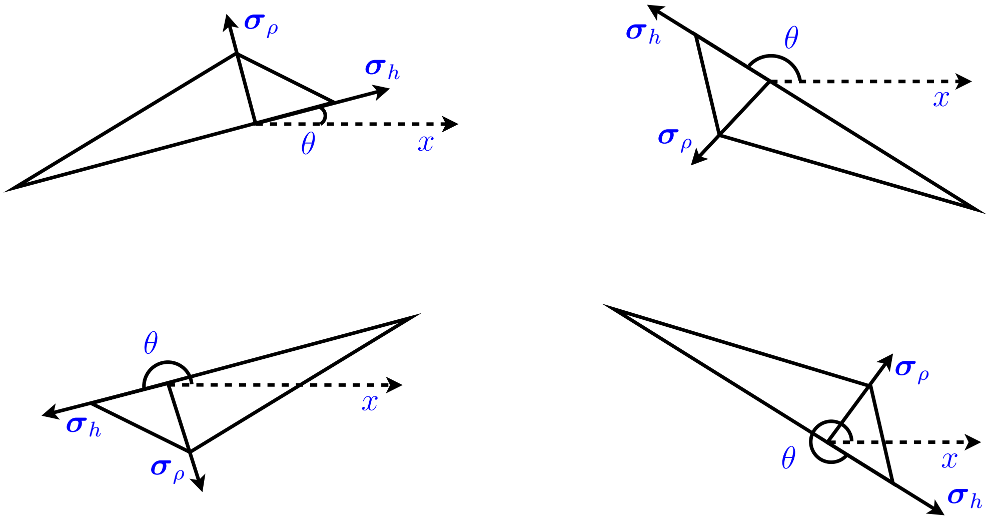


Figure 2.3: Several possible positions of the triangle $T$, the unit vector $\boldsymbol{\sigma}_h$ and $\boldsymbol{\sigma}_\rho$, and the angle of rotation $\theta$ in (2.9).

**Remark 2.1.2.** In Figure 2.3 the edges of $T$ are counterclockwise oriented, and from the definitions of $\boldsymbol{\sigma}_\rho$ and $\boldsymbol{\sigma}_h$ on page 18, we see that $\boldsymbol{\sigma}_\rho$ is obtained from $\boldsymbol{\sigma}_h$ by a counterclockwise rotation of angle $\frac{\pi}{2}$.

### 2.1.4 Local approximations on triangles

We shall now start estimating the local approximation error on a given triangle $T$. Let $I_T$ denote the linear polynomial interpolation operator on $T$ so that

$$I_T : f \in C(T) \mapsto I_T f := I_T(f) \in C(T),$$

where $(I_T f)(v) = f(v)$ for any vertex $v$ of $T$. Given an invertible affine map $\psi$, the vertices of the triangle $\psi(T)$ are the images of the vertices of $T$. This means that, for any vertex $v$ of $T$, $v' = \psi(v)$ is a vertex of $\psi(T)$, and therefore $I_T$ and $\psi$ are commutative in the following sense,

$$(I_{\psi(T)} f) \circ \psi(v) = (I_{\psi(T)} f)(v') = f(v') = f \circ \psi(v) = I_T(f \circ \psi)(v). \qquad (2.11)$$

It follows that $(I_{\psi(T)}f)\circ\psi$ and $I_T(f\circ\psi)$ coincide at three vertices of $T$, and since both are linear polynomials, they are necessarily the same function, that is,

$$I_T(f\circ\psi)(z) = (I_{\psi(T)}f)\circ\psi(z), \quad z\in\mathbb{R}. \tag{2.12}$$

Considering a homogeneous quadratic polynomial $\pi$, the norm $\|I_T\pi\|_{L_\infty(T)}$ is attained at one of the vertices of the triangle $T$ since $I_T\pi$ defines a plane. Recalling that $I_T\pi$ interpolates $\pi$ at the vertices of $T$, clearly

$$\|I_T\pi\|_{L_\infty(T)} \le \|\pi\|_{L_\infty(T)},$$

which shows that the norm of the operator $I_T$ is one independently of $T$. We thus obtain

$$\|I_T\pi\|_{L_p(T)} \le |T|^{\frac{1}{p}}\|I_T\pi\|_{L_\infty(T)} \le |T|^{\frac{1}{p}}\|\pi\|_{L_\infty(T)}. \tag{2.13}$$

We denote by $e_T(f)$ the $L_p$-norm of the error on $T$, with $1\le p\le\infty$,

$$e_T(f) := \|f - I_Tf\|_{L_p(T)}. \tag{2.14}$$

The result below is inspired from some proof of the results in [29]. However, thanks to (2.5) we are able to provide a more detailed expression of the terms in the right hand side of (2.15).

**Lemma 2.1.2.** *There exists a constant $C$ depending only on $p$ such that, for any homogeneous quadratic polynomials $\pi,\pi'\in\mathbb{H}_2$ and any triangle $T$ possessing a vertex at the origin,*

$$|e_T(\pi) - e_T(\pi')| \le Ch_T^2|T|^{\frac{1}{p}}\|\pi-\pi'\|, \tag{2.15}$$

*where the norm $\|\cdot\|$ is defined in (2.3).*

*Proof.* With the triangle $T$ fixed, it is easy to prove that the function $\pi \mapsto$

$e_T(\pi)$ defines a norm on $\mathbb{H}_2$. It is only a seminorm on the space of all quadratic polynomials. Hence, by considering the fixed reference triangle $\hat{T}$ on the left of Figure 2.2, the existence of a constant $C$ depending only on $p$ is guaranteed by the equivalence of the norms $\|\cdot\|_{L_p(\hat{T})}$ and $e_{\hat{T}}$ on $\mathbb{H}_2$, that is,

$$e_{\hat{T}}(\pi) \leq C\|\pi\|_{L_p(\hat{T})}, \quad \pi \in \mathbb{H}_2. \tag{2.16}$$

We show that the constant $C$ above remains absolute for the triangle $T$: Consider the affine map $\psi$ for which $T = \psi(\hat{T})$ holds, where $\psi$ is given in (2.8). By simple change of variables, clearly

$$e_T(\pi) = \left(\int_{\hat{T}} |\det M||\pi \circ \psi(z) - (I_T\pi) \circ \psi(z)|^p \,\mathrm{d}z\right)^{\frac{1}{p}} = |\det M|^{\frac{1}{p}} e_{\hat{T}}(\pi \circ \psi),$$

where $M$ is the invertible matrix occurring in (2.8), with $|T| = |\det M||\hat{T}| = \frac{1}{2}|\det M|$. Since $T$ has a vertex at the origin, the map $\psi$ is designed so that its translation vector is null, $\mathbf{t} = 0$: With the vertices of $T$ being $(0,0), (x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$, $\psi$ is the linear map associated with the matrix $M$ that maps $(0,0)$ to itself, $(1,0)$ to $(x_1, y_1)$ and $(\frac{1}{2}, 1)$ to $(x_2, y_2)$, more precisely $M = \begin{bmatrix} x_1 & x_2 - x_1/2 \\ y_1 & y_2 - y_1/2 \end{bmatrix}$. The fact that $\pi$ is a homogeneous quadratic polynomial and $\psi$ a linear map proves that $\pi \circ \psi \in \mathbb{H}_2$. We deduce from (2.16) that $e_{\hat{T}}(\pi \circ \psi) \leq C\|\pi \circ \psi\|_{L_p(\hat{T})}$. Thus

$$e_T(\pi) \leq C(2|T|)^{\frac{1}{p}}\|\pi \circ \psi\|_{L_p(\hat{T})} = C\|\pi\|_{L_p(T)}, \tag{2.17}$$

with $C$ being the absolute constant in (2.16). For any homogeneous quadratic polynomial $\pi' \in \mathbb{H}_2$, we use the triangular inequality to obtain

$$\begin{aligned} e_T(\pi) &= \|(\pi' - I_T\pi') + (\pi - \pi') - I_T(\pi - \pi')\|_{L_p(T)} \\ &\leq e_T(\pi') + e_T(\pi - \pi'), \end{aligned}$$

which, together with (2.17), yields

$$|e_T(\pi) - e_T(\pi')| \leq e_T(\pi - \pi') \leq C\|\pi - \pi'\|_{L_p(T)},$$

proving that $e_T$ is Lipschitz. Combining the above result with (2.5) yields the desired result. □

The estimation in the above lemma will be used at a later stage in Chapter 3 when estimating the error bounds on the so-called regular triangles.

## 2.2 Standard estimations

In this section, standard estimations are provided which are of great use for the rest of this thesis, namely in Section 2.5-2.6, as well as in Section 3.4 of Chapter 3.

Let $\hat{T}$ be the reference triangle shown on the left of Figure 2.2, and $T$ be an arbitrary triangle such that $T = \psi(\hat{T})$, where $\psi$ is the invertible affine map as in (2.8). Recall that $h$ and $\rho$ are respectively the diameter and smallest height of the triangle $T$.

**Lemma 2.2.1** ([13, Theorem 3.1.5])**.** *There exists a constant $C$ such that, for all functions $v \in W_p^2(T)$, the error $v - I_T v$ satisfies*

$$|v - I_T v|_{W_p^m(T)} \leq C\frac{h^2}{\rho^m}|v|_{W_p^2(T)}, \quad m = 0, 1. \tag{2.18}$$

For $m = 0$, the error bound $Ch^2|v|_{W_p^2(T)}$ on the right hand side of (2.18) is commonly used when approximating $v$ on either isotropic or anisotropic triangles. In the simple case where $v$ is a quadratic polynomial of the form $v(x, y) = ax^2 + by^2$, with $a, b \in \mathbb{R}$, the fact that

$$\begin{aligned} |v|_{W_p^2(T)} &= \left(\int_T \left(|D^2_{xx}v(z)|^p + 2|D^2_{xy}v(z)|^p + |D^2_{yy}v(z)|^p\right)\mathrm{d}z\right)^{\frac{1}{p}} \\ &= 2|T|^{\frac{1}{p}}(|a|^p + |b|^p)^{\frac{1}{p}} \\ &\leq 2|T|^{\frac{1}{p}}(|a| + |b|), \end{aligned}$$

implies that the error satisfies $\|v - I_T v\|_{L_p(T)} \leq Ch_T^2|T|^{\frac{1}{p}}(|a| + |b|)$. For $m = 1$, the error bound $C\frac{h^2}{\rho}|v|_{W_p^2(T)}$ on the right hand side of (2.18) may be coarse when the aspect ratio $\frac{h}{\rho}$ is not controlled. In (2.20) below, an alternative (sharper) bound is given subject to the condition that the interior angles of $T$ are far from $\pi$. The constant in the estimation depends on the maximum interior angle $\gamma(T)$, though the dependency of that constant to $\gamma(T)$ is unknown.

**Lemma 2.2.2** ([1, Theorem 5.5])**.** *For all functions $v \in W_p^2(T)$,*

$$\|v - I_T v\|_{L_p(T)} \lesssim h^2 \|D^2_{\boldsymbol{\sigma}_h \boldsymbol{\sigma}_h} v\|_{L_p(T)} + h\rho \|D^2_{\boldsymbol{\sigma}_h \boldsymbol{\sigma}_\rho} v\|_{L_p(T)} + \rho^2 \|D^2_{\boldsymbol{\sigma}_\rho \boldsymbol{\sigma}_\rho} v\|_{L_p(T)}, \tag{2.19}$$

$$|v - I_T v|_{W_p^1(T)} \lesssim h |D_{\boldsymbol{\sigma}_h} v|_{W_p^1(T)} + \rho |D_{\boldsymbol{\sigma}_\rho} v|_{W_p^1(T)}, \tag{2.20}$$

*with the constant in the inequalities depending only on $\gamma(T)$, and where $\boldsymbol{\sigma}_h$ and $\boldsymbol{\sigma}_\rho$ are defined on page 18.*

Note that in the above estimations, the constants do not depend on $p$ since $p = q$ in the settings of [1].

As we have already mentioned in the introduction (see (1.6)), the right hand side of (2.19) is a better estimation compared to (2.18) for $m = 0$, however the latter can remain advantageous since the dependency on $\gamma(T)$ of the constants in Lemma 2.2.2 is unknown. The estimation in (2.20) is widely used in order to estimate the derivatives of the error on a strongly anisotropic triangle, that is, when $\frac{h}{\rho}$ is large. It can be re-written as follow,

$$|v - I_T v|_{W_p^1(T)} \lesssim h \|D^2_{\boldsymbol{\sigma}_h \boldsymbol{\sigma}_h} v\|_{L_p(T)} + h \|D^2_{\boldsymbol{\sigma}_\rho \boldsymbol{\sigma}_h} v\|_{L_p(T)} + \rho \|D_{\boldsymbol{\sigma}_\rho \boldsymbol{\sigma}_\rho} v\|_{L_p(T)}. \tag{2.21}$$

For $p = 2$, an improvement of (2.18) is provided in Lemma 2.2.3 below where the estimation is independent of the maximum interior angle of the triangle.

Let $M$ denote the matrix occurring in the map $\psi$ in (2.8) for which $T = \psi(\hat{T})$. We can write the matrix $M$ into its polar form

$$M = BU, \quad B = [\mathbf{r}_1 \ \mathbf{r}_2] \begin{bmatrix} \nu_1 & 0 \\ 0 & \nu_2 \end{bmatrix} [\mathbf{r}_1 \ \mathbf{r}_2]^t, \tag{2.22}$$

where $U$ is orthonormal and $B$ a symmetric positive definite matrix whose eigenvalues and eigenvectors are respectively $\nu_1 \geq \nu_2 \geq 0$ and $\mathbf{r}_1, \mathbf{r}_2$.

The result below is proved in [23] (see also [10] for a similar method) for $p = 2$.

**Lemma 2.2.3** ([23, Proposition 2.1]). *There is a constant $C$ such that, for all functions $v \in W_2^2(T)$, the estimations*

$$\begin{aligned}
\|v - I_T v\|_{L_2(T)} \leq & C\Big(\nu_1^4\big(|\boldsymbol{r}_1^t|N_v|\boldsymbol{r}_1|\big)^2 + \nu_2^4\big(|\boldsymbol{r}_2^t|N_v|\boldsymbol{r}_2|\big)^2 + 2\nu_1^2\nu_2^2\big(|\boldsymbol{r}_1^t|N_v|\boldsymbol{r}_2|\big)^2\Big)^{\frac{1}{2}}, \\
|v - I_T v|_{W_2^1(T)} \leq & C\nu_2\Big(\frac{\nu_1^4}{\nu_2^4}\big(|\boldsymbol{r}_1^t|N_v|\boldsymbol{r}_1|\big)^2 + \big(|\boldsymbol{r}_2^t|N_v|\boldsymbol{r}_2|\big)^2 + 2\frac{\nu_1^2}{\nu_2^2}\big(|\boldsymbol{r}_1^t|N_v|\boldsymbol{r}_2|\big)^2\Big)^{\frac{1}{2}},
\end{aligned}$$

*hold, where $\nu_i$ and $|\boldsymbol{r}_i| = \Big|[r_{i1}\ r_{i2}]^t\Big| = \Big[|r_{i1}|\ |r_{i2}|\Big]^t$, $i = 1, 2$, are defined as in* (2.22) *and $N_v = \begin{bmatrix} \|D_{xx}^2 v\|_{L_2(T)} & \|D_{xy}^2 v\|_{L_2(T)} \\ \|D_{xy}^2 v\|_{L_2(T)} & \|D_{yy}^2 v\|_{L_2(T)} \end{bmatrix}$.*

To compare the results of Lemma 2.2.2 and Lemma 2.2.3, we provide estimations involving the eigenvalues and eigenvectors of the matrix $B$ occurring in (2.22). The estimations in Lemma 2.2.3, however, are not invariant with respect to the coordinate system which means that we obtain different formulas when replacing $\mathbf{r}_i$ by $R_\vartheta \mathbf{r}_i$, $i = 1, 2$, with $R_\vartheta$ being a counterclockwise rotation of angle $\vartheta$.

The eigenvalues $\nu_1$ and $\nu_2$ can be expressed by using the diameter $h$ and smallest height $\rho$ of $T$. Consider the singular value decomposition $B_T = U_0 S V_0^t$ of the matrix $B_T$ occurring in (2.9), where $U_0$ and $V_0$ are orthonormal matrices and $S$ a diagonal matrix. The column vectors of $U_0$ are the normalized eigenvectors of the matrix $B_T B_T^t$ given by

$$B_T B_T^t = \begin{bmatrix} h & \delta - h/2 \\ 0 & \rho \end{bmatrix} \begin{bmatrix} h & 0 \\ \delta - h/2 & \rho \end{bmatrix} = \begin{bmatrix} h^2 + (\delta - h/2)^2 & \rho(\delta - h/2) \\ \rho(\delta - h/2) & \rho^2 \end{bmatrix},$$

whose characteristic polynomial $P(\Lambda)$ in the variable $\Lambda$ satisfies

$$\begin{aligned}
P(\Lambda) =& (h^2 + (\delta - h/2)^2 - \Lambda)(\rho^2 - \Lambda) - \rho^2(\delta - h/2)^2, \\
=& \Lambda^2 - (h^2 + \rho^2 + (\delta - h/2)^2)\Lambda + \rho^2 h^2. \qquad (2.23)
\end{aligned}$$

With $\Delta = (h^2+\rho^2+\mu_h^2)^2 - 4\rho^2h^2 \leq 13h^4$ where $\mu_h = \delta - h/2$, the eigenvalues $\Lambda_1, \Lambda_2$ are the solutions to $P(\Lambda) = 0$, with

$$\Lambda_1 = \frac{h^2+\rho^2+\mu_h^2+\sqrt{\Delta}}{2}, \tag{2.24}$$

$$\Lambda_2 = \frac{h^2+\rho^2+\mu_h^2-\sqrt{\Delta}}{2}. \tag{2.25}$$

Then, the singular values of $B$ are exactly given by $\nu_1 = \sqrt{\Lambda_1}$, $\nu_2 = \sqrt{\Lambda_2}$, whereas the corresponding eigenvectors $\mathbf{v}_1 = [x_1\ y_1]^t$ and $\mathbf{v}_2 = [x_2\ y_2]^t$ satisfy

$$\begin{bmatrix} h^2+\mu_h^2-\Lambda_i & \rho\mu_h \\ \rho\mu_h & \rho^2-\Lambda_i \end{bmatrix} [x_i\ y_i]^t = 0, \quad i = 1, 2,$$

or, equivalently, $(h^2+\mu_h^2-\Lambda_i)x_i + \rho\mu_h y_i = 0$. After choosing $x_1 = x_2 = 1$,

$$y_1 = -\frac{h^2+\mu_h^2-\Lambda_1}{\rho\mu_h}x_1 = \frac{\rho^2-h^2-\mu_h^2+\sqrt{\Delta}}{2\rho\mu_h},$$
$$y_2 = -\frac{h^2+\mu_h^2-\Lambda_2}{\rho\mu_h}x_2 = \frac{\rho^2-h^2-\mu_h^2-\sqrt{\Delta}}{2\rho\mu_h}.$$

In the case where the triangle $T$ is isosceles such that its largest interior angle is formed by its two edges of the same length, we have $\delta = h/2$ which leads to simpler expressions for the values of $\nu_1, \nu_2$ and their corresponding eigenvectors $\mathbf{v}_1, \mathbf{v}_2$. Indeed, we have $\nu_1 = h$, $\nu_2 = \rho$ but also $B_T = \begin{bmatrix} h & 0 \\ 0 & \rho \end{bmatrix}$. Thus from (2.9), we have

$$M = \left(R_\theta B_T R_\theta^t\right) R_\theta = \left(R_\theta \begin{bmatrix} h & 0 \\ 0 & \rho \end{bmatrix} R_\theta^t\right) R_\theta, \tag{2.26}$$

for some angle $\theta$. Since the polar decomposition is unique, by identifying the matrices in (2.26) with those in (2.22), we find that $[\mathbf{r}_1\ \mathbf{r}_2] = R_\theta = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. We thus obtain the following expressions:

$$\begin{aligned} |\mathbf{r}_1^t|N_v|\mathbf{r}_1| &= [|\cos\theta|\ |\sin\theta|]N_v \begin{bmatrix} |\cos\theta| \\ |\sin\theta| \end{bmatrix} \\ &= |\cos\theta|^2 \|D_{xx}^2 v\|_{L_2(T)} + 2|\sin\theta\cos\theta| \|D_{xy}^2 v\|_{L_2(T)} + |\sin\theta|^2 \|D_{yy}^2 v\|_{L_2(T)}; \end{aligned}$$

$$\begin{aligned}|\mathbf{r}_2^t|N_v|\mathbf{r}_2| &= [\sin\theta|\ |\cos\theta|]N_v\begin{bmatrix}|\sin\theta|\\|\cos\theta|\end{bmatrix}\\&=|\sin\theta|^2\|D^2_{xx}v\|_{L_2(T)}+2|\sin\theta\cos\theta|\|D^2_{xy}v\|_{L_2(T)}+|\cos\theta|^2\|D^2_{yy}v\|_{L_2(T)};\end{aligned}$$

$$\begin{aligned}|\mathbf{r}_1^t|N_v|\mathbf{r}_2| =&[|\cos\theta|\ |\sin\theta|]N_v\begin{bmatrix}|\sin\theta|\\|\cos\theta|\end{bmatrix}\\=&|\sin\theta\cos\theta|\Big(\|D^2_{xx}v\|_{L_2(T)}+\|D^2_{yy}v\|_{L_2(T)}\Big)+\|D^2_{xy}v\|_{L_2(T)}.\end{aligned}$$

Substituting the expressions of $|\mathbf{r}_1^t|N_v|\mathbf{r}_1|$, $|\mathbf{r}_2^t|N_v|\mathbf{r}_2|$ and $|\mathbf{r}_1^t|N_v|\mathbf{r}_2|$ into the first estimation in Lemma 2.2.3, we find that

$$\begin{aligned}\|v-I_Tv\|_{L_2(T)}\le& C\Big(\nu_1^2|\mathbf{r}_1^t|N_v|\mathbf{r}_1|+\nu_2^2|\mathbf{r}_2^t|N_v|\mathbf{r}_2|+2\nu_1\nu_2|\mathbf{r}_1^t|N_v|\mathbf{r}_2|\Big)\\=&C\Big(\Big(h^2|\cos\theta|^2+\rho^2|\sin\theta|^2+2h\rho|\sin\theta\cos\theta|\Big)\|D^2_{xx}v\|_{L_2(T)}\\&+2\Big((h^2+\rho^2)|\sin\theta\cos\theta|+\rho h\Big)\|D^2_{xy}v\|_{L_2(T)}\\&+\Big(h^2|\sin\theta|^2+\rho^2|\cos\theta|^2+2h\rho|\sin\theta\cos\theta|\Big)\|D^2_{yy}v\|_{L_2(T)}\Big)\\=&C\Big(\Big(h|\cos\theta|+\rho|\sin\theta|\Big)^2\|D^2_{xx}v\|_{L_2(T)}\\&+2\Big(h|\cos\theta|+\rho|\sin\theta|\Big)\Big(h|\sin\theta|+\rho|\cos\theta|\Big)\|D^2_{xy}v\|_{L_2(T)}\\&+\Big(h|\sin\theta|+\rho|\cos\theta|\Big)^2\|D^2_{yy}v\|_{L_2(T)}\Big).\end{aligned}\tag{2.27}$$

In a similar way, the second estimation in Lemma 2.2.3 reads

$$\begin{aligned}|v-I_Tv|_{W_2^1(T)}\le& C\nu_2\Big(\frac{\nu_1^2}{\nu_2^2}|\mathbf{r}_1^t|N_v|\mathbf{r}_1|+|\mathbf{r}_2^t|N_v|\mathbf{r}_2|+2\frac{\nu_1}{\nu_2}|\mathbf{r}_1^t|N_v|\mathbf{r}_2|\Big)\\=&C\rho\Big(\Big(\frac{h^2}{\rho^2}|\cos\theta|^2+|\sin\theta|^2+2\frac{h}{\rho}|\sin\theta\cos\theta|\Big)\|D^2_{xx}v\|_{L_2(T)}\\&+2\Big(\Big(\frac{h^2}{\rho^2}+1\Big)+\frac{h}{\rho}\Big)\|D^2_{xy}v\|_{L_2(T)}\\&+\Big(\frac{h^2}{\rho^2}|\sin\theta|^2+|\cos\theta|^2+2\frac{h}{\rho}|\sin\theta\cos\theta|\Big)\|D^2_{yy}v\|_{L_2(T)}\Big)\end{aligned}$$

$$
\begin{aligned}
=&\frac{C}{\rho}\Big(\Big(h|\cos\theta|+\rho|\sin\theta|\Big)^2\|D^2_{xx}v\|_{L_2(T)}\\
&+2\Big(h|\cos\theta|+\rho|\sin\theta|\Big)\Big(h|\sin\theta|+\rho|\cos\theta|\Big)\|D^2_{xy}v\|_{L_2(T)}\\
&+\Big(h|\sin\theta|+\rho|\cos\theta|\Big)^2\|D^2_{yy}v\|_{L_2(T)}\Big).
\end{aligned} \tag{2.28}
$$

We then have the following result for isosceles triangles.

**Corollary 2.2.4.** *For any isosceles triangles $T$ whose maximum interior angle $\gamma(T)$ is the angle between the two edges of equal length, the estimations in Lemma 2.2.3 read, there exists a constant $C$ such that, for all functions $v \in W^2_p(T)$,*

$$
\begin{aligned}
&\|v-I_Tv\|_{L_2(T)}\\
&\leq C\Big(\beta^2_{1,T}\|(\cos\theta)^2D^2_{\boldsymbol{\sigma}_h\boldsymbol{\sigma}_h}v+2\cos\theta\sin\theta D^2_{\boldsymbol{\sigma}_h\boldsymbol{\sigma}_\rho}v+(\sin\theta)^2D^2_{\boldsymbol{\sigma}_\rho\boldsymbol{\sigma}_\rho}v\|_{L_2(T)}\\
&+2\beta_{1,T}\beta_{2,T}\|\sin\theta\cos\theta(D^2_{\boldsymbol{\sigma}_h\boldsymbol{\sigma}_h}v-D^2_{\boldsymbol{\sigma}_\rho\boldsymbol{\sigma}_\rho}v)+(\cos 2\theta)D^2_{\boldsymbol{\sigma}_h\boldsymbol{\sigma}_\rho}v\|_{L_2(T)}\\
&+\beta^2_{2,T}\|(\sin\theta)^2D^2_{\boldsymbol{\sigma}_h\boldsymbol{\sigma}_h}v-2\cos\theta\sin\theta D^2_{\boldsymbol{\sigma}_h\boldsymbol{\sigma}_\rho}v+(\cos\theta)^2D^2_{\boldsymbol{\sigma}_\rho\boldsymbol{\sigma}_\rho}v\|_{L_2(T)}\Big);
\end{aligned} \tag{2.29}
$$

$$
\begin{aligned}
&|v-I_Tv|_{W^1_2(T)}\\
&\leq \frac{C}{\rho}\Big(\beta^2_{1,T}\|(\cos\theta)^2D^2_{\boldsymbol{\sigma}_h\boldsymbol{\sigma}_h}v+2\cos\theta\sin\theta D^2_{\boldsymbol{\sigma}_h\boldsymbol{\sigma}_\rho}v+(\sin\theta)^2D^2_{\boldsymbol{\sigma}_\rho\boldsymbol{\sigma}_\rho}v\|_{L_2(T)}\\
&+2\beta_{1,T}\beta_{2,T}\|\sin\theta\cos\theta(D^2_{\boldsymbol{\sigma}_h\boldsymbol{\sigma}_h}v-D^2_{\boldsymbol{\sigma}_\rho\boldsymbol{\sigma}_\rho}v)+(\cos 2\theta)D^2_{\boldsymbol{\sigma}_h\boldsymbol{\sigma}_\rho}v\|_{L_2(T)}\\
&+\beta^2_{2,T}\|(\sin\theta)^2D^2_{\boldsymbol{\sigma}_h\boldsymbol{\sigma}_h}v-2\cos\theta\sin\theta D^2_{\boldsymbol{\sigma}_h\boldsymbol{\sigma}_\rho}v+(\cos\theta)^2D^2_{\boldsymbol{\sigma}_\rho\boldsymbol{\sigma}_\rho}v\|_{L_2(T)}\Big),
\end{aligned} \tag{2.30}
$$

*where $\beta_{1,T} := h|\cos\theta|+\rho|\sin\theta|$, $\beta_{2,T} := h|\sin\theta|+\rho|\cos\theta|$, with $\theta$ being the angle of rotation of $R_\theta$ in (2.26).*

*Proof.* Since $D_xv = \cos\theta D_{\boldsymbol{\sigma}_h}v + \sin\theta D_{\boldsymbol{\sigma}_\rho}v$ and $D_yv = -\sin\theta D_{\boldsymbol{\sigma}_h}v + \cos\theta D_{\boldsymbol{\sigma}_\rho}v$, with $\theta$ being the angle of rotation of $R_\theta$ in (2.26), we easily prove that

$$
\begin{aligned}
D^2_{xx}v &= (\cos\theta)^2D^2_{\boldsymbol{\sigma}_h\boldsymbol{\sigma}_h}v+2\cos\theta\sin\theta D^2_{\boldsymbol{\sigma}_h\boldsymbol{\sigma}_\rho}v+(\sin\theta)^2D^2_{\boldsymbol{\sigma}_\rho\boldsymbol{\sigma}_\rho}v,\\
D^2_{xy}v &= \sin\theta\cos\theta(D^2_{\boldsymbol{\sigma}_h\boldsymbol{\sigma}_h}v-D^2_{\boldsymbol{\sigma}_\rho\boldsymbol{\sigma}_\rho}v)+(\cos 2\theta)D^2_{\boldsymbol{\sigma}_h\boldsymbol{\sigma}_\rho}v,\\
D^2_{yy}v &= (\sin\theta)^2D^2_{\boldsymbol{\sigma}_h\boldsymbol{\sigma}_h}v-2\cos\theta\sin\theta D^2_{\boldsymbol{\sigma}_h\boldsymbol{\sigma}_\rho}v+(\cos\theta)^2D^2_{\boldsymbol{\sigma}_\rho\boldsymbol{\sigma}_\rho}v.
\end{aligned}
$$

The result is proved by combining the above with (2.27) and (2.28). □

As discussed in [23, Section 2.1], while using the estimation (2.30), the ratio

$$\frac{\text{actual error}}{\text{estimated error}},$$

can be bounded in such a way that the maximum angle condition on $T$ is not required. As a result, (2.30) may suffer from an over-estimation as compared to (2.20). In particular, given an isosceles triangle $T$ whose maximum interior angle is formed by its two edges of the same length, and such that its longest edge is chosen to be on the $x$-axis, that is $\theta = 0$, we have $\beta_{1,T} = h$ and $\beta_{2,T} = \rho$. Hence from (2.30),

$$|v - I_T v|_{W_2^1(T)} \leq \frac{C}{\rho}\Big(h^2 \|D^2_{\boldsymbol{\sigma}_h \boldsymbol{\sigma}_h} v\|_{L_2(T)} + 2h\rho \|D^2_{\boldsymbol{\sigma}_h \boldsymbol{\sigma}_\rho} v\|_{L_2(T)} + \rho^2 \|D^2_{\boldsymbol{\sigma}_\rho \boldsymbol{\sigma}_\rho} v\|_{L_2(T)}\Big), \tag{2.31}$$

which, unless using an additional assumption such as

$$\|D^2_{\boldsymbol{\sigma}_h \boldsymbol{\sigma}_h} v\|_{L_2(T)} = \|D^2_{xx} v\|_{L_2(T)} = 0,$$

is an over-estimation as compared to (2.20), for $p = 2$, due to the additional factor $\frac{h}{\rho}$ of $\|D^2_{\boldsymbol{hh}} v\|_{L_2(T)}$. Note that for isosceles triangles $T$ whose interior angle $\gamma(T)$ is defined by its two edges which have the same length, the tangent of $\gamma(T)$ is comparable to the aspect ratio $\frac{h}{\rho}$, more precisely, $\tan(\gamma(T)/2) = \frac{h}{2\rho}$. Since Lemma 2.2.2 does not specify how the constants depend on $\gamma(T)$, the estimation (2.31) provides a stronger result on isosceles triangles of the type described in the Corollary above, as compared to (2.20).

## 2.3 Approximation of quadratic polynomials

We present here local error bounds when approximating a quadratic polynomial $\pi$ on a given triangle $T$. The interpolation operator $I_T$ being linear, the function error $\pi - I_T\pi$ is identically null if $\pi$ is a linear polynomial. We therefore assume that $\pi$ does not possess any linear part, that is, $\pi \in \mathbb{H}_2$ is a homogeneous quadratic polynomial of the form

$$\pi(x,y) = ax^2 + 2bxy + cy^2, \quad \text{where} \quad a, b, c \in \mathbb{R},$$

which can also be written in a matrix formulation as follows, in the form of (2.1),

$$\pi(x,y) = [x\ y]U_\pi \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} U_\pi^t [x\ y]^t, \tag{2.32}$$

with $\lambda_1, \lambda_2$ and $U_\pi$ being as described in (2.1) and (2.2).

### 2.3.1 Estimations on the reference triangle

Let $\hat{T}$ be the reference triangle on the left of Figure 2.2, and consider the invertible affine map $\psi$ defined in (2.8) for which $T = \psi(\hat{T})$. The results in Lemma 2.3.1 below are useful for the proof of Proposition 2.3.5.

**Lemma 2.3.1.** *Consider the quadratic polynomial $\hat{\pi} = \pi \circ \psi$. Given a unit vector $\hat{\boldsymbol{\sigma}}$, consider $\boldsymbol{\sigma} = \frac{M\hat{\boldsymbol{\sigma}}}{|M\hat{\boldsymbol{\sigma}}|}$ where $M$ is the invertible matrix occurring in (2.8). Then,*

$$\|\hat{\pi}\|_{L_p(\hat{T})} = |\det M|^{-\frac{1}{p}} \|\pi\|_{L_p(T)}, \tag{2.33}$$

$$\|D_{\hat{\boldsymbol{\sigma}}}\hat{\pi}\|_{L_p(\hat{T})} = |M\hat{\boldsymbol{\sigma}}||\det M|^{-\frac{1}{p}} \|D_{\boldsymbol{\sigma}}\pi\|_{L_p(T)}, \tag{2.34}$$

$$\|D_{\hat{\boldsymbol{\sigma}}}(\hat{\pi} - I_{\hat{T}}\hat{\pi})\|_{L_p(\hat{T})} = |M\hat{\boldsymbol{\sigma}}||\det M|^{-\frac{1}{p}} \|D_{\boldsymbol{\sigma}}(\pi - I_T\pi)\|_{L_p(T)}. \tag{2.35}$$

*Proof.* By using the differentiation rule for composite functions, for $x, y \in \mathbb{R}$,

$$D_x(\pi \circ \psi)(x,y) = D_x\psi_1(x,y) \cdot (D_x\pi) \circ \psi(x,y) + D_x\psi_2(x,y) \cdot (D_y\pi) \circ \psi(x,y),$$
$$D_y(\pi \circ \psi)(x,y) = D_y\psi_1(x,y) \cdot (D_x\pi) \circ \psi(x,y) + D_y\psi_2(x,y) \cdot (D_y\pi) \circ \psi(x,y),$$

where $\psi(x,y) = [\psi_1(x,y)\ \psi_2(x,y)]^t$.

Given a unit vector $\boldsymbol{\xi}_0 = [x_0\ y_0]^t$, we show that the derivatives $D_{\boldsymbol{\xi}_0}(\pi \circ \psi)$ and

$\Big(D_{M\boldsymbol{\xi}_0}\pi\Big)\circ\psi$ are related as shown in (2.37) below. First observe that for $x, y \in \mathbb{R}$,

$$\begin{aligned}D_{\boldsymbol{\xi}_0}(\pi\circ\psi)(x,y) &= x_0 D_x(\pi\circ\psi)(x,y) + y_0 D_y(\pi\circ\psi)(x,y)\\ &= x_0\Big(D_x\psi_1(x,y)\cdot(D_x\pi)\circ\psi(x,y) + D_x\psi_2(x,y)\cdot(D_y\pi)\circ\psi(x,y)\Big)\end{aligned}$$

$$\begin{aligned}&+ y_0\Big(D_y\psi_1(x,y)\cdot(D_x\pi)\circ\psi(x,y) + D_y\psi_2(x,y)\cdot(D_y\pi)\circ\psi(x,y)\Big)\\ =&\Big(x_0 D_x\psi_1(x,y) + y_0 D_y\psi_1(x,y)\Big)(D_x\pi)\circ\psi(x,y)\\ &+ \Big(x_0 D_x\psi_2(x,y) + y_0 D_y\psi_2(x,y)\Big)(D_y\pi)\circ\psi(x,y). \qquad (2.36)\end{aligned}$$

Next, with $M$ being the invertible matrix in (2.8) which we write $M = \begin{bmatrix} a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix}$ for some $a_{ij}$, $i, j = 1, 2$ satisfying $a_{11}a_{22} - a_{21}a_{12} \neq 0$, clearly

$$M[x\ y]^t = [a_{11}x + a_{12}y \quad a_{21}x + a_{22}y]^t =: [M_1(x,y)\ M_2(x,y)]^t.$$

With the expressions of $M_1$ and $M_2$ satisfying

$$\begin{aligned}[x_0 D_x\psi_1(x,y) + y_0 D_y\psi_1(x,y) \quad & x_0 D_x\psi_2(x,y) + y_0 D_y\psi_2(x,y)]^t\\ &= [a_{11}x_0 + a_{12}y_0 \quad a_{21}x_0 + a_{22}y_0]^t\\ &= [M_1(x_0,y_0)\ M_2(x_0,y_0)]^t,\end{aligned}$$

we deduce from (2.36) that

$$\begin{aligned}D_{\boldsymbol{\xi}_0}(\pi\circ\psi)(x,y) &= \Big(M_1(x_0,y_0)(D_x\pi) + M_2(x_0,y_0)(D_y\pi)\Big)\circ\psi(x,y)\\ &= (D_{M\boldsymbol{\xi}_0}\pi)\circ\psi(x,y). \qquad (2.37)\end{aligned}$$

We shall now prove (2.33). Writing $z = \psi(\hat{z})$ where $\hat{z} \in \hat{T}$, we have

$$\hat{\pi}(\hat{z}) = \pi\circ\psi(\hat{z}) = \pi(z), \qquad (2.38)$$

and since $\mathrm{d}z = |\det M| \mathrm{d}\hat{z}$, by a change of variables, clearly

$$\|\hat{\pi}\|_{L_p(\hat{T})} = \left( \int_{\hat{T}} |\hat{\pi}(\hat{z})|^p \mathrm{d}\hat{z} \right)^{\frac{1}{p}} = |\det M|^{-\frac{1}{p}} \left( \int_{T} |\pi(z)|^p \mathrm{d}z \right)^{\frac{1}{p}} = |\det M|^{-\frac{1}{p}} \|\pi\|_{L_p(T)}.$$

In order to prove (2.34), we first combine (2.37) and (2.38) to obtain the equalities

$$D_{\hat{\boldsymbol{\sigma}}}\hat{\pi}(\hat{z}) = D_{\hat{\boldsymbol{\sigma}}}(\pi \circ \psi)(\hat{z}) = (D_{M\hat{\boldsymbol{\sigma}}}\pi) \circ \psi(\hat{z}) = D_{M\hat{\boldsymbol{\sigma}}}\pi(z) = |M\hat{\boldsymbol{\sigma}}| D_{\boldsymbol{\sigma}}\pi(z).$$

By using again a simple change of variables, we find that

$$\begin{aligned} \|D_{\hat{\boldsymbol{\sigma}}}\hat{\pi}\|_{L_p(\hat{T})} &= \left( \int_{\hat{T}} |D_{\hat{\boldsymbol{\sigma}}}\hat{\pi}(\hat{z})|^p \mathrm{d}\hat{z} \right)^{\frac{1}{p}} = |M\hat{\boldsymbol{\sigma}}| |\det M|^{-\frac{1}{p}} \left( \int_{T} |D_{\boldsymbol{\sigma}}\pi(z)|^p \mathrm{d}z \right)^{\frac{1}{p}} \\ &= |M\hat{\boldsymbol{\sigma}}| |\det M|^{-\frac{1}{p}} \|D_{\boldsymbol{\sigma}}\pi\|_{L_p(T)}. \end{aligned}$$

The proof of (2.35) goes as follows: By using (2.12), with $z = \psi(\hat{z})$ where $\hat{z} \in \hat{T}$, it holds that

$$I_{\hat{T}}\hat{\pi}(\hat{z}) = I_{\hat{T}}(\pi \circ \psi)(\hat{z}) = (I_{\psi(\hat{T})}\pi) \circ \psi(\hat{z}) = I_T \pi(z)$$

and therefore $(\hat{\pi} - I_{\hat{T}}\hat{\pi})(\hat{z}) = (\pi - I_T\pi)(z)$. We then deduce from (2.34) that

$$\|D_{\boldsymbol{\sigma}}(\pi - I_T\pi)\|_{L_p(T)} = \frac{|\det M|^{\frac{1}{p}}}{|M\hat{\boldsymbol{\sigma}}|} \|D_{\hat{\boldsymbol{\sigma}}}(\hat{\pi} - I_{\hat{T}}\hat{\pi})\|_{L_p(\hat{T})},$$

and the result is proved. □

### 2.3.2 Measure of non-degeneracy

The alignment and shape of a triangle are important characteristics in order to obtain a sharper error estimation. This is commonly known especially in the case where the target function presents singularities and fast changing behavior at some points. In this section, given a homogeneous quadratic polynomial $\pi$, we shall characterize a triangle $T$ by its *measure of non-degeneracy* $\rho_\pi(T)$ [31] (see also [14]) defined as follows,

$$\rho_\pi(T) := \frac{\max\{|\pi(\mathbf{e})| : \mathbf{e} \text{ edge of } T\}}{|T|\sqrt{|\det \pi|}}, \quad \pi \in \mathbb{H}_2. \tag{2.39}$$

Observe that in the numerator of (2.39), the difference of vertices is used rather than the vertices. Also, if $\mathbf{e}_1, \mathbf{e}_2$ and $\mathbf{e}_3$ are counterclockwise oriented edge-vectors of $T$ such that $\mathbf{e}_1 + \mathbf{e}_2 + \mathbf{e}_3 = 0$, then their images under $\pi$ do not necessarily form a triangle since $\pi(\mathbf{e}_1) + \pi(\mathbf{e}_2) + \pi(\mathbf{e}_3)$ might not be zero. It is easy to see that $\rho_\pi$ is invariant under translation and scaling of $T$ by a constant. It can also be generalized into a wider space of functions.

In the example below, we explain the relation between the measure $\rho_\pi(T)$ and the aspect ratio of the triangle.

**Example 2.3.1.** Let $\pi_0(x, y) := x^2 + y^2$ and $\pi_1(x, y) = x^2 - y^2$. Since $\det \pi_0 = 1$ and $\pi_0(\mathbf{e}) = |\mathbf{e}|^2$ for any edge $\mathbf{e}$ of $T$, we obtain

$$\rho_{\pi_0}(T) = \frac{\operatorname{diam}(T)^2}{|T|} = \frac{h^2}{\frac{h\rho}{2}} = \frac{2h}{\rho},$$

which is twice the aspect ratio of $T$. As a result, the minimum value of $\rho_{\pi_0}(T)$ is attained when $T$ is an equilateral triangle for which $\rho = \frac{\sqrt{3}}{2}h$ and $\rho_{\pi_0}(T) = \frac{4}{\sqrt{3}} = \frac{4\sqrt{3}}{3}$. For any triangle $T$, it is discussed in [31] that the minimum value for $\rho_{\pi_1}(T)$ is attained when $T$ is a half of a square whose edges are parallel to the $x$- and $y$- axes of the Cartesian coordinate system. Observe that, since $|\pi_1(\mathbf{e})| \leq |\pi_0(\mathbf{e})|$ for any vector $\mathbf{e}$, the inequality

$$\max\{\rho_{\pi_0}(T), \rho_{\pi_1}(T)\} \leq \frac{2h}{\rho}, \tag{2.40}$$

holds for any triangle $T$.

Other characterization of triangles can be found in [31]. In particular, the measure of non-degeneracy helps in characterizing the triangles that may present big local errors, they are then bisected into two.

Given a linear map $\phi$ and a homogeneous quadratic polynomial $\pi \in \mathbb{H}_2$, the polynomial $\pi \circ \phi$ belongs to $\mathbb{H}_2$ and $\rho_{\pi\circ\phi}$ is well defined. This is not the case if $\phi$ was an affine map because in general $\pi \circ \psi \notin \mathbb{H}_2$. However, if $\psi$ is an affine map of the form $\psi = \phi + \phi_{\mathbf{t}}$ where $\phi_{\mathbf{t}}$ is a translation with vector $\mathbf{t} \in \mathbb{R}^2$, then for any triangle $T$, by invariance of $\rho_\pi$ by translation, we have

$$\rho_\pi(\psi(T)) = \rho_\pi(\phi(T)).$$

The right hand side is equal to the measure $\rho_{\pi\circ\phi}(T)$, as proved in the result below which shows the commutativity of $\rho_\pi$ and the linear map $\phi$.

**Lemma 2.3.2.** *For any invertible linear map $\phi$,*

$$\rho_{\pi\circ\phi}(T) = \rho_\pi(\phi(T)). \tag{2.41}$$

*Proof.* Assume that $\phi$ is a linear map of the form $\phi(x, y) = (\alpha x+\beta y, \gamma x+\delta y)$ with $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ satisfying $\alpha\delta - \beta\gamma \neq 0$. The homogeneous quadratic polynomial $\pi \circ \phi$ is expressed as follows:

$$\begin{aligned}\pi \circ \phi(x, y) &= a(\alpha x + \beta y)^2 + 2b(\alpha x + \beta y)(\gamma x + \delta y) + c(\gamma x + \delta y)^2\\ &= Ax^2 + 2Bxy + Cy^2,\end{aligned}$$

where $A = a\alpha^2+2b\alpha\gamma+c\gamma^2$, $B = a\alpha\beta+b(\alpha\delta+\beta\gamma)+c\delta\gamma$ and $C = a\beta^2+2b\beta\delta+c\delta^2$. Since $Q_\pi = \begin{bmatrix} a & b\\ b & c\end{bmatrix}$, we immediately verify that $\begin{bmatrix}\alpha & \beta\\ \gamma & \delta\end{bmatrix}^t Q_\pi \begin{bmatrix}\alpha & \beta\\ \gamma & \delta\end{bmatrix} = \begin{bmatrix} A & B\\ B & C\end{bmatrix} = Q_{\pi\circ\phi}$ from which it follows that

$$\det(\pi \circ \phi) = AC - B^2 = (\det \pi)(\det \phi)^2.$$

The following equalities then hold,

$$\begin{aligned}\rho_{\pi\circ\phi}(T) &= \frac{\max\{|\pi \circ \phi(\mathbf{e})| : \mathbf{e} \text{ edge of } T\}}{|T|\sqrt{|\det \pi \circ \phi|}} = \frac{\max\{|\pi(\phi(\mathbf{e}))| : \mathbf{e} \text{ edge of } T\}}{|T|\sqrt{|\det \pi||\det \phi|^2}}\\ &= \frac{\max\{|\pi(\mathbf{e})| : \mathbf{e} \text{ edge of } \phi(T)\}}{|\phi(T)|\sqrt{|\det \pi|}} = \rho_\pi(\phi(T)),\end{aligned}$$

thereby proving the result. □

Given a non-degenerate homogeneous quadratic polynomial $\pi$, consider the invertible linear map $\phi_\pi : [x\ y]^t \mapsto [X\ Y]^t$ defined by

$$\phi_\pi(x,y) := |\det \pi|^{\frac{1}{4}} U_\pi \begin{bmatrix} |\lambda_1|^{-\frac{1}{2}} & 0 \\ 0 & |\lambda_2|^{-\frac{1}{2}} \end{bmatrix} [x\ y]^t = [X\ Y]^t, \tag{2.42}$$

where $\lambda_1, \lambda_2$ are the eigenvalues of $Q_\pi$ as defined in (2.2), with $|\lambda_1| \leq |\lambda_2|$. Define also the quadratic polynomial $\varpi_\pi$ by

$$\varpi_\pi(x,y) := x^2 + \varepsilon_\pi y^2, \quad x, y \in \mathbb{R}, \tag{2.43}$$

where $\varepsilon_\pi := \operatorname{sign}(\det \pi)$. By using (2.32), we can express the homogeneous quadratic polynomial $\pi \circ \phi_\pi$ in terms of the determinant $\det \pi$ and $\varpi_\pi(x,y)$, for all $(x,y) \in \mathbb{R}^2$,

$$\begin{aligned}
\pi\circ\phi_\pi(x,y) &= [X\ Y] U_\pi \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} U_\pi^t [X\ Y]^t \\
&= |\det \pi|^{\frac{1}{2}} [x\ y] \begin{bmatrix} |\lambda_1|^{-\frac{1}{2}} & 0 \\ 0 & |\lambda_2|^{-\frac{1}{2}} \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} |\lambda_1|^{-\frac{1}{2}} & 0 \\ 0 & |\lambda_2|^{-\frac{1}{2}} \end{bmatrix} [x\ y]^t \\
&= |\det \pi|^{\frac{1}{2}} [x\ y] \begin{bmatrix} \operatorname{sign}(\lambda_1) & 0 \\ 0 & \operatorname{sign}(\lambda_2) \end{bmatrix} [x\ y]^t \\
&= \operatorname{sign}(\lambda_1) |\det \pi|^{\frac{1}{2}} \varpi_\pi(x,y).
\end{aligned} \tag{2.44}$$

The result below is obtained by applying Lemma 2.3.2 to $\phi_\pi$.

**Corollary 2.3.3.** *Let $\phi_\pi$ and $\varpi_\pi$ be defined as in (2.42) and in (2.43). Then*

$$\rho_\pi(T) = \rho_{\varpi_\pi}(\phi_\pi^{-1}(T)) = \rho_{\varpi_\pi}\left( \begin{bmatrix} |\lambda_1/\lambda_2|^{\frac{1}{4}} & 0 \\ 0 & |\lambda_2/\lambda_1|^{\frac{1}{4}} \end{bmatrix} U_\pi^t(T) \right). \tag{2.45}$$

*Proof.* The result is easily proved by using (2.41), (2.44) and the easily proved equalities below, for any $\alpha \neq 0$,

$$\rho_{\alpha\pi}(T) = \frac{\max\{|\alpha||\pi(\mathbf{e})| : \mathbf{e} \text{ edge of } T\}}{|T|\sqrt{|\alpha|^2| \det \pi|}} = \rho_\pi(T).$$

Indeed, writing $\bar{T} = \phi_\pi^{-1}(T)$ and choosing $\alpha = \operatorname{sign}(\lambda_1)|\det \pi|^{\frac{1}{2}}$, from (2.41) and (2.44), the following equalities hold

$$\rho_\pi(\phi_\pi(\bar{T})) = \rho_{\pi\circ\phi_\pi}(\bar{T}) = \rho_{\alpha\varpi_\pi}(\bar{T}) = \rho_{\varpi_\pi}(\bar{T}) = \rho_{\varpi_\pi}(\phi_\pi^{-1}(T)).$$

The result is obtained by noticing that $\phi_\pi^{-1} = \begin{bmatrix} |\lambda_1/\lambda_2|^{\frac{1}{4}} & 0 \\ 0 & |\lambda_2/\lambda_1|^{\frac{1}{4}} \end{bmatrix} U_\pi^t.$ □

With $\varpi_\pi$ defined in (2.43), we deduce from (2.40) that, for any triangle $\bar{T}$,

$$\rho_{\varpi_\pi}(\bar{T}) \leq \frac{2h_{\bar{T}}}{\rho_{\bar{T}}}. \tag{2.46}$$

Therefore, $\rho_{\varpi_\pi}(\bar{T})$ is bounded as long as $\bar{T}$ is isotropic. Combining this with (2.45), with $T = \phi_\pi(\bar{T})$, the measure of non-degeneracy

$$\rho_\pi(T) = \rho_{\varpi_\pi}(\bar{T}) \leq \frac{2h_{\bar{T}}}{\rho_{\bar{T}}}, \tag{2.47}$$

is bounded whenever $\phi_\pi^{-1}(T) = \bar{T}$ is isotropic. In [31], it is shown that the minimum value of $\rho_\pi(T)$ is attained on triangles which are isotropic with respect to the induced metric $|\cdot|_\pi$ defined by $|v|_\pi := \sqrt{|\pi(v)|}$.

### 2.3.3 Error bounds for directional derivatives

Consider again a homogeneous quadratic polynomial $\pi$. Our goal is to derive alternative estimations to the estimations in [1] and [13] in the case where the measure of non-degeneracy $\rho_\pi(T)$ is bounded. We estimate the derivatives of the approximation error obtained by interpolating homogeneous quadratic polynomials. With $T$ being a fixed triangle, the interpolant $I_T$ reproduces linear polynomials and thus we can consider only homogeneous quadratic polynomials.

Let us first recall the Markov-type inequality below.

**Lemma 2.3.4** ([28, Theorem 2.32]). *Let $\pi$ be a polynomial of total degree $n \geq 0$. Given any non-negative integers $\nu, \eta$ such that $\nu + \eta \leq n$, it holds that, for any triangle $T$,*

$$\|D_x^\nu D_y^\eta \pi\|_{L_p(T)} \leq \frac{C_n}{\rho^{\nu+\eta}} \|\pi\|_{L_p(T)}, \tag{2.48}$$

*with $C_n$ being a constant depending only on $n$.*

Note that on the reference triangle $\hat{T}$ shown in Figure 2.2, the factor $\hat{C} = \frac{C_n}{\hat{\rho}^{\nu+\eta}}$ in (2.48) remains an absolute constant for quadratic polynomials since the smallest height $\hat{\rho}$ of $\hat{T}$ is constant.

We now prove the following result.

**Proposition 2.3.5.** *There exist absolute constants $C_1$ and $C_2$ such that, given any triangle $T$ and any non-degenerate polynomial $\pi \in \mathbb{H}_2$, the estimations*

(i) $\|\pi - I_T\pi\|_{L_p(T)} \leq C_1 \rho_\pi(T) |T|^{\frac{1}{p}} h\rho\sqrt{|\det \pi|}$;

(ii) $\|D_{\boldsymbol{\sigma}_h}(\pi - I_T\pi)\|_{L_p(T)} \leq C_2 \rho_\pi(T) |T|^{\frac{1}{p}} \rho\sqrt{|\det \pi|}$;

(iii) $\|D_{\boldsymbol{\sigma}_\rho}(\pi - I_T\pi)\|_{L_p(T)} \leq C_2 \rho_\pi(T) |T|^{\frac{1}{p}} h\sqrt{|\det \pi|}$,

*hold, where $\boldsymbol{\sigma}_h$ and $\boldsymbol{\sigma}_\rho$ denote the unit vectors defined on page 18, and where $\rho_\pi$ is the measure of non-degeneracy defined in (2.39).*

*Proof.* Let $\hat{T}$ be the reference triangle in Figure 2.2. Given a triangle $T$, consider the function $\psi$ in (2.8) for which $T = \psi(\hat{T})$, and let $\hat{\pi} := \pi \circ \psi$. Note that, except when using the correspondence (2.38), the variable of $\hat{\pi}$ remains $z = (x, y)$.

(i): By using (2.48), there exists an absolute constant $\hat{C}_1$ such that

$$|\hat{\pi}|_{W_p^2(\hat{T})} = \left(\|D_{xx}^2\hat{\pi}\|_{L_p(\hat{T})}^p + \|D_{xy}^2\hat{\pi}\|_{L_p(\hat{T})}^p + \|D_{yy}^2\hat{\pi}\|_{L_p(\hat{T})}^p\right)^{\frac{1}{p}} \leq \hat{C}_1 \|\hat{\pi}\|_{L_p(\hat{T})},$$

which, together with (2.18) for $m = 0$, yields

$$\|\hat{\pi} - I_{\hat{T}}\hat{\pi}\|_{L_p(\hat{T})} \leq \hat{C}_2 |\hat{\pi}|_{W_p^2(\hat{T})} \leq (\hat{C}_1\hat{C}_2)\|\hat{\pi}\|_{L_p(\hat{T})}, \tag{2.49}$$

where $\hat{C}_2$ is an absolute constant.

We now estimate the error $\|\pi - I_T\pi\|$ as follows. By (2.7), there exists an absolute constant $c_2$ such that

$$\begin{aligned} \|\pi\|_{L_p(T)} &\leq c_2\|\pi\|_T = c_2|T|^{\frac{1}{p}} \max\{|\pi(\mathbf{e})| : \mathbf{e} \text{ edge of } T\} \\ &= c_2\rho_\pi(T)|T|^{1+\frac{1}{p}}\sqrt{|\det \pi|}, \end{aligned} \tag{2.50}$$

where $\rho_\pi$ is defined as in (2.39). Combining the above inequality with (2.33) and (2.49), we can estimate $\|\pi - I_T\pi\|_{L_p(T)}$ as follows:

$$\begin{aligned} \|\pi - I_T\pi\|_{L_p(T)} &= |\det M|^{\frac{1}{p}}\|\hat{\pi} - I_{\hat{T}}\hat{\pi}\|_{L_p(\hat{T})} \leq (\hat{C}_1\hat{C}_2)|\det M|^{\frac{1}{p}}\|\hat{\pi}\|_{L_p(\hat{T})} \\ &= (\hat{C}_1\hat{C}_2)\|\pi\|_{L_p(T)} \leq (c_2\hat{C}_1\hat{C}_2)\rho_\pi(T)|T|^{1+\frac{1}{p}}\sqrt{|\det \pi|}, \end{aligned}$$

which, since $|T| = \frac{h\rho}{2}$, proves (i) with $C_1 = \frac{c_2\hat{C}_1\hat{C}_2}{2}$.

(ii) and (iii): Given a unit vector $\hat{\boldsymbol{\sigma}}$, combining (2.48) with (2.49) and (2.33) yields

$$\begin{aligned} \|D_{\hat{\boldsymbol{\sigma}}}(\hat{\pi} - I_{\hat{T}}\hat{\pi})\|_{L_p(\hat{T})} &\leq \hat{C}_1\|\hat{\pi} - I_{\hat{T}}\hat{\pi}\|_{L_p(\hat{T})} \\ &\leq (\hat{C}_1^2\hat{C}_2)\|\hat{\pi}\|_{L_p(\hat{T})} \\ &= (\hat{C}_1^2\hat{C}_2)|\det M|^{-\frac{1}{p}}\|\pi\|_{L_p(T)}. \end{aligned}$$

Combining this with (2.35) and (2.50), with $\boldsymbol{\sigma} = \frac{M\hat{\boldsymbol{\sigma}}}{|M\hat{\boldsymbol{\sigma}}|}$, we find that

$$\begin{aligned} \|D_{\boldsymbol{\sigma}}(\pi - I_T\pi)\|_{L_p(T)} &\leq (\hat{C}_1^2\hat{C}_2)|M\hat{\boldsymbol{\sigma}}|^{-1}\|\pi\|_{L_p(T)} \\ &\leq (c_2\hat{C}_1^2\hat{C}_2)|M\hat{\boldsymbol{\sigma}}|^{-1}\rho_\pi(T)|T|^{1+\frac{1}{p}}\sqrt{|\det \pi|}. \end{aligned} \tag{2.51}$$

By virtue of (2.10), we have

$$\boldsymbol{\sigma}_h = \frac{M\hat{\boldsymbol{\sigma}}}{|M\hat{\boldsymbol{\sigma}}|} \quad \text{and} \quad \boldsymbol{\sigma}_\rho = \frac{M\hat{\boldsymbol{\tau}}}{|M\hat{\boldsymbol{\tau}}|}, \tag{2.52}$$

where $\hat{\boldsymbol{\sigma}} = [1\ 0]^t$ and $\hat{\boldsymbol{\tau}} = \left[\frac{1}{2} - \frac{\delta}{h}\ 1\right]^t$. Note that $\hat{\boldsymbol{\tau}}$ is not a unit vector, however the second equality in (2.52) still holds if $\hat{\boldsymbol{\tau}}$ is replaced by $\alpha\hat{\boldsymbol{\tau}}$, $\alpha \neq 0$. Hence, by using the unit vector $\bar{\boldsymbol{\tau}} = \frac{\hat{\boldsymbol{\tau}}}{\|\hat{\boldsymbol{\tau}}\|}$ we have $\boldsymbol{\sigma}_\rho = \frac{M\bar{\boldsymbol{\tau}}}{|M\bar{\boldsymbol{\tau}}|}$. Noting that $|M\hat{\boldsymbol{\sigma}}| = h$ and $|M\bar{\boldsymbol{\tau}}| = \|\hat{\boldsymbol{\tau}}\|^{-1}\rho$, applying (2.51) for both $\boldsymbol{\sigma} = \boldsymbol{\sigma}_h$ and $\boldsymbol{\sigma} = \boldsymbol{\sigma}_\rho$, respectively, we obtain

$$\|D_{\boldsymbol{\sigma}_h}(\pi - I_T\pi)\|_{L_p(T)} \leq \frac{2C_2}{h}\rho_\pi(T)|T|^{1+\frac{1}{p}}\sqrt{|\det \pi|} = C_2\rho_\pi(T)|T|^{\frac{1}{p}}\rho\sqrt{|\det \pi|},$$
$$\|D_{\boldsymbol{\sigma}_\rho}(\pi - I_T\pi)\|_{L_p(T)} \leq \frac{2C_2\|\hat{\boldsymbol{\tau}}\|}{\rho}\rho_\pi(T)|T|^{1+\frac{1}{p}}\sqrt{|\det \pi|} = C_2\rho_\pi(T)|T|^{\frac{1}{p}}h\sqrt{|\det \pi|},$$

thereby proving (ii) and (iii) with $C_2 = C\hat{C}_1^2\hat{C}_2\|\hat{\boldsymbol{\tau}}\|$. □

In the above result, clearly $\rho_\pi(T)$ needs to be bounded. The proof uses the reference triangle $\hat{T}$ and the affine map $\psi$ for which $T = \psi(\hat{T})$ holds, as well as simple change of variables in the computations of the $L_p$-norms.

The presence of the measure of non-degeneracy $\rho_\pi(T)$ in the estimations (i)-(iii) is a novel feature. The example below demonstrates the effectiveness of these error bounds on a triangle which has an interior angle approaching $\pi$ but has its measure of non-degeneracy bounded.

**Example 2.3.2.** Consider the quadratic polynomial $\pi(x, y) = ax^2 + by^2$ such that $|a| \leq |b|$ (the case where $|a| \geq |b|$ is obtained by the change of axes). Let $T$ be an isosceles triangle with vertices at $(0, 0), (h_1, 0)$ and $(\frac{1}{2}h_1, h_2)$ where $h_1, h_2 > 0$. In the case where $h_2$ is small and $h_1$ large, the triangle $T$ is strongly anisotropic with a big interior angle. The eigenvalues of $Q_\pi$ are exactly $\lambda_1 = a$ and $\lambda_2 = b$, with associated eigenvectors $[1\ 0]^t$ and $[0\ 1]^t$. Observe that $T$ is aligned in the direction orthogonal to the eigenvector corresponding to the largest eigenvalue.

The determinant of $\pi$ and the linear map $\phi_\pi$ as defined in (2.42) are given by

$$\det \pi = ab \quad \text{and} \quad \phi_\pi = \begin{bmatrix} \left|\frac{b}{a}\right|^{\frac{1}{4}} & 0 \\ 0 & \left|\frac{a}{b}\right|^{\frac{1}{4}} \end{bmatrix}.$$

The image $\bar{T} = \phi_\pi^{-1}(T)$ has the coordinates $(0,0)$, $\left(\left|\frac{a}{b}\right|^{\frac{1}{4}} h_1, 0\right)$ and $\left(\left|\frac{a}{b}\right|^{\frac{1}{4}} \frac{h_1}{2}, \left|\frac{b}{a}\right|^{\frac{1}{4}} h_2\right)$. By choosing $h_1 = \left|\frac{b}{a}\right|^{\frac{1}{4}}$ and $h_2 = \left|\frac{a}{b}\right|^{\frac{1}{4}}$, we deduce from (2.47) and (2.39) that

$$\rho_\pi(T) = \rho_{\varpi_\pi}(\bar{T}) = \frac{1+\frac{1}{4}}{\frac{1}{2}} = \frac{5}{2},$$

by virtue of the fact that $|T| = \frac{h_1 h_2}{2} = \frac{1}{2}$.

Noting that $h_T \sim h_1$ and $\rho_T \sim h_2$, we deduce from Proposition 2.3.5 that

(i) $\|\pi - I_T\pi\|_{L_p(T)} \le \frac{5C_1}{2}|T|^{\frac{1}{p}} h_T \rho_T |ab|^{\frac{1}{2}} \le C_1'|T|^{\frac{1}{p}}|ab|^{\frac{1}{2}}$;

(ii) $\|D_{\boldsymbol{\sigma}_{h_T}}(\pi - I_T\pi)\|_{L_p(T)} \le \frac{5C_2}{2}|T|^{\frac{1}{p}} \rho_T |ab|^{\frac{1}{2}} \le C_2'|T|^{\frac{1}{p}}|a|^{\frac{3}{4}}|b|^{\frac{1}{4}}$;

(iii) $\|D_{\boldsymbol{\sigma}_{\rho_T}}(\pi - I_T\pi)\|_{L_p(T)} \le \frac{5C_2}{2}|T|^{\frac{1}{p}} h_T |ab|^{\frac{1}{2}} \le C_2'|T|^{\frac{1}{p}}|a|^{\frac{1}{4}}|b|^{\frac{3}{4}}$,

where $C_1'$ and $C_2'$ are absolute constants.

The above example shows that, although a triangle may present an interior angle near to $\pi$, we can still obtain sharp estimations if the triangle is aligned (by this we mean the longest edge) in the direction orthogonal to the eigenvector corresponding to the largest eigenvalue.

## 2.4 Optimal triangles

In this section, we study triangles which minimize the $L_p$-norm of the error $\pi - I_T\pi$ for a given homogeneous quadratic polynomials $\pi \in \mathbb{H}_2$. Given two triangles $T, T'$ having the same area, clearly one of the errors $e_T(\pi), e_{T'}(\pi)$ may be smaller than the other. Assume that both triangles contain the origin. We show that translating the triangles will not change the error (also proved in [2, 29]). Given a constant $\mathbf{a} \in \mathbb{R}^2$, $\pi(\mathbf{x}+\mathbf{a}) = \pi(\mathbf{x}) + \ell(\mathbf{x}+\mathbf{a})$ for some linear polynomial $\ell$. Recalling that the linear interpolation is exact for linear polynomials, for any triangle $T$, we have

$$\begin{aligned}
\|\pi - I_{T+\mathbf{a}}\pi\|_{L_p(T+\mathbf{a})} &= \|\pi(\cdot+\mathbf{a}) - (I_{T+\mathbf{a}}\pi)(\cdot+\mathbf{a})\|_{L_p(T)} \\
&= \|\pi(\cdot+\mathbf{a}) - I_T(\pi(\cdot+\mathbf{a}))\|_{L_p(T)} \\
&= \|\pi - I_T\pi\|_{L_p(T)},
\end{aligned}$$

by virtue of (2.12).

To obtain a deeper study of the approximation error on a triangle, following [29], we define the *shape function* $K_p$ as follows: for any $\pi \in \mathbb{H}_2$,

$$K_p(\pi) := \inf_{\substack{|T|=1 \\ 0 \in T}} e_T(\pi), \tag{2.53}$$

with $e_T$ defined in (2.14). The functional $K_p$ can be extended to a wider space of functions, however, we shall only restrict to homogeneous quadratic polynomials where its expression is known. Indeed, it is shown in [29] that $K_p(\pi) = 0$ if $\pi$ is degenerate, whereas, if $\pi$ is non-degenerate,

$$K_p(\pi) = \sqrt{|\det \pi|}K_p(\varpi_\pi), \tag{2.54}$$

where $\varpi_\pi$ is defined in (2.43).

Following [29], a *tempered* shape function $K_{p,L}$, with $L > 0$, is defined in order to bound the diameter of the triangle for which (2.53) is attained: for any homogeneous quadratic polynomial $\pi$, define

$$K_{p,L}(\pi) = \inf_{T \in \mathbb{T}_L} e_T(\pi), \tag{2.55}$$

where $\mathbb{T}_L$ is the set of triangles of unit area, containing the origin and of diameter less than or equal to $L$, that is,

$$\mathbb{T}_L := \{T : |T| = 1,\, 0 \in T,\, h_T \leq L\}.$$

We refer to [29] for details of the properties between $K_p$ and $K_{p,L}$.

Given a homogeneous quadratic polynomial $\pi$, we define the set $\Delta_p(\pi)$ of all triangles $T$ of unit area, having a vertex at the origin and such that the infimum in (2.53) is attained. A triangle $T \in \Delta_p(\pi)$ is called an *optimal* triangle for $\pi$. In Sections 2.4.1, 2.4.2, 2.4.3 we discuss the properties of the triangles in $\Delta_p(\pi)$. In particular, the case where $\det \pi < 0$ with $p < \infty$ remains an open question.

The following result follows from Lemma 2.1.2.

**Lemma 2.4.1.** *There is a constant $C$ such that, for any $\pi, \pi' \in \mathbb{H}_2$ such that $\Delta_p(\pi) \neq \emptyset$ and $\Delta_p(\pi') \neq \emptyset$, we have*

$$|K_p(\pi) - K_p(\pi')| \leq C h_{T_0}^2 \|\pi - \pi'\|, \tag{2.56}$$

*with some $T_0 \in \Delta_p(\pi) \cup \Delta_p(\pi')$, and where the norm $\|\cdot\|$ is defined in (2.3).*

*Proof.* From (2.53) and (2.15), there is a constant $C$ such that $K_p(\pi) \leq e_T(\pi') + C h_T^2 \|\pi - \pi'\|$ holds for any triangle $T$ of unit area and having a vertex at the origin. In particular, for a $T_0 \in \Delta_p(\pi')$,

$$K_p(\pi) \leq K_p(\pi') + C h_{T_0}^2 \|\pi - \pi'\|.$$

In a similar way, we easily prove that

$$K_p(\pi') \leq K_p(\pi) + C h_{T_0}^2 \|\pi - \pi'\|,$$

holds whenever $T_0 \in \Delta_p(\pi)$. The result directly follows. □

Before we start discussing optimal triangles for the canonic quadratic polynomials $\pi_0(x, y) = x^2 + y^2$ and $\pi_1(x, y) = x^2 - y^2$ defined in Example 2.3.1, we first present the following result about *invariance property.*

**Lemma 2.4.2** ([32, Theorem 7]). *Given a homogeneous quadratic polynomial $\pi \in \mathbb{H}_2$ and an affine map $\psi(\mathbf{x}) = A\mathbf{x} + \mathbf{b}$ for which $A$ is $Q_\pi$-orthogonal, i.e $A^t Q_\pi A = Q_\pi$, the equality*

$$\|\pi - I_T\pi\|_{L^p(T)} = \|\pi - I_{T'}\pi\|_{L^p(T')}, \tag{2.57}$$

*holds for any triangle $T$, with $T' = \psi(T)$.*

### 2.4.1 For $x^2 + y^2$

The result below shows that equilateral triangles of unit area are optimal for $\pi_0(x, y) = x^2 + y^2$.

**Lemma 2.4.3** ([3]). *A triangle $T$ belongs to $\Delta_p(\pi_0)$ if and only if it is an equilateral triangle of unit area.*

*Sketch of proof.* We provide here a sketch of proof for triangles of arbitrary area.

For $p < \infty$, the value of $K_p(\pi_0)$ is equal to

$$C_p^+ = \inf_T \frac{\|\pi_0 - I_T\pi_0\|_{L_p(T)}}{|T|^{1+\frac{1}{p}}}, \tag{2.58}$$

which, by virtue of (2.11), is easily proved to be invariant under scaling, translation and rotation of the triangle $T$. Given a non-equilateral triangle $T$, the proof consists in finding a triangle $\bar{T}$ for which

$$\frac{\|\pi_0 - I_T\pi_0\|_{L_p(T)}}{|T|^{1+\frac{1}{p}}} > \frac{\|\pi_0 - I_{\bar{T}}\pi_0\|_{L_p(\bar{T})}}{|\bar{T}|^{1+\frac{1}{p}}}. \tag{2.59}$$

By virtue of the invariance by scaling, translation and rotation, we can assume $T$ to have the vertices $(-1, 0), (a, b), (1, 0)$ where $a \in \mathbb{R}$ and $b > 0$. Then direct computations show that the isosceles triangle $\bar{T}$ with vertices $(-1, 0), (0, b), (1, 0)$ satisfies (2.59). This means that for $a = 0$ the triangles $T$ and $\bar{T}$ are identical. If

$b \neq \sqrt{3}$ in which case $T$ is an equilateral triangle, we use a rotation and a similar argument as before to find a triangle $\bar{T}'$ such that

$$\frac{\|\pi_0 - I_T\pi_0\|_{L_p(T)}}{|T|^{1+\frac{1}{p}}} > \frac{\|\pi_0 - I_{\bar{T}'}\pi_0\|_{L_p(\bar{T}')}}{|\bar{T}'|^{1+\frac{1}{p}}}.$$

This means that in order for $T$ to yield $K_p(\pi_0)$, it must be symmetric with respect to each of its bisectors. In other words, equilateral.

A similar argument applies for $p = \infty$, with $K_\infty(\pi_0)$ equal to

$$C_\infty^+ = \inf_T \frac{\|\pi_0 - I_T\pi_0\|_{L_\infty(T)}}{|T|} = \frac{4}{3\sqrt{3}}. \tag{2.60}$$

Given an arbitrary triangle $T$, the set of points $z$ for which $(\pi_0 - I_T\pi_0)(z) = 0$ is the circle $\mathcal{C}_T$ with center $\mathbf{m}$ and radius $R$ defined as follows: Let $\mathcal{C}_T^0$ be the circumscribed circle and denote by $\mathbf{m}^0$ its center. If $\mathbf{m}^0$ is contained in $T$, then $\mathcal{C}_T = \mathcal{C}_T^0$. Otherwise $\mathcal{C}_T$ is the circle centered at the mid-point of the longest edge of $T$ and with radius $\frac{h_T}{2}$. Observe that the center of $\mathcal{C}_T$ is always contained in $T$.

It is proved that (see [32]) on a segment $[z_1, z_2]$, the maximum error between $\pi_0$ and a linear polynomial $\ell$ such that $\pi_0(z_1) - \ell(z_1) = \pi_0(z_2) - \ell(z_2) = 0$ is attained at the mid-point of $[z_1, z_2]$. Using this to any segment contained in $T$, clearly $\|\pi_0 - I_T\pi_0\|_{L_\infty(T)}$ is attained at the center $\mathbf{m}$ of $\mathcal{C}_T$. It follows that, for any triangle $\bar{T}$ contained in $\mathcal{C}_T$ and such that $\mathbf{m} \in \bar{T}$,

$$\|\pi_0 - I_T\pi_0\|_{L_\infty(T)} = \|\pi_0 - I_{\bar{T}}\pi_0\|_{L_\infty(\bar{T})},$$

that is, the errors on $\bar{T}$ and $T$ are the same. However, the area of both triangle are not the same, and if $|\bar{T}| \geq |T|$, then

$$\frac{\|\pi_0 - I_T\pi_0\|_{L_\infty(T)}}{|T|} \geq \frac{\|\pi_0 - I_{\bar{T}}\pi_0\|_{L_\infty(\bar{T})}}{|\bar{T}|}.$$

Only equilateral triangles with vertices on $\mathcal{C}_T$ have the largest area in $\mathcal{C}_T$, and thus (2.60) can only be attained on equilateral triangles. □

The general multivariate case ($d \geq 2$) is given in [32] with a very similar proof. Observe that the placement of the vertices of an optimal triangle for $\pi_0$ is not important as long as it is an equilateral triangle of unit area. By this reason, we are able to present Algorithm 3.1 in Section 3.2.2 which provides an isosceles optimal triangle for a homogeneous quadratic polynomial $\pi$ whose determinant is positive.

### 2.4.2 For $x^2 - y^2$

Obtaining an optimal triangle for $\pi_1(x,y) = x^2 - y^2$ is not as straightforward as for $\pi_0$. Up to date, optimal triangles for $\pi_1$ are known only for $p = \infty$.

#### When $p = \infty$

For the study of optimal triangles for a $\pi \in \mathbb{H}_2$ whose determinant is negative, only the case $p = \infty$ is discussed in [2], the generalization for $p < \infty$ is still an open question. We present below the case where the quadratic polynomial is $\pi_1$.

Given two numbers $a, b > 0$ such that $\frac{3\sqrt{5}-5}{4}ab = 1$, the triangle $T_{a,b}$ is defined by the vertices given by

$$(0,0),\quad \frac{1}{2}\Big(c_0 a + b, c_0 a - b\Big),\quad \frac{1}{2}\Big(a + c_0 b, a - c_0 b\Big), \tag{2.61}$$

where $c_0 = \frac{3-\sqrt{5}}{2}$. In Figure 2.4, we plot a few triangles $T_{a,b}$ for different values of $a$ and $b$. We define as well the set $\mathcal{I}_0 = \left\{ \begin{bmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{bmatrix} \right\}$ whose elements are matrices which represent symmetry with respect to the $x$- and or $y$-axis.

**Lemma 2.4.4** ([3, Lemma 9])**.** *A triangle $T$ belongs to $\Delta_\infty(\pi_1)$ if and only if $T = \phi_{M_0} T_{a,b}$, $a, b > 0$, where $\phi_{M_0}$ is the linear map associated with a matrix $M_0 \in \mathcal{I}_0$.*

*Sketch of proof.* We provide here a sketch of proof for triangles of arbitrary area, and for any quadratic polynomial $\pi$ with negative determinant. Let $\lambda_{\min} <$

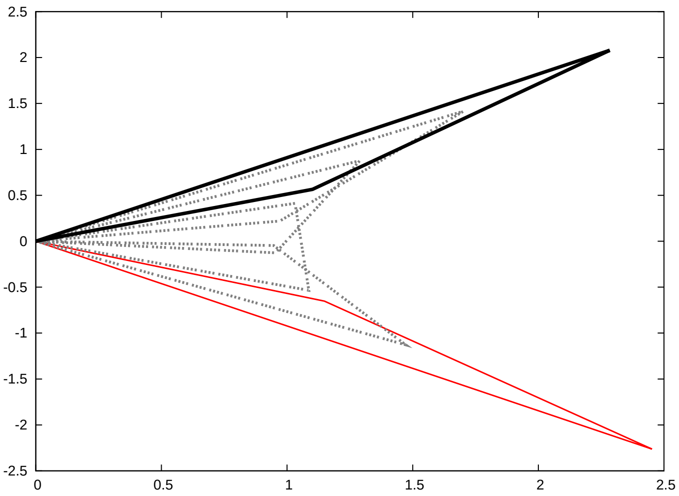


Figure 2.4: Family of optimal triangles for $\pi_1$, with $p = \infty$, whose vertices are given by (2.61).

$0 < \lambda_{\max}$ denote the eigenvalues of $\pi$. Denoting by $(\xi_1, \xi_2)$ a unit eigenvector associated with $\lambda_{\max}$, the unit vector $(\xi_2, -\xi_1)$ is an eigenvector corresponding to $\lambda_{\min}$.

Considering the linear change of variables,

$$\begin{aligned} F_1(x,y) &:= \left[\xi_1 x + \xi_2 y \;\; \xi_2 x - \xi_1 y\right]^t, \\ G_2(x,y) &:= \left[\sqrt{\lambda_{\max}}x - \sqrt{|\lambda_{\min}|}y \quad \sqrt{\lambda_{\max}}x + \sqrt{|\lambda_{\min}|}y\right]^t, \end{aligned}$$

we obtain the function $\tilde{\pi}$ given by

$$\tilde{\pi}(x,y) := \pi \circ F_1^{-1} \circ G_2^{-1}(x,y) = xy. \tag{2.62}$$

The interpolation error for $\pi$ does not change under the above transformation. However, $|G_2(T)| = 2\sqrt{|\lambda_{\max}\lambda_{\min}|}|T|$. Moreover, it has been proved in [3, Lemma 6] that $\|\tilde{\pi} - I_T\tilde{\pi}\|_{L_\infty(T)}$ is attained on the boundary of $T$.

Given a triangle $T$, consider the rectangle of minimal area containing $T$, whose sides are parallel to the coordinate axes, with side lengths $a$ and $b$. By translation, we can assume that $T$ has a vertex at the origin. The complement of $T$ with respect to the rectangle is formed by three right triangles $T_1, T_2, T_3$ with respective area $S_1, S_2, S_3$. Observe that each side of $T$ is the longest side of one of the three right triangles. Also, the interpolation error on each side of $T$ is equal to $\frac{1}{2}|T_i|$ where $T_i$ has the same edge as $T$. Hence, the problem of minimizing $\frac{\|\pi - I_T\pi\|_{L_\infty(T)}}{|T|}$ is equivalent to finding triangles which solve the minimization problem

$$\frac{\max\{S_1, S_2, S_3\}}{2(ab - (S_1 + S_2 + S_3))} \to \min . \tag{2.63}$$

Since the triangle is contained in the rectangle, and two of its vertices on the sides of the rectangle, we can assume that $T$ has vertices $(0,0), (x,b), (a,y)$. Then (2.63) reads

$$\frac{1}{2} \max \left\{ \frac{bx}{ab - xy}, \frac{ay}{ab - xy}, \frac{(a-x)(b-y)}{ab - xy} \right\} \to \min . \tag{2.64}$$

It is proved that for any triangle $T$ with vertices $(0,0), (x,b), (a,y)$ where $x \in (0,a)$ and $y \in (0,b)$, we have

$$\frac{1}{2} \max \left\{ \frac{bx}{ab - xy}, \frac{ay}{ab - xy}, \frac{(a-x)(b-y)}{ab - xy} \right\} \geq \frac{1}{2\sqrt{5}}.$$

Also, with $c_0 = \frac{3-\sqrt{5}}{2}$, studying each of the cases

1. $x \geq c_0 a, y \geq c_0 b$;
2. $x \leq c_0 a, y \leq c_0 b$;
3. $x \geq c_0 a, y \leq c_0 b$ or $x \leq c_0 a, y \geq c_0 b$,

shows that (2.64) is attained on triangles $\bar{T}$ with vertices given by

$$(0,0), \quad \Big(c_0 a, b\Big), \quad \Big(a, c_0 b\Big), \tag{2.65}$$

where $c_0 = \frac{3-\sqrt{5}}{2}$. Finally, the triangles which minimize $\|\pi - I_T\pi\|_{L_\infty(T)}$ are given by $T = (F_1^{-1} \circ G_2^{-1})(\bar{T})$, or symmetric to it with respect to any coordinate axis, and only for such triangles.

Now, we take $\pi(x,y) = \pi_1(x,y) = x^2 - y^2$. We have $\lambda_{\max} = 1$, $\lambda_{\min} = -1$

and the eigenvalues of $Q_{\pi_1}$, with $(\xi_1, \xi_2) = (1, 0)$ being the eigenvector associated with $\lambda_{\max}$. Clearly $F_1^{-1} \circ G_2^{-1}(x, y) = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}[x\ y]^t$. The images of the vertices $(0, 0)$, $(c_0 a, b)$ and $(a, c_0 b)$ under $F_1^{-1} \circ G_2^{-1}$ are the vertices (2.61) of $T_{a,b}$ whose area is given by

$$|T_{a,b}| = \frac{1}{2}\left|\det \begin{bmatrix} (c_0 a + b)/2 & (a + c_0 b)/2 & 0 \\ (c_0 a - b)/2 & (a - c_0 b)/2 & 0 \\ 1 & 1 & 1 \end{bmatrix}\right| = \frac{3\sqrt{5} - 5}{4} ab.$$

Since an optimal triangle must have a unit area, we require that $\frac{3\sqrt{5}-5}{4} ab = 1$. □

Observe from Figure 2.4 that the vertices of an optimal triangle for $\pi_1$ are either given by (2.61) or given by the symmetry of it with respect to the $x$- and or $y$-axis of these vertices. Since $\|\pi_1 - I_T \pi\|_{L_p(T)} \leq |T|^{\frac{1}{p}} \|\pi - I_T \pi\|_{L_\infty(T)}$ for any triangle $T$, we immediately deduce from the definition of $K_p$ in (2.53) that

$$K_p(\pi_1) \leq K_\infty(\pi_1), \quad p \in [1, \infty]. \tag{2.66}$$

More discussions (similar to above) on the optimal triangles for $\pi_1(x, y) = x^2 - y^2$ can be found in [32], however by using the representation $\pi_1'(x, y) = xy$. We also refer to [32] for the homogeneous quadratic polynomial $\pi_2(x, y) = x^2$ and for more than two variables homogeneous quadratic polynomials. However, the case $\pi_1(x, y) = x^2 - y^2$ with $p < \infty$ still remains to be investigated.

### When $p < \infty$

In this section, we shall discuss a method for obtaining an optimal triangle for $\pi_1$ by using Lemma 2.4.2. In particular, we study the existence of an optimal triangle which is isotropic which still remains an open question.

Let us begin by investigating the families of optimal triangles for the quadratic polynomials $\pi_0$ and $\pi_1$ defined in Example 2.3.1. Given a matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, we have that

$$\begin{aligned} A^tQ_{\pi_0}A = Q_{\pi_0} &\iff A^tA = I \\ &\iff a^2 + c^2 = 1,\ ab + cd = 0,\ b^2 + d^2 = 1. \end{aligned} \tag{2.67}$$

The set of matrices which satisfy (2.67) is the set of all 2 × 2 orthogonal matrices which is given by

$$\mathcal{A}_0 := \left\{ \begin{bmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{bmatrix}, \begin{bmatrix} 0 & \pm 1 \\ \pm 1 & 0 \end{bmatrix} \right\} \bigcup \left\{ \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} : \theta \in [0, 2\pi] \right\}. \tag{2.68}$$

The set $\mathcal{A}_0$ in (2.68) is the union of two sets: On the left is the set of matrices that represent symmetry with respect to the $x$-, $y$-axis and symmetry with respect to the lines $x+y = 0$ and $x-y = 0$, whereas on the right is the set of counterclockwise rotation matrices.

In a similar way, we have that

$$\begin{aligned} A^tQ_{\pi_1}A = Q_{\pi_1} &\iff \begin{bmatrix} a & c \\ b & d \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \\ &\iff a^2 - c^2 = 1,\ ab - cd = 0,\ b^2 - d^2 = -1. \end{aligned} \tag{2.69}$$

The set of matrices which satisfy (2.69) is given by

$$\mathcal{A}_1 := \left\{ \begin{bmatrix} \pm 1 & 0 \\ 0 & \pm 1 \end{bmatrix}, \begin{bmatrix} 0 & \pm 1 \\ \pm 1 & 0 \end{bmatrix} \right\} \bigcup \left\{ \begin{bmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{bmatrix} : \theta \in \mathbb{R} \right\}. \tag{2.70}$$

The set $\mathcal{A}_1$ in (2.70) is the union of two sets: On the left is the set of matrices that represent symmetry with respect to the $x$- and $y$-axis and symmetry with respect to the lines $x + y = 0$ and $x - y = 0$ (clearly (2.57) also holds for $A^tQ_\pi A = -Q_\pi$), and on the right is the set of matrices with hyperbolic entries.

Suppose that $T_1$ is an optimal triangle for $\pi_1$, and let $\mathbf{e}_i = [x_i\, y_i]^t$, $\mathbf{e}_j = [x_j\, y_j]^t$ be two of its edge-vectors such that its maximum interior angle satisfies $\gamma(T_1) = \widehat{\mathbf{e}_i\mathbf{e}_j}$. Considering a matrix $S_\theta = \left[\begin{smallmatrix} \sinh\theta & \cosh\theta \\ \cosh\theta & \sinh\theta \end{smallmatrix}\right]$, with $\theta \in \mathbb{R}$, we shall study the positions of the images $S_\theta\mathbf{e}_i$ and $S_\theta\mathbf{e}_j$ by introducing the functions $X(\theta) =$

$a\cosh\theta + b\sinh\theta = (a + t_\theta b)\cosh\theta$ and $Y(\theta) = a\sinh\theta + b\cosh\theta = (at_\theta + b)\cosh\theta$ where $a, b \in \mathbb{R}$ and $t_\theta = \tanh\theta \in (-1, 1)$.

Obviously, the curve defined by $X(\theta)$ and $Y(\theta)$ is hyperbolic by virtue of the easily provable equality $X(\theta)^2 - Y(\theta)^2 = a^2 - b^2$. Since

$$\frac{Y(\theta)}{X(\theta)} = \frac{at_\theta + b}{a + t_\theta b},$$

we observe that for $\theta \to \infty$ we have $t_\theta \to 1$ and $\lim_{\theta\to\infty} \frac{Y(\theta)}{X(\theta)} = 1$. In a similar way, we show that $\lim_{\theta\to-\infty} \frac{Y(\theta)}{X(\theta)} = -1$ and thus the lines $y = x$ and $y = -x$ are asymptotic to the hyperbolic curve. In Figure 2.5 we illustrate the hyperbolic curves obtained from the cases 1-2 which we discuss below.

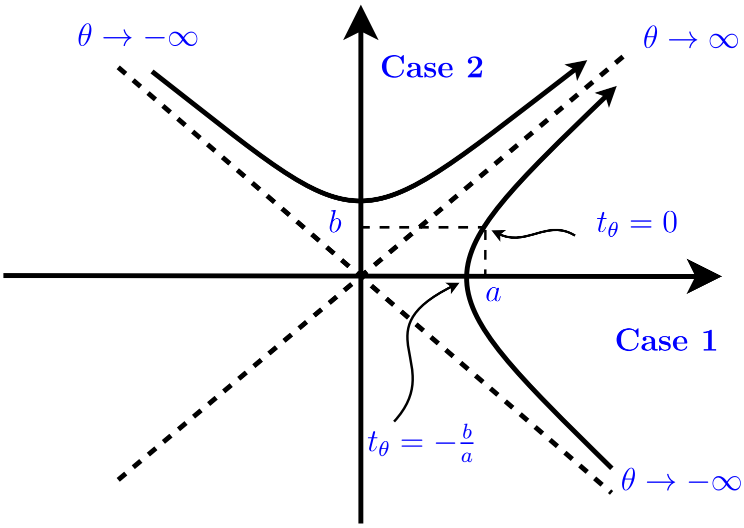


Figure 2.5: Hyperbolic curves defined by $X(\theta)$ and $Y(\theta)$ for **Case 1** and **Case 2**.

**Case 1: $a, b \geq 0$ such that $a \geq b$.** In the case where $a \geq b$, clearly $X(\theta)$ is always positive. Also, since $D_\theta X(\theta) = a\sinh\theta + b\cosh\theta$, the derivative of $X(\theta)$ changes sign when $t_\theta = -\frac{b}{a}$, whereas $Y(\theta)$ is always increasing.

**Case 2: $a, b \geq 0$ such that $a \leq b$.** In a similar way as in the above case, for $a \leq b$, it is easy to show that $X(\theta)$ is always increasing whereas the derivative of $Y(\theta)$ changes sign when $t_\theta = -\frac{a}{b}$.

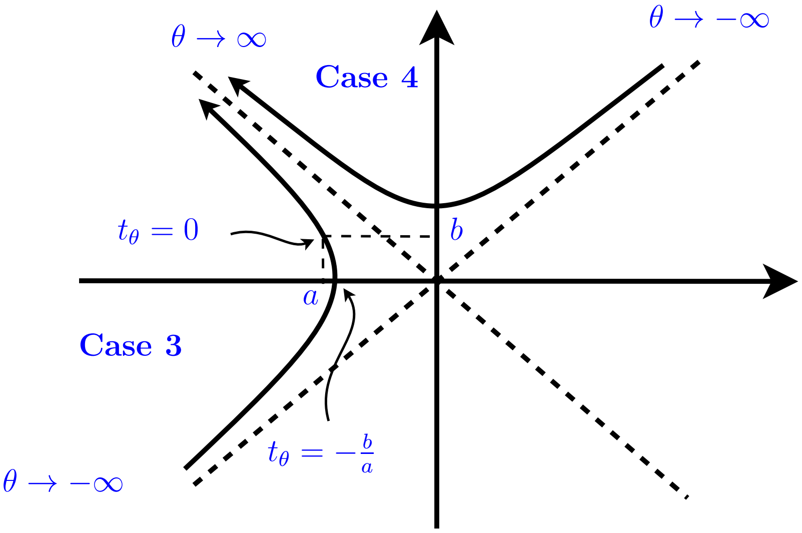


Figure 2.6: Hyperbolic curves defined by $X(\theta)$ and $Y(\theta)$ for **Case 3** and **Case 4**.

**Case 3: $a \leq 0, \quad b \geq 0$ such that $b \leq -a$.** In this case, $Y(\theta)$ is always increasing whereas the derivative of $X(\theta)$ changes sign when $t_\theta = -\frac{b}{a}$.

**Case 4: $a \leq 0, \quad b \geq 0$ such that $-a \leq b$.** It is easy to see that $X(\theta)$ is always decreasing whereas the derivative of $Y(\theta)$ changes sign when $t_\theta = -\frac{a}{b}$. **Cases 3-4** are illustrated in Figure 2.6.

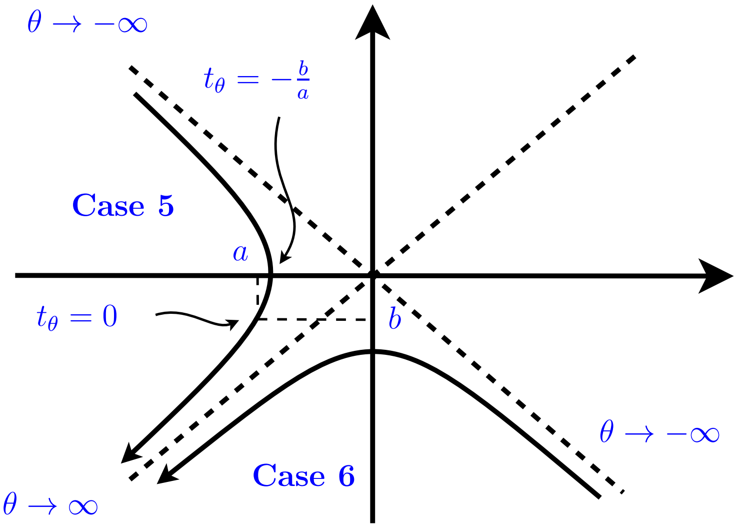


Figure 2.7: Hyperbolic curves defined by $X(\theta)$ and $Y(\theta)$ for **Case 5** and **Case 6**.

**Case 5: $a, b \leq 0$ such that $-b \leq -a$.** This case is obtained from **Case 1** by taking its symmetry with respect to the origin.

**Case 6: $a, b \leq 0$ such that $-a \leq -b$.** This case is obtained from **Case 2** by taking its symmetry with respect to the origin. **Cases 5-6** are illustrated in Figure 2.7.

**Case 7: $a \geq 0, \quad b \leq 0$ such that $-b \leq a$.** This case is obtained from **Case 3** by taking its symmetry with respect to the origin.

**Case 8: $a \geq 0, \quad b \leq 0$ such that $a \leq -b$.** This case is obtained from **Case 4** by taking its symmetry with respect to the origin. **Cases 7-8** are illustrated in Figure 2.8.

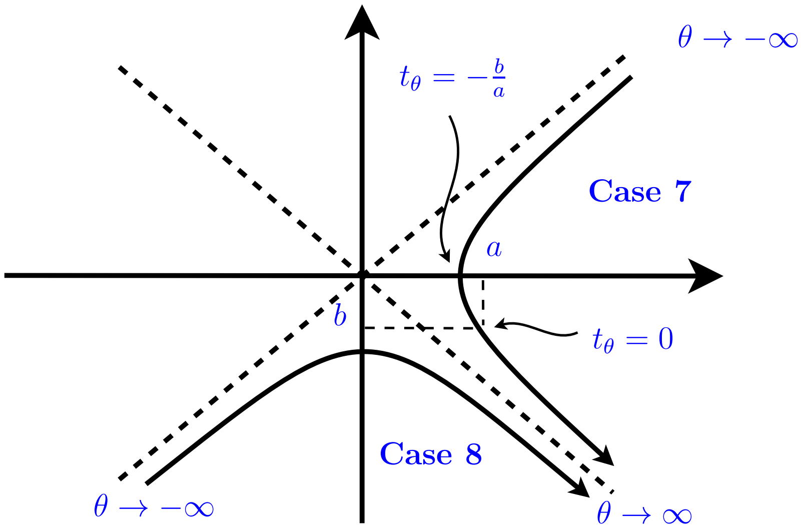


Figure 2.8: Hyperbolic curves defined by $X(\theta)$ and $Y(\theta)$ for **Case 7** and **Case 8**.

Observe that $\|\pi_1 - I_T\pi_1\|_{L_p(T)}$ is invariant under a linear transform $\phi_{M_0}$ where $M_0 \in \mathcal{I}_0$. Thus, we can assume that $x_i, y_i \geq 0$. There are four different cases for $x_j, y_j$.

$a$ If $x_j, y_j \geq 0$, then $T_1$ is an acute triangle;

*b* If $x_j \leq 0$ and $y_j \geq 0$, then it is easy to show that the angle $\widehat{S_\theta \mathbf{e}_i\, S_\theta \mathbf{e}_j} \to \frac{\pi}{2}$ as $\theta \to \infty$, thus a nearly-acute optimal triangle can be obtained. This is shown in Figure 2.9 by combining **Cases 1-4**;

*c* If $x_j \geq 0$ and $y_j \leq 0$, then as in *b*, $\widehat{S_\theta \mathbf{e}_i\, S_\theta \mathbf{e}_j} \to \frac{\pi}{2}$ as $\theta \to \infty$, and a nearly-acute optimal triangle can be obtained. This is shown in Figure 2.10 by combining **Cases 1-2** and **Cases 7-8**;

*d* If $x_j, y_j \leq 0$, then

*i* If $|x_i| \geq |y_i|$ and $|y_j| \geq |x_i|$, by letting $\theta \to -\infty$ the angle $\widehat{S_\theta \mathbf{e}_i\, S_\theta \mathbf{e}_j}$ decreases and a nearly-acute optimal triangle can be obtained. This is also the case if $|y_i| \geq |x_i|$ and $|x_j| \geq |x_i|$;

*ii* Otherwise, i.e if $|x_i| \geq |y_i|$ and $|x_j| \geq |y_j|$, or $|x_i| \leq |y_i|$ and $|x_j| \leq |y_j|$, then it may be impossible to decrease the angle $\widehat{S_\theta \mathbf{e}_i\, S_\theta \mathbf{e}_j}$. This case is the reason why designing a nearly-acute optimal triangle remains an open question. We illustrate this in Figure 2.11.

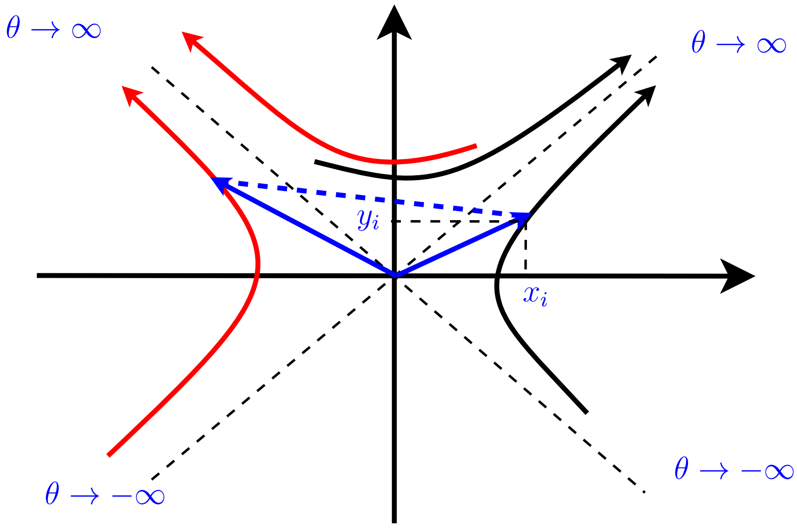


Figure 2.9: Obtaining nearly-acute optimal triangle for $\pi_1$, case b.

In brief, part *ii* of case *d* shows that characterizing optimal triangles for $\pi_1$ may be difficult since they could essentially take arbitrary shape, and the existence of an isotropic optimal triangle is not guaranteed.

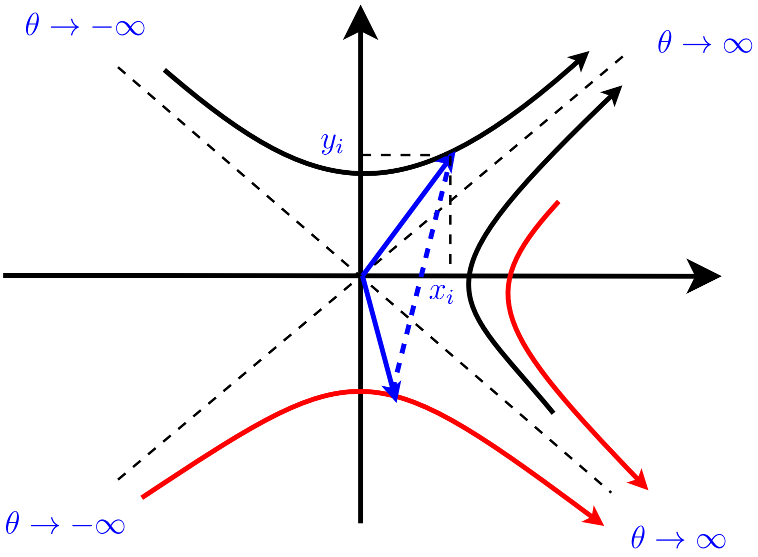


Figure 2.10: Obtaining nearly-acute optimal triangle for $\pi_1$, case c.

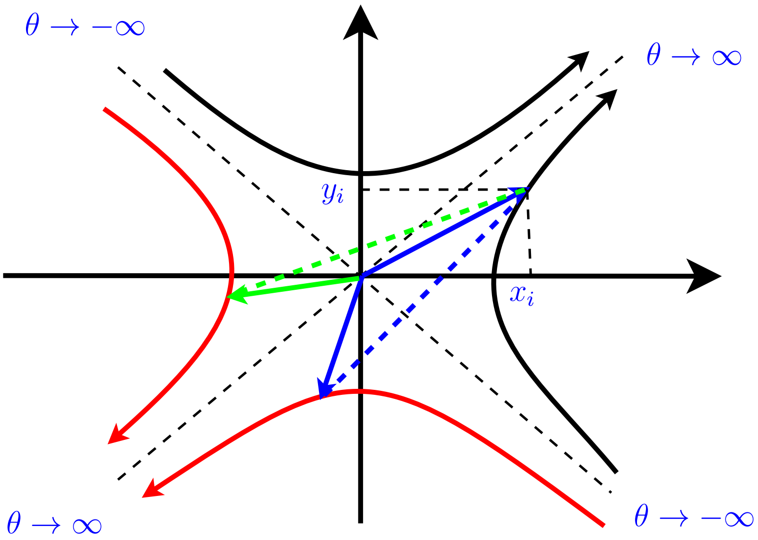


Figure 2.11: In blue for $i$ of case $d$: obtaining nearly-acute optimal triangle for $\pi_1$; and in green for $ii$ of case $d$: the angle $\widehat{S_\theta \mathbf{e}_i \, S_\theta \mathbf{e}_j}$ approaching $\pi$ if $\theta \to \pm\infty$.

### 2.4.3 For general quadratic polynomials

We show in Lemma 2.4.5 below that to obtain an optimal triangle for a quadratic polynomial $\pi$, it is sufficient to know an optimal triangle for $\varpi_\pi$. In addition, the length scales of an optimal triangle can be estimated by using the condition number of $Q_\pi$.

The result below is partially extracted from the proof of [29, Proposition 2.2], here we present a more elaborated construction of an optimal triangle.

**Lemma 2.4.5.** *Given a quadratic polynomial $\pi \in \mathbb{H}_2$, define $\phi_\pi$ and $\varpi_\pi$ respectively by* (2.42) *and* (2.43). *For any optimal triangle $T_0 \in \Delta_p(\varpi_\pi)$, its image $\phi_\pi(T_0)$ under $\phi_\pi$ is an optimal triangle for $\pi$. Moreover,*

$$c_1\left|\frac{\lambda_2}{\lambda_1}\right|^{\frac{1}{4}} \leq h_{\phi_\pi(T_0)} \leq c_2\left|\frac{\lambda_2}{\lambda_1}\right|^{\frac{1}{4}} \quad \textit{and} \quad c_1'\left|\frac{\lambda_1}{\lambda_2}\right|^{\frac{1}{4}} \leq \rho_{\phi_\pi(T_0)} \leq c_2'\left|\frac{\lambda_1}{\lambda_2}\right|^{\frac{1}{4}} \tag{2.71}$$

*hold, with $c_1, c_2$ and $c_1', c_2'$ being constants depending only on the triangle $T_0$, and $\lambda_1, \lambda_2$ the eigenvalues of $Q_\pi$ as described in* (2.32) *and* (2.2).

*Proof.* By using (2.44), we have

$$\varpi_\pi = \operatorname{sign}(\lambda_1)|\det \pi|^{-\frac{1}{2}}\pi \circ \phi_\pi.$$

Since the determinant of $\phi_\pi$ is equal to one, the image $\bar{T} = \phi_\pi(T_0)$ of a triangle $T_0$ of unit area is also of unit area. Next, since $\operatorname{sign}(\lambda_1)|\det \pi|^{-\frac{1}{2}}$ is a constant, by using the linearity of $I_T$ and its property in (2.12), we find that

$$\begin{aligned} e_{T_0}(\varpi_\pi) &= e_{T_0}\Big(\operatorname{sign}(\lambda_1)|\det \pi|^{-\frac{1}{2}}\pi \circ \phi_\pi\Big) = |\det \pi|^{-\frac{1}{2}} e_{T_0}(\pi \circ \phi_\pi) \\ &= |\det \pi|^{-\frac{1}{2}} \|\pi \circ \phi_\pi - I_{T_0}(\pi \circ \phi_\pi)\|_{L_p(T_0)} = |\det \pi|^{-\frac{1}{2}} \|(\pi - I_{\bar{T}}\pi) \circ \phi_\pi\|_{L_p(T_0)} \\ &= |\det \pi|^{-\frac{1}{2}} \|\pi - I_{\bar{T}}\pi\|_{L_p(\bar{T})}, \end{aligned}$$

from which we immediately see that

$$\inf_{|T|=1} e_T(\varpi_\pi) = |\det \pi|^{-\frac{1}{2}} \inf_{|\bar{T}|=1} e_{\bar{T}}(\pi),$$

thereby proving that for any optimal triangle $T_0$ for $\varpi_\pi$, its image $\bar{T} = \phi_\pi(T_0)$ is an optimal triangle for $\pi$.

Given an edge-vector $\mathbf{e}_0 = [x_0 \ y_0]^t$ of an optimal triangle $T_0 \in \Delta_p(\varpi_\pi)$, we have

$$|\phi_\pi(\mathbf{e}_0)| = \sqrt{\left|\frac{\lambda_2}{\lambda_1}\right|^{\frac{1}{2}} x_0^2 + \left|\frac{\lambda_1}{\lambda_2}\right|^{\frac{1}{2}} y_0^2} \le \left|\frac{\lambda_2}{\lambda_1}\right|^{\frac{1}{4}} \sqrt{x_0^2 + y_0^2} \le \left|\frac{\lambda_2}{\lambda_1}\right|^{\frac{1}{4}} h_0,$$

where $h_0$ is the diameter of $T_0$, thereby proving that $h_{\phi_\pi(T_0)} \le c_2 \left|\frac{\lambda_2}{\lambda_1}\right|^{\frac{1}{4}}$, with $c_2 = h_0$. Choosing an edge $\mathbf{e}_0$ such that $|x_0| \ne 0$, we obtain

$$\left|\frac{\lambda_2}{\lambda_1}\right|^{\frac{1}{4}} |x_0| \le |\phi_\pi(\mathbf{e}_0)| \le h_{\phi_\pi(T_0)}.$$

Since $|x_0| \ge \rho_0$ always holds for an appropriate edge $\mathbf{e}_0$, we prove that $c_1 \left|\frac{\lambda_2}{\lambda_1}\right|^{\frac{1}{4}} \le h_{\phi_\pi(T)}$, with $c_1 = \rho_0$. The rest of the proof follows easily by noticing that, for an optimal triangle $T_0$ whose area is one, we have $\rho_{\phi_\pi(T_0)} = 2/h_{\phi_\pi(T_0)}$. □

Observing from the definition of $\varpi_\pi$ in (2.42), an optimal triangle for $\pi$ is related to the eigenvalues and eigenvectors of $Q_\pi$. The same features are found in the constructions in [2, 3, 27].

For any non-degenerate homogeneous polynomial $\pi$, define the set $\mathcal{A}_\pi$ by

$$\mathcal{A}_\pi := \begin{cases} \mathcal{A}_0, & \text{if } \det \pi > 0; \\ \mathcal{A}_1, & \text{otherwise.} \end{cases} \tag{2.72}$$

where $\mathcal{A}_0$ and $\mathcal{A}_1$ are defined in (2.68) and (2.70). Since $Q_{\varpi_\pi} = \begin{bmatrix} 1 & 0 \\ 0 & \varepsilon_\pi \end{bmatrix}$ where $\varepsilon_\pi = \text{sign}(\det \pi)$, we deduce from (2.67) and (2.69) that, for any matrix $A \in \mathcal{A}_\pi$,

$$\begin{aligned} A^t Q_{\varpi_\pi} A &= \begin{cases} A^t Q_{\pi_0} A, & \text{with } A \in \mathcal{A}_0, \text{ if } \det \pi > 0, \\ A^t Q_{\pi_1} A, & \text{with } A \in \mathcal{A}_1, \text{ otherwise,} \end{cases} \\ &= Q_{\varpi_\pi}, \end{aligned} \tag{2.73}$$

that is, $A$ is $Q_{\varpi_\pi}$-orthogonal.

**Corollary 2.4.6.** *Given a quadratic polynomial $\pi$, let $\varpi_\pi$ be defined as in* (2.43) *and $\phi_\pi$ as in* (2.42)*. For any optimal triangle $T_0 \in \Delta_p(\varpi_\pi)$ and any linear map $\phi_A$, with $A \in \mathcal{A}_\pi$, the triangle $(\phi_\pi \circ \phi_A)(T_0)$ is optimal for $\pi$.*

*Proof.* Since any matrix $A \in \mathcal{A}_\pi$ is $Q_{\varpi_\pi}$-orthogonal, Lemma 2.4.2 yields that

$$\|\varpi_\pi - I_{T_0}\varpi_\pi\|_{L_p(T_0)} = \|\varpi_\pi - I_{T_0'}\varpi_\pi\|_{L_p(T_0')}$$

where $T_0' = AT_0 = \phi_A(T_0)$. Since $\varpi_\pi \in \Delta_p(\varpi_\pi)$, we have that $\phi_A(T_0) \in \Delta_p(\varpi_\pi)$. We deduce from Lemma 2.4.5 that $\phi_\pi\Big(\phi_A(T_0)\Big)$ is an optimal triangle for $\pi$. □

## 2.5 Approximation on nearly optimal triangles

In this section, we present estimations in $L_p$-norm and $W_p^1$-seminorm for the approximation error of a function on nearly optimal triangles. Our method for local approximation is similar to the one from [29], we use the Hessian of the function to perform a spectral analysis, so that the eigenvectors specify the stretching directions, the eigenvalues dictate the aspect ratio, of the nearly-optimal triangles.

Given a point $z \in \Omega$ and a function $f \in C^2(\Omega)$, where $\Omega$ is a domain in $\mathbb{R}^2$, we define the quadratic polynomial $\pi_z$ by

$$\pi_z(x,y) := \frac{1}{2}D^2_{xx}f(z)x^2 + D^2_{xy}f(z)xy + \frac{1}{2}D^2_{yy}f(z)y^2, \quad x, y \in \Omega, \tag{2.74}$$

and the corresponding *modulus of continuity* of the function $z \mapsto \pi_z$ by

$$\omega(r) := \sup_{\substack{\|z-z'\|\le r \\ z,z'\in\Omega}} \|\pi_z - \pi_{z'}\|, \quad r \ge 0. \tag{2.75}$$

The non-decreasing function $\omega(r)$ is of great use in Section 3.2 for the triangulation of $\Omega$. Also, it will appear in Section 3.4 for the study of asymptotic error estimates. Observe that the matrix $Q_{\pi_z}$ (defined in (2.1)) associated with $\pi_z$ is equal to $2H_f(z)$, with $H_f$ denoting the Hessian matrix of $f$.

We say that a triangle $T \subset \Omega$ is *nearly-optimal* if it is an optimal triangle for the quadratic polynomial $\pi_z$, with some $z \in \Omega$ close to the barycenter $b_T$ of $T$. Recall that the error does not change by translation of the triangle. Thus we can design and use nearly-optimal triangles in the neighborhood of the origin. Except in the case of estimating the derivatives of the approximation errors on nearly-optimal triangles, we shall not use classical estimations.

### 2.5.1 $L_p$-norm error bounds

Considering a fixed triangle $T$, let $e_T$ be defined as in (2.14). For any point $z \in \mathbb{R}^2$, a simple triangular inequality shows that

$$\begin{aligned} \|f - I_T f\|_{L_p(T)} &= \|(f-\pi_z) - I_T(f-\pi_z)\|_{L_p(T)} + \|\pi_z - I_T\pi_z\|_{L_p(T)} \\ &\le e_T(f-\pi_z) + e_T(\pi_z). \end{aligned} \tag{2.76}$$

The first term on the right hand side of (2.76) can be estimated by using Lemma 2.5.2 below, whereas the second term can be estimated by using Lemma 2.5.3. Both of these results have some similarities with the results in [2, 29].

**Lemma 2.5.1.** *Let $\varphi_z$ denote the linear polynomial in the Taylor expansion of $f \in C^2(\Omega)$ at a point $z \in T$. With $\pi_z$ defined as in (2.74), we have*

$$\|f - (\varphi_z + \pi_z)\|_{L_\infty(T)} \le h_T^2 \omega(h_T).$$

*Proof.* Let $z_0, z_1 \in T$ be fixed and consider the function $g(t) = f(z_0 + t(z_1 - z_0))$. Since $f \in C^2(T)$, it is clear that $g \in C^2([0,1])$. Denoting by $g^{(m)}$, $m = 0, 1$, the $m$-th derivative of $g$, the integral form of the Taylor expansion at 0 is given by

$$g(t) = g(0) + g^{(1)}(0)t + \frac{1}{2}\int_0^t g^{(2)}(s)(t-s)\mathrm{d}s, \tag{2.77}$$

where, by writing $z_0 = (x_0, y_0)$, $z_1 = (x_1, y_1)$ and $z_s = z_0 + s(z_1 - z_0)$, $s \in [0, 1]$,

$$g^{(1)}(s) =(x_1 - x_0)D_x f(z_s) + (y_1 - y_0)D_y f(z_s), \tag{2.78}$$

$$\begin{aligned} g^{(2)}(s) =&(x_1 - x_0)^2 D^2_{xx} f(z_s) + 2(x_1 - x_0)(y_1 - y_0)D^2_{xy} f(z_s) \\ &+ (y_1 - y_0)^2 D^2_{yy} f(z_s). \end{aligned} \tag{2.79}$$

Noting that $\varphi_{z_0}(x_1, y_1) = f(z_0) + (x_1 - x_0)D_x f(z_0) + (y_1 - y_0)D_y f(z_0) = g^{(1)}(0)$ and that $f(z_1) = g(1)$, we deduce from (2.77) that

$$f(z_1) - \varphi_{z_0}(z_1) = \int_0^1 g^{(2)}(s)(1 - s)\mathrm{d}s. \tag{2.80}$$

With $z_s = (x_s, y_s)$, define the function $\psi_{z_s}(x, y) = \pi_{z_s}(x - x_s, y - y_s)$. We use (2.79) to deduce that

$$\begin{aligned} f(z_1) - \varphi_{z_0}(z_1) - \psi_{z_0}(z_1) =&2 \int_0^1 \Big(\pi_{z_s}(z_1 - z_0) - \psi_{z_0}(z_1)\Big)(1 - s)\mathrm{d}s \\ =&2 \int_0^1 \Big(\pi_{z_s}(z_1 - z_0) - \pi_{z_0}(z_1 - z_0)\Big)(1 - s)\mathrm{d}s. \end{aligned} \tag{2.81}$$

Observe that for any $\pi \in \mathbb{H}_2$ and any $z = (x, y) \in \Omega$, clearly $|\pi(z)| \leq \|\pi\|(|x|^2 + |xy| + |y|^2) \leq \|\pi\||z|^2$. Thus, since $|z_1 - z_0| \leq \operatorname{diam}(T)$, for all $s \in [0, 1]$,

$$|\pi_{z_s}(z_1 - z_0) - \pi_{z_0}(z_1 - z_0)| \leq \|\pi_{z_s} - \pi_{z_0}\||z_1 - z_0|^2 \leq \omega(\operatorname{diam}(T)) \operatorname{diam}(T)^2.$$

Noticing that (2.81) holds for any $z_1 \in T$, and also that $2\int_0^1 (1 - s)ds = 1$, we deduce that $\|f - \varphi_{z_0} - \psi_{z_0}\|_{L_\infty(T)} \leq \omega(\operatorname{diam}(T)) \operatorname{diam}(T)^2$. □

The proof of Lemma 2.5.2 below is inspired from the proofs of results in [29].

**Lemma 2.5.2.** *For any $f \in C^2(\Omega)$ and any triangle $T \subset \Omega$,*

$$e_T(f - \pi_z) \leq 2h^2\omega(h)|T|^{\frac{1}{p}}, \quad z \in T.$$

*Proof.* Let $\varphi_z$ be the linear polynomial in the Taylor expansion of $f$ at $z$, $z \in T$, it satisfies $\varphi_z = I_T\varphi_z$. We thereby have that

$$\begin{aligned} e_T(f-\pi_z) =& \|(f-\pi_z-\varphi_z) - I_T(f-\pi_z-\varphi_z)\|_{L_p(T)} \\ \leq& \|f-\pi_z-\varphi_z)\|_{L_p(T)} + \|I_T(f-\pi_z-\varphi_z)\|_{L_p(T)} \\ \leq& 2|T|^{\frac{1}{p}}\|f-\pi_z-\varphi_z\|_{L_\infty(T)} \\ \leq& 2|T|^{\frac{1}{p}}h^2\omega(h), \end{aligned}$$

by virtue of Lemma 2.5.1 and (2.13). □

In Lemma 2.5.2, the difference $f-\pi_z$ is bounded by a factor times $\omega(h_T)$ which may be very small if $f$ behaves like $\pi_z$ in the neighborhood of $z$. We shall come back to this in Chapter 3.

We are now interested in the approximation error $e_T(\pi_z)$ with the condition that $T$ is an scaled and translated version of an optimal triangle for $\pi_t$ for some $t \in \Omega$. Generally $T$ is not optimal for $\pi_t$, however if $z$ and $t$ are close to one another, we can bound $e_T(\pi_z)$ by using the result below which is also inspired from the proof of [29, Proposition 3.2].

**Lemma 2.5.3.** *Given $t \in \Omega$, let $T = c_1T_0 + \mathbf{t}_1$ where $T_0 \in \Delta_p(\pi_t)$, with $c_1$ a non-zero scaling factor and $\mathbf{t}_1$ a translation vector. Then, for any $z \in \Omega$,*

$$e_T(\pi_z) \leq \Big(K_p(\pi_z) + Ch_{T_0}^2\omega(|z-t|)\Big)|T|^{1+\frac{1}{p}}, \tag{2.82}$$

*where the shape function $K_p$ is defined in (2.53) and $C$ an absolute constant.*

*Proof.* Note that $T_0$ is of unit area. By using Lemma 2.1.2 and Lemma 2.4.1, there is a constant $C$ such that

$$\begin{aligned} e_{T_0}(\pi_z) \leq e_{T_0}(\pi_t) + Ch_{T_0}^2\|\pi_z-\pi_t\| &= K_p(\pi_t) + Ch_{T_0}^2\|\pi_z-\pi_t\| \\ \leq K_p(\pi_z) + 2Ch_{T_0}^2\|\pi_z-\pi_t\| &\leq K_p(\pi_z) + 2Ch_{T_0}^2\omega(|z-t|). \end{aligned} \tag{2.83}$$

Note also that $|T| = |c_1T_0| = c_1^2$. Writing $\mathbf{t}_1 = [x_1\ y_1]^t$, the change of variable

$\phi(x,y) = (c_1 x + x_1, c_1 y + y_1)$ has its Jacobian equal to $c_1^2 = |T|$, hence

$$e_T(\pi_z) = \Big( \int_{T_0} c_1^{2p} |\pi_z(x,y) - I_{T_0}\pi_z(x,y)|^p |T| \mathrm{d}x\mathrm{d}y \Big)^{\frac{1}{p}} = e_{T_0}(\pi_z)|T|^{1+\frac{1}{p}},$$

which, together with (2.83), yields (2.82). □

**Remark 2.5.1.** Choosing $z = t$ in (2.82) and using (2.54), and with $\varpi_{\pi_z}$ defined as in (2.43), we obtain

$$e_T(\pi_z) \le K_p(\varpi_{\pi_z})|T|^{1+\frac{1}{p}} \sqrt{|\det \pi_z|} = \frac{1}{2} K_p(\varpi_{\pi_z})|T|^{\frac{1}{p}} h\rho \sqrt{|\det \pi_z|}. \tag{2.84}$$

In the above estimation, the triangle $T = c_1 T_0 + \mathbf{t}_1$ is a scaled and translated version of an optimal triangle $T_0 \in \Delta_p(\pi_z)$. On such a triangle, a similar estimation to (2.84) is given by part (i) of Proposition 2.3.5,

$$e_T(\pi_z) \le C_1 \rho_{\pi_z}(T)|T|^{\frac{1}{p}} h\rho \sqrt{|\det \pi_z|}, \tag{2.85}$$

where $\rho_{\pi_z}(T)$ is a constant since the measure $\rho_{\pi_z}$, which is defined in (2.39), is invariant under translation and scaling by a constant of the triangle so that, by using (2.45),

$$\begin{aligned} \rho_{\pi_z}(T) &= \rho_{\varpi_{\pi_z}}\Big(\phi_{\pi_z}^{-1}(c_1 T_0 + \mathbf{t}_1)\Big) = \rho_{\varpi_{\pi_z}}\Big(\phi_{\pi_z}^{-1}\Big(c_1 \phi_{\pi_z}(\bar{T}) + \mathbf{t}_1\Big)\Big) \\ &= \rho_{\varpi_{\pi_z}}\Big(c_1 \bar{T} + \phi_{\pi_z}^{-1}(\mathbf{t}_1)\Big) = \rho_{\varpi_{\pi_z}}(\bar{T}), \end{aligned}$$

where $T_0 = \phi_{\pi_z}(\bar{T})$, with $\bar{T} \in \Delta_p(\varpi_{\pi_z})$. Observe that, if $\det \pi_z > 0$ for all $z$, then $\rho_{\varpi_{\pi_z}}(\bar{T})$ is a constant since $\bar{T}$ is equilateral.

Coming back now to estimating the error on a nearly-optimal triangle, we have the following result. A similar estimation to (2.86) can be found in [29].

**Proposition 2.5.4.** *Given $t \in \Omega$, let $T = c_1 T_0 + \mathbf{t}_1$ where $c_1$ is a non-zero constant, $\mathbf{t}_1$ a translation vector and $T_0 \in \Delta_p(\pi_t)$. Then, for any $z \in T$,*

$$\|f - I_T f\|_{L_p(T)} \le \Big( K_p(\pi_z) + C h_{T_0}^2 \omega\Big(\max\{|z - t|, h_T\}\Big)\Big)|T|^{1+\frac{1}{p}}, \tag{2.86}$$

*with the function $\omega$ defined in (2.75), and $C$ an absolute constant.*

*Proof.* By using (2.76), together with Lemmas 2.5.3 and 2.5.2,

$$\|f - I_T f\|_{L_p(T)} \leq \Big(K_p(\pi_z) + Ch_{T_0}^2 \omega(|z-t|)\Big)|T|^{1+\frac{1}{p}} + 2h_T^2 \omega(h_T)|T|^{\frac{1}{p}},$$

holds. Since $c_1 = |T|^{\frac{1}{2}}$ and $h_T = c_1 h_{T_0}$, we find that $h_T^2 = |T| h_{T_0}^2$. Inserting this into the above inequality yields

$$\begin{aligned}\|f - I_T f\|_{L_p(T)} &\leq \Big(K_p(\pi_z) + Ch_{T_0}^2 \omega(|z-t|)\Big)|T|^{1+\frac{1}{p}} + 2h_{T_0}^2 \omega(h_T)|T|^{1+\frac{1}{p}} \\ &\leq \Big(K_p(\pi_z) + (C+2)h_{T_0}^2 \omega\Big(\max\{|z-t|, h_T\}\Big)\Big)|T|^{1+\frac{1}{p}},\end{aligned}$$

which proves (2.86). □

The result in Proposition 2.5.4 is important in Section 3.4.1 in order to obtain the asymptotic $L_p$-norm of the error resulting from the approximation of $f \in C^2(\Omega)$ on an anisotropic triangulation $\Delta_{s,r}$ of $\Omega$.

### 2.5.2 Sobolev seminorm error bounds

Many areas in numerical analysis require the study of derivatives for a given approximation problem. In this view, we are interested in estimating the $W_p^1$-seminorm of the error resulting from the approximation of a function on a given triangle. The results here are essential for Chapter 3 in order to derive the asymptotic estimation in $W_p^1$-seminorm.

First we impose no restriction on a given triangle $T$. Recall that $h, \rho$ are the diameter and smallest height of $T$, and the unit vectors $\boldsymbol{\sigma}_h, \boldsymbol{\sigma}_\rho$ are defined on page *18*. The following lemma is a consequence of Lemma 2.2.2.

**Lemma 2.5.5.** *Consider a function $f \in W_p^2(T)$ and a pair $(\boldsymbol{\sigma}, \boldsymbol{\tau})$ of orthonormal vectors. With $\eta \in [0, 2\pi]$ denoting the angle between $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma}_h$,*

$$|f - I_T f|_{W_p^1(T)} \lesssim \Big(h + \rho|\sin\eta|\Big)|D_{\boldsymbol{\sigma}} f|_{W_p^1(T)} + \Big(\rho + h|\sin\eta|\Big)|D_{\boldsymbol{\tau}} f|_{W_p^1(T)}, \quad (2.87)$$

*with constant depending only on the maximum interior angle $\gamma(T)$ of $T$.*

*Proof.* From the definition of directional derivatives, we have that

$$D_{\boldsymbol{\sigma}_h} f = \cos\theta_1 D_x f + \sin\theta_1 D_y f, \quad D_{\boldsymbol{\sigma}_\rho} f = -\sin\theta_1 D_x f + \cos\theta_1 D_y f,$$

with $\theta_1 \in [0, 2\pi]$ being the angle between the $x$-axis and $\boldsymbol{\sigma}_h$. Observe that the derivatives in $x$- and $y$- directions can be expressed in terms of the derivatives in the directions of a given pair of orthonormal vectors $(\boldsymbol{\sigma}, \boldsymbol{\tau})$,

$$D_x f = \cos\theta_2 D_{\boldsymbol{\sigma}} f - \sin\theta_2 D_{\boldsymbol{\tau}} f, \quad D_y f = \sin\theta_2 D_{\boldsymbol{\sigma}} f + \cos\theta_2 D_{\boldsymbol{\tau}} f,$$

where $\theta_2 \in [0, 2\pi]$ is the angle between the $x$-axis and $\boldsymbol{\sigma}$. It follows that

$$\begin{aligned} D_{\boldsymbol{\sigma}_h} f &= \cos(\theta_1 - \theta_2) D_{\boldsymbol{\sigma}} f - \sin(\theta_1 - \theta_2) D_{\boldsymbol{\tau}} f \\ D_{\boldsymbol{\sigma}_\rho} f &= \sin(\theta_1 - \theta_2) D_{\boldsymbol{\sigma}} f + \cos(\theta_1 - \theta_2) D_{\boldsymbol{\tau}} f. \end{aligned}$$

With $\eta$ denoting the angle between $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma}_h$, simple triangular inequalities yield

$$\begin{aligned} |D_{\boldsymbol{\sigma}_h} f|_{W_p^1(T)} &\leq |D_{\boldsymbol{\sigma}} f|_{W_p^1(T)} + |\sin\eta| |D_{\boldsymbol{\tau}} f|_{W_p^1(T)}, \\ |D_{\boldsymbol{\sigma}_\rho} f|_{W_p^1(T)} &\leq |\sin\eta| |D_{\boldsymbol{\sigma}} f|_{W_p^1(T)} + |D_{\boldsymbol{\tau}} f|_{W_p^1(T)}. \end{aligned}$$

Substituting the above expressions into (2.20) yields

$$\begin{aligned} |f - I_T f|_{W_p^1(T)} \lesssim & h\Big(|D_{\boldsymbol{\sigma}} f|_{W_p^1(T)} + |\sin\eta| |D_{\boldsymbol{\tau}} f|_{W_p^1(T)}\Big) \\ & + \rho\Big(|\sin\eta| |D_{\boldsymbol{\sigma}} f|_{W_p^1(T)} + |D_{\boldsymbol{\tau}} f|_{W_p^1(T)}\Big) \\ = & \Big(h + \rho|\sin\eta|\Big) |D_{\boldsymbol{\sigma}} f|_{W_p^1(T)} + \Big(\rho + h|\sin\eta|\Big) |D_{\boldsymbol{\tau}} f|_{W_p^1(T)}, \end{aligned}$$

thereby proving the result. □

Our first objective is to express the terms $|D_{\boldsymbol{\sigma}} f|_{W_p^1(T)}$ and $|D_{\boldsymbol{\tau}} f|_{W_p^1(T)}$ in (2.87) by using the eigenvalues of a quadratic polynomial $\pi$ whose Hessian is sufficiently close to the Hessian of $f$ on a triangle $T$. To this end, we provide below some relations between the derivatives in $\boldsymbol{\sigma}, \boldsymbol{\tau}$ directions and the derivatives in $x$- and $y$- directions, the latter being the axes of the standard Cartesian system.

**Lemma 2.5.6.** *Given a counterclockwise rotation matrix $R_\theta$ of angle $\theta$, denote its columns vectors by $\boldsymbol{\sigma}, \boldsymbol{\tau}$. The following equalities of functions hold,*

$$\begin{aligned} D_{xx}(f \circ R_\theta) &= (D_{\boldsymbol{\sigma\sigma}} f) \circ R_\theta, \\ D_{yy}(f \circ R_\theta) &= (D_{\boldsymbol{\tau\tau}} f) \circ R_\theta, \\ D_{xy}(f \circ R_\theta) &= (D_{\boldsymbol{\sigma\tau}} f) \circ R_\theta. \end{aligned}$$

*Proof.* Using the differentiation rule for composite functions and using the notations $\bar{x} = \cos\theta x - \sin\theta y$ and $\bar{y} = \sin\theta x + \cos\theta y$ where $x, y \in \mathbb{R}$, we have

$$\begin{aligned} D_x(f \circ R_\theta)(x,y) =& D_x f(\bar{x},\bar{y}) = \cos\theta (D_x f)(\bar{x},\bar{y}) + \sin\theta (D_y f)(\bar{x},\bar{y}) \\ =& (D_{\boldsymbol{\sigma}} f)(\cos\theta x - \sin\theta y, \sin\theta x + \cos\theta y), \end{aligned}$$

by virtue of the fact that $D_{\boldsymbol{\sigma}} f = \cos\theta D_x f + \sin\theta D_y f$. A repeated process of the above equality yields

$$D_x(D_x(f \circ R_\theta)) = D_x((D_\sigma f) \circ R_\theta) = (D_\sigma(D_\sigma f)) \circ R_\theta.$$

The same argument is applied to prove the rest of the result. □

Let us now impose a condition on the target function $f \in C^2(\Omega)$ which we shall use in Chapter 3. Let $(\boldsymbol{\sigma}, \boldsymbol{\tau})$ be any pair of orthonormal vectors, $z_0 \in \Omega$ and $\mathcal{B}(z_0, d)$ the ball centered at $z_0$ and with radius $d \geq 0$. Suppose that there exists a small number $\nu > 0$ such that, for any $z \leq \mathcal{B}(z_0, d)$,

$$|D^2_{\mathbf{ij}} f(z_0) - D^2_{\mathbf{ij}} f(z)| \leq \nu, \tag{2.88}$$

holds, with $\mathbf{i}, \mathbf{j} \in \{\boldsymbol{\sigma}, \boldsymbol{\tau}\}$. The above condition implies that the second derivatives of $f$ do not significantly change in the neighborhood of $z_0$ (also assumed in [2, 3, 29]), allowing us to view $f$ locally as the homogeneous quadratic polynomial

$\pi = \pi_{z_0}$. In the result below, we derive estimations involving the eigenvalues of $Q_\pi$.

**Proposition 2.5.7.** *Consider a function $f \in C^2(T)$ where $T \subset \mathcal{B}(z_0, d)$ for some $z_0 \in \Omega$ and $d > 0$, and assume that (2.88) holds for a sufficiently small number $\nu > 0$. Then, with $\pi := \pi_{z_0}$, we have*

$$\|\pi - \pi_z\| \leq \nu, \quad \text{for all} \ \ z \in T, \tag{2.89}$$

*where $\|\cdot\|$ is the norm on $\mathbb{H}_2$ defined in (2.3). With $\lambda_1, \lambda_2$ such that $|\lambda_1| \leq |\lambda_2|$ being the eigenvalues of the matrix $Q_\pi$ defined in (2.2), we have*

$$|f - I_T f|_{W_p^1(T)} \lesssim \Big(2h|\lambda_1| + \rho|\lambda_2| + 8\nu h\Big)|T|^{\frac{1}{p}}. \tag{2.90}$$

*If, moreover, $\frac{\rho}{h} \sim \left|\frac{\lambda_1}{\lambda_2}\right|^{\frac{1}{2}}$, then*

$$|f - I_T f|_{W_p^1(T)} \lesssim \Big(\sqrt{|\det \pi|} + 8\nu\Big) h|T|^{\frac{1}{p}}. \tag{2.91}$$

*The constants in both (2.90) and (2.91) depend only on $\gamma(T)$.*

*Proof.* The result in (2.89) is straightforward. Consider a pair $(\boldsymbol{\sigma}, \boldsymbol{\tau})$ of orthonormal vectors. Given a point $z \in T$, (2.88) implies that $|D^2_{\mathbf{ij}} f(z) - D^2_{\mathbf{ij}} \pi(z)| \leq \nu$ so that $|D^2_{\mathbf{ij}} f(z)| \leq |D^2_{ij}\pi(z)| + \nu$ for all $\mathbf{i}, \mathbf{j} \in \{\boldsymbol{\sigma}, \boldsymbol{\tau}\}$. With $\eta \in [0, 2\pi]$ denoting the angle between $\boldsymbol{\sigma}$ and $\boldsymbol{\sigma}_h$, and since from (1.22) $\|\nabla D_{\boldsymbol{\sigma}} f\|_{L_p(T)} \leq |D_{\boldsymbol{\sigma}} f|_{W_p^1(T)}$, Lemma 2.5.5 yields

$$\begin{aligned}
|f - I_T f|_{W_p^1(T)} \lesssim &\,(h + \rho|\sin\eta|)\Big(\int_T \Big(|D^2_{\boldsymbol{\sigma\sigma}} f(z)|^2 + |D^2_{\boldsymbol{\sigma\tau}} f(z)|^2\Big)^{\frac{p}{2}} \mathrm{d}z\Big)^{\frac{1}{p}} \\
&+ (\rho + h|\sin\eta|)\Big(\int_T \Big(|D^2_{\boldsymbol{\sigma\tau}} f(z)|^2 + |D^2_{\boldsymbol{\tau\tau}} f(z)|^2\Big)^{\frac{p}{2}} \mathrm{d}z\Big)^{\frac{1}{p}} \\
\lesssim &(h + \rho|\sin\eta|)\Big(\int_T \Big((|D^2_{\boldsymbol{\sigma\sigma}} \pi(z)| + \nu)^2 + (|D^2_{\boldsymbol{\sigma\tau}} \pi(z)| + \nu)^2\Big)^{\frac{p}{2}} \mathrm{d}z\Big)^{\frac{1}{p}} \\
&+ (\rho + h|\sin\eta|)\Big(\int_T \Big((|D^2_{\boldsymbol{\sigma\tau}} \pi(z)| + \nu)^2 + (|D^2_{\boldsymbol{\tau\tau}} \pi(z)| + \nu)^2\Big)^{\frac{p}{2}} \mathrm{d}z\Big)^{\frac{1}{p}}.
\end{aligned}$$

Observe that, with $\lambda_1, \lambda_2$ and $U_\pi$ being as in (2.1) and (2.2), for any $x, y \in \mathbb{R}$,

$$\pi \circ U_\pi(x,y) = [x\ y]\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}[x\ y]^t = \lambda_1 x^2 + \lambda_2 y^2,$$

which, together with Lemma 2.5.6, yields

$$\begin{aligned}
D^2_{xx}(\pi \circ U_\pi)(x,y) &= (D^2_{\boldsymbol{\sigma}\boldsymbol{\sigma}}\pi)\circ U_\pi(x,y) = \lambda_1,\\
D^2_{yy}(\pi \circ U_\pi)(x,y) &= (D^2_{\boldsymbol{\tau}\boldsymbol{\tau}}\pi)\circ U_\pi(x,y) = \lambda_2,\\
D^2_{xy}(\pi \circ U_\pi)(x,y) &= (D^2_{\boldsymbol{\sigma}\boldsymbol{\tau}}\pi)\circ U_\pi(x,y) = 0.
\end{aligned}$$

By considering the triangle $T_0 = U_\pi^{-1}(T)$, we have $|T_0| = |T|$ and

$$\begin{aligned}
|f - I_T f|_{W^1_p(T)} \lesssim{}& (h + \rho|\sin\eta|)\Big(\int_{T_0}\Big((|(D^2_{\boldsymbol{\sigma}\boldsymbol{\sigma}}\pi)\circ U_\pi(z)| + \nu)^2\\
&+ (|(D^2_{\boldsymbol{\sigma}\boldsymbol{\tau}}\pi)\circ U_\pi(z)| + \nu)^2\Big)^{\frac{p}{2}}\mathrm{d}z\Big)^{\frac{1}{p}} + (\rho + h|\sin\eta|)\\
&\Big(\int_{T_0}\Big((|(D^2_{\boldsymbol{\sigma}\boldsymbol{\tau}}\pi)\circ U_\pi(z)| + \nu)^2 + (|(D^2_{\boldsymbol{\tau}\boldsymbol{\tau}}\pi)\circ U_\pi(z)| + \nu)^2\Big)^{\frac{p}{2}}\mathrm{d}z\Big)^{\frac{1}{p}}
\end{aligned}$$

$$\begin{aligned}
=&|T_0|^{\frac{1}{p}}\Big(\Big(h + \rho|\sin\eta|\Big)\Big((|\lambda_1| + \nu)^2 + \nu^2\Big)^{\frac{1}{2}} + \Big(\rho + h|\sin\eta|\Big)\Big((|\lambda_2| + \nu)^2 + \nu^2\Big)^{\frac{1}{2}}\Big)\\
\leq&|T|^{\frac{1}{p}}\Big(\Big(h + \rho|\sin\eta|\Big)|\lambda_1| + \Big(\rho + h|\sin\eta|\Big)|\lambda_2| + 2\nu(h+\rho)\Big(1 + |\sin\eta|\Big)\Big),
\end{aligned}$$

with constant depending only on $\gamma(T)$. Since the above result holds for any pair $(\boldsymbol{\sigma}, \boldsymbol{\tau})$, we can choose $(\boldsymbol{\sigma}, \boldsymbol{\tau}) = (\boldsymbol{\sigma}_\rho, \boldsymbol{\sigma}_h)$ so that $\eta = 0$. Hence (2.90).

By factorizing $h$ in (2.90), we obtain

$$\begin{aligned}
|f - I_T f|_{W^1_p(T)} &\lesssim \Big(2|\lambda_1| + \frac{\rho}{h}|\lambda_2| + 8\nu\Big)h|T|^{\frac{1}{p}}\\
&\lesssim \Big(2|\lambda_1\lambda_2|^{\frac{1}{2}} + C\Big|\frac{\lambda_1}{\lambda_2}\Big|^{\frac{1}{2}}|\lambda_2| + 8\nu\Big)h|T|^{\frac{1}{p}}\\
&\lesssim \Big((2 + C)|\lambda_1\lambda_2|^{\frac{1}{2}} + 8\nu\Big)h|T|^{\frac{1}{p}},
\end{aligned}$$

which proves the result in (2.91). □

The condition $\frac{\rho}{h} \sim \left|\frac{\lambda_1}{\lambda_2}\right|^{\frac{1}{2}}$ is satisfied by triangles obtained by Lemma 2.4.5 by using optimal or nearly-optimal triangles. This condition is also satisfied for the triangle considered in Example 2.3.2, with $h_1 = h = \left|\frac{\lambda_2}{\lambda_1}\right|^{\frac{1}{4}}$ and $\rho = h_2 = \left|\frac{\lambda_1}{\lambda_2}\right|^{\frac{1}{4}}$. We obtain the following results for nearly-optimal triangles.

**Corollary 2.5.8.** *Under the conditions of Proposition 2.5.7, suppose that $T = c_1T_0 + \mathbf{t}_1$ is a scaled and translated version of an optimal triangle $T_0 \in \Delta_p(\pi)$, where $c_1$ is a non-zero constant and $\mathbf{t}_1$ a translation vector. Then,*

$$|f - I_T f|_{W_p^1(T)} \lesssim \left(\sqrt{|\det \pi|} + 8\nu\right) h|T|^{\frac{1}{p}}. \tag{2.92}$$

*If, moreover, $\nu \leq \sqrt{|\det \pi|}$, then*

$$|f - I_T f|_{W_p^1(T)} \lesssim |\det \pi|^{\frac{1}{4}} \|Q_\pi\|_2^{\frac{1}{2}} |T|^{\frac{1}{2}+\frac{1}{p}}. \tag{2.93}$$

*The constants in both (2.92) and (2.93) depend only on $\gamma(T)$.*

*Proof.* Since $h = c_1 h_{T_0}$ and $\rho = c_1\rho_{T_0}$, we deduce from (2.71) that

$$\frac{\rho}{h} = \frac{\rho_{T_0}}{h_{T_0}} \sim \left|\frac{\lambda_1}{\lambda_2}\right|^{\frac{1}{2}},$$

where $\lambda_1, \lambda_2$ are the eigenvalues of the matrix $Q_\pi$. Using Proposition 2.5.7, we deduce from (2.91) that

$$|f - I_T f|_{W_p^1(T)} \lesssim \left(\sqrt{|\det \pi|} + 8\nu\right) h|T|^{\frac{1}{p}},$$

with constant depending only on $\gamma(T)$.

In the case where $\nu \leq \sqrt{|\det \pi|}$, the fact that $|\det \pi| = |\lambda_1\lambda_2|$ and

$$h|T|^{\frac{1}{p}} = \sqrt{2}\left(\frac{h}{\rho}\right)^{\frac{1}{2}} |T|^{\frac{1}{2}+\frac{1}{p}} \sim \left|\frac{\lambda_2}{\lambda_1}\right|^{\frac{1}{4}} |T|^{\frac{1}{2}+\frac{1}{p}} \tag{2.94}$$

shows that

$$|f - I_T f|_{W^1_p(T)} \lesssim \left|\frac{\lambda_2}{\lambda_1}\right|^{\frac{1}{4}} |\lambda_1\lambda_2|^{\frac{1}{2}} |T|^{\frac{1}{2}+\frac{1}{p}} = |\det \pi|^{\frac{1}{4}} \|Q_\pi\|_2^{\frac{1}{2}} |T|^{\frac{1}{2}+\frac{1}{p}},$$

thereby proving the result in (2.93). □

The condition in (2.88) characterizes the partition into sub-squares of a square domain $\Omega$ in Section 3.2.2.

The estimations in Corollary 2.5.8 use a nearly-optimal triangle $T$ whose diameter is $h \sim \left|\frac{\lambda_2}{\lambda_1}\right|^{\frac{1}{4}}$ where $\lambda_1, \lambda_2$ are the eigenvalues of $\pi$. In the case where $f$ is actually a homogeneous quadratic polynomial, as such $f = \pi$, due to the equivalence in (2.94) and the fact that $\nu = 0$, it is sufficient to use (2.92) to obtain the estimation

$$|\pi - I_T \pi|_{W^1_p(T)} \lesssim |\lambda_1|^{\frac{1}{4}} |\lambda_2|^{\frac{3}{4}} |T|^{\frac{1}{p}}. \tag{2.95}$$

On the other hand, for a nearly-optimal triangle $T$ whose measure of non-degeneracy $\rho_\pi(T)$ is bounded by a constant, we deduce from Proposition 2.3.5 that

$$|\pi - I_T \pi|_{W^1_p(T)} \leq C(h^p + \rho^p)^{\frac{1}{p}} \sqrt{|\det \pi|} |T|^{\frac{1}{p}} \leq C' |\lambda_1|^{\frac{1}{4}} |\lambda_2|^{\frac{3}{4}} |T|^{\frac{1}{p}},$$

for some constants $C$ and $C'$, thereby achieving a similar estimation as in (2.95).

## 2.6 Approximation on non-optimal triangles.

In general, there is no method as to know whether a given triangle $T$ is nearly-optimal for some $\pi_z$ or not. The $L_p$-norm estimations do not suffer from this since we use (2.18) for $m = 0$ to obtain

$$\|f - I_T f\|_{L_p(T)} \leq Ch^2 |f|_{W^2_\infty(T)}, \tag{2.96}$$

where $C$ is an absolute constant.

Our goal in this section is to provide $W^1_p$-seminorm estimations for the approximation error on triangles which are not nearly-optimal. The results in this section are of great use in Section 3.4.4 in order to estimate the $W^1_p$-seminorm of the errors on the so-called *interface* triangles.

### 2.6.1 Using invertible affine maps

Instead of ensuring that the maximum interior angle $\gamma(T)$ of a triangle $T$ is far from the flat angle, suppose that there exists a linear map $\varphi$ such that $\operatorname{cond}(\varphi)$ is bounded and $\gamma(\varphi^{-1}(T))$ is far from the flat angle. It is necessary to assume that the condition number is bounded since otherwise such linear map always exists. We have the following result.

**Lemma 2.6.1.** *Given a triangle $T$, let $\varphi$ be an invertible affine map. Then, for any function $f \in C^2(T)$, we have*

$$|f - I_T f|_{W^1_p(T)} \lesssim \operatorname{cond}(\varphi)^2 h |f|_{W^2_p(T)}, \tag{2.97}$$

*with constant depending only on the maximum angle $\gamma\Big(\varphi^{-1}(T)\Big)$.*

*Proof.* Let the matrix associated with $\phi$ be written as $M = U\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}V^t$ where $U$ and $V$ are rotation matrices, and $|a|, |b|$ the singular values of $M$. Since rotation does not change the interior angles of a triangle, we can assume without loss of generality that $V = I$. Consider the change of variables $(x, y) = \varphi(\bar{x}, \bar{y})$. Since $U$ is a rotation matrix of some angle $\theta$,

$$\varphi(\bar{x}, \bar{y}) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}[\bar{x}\ \bar{y}]^t + [t_1\ t_2]^t = \Big[\varphi_1(\bar{x}, \bar{y}) \quad \varphi_2(\bar{x}, \bar{y})\Big]^t, \tag{2.98}$$

where $\varphi_1(\bar{x}, \bar{y}) = a\cos\theta\bar{x} - b\sin\theta\bar{y} + t_1$ and $\varphi_2(\bar{x}, \bar{y}) = a\sin\theta\bar{x} + b\cos\theta\bar{y} + t_2$, with $\mathbf{t} = [t_1\ t_2]^t \in \mathbb{R}^2$ is the translation vector associated with $\phi$. We define the function $\hat{f} := f - I_T f$ which satisfies

$$\begin{aligned}|\hat{f}|_{W_p^1(T)} &\lesssim \Big(\int_T \Big(|D_x\hat{f}(x,y)|^2 + |D_y\hat{f}(x,y)|^2\Big)^{\frac{p}{2}} \mathrm{d}x\mathrm{d}y\Big)^{\frac{1}{p}} \\ &= \Big(\int_{\bar{T}} |ab|\Big(|(D_x\hat{f})\circ\varphi(\bar{x},\bar{y})|^2 + |(D_y\hat{f})\circ\varphi(\bar{x},\bar{y})|^2\Big)^{\frac{p}{2}} \mathrm{d}\bar{x}\mathrm{d}\bar{y}\Big)^{\frac{1}{p}}, \qquad (2.99)\end{aligned}$$

where $\bar{T} = \varphi^{-1}(T)$. Using the differentiation rules for the composite function $\hat{f}\circ\varphi$, whose variables are $\bar{x}, \bar{y}$, we have

$$\begin{cases} D_{\bar{x}}(\hat{f}\circ\varphi)(\bar{x},\bar{y}) &= (D_x\hat{f})\circ\varphi(\bar{x},\bar{y})\cdot a\cos\theta + (D_y\hat{f})\circ\varphi(\bar{x},\bar{y})\cdot a\sin\theta, \\ D_{\bar{y}}(\hat{f}\circ\varphi)(\bar{x},\bar{y}) &= -(D_x\hat{f})\circ\varphi(\bar{x},\bar{y})\cdot b\sin\theta + (D_y\hat{f})\circ\varphi(\bar{x},\bar{y})\cdot b\cos\theta, \end{cases}$$

or, equivalently,

$$\begin{cases} (D_x\hat{f})\circ\varphi(\bar{x},\bar{y}) &= D_{\bar{x}}(\hat{f}\circ\varphi)(\bar{x},\bar{y})\cdot b\cos\theta - D_{\bar{y}}(\hat{f}\circ\varphi)(\bar{x},\bar{y})\cdot a\sin\theta, \\ (D_y\hat{f})\circ\varphi(\bar{x},\bar{y}) &= D_{\bar{x}}(\hat{f}\circ\varphi)(\bar{x},\bar{y})\cdot b\sin\theta + D_{\bar{y}}(\hat{f}\circ\varphi)(\bar{x},\bar{y})\cdot a\cos\theta. \end{cases}$$

Recalling that $\hat{f} = f - I_T f$, substituting the above expressions into (2.99) yields

$$\begin{aligned}|f - I_T f|_{W_p^1(T)} &\lesssim \Big(\int_{\bar{T}} |ab|\Big(|D_{\bar{x}}(\hat{f}\circ\varphi)(\bar{x},\bar{y})|^2 b^2 + |D_{\bar{y}}(\hat{f}\circ\varphi)(\bar{x},\bar{y})|^2 a^2\Big)^{\frac{p}{2}} \mathrm{d}\bar{x}\mathrm{d}\bar{y}\Big)^{\frac{1}{p}} \\ &\lesssim \max\{|a|,|b|\}|ab|^{\frac{1}{p}}|\hat{f}\circ\varphi|_{W_p^1(\bar{T})}. \qquad (2.100)\end{aligned}$$

Since $\hat{f}\circ\varphi = (f\circ\varphi) - I_{\bar{T}}(f\circ\varphi)$, we can estimate $|\hat{f}\circ\varphi|_{W_p^1(\bar{T})}$ by using (2.20) with $(\bar{\sigma},\bar{\tau}) := (\sigma_{h_{\bar{T}}}, \sigma_{\rho_{\bar{T}}})$, and obtain from (2.100) that,

$$\begin{aligned}|f - I_T f|_{W_p^1(T)} \lesssim &\max\{|a|,|b|\}|ab|^{\frac{1}{p}}|(f\circ\varphi) - I_{\bar{T}}(f\circ\varphi)|_{W_p^1(\bar{T})} \\ \lesssim &\max\{|a|,|b|\}|ab|^{\frac{1}{p}}\Big(h_{\bar{T}}\|D^2_{\bar{\sigma}\bar{\sigma}}(f\circ\varphi)\|_{L_p(\bar{T})} \\ &+ h_{\bar{T}}\|D^2_{\bar{\sigma}\bar{\tau}}(f\circ\varphi)\|_{L_p(\bar{T})} + \rho_{\bar{T}}\|D^2_{\bar{\tau}\bar{\tau}}(f\circ\varphi)\|_{L_p(\bar{T})}\Big), \qquad (2.101)\end{aligned}$$

with constant depending on $\gamma(\bar{T})$. The norms $\|D^2_{\bar{\sigma}\bar{\sigma}}(f\circ\varphi)\|_{L_p(\bar{T})}$, $\|D^2_{\bar{\sigma}\bar{\tau}}(f\circ\varphi)\|_{L_p(\bar{T})}$ and $\|D^2_{\bar{\tau}\bar{\tau}}(f\circ\varphi)\|_{L_p(\bar{T})}$ need to be evaluated. First, we express $D_{\bar{\sigma}}(f\circ\varphi)$ and $D_{\bar{\tau}}(f\circ\varphi)$ as linear combinations of $D_{\bar{x}}(f\circ\varphi)$ and $D_{\bar{y}}(f\circ\varphi)$, that is, there is an angle $\bar{\theta}$ such that

$$\begin{aligned} D_{\bar{\sigma}}(f \circ \varphi) &= D_{\bar{x}}(f \circ \varphi) \cdot \cos \bar{\theta} + D_{\bar{y}}(f \circ \varphi) \cdot \sin \bar{\theta}, \\ D_{\bar{\tau}}(f \circ \varphi) &= -D_{\bar{x}}(f \circ \varphi) \cdot \sin \bar{\theta} + D_{\bar{y}}(f \circ \varphi) \cdot \cos \bar{\theta}. \end{aligned}$$

Recalling from (2.98) that $\varphi = (\varphi_1, \varphi_2)$, we find that

$$\begin{aligned} D_{\bar{x}}(f \circ \varphi)(\bar{x}, \bar{y}) &= \frac{\partial \varphi_1(\bar{x}, \bar{y})}{\partial \bar{x}}(D_x f) \circ \varphi(\bar{x}, \bar{y}) + \frac{\partial \varphi_2(\bar{x}, \bar{y})}{\partial \bar{x}}(D_y f) \circ \varphi(\bar{x}, \bar{y}) \\ &= a \cos \theta (D_x f) \circ \varphi(\bar{x}, \bar{y}) + a \sin \theta (D_y f) \circ \varphi(\bar{x}, \bar{y}), \\ D_{\bar{y}}(f \circ \varphi)(\bar{x}, \bar{y}) &= \frac{\partial \varphi_1(\bar{x}, \bar{y})}{\partial \bar{y}}(D_x f) \circ \varphi(\bar{x}, \bar{y}) + \frac{\partial \varphi_2(\bar{x}, \bar{y})}{\partial \bar{y}}(D_y f) \circ \varphi(\bar{x}, \bar{y}) \\ &= -b \sin \theta (D_x f) \circ \varphi(\bar{x}, \bar{y}) + b \cos \theta (D_y f) \circ \varphi(\bar{x}, \bar{y}). \end{aligned}$$

With $A_{11} = a \cos \theta \cos \bar{\theta} - b \sin \theta \sin \bar{\theta}$, $A_{12} = a \sin \theta \cos \bar{\theta} + b \cos \theta \sin \bar{\theta}$ and also $A_{21} = -(a \cos \theta \sin \bar{\theta} + b \sin \theta \cos \bar{\theta})$, $A_{22} = -a \sin \theta \sin \bar{\theta} + b \cos \theta \cos \bar{\theta}$,

$$\begin{aligned} D_{\bar{\sigma}}(f \circ \varphi) &= A_{11}(D_x f) \circ \varphi + A_{12}(D_y f) \circ \varphi, \\ D_{\bar{\tau}}(f \circ \varphi) &= A_{21}(D_x f) \circ \varphi + A_{22}(D_y f) \circ \varphi. \end{aligned}$$

It immediately follows that

$$\begin{aligned} D^2_{\bar{\sigma}\bar{\sigma}}(f \circ \varphi) =& D_{\bar{\sigma}}\Big(A_{11}(D_x f) \circ \varphi + A_{12}(D_y f) \circ \varphi\Big) \\ =& A_{11}\Big(A_{11}(D^2_{xx} f) \circ \varphi + A_{12}(D^2_{xy} f) \circ \varphi\Big) \\ & + A_{12}\Big(A_{11}(D^2_{xy} f) \circ \varphi + A_{12}(D^2_{yy} f) \circ \varphi\Big) \\ =& A_{11}^2 (D^2_{xx} f) \circ \varphi + 2A_{11}A_{12}(D^2_{xy} f) \circ \varphi + A_{12}^2 (D^2_{yy} f) \circ \varphi, \qquad (2.102) \end{aligned}$$

$$\begin{aligned}D^2_{\bar{\sigma}\bar{\tau}}(f\circ\varphi) =&D_{\bar{\sigma}}\Big(A_{21}(D_xf)\circ\varphi+A_{22}(D_yf)\circ\varphi\Big)\\ =&A_{21}\Big(A_{11}(D^2_{xx}f)\circ\varphi+A_{12}(D^2_{xy}f)\circ\varphi\Big)\\ &+A_{22}\Big(A_{11}(D^2_{xy}f)\circ\varphi+A_{12}(D^2_{yy}f)\circ\varphi\Big)\\ =&A_{11}A_{21}(D^2_{xx}f)\circ\varphi+(A_{21}A_{12}+A_{11}A_{22})(D^2_{xy}f)\circ\varphi\\ &+A_{12}A_{22}(D^2_{yy}f)\circ\varphi,\end{aligned}\tag{2.103}$$

$$\begin{aligned}D^2_{\bar{\tau}\bar{\tau}}(f\circ\varphi) =&D_{\bar{\tau}}\Big(A_{21}(D_xf)\circ\varphi+A_{22}(D_yf)\circ\varphi\Big)\\ =&A_{21}\Big(A_{21}(D^2_{xx}f)\circ\varphi+A_{22}(D^2_{xy}f)\circ\varphi\Big)\\ &+A_{22}\Big(A_{21}(D^2_{xy}f)\circ\varphi+A_{22}(D^2_{yy}f)\circ\varphi\Big)\\ =&A_{21}^2(D^2_{xx}f)\circ\varphi+2A_{21}A_{22}(D^2_{xy}f)\circ\varphi+A_{22}^2(D^2_{yy}f)\circ\varphi.\end{aligned}\tag{2.104}$$

Since $\max\{|A_{11}|,|A_{12}|,|A_{21}|,|A_{22}|\}\leq 2\max\{|a|,|b|\}$, we easily prove that

$$\begin{aligned}&\max\Big\{\|D^2_{\bar{\sigma}\bar{\sigma}}(f\circ\varphi)\|_{L_p(\bar{T})},\|D^2_{\bar{\sigma}\bar{\tau}}(f\circ\varphi)\|_{L_p(\bar{T})},\|D^2_{\bar{\tau}\bar{\tau}}(f\circ\varphi)\|_{L_p(\bar{T})}\Big\}\\ &\leq 16\max\{|a|,|b|\}^2\\ &\quad\max\Big\{\|(D^2_{xx}f)\circ\varphi\|_{L_p(\bar{T})},\|(D^2_{xy}f)\circ\varphi\|_{L_p(\bar{T})},\|(D^2_{yy}f)\circ\varphi\|_{L_p(\bar{T})}\Big\}\\ &=16\max\{|a|,|b|\}^2|ab|^{-\frac{1}{p}}\max\Big\{\|D^2_{xx}f\|_{L_p(T)},\|D^2_{xy}f\|_{L_p(T)},\|D^2_{yy}f\|_{L_p(T)}\Big\}\\ &\leq 16\max\{|a|,|b|\}^2|ab|^{-\frac{1}{p}}|f|_{W^2_p(T)}.\end{aligned}$$

The estimation in (2.101) now reads

$$|f-I_Tf|_{W^1_p(T)}\lesssim\max\{|a|,|b|\}^3h_{\bar{T}}|f|_{W^2_p(T)}.$$

We are now left to estimate $h_{\bar{T}}$. By virtue of the fact that translation vectors do not change the edge-vectors of a triangle, for any edge-vector $\mathbf{e}$ of $T$, we observe that $\varphi^{-1}(\mathbf{e})=\frac{1}{ab}\begin{bmatrix}1/a & 0\\ 0 & 1/b\end{bmatrix}U^{-1}(\mathbf{e})$. We deduce that $\|\varphi^{-1}(\mathbf{e})\|\leq$

$\frac{1}{|ab|}\max\left\{\frac{1}{|a|},\frac{1}{|b|}\right\}\|\mathbf{e}\|$. Hence

$$|f - I_T f|_{W_p^1(T)} \lesssim \frac{\max\{|a|,|b|\}}{\min\{|a|,|b|\}}\frac{\max\{|a|,|b|\}^2}{|ab|}h_T|f|_{W_p^2(T)}.$$

Combining this with the above estimation yields the result in (2.97). □

In the case where the interior angles of $T$ are far from $\pi$, the map $\varphi$ is the identity function so that $a = b = 1$ and thus

$$|f - I_T f|_{W_p^1(T)} \lesssim h|f|_{W_p^2(T)},$$

which is obtainable from (2.20).

In the example below, we show what happens when we use the map $\varphi$ given by (2.8).

**Example 2.6.1.** Let $T$ be a fixed triangle. Given a triangle $T_0$, there always exists an affine map $\varphi$ such that $\varphi(T_0) = T$. Choosing $T_0$ as the reference triangle $\hat{T}$ in Figure 2.2, the affine map $\varphi$ is given by (2.8). For the associated matrix $M$ given in (2.9), its singular values are given by $\sqrt{\Lambda_1}$, $\sqrt{\Lambda_2}$ in (2.24) and (2.25), they satisfy $\sqrt{\Lambda_1} \sim h$ and $\sqrt{\Lambda_2} \sim \rho$ where $h$ is the diameter of $T$ and $\rho$ its smallest height. Since $\hat{T}$ is fixed, the constant depending on its maximum interior angle as shown in (2.97) becomes an absolute constant. Thus

$$|f - I_T f|_{W_p^1(T)} \leq C\frac{\Lambda_1}{\Lambda_2}h|f|_{W_p^2(T)} \leq C'\frac{h^3}{\rho^2}|f|_{W_p^2(T)}, \tag{2.105}$$

where $C$ and $C'$ are absolute constants. This is weaker than the standard estimation in (2.18).

### 2.6.2 Using quadratic polynomials

Given a triangle $T \subset \Omega$ and a function $f \in C^2(T)$, our aim is to estimate the $W_p^1$-seminorm of the error $f - I_T f$ by using homogeneous quadratic polynomials as shown in (2.106) below, instead of ensuring that the maximum angle $\gamma(T)$

is far from the angle $\pi$. The methods presented here can be applied when the Hessian of $f$ does not vary much on $T$ (e.g. if $f$ satisfies (2.88)), that is, the Hessian $H_f$ can be represented by a homogeneous quadratic polynomial in $\mathbb{H}_2$.

Given a homogeneous quadratic polynomial $\pi \in \mathbb{H}_2$,

$$f - I_T f = (f - \pi) - I_T(f - \pi) + \pi - I_T \pi. \tag{2.106}$$

By using (2.18), there is a constant $C$ such that

$$|(f - \pi) - I_T(f - \pi)|_{W_p^1(T)} \leq C \frac{h^2}{\rho} |f - \pi|_{W_p^2(T)}, \tag{2.107}$$

with $h$ and $\rho$ being the length scales of $T$. Since we assume that $T$ is an arbitrary triangle which is not nearly-optimal, the ratio $\frac{h}{\rho}$ can be unbounded.

We now use Proposition 2.3.5 to obtain that

$$|\pi - I_T \pi|_{W_p^1(T)} \leq C_2 \rho_\pi(T) h |T|^{\frac{1}{p}} \sqrt{|\det \pi|}, \tag{2.108}$$

for some constant $C_2$, where the measure of non-degeneracy $\rho_\pi$ is defined in (2.39). For convenience, we denote by $C$ the maximum between $C$ and $C_2$, so that combining (2.107) and (2.108) yields

$$\begin{aligned} |f - I_T f|_{W_p^1(T)} &\leq C \Big( \frac{h}{\rho} |f - \pi|_{W_\infty^2(T)} + \rho_\pi(T) \sqrt{|\det \pi|} \Big) h |T|^{\frac{1}{p}} \\ &= C \Big( \frac{h}{\rho} \max_{z' \in T} \|\pi_{z'} - \pi\| + \rho_\pi(T) \sqrt{|\det \pi|} \Big) h |T|^{\frac{1}{p}} \end{aligned} \tag{2.109}$$

with $\|\pi\|$ denoting the maximum coefficient of $\pi$. The result below is obtained by choosing $\pi = \pi_z$ for some point $z \in T$.

**Proposition 2.6.2.** *Let $z \in T$ be any point of a triangle $T$. Given a function $f \in C^2(T)$, there is an absolute constant $C$ such that*

$$|f - I_T f|_{W_p^1(T)} \leq C \Big( \frac{h}{\rho} \omega(h) + \rho_{\pi_z}(T) \sqrt{|\det \pi_z|} \Big) h |T|^{\frac{1}{p}}. \tag{2.110}$$

*Proof.* Noting that

$$\max_{z'\in T} \|\pi_{z'} - \pi_z\| \le \omega(\max_{z'\in T} |z' - z|) \le \omega(h), \tag{2.111}$$

the result in (2.110) follows from (2.109). □

Similar to the estimation $|f - I_T f|_{W_p^1(T)} \le C\frac{h^2}{\rho}|f|_{W_p^2(T)}$ obtainable from (2.18), the ratio $\frac{h}{\rho}\omega(h)$ in (2.110) can be problematic when the smallest height $\rho$ is small. However it does give $O(h)$ estimates for moderately anisotropic triangles for which $h\omega(h) \le C_1\rho$ and $\rho_{\pi_z}(T) \le C_2$. For example when $f \in C_p^3(T)$ the mean value theorem for the second derivatives of $f$ yields

$$\|\pi_z - \pi_{z'}\| \le |z - z'||f|_{W_\infty^3(T)}, \quad z, z' \in T.$$

This implies that for all $r > 0$, we have $\omega(r) \le r|f|_{W_\infty^3(T)}$. Hence $\frac{h}{\rho}\omega(h) \le \frac{h^2}{\rho}|f|_{W_\infty^3(T)}$ and $O(h)$ estimate holds when $\frac{h^2}{\rho}$ is bounded which allows aspect ratio up to $\frac{h}{\rho} \sim \frac{h}{h^2} = \frac{1}{h}$.

In the example below, we illustrate (2.110) in the setting of Example 2.3.2.

**Example 2.6.2.** Suppose that $\pi_z(x, y) = ax^2 + by^2$ at a point $z \in A$. We use the same settings as in Example 2.3.2 but with $\pi = \pi_z$. In addition, suppose that there is a constant $C_0$ so that $\omega(h) \le C_0|\lambda_1| = C_0|a|$ holds, with $|\lambda_1| \ne 0$ being the smallest (in absolute value) eigenvalues of $Q_{\pi_z}$. Thus

$$\frac{h}{\rho}\omega(h) \lesssim \left|\frac{b}{a}\right|^{\frac{1}{2}}|a| = |ab|^{\frac{1}{2}}.$$

By using Proposition 2.6.2, with $h \sim \left|\frac{b}{a}\right|^{\frac{1}{4}}$,

$$|f - I_T f|_{W_p^1(T)} \lesssim |ab|^{\frac{1}{2}}\left|\frac{b}{a}\right|^{\frac{1}{4}}|T|^{\frac{1}{p}} \lesssim |a|^{\frac{1}{4}}|b|^{\frac{3}{4}}|T|^{\frac{1}{p}},$$

the constant in the inequalities are absolute. For comparison, the estimation from

(2.18) gives

$$|f - I_T f|_{W_p^1(T)} \leq C\frac{h^2}{\rho}|f|_{W_p^2(T)} \lesssim \Big|\frac{b}{a}\Big|^{\frac{3}{4}}|f|_{W_p^2(T)},$$

which is unbounded as $a \to 0$. Moreover, the estimation in (2.87) cannot be applied when $a$ is very small and $b$ large in which case the triangle $T$ presents a big interior angle, although the expected bound shows

$$|f - I_T f|_{W_p^1(T)} \lesssim \Big|\frac{b}{a}\Big|^{\frac{1}{4}}|D_x^2 f|_{W_p^1(T)} + \Big|\frac{a}{b}\Big|^{\frac{1}{4}}|D_y^2 f|_{W_p^1(T)},$$

with constant depending on the maximum interior angle of $T$.

# Chapter 3

# Asymptotically optimal interpolation by piecewise linear polynomials

In this chapter, the target function $f$ is approximated on a square domain $\Omega \subset \mathbb{R}^2$ which we triangulate according to the properties of $f$, that is, by using the eigenvalues and eigenvectors of the Hessian $H_f$ at some pre-selected points. We assume that $f \in C^2(\Omega)$ is strictly convex (or concave) on $\Omega$.

It is known that the approximation order in $L_p$-norm ($p \in [1, \infty]$) of a piecewise linear approximation on a triangulation $\Delta_N$ cannot be better than $O(N^{-1})$ (see [16]), where $N$ is the bound on the number of triangles. This order can easily be achieved (by using the estimation (2.18)), for instance, on uniform triangulations where the diameter $h_N := \max_{T \in \Delta_N} h_T$ satisfies $h_N^2 \sim N^{-1}$,

$$\|f - f_N\|_{L_p(\Omega)} \leq CN^{-1} \|H_f\|_{L_p(\Omega)}, \tag{3.1}$$

where $f_N$ is the approximant of $f$ by using $\Delta_N$, and where $\|H_f\|_{L_p(\Omega)} := |f|_{W_p^2(\Omega)}$. The improvement of the estimation with respect to the $L_p$-norm on the right hand side of (3.1) has been addressed by many papers. For instance, in [12], the estimation below is achieved for $f \in C^2(\Omega)$,

$$\|f - f_N\|_{L_p(\Omega)} \leq CN^{-\frac{2}{d}} \| \sqrt[d]{\det \mathcal{H}_f} \|_{L_\sigma(\Omega)}, \tag{3.2}$$

where $\Omega \subset \mathbb{R}^d$, $d \geq 2$, and $C$ is independent of $f$, with $\frac{1}{\sigma} = \frac{2}{d} + \frac{1}{p}$, and $\mathcal{H}_f$ denoting a *majorizing* Hessian matrix for $f$ which is the same as the Hessian matrix $H_f$ if $f$ is a strictly convex function. In the case where $d = 2$ and $f$ is a strictly convex function, clearly, (3.2) improves on (3.1) by virtue of

$$\| \sqrt{\det H_f} \|_{L_q(\Omega)} \leq \| \sqrt{\det H_f} \|_{L_p(\Omega)} \leq \|H_f\|_{L_p(\Omega)},$$

where $\frac{1}{q} = 1 + \frac{1}{p}$. However, the expression of the constant $C$ in (3.2) is generally unknown. For $d = 2$, several papers considered the question of designing triangulations where the error of interpolation has the best possible asymptotic constant in the sense of lim sup. For strictly convex functions, the smallest possible constant and its exact expression has been found in [3]: Amongst all triangulations $\Delta_N$ of at most $N$ triangles, it is proved that

$$\lim_{N \to \infty} N \Big( \inf_{\Delta_N} \|f - f_N\|_{L_p(\Omega)} \Big) = \frac{C_p^+}{2} \| \sqrt{\det H_f} \|_{L_q(\Omega)}, \tag{3.3}$$

where $C_p^+$ is the value of the $L_p$-norm best approximation of the polynomial $\pi_0(x, y) = x^2 + y^2$ over all equilateral triangles of unit area. The case $p = \infty$ is treated in [2] where slight changes occur on the right hand side of (3.3). Namely, $q$ becomes 1 and $C_p^+$ becomes a constant equivalent to $1 + o(1)$ as $N \to \infty$.

The extension of (3.3), for the case where the approximant[1] $f_{m,N}$ is a piecewise polynomial of order $m - 1$, with $m \geq 2$, and where the function $f \in C^m(\Omega)$ is not necessarily convex, has been developed in [29] where the estimation is of the form

$$\limsup_{N \to \infty} N^{\frac{m}{2}} \|f - f_{m,N}\|_{L_p(\Omega)} \leq \left\| K_{m,p}\Big(\frac{d^m f}{m}\Big) \right\|_{L_\varrho(\Omega)}, \tag{3.4}$$

[1] Interpolating the target function at specific points of the triangulation by using barycentric coordinates.

where $\frac{1}{\varrho} = \frac{m}{2} + \frac{1}{p}$, the notation $d^m f$ denotes the $m$-th derivative of $f$, and where $K_{m,p}(\pi)$ is the value of the $L_p$-norm best approximation over all triangles of unit area of a homogeneous polynomial $\pi$ of degree $m$. Note that (3.4) is *optimal* in the sense that its right hand side is a lower bound for a large class of *admissible* triangulations. Note also that in the case $m = 2$, (3.4) coincides with (3.3). Another extension, but with a different triangulation method (although still using the patching strategy described in Section 3.1.1 below), is developed in [30] for the $W_p^1$-seminorm of the error,

$$\limsup_{N\to\infty} N^{\frac{m-1}{2}} |f - f_{m,N}|_{W_p^1(\Omega)} \leq \left\| L_{m,p}\Big(\frac{d^m f}{m}\Big) \right\|_{L_\tau(\Omega)}, \tag{3.5}$$

where $\frac{1}{\tau} = \frac{m-1}{2} + \frac{1}{p}$, and where $L_{m,p}(\pi)$ is the value of the $W_p^1$-seminorm best approximation over all triangles of unit area of a homogeneous polynomial $\pi$ of degree $m$. It is proved that (3.5) cannot be further improved.

In [2, 3], it is discussed that the problem of approximating a convex function $f \in C^2(\Omega)$ by piecewise linear polynomials is related to the problem of approximating convex bodies in $\mathbb{R}^3$ by inscribed polytopes. That is, designing a triangulation is equivalent to inscribing a polytope. We refer the reader to [6, 7, 25] for more details on this problem.

In [14, 31], the triangulation method is based on the so-called *greedy algorithm* which iteratively constructs a sequence of nested triangulations $(\Delta_N)_{N\geq N_0}$ from a given triangulation $\Delta_{N_0}$. The general idea consists in obtaining triangulations which *equidistribute* the local errors between triangles. A triangle is then bisected if it gives a local error greater than a prescribed tolerance. This, however, results in a non-conforming triangulation. Nevertheless, for strictly convex functions, the estimation (3.2) is also achieved (with $d = 2$), where the constant $C$ has no exact expression.

We refer to [5] for a review of the rich literature on mesh generations and optimal triangulations aimed at solving partial differential equations. In many papers, the optimal triangulation is characterized by the *metric* induced by the

Hessian [11, 12, 14, 31, 32]. The triangles of the triangulation are required to have regular (or isotropic) shapes with respect to the metric used. The problem of whether or not such a characterization can be applied to obtain asymptotically optimal triangulations is still open.

It is also an open question whether both optimal estimates (3.4) and (3.5) can be achieved simultaneously on the same sequence of triangulations. In this regard, the purpose of this chapter is to design a sequence $(\Delta_N)_{N\geq N_0}$ of anisotropic triangulations (see Section 3.2) on which the optimal result in (3.3) is achieved, and such that an asymptotic error estimation in $W_p^1$-seminorm of optimal order $O(N^{-1/2})$ can be derived. For any strictly convex $f \in C^2(\Omega)$, we obtain the asymptotic estimations

$$\limsup_{N\to\infty} N\|f - f_N\|_{L_p(\Omega)} \leq \left\|K_p\Big(\frac{d^2 f}{2}\Big)\right\|_{L_q(\Omega)}, \tag{3.6}$$

$$\limsup_{N\to\infty}\; N^{\frac{1}{2}}|f - f_N|_{W_p^1(\Omega)} \leq C_p|f|^{\frac{1}{2}}_{W_p^2(\Omega)}\left\|K_p\Big(\frac{d^2 f}{2}\Big)\right\|^{\frac{1}{2}}_{L_q(\Omega)}, \tag{3.7}$$

where $1 \leq p < \infty$, $\frac{1}{q} = 1+\frac{1}{p}$, $C_p$ is a constant depending on $p$, and $K_p = K_{2,p}$. The asymptotic estimation in (3.6) is exactly the same as in (3.4) for $m = 2$. However, the estimation in (3.7) is the first in $W_p^1$-seminorm estimation to be obtained when using a sequence of triangulations which are designed to be optimal with respect to the $L_p$-norm.

The triangulation method which we present can be divided into two main tasks: The first task consists in obtaining the so-called *regular regions*, a similar method as in [2, 3, 29, 30]. Obtaining regular regions consists in grouping triangles that fit into initially prescribed sub-squares of $\Omega$. In our approach, the triangles contained in a regular region are designed to be isosceles and such that their directions of alignments, which are the directions of the eigenvalues of $H_f$ at the centers of the subs-squares, are well-conditioned. The triangles obtained from regular regions are slightly modified optimal triangles and thus local errors are easy to estimate. The second task, which is much more difficult and more delicate, consists in obtaining the so-called *irregular* polygons by extending the segments defining the regular regions (see Section 3.2.3). The triangles obtained from irregular polygons can have various shapes, and we use the results in Section 2.6 to estimate the local errors. Various properties of our triangulation are provided in Section 3.3.

The chapter is organized as follows. In Section 3.1, we review the patching strategy and discuss the optimality of (3.3) and (3.4). In Section 3.2, we present our construction method in order to produce a sequence of optimal triangulations. It involves designing regular and irregular regions mentioned above. The properties of the constructed triangulations are discussed in Section 3.3. They include the estimation of the longest edge of an irregular triangle, the area covered by irregular regions, and the interior angles of the irregular regions after the so-called *back transformation*. These properties are essential for the analytical proof of our asymptotic error estimations in Section 3.4, in both $L_p$-norm and $W_p^1$-seminorm, as shown in (3.6) and (3.7). We conclude the chapter with a numerical illustration in Section 3.5.

## 3.1 Background

Let $\Omega \subset \mathbb{R}^2$ be a bounded domain. In order to obtain (3.6), the anisotropic triangulation $\Delta_N$ of $\Omega$ is constructed according to the properties of $f$. We present here the principal ideas, as found in [2, 3] and [29] but using our settings, for the construction of the anisotropic triangulation $\Delta_N$. For simplicity, the domain $\Omega$ is assumed to be a square.

### 3.1.1 Triangulation by patching strategy

The domain $\Omega$ is divided into $m^2$ sub-squares $S_i$, $i = 1, \ldots, m^2$ of side length $r > 0$. The parameter $r$ is chosen small enough so that the second derivatives of $f$ do not significantly change on each $S_i$ (e.g. by using (2.88) with $d = \sqrt{2}r$). On each sub-square $S_i$, the function $f$ is replaced by the quadratic polynomial

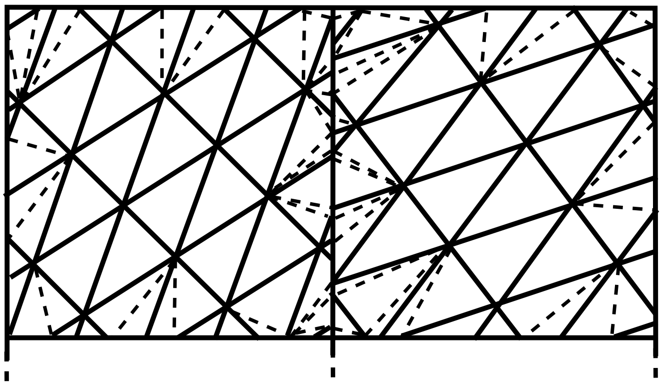


Figure 3.1: Regular and irregular triangles obtained by patching strategy.

$\pi_{b_i}$, where $b_i$ is the barycenter of $S_i$. Recall that an optimal triangle for $\pi_{b_i}$ is a triangle on which the infimum

$$\inf_{|T|=1} \|\pi_{b_i} - I_T \pi_{b_i}\|_{L_p(T)}, \tag{3.8}$$

is attained. The local error analysis developed in Chapter 2 is applied.

The first step consists in partitioning each sub-square $S_i$, $i = 1, \ldots, m^2$ into polygons by copying and aligning side by side into $S_i$ the scaled versions of an optimal triangle for $\pi_{b_i}$. The areas covered by the copies of the scaled optimal triangle are called *regular* regions, and the remaining non-covered areas are called the *irregular* regions. One of the principal goals is to ensure that the area covered by irregular regions is significantly smaller than the area covered by regular regions.

Naturally, the shapes and directions of the triangles inside the regular regions are chosen depending on the Hessian $H_f$ of the function. The irregular regions are then partitioned by some method. For instance, in [29] it is suggested to use the Delaunay triangulation, (see Figure 3.1). The emerging *irregular* triangles can have arbitrary shapes and this is one of the reasons that deriving $W_p^1$-seminorm estimations are problematic: The diameters of the irregular triangles are easily bounded and thus $L_p$-norm estimations of the error can be obtained, however there is no guarantee for the aspect ratios of the irregular triangles to be bounded, nor that their interior angles should be far from the flat angle.

In [29], the study of $L_p$-norm asymptotic estimations does not depend on the knowledge of the optimal triangles' shapes, and it is only assumed that $f \in C^2(\Omega)$. However, as shown in [3], when the Hessian $H_f$ is positive definite (which is also our case in Section 3.2), it is known that (3.8) is attained on triangles that are stretched and aligned in the directions of the eigenvectors of the matrix associated with $\pi_{b_i}$, thereby allowing for more practical triangulation algorithms to be performed. Such a favorable attribute is difficult to achieve in the case of indefinite $H_f$ since the shapes of triangles satisfying (3.8) are not fully known (see Section 2.4.2).

### 3.1.2 Optimality

The optimality of (3.6) means that equality occurs when using a certain family of triangulations $(\Delta_N)_{N \geq N_0}$. In [29], the right hand side of (3.6) is proved to be a lower bound when using a family of triangulations $(\Delta_N)_{N \geq N_0}$, termed *admissible*, satisfying the condition

$$\sup_{T \in \Delta_N} h_T^2 \leq C N^{-1}, \tag{3.9}$$

where $C$ is a constant independent of $N$, and recalling that $h_T$ denotes the diameter of the triangle $T$. Clearly the mesh size of a triangulation $\Delta_N$ belonging to such family goes to zero as $N$ grows. The condition in (3.9) is piqued by the estimation of the difference of errors, for each $T \in \Delta_N$,

$$|e_T(\pi_z) - e_T(f)| \leq 2|T|^{\frac{1}{p}} h_T^2 \omega(h_T), \quad z \in T,$$

with $\omega$ defined in (2.75). The above estimation is the pillar to proving the lower asymptotic estimation which is obtained by summing up the errors over all triangles in the triangulation.

By a different method, the optimality of (3.6) is proved in [3] by using a family of triangulations $(\Delta_N)_{N\geq N_0}$ satisfying, for a given $\varepsilon > 0$,

$$\sum_{T\in I_N(\varepsilon)\cup J_N(\varepsilon)} |T| < \varepsilon, \tag{3.10}$$

where the subsets $I_N(\varepsilon)$ and $J_N(\varepsilon)$ of $\Delta_N$ are defined by

$$I_N(\varepsilon) := \left\{T : \tfrac{h_T}{\rho_T} \geq \tfrac{C_p^+\varepsilon}{16\omega(h_T)}\sqrt{|\det \pi_{z_T}|}\right\} \text{ and } J_N(\varepsilon) := \left\{T : h_T \geq \tfrac{\varepsilon}{N^{\frac{1}{4}}}\right\},$$

with $C_p^+ = \inf_{|T|=1} \|\pi_0 - I_T\pi_0\|_{L_p(T)}$ where $\pi_0(x,y) = x^2 + y^2$, and with $z_T$ being a point on the longest edge of $T$. The condition in (3.10) shows that the area covered by triangles whose aspect ratios or diameters are uncontrollable is small.

Note that (3.9) and (3.10) are not necessarily linked to one another. In the example below, we show that a uniform triangulation can satisfy both of the conditions (3.9) and (3.10).

**Example 3.1.1.** Consider a family of uniform triangulations $(\Delta_N)_{N\geq 2}$, where each $\Delta_N$ is described as follows: Divide $\Omega$ into $m^2$ sub-squares of side length $r > 0$, then divide each sub-square by its diagonal parallel to the vector $[1\ 1]^t$. Then the family of triangulations $(\Delta_N)_{N\geq 2}$ satisfies (3.9). Indeed, since $|\Omega| = \sum_{T\in\Delta_N} |T| = N\frac{r^2}{2}$ and for each triangle $T \in \Delta_N$ we have $h_T = \sqrt{2}r$, clearly

$$\sup_{T\in\Delta_N} h_T^2 = 4|\Omega|N^{-1}.$$

The aspect ratio of each triangle $T \in \Delta_N$ satisfies $\frac{h_T}{\rho_T} = 2$, thus given $\varepsilon > 0$ clearly $\frac{h_T}{\rho_T} = 2 \geq \frac{C_p^+\varepsilon}{16\omega(h_T)}\sqrt{|\det \pi_{z_T}|}$ holds whenever $\varepsilon$ satisfies

$$\varepsilon \leq \frac{32\omega(h_T)}{C_p^+\sqrt{|\det \pi_{z_T}|}}. \tag{3.11}$$

The term $\omega(h_T) = \omega(2\sqrt{|\Omega|}N^{-\frac{1}{2}})$ which depends on the Hessian $H_f$ does not necessarily have an explicit formula. In the case where $f$ is a quadratic polynomial of the form $f(x, y) = x^2 + y^2$, $\det \pi_z = 1$ for all $z \in \Omega$, and since for any $z' \in \Omega$,

$$\|\pi_z - \pi_{z'}\| = \max\{|D^2_{xx}f(z) - D^2_{xx}f(z')|, |D^2_{yy}f(z) - D^2_{yy}f(z')|\} = 0,$$

clearly $\omega(r) = 0$ for all $r$ and thus (3.11) can not hold for any $\varepsilon > 0$, that is, $I_N(\varepsilon) = \emptyset$.

On the other hand, the condition for a triangle $T$ to be in $J_N(\varepsilon)$ is $h_T = 2\sqrt{|\Omega|}N^{-\frac{1}{2}} \geq \varepsilon N^{-\frac{1}{4}}$, that is, if $\varepsilon$ satisfies,

$$\varepsilon \leq 2\sqrt{|\Omega|}N^{-\frac{1}{4}}. \tag{3.12}$$

Letting $N_\varepsilon = \lceil \frac{16|\Omega|^2}{\varepsilon^4} \rceil$, for any $N > N_\varepsilon$ we have $\varepsilon > 2\sqrt{|\Omega|}N^{-\frac{1}{4}}$, so that (3.12) does not hold, that is, $J_N(\varepsilon) = \emptyset$ independently of the properties of $H_f$. If $f$ is the quadratic polynomial mentioned above, then $(\Delta_N)_{N\geq 2}$ clearly satisfies (3.10) for any $\varepsilon > 0$.

It is well known that isotropic triangulations (including the uniform ones) do not necessarily provide the sharpest asymptotic estimations. In this regard, anisotropic triangulations find interest in our study and we shall now present our triangulation method.

## 3.2 Triangulation of the domain

In this section, we present a novel technique of anisotropic triangulation. The initial procedures are similar to those found in [2, 3, 29, 30] in order to obtain regular regions. However, the triangles that form these regions will be isosceles. Also, instead of using the patching strategy (see Section 3.1.1) that connects the vertices of non-regular regions, the vertices of the regular regions will be connected to other ones by extending the segments that define the regular regions.

The square domain $\Omega$ is divided into $m^2$ auxiliary sub-squares $S_i$, $i = 1, \ldots, m^2$, of side length $r > 0$. For each $i = 1, \ldots, m^2$, we denote by $b_i$ the barycenter of the sub-square $S_i$. As before, the parameter $r$ is chosen small enough so that the second derivatives of $f$ do not significantly change on each sub-square. The regular region contained in $S_i$ is described in Section 3.2.2 by using the shifts of a scaled version of an optimal triangle $T_i$ for $\pi_{b_i}$. The characteristics of each $T_i$ are described in the sections below.

We denote by $\lambda_{1,i}, \lambda_{2,i}$ the eigenvalues of the matrix $Q_{\pi_{b_i}}$, with $i \in \{1, \dots, m^2\}$, described as in (2.1) and (2.2). We assume that there is a constant $\delta_f \in (0, 1)$ such that, for all $i \in \{1, \dots, m^2\}$,

$$\delta_f \leq |\lambda_{1,i}|. \tag{3.13}$$

The above condition simply means that $f$ is strictly convex. Such an assumption can be found in [2, 3, 31], but not in [29, 30] where the results hold for any smooth function $f \in C^2(\Omega)$.

We also use the following assumption (same as in (2.88) with $\nu = \omega(\sqrt{2}r)$): Given a pair of orthonormal vectors $(\boldsymbol{\sigma}, \boldsymbol{\tau})$, for any $z \leq \mathcal{B}(b_k, \sqrt{2}r)$, $k = 1, \dots, m^2$,

$$|D^2_{\mathbf{ij}} f(b_k) - D^2_{\mathbf{ij}} f(z)| \leq \omega(\sqrt{2}r), \tag{3.14}$$

with $\mathbf{i}, \mathbf{j} \in \{\boldsymbol{\sigma}, \boldsymbol{\tau}\}$, and where $\omega$ is defined in (2.75). Observe that the radius $\sqrt{2}r$ of the ball $\mathcal{B}(b_k, \sqrt{2}r)$ is the maximum distance of two neighboring barycenters. By choosing $r$ to be small enough, we can ensure that the second derivatives of $f$ are still close to one another (or nearly constant) between neighboring barycenters.

The parameter $r$ is chosen to be sufficiently small in such a way that

$$\omega(\sqrt{2}r) \leq \left( \frac{\mathcal{K}_f \delta_f}{\max_{z \in \Omega} \|H_f(z)\|_2} \right)^2, \tag{3.15}$$

where $C_{\delta_f} = 33\pi + 2\delta_f^{-1/2}$ and $\mathcal{K}_f = \left(10^5 \frac{3}{2} C^2_{\delta_f} |f|^{\frac{1}{2}}_{W^2_\infty(\Omega)}\right)^{-1}$. The presence of the two constants will be justified later. We can relate the above condition to the eigenvalues of $Q_{\pi_i}$, $i = 1, \dots, m^2$: For each $i = 1, \dots, m^2$, we combine (3.13) with the fact that $|\lambda_{2,i}| \leq \max_{z\in\Omega} \|H_f(z)\|_2$ to deduce that

$$\frac{\delta_f}{\max_{z\in\Omega} \|H_f(z)\|_2} \leq \min_i \Big|\frac{\lambda_{1,i}}{\lambda_{2,i}}\Big|, \tag{3.16}$$

and thus a consequence of (3.15) is that

$$\omega(\sqrt{2}r) \leq \min_{i\in\{1,\dots,m^2\}} \Big\{\mathcal{K}_f^2 \Big|\frac{\lambda_{1,i}}{\lambda_{2,i}}\Big|^2\Big\}. \tag{3.17}$$

It is essential that $r$ is chosen small enough as we shall see in Section 3.4 where we derive asymptotic error estimations in $L_p$-norm and $W_p^1$-seminorm.

### 3.2.1 Conditioning angles of rotation

Recall from the previous chapter that we use homogeneous quadratic polynomials as intermediate steps in order to estimate local errors. In fact, as we shall see later, we shall instead use slightly perturbed homogeneous quadratic polynomials.

Given a non-degenerate polynomial $\pi \in \mathbb{H}_2$, the eigenvectors of $U_\pi = R_\mu$ as described in (2.2) provide the alignment directions of an optimal triangle $T \in \Delta_\pi$. This is clear from Lemma 2.4.5 and the properties of the linear map $\phi_\pi$ as defined in (2.42). Unfortunately, these eigenvectors are ill-conditioned when the eigenvalues of $Q_\pi$ are close to one another (see [20, 35]). That is to say, for a matrix $\bar{Q}_\pi$ resulting from small perturbations in the coefficients of $Q_\pi$ (or those of $\pi$), the alignment directions of its eigenvalues can be abruptly altered from those of $Q_\pi$.

In the example below, we illustrate the effect of close eigenvalues that causes ill-conditioning to the corresponding eigenvectors:

**Example 3.2.1.** Consider the function $f(x, y) = x^2 + (1 + \delta x)y^2$, $\delta \geq 0$, whose Hessian is given by $H_f(x, y) = \begin{bmatrix} 2 & 2\delta y \\ 2\delta y & 2(1+\delta x)\end{bmatrix}$. For $x_0 \in \mathbb{R}$, the eigenvectors of $H_f(x_0, 0) = \begin{bmatrix} 2 & 0 \\ 0 & 2(1+\delta x_0)\end{bmatrix}$ are exactly the unit vectors $\mathbf{e}_1 = [1\ \ 0]^t$ and $\mathbf{e}_2 = [0\ \ 1]^t$. These eigenvectors are fixed whenever $y = 0$. For $y_0 > 0$, the

Hessian given by $H_f(0, y_0) = \begin{bmatrix} 2 & 2\delta y_0 \\ 2\delta y_0 & 2 \end{bmatrix}$ has eigenvalues $\lambda_1 = 2 - 2\delta y_0$ and $\lambda_2 = 2 + 2\delta y_0$. For each $i = 1, 2$, the eigenvector $\mathbf{v}_i = [x_i \; y_i]^t$ corresponding to $\lambda_i$ satisfies $2x_i + 2\delta y_0 y_i = \lambda_i x_i$ which yields, for any $\alpha \neq 0$, $[x_i \; y_i] = \begin{bmatrix} \alpha & \frac{(\lambda_i - 2)\alpha}{2\delta y_0} \end{bmatrix}$. It is easily shown that

$$\mathbf{v}_1 = [1 \; 1]^t \;\text{ and }\; \mathbf{v}_2 = [-1 \;\; 1]^t.$$

If $\delta$ is small, we observe the abrupt change of eigenvectors in the neighborhood of the point $(0, 0)$, which shows the ill-conditioning of eigenvectors when eigenvalues are close to one another.

Such big variations can cause an extensive disadvantage in the quality of the triangulations constructed in Section 3.2 if the alignment directions of the optimal triangles $T_i$, $i = 1, \dots, m^2$ remain the same as those of the eigenvectors of $Q_{\pi_{b_i}}$. To prevent this issue, we introduce *adjustment angles* for quadratic polynomials whose eigenvalues are close to one another.

Consider a non-degenerate polynomial $\pi \in \mathbb{H}_2$ such that the difference of its eigenvalues satisfy $|\lambda_2 - \lambda_1| \leq \varepsilon$, for some small number $\varepsilon > 0$. Let $\mathbf{v}_1, \mathbf{v}_2$ denote the eigenvectors associated with $\lambda_1, \lambda_2$, respectively. For any angle $\vartheta \in [0, 2\pi]$, with $R_{\mu-\vartheta} = R_\mu \circ R_{-\vartheta}$, the eigenvalues of the matrix

$$\bar{Q}_\pi = R_{\mu-\vartheta} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} R^t_{\mu-\vartheta}, \tag{3.18}$$

are exactly $\lambda_1$ and $\lambda_2$, with easily provable corresponding eigenvectors $\bar{\mathbf{v}}_1 = R_{-\vartheta}\mathbf{v}_1$ and $\bar{\mathbf{v}}_2 = R_{-\vartheta}\mathbf{v}_2$ due to the commutativity $R_\mu \circ R_{-\vartheta} = R_{-\vartheta} \circ R_\mu$ from which it holds that

$$\bar{Q}_\pi \bar{\mathbf{v}}_i = \left( R_{-\vartheta} U_\pi \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} U^t_\pi R_\vartheta \right) R_{-\vartheta}\mathbf{v}_i = \lambda_i R_{-\vartheta}\mathbf{v}_i = \lambda_i \bar{\mathbf{v}}_i, \quad i = 1, 2.$$

The adjustment angle $\vartheta$ defined in Proposition 3.2.4 is chosen such that the angle $\mu - \vartheta$ is well conditioned in the sense of (3.28), where $\mu = \mu(\pi) \in [0, \pi)$ is as described in (2.2) and in Figure 2.1. Recall that the optimal triangle for

$\pi$ is obtained from an optimal triangle for $\varpi_\pi$ by stretching and aligning it in the directions of the eigenvectors of $Q_\pi$. The angle $\vartheta$ is used to adjust these directions so that abrupt changes of alignments may not occur. In doing so, the new alignment directions are well-conditioned and the error on the adjusted triangles can be bounded as shown in part $c$. of Proposition 3.2.4.

Let $U_\pi = [\mathbf{v}_1 \ \mathbf{v}_2]$ where $\mathbf{v}_1 = [\cos\mu \quad \sin\mu]^t$ and $\mathbf{v}_2 = [-\sin\mu \quad \cos\mu]^t$ denote the normalized eigenvectors associated with $\lambda_1$ and $\lambda_2$. Let $\pi' \in \mathbb{H}_2$ be a quadratic polynomial obtained by small perturbations of the coefficients of $\pi$. Denote by $\mathbf{v}'_1, \mathbf{v}'_2$ the eigenvectors corresponding to the eigenvalues $\lambda'_1, \lambda'_2$ of $\pi'$.

The following result is found in [35, Corollary 5.5.6].

**Lemma 3.2.1.** *For any eigenvalue $\lambda'_i$, $i = 1, 2$, of $Q_{\pi'}$, there is an eigenvalue $\lambda_j$ of $Q_\pi$, with $j \in \{1, 2\}$, such that*

$$|\lambda'_i - \lambda_j| \leq \|Q_{\pi'} - Q_\pi\|_2. \tag{3.19}$$

Lemma 3.2.1 shows that the eigenvalues of $Q_\pi$ are always well-conditioned. With $\mu'$ denoting the angle of $U_{\pi'}$, writing $\mathbf{v}'_1 = [\cos\mu' \quad \sin\mu']^t$ and $\mathbf{v}'_2 = [-\sin\mu' \quad \cos\mu']^t$, simple computations show that

$$\begin{aligned} \|\mathbf{v}_k - \mathbf{v}'_k\|_2 &= \sqrt{(\cos\mu - \cos\mu')^2 + (\sin\mu - \sin\mu')^2} = \sqrt{2 - 2\cos|\mu - \mu'|} \\ &= 2\sin\Big(\frac{|\mu - \mu'|}{2}\Big), \end{aligned} \tag{3.20}$$

hold for each $k = 1, 2$. Thus, the 2-norm of the perturbation vector $\mathbf{v}_k - \mathbf{v}'_k$ depends on the difference $\mu - \mu'$. Let $\mathbf{w}_1, \mathbf{w}_2$ denote two normalized *left* eigenvectors of $Q_\pi$ associated with $\lambda_1, \lambda_2$, that is, satisfying $\mathbf{w}_i^t Q_\pi = \lambda_i \mathbf{w}_i^t$, $i = 1, 2$, and define $s_i = \mathbf{w}_i^t \mathbf{v}_i$.

**Lemma 3.2.2** ([35]). *For each $i = 1, 2$, define $\kappa_i = |s_i|^{-1}$ as the condition number associated with the eigenvalue $\lambda_i$. Then, for each $k = 1, 2$,*

$$\|\mathbf{v}_k - \mathbf{v}'_k\|_2 \leq \Big(\sum_{i \neq k} \frac{\kappa_i}{|\lambda_k - \lambda_i|}\Big) \|Q_\pi - Q_{\pi'}\|_2 + O\Big(\|Q_\pi - Q_{\pi'}\|_2^2\Big). \tag{3.21}$$

The result in Lemma 3.2.2 shows that if $Q_\pi$ has distinct eigenvalues, all of which being always well-conditioned, then the eigenvectors of $Q_\pi$ are well-conditioned. Note that since $Q_\pi$ is symmetric, we have $\mathbf{w}_i = \mathbf{v}_i$ so that $\kappa_i = 1$, $i = 1, 2$, and (3.21) takes the form

$$\|\mathbf{v}_k - \mathbf{v}_{k'}\|_2 \leq \frac{1}{|\lambda_2 - \lambda_1|}\|Q_\pi - Q_{\pi'}\|_2 + O\Big(\|Q_\pi - Q_{\pi'}\|_2^2\Big). \tag{3.22}$$

Our objective is to replace the angle of rotation $\mu$ of $Q_\pi$ by the angle $\mu - \vartheta$, with $\vartheta$ appropriately chosen in such a way that the triangle is still near optimal in the sense of part $c$. of Proposition 3.2.4. Writing $\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix} = \lambda_1 I + \begin{bmatrix}0 & 0\\ 0 & \lambda_2 - \lambda_1\end{bmatrix}$, the difference between the matrices $Q_\pi$ and $\bar{Q}_\pi$ can be expressed in terms of matrices involving the difference $\lambda_2 - \lambda_1$,

$$\begin{aligned}
Q_\pi - \bar{Q}_\pi =& U_\pi\Big(\lambda_1 I + \begin{bmatrix}0 & 0\\ 0 & \lambda_2 - \lambda_1\end{bmatrix}\Big)U_\pi^t - R_{\mu-\vartheta}\Big(\lambda_1 I + \begin{bmatrix}0 & 0\\ 0 & \lambda_2 - \lambda_1\end{bmatrix}\Big)R_{\mu-\vartheta}^t\\
=& U_\pi\begin{bmatrix}0 & 0\\ 0 & \lambda_2 - \lambda_1\end{bmatrix}U_\pi^t - U_\pi R_{-\vartheta}\begin{bmatrix}0 & 0\\ 0 & \lambda_2 - \lambda_1\end{bmatrix}U_\pi^t\\
& - U_\pi R_{-\vartheta}\begin{bmatrix}0 & 0\\ 0 & \lambda_2 - \lambda_1\end{bmatrix}R_{-\vartheta}^t U_\pi^t + U_\pi R_{-\vartheta}\begin{bmatrix}0 & 0\\ 0 & \lambda_2 - \lambda_1\end{bmatrix}U_\pi^t\\
=& U_\pi\Big(I - R_{-\vartheta}\Big)\begin{bmatrix}0 & 0\\ 0 & \lambda_2 - \lambda_1\end{bmatrix}U_\pi^t + U_\pi R_{-\vartheta}\begin{bmatrix}0 & 0\\ 0 & \lambda_2 - \lambda_1\end{bmatrix}\Big(I - R_{-\vartheta}^t\Big)U_\pi^t.
\end{aligned}$$

Since $\Big(I - R_{-\vartheta}\Big)^t = I - R_{-\vartheta}^t = \begin{bmatrix}1 - \cos\vartheta & -\sin\vartheta\\ \sin\vartheta & 1 - \cos\vartheta\end{bmatrix}$, the 2-norm of $Q_\pi - \bar{Q}_\pi$ is bounded by using terms involving the difference $\lambda_2 - \lambda_1$ and $\vartheta$, that is,

$$\begin{aligned}
\|Q_\pi - \bar{Q}_\pi\|_2 \leq & \left\|\Big(I - R_{-\vartheta}\Big)\begin{bmatrix}0 & 0\\ 0 & \lambda_2 - \lambda_1\end{bmatrix}\right\|_2 + \left\|\begin{bmatrix}0 & 0\\ 0 & \lambda_2 - \lambda_1\end{bmatrix}\Big(I - R_{-\vartheta}^t\Big)\right\|_2\\
=& \left\|\begin{bmatrix}0 & (\lambda_2 - \lambda_1)\sin\vartheta\\ 0 & (\lambda_2 - \lambda_1)(1 - \cos\vartheta)\end{bmatrix}\right\|_2 + \left\|\begin{bmatrix}0 & 0\\ (\lambda_2 - \lambda_1)\sin\vartheta & (\lambda_2 - \lambda_1)(1 - \cos\vartheta)\end{bmatrix}\right\|_2.
\end{aligned}$$

The two matrices on the right hand side are transpose to each other, their eigenvalues are exactly 0 and $\eta = (\lambda_2 - \lambda_1)(1 - \cos\vartheta)$, and we have

$$\|Q_\pi - \bar{Q}_\pi\|_2 \leq 2|\eta| = 4|\lambda_2 - \lambda_1|\sin^2(\vartheta/2). \tag{3.23}$$

We are interested in the choice of the angle $\vartheta \in [0, 2\pi]$ when the difference of eigenvalues $|\lambda_2 - \lambda_1|$ is small. The difference of eigenvalues $|\lambda_2 - \lambda_1|$ can be expressed in terms of the coefficients of $\pi$, a useful tool to determine the adjustment angle $\vartheta$ in (3.26).

**Lemma 3.2.3.** *Given a quadratic polynomial $\pi(x, y) = ax^2 + 2bxy + cy^2$ such that $ac - b^2 \neq 0$, denote by $\lambda_1, \lambda_2$ its eigenvalues, and by $\mu$ the angle of rotation of the matrix $U_\pi$ defined in (2.2). Then*

$$|\lambda_2 - \lambda_1| = \sqrt{(c-a)^2 + 4b^2}, \quad \textit{and if } b \neq 0, \ \tan\mu = \frac{c - a - (\lambda_2 - \lambda_1)}{2b}.$$

*Proof.* Let $Q_\pi = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ be the matrix associated with $\pi$. With $\mathbf{v}_1 = [\cos\mu \ \ \sin\mu]^t$ and $\mathbf{v}_2 = [-\sin\mu \ \ \cos\mu]^t$ denoting the normalized eigenvectors associated with $\lambda_1$ and $\lambda_2$, from $Q_\pi \mathbf{v}_i = \lambda_i \mathbf{v}_i$, $i = 1, 2$, we deduce the following system of equations:

$$a\cos\mu + b\sin\mu = \lambda_1\cos\mu \ \text{ and } \ b\cos\mu + c\sin\mu = \lambda_1\sin\mu, \tag{3.24}$$

$$-a\sin\mu + b\cos\mu = -\lambda_2\sin\mu \ \text{ and } \ -b\sin\mu + c\cos\mu = \lambda_2\cos\mu. \tag{3.25}$$

We easily deduce that

$$(c-a)\sin\mu + 2b\cos\mu = (\lambda_1 - \lambda_2)\sin\mu$$
$$(c-a)\cos\mu - 2b\sin\mu = (\lambda_2 - \lambda_1)\cos\mu.$$

By taking the squares and adding the two equations to one another, we deduce that $(\lambda_2 - \lambda_1)^2 = (c-a)^2 + 4b^2$. Note that if $b = 0$, the above equations yield

$$\Big((c-a) + (\lambda_2 - \lambda_1)\Big)\sin\mu = 0 \ \text{ and } \ \Big((c-a) - (\lambda_2 - \lambda_1)\Big)\cos\mu = 0,$$

which imply that $|\lambda_2 - \lambda_1| = |c - a|$ as before, and $\mu = 0$ if $\lambda_2 - \lambda_1 = c - a$, whereas $\mu = \frac{\pi}{2}$ if $\lambda_2 - \lambda_1 = a - c$. If $b \neq 0$, by dividing the left and right hand sides by $\cos\mu$, the second equation implies that $\tan\mu = \frac{(c-a)-(\lambda_2-\lambda_1)}{2b}$. □

The following result shows how the adjustment angle $\vartheta$ (which is firstly introduced in (3.18) and for which (3.23) holds) is chosen in such a way that the alignment directions defined by the column vectors of the matrix $R_{\mu-\vartheta}$ occurring in (3.18) are well-conditioned in the sense of (3.28).

**Proposition 3.2.4.** *Let $f \in C^2(\Omega)$ and $z \in \Omega$ be such that $\pi_z$ is non-degenerate. Given $\varepsilon > 0$, we define the adjustment angle $\vartheta$ by*

$$\vartheta = \begin{cases} \dfrac{\mu\Big(\varepsilon^{\frac{1}{2}} - |\lambda_2 - \lambda_1|\Big)}{\varepsilon^{\frac{1}{2}} + |\lambda_2 - \lambda_1|}, & \text{if } |\lambda_2 - \lambda_1| \leq \varepsilon^{\frac{1}{2}}, \\ \\ 0, & \text{otherwise}, \end{cases} \tag{3.26}$$

*with $\lambda_1, \lambda_2$ denoting the eigenvalues of $\pi_z$ and $\mu \in [0, \pi)$ being the angle of rotation of $U_{\pi_z}$. The following statements hold:*

*a. $\vartheta = \mu$ if $\lambda_2 = \lambda_1$.*

*b. $\|Q_{\pi_z} - \bar{Q}_{\pi_z}\|_2 \leq \mu^2 \varepsilon^{\frac{1}{2}}$, where $\bar{Q}_{\pi_z}$ is defined in (3.18);*

*c. Given a triangle $T$ containing the origin, defining $\bar{T} := R_{-\vartheta}(T)$, we have*

$$\|\pi_z - I_{\bar{T}}\pi_z\|_{L_p(\bar{T})} \leq \|\pi_z - I_T\pi_z\|_{L_p(T)} + 9h^2|T|^{\frac{1}{p}}\mu^2\varepsilon^{\frac{1}{2}}, \tag{3.27}$$

*with $h$ being the diameter of $T$.*

*d. Suppose that $|\lambda_2 - \lambda_1| \leq \varepsilon^{\frac{1}{2}}$. Then, for any $z' \in \Omega$ such that $\det \pi_{z'} \neq 0$, $|\lambda_2' - \lambda_1'| \leq \varepsilon^{\frac{1}{2}}$ and $\|Q_{\pi_{z'}} - Q_{\pi_z}\|_2 \leq \varepsilon$, there exists a constant $C_1$ such that*

$$|\Delta(\mu - \vartheta)| = |(\mu - \vartheta) - (\mu' - \vartheta')| \leq C_1\varepsilon^{\frac{1}{2}} + O(\varepsilon^2), \tag{3.28}$$

*where $\lambda_1', \lambda_2'$ are the eigenvalues of $\pi_{z'}$, with $\mu'$ being the angle of rotation of $U_{\pi_{z'}}$, $\vartheta'$ the corresponding adjustment angle, and where the constants in (3.28) are absolute.*

In the result (3.46) of Proposition 3.3.2, we show that (3.28) of part *d.* of Proposition 3.2.4 holds without any assumptions on $|\lambda_2 - \lambda_1|$ and $|\lambda_2' - \lambda_1'|$.

*Proof of Proposition 3.2.4.* The proof of $a$. is straightforward.

$b$. From the definition of $\vartheta$, we have $\vartheta = 0$ when $|\lambda_2 - \lambda_1| \geq \varepsilon^{\frac{1}{2}}$, whereas when $|\lambda_2 - \lambda_1| \leq \varepsilon^{\frac{1}{2}}$, we find that

$$\vartheta \leq \frac{\mu\Big(\varepsilon^{\frac{1}{2}} - |\lambda_2 - \lambda_1|\Big)}{|\lambda_2 - \lambda_1|} \leq \frac{\mu\varepsilon^{\frac{1}{2}}}{|\lambda_2 - \lambda_1|}.$$

Since by definition the condition $\vartheta \leq \mu$ always holds, we have $\sin^2(\frac{\vartheta}{2}) \leq \frac{\vartheta^2}{4} \leq \frac{\vartheta}{4}\mu$, so that the above inequality and (3.23) yield

$$\|Q_\pi - \bar{Q}_\pi\|_2 \leq |\lambda_2 - \lambda_1|\vartheta\mu \leq \mu^2\varepsilon^{\frac{1}{2}}.$$

$c$. For convenience of notation, let $\pi := \pi_z$ and consider the homogeneous quadratic polynomial $\bar{\pi} := \pi \circ R_\vartheta$. After first noticing from (2.12) that $(I_{\bar{T}}\bar{\pi}) \circ R_{-\vartheta} = (I_{R_{-\vartheta}(T)}\bar{\pi}) \circ R_{-\vartheta} = I_T(\bar{\pi} \circ R_{-\vartheta})$, we have

$$\begin{aligned}
\|\bar{\pi} - I_{\bar{T}}\bar{\pi}\|^p_{L_p(\bar{T})} &= \int_{R_{-\vartheta}(T)} \Big|\bar{\pi}(x,y) - I_{\bar{T}}\bar{\pi}(x,y)\Big|^p \mathrm{d}x\mathrm{d}y \\
&= \int_T \Big|\bar{\pi} \circ R_{-\vartheta}(x,y) - (I_{\bar{T}}\bar{\pi}) \circ R_{-\vartheta}(x,y)\Big|^p \mathrm{d}x\mathrm{d}y \\
&= \int_T |\pi(x,y) - I_T\pi(x,y)|^p \mathrm{d}x\mathrm{d}y \\
&= \|\pi - I_T\pi\|^p_{L_p(T)}.
\end{aligned}$$

We can now estimate the error on $\bar{T}$: By using (2.5) and (2.13), we have

$$\begin{aligned}
\|\pi - I_{\bar{T}}\pi\|_{L_p(\bar{T})} &= \|(\pi - \bar{\pi}) - I_{\bar{T}}(\pi - \bar{\pi}) + (\bar{\pi} - I_{\bar{T}}\bar{\pi})\|_{L_p(\bar{T})} \\
&\leq \|\pi - \bar{\pi}\|_{L_p(\bar{T})} + \|I_{\bar{T}}(\pi - \bar{\pi})\|_{L_p(\bar{T})} + \|\bar{\pi} - I_{\bar{T}}\bar{\pi}\|_{L_p(\bar{T})} \\
&\leq 3h^2_{\bar{T}}|\bar{T}|^{\frac{1}{p}}\|\pi - \bar{\pi}\| + |\bar{T}|^{\frac{1}{p}}\|\pi - \bar{\pi}\|_{L_\infty(\bar{T})} + \|\bar{\pi} - I_{\bar{T}}\bar{\pi}\|_{L_p(\bar{T})} \\
&\leq 6h^2_{\bar{T}}|\bar{T}|^{\frac{1}{p}}\|\pi - \bar{\pi}\| + \|\pi - I_T\pi\|_{L_p(T)}, \qquad (3.29)
\end{aligned}$$

by virtue of the fact that $\bar{T}$ contains the origin so that

$$|\bar{T}|^{\frac{1}{p}}\|\pi - \bar{\pi}\|_{L_\infty(\bar{T})} \leq 3h^2_{\bar{T}}|\bar{T}|^{\frac{1}{p}}\|\pi - \bar{\pi}\|.$$

From the identities $R_\vartheta[x\ y]^t = R^t_{-\vartheta}[x\ y]^t$ and $\left(R_\vartheta[x\ y]^t\right)^t = [x\ y]R_{-\vartheta}$, we obtain

$$\begin{aligned}\bar{\pi}(x,y) = \pi \circ R_\vartheta(x,y) &= [x\ y]R_{-\vartheta}U_\pi \begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix} U^t_\pi R^t_{-\vartheta}[x\ y]^t\\ &= [x\ y]\bar{Q}_\pi[x\ y]^t,\end{aligned}$$

from which we conclude that $Q_{\bar{\pi}} = \bar{Q}_\pi$, with $\bar{Q}_\pi$ defined in (3.18). Moreover, $Q_{\pi-\bar{\pi}} = Q_\pi - Q_{\bar{\pi}} = Q_\pi - \bar{Q}_\pi$ so that $\|\pi - \bar{\pi}\|_2 = \|Q_\pi - \bar{Q}_\pi\|_2$. Now, by using the equivalence of norms in (2.4), we find that $\|\pi - \bar{\pi}\| \le \frac{3}{2}\|Q_\pi - \bar{Q}_\pi\|_2$. Combining this with (3.29), together with the immediate fact from $\bar{T} = R_{-\vartheta}(T)$ that $h_{\bar{T}} = h$ and $|\bar{T}| = |T|$, and the result in part $b$, we obtain

$$\begin{aligned}\|\pi - I_{\bar{T}}\pi\|_{L_p(\bar{T})} &\le 9h^2|T|^{\frac{1}{p}}\|Q_\pi - \bar{Q}_\pi\|_2 + \|\pi - I_T\pi\|_{L_p(T)}\\ &\le 9h^2|T|^{\frac{1}{p}}\mu^2\varepsilon^{\frac{1}{2}} + \|\pi - I_T\pi\|_{L_p(T)}.\end{aligned}$$

*d*. Suppose now that $|\lambda_2 - \lambda_1| \le \varepsilon^{\frac{1}{2}}$. We have

$$\mu - \vartheta = \frac{2\mu|\lambda_2 - \lambda_1|}{\varepsilon^{\frac{1}{2}} + |\lambda_2 - \lambda_1|}. \tag{3.30}$$

With $\bar{\mu} = 2\mu|\lambda_2 - \lambda_1|$ and $\lambda = |\lambda_2 - \lambda_1|$, we see that $\mu - \vartheta$ is given by the function $\ell(\bar{\mu}, \lambda) = \dfrac{\bar{\mu}}{\varepsilon^{\frac{1}{2}} + \lambda}$. With $\mu'$ denoting the angle of rotation of $U_{\pi_{z'}}$ and $\vartheta'$ the corresponding adjustment angle, both of which depending on $z'$, the total variation between $z$ and $z'$ satisfies

$$|\Delta(\mu - \vartheta)| = |(\mu - \vartheta) - (\mu' - \vartheta')| \le \left|\frac{\partial\ell(\hat{\mu}, \hat{\lambda})}{\partial\bar{\mu}}\right||\Delta\bar{\mu}| + \left|\frac{\partial\ell(\hat{\mu}, \hat{\lambda})}{\partial\lambda}\right||\Delta\lambda|, \tag{3.31}$$

for some $\hat{\mu}$ between $\bar{\mu}$ and $\bar{\mu}'$, and some $\hat{\lambda}$ between $\lambda$ and $\lambda'$, where $\lambda' = |\lambda'_2 - \lambda'_1|$ is the difference between the eigenvalues of $\pi_{z'}$. First, simple computations show that, with $\hat{\mu} \le \max\{\bar{\mu}, \bar{\mu}'\} \le 2\varepsilon^{\frac{1}{2}}\max\{\mu, \mu'\} \le 4\pi\varepsilon^{\frac{1}{2}}$,

$$\left|\frac{\partial\ell(\hat{\mu}, \hat{\lambda})}{\partial\bar{\mu}}\right| = \left|\frac{1}{\varepsilon^{\frac{1}{2}} + \hat{\lambda}}\right| \le \varepsilon^{-\frac{1}{2}} \text{ and } \left|\frac{\partial\ell(\hat{\mu}, \hat{\lambda})}{\partial\lambda}\right| = \left|\frac{-\hat{\mu}}{\left(\varepsilon^{\frac{1}{2}} + \hat{\lambda}\right)^2}\right| \le 4\pi\varepsilon^{-\frac{1}{2}}. \tag{3.32}$$

We continue by evaluating $|\Delta\bar{\mu}| = 2|\mu\lambda - \mu'\lambda'|$ and $|\Delta\lambda| = |\lambda - \lambda'|$. From (3.19) there are $i, j \in \{1, 2\}$ for which $\max\{|\lambda_1' - \lambda_i|, |\lambda_2' - \lambda_j|\} \leq \|Q_{\pi_{z'}} - Q_{\pi_z}\|_2$ holds, with $i$ not necessarily different of $j$. By using simple triangular inequalities, we find that

$$\begin{aligned} |\Delta\lambda| &= ||\lambda_2 - \lambda_1| - |\lambda_2' - \lambda_1'|| \\ &\leq \|(\lambda_2 - \lambda_1) + (\lambda_2' - \lambda_1')| \\ &= |(\lambda_2 - \lambda_j) + (\lambda_i - \lambda_1) + (\lambda_j - \lambda_2') + (\lambda_1' - \lambda_i)| \\ &\leq 4\|Q_{\pi_{z'}} - Q_{\pi_z}\|_2. \end{aligned} \tag{3.33}$$

Next, let $\mathbf{v}_1', \mathbf{v}_2'$ denote the orthonormal eigenvectors associated with $\lambda_1', \lambda_2'$. Note that $\lambda\mathbf{v}_1$ is an eigenvector of $Q_{\pi_z}$ associated with $\lambda_1$. The eigenvectors $\lambda\mathbf{v}_1$ and $\lambda\mathbf{v}_2$ are well-conditioned. To see this,

$$\begin{aligned} \|\lambda\mathbf{v}_1 - \lambda'\mathbf{v}_1'\|_2 &= \|\lambda(\mathbf{v}_1 - \mathbf{v}_1') + (\lambda - \lambda')\mathbf{v}_1'\|_2 \\ &\leq \|\lambda(\mathbf{v}_1 - \mathbf{v}_1')\|_2 + |\lambda - \lambda'|\|\mathbf{v}_1'\|_2, \end{aligned} \tag{3.34}$$

where $\|\mathbf{v}_1'\|_2 = 1$. The inequality below is obtained by multiplying (3.22) with $\lambda$,

$$\|\lambda(\mathbf{v}_1' - \mathbf{v}_1)\|_2 \leq \|Q_{\pi_{z'}} - Q_{\pi_z}\|_2 + \lambda \cdot O\Big(\|Q_{\pi_{z'}} - Q_{\pi_z}\|_2^2\Big),$$

which, together with (3.33) and (3.34), yields

$$\|\lambda\mathbf{v}_1 - \lambda'\mathbf{v}_1'\|_2 \leq 5\|Q_{\pi_{z'}} - Q_{\pi_z}\|_2 + \lambda \cdot O\Big(\|Q_{\pi_{z'}} - Q_{\pi_z}\|_2^2\Big). \tag{3.35}$$

This shows that $\lambda\mathbf{v}_1$ is well-conditioned. A similar approach is used to show that $\lambda\mathbf{v}_2$ is also well-conditioned. As a consequence, the angle $\lambda\mu$ must be well-conditioned: Since $\frac{2}{\pi}x \leq \sin x$ for $x \in [0, \frac{\pi}{2}]$, we deduce from (3.20) that $|\mu - \mu'| \leq \frac{\pi}{2}\|\mathbf{v}_1 - \mathbf{v}_1'\|_2$ which together with (3.22) yields

$$\lambda|\mu - \mu'| \leq \frac{\pi}{2}\|Q_{\pi_z} - Q_{\pi_{z'}}\|_2 + \frac{\pi}{2}\lambda O\Big(\|Q_{\pi_z} - Q_{\pi_{z'}}\|_2^2\Big). \tag{3.36}$$

Since

$$\frac{1}{2}|\Delta\bar{\mu}| = |\lambda\mu - \lambda'\mu'| = |\lambda\mu - \lambda\mu' + \lambda\mu' - \lambda'\mu'| \leq \lambda|\mu - \mu'| + |\lambda - \lambda'|\mu',$$

we deduce from (3.33) and (3.36) that

$$\frac{1}{2}|\Delta\bar{\mu}| \leq (\frac{\pi}{2} + 4\mu')\|Q_{\pi_{z'}} - Q_{\pi_z}\|_2 + \pi\lambda \cdot O\Big(\|Q_{\pi_{z'}} - Q_{\pi_z}\|_2^2\Big). \tag{3.37}$$

Combining all together (3.31), (3.32), (3.33) and (3.37), we obtain

$$\begin{aligned}|\Delta(\mu - \vartheta)| &\leq \Big(\pi + 8\mu' + 16\pi\Big)\varepsilon^{-\frac{1}{2}}\|Q_{\pi_{z'}} - Q_{\pi_z}\|_2 + 2\pi\lambda\varepsilon^{-\frac{1}{2}} \cdot O\Big(\|Q_{\pi_{z'}} - Q_{\pi_z}\|_2^2\Big)\\ &\leq C_1\varepsilon^{\frac{1}{2}} + O(\varepsilon^2), \end{aligned} \tag{3.38}$$

by virtue of the facts that $\|Q_{\pi_{z'}} - Q_{\pi_z}\|_2 \leq \varepsilon$ and $\lambda\varepsilon^{-\frac{1}{2}} \leq 1$. In the above estimation, $C_1$ satisfies $C_1 = 33\pi \geq \pi + 8\mu' + 16\pi$. □

On a side note, part $a$. of Proposition 3.2.4 indicates that when the eigenvalues of $\pi_z$ are equal, the angle of adjustment $\vartheta$ for an optimal triangle $T \in \Delta_{\pi_z}$ is none other than the angle $\mu$. Moreover, any pair of orthonormal vectors are eigenvectors of $\pi_z$, therefore a choice of particular directions need to be set, which is discussed in Section 3.2.2 below. Also, the condition that $\|Q_{\pi_{z'}} - Q_{\pi_z}\|_2 \leq \varepsilon$ in the fourth statement of Proposition 3.2.4 is similar to the condition (2.88) with $\varepsilon = \nu$. Such a condition is the first to be taken into account when choosing the parameter $r$ (which is described in Section 3.1.1 and in the second paragraph of Section 3.2) for the partition of the domain $\Omega$ as seen in Section 3.2.2 below.

### 3.2.2 Regular regions

After briefly presenting algorithms to obtain isosceles nearly-optimal triangles for a given quadratic polynomial $\pi$, we shall discuss the initial steps to partition the square domain $\Omega$. We recall that the $x$- and $y$-axes represent the Cartesian coordinate system.

Step 1 of Algorithm 3.1 is justified by Lemma 2.4.3, whereas Step 2 is motivated by Lemma 2.4.5 and the adjustment of alignment directions in Proposition 3.2.4.

**Algorithm 3.1: Near optimal isosceles triangles for $\det \pi > 0$.**

Input $\pi \in \mathbb{H}_2$ such that $\det \pi > 0$, and let $\varepsilon > 0$ be given;

1. Let $T_0$ be the equilateral triangle of unit area such that one of its vertices coincides with the origin and the bisector passing through it lies on the right half of the $x$-axis;
2. Obtain the isosceles triangle $T = \varphi_\pi(T_0)$ where

$$\varphi_\pi = R_{-\vartheta} \circ \phi_\pi, \tag{3.39}$$

where $\phi_\pi$ is defined in (2.42) and $\vartheta$ is as in Proposition 3.2.4.

For $p = \infty$ and $\det \pi < 0$, a similar algorithm can be presented since from Lemma 2.4.4, we can design an isosceles optimal triangle $T_1$ for $\pi_1(x, y) = x^2 - y^2$ as follows: By using triangles whose vertices are given in (2.61) by

$$(0,0), \quad \frac{1}{2}\Big(c_0 a + b, c_0 a - b\Big), \quad \frac{1}{2}\Big(a + c_0 b, a - c_0 b\Big),$$

where $c_0 = \frac{3-\sqrt{5}}{2}$ and $a, b > 0$ such that $\frac{3\sqrt{5}-5}{4}ab = 1$, we can choose $a = b = 2(3\sqrt{5}-5)^{-\frac{1}{2}}$ and obtain an isosceles triangle $T_1$ whose coordinates are given by

$$(0,0), \quad \frac{1}{2}(1 + c_0, c_0 - 1)a, \quad \frac{1}{2}(1 + c_0, 1 - c_0)a. \tag{3.40}$$

Note that one vertex of $T_1$ thus coincides with the origin, the bisector passing through that vertex lies on the right half of the $x$-axis. Also, Step 2 of the algorithm results from Lemma 2.4.5. Note that the adjustment of alignment directions in Proposition 3.2.4. is not needed for $\det \pi < 0$ since the eigenvalues are of different sign and hence the corresponding eigenvectors are well-conditioned.

**Algorithm 3.2: Near optimal isosceles triangles for $\det \pi < 0$ and $p = \infty$.**

Input $\pi \in \mathbb{H}_2$ such that $\det \pi < 0$, and let $\varepsilon > 0$ be given;

1. Choose $T_1 \in \Delta_\infty(\pi_1)$ to have its vertices given by (3.40);
2. Obtain the isosceles triangle $T = \phi_\pi(T_1)$ where $\phi_\pi$ is defined in (2.42).

Observe that the isosceles triangles obtained through Algorithm 3.1 and Algorithm 3.2 always have their smallest interior angle defined by the two edges having the same length.

**Obtaining regular regions.** We now present the initial steps for obtaining regular regions. Assume that $p < \infty$. Let $\eta$ be such that

$$1 + \frac{1}{8p(p+1)} < \eta < 1 + \frac{1}{2p} \tag{3.41}$$

The following steps are set:

(i) For each $k = 1, \ldots, m^2$, obtain an isosceles near optimal triangle $T$ by applying Algorithm 3.1 to $\pi_k = \pi_{b_k}$ where $b_k$ is the barycenter of the sub-square $S_k$. [2] Obtain a scaled and shifted version of $T$

$$T_k := \Lambda_k T + \mathbf{t}_k, \tag{3.42}$$

where

$$\Lambda_k = s^\eta \Big( K_p(\pi_k) + 2Ch_{\pi_k}^2 \omega(r) \Big)^{-\frac{q}{2}} \tag{3.43}$$

where $\frac{1}{q} := 1 + \frac{1}{p}$, $C$ is the constant occurring in (2.56), $h_{\pi_k}$ denotes the diameter of the optimal triangle $T$, the modulus of continuity $\omega$ is defined in

[2] Here $\pi_1 = \pi_{b_1}$ is has nothing to do with the quadratic polynomial $x^2 - y^2$ in Example 2.3.2 which has a negative determinant.

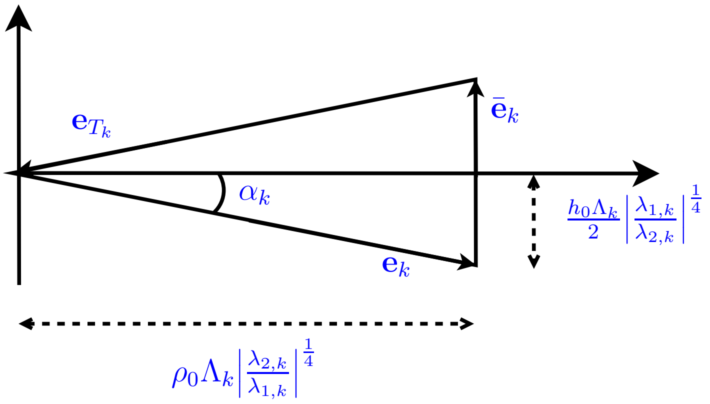

$\mathbf{e}_{T_k}$
$\bar{\mathbf{e}}_k$
$\alpha_k$
$\frac{h_0\Lambda_k}{2}\left|\frac{\lambda_{1,k}}{\lambda_{2,k}}\right|^{\frac{1}{4}}$
$\mathbf{e}_k$
$\rho_0\Lambda_k\left|\frac{\lambda_{2,k}}{\lambda_{1,k}}\right|^{\frac{1}{4}}$

Figure 3.2: Triangle $T_k$ before alignments to the directions of eigenvectors of $\pi_{b_k}$ and before translation into the sub-square $S_k$. Also, $\lambda_{1,k}, \lambda_{2,k}$ denote the eigenvectors of $\pi_{b_k}$ whereas $\alpha_k$ is the initial angle which $\mathbf{e}_k$ makes with the $x$-axis.

(2.75), and $\mathbf{t}_k$ is a translation vector so that the barycenter of $T_k$ coincides with $b_k$. The parameter $s$ is chosen small enough so that $T_k \subset S_k$,

(ii) For each $k = 1, \ldots, m^2$, obtain the *micro-parallelogram* $P_k := P_{\pi_k}$ formed from $T_k$ as follows: The edges of $T_k$ being counterclockwise oriented as shown in Figure 3.2, let $\bar{\mathbf{e}}_k$ be the shortest edge-vector, $\mathbf{e}_{T_k}$ the edge following $\bar{\mathbf{e}}_k$ with respect to the counterclockwise orientation, and $\mathbf{e}_k$ the remaining edge-vector. Note that $\mathbf{e}_k$ is not perpendicular to $\bar{\mathbf{e}}_k$. Then $P_k$ is formed from $T_k$ and its reflexion about the midpoint of $\mathbf{e}_{T_k}$, i.e it is defined by the two edge-vectors $\mathbf{e}_k$ and $\bar{\mathbf{e}}_k$;

(iii) For each $k = 1, \ldots, m^2$, define the polygon $R_k \subset S_k$ as the union of all parallelograms $P \subset R_k$ obtained by aligning side by side the shifted versions of $P_k$, such that for each vertex $z$ of $P$ the four points $z \pm \mathbf{e}_k$, $z \pm \bar{\mathbf{e}}_k$ belong to $S_k$. The reasons for the last condition are discussed in Section 3.2.4 for the study of the interior angles of the so-called *irregular* polygons.

Although in step (iii) we ensure that the vertices $z \pm \mathbf{e}_k$, $z \pm \bar{\mathbf{e}}_k$ belong to the sub-square $S_k$ (these are necessary requirements in Section 3.3), for practical illustration in Figure 3.3 we only ensure that $z \pm \frac{1}{2}\mathbf{e}_k$, $z \pm \frac{1}{2}\bar{\mathbf{e}}_k$ for any vertex $z \in P$ of a parallelogram $P \in R_k$. The remaining figures in this section will use this setting.

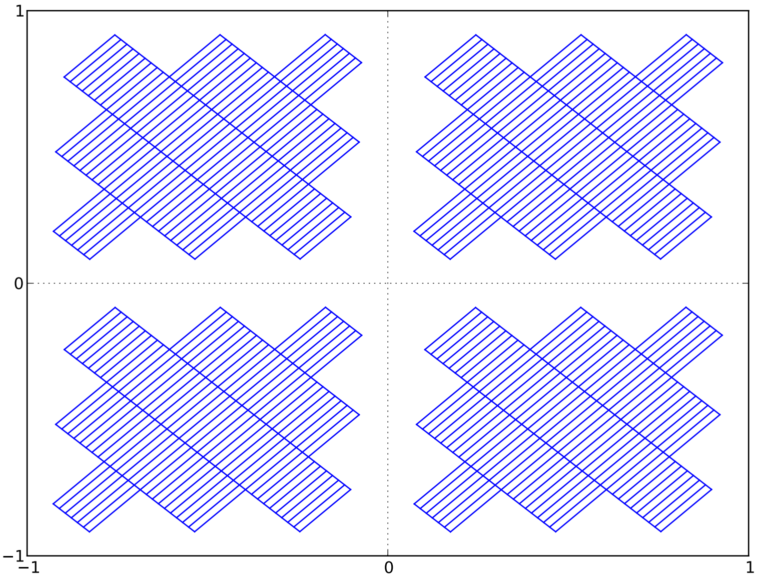


Figure 3.3: Partition into sub-squares and regular regions.

For each $k = 1, \ldots, m^2$, the polygon $R_k$ is termed a *regular* region (see Figure 3.3), it can also be defined by two systems of parallel *segments* $\mathcal{L}_k$ and $\bar{\mathcal{L}}_k$ whose directional vectors are $\mathbf{e}_k$ and $\bar{\mathbf{e}}_k$: A *segment* $\ell$ of $\mathcal{L}_k$ (or $\bar{\mathcal{L}}_k$) is a straight line segment defined by the sides of a collection of parallelograms of $R_k$ which are aligned side by side in $\mathbf{e}_k$ (or $\bar{\mathbf{e}}_k$) direction, such that the end-points are at the boundary of $R_k$ and $\ell$ cuts $R_k$ in two regions unless $\ell$ is part of its boundary. A *vertex* of $R_k$ is a vertex of a parallelogram such that it is on the boundary of $R_k$.

The regular region $R_k$ is *directionally convex* in the direction of $\mathbf{e} = \mathbf{e}_k, \bar{\mathbf{e}}_k$, in the sense described as follows: $R_k$ can be cut in the direction of $\mathbf{e}$, thereby creating layers of blocks $(D_{k,n})_n$ where each $D_{k,n}$ is an union of parallelograms aligned side by side in $\mathbf{e}$ direction. Moreover, since the square $S_k$ is convex, the region between non-consecutive parallelograms inside $D_{k,n}$ is still inside $D_{k,n}$, hence showing that $D_{k,n}$ is convex.

Each regular region has a natural triangulation by drawing the shortest diagonals of its parallelograms, the triangles thus obtained are called *regular* triangles. Observe that regular triangles are therefore near-optimal triangles, the study of the local errors on such triangles is already given in Section 2.5.

**Remark 3.2.1** (Removal of vertices in (iii))**.** For the extension of segments discussed in Section 3.2.3 below, it is preferable to control the minimum distance between the vertices of $R_i$ and the sides of the sub-square $S_i$, with $i = 1, \ldots, m^2$. Therefore in (iii) we only include parallelograms whose vertices are at a certain distance from the sides of $S_i$. Without analytical proof, numerical implementations show that vertices which are too close to the sides of the sub-squares can cause big angles in the final triangulation of $\Omega$. Moreover, their removal allows us to avoid unnecessary processes such as identification of vertices which are very close to one another.

### 3.2.3 Extension of segments

Following the initial steps in the previous section, we proceed to split the remaining non-covered space $\Omega^{\text{irr}} := \Omega \backslash \cup_i R_i$ by using either of the Settings 1, 2, 3, 4 or 5 given hereafter. The splitting is characterized by the extensions of the segments in $\mathcal{L}_i \cup \bar{\mathcal{L}}_i$ up to a neighboring regular region $R_j$ or up to the boundary, and thereby generating the so-called *irregular* polygons which have six edges at most. Extensions follow directions parallel to either of $\mathbf{e}_i$ or $\bar{\mathbf{e}}_i$, with $i = 1, \ldots, m^2$. Due to similarity, we shall only present algorithms that extend the segments in $\mathcal{L}_i$, and give additional details when ambiguity may occur.

For each $k = 1, \ldots, m^2$, since the segments in $\mathcal{L}_k$ are all parallel to $\mathbf{e}_k$, we can order them in the direction of $\bar{\mathbf{e}}_k$, with $\mathbf{e}_k$ and $\bar{\mathbf{e}}_k$ as defined in (iii) of Section 3.2.2. The first and last segments with respect to such order are called the *end-segments* of $\mathcal{L}_k$. The end-segments of $\bar{\mathcal{L}}_k$ are defined in a similar way. Observe that end-segments are part of the boundary of the regular region.

We say that two squares $S_i, S_j$, with $i, j \in \{1, \ldots, m^2\}$ are *neighbors* if they share a vertex or an edge. Two regular regions $R_i, R_j$ are termed *neighbors* if the squares that contain them are neighbors. Note that, apart from $S_i$ itself, there are at most eight squares that are neighbors to $S_i$. Some of them share edges or corners with $S_i$.

For each $i = 1, \ldots, m^2$, consider the set of parallel segments $\mathcal{L}_i$. A given segment $\ell_0 \in \mathcal{L}_i$ is oriented according to the direction of the vector $\mathbf{e}_i$ defined in (iii) of Section 3.2.2. By using either of the directions $\mathbf{e}_i$ or $-\mathbf{e}_i$ and one of the settings presented hereafter, the segment $\ell_0$ is *extended* up to a neighboring region $R_j$ or to the boundary. Once more, due to similarity, we shall only describe extensions by using the direction $\mathbf{e}_i$. We denote by $z_0$ the end-point of $\ell_0$ in the direction of $\mathbf{e}_i$, and connect it to another point $z_1$ of a neighboring region $R_j$, or to the boundary. Note that the extended segments may overlap, for example by extending from $R_i$ to $R_j$ in the direction of $\mathbf{e}_i$ and from $R_j$ to $R_i$ in the direction of $-\mathbf{e}_j$ we may obtain the same extension twice.

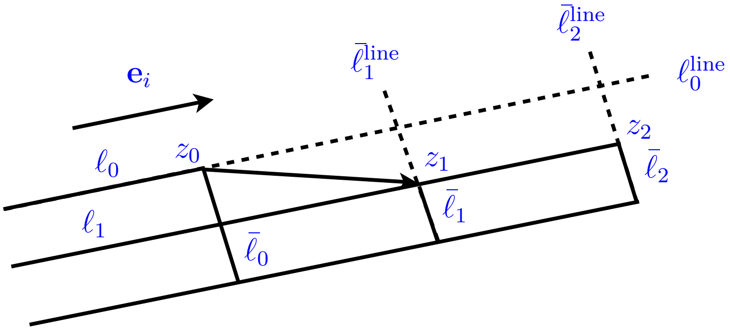


Figure 3.4: Connection to own square.

```
Setting 1: Extend $\ell_0 \in \mathcal{L}_i$ into own square.
if the following conditions hold: For some $\bar{\ell}_1, \bar{\ell}_2 \in \bar{\mathcal{L}}_i$;
    1. The ray originating at $z_0$ in the direction of $\mathbf{e}_i$ intersects the straight lines $\bar{\ell}_1^{\text{line}}$, $\bar{\ell}_2^{\text{line}}$ containing $\bar{\ell}_1$ and $\bar{\ell}_2$ outside of $R_i$, see Figure 3.4. There is no segment in $\bar{\mathcal{L}}_i$ between $\bar{\ell}_1$ and $\bar{\ell}_0$ that contains $z_0$;
    2. With $z_1 \in \bar{\ell}_1$ and $z_2 \in \bar{\ell}_2$ being the closest end-points to $z_0$, $[z_1, z_2]$ is part of a single segment $\ell_1$ of $\mathcal{L}_i$;
    3. There is no segment in $\mathcal{L}_i$ between $\ell_0$ and $\ell_1$.
then
    Connect the point $z_0$ with $z_1$.
```

The connection in Setting 1, as illustrated in Figure 3.4, is a cautious step in order to avoid a possible very long extension of $\ell_0$ due to Setting 2, or crossing of regular regions by extended segments resulting from either of Setting 2 or 3.

If Setting 1 does not apply for $z_0$, we apply Setting 2 where $S_j$ is the neighbor of $S_i$ which the extension $\ell_0^{\text{line}}$ of $\ell_0$ in the direction of $\mathbf{e}_i$ intersects first. If $\ell_0^{\text{line}}$ passes through one of the corners of $S_i$, we choose $S_j$ to be any of the two squares sharing that corner and an edge with $S_i$. We denote $h_i = |\mathbf{e}_i|$ and $d_i = |\bar{\mathbf{e}}_i|$.

```
Setting 2-a: Extend $\ell_0 \in \mathcal{L}_i$ up to $R_j$ in $\mathbf{e}_i$ direction.
if the line $\ell_0^{\text{line}}$ extending $\ell_0$ intersects the regular region $R_j$.
    Consider the first segment $\bar{\ell}_1 \in \bar{\mathcal{L}}_j$ that $\ell_0^{\text{line}}$ intersects in the direction of
    $\mathbf{e}_i$. Denote by $z_1 \in \bar{\ell}_1$ the vertex of $R_j$ which is the closest to the
    intersection point $\ell_0^{\text{line}} \cap \bar{\ell}_1$ and which is not a re-entrant corner of $R_j$, see
    Figure 3.6;
then
    Connect $z_0$ with the vertex $z_1$.
```

Note that the existence of a segment in $\bar{\mathcal{L}}_i$ intersected by $\ell_0^{\text{line}}$ is guaranteed by the fact that the angle between $\mathbf{e}_i$ and $\mathbf{e}_j$ is small due to the assumption (3.17) and in view of (3.45) and Proposition 3.3.2 with $\varepsilon = \frac{3}{2}\omega(\sqrt{2}r)$.

We use Setting 2-b below only if Setting 2-a does not apply, that is, the line $\ell_0^{\text{line}}$ does not intersect the regular region $R_j$.

```
Setting 2-b: Extend $\ell_0 \in \mathcal{L}_i$ up to $R_j$ in $\mathbf{e}_i$ direction, with tolerance.
Initialize the set $\bar{\mathcal{L}}_j^*$ to be empty, and compute $d_j = |\bar{\mathbf{e}}_j|$, $h_i = |\mathbf{e}_i|$;
foreach $\bar{\ell} \in \bar{\mathcal{L}}_j$ do
    i. Compute $L = d(z_0, \bar{z})$ which is the distance between $z_0$ and
       $\bar{z} = \ell_0^{\text{line}} \cap \bar{\ell}^{\text{line}}$, with $\bar{\ell}^{\text{line}}$ being the straight line extending $\bar{\ell}$;
    ii. Add to $\bar{\mathcal{L}}_j^*$ the segment $\bar{\ell}^*$ defined by extending $\bar{\ell}$ by segments of
        the length $\frac{d_j L}{8h_i}$ at both ends (see Figure 3.5);
if (C1) and either of (C2)-($i$), (C2)-($ii$) below hold:
(C1) The straight line $\ell_0^{\text{line}}$ extending $\ell_0$ does not intersect $R_j$ but intersects a segment in $\bar{\mathcal{L}}_j^*$;
        Consider the first segment $\bar{\ell}_1^* \in \bar{\mathcal{L}}_j^*$ that $\ell_0^{\text{line}}$ intersects in the direction of $\mathbf{e}_i$. Denote by $z_1 \in \bar{\ell}_1^* \cap R_j$ the closest vertex of $R_j$ to the intersection point $\ell_0^{\text{line}} \cap \bar{\ell}_1^*$, which is not a re-entrant corner of $R_j$;
(C2) ($i$) $z_1 \in R_j$ does not belong to the interior of an end-segment in $\mathcal{L}_j$;
     ($ii$) The second sub-square which $\ell_0^{\text{line}}$ intersects in $\mathbf{e}_i$ direction is not a neighbor sub-square to $R_i$.
then
    Connect $z_0$ with the vertex $z_1$.
```

The *enlarged* segment $\bar{\ell}^*$ as described in ii. of Setting 2-b is an extension of $\bar{\ell}$ into a longer segment characterized by the value $\frac{d_j L}{8h_i} = \frac{L}{8h_i}|\mathbf{e}_j|$ which we call *tolerance*, as illustrated in Figure 3.6. Analogously, when extending a segment $\ell_0 \in \bar{\mathcal{L}}_i$ up to $R_j$ in the direction parallel to $\bar{\mathbf{e}}_i$, the tolerance is $\frac{h_j L}{8d_i}$ where $L$ remains the distance between $z_0 \in \ell_0$ and the intersection point $z^* = \ell_0^{\text{line}} \cap \bar{\ell}^{\text{line}}$, with $\bar{\ell} \in \mathcal{L}_j$. In the second condition of Setting 2-a, the vertex $z_1$ may be an end-point of an end-segment of $\mathcal{L}_j$.

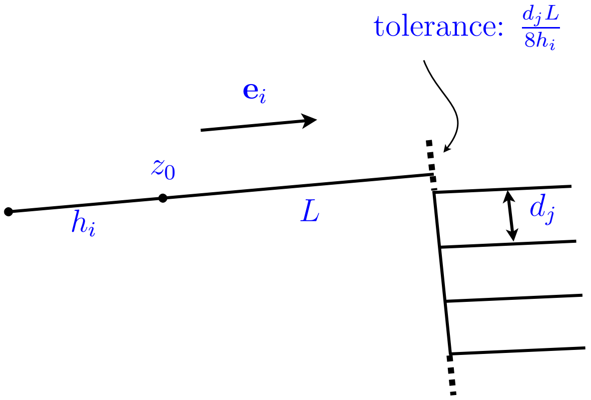


Figure 3.5: Illustration of the tolerance $\frac{d_j L}{8h_i}$.

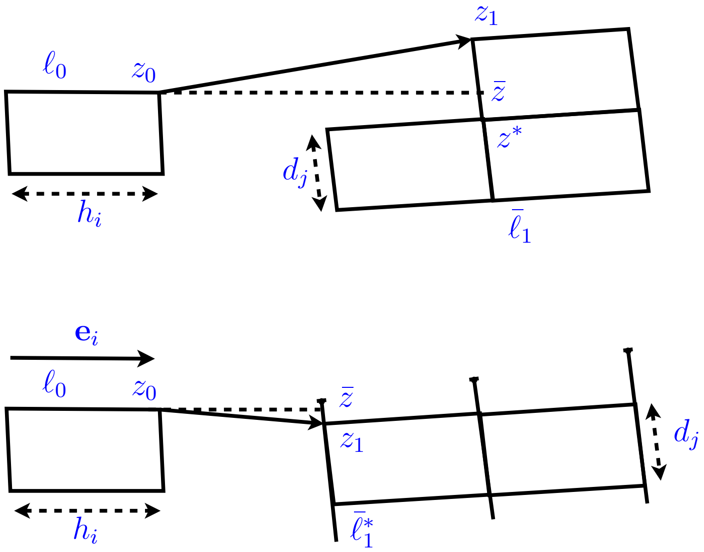


Figure 3.6: Comparison of connections in $\mathbf{e}_i$ direction: on top by using Setting 2-a, with no connection to the re-entrant corner $z^*$. At the bottom by using Setting 2-b, with tolerance $\frac{d_j L}{8h_i}$ (see also Figure 3.5).

Many properties of the final triangulation depend on the choice of the tolerance. Namely, in Section 3.2.4 the factor $\frac{1}{8}$ in the tolerance will play a crucial role.

If either of Setting 1, Setting 2-a or Setting 2-b applies to $z_0$, then $z_0$ is termed a *connected* vertex, otherwise it is called a *non-connected* vertex of $\ell_0$ with direction $\mathbf{e}_i$, and we apply Setting 3. In Figure 3.7 we provide an illustration of the configurations after applying Setting 1 and 2. For each regular region $R_i$, $i = 1, \dots, m^2$, we associate four sets of non-connected vertices $V_i^{\mathrm{u}}, V_i^{\mathrm{d}}$ and $\bar{V}_i^{\mathrm{u}}, \bar{V}_i^{\mathrm{d}}$

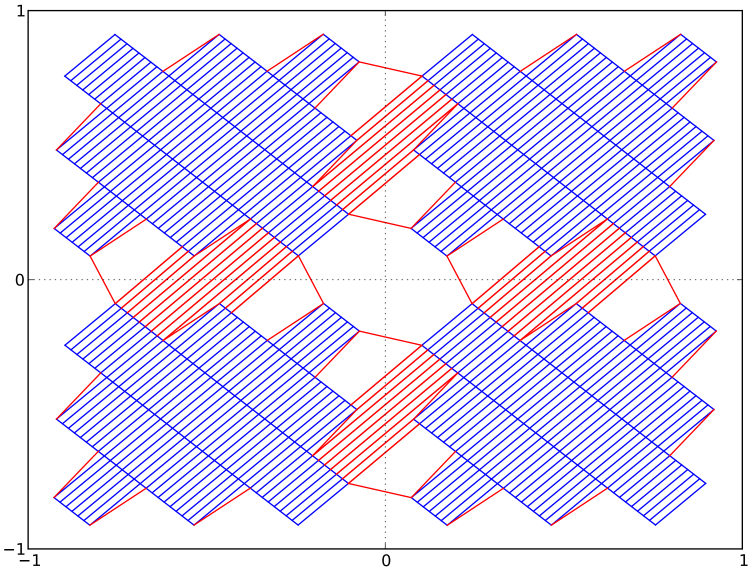


Figure 3.7: Connections after Setting 1 and Setting 2.

which we illustrate in Figure 3.8 and which are defined as follows.

- Let $V_i^{\mathrm{u}}$ (resp. $V_i^{\mathrm{d}}$) denote the set of non-connected vertices of all segments in $\mathcal{L}_i$, all of them in the direction of $\mathbf{e}_i$ (resp. $-\mathbf{e}_i$).
- Similarly, let $\bar{V}_i^{\mathrm{u}}$ (resp. $\bar{V}_i^{\mathrm{d}}$) denote the set of non-connected vertices of all segments in $\bar{\mathcal{L}}_i$, all of them in the direction of $\bar{\mathbf{e}}_i$ (resp. $-\bar{\mathbf{e}}_i$).

The segments of $\mathcal{L}_i$ which contain the non-connected vertices of $V_i^{\mathrm{u}}$ are ordered in the direction of $\bar{\mathbf{e}}_i$, and similarly, the segments of $\mathcal{L}_j$ which contain the non-connected vertices of $V_j^{\mathrm{d}}$ are ordered in the direction of $\bar{\mathbf{e}}_j$. Note that the four sets are not necessarily disjoint to one another, however, one vertex may belong to two sets at most. Only a vertex which belongs to just one parallelogram of $R_i$ may belong to two sets at once, see Figure 3.8.

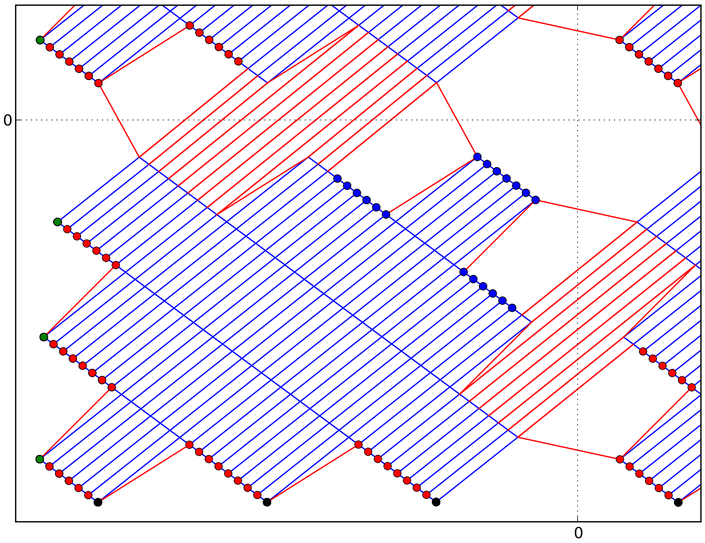


Figure 3.8: The sets of vertices at the bottom-left sub-square of Figure 3.7: $V_i^{\text{u}}$ colored in blue; $V_i^{\text{d}}$ colored in red; $\bar{V}_i^{\text{u}}$ colored in green; And $\bar{V}_i^{\text{d}}$ colored in black. Only a vertex which belongs to only one parallelogram can belong to two sets.

If neither Setting 1 nor Setting 2-a, 2-b applies to $z_0$, then the extension of $\ell_0$ in the direction of $\mathbf{e}_i$ either crosses the boundary of $\Omega$ at a point on the boundary of $S_j$, or the second sub-square intersected by it is also a neighbor of $S_i$ sharing with it a single vertex. Indeed, by making $s$ small enough in (3.43) so that there are enough segments in $\bar{\mathcal{L}}_j$ from the regular region $R_j$, there must be at least one enlarged segment $\bar{\ell}^*$ that intersect the side of $S_j$. Hence, the failure of an extension by Setting 2-b means that the second sub-square that the line $\ell_0^{\text{line}}$ intersect must be a neighbor of $S_i$ sharing with it a single vertex.

Note that if $z_0$ belongs to just one parallelogram of a regular region $R_i$, then it is the originating point of extensions of exactly two segments $\ell_0 \in \mathcal{L}_i$ and $\ell_1 \in \bar{\mathcal{L}}_i$.

Let $S_j$ denote the second sub-square intersected by the extension of $\ell_0$ in the direction of $\mathbf{e}_i$. As mentioned before, $S_j$ is necessarily a neighbor of $S_i$ and shares one vertex with it, denoted by $v_{ij}$.

**Setting 3: Connection across.**

Let $S_i$ and $S_j$ be neighboring sub-squares sharing exactly one vertex $v_{ij}$;

Compute $d_j = |\bar{\mathbf{e}}_j|$;

Extend each segment $\ell_0^* \in \mathcal{L}_i$ containing a vertex $z_0^* \in V_i^{\mathrm{u}}$ into $R_j$, which is done by using either Setting 2-a or Setting 2-b with direction $\mathbf{e}_i$, with the tolerance $\frac{d_j L}{8h_i}$ being replaced by $\frac{d_j}{2}$.

If $z_0^*$ remains unconnected to the regular region $R_j$, then apply the extension in Setting 2-b by using only the condition (C1), with the direction $\mathbf{e}_i$ and the tolerance $\frac{d_j L}{8h_i}$.

Note that the segments containing some vertices in $\in V_i^{\mathrm{d}}, \in \bar{V}_i^{\mathrm{u}}$ and $\in \bar{V}_i^{\mathrm{d}}$ are also extended by using Setting 3 by using the directions $-\mathbf{e}_i, \bar{\mathbf{e}}_i$ and $-\bar{\mathbf{e}}_i$. The goal of Setting 3 is to make sure that there are no hanging vertices. It applies to vertices which are non-connected (resulting from the failures of Setting 1-2). Note that the vertex $z_0^*$ may belong to only one parallelogram, and it can thus be a connected vertex with respect to one of the directions $\bar{\mathbf{e}}_i$ or $-\bar{\mathbf{e}}_i$.

We show in Figure 3.9 an example of connections obtained by using Setting 3. Observe, in particular, that the extensions by Setting 3 always connect two vertices $z_0$ and $z_1$ from neighboring sub-squares, unless the line originating at $z_0$ cross the boundary of $\Omega$. This is where Setting 4 described below applies.

**Setting 4: Connection to the boundary of $\Omega$.**

```
Suppose that $S_i$ has a side overlapping the boundary of $\Omega$;
foreach $\ell_0 \in \mathcal{L}_i$ do
    if the end-point $z_0 \in \ell_0$ belongs to any of the sets $V_i^{\mathrm{u}}, V_i^{\mathrm{d}}, \bar{V}_i^{\mathrm{u}}$ or $\bar{V}_i^{\mathrm{d}}$;
    then
        Connect $z_0 \in \ell_0$ with the intersection point $\ell_0^{\mathrm{line}} \cap E$, where $E$ is the boundary of $\Omega$.
```

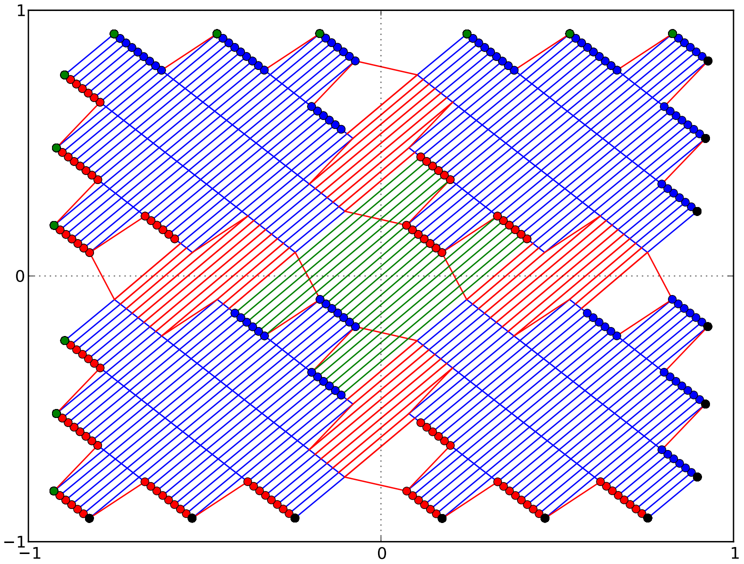


Figure 3.9: Connections across using Setting 3. The colored dots are the non-connected vertices already shown in Figure 3.8.

Setting 4 is applied when the sub-square $S_i$ has one or two of its sides as a part of the boundary of $\Omega$, and where Setting 1, 2-a, 2-b and Setting 3 cannot be applied.

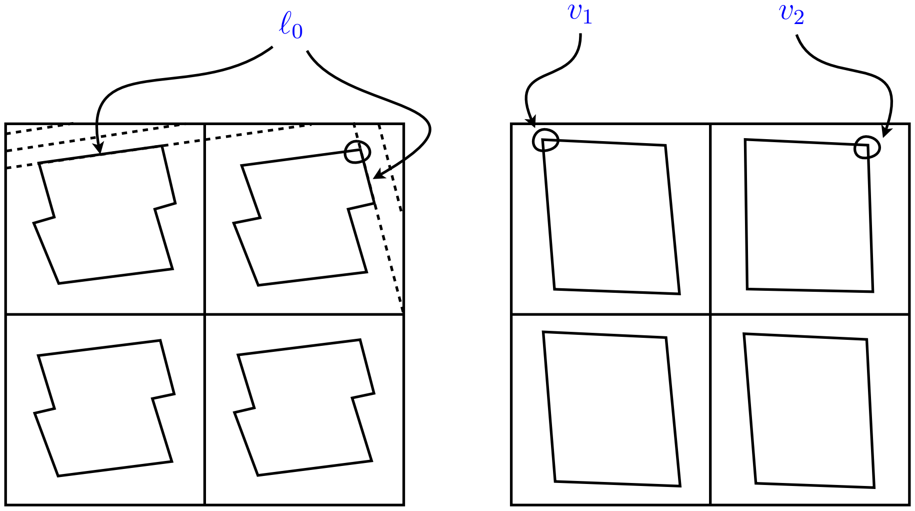
Suppose now that all the segments in $\mathcal{L}_i \cup \bar{\mathcal{L}}_i$, for $i = 1, \ldots, m^2$, are extended by using the Settings 1-4. In particular, by Setting 4, every end-segment is extended from both of its end-point. This can create big area irregular regions in the neighborhood of the corners that are near the end-points, more precisely, at the corners of the domain $\Omega$.

The following Setting 5 will partition the irregular regions in the neighborhood of the four corners $c_k$, $k = 1, 2, 3, 4$, of $\Omega$. For each corner $c_k$, consider the sub-square $S_i$ with a vertex at $c_k$. There is at most one end-segment $\ell_{0,k}$ of $\mathcal{L}_i \cup \bar{\mathcal{L}}_i$ possessing the following properties (see left of Figure 3.10):

(1) one of the two end-points of $\ell_{0,k}$ is the vertex of $R_i$ which is the closest to the corner $c_k$;

(2) the straight line extending $\ell_{0,k}$ does not intersect two opposite sides of $S_i$.

Note that such a segment does not necessarily exist. In the right hand side of



(a) End-segments in the neighborhood of corners

(b) Vertices $v_1, v_2$ belonging to two end-segments

Figure 3.10: (a) Splittings by using Setting 4-5. (b) No splitting from Setting 5

Figure 3.10, on the top left, the vertex $v_1$ belongs to two end-segments, yet none of the two satisfies the above conditions.

If such an end-segment exists (see left of Figure 3.10), denote by $z_{1,k}$ and $z_{2,k}$ the intersection points of the line extending $\ell_0$ with the boundary of $\Omega$. Then define the triangle $P_k$ by the points $c_k, z_{1,k}, z_{2,k}$.

The condition (2) is to avoid the following situation: In the case where the sides of parallelograms are nearly parallel to the sides of the sub-square that contains it, $P_k$ is the trapezoid defined by $c_k, c_{k'}, z_{1,k}, z_{2,k}$ (see right of Figure 3.10), where $c_{k'}$ is the corner of $\Omega$ which is the closest to one of $z_{1,k}, z_{2,k}$. Note that such a trapezoid may create undesirable and uncontrollable irregular regions.

We use Setting 5 below to partition $P_k$, see Figure 3.10.

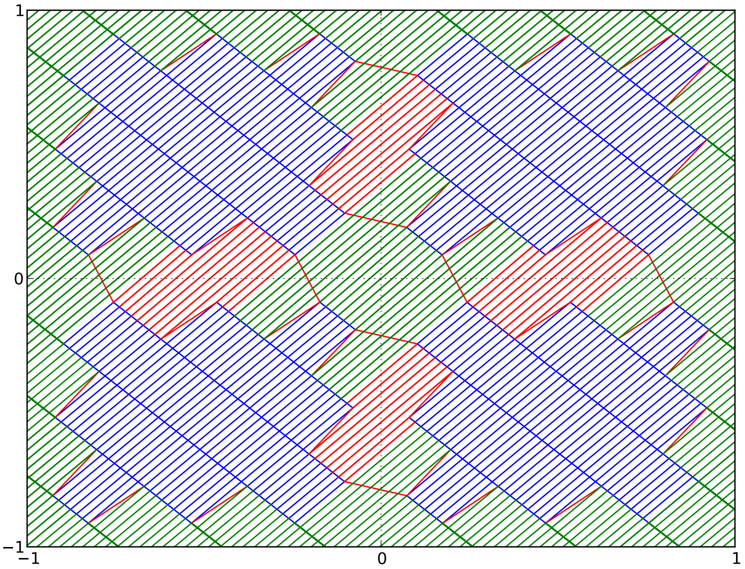


Figure 3.11: Splittings after Setting 3-4 and 5.

**Setting 5: In the neighbourhood of the corners of $\Omega$.**

Given a corner $c_k$ of $\Omega$, suppose there exists an end-segment $\ell_{0,k}$ satisfying the properties (1), (2) above and let $P_k$ be the above defined triangle associated with $c_k$;

**then**

> From $\ell_{0,k}$ up to $c_k$, partition $P_k$ by using straight line segments parallel to $\ell_{0,k}$ and equidistant by $d_i$ if $\ell_0$ is parallel to $\bar{\mathbf{e}}_i$, otherwise by $h_i$. Here $d_i, h_i$ denote the lengths of the smallest and largest edge of the parallelogram containing $\ell_{0,k}$.

The above setting, together with Setting 1, 2-a, 2-b, 3 and 4, concludes the partition of $\Omega^{\text{irr}}$ into irregular polygons (see Figure 3.11).

By construction, irregular polygons are obtained from the intersections of segments belonging to one of the three families of segments formed by:

- the four boundary edges of $\Omega$;
- the extended segments obtained from Settings 1-5 by using either of the directions $\pm\mathbf{e}_i$, $i = 1, \ldots, m^2$;

- the extended segments obtained from Settings 1-5 by using either of the directions $\pm\bar{\mathbf{e}}_i$, $i = 1, \dots, m^2$.

Each irregular polygon has at most two edges from the same family. To prove this, we shall define *polygonal chains* from the three families of segments described above. A polygonal chain $\zeta$ cutting a sub-square $S_i$ into two parts is a collection of line segments generated by any segment $\ell_0 \in \mathcal{L}_i \cup \bar{\mathcal{L}}_i$ as follows:

Case 1: A segment obtained from Setting 5 (in the neighborhood of the corners of $\Omega$) is a polygonal chain which does not contain any part of a regular region. It is a straight line segment whose end-points belong to the boundary of $\Omega$;

Case 2: Otherwise, the extensions of $\ell_0$ by Settings 1-4 are only using its two end-points $z_0, z_0'$ and the directions $\mathbf{e}_i$ and $-\mathbf{e}_i$ (a similar construction is used for a polygonal chain whose first segment belongs to $\bar{\mathcal{L}}_i$, with directions of extensions $\bar{\mathbf{e}}_i$ and $-\bar{\mathbf{e}}_i$). There are three cases for which $z_0$ is connected to, either to a vertex $z_1 \in R_i$ (Setting 1), to a vertex $z_1 \in R_j$ where $R_j$ is a neighboring regular region (Settings 2-3), or to $z_1$ which is on the boundary of $\Omega$ (Setting 4). We then add the extended segment $[z_0, z_1]$ to the polygonal chain $\zeta$. In the case where $z_1 \in R_i$, there is a segment $\ell_0' \in \mathcal{L}_i$ which contains $z_1$: we then add $[z_1, z_1']$ to the polygonal chain $\zeta$ where $z_1'$ is the end-point of $\ell_0'$ in the direction of $\mathbf{e}_i$. Originating from the point $z_1'$ we extend $\ell_0'$ by using one of the Settings 1-4 and keep repeating this procedure until one of the extended segments intersects the boundary of the sub-square $S_i$. We do the same procedure of extension for the other end-point $z_0'$ of $\ell_0$, and add to $\zeta$ the obtained segments.

The polygonal chains described above, see Figure 3.12, are divided into two groups:

- The group $\mathcal{P}_i$ where a polygonal chain contains only two types of segments: some part of segments of $\mathcal{L}_i$ and some extended segments using the directions of $\mathbf{e}_i$ and $-\mathbf{e}_i$;

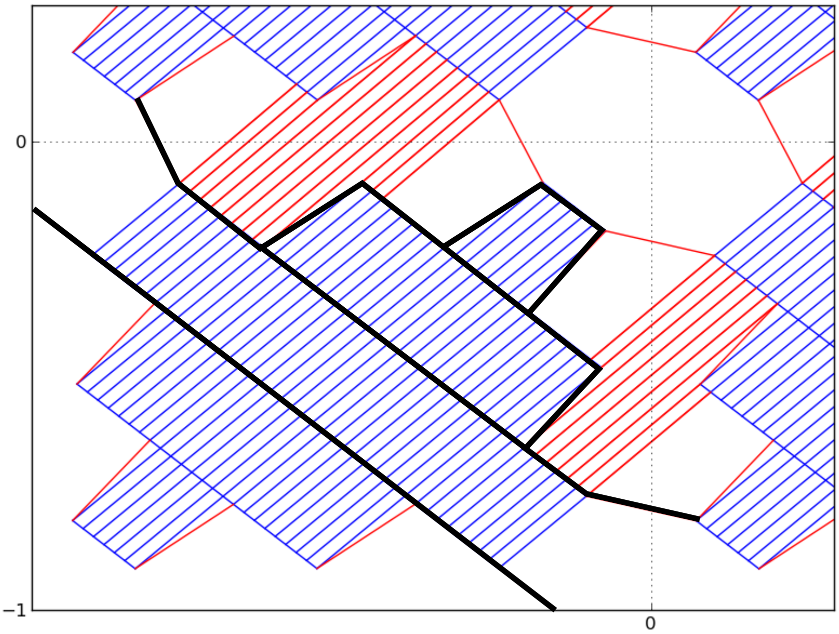


Figure 3.12: Examples of polygonal chains (in black) that have common segments.

- The group $\bar{\mathcal{P}}_i$ where a polygonal chain contains only two types of segments: some part of segments of $\bar{\mathcal{L}}_i$ and some extended segments using the directions of $\bar{\mathbf{e}}_i$ and $-\bar{\mathbf{e}}_i$;

A segment of a polygonal chain can be intersected by another one from a different group. The properties below are also observed, they follow from the results in Lemmas 3.3.4, 3.3.5 and 3.3.6 which are proved in Section 3.3.2:

$(i)$ If two polygonal chains from the same group intersect at a point $z$, then $z$ is the end-point of a common segment of the two polygonal chains. That segment is a part of a segment in $\mathcal{L}_i \cup \bar{\mathcal{L}}_i$ or part of an extended segment;

$(ii)$ Two polygonal chains from two different groups cannot have a common segment, but they can intersect;

$(iii)$ Each vertex of a polygonal chain is the end-point of at least two segments of some polygonal chains from different groups.

We have the result below.

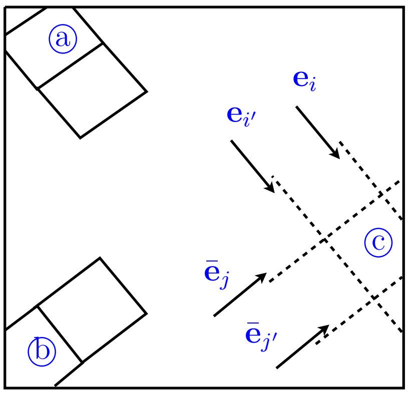


Figure 3.13: Interface polygons. In ⓐ a hexagon which can only appear in the neighborhood of a corner; in ⓑ a pentagon possessing two edges overlapping the boundary; and in ⓒ a pentagon obtained from intersections of segments belonging to the above three families.

**Lemma 3.2.5.** *Let $r$ be sufficiently small so that (3.17) holds. Then, each irregular polygon $P$ has at most two edges from the same family, it cannot have more than six edges, of which at most four lie in the interior of $\Omega$ and at most two on its boundary.*

*Proof.* Let $P$ be an irregular polygon. Since the extension of any segment never crosses into a non-neighboring sub-square, $P$ is contained in the union $\tilde{S}$ of at most four sub-squares with a common vertex. The set $\tilde{S}$ is divided into two subsets by any polygonal chain $\zeta$ obtained by extending (if necessary) a polygonal chain $\zeta_0$ from one sub-square up to the neighboring sub-square that intersects $\zeta_0$: From the above construction of a polygonal chain, there is an extended segment $[z_0, z_1]$ of $\zeta_0$ which intersects the boundary of $S_i$. If $z_1$ is on the boundary (i.e. obtained from Setting 4), then we do not extend $\zeta_0$. Otherwise, that is if $z_1 \in R_j$, we add to $\zeta_0$ the segment of $\mathcal{L}_j$ having $z_1$ as an end-point, and extend the other end-point in the direction of $\mathbf{e}_j$ by using the same method as for obtaining a polygonal chain. The only case where the resulting polygonal chain may not divide the four sub-squares is when its end-point still belong to the union of the sub-squares. In this case, we apply a second time the extension of $\zeta_0$ until it divides the set $\tilde{S}$.

Thus, we obtain two families $\mathcal{P}$ and $\bar{\mathcal{P}}$ of polygonal chains in $\tilde{S}$. The partition of $\tilde{S}$ into regular parallelograms and irregular polygons is generated by the families

$\mathcal{P}$ and $\bar{\mathcal{P}}$, where each cell of the partition is obtained by the intersection of a stripe between a pair of consecutive chains of $\mathcal{P}$ with a stripe between a pair of consecutive chains of $\bar{\mathcal{P}}$. In particular, at most two edges of $P$ are formed by the segments of $\mathcal{P}$, and at most two edges by the segments of $\bar{\mathcal{P}}$. In addition, at most two more edges can be obtained from the boundary of $\Omega$ as illustrated in Figure 3.13. □

### 3.2.4 Back transformation and triangulation

Given $i \in \{1, \dots, m^2\}$, consider the invertible affine map $\psi_i(x, y) := \varphi_i(x, y) + \mathbf{t}_{b_i}$ where $\varphi_i$ is defined in (3.39) and $\mathbf{t}_{b_i}$ is the position vector of $b_i$ which is the barycenter of the sub-square $S_i$. Observe that $\psi_i$ maps the equilateral triangle described in Algorithm 3.1 to the near-optimal triangle with a vertex at $b_i$ (see (i) of Section 3.2.2).

A *back transformation* of a polygon $P$ is the polygon $\psi_i^{-1}(P)$ for some $i \in \{1, \dots, m^2\}$. If $P$ is a fixed polygon, we can choose $i$ as the index of a sub-square $S_i$ which is one of the closest sub-squares to the barycenter $b_P$ of $P$.

**Triangulation of $\Omega$.** The domain $\Omega$ being partitioned into polygons of at most six edges, the triangulation of $\Omega$ which we denote by $\Delta_{s,r}$, is obtained as follows: Triangles remain the same, whereas each quadrilateral which does not have a vertex on the boundary of $\Omega$ is divided by drawing a diagonal crossing its largest interior angle. For a quadrilateral $P$ possessing a vertex on the boundary of $\Omega$, we draw a diagonal $\ell$ such that $\psi_i^{-1}(\ell)$ is a diagonal crossing the largest interior angle of $\psi_i^{-1}(P)$, with $i$ being the index of one of the closest sub-squares to the barycenter of $P$. Similarly, a pentagon $P$ possessing a vertex on the boundary is divided by two diagonals whose images by $\psi_i^{-1}$ cross the largest interior angles of $\psi_i^{-1}(P)$. A hexagon $P$ is divided into four triangles by three diagonals whose images by $\psi_i^{-1}$ cross the largest interior angle of $\psi_i^{-1}(P)$.

Triangles obtained by dividing the parallelograms of the regular regions are called regular triangles, the rest of the triangles are called *irregular* triangles. The resulting triangulation of $\Omega$ is denoted by $\Delta_{s,r}$.

We can also consider the back transformation of a triangle $T$. Given a fixed index $i \in \{1, \ldots, m^2\}$, consider the triangulation $\Delta_i$ given by

$$\Delta_i = \{\psi_i^{-1}(T) : T \in \Delta_{s,r}\}, \tag{3.44}$$

which is called a *back transformation* of $\Delta_{s,r}$. Since $R_i$ is uniformly triangulated and $\psi$ an affine map, by virtue of Algorithm 3.1 and step (ii) of Section 3.2.2, its image $\psi_i^{-1}(R_i)$ is uniformly triangulated with equilateral triangles which are shifted versions of $\Lambda_i T_0$, with $\Lambda_k$ defined in (3.43), and $T_0$ being the equilateral triangle of unit area as described in Algorithm 3.1.

## 3.3 Properties of the triangulation

By construction, regular triangles are isosceles and their alignment directions are as specified by Algorithm 3.1 of the previous section. We analyze the change of these directions when segments are extended according to Settings 1-4. The segments obtained from Setting 5 are left over since they are parallel to the edges of some parallelograms. In particular, we discuss the properties of the intersections of extended segments which produce irregular triangles. Note that unlike regular triangles, irregular triangles may possess various shapes, this is where the "back transformation" helps out to bound the $W_p^1$-seminorm of the interpolation error.

### 3.3.1 Conformity of alignment directions

Following the discussion in Section 3.2.1 about the sensitivity of eigenvectors, given two neighbor squares $R_i$ and $R_j$, where $i, j \in \{1, \ldots, m^2\}$, we investigate the angles formed from the associated systems of segments $\mathcal{L}_i, \bar{\mathcal{L}}_i$ and $\mathcal{L}_j, \bar{\mathcal{L}}_j$. Recall that they determine the alignment directions of the parallelograms in $R_i$

and $R_j$, and we here study the conformity of these alignments in terms of the difference of angles that they form and using their respective directional vectors $\mathbf{e}_i, \bar{\mathbf{e}}_i$ and $\mathbf{e}_j, \bar{\mathbf{e}}_j$ which are as defined in (iii) of Section 3.2.2.

Recall that $b_i$ is the barycenter of the sub-square $S_i$, with $i \in \{1, \ldots, m^2\}$, and $\pi_{b_i} \in \mathbb{H}_2$ the quadratic polynomial whose coefficients are the second derivatives of $f$ at the point $b_i$. We denote by $\mu_i$ the angle of rotation of the matrix $U_{\pi_{b_i}}$ which is as defined in (2.2). Then the smallest angle $\bar{\theta}_{ij}$ between $\bar{\mathbf{e}}_i$ and $\bar{\mathbf{e}}_j$ is exactly the difference $|\mu_i - \mu_j|$. Taking into account the ill-conditioning of eigenvectors discussed in Section 3.2.1, this difference can be large if the eigenvalues $\lambda_{1,i}, \lambda_{2,i}$ associated with $\pi_{b_i}$ are close to one another. We assume that $|\lambda_{1,i}| \leq |\lambda_{2,i}|$.

Apart from the standard estimation (3.19) in Lemma 3.2.1, the difference of eigenvalues $|\lambda_{2,i} - \lambda_{1,i}|$ can be estimated.

**Lemma 3.3.1.** *Suppose that* $\max\{|\lambda_{1,i} - \lambda_{2,j}|, |\lambda_{2,i} - \lambda_{1,j}|\} \leq \varepsilon$. *If all the eigenvalues* $\lambda_{1,i}, \lambda_{2,i}$ *and* $\lambda_{1,j}, \lambda_{2,j}$ *are positive, then*

$$\max\{|\lambda_{2,i} - \lambda_{1,i}|, |\lambda_{2,j} - \lambda_{1,j}|\} < 2\varepsilon.$$

*Proof.* The fact that the distance between $\lambda_{2,j}$ and $\lambda_{1,i}$ is less than $\varepsilon$, that the distance between $\lambda_{2,i}$ and $\lambda_{1,j}$ is less than $\varepsilon$, and that $\lambda_{2,j}$ is greater than $\lambda_{1,j}$, we necessarily have that the distance between $\lambda_{2,i}$ and $\lambda_{1,i}$ is strictly less than $2\varepsilon$. Similar approach is used to prove that $|\lambda_{2,j} - \lambda_{1,j}| < 2\varepsilon$. □

Recall that $Q_\pi$ is the matrix associated (defined in (2.1)) with a homogeneous quadratic polynomial $\pi$. By using (2.4) and (2.75), we obtain

$$\|Q_{\pi_{b_i}} - Q_{\pi_{b_j}}\|_2 \leq \frac{3}{2}\omega(\sqrt{2}r), \tag{3.45}$$

where $\omega$ is defined in (2.75) and $r$ is chosen to so that (3.17) holds.

The constant $C_{\delta_f}$ in (3.15) originates from the result below. Recall that, as in Section 3.2.2, we assume that the Hessian $H_f$ is positive definite. Also, in view of (3.45), a natural choice for $\varepsilon$ is that $\varepsilon = \frac{3}{2}\omega(\sqrt{2}r)$.

**Proposition 3.3.2.** *Let $S_i$ and $S_j$ be neighboring sub-squares. If $\|Q_{\pi_{b_i}} - Q_{\pi_{b_j}}\|_2 \leq \varepsilon$ for a sufficiently small number $\varepsilon > 0$, then the angle $\bar{\theta}_{ij} \in [0, \pi)$ between $\bar{\mathbf{e}}_i$ and $\bar{\mathbf{e}}_j$ satisfies*

$$\bar{\theta}_{ij} \leq C_1 \varepsilon^{\frac{1}{2}} + O(\varepsilon^2), \tag{3.46}$$

*where $C_1 = 33\pi$ is a constant. Also, the angle $\theta_{ij}$ between $\mathbf{e}_i$ and $\mathbf{e}_j$ satisfies*

$$\theta_{ij} \leq C_{\delta_f} \varepsilon^{\frac{1}{2}} + O(\varepsilon^2), \tag{3.47}$$

*where $C_{\delta_f} = 33\pi + 2\delta_f^{-\frac{1}{2}}$ is a constant depending on $\delta_f$, with $\delta_f$ being defined in (3.13).*

*Proof.* Let $\mathbf{v}_{1,i}, \mathbf{v}_{2,i}$ be the unit eigenvectors corresponding to the eigenvalues $\lambda_{1,i}, \lambda_{2,i}$ of $\pi_{b_i}$. In view of Proposition 3.2.4, for a given small number $\varepsilon > 0$, we shall distinguish the cases where $|\lambda_{2,i} - \lambda_{1,i}| \leq \varepsilon^{\frac{1}{2}}$ and $|\lambda_{2,i} - \lambda_{1,i}| \geq \varepsilon^{\frac{1}{2}}$. By this we can distinguish whether or not the optimal triangles for $\pi_{b_i}$ and $\pi_{b_j}$ need the adjustment angles described in Proposition 3.2.4. We have three cases:

i. If $|\lambda_{2,i} - \lambda_{1,i}| \geq \varepsilon^{\frac{1}{2}}$ and $|\lambda_{2,j} - \lambda_{1,j}| \geq \varepsilon^{\frac{1}{2}}$. Since $\frac{2}{\pi}x \leq \sin x$ for all $x \in [0, \frac{\pi}{2}]$, we deduce from (3.20) and (3.22) that

$$\begin{aligned}\bar{\theta}_{ij} &\leq \frac{\pi}{2} \max\{\|\mathbf{v}_{1,i} - \mathbf{v}_{1,j}\|_2, \|\mathbf{v}_{2,i} - \mathbf{v}_{2,j}\|_2\} \\ &\leq \frac{\pi}{2} \varepsilon^{-\frac{1}{2}} \|Q_{\pi_{b_i}} - Q_{\pi_{b_j}}\|_2 + O(\|Q_{\pi_{b_i}} - Q_{\pi_{b_j}}\|_2^2) \\ &\leq \frac{\pi}{2} \varepsilon^{\frac{1}{2}} + O(\varepsilon^2).\end{aligned}$$

ii. If $|\lambda_{2,i} - \lambda_{1,i}| \leq \varepsilon^{\frac{1}{2}}$ and $|\lambda_{2,j} - \lambda_{1,j}| \leq \varepsilon^{\frac{1}{2}}$. We deduce from (3.28) of Proposition 3.2.4 that $\bar{\theta}_{ij} \leq C_1 \varepsilon^{\frac{1}{2}} + O(\varepsilon^2)$, with $C_1 = 33\pi$ according to (3.38).

iii. If $|\lambda_{2,i} - \lambda_{1,i}| \leq \varepsilon^{\frac{1}{2}}$ and $|\lambda_{2,j} - \lambda_{1,j}| \geq \varepsilon^{\frac{1}{2}}$. With $\vartheta_i$ being the adjustment angle for an optimal triangle for $\pi_{b_i}$, we have $\mu_i - \vartheta_i = \frac{2\mu_i |\lambda_{2,i} - \lambda_{1,i}|}{\varepsilon^{\frac{1}{2}} + |\lambda_{2,i} - \lambda_{1,i}|}$ and

$$\begin{aligned}\bar{\theta}_{ij} = |(\mu_i - \vartheta_i) - \mu_j| &= \left|\frac{2\mu_i|\lambda_{2,i}-\lambda_{1,i}| - \mu_j\varepsilon^{\frac{1}{2}} - \mu_j|\lambda_{2,i}-\lambda_{1,i}|}{\varepsilon^{\frac{1}{2}} + |\lambda_{2,i}-\lambda_{1,i}|}\right| \\ &\leq \frac{\lambda|\mu_i-\mu_j| + |\mu_i\lambda - \mu_j\varepsilon^{\frac{1}{2}}|}{\varepsilon^{\frac{1}{2}}+\lambda}\end{aligned}$$

where $\lambda = |\lambda_{2,i}-\lambda_{1,i}|$. Observe that, since $\frac{2}{\pi}x \leq \sin x$ for all $x \in [0, \frac{\pi}{2}]$, we deduce from (3.20) that $|\mu_i - \mu_j| \leq \frac{\pi}{2}\|\mathbf{v}_i - \mathbf{v}_j\|_2$ which together with (3.22), yields

$$|\lambda_{2,i}-\lambda_{1,i}||\mu_i-\mu_j| \leq \frac{\pi}{2}\|Q_{\pi_{b_i}} - Q_{\pi_{b_j}}\|_2 + \frac{\pi}{2}|\lambda_{2,i}-\lambda_{1,i}|O(\|Q_{\pi_{b_i}} - Q_{\pi_{b_j}}\|_2^2).$$

Also, by denoting $\lambda = |\lambda_{2,i}-\lambda_{1,i}|$, it is clear that

$$\begin{aligned}|\mu_i\lambda - \mu_j\varepsilon^{\frac{1}{2}}| &= |\mu_i\lambda - \mu_j\varepsilon^{\frac{1}{2}} + \mu_j\lambda - \mu_j\lambda| \\ &\leq |\mu_i-\mu_j|\lambda + \mu_j|\varepsilon^{\frac{1}{2}} - \lambda|,\end{aligned}$$

where, with $\lambda' := |\lambda_{2,j}-\lambda_{1,j}| \geq \varepsilon^{\frac{1}{2}}$ and noting that (3.33) holds, clearly $|\varepsilon^{\frac{1}{2}} - \lambda| \leq |\lambda' - \lambda| \leq 2\varepsilon$. We thus obtain

$$\begin{aligned}\bar{\theta}_{ij} &\leq \frac{2\varepsilon + \varepsilon^{\frac{1}{2}}O(\varepsilon^2) + |\mu_i-\mu_j|\lambda + \mu_j\varepsilon}{\varepsilon^{\frac{1}{2}}} \leq \frac{8\varepsilon + 2\varepsilon^{\frac{1}{2}}O(\varepsilon^2) + 2\pi\varepsilon}{\varepsilon^{\frac{1}{2}}} \\ &= \frac{(8+2\pi)\varepsilon + \varepsilon^{\frac{1}{2}}O(\varepsilon^2)}{\varepsilon^{\frac{1}{2}}} = (8+2\pi)\varepsilon^{\frac{1}{2}} + O(\varepsilon^2),\end{aligned}$$

thereby proving (3.46).

Let $h_0$ and $\rho_0$ denote the diameter and smallest height of an equilateral triangle of unit area. Then, $\mathbf{e}_i$ makes with the $x$-axis the angle $\mu_i - \vartheta_i - \alpha_i$, where the angle $\alpha_i$ is the initial angle which $\mathbf{e}_i$ makes with the $x$-axis before rotation to the alignment direction $\mu_i - \vartheta_i$, see Figure 3.2, with

$$\tan\alpha_i = \left(\frac{h_0}{2}\left|\frac{\lambda_{1,i}}{\lambda_{2,i}}\right|^{\frac{1}{4}}\right)\Big/\left(\rho_0\left|\frac{\lambda_{2,i}}{\lambda_{1,i}}\right|^{\frac{1}{4}}\right) = \frac{h_0}{2\rho_0}\left|\frac{\lambda_{1,i}}{\lambda_{2,i}}\right|^{\frac{1}{2}}. \tag{3.48}$$

To estimate the angle $\theta_{ij}$ between $\mathbf{e}_i$ and $\mathbf{e}_j$, we shall only need to estimate the difference of angle $|\alpha_j - \alpha_i|$ since

$$\begin{aligned}\theta_{ij} &= |(\mu_i - \vartheta_i - \alpha_i) - (\mu_j - \vartheta_j - \alpha_j)| \\ &\leq |(\mu_i - \vartheta_i) - (\mu_j - \vartheta_j)| + |\alpha_j - \alpha_i|, \\ &\leq \bar{\theta}_{ij} + |\alpha_j - \alpha_i|, \qquad (3.49)\end{aligned}$$

with $\bar{\theta}_{ij}$ being estimated by (3.46). Observe that $0 < \alpha_i \leq \frac{\pi}{6}$ always holds. Since $x \leq \tan x$ on $[0, \frac{\pi}{2}]$, clearly $|\alpha_j - \alpha_i| \leq \tan|\alpha_j - \alpha_i|$. Hence, from the fact that

$$\tan|\alpha_j - \alpha_i| = \frac{|\tan\alpha_j - \tan\alpha_i|}{1 + \tan\alpha_j \tan\alpha_i} \leq |\tan\alpha_j - \tan\alpha_i|,$$

combined with (3.48), we deduce that $|\alpha_j - \alpha_i|$ satisfies

$$\begin{aligned}|\alpha_j - \alpha_i| &\leq \frac{h_0}{2\rho_0}\left|\left|\frac{\lambda_{1,j}}{\lambda_{2,j}}\right|^{\frac{1}{2}} - \left|\frac{\lambda_{1,i}}{\lambda_{2,i}}\right|^{\frac{1}{2}}\right| \leq \frac{h_0}{2\rho_0}\left|\frac{\lambda_{1,j}}{\lambda_{2,j}} - \frac{\lambda_{1,i}}{\lambda_{2,i}}\right|^{\frac{1}{2}} \\ &\leq \frac{h_0}{2\rho_0}\left|\frac{\lambda_{1,j}\lambda_{2,i} - \lambda_{1,i}\lambda_{2,j}}{\lambda_{2,i}\lambda_{2,j}}\right|^{\frac{1}{2}}, \qquad (3.50)\end{aligned}$$

which we shall estimate. We prove that $|\lambda_{1,j}\lambda_{2,i} - \lambda_{1,i}\lambda_{2,j}| \leq 4\varepsilon \max\{|\lambda_{2,i}, \lambda_{2,j}|\}$ as follows. By Lemma 3.2.1, there exist $k, k' \in \{1, 2\}$ such that $\max\{|\lambda_{1,i} - \lambda_{k,j}|, |\lambda_{2,i} - \lambda_{k',j}|\} \leq \varepsilon$. We now have the following cases:

- If $(k, k') = (1, 1)$, then we easily prove that $|\lambda_{2,j} - \lambda_{2,i}| \leq 3\varepsilon$: If $\lambda_{2,i}$ is the closest eigenvalue to $\lambda_{2,j}$, then $|\lambda_{2,j} - \lambda_{2,i}| \leq \varepsilon \leq 3\varepsilon$. Otherwise, if $\lambda_{1,i}$ is the closest eigenvalue to $\lambda_{2,j}$ so that $|\lambda_{2,j} - \lambda_{1,i}| \leq \varepsilon$, then a simple triangular inequality shows that

  $$|\lambda_{2,j} - \lambda_{2,i}| \leq |\lambda_{2,j} - \lambda_{1,i}| + |\lambda_{1,i} - \lambda_{1,j}| + |\lambda_{1,j} - \lambda_{2,i}| \leq 3\varepsilon,$$

  from which we deduce that

  $$\begin{aligned}|\lambda_{1,j}\lambda_{2,i} - \lambda_{1,i}\lambda_{2,j}| &\leq |\lambda_{2,i}||\lambda_{1,j} - \lambda_{1,i}| + |\lambda_{1,i}||\lambda_{2,j} - \lambda_{2,i}| \\ &\leq 4\varepsilon \max\{|\lambda_{2,i}|, |\lambda_{2,j}|\};\end{aligned}$$

- If $(k,k')=(2,2)$, using the same argument as above shows that $|\lambda_{1,j}-\lambda_{1,i}|\leq 3\varepsilon$: If $\lambda_{1,i}$ is the closest eigenvalue to $\lambda_{1,j}$, then $|\lambda_{1,j}-\lambda_{1,i}|\leq\varepsilon\leq 3\varepsilon$. Otherwise, if $\lambda_{2,i}$ is the closest to $\lambda_{1,j}$ so that $|\lambda_{1,j}-\lambda_{2,i}|\leq\varepsilon$, then a simple triangular inequality shows that

  $$|\lambda_{1,j}-\lambda_{1,i}|\leq|\lambda_{1,j}-\lambda_{2,i}|+|\lambda_{2,i}-\lambda_{2,j}|+|\lambda_{2,j}-\lambda_{2,i}|\leq 3\varepsilon,$$

  and we have the following,

  $$\begin{aligned}|\lambda_{1,j}\lambda_{2,i}-\lambda_{1,i}\lambda_{2,j}| &\leq |\lambda_{2,i}||\lambda_{1,j}-\lambda_{1,i}|+|\lambda_{1,i}||\lambda_{2,i}-\lambda_{2,j}|\\ &\leq 4\varepsilon\max\{|\lambda_{2,i}|,|\lambda_{2,j}|\};\end{aligned}$$

- If $(k,k')=(1,2)$, then we write

  $$\begin{aligned}|\lambda_{1,j}\lambda_{2,i}-\lambda_{1,i}\lambda_{2,j}| &= |\lambda_{1,j}(\lambda_{2,i}-\lambda_{k',j})+\lambda_{2,j}(\lambda_{k,j}-\lambda_{1,i})|\\ &\leq 2\varepsilon\max\{|\lambda_{2,i}|,|\lambda_{2,j}|\};\end{aligned}$$

- The case $(k,k')=(2,1)$ means that $|\lambda_{1,i}-\lambda_{2,j}|\leq\varepsilon$ and $|\lambda_{2,i}-\lambda_{1,j}|\leq\varepsilon$. Since all eigenvalues are positive, we deduce from Lemma 3.3.1 that $\max\{|\lambda_{2,i}-\lambda_{1,i}|,|\lambda_{2,j}-\lambda_{1,j}|\}<2\varepsilon$. This implies that

  $$|\lambda_{2,j}-\lambda_{2,i}|\leq|\lambda_{2,j}-\lambda_{1,i}|+|\lambda_{2,i}-\lambda_{1,i}|\leq 3\varepsilon.$$

  It follows that

  $$|\lambda_{1,j}\lambda_{2,i}-\lambda_{1,i}\lambda_{2,j}|\leq|\lambda_{1,j}||\lambda_{2,i}-\lambda_{2,j}|+|\lambda_{2,j}||\lambda_{1,j}-\lambda_{1,i}|\leq 4\varepsilon\max\{|\lambda_{2,i}|,|\lambda_{2,j}|\}.$$

Combining all these cases, we prove that

$$|\lambda_{1,j}\lambda_{2,i}-\lambda_{1,i}\lambda_{2,j}|\leq 4\varepsilon\max\{|\lambda_{2,i}|,|\lambda_{2,j}|\}. \tag{3.51}$$

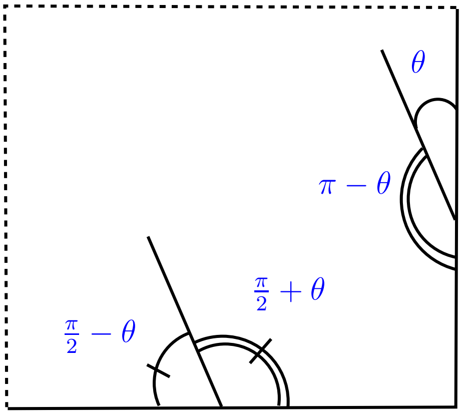


Figure 3.14: Angles formed from the intersections of a segment to the vertical and horizontal lines.

We easily deduce from (3.13) that

$$\left|\frac{\lambda_{1,j}\lambda_{2,i}-\lambda_{1,i}\lambda_{2,j}}{\lambda_{2,i}\lambda_{2,j}}\right|^{\frac{1}{2}} \leq 2\varepsilon^{\frac{1}{2}}\frac{\max\{|\lambda_{2,i}|,|\lambda_{2,j}|\}^{\frac{1}{2}}}{|\lambda_{2,i}\lambda_{2,j}|^{\frac{1}{2}}} \leq 2\delta_f^{-\frac{1}{2}}\varepsilon^{\frac{1}{2}}. \tag{3.52}$$

It follows from (3.50) that

$$|\alpha_j-\alpha_i| \leq \frac{h_0}{\rho_0}\delta_f^{-\frac{1}{2}}\varepsilon^{\frac{1}{2}},$$

which, together with (3.49) and the fact that $\frac{h_0}{\rho_0}=\frac{2}{\sqrt{3}}\leq 2$, proves the estimation (3.47) with the constant $C_{\delta_f}=33\pi+2\delta_f^{-\frac{1}{2}}\geq C_1+\frac{h_0}{\rho_0}\delta_f^{-\frac{1}{2}}$. □

Note that a straight line segment forms four angles with the vertical and horizontal lines (see Figure 3.14). The angles belong to the following two sets $\Theta_i,\bar{\Theta}_i$, $i=1,\dots,m^2$, obtainable by replacing $\theta$ in Figure 3.14 by $\mu_i-\vartheta_i$ and $\mu_i-\alpha_i-\vartheta_i$:

$$\Theta_i := \{\mu_i-\vartheta_i,\ \frac{\pi}{2}-\mu_i+\vartheta_i,\ \pi-\mu_i+\vartheta_i,\ \frac{\pi}{2}+\mu_i-\vartheta_i\}; \tag{3.53}$$

$$\bar{\Theta}_i := \{\mu_i-\alpha_i-\vartheta_i,\ \frac{\pi}{2}-\mu_i+\alpha_i+\vartheta_i,\ \pi-\mu_i+\alpha_i+\vartheta_i,\ \frac{\pi}{2}+\mu_i-\alpha_i-\vartheta_i\}. \tag{3.54}$$

where $\mu_i$ is the angle of rotation of $Q_{\pi_i}$, whereas $\vartheta_i$ is the corresponding angle of adjustment. For any neighboring sub-square $S_j$, the angles may change by at most $\max\{\theta_{ij}, \bar{\theta}_{ij}\}$.

We formulate the following statement, obvious from Figure 3.2.

**Lemma 3.3.3.** *For each $i = 1, \ldots, m^2$, the interior angles of a regular triangle $T \in R_i$ are exactly $2\alpha_i$ and $\beta_i$ where $\alpha_i$ is defined by (3.48) and $\beta_i$ by*

$$\tan \beta_i = \frac{1}{\tan \alpha_i}. \tag{3.55}$$

*Consequently, the interior angles of any parallelogram obtained by using step (ii) or (iii.) of Section 3.2.2 are exactly $2\alpha_i + \beta_i$ and $\beta_i$.*

### 3.3.2 Intersections of extended segments

In this section, we present the properties of extended segments obtained by Settings 1-4. Recall that their intersections create polygons of at most five edges which we call irregular polygons.

By virtue of Proposition 3.3.2, by making the parameter $r$ small enough, the segments in $\mathcal{L}_i$ and $\mathcal{L}_j$, and in $\bar{L}_i$ and $\bar{L}_j$ for neighboring regular regions $R_i, R_j$, are "almost" parallel in the sense that the angle between them does not exceed $\theta$, where

$$\theta^* = \max\{\theta_{ij}, \bar{\theta}_{ij} : i, j = 1, \ldots, m^2\} \tag{3.56}$$

with $\theta_{ij}$ and $\bar{\theta}_{ij}$ estimated in Proposition 3.3.2.

The result below is straightforward by construction.

**Lemma 3.3.4.** *Let $S_i, S_j$ be neighbors that share an edge. For any segment $\ell_0 \in \mathcal{L}_i \cup \bar{\mathcal{L}}_i$ extended by using Setting 1, and any $\ell_1 \in \mathcal{L}_j \cup \bar{\mathcal{L}}_j$ extended into $R_i$ by using either of Setting 2-a, 2-b or 3, the extended segments $\ell_0^{ext}$ and $\ell_1^{ext}$ can only intersect at their end-points.*

*Proof.* The result is clear since we do not allow connection to a re-entrant corner (see Figure 3.15). □

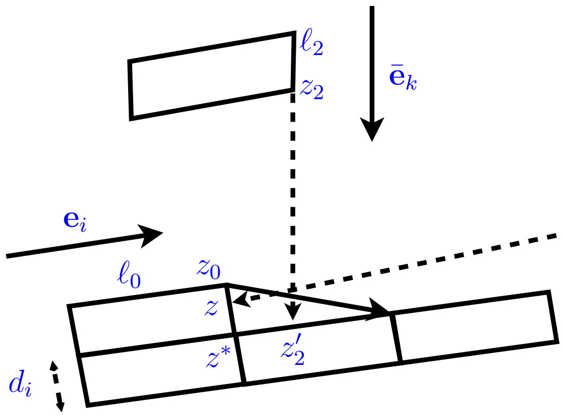


Figure 3.15: Extension of segments

**Lemma 3.3.5.** *Let $r$ be sufficiently small so that (3.17) holds. For any neighboring sub-squares $S_i, S_j$ and any two segments $\ell_0 \in \mathcal{L}_i$ (resp. $\ell_0 \in \bar{\mathcal{L}}_i$) and $\ell_1 \in \mathcal{L}_j$ (resp. $\ell_1 \in \bar{\mathcal{L}}_i$), the associated extended segments $\ell_0^{ext}, \ell_1^{ext}$ can only intersect at their end-points.*

*Proof.* We shall prove the result by contradiction. Suppose that there are two extended segments $\ell_0^{\text{ext}}, \ell_1^{\text{ext}}$ which intersect at a point which is not an end-point. Clearly, neither of $\ell_0^{\text{ext}}, \ell_1^{\text{ext}}$ can be obtained from Setting 1, 4 and 5. Indeed, this is obvious for Setting 4 and 5, and for Setting 1 we have Lemma 3.3.4.

Denote by $z$ the intersection of the lines $\ell_0^{\text{line}}, \ell_1^{\text{line}}$. Without loss of generality, assume that $z \in S_j$. Consider the triangle formed by the three points $z, z_1, \bar{z}$ (see Figure 3.16) where $z_1$ is the common end-point of $\ell_1$ and $\ell_1^{\text{ext}}$, and $\bar{z}$ the intersection point of $\ell_0^{\text{line}}$ with the segment of $\bar{L}_j$ containing $z_1$. Denote by $\gamma, \beta$ the interior angles at $z, \bar{z}$, respectively. Also, recall that $d_j, h_j$ denote the lengths of the shortest and longest edge of a parallelogram in $R_j$.

Let us first estimate the angle $\gamma$. Recall that $\lambda_{1,j}, \lambda_{2,j}$ are the eigenvalues of the matrix $Q_{\pi_j}$, and that $r$ is small enough so that (3.17) holds. Then, by combining (3.45) and (3.17) with the results in Proposition 3.3.2 with $\varepsilon = \frac{3}{2}\omega(\sqrt{2}r)$, we have

$$\begin{aligned}\gamma \le \theta \le C_{\delta_f}\varepsilon^{\frac{1}{2}} &= C_{\delta_f}[\frac{3}{2}\omega(\sqrt{2}r)]^{\frac{1}{2}} + O(\varepsilon^2) \\ &\le \frac{1}{10^5(\frac{3}{2})^{\frac{1}{2}}C_{\delta_f}|f|^{\frac{1}{2}}_{W^2_\infty(\Omega)}}\Big|\frac{\lambda_{1,j}}{\lambda_{2,j}}\Big|. \qquad (3.57)\end{aligned}$$

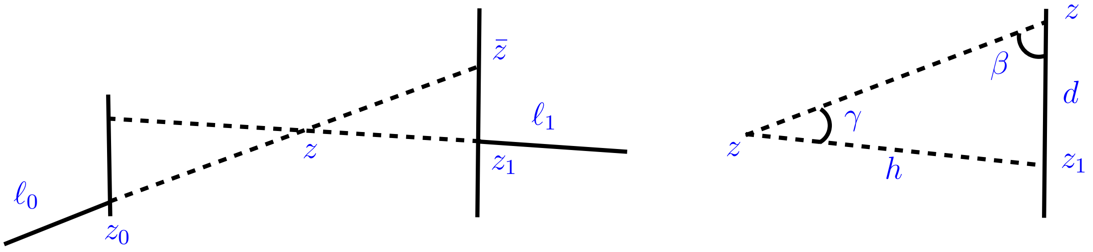


Figure 3.16: An intersection of two extended segments.

Since Setting 3 is based on Setting 2-a, 2-b, we only have three cases:

**Case 1:** both $\ell_0^{\mathrm{ext}}, \ell_1^{\mathrm{ext}}$ are obtained from Setting 2-a. Since $\ell_0 \in \mathcal{L}_i$ and $\ell_1 \in \mathcal{L}_j$, the directions of extension are two of $\mathbf{e}_i, -\mathbf{e}_i, \mathbf{e}_j$ and $-\mathbf{e}_j$. We thus have

$$|z_1 - \bar{z}| = d \geq \frac{d_j}{2}, \tag{3.58}$$

otherwise the end-points of $\ell_0^{\mathrm{ext}}$ would be $z_0$ and $z_1$. Note that $z_1$ must be on the boundary of $R_j$, thus by construction $|z - z_1| = h \geq h_j$. By using the sine rule $\frac{\sin\gamma}{d} = \frac{\sin\beta}{h}$, we obtain

$$\sin\gamma \geq \frac{d_j}{2h_j} \sin\beta. \tag{3.59}$$

We now need some lower bound for the angle $\beta$. Recall that given a parallelogram $P$ of a regular region, there is an angle $\frac{\pi}{3} \leq \vartheta \leq \frac{\pi}{2}$ such that the interior angles of $P$ are exactly $\vartheta$ and $\pi - \vartheta$. Then the angle $\beta$ is bounded by $\vartheta - \gamma$ and $\pi - \vartheta + \gamma$. By virtue of (3.57) we have that $\gamma < \frac{\pi}{6}$. Hence

$$\sin\beta \geq \sin\frac{\pi}{6}. \tag{3.60}$$

Note that $\frac{d_j}{h_j} = \frac{h_0}{2\rho_0}\left|\frac{\lambda_{1,j}}{\lambda_{2,j}}\right|^{\frac{1}{2}}$, with $\rho_0, h_0$ being the length scales of an equilateral triangle of unit area, with $h_0 = 2/3^{1/4}$ and $\rho_0 = 3^{1/4}$. We deduce from (3.59) that

$$\sin\gamma \geq \frac{h_0}{8\rho_0}\left|\frac{\lambda_{1,j}}{\lambda_{2,j}}\right|^{\frac{1}{2}} = \frac{1}{4\sqrt{3}}\left|\frac{\lambda_{1,j}}{\lambda_{2,j}}\right|^{\frac{1}{2}}, \tag{3.61}$$

which also contradicts the estimation in (3.57).

In the case where $\ell_0 \in \bar{\mathcal{L}}_i$ and $\ell_1 \in \bar{\mathcal{L}}_j$, the directions of extension are two of $\bar{\mathbf{e}}_i, -\bar{\mathbf{e}}_i, \bar{\mathbf{e}}_j$ and $-\bar{\mathbf{e}}_j$, and the values of $d, h$ change: in (3.58) we have $d \geq \frac{h_j}{2}$, but also that $h \leq d_j$. Thus the lower bound in (3.59) becomes $\frac{h_j}{2d_j} \sin\beta \geq \frac{1}{2}\sin\beta$. Combining this with (3.60) yields

$$\sin\gamma \geq \frac{1}{2}\sin\frac{\pi}{6}, \tag{3.62}$$

which also contradicts (3.57) because $\left|\frac{\lambda_{1,j}}{\lambda_{2,j}}\right| \leq 1$.

**Case 2:** one of $\ell_0^{\text{ext}}, \ell_1^{\text{ext}}$ is obtained from Setting 2-a and the other from Setting 2-b. Suppose that $\ell_0 \in \mathcal{L}_i$ and $\ell_1 \in \mathcal{L}_j$. Since $z \in S_j$, the same setting as in **Case 1** applies if $\ell_0^{\text{ext}}$ is obtained from Setting 2-a. If $\ell_0^{\text{ext}}$ is obtained from Setting 2-b, then the value of $d$ is greater than $d_j$, making the lower bound in (3.61) larger. In the case $\ell_0 \in \bar{\mathcal{L}}_i$ and $\ell_1 \in \bar{\mathcal{L}}_j$, we also have the same setting as above, and if $\ell_0^{\text{ext}}$ is obtained from Setting 2-b then $d$ is greater than $h_j$, which makes the lower bound in (3.62) larger.

**Case 3:** both $\ell_0^{\text{ext}}, \ell_1^{\text{ext}}$ are obtained from Setting 2-b. Suppose that $\ell_0 \in \mathcal{L}_i$ and $\ell_1 \in \mathcal{L}_j$. The same setting as in **Case 2** applies: the value of $d$ is greater than $d_j$, making the lower bound in (3.61) larger. If $\ell_0 \in \bar{\mathcal{L}}_i$ and $\ell_1 \in \bar{\mathcal{L}}_j$, then $d$ is greater than $h_j$, which makes the lower bound in (3.62) larger. □

**Lemma 3.3.6.** *An extended segment obtained by Setting 1 is necessarily a diagonal of a shifted version of a regular parallelogram. Moreover, two extended segments obtained by using Setting 1 cannot intersect.*

*Proof.* The first statement is obvious by construction. We prove the second part by contradiction. Let $\ell_0^{\text{ext}} = [z_0, z_1]$ and $\ell_1^{\text{ext}}$ be two extended segments obtained from Setting 1 such that both connect the vertices of the same regular region $R_i$, for some $i \in \{1, \dots, m^2\}$. Supposing that they intersect, by construction, either they are identical or they share an end-point.

Suppose that the first case is possible. Then one of the segments $\ell_0, \ell_1$ must

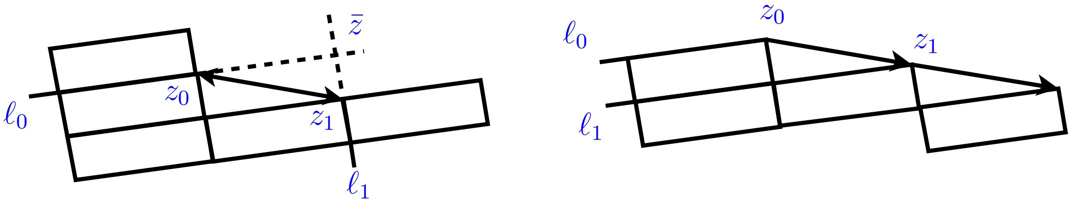


Figure 3.17: Impossible configurations of extended segments of Setting 1.

belong to $\mathcal{L}_i$ and the other to $\bar{\mathcal{L}}_i$. But then, once again by construction, the intersection point $\bar{z}$ of $\ell_0^{\text{line}}$ with $\ell_1^{\text{line}}$ must be a vertex of $R_i$ (see left of Figure 3.17), and therefore should belong to both $\ell_0$ and $\ell_1$, thus a contradiction.

Suppose that the second case is possible, that is, $\ell_0^{\text{ext}}$ and $\ell_1^{\text{ext}}$ share an endpoint. Then necessarily both $\ell_0, \ell_1$ must belong to either $\mathcal{L}_i$ or $\bar{\mathcal{L}}_i$ (see right of Figure 3.17). However, by the second condition in Setting 1, one of the segments $\ell_0, \ell_1$ cannot be extended by Setting 1, thus again a contradiction. Hence, there is no intersection of extended segments from Setting 1. □

### 3.3.3 Interior angles of triangles

Let $i \in \{1, \ldots, m^2\}$ be fixed and consider a neighboring sub-square $S_j$ to $S_i$. Then $\psi_i^{-1}(R_j)$ is uniformly triangulated, with triangles not necessarily equilateral but still isotropic. We show this below.

Let $R_j$ be a neighboring regular region to $R_i$ and consider a triangle $T' \in R_j$. Then, there is a vector $\mathbf{t}'$ such that $T' = \Lambda_j \varphi_j(T_0) + \mathbf{t}'$ where $\Lambda_j$ is as defined in (3.43). Since a translation does not change the shape of triangles, the triangle $\psi_i^{-1}(T')$ has the same shape as $\varphi_i^{-1}(T')$, see Figure 3.18, which is given by

$$\begin{aligned}\varphi_i^{-1} \circ (\Lambda_j \varphi_j)(T_0) &= \begin{bmatrix} \left|\frac{\lambda_{1,i}}{\lambda_{2,i}}\right|^{\frac{1}{4}} & 0 \\ 0 & \left|\frac{\lambda_{2,i}}{\lambda_{1,i}}\right|^{\frac{1}{4}} \end{bmatrix} R_{(\mu_j - \vartheta_j) - (\mu_i - \vartheta_i)} \begin{bmatrix} \Lambda_j \left|\frac{\lambda_{2,j}}{\lambda_{1,j}}\right|^{\frac{1}{4}} & 0 \\ 0 & \Lambda_j \left|\frac{\lambda_{1,j}}{\lambda_{2,j}}\right|^{\frac{1}{4}} \end{bmatrix} (T_0) \\ &= D_i^{-1} \circ R_{\bar{\theta}_{ij}} \circ (\Lambda_j D_j)(T_0),\end{aligned}$$

where $D_i = \operatorname{diag}\left(\left|\frac{\lambda_{2,i}}{\lambda_{1,i}}\right|^{\frac{1}{4}}, \left|\frac{\lambda_{1,i}}{\lambda_{2,i}}\right|^{\frac{1}{4}}\right)$ with $\lambda_{1,i}, \lambda_{2,i}$ being the eigenvalues of $Q_{\pi_i}$, and $\bar{\theta}_{ij} = (\mu_j - \vartheta_j) - (\mu_i - \vartheta_i)$ which is estimated in (3.46).

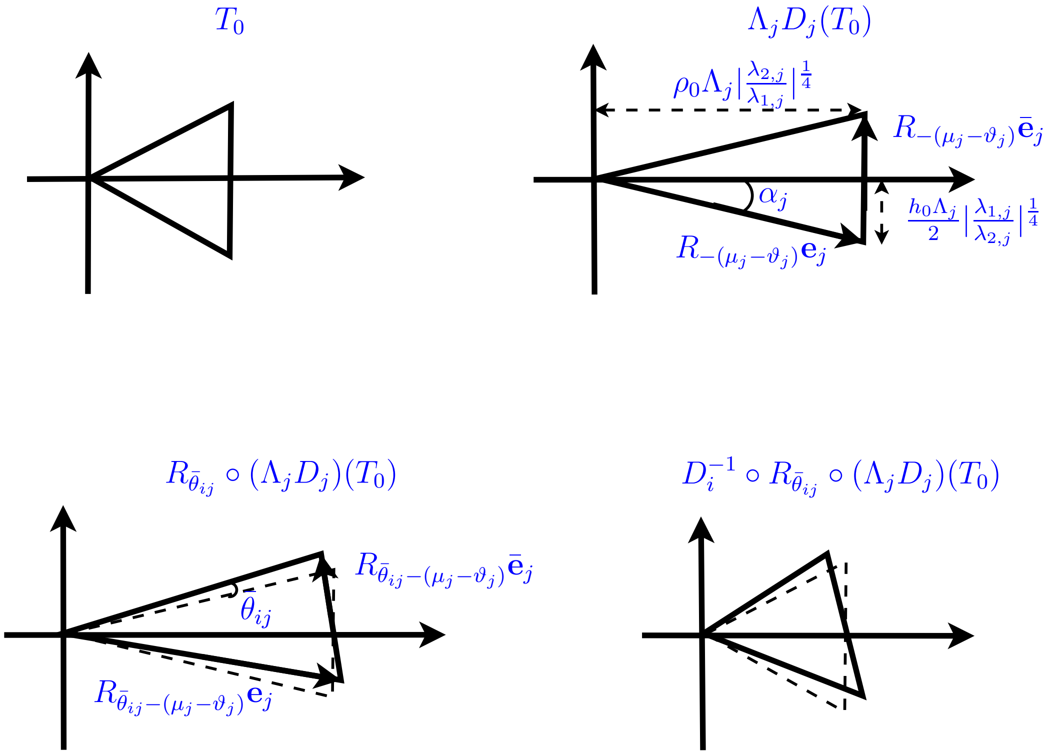

$T_0$
$\Lambda_j D_j(T_0)$
$\rho_0\Lambda_j|\frac{\lambda_{2,j}}{\lambda_{1,j}}|^{\frac{1}{4}}$
$R_{-(\mu_j-\vartheta_j)}\bar{\mathbf{e}}_j$
$\alpha_j$
$\frac{h_0\Lambda_j}{2}|\frac{\lambda_{1,j}}{\lambda_{2,j}}|^{\frac{1}{4}}$
$R_{-(\mu_j-\vartheta_j)}\mathbf{e}_j$
$R_{\bar{\theta}_{ij}} \circ (\Lambda_j D_j)(T_0)$
$R_{\bar{\theta}_{ij}-(\mu_j-\vartheta_j)}\bar{\mathbf{e}}_j$
$\bar{\theta}_{ij}$
$R_{\bar{\theta}_{ij}-(\mu_j-\vartheta_j)}\mathbf{e}_j$
$D_i^{-1} \circ R_{\bar{\theta}_{ij}} \circ (\Lambda_j D_j)(T_0)$

Figure 3.18: Back transformation of a regular triangle from a neighboring regular region.

Recall the edge-vectors $\mathbf{e}_i$, $\bar{\mathbf{e}}_i$ of $T_i$ which are described in step (iii) of Section 3.2.2. In Figure 3.18, the vectors $\mathbf{e}_j$ and $\bar{\mathbf{e}}_j$ are as defined in step (iii) of Section 3.2.2, the angle $\alpha_i$ has its tangent given in (3.48), and $\Lambda_j$ is defined as in (3.42) . Recall also that $h_0 = 2/3^{\frac{1}{4}}$ denotes the diameter of an equilateral triangle of unit area.

We have the following result.

**Lemma 3.3.7.** *Let $r$ be sufficiently small so that (3.17) holds. The diameter of $\psi_i^{-1}(T')$ is less than $(1+\frac{1}{10^4(\frac{3}{2})^{\frac{1}{2}}})(1+\xi)h_0\Lambda_j$, with $\Lambda_j$ as defined in (3.43), and where*

$$\xi = 6^{\frac{1}{4}}\varepsilon^{\frac{1}{4}}\delta_f^{-\frac{1}{2}}|f|^{\frac{1}{4}}_{W^2_\infty(\Omega)}, \tag{3.63}$$

*with $\varepsilon = \frac{3}{2}\omega(\sqrt{2}r)$ such that $\|Q_{\pi_i} - Q_{\pi_j}\|_2 \le \varepsilon$ and $\delta_f$ satisfying (3.13).*

*Proof.* The two edge vectors $\mathbf{e}_j^*$ and $\bar{\mathbf{e}}_j^*$ of the triangle $R_{\bar{\theta}_{ij}} \circ (\Lambda_j D_j)(T_0)$ are respectively given by (3.64) and (3.72) below. The first edge-vector is given by

$$\mathbf{e}_j^* = R_{\bar{\theta}_{ij}-(\mu_j-\vartheta_j)}\mathbf{e}_j = \Big[|\mathbf{e}_j|\cos(\alpha_j-\bar{\theta}_{ij})\ ,\ -|\mathbf{e}_j|\sin(\alpha_j-\bar{\theta}_{ij})\Big]^t, \tag{3.64}$$

where its image by $D_i^{-1}$ can be easily computed,

$$D_i^{-1}\mathbf{e}_j^* = |\mathbf{e}_j|\Big[\Big|\frac{\lambda_{1,i}}{\lambda_{2,i}}\Big|^{\frac{1}{4}}\cos(\alpha_j-\bar{\theta}_{ij})\ ,\ -\Big|\frac{\lambda_{2,i}}{\lambda_{1,i}}\Big|^{\frac{1}{4}}\sin(\alpha_j-\bar{\theta}_{ij})\Big]^t.$$

The length of the above vector is estimated as follows: From the fact that $\sin\alpha_j = \frac{h_0\Lambda_j}{2|\mathbf{e}_j|}\Big|\frac{\lambda_{1,j}}{\lambda_{2,j}}\Big|^{\frac{1}{4}}$ and $\cos\alpha_j = \frac{\rho_0\Lambda_j}{|\mathbf{e}_j|}\Big|\frac{\lambda_{2,j}}{\lambda_{1,j}}\Big|^{\frac{1}{4}}$ we easily get

$$\begin{aligned}\cos(\alpha_j-\bar{\theta}_{ij}) &= \cos\alpha_j\cos\bar{\theta}_{ij}+\sin\alpha_j\sin\bar{\theta}_{ij}\\ &= \frac{\rho_0\Lambda_j}{|\mathbf{e}_j|}\Big|\frac{\lambda_{2,j}}{\lambda_{1,j}}\Big|^{\frac{1}{4}}\cos\bar{\theta}_{ij}+\frac{h_0\Lambda_j}{2|\mathbf{e}_j|}\Big|\frac{\lambda_{1,j}}{\lambda_{2,j}}\Big|^{\frac{1}{4}}\sin\bar{\theta}_{ij},\end{aligned}$$

whereas,

$$\begin{aligned}\sin(\alpha_j-\bar{\theta}_{ij}) &= \sin\alpha_j\cos\bar{\theta}_{ij}-\cos\alpha_j\sin\bar{\theta}_{ij}\\ &= \frac{h_0\Lambda_j}{2|\mathbf{e}_j|}\Big|\frac{\lambda_{1,j}}{\lambda_{2,j}}\Big|^{\frac{1}{4}}\cos\bar{\theta}_{ij}-\frac{\rho_0\Lambda_j}{|\mathbf{e}_j|}\Big|\frac{\lambda_{2,j}}{\lambda_{1,j}}\Big|^{\frac{1}{4}}\sin\bar{\theta}_{ij}.\end{aligned}$$

We now deduce that

$$\begin{aligned}|D_i^{-1}\mathbf{e}_j^*| =&\Lambda_j\Big[\Big(\rho_0\cos\bar{\theta}_{ij}\Big|\frac{\lambda_{1,i}}{\lambda_{2,i}}\frac{\lambda_{2,j}}{\lambda_{1,j}}\Big|^{\frac{1}{4}}+\frac{h_0}{2}\sin\bar{\theta}_{ij}\Big|\frac{\lambda_{1,i}}{\lambda_{2,i}}\frac{\lambda_{1,j}}{\lambda_{2,j}}\Big|^{\frac{1}{4}}\Big)^2\\ &+\Big(\frac{h_0}{2}\cos\bar{\theta}_{ij}\Big|\frac{\lambda_{2,i}}{\lambda_{1,i}}\frac{\lambda_{1,j}}{\lambda_{2,j}}\Big|^{\frac{1}{4}}-\rho_0\sin\bar{\theta}_{ij}\Big|\frac{\lambda_{2,i}}{\lambda_{1,i}}\frac{\lambda_{2,j}}{\lambda_{1,j}}\Big|^{\frac{1}{4}}\Big)^2\Big]^{\frac{1}{2}}\\ =&\Lambda_j\Big[\rho_0^2\cos^2\bar{\theta}_{ij}\Big|\frac{\lambda_{1,i}}{\lambda_{2,i}}\frac{\lambda_{2,j}}{\lambda_{1,j}}\Big|^{\frac{1}{2}}+\frac{h_0^2}{4}\sin^2\bar{\theta}_{ij}\Big|\frac{\lambda_{1,i}}{\lambda_{2,i}}\frac{\lambda_{1,j}}{\lambda_{2,j}}\Big|^{\frac{1}{2}}\\ &+\frac{h_0^2}{4}\cos^2\bar{\theta}_{ij}\Big|\frac{\lambda_{2,i}}{\lambda_{1,i}}\frac{\lambda_{1,j}}{\lambda_{2,j}}\Big|^{\frac{1}{2}}+\rho_0^2\sin^2\bar{\theta}_{ij}\Big|\frac{\lambda_{2,i}}{\lambda_{1,i}}\frac{\lambda_{2,j}}{\lambda_{1,j}}\Big|^{\frac{1}{2}}\\ &+2\cos\bar{\theta}_{ij}\sin\bar{\theta}_{ij}\Big(\Big|\frac{\lambda_{1,i}}{\lambda_{2,i}}\Big|^{\frac{1}{2}}-\Big|\frac{\lambda_{2,i}}{\lambda_{1,i}}\Big|^{\frac{1}{2}}\Big)\Big]^{\frac{1}{2}},\end{aligned} \tag{3.65}$$

since $\dfrac{h_0\rho_0}{2}=1$. Recall that $\delta_f$ is defined in (3.13) and that $\varepsilon=\frac{3}{2}\omega(\sqrt{2}r)$, with

$\|Q_{\pi_i} - Q_{\pi_j}\|_2 \leq \varepsilon$. We also deduce from (2.4) the inequalities

$$\max\{|\lambda_{2,i}|, |\lambda_{2,j}|\} \leq \frac{3}{2}\max\{\|\pi_i\|, \|\pi_j\|\} \leq \frac{3}{2}|f|_{W^2_\infty(\Omega)}.$$

Hence, from (3.52) (see proof of Proposition 3.3.2) we also have

$$\Big|\frac{\lambda_{1,j}}{\lambda_{2,j}} - \frac{\lambda_{1,i}}{\lambda_{2,i}}\Big|^{\frac{1}{2}} \leq 2(\frac{3}{2})^{\frac{1}{2}}\varepsilon^{\frac{1}{2}}\delta_f^{-1}|f|^{\frac{1}{2}}_{W^2_\infty(\Omega)}. \tag{3.66}$$

We also deduce from (3.51) that

$$\Big|\frac{\lambda_{2,i}}{\lambda_{1,i}} - \frac{\lambda_{2,j}}{\lambda_{1,j}}\Big|^{\frac{1}{2}} = \frac{|\lambda_{1,j}\lambda_{2,i} - \lambda_{1,i}\lambda_{2,j}|^{\frac{1}{2}}}{|\lambda_{1,i}\lambda_{1,j}|^{\frac{1}{2}}} \leq 2(\frac{3}{2})^{\frac{1}{2}}\varepsilon^{\frac{1}{2}}\delta_f^{-1}|f|^{\frac{1}{2}}_{W^2_\infty(\Omega)}. \tag{3.67}$$

Taking the square root, from (3.67) it immediately follows that

$$\Big|\frac{\lambda_{2,i}}{\lambda_{1,i}}\Big|^{\frac{1}{4}}\Big|\frac{\lambda_{1,i}\lambda_{2,j}}{\lambda_{2,i}\lambda_{1,j}} - 1\Big|^{\frac{1}{4}} = \Big|\frac{\lambda_{2,j}}{\lambda_{1,j}} - \frac{\lambda_{2,i}}{\lambda_{1,i}}\Big|^{\frac{1}{4}} \leq 6^{\frac{1}{4}}\varepsilon^{\frac{1}{4}}\delta_f^{-\frac{1}{2}}|f|^{\frac{1}{4}}_{W^2_\infty(\Omega)},$$

which yields the estimation $\Big|\dfrac{\lambda_{1,i}}{\lambda_{2,i}}\dfrac{\lambda_{2,j}}{\lambda_{1,j}}\Big|^{\frac{1}{4}} \leq 1 + \xi$ with $\xi = 6^{\frac{1}{4}}\varepsilon^{\frac{1}{4}}\delta_f^{-\frac{1}{2}}|f|^{\frac{1}{4}}_{W^2_\infty(\Omega)}$. The estimation $\Big|\dfrac{\lambda_{1,j}}{\lambda_{2,j}}\dfrac{\lambda_{2,i}}{\lambda_{1,i}}\Big|^{\frac{1}{4}} \leq 1 + \xi$ is proved in a similar way.

Next, we shall estimate $\rho_0^2(\sin\bar{\theta}_{ij})^2\Big|\frac{\lambda_{2,i}}{\lambda_{1,i}}\frac{\lambda_{2,j}}{\lambda_{1,j}}\Big|^{\frac{1}{2}}$. By using (3.46) and (3.17), and recalling that $\varepsilon = \frac{3}{2}\omega(\sqrt{2}r)$, we have

$$\begin{aligned}(\sin\bar{\theta}_{ij})^2 \leq \bar{\theta}_{ij}^2 &\leq \bar{\theta}_{ij}C_{\delta_f}[\frac{3}{2}\omega(\sqrt{2}r)]^{\frac{1}{2}} \\ &\leq \bar{\theta}_{ij}\frac{1}{10^5(\frac{3}{2})^{\frac{1}{2}}C_{\delta_f}|f|^{\frac{1}{2}}_{W^2_\infty(\Omega)}}\Big|\frac{\lambda_{1,j}}{\lambda_{2,j}}\Big|,\end{aligned}$$

and therefore, since $C_1 \leq C_{\delta_f}$, we deduce from (3.47) that

$$\rho_0^2(\sin\bar{\theta}_{ij})^2\Big|\frac{\lambda_{2,i}}{\lambda_{1,i}}\frac{\lambda_{2,j}}{\lambda_{1,j}}\Big|^{\frac{1}{2}} \leq \bar{\theta}_{ij}\frac{\rho_0^2}{10^5(\frac{3}{2})^{\frac{1}{2}}C_{\delta_f}|f|^{\frac{1}{2}}_{W^2_\infty(\Omega)}}\Big|\frac{\lambda_{2,i}}{\lambda_{1,i}}\frac{\lambda_{1,j}}{\lambda_{2,j}}\Big|^{\frac{1}{2}}$$

$$\begin{aligned}&\leq \frac{\rho_0^2 \varepsilon^{\frac{1}{2}}}{10^5 (\frac{3}{2})^{\frac{1}{2}} |f|^{\frac{1}{2}}_{W^2_\infty(\Omega)}} (1+\xi)^2 \\ &= \frac{\omega(\sqrt{2}r)^{\frac{1}{2}}}{10^5 |f|^{\frac{1}{2}}_{W^2_\infty(\Omega)}} \rho_0^2 (1+\xi)^2, \end{aligned} \tag{3.68}$$

by virtue of $\varepsilon = \frac{3}{2}\omega(\sqrt{2}r)$.

We observe that, since $\delta_f \in (0,1)$ (see (3.13)), clearly

$$\frac{1}{C^2_{\delta_f}} = \frac{\delta_f^2}{(2+33\pi\delta_f)^2} \leq \frac{\delta_f}{4}.$$

Also, since from (2.4) we have that $\delta_f \leq \frac{3}{2}|f|_{W^2_\infty(\Omega)}$, we obtain $\frac{1}{C^2_{\delta_f}} \leq \frac{3}{8}|f|_{W^2_\infty(\Omega)} \leq |f|_{W^2_\infty(\Omega)}$. It is now easy to show from (3.15) that

$$\frac{\omega(\sqrt{2}r)^{\frac{1}{2}}}{10^5 |f|_{W^2_\infty(\Omega)}} \leq \frac{1}{10^8 \frac{3}{2}}. \tag{3.69}$$

Next, from the fact that $\left|\frac{\lambda_{1,i}}{\lambda_{2,i}}\right|^{\frac{1}{2}} \leq 1$, we deduce from (3.15) that

$$\begin{aligned}\left|\sin\bar{\theta}_{ij}\left(\left|\frac{\lambda_{1,i}}{\lambda_{2,i}}\right|^{\frac{1}{2}} - \left|\frac{\lambda_{2,i}}{\lambda_{1,i}}\right|^{\frac{1}{2}}\right)\right| &\leq |\sin\bar{\theta}_{ij}| \left|\frac{\lambda_{2,i}}{\lambda_{1,i}}\right|^{\frac{1}{2}} \\ &\leq C_{\delta_f}[\frac{3}{2}\omega(\sqrt{2}r)]^{\frac{1}{2}} \left|\frac{\lambda_{2,i}}{\lambda_{1,i}}\right|^{\frac{1}{2}} \\ &\leq \frac{1}{10^5 (\frac{3}{2})^{\frac{1}{2}} C_{\delta_f} |f|^{\frac{1}{2}}_{W^2_\infty(\Omega)}}.\end{aligned}$$

By virtue of $\frac{1}{C_{\delta_f}} \leq \frac{1}{2}\delta_f^{\frac{1}{2}} \leq \frac{(\frac{3}{2})^{\frac{1}{2}}}{2}|f|^{\frac{1}{2}}_{W^2_\infty(\Omega)} \leq |f|^{\frac{1}{2}}_{W^2_\infty(\Omega)}$, we obtain

$$\left|\sin\bar{\theta}_{ij}\left(\left|\frac{\lambda_{1,i}}{\lambda_{2,i}}\right|^{\frac{1}{2}} - \left|\frac{\lambda_{2,i}}{\lambda_{1,i}}\right|^{\frac{1}{2}}\right)\right| \leq \frac{1}{10^5(\frac{3}{2})^{\frac{1}{2}}} \leq \frac{1}{10^4(\frac{3}{2})^{\frac{1}{2}}}(1+\xi)^2. \tag{3.70}$$

Now using (3.65), (3.68) and (3.69), together with (3.70) and the fact that

$\rho_0^2 + \frac{h_0^2}{4} = h_0^2$, we deduce from (3.65) that

$$\begin{aligned} |D_i^{-1}\mathbf{e}_j^*| &\le (1+\xi)\Lambda_j\Big(\rho_0^2\cos^2\bar\theta_{ij} + \frac{h_0^2}{4} + \frac{\rho_0^2}{10^8\frac{3}{2}} + \frac{2|\cos\bar\theta_{ij}|}{10^4(\frac{3}{2})^{\frac{1}{2}}}\Big)^{\frac{1}{2}} \\ &\le (1+\xi)h_0\Lambda_j(1 + \frac{1}{10^8\frac{3}{2}} + \frac{2}{10^4(\frac{3}{2})^{\frac{1}{2}}})^{\frac{1}{2}} \\ &= (1+\xi)(1+\frac{1}{10^4(\frac{3}{2})^{\frac{1}{2}}})h_0\Lambda_j. \end{aligned} \tag{3.71}$$

The second edge-vector of $R_{\bar\theta_{ij}} \circ (\Lambda_j D_j)(T_0)$ is given by

$$\bar{\mathbf{e}}_j^* = R_{\bar\theta_{ij}-(\mu_j-\vartheta_j)}\bar{\mathbf{e}}_j = [-|\bar{\mathbf{e}}_j|\cos\bar\theta_{ij}\ ,\ |\bar{\mathbf{e}}_{ij}|\sin\bar\theta_{ij}]^t, \tag{3.72}$$

with its image by $D_i^{-1}$ being

$$D_i^{-1}\bar{\mathbf{e}}_j^* = |\bar{\mathbf{e}}_j|\Big[-\Big|\frac{\lambda_{1,i}}{\lambda_{2,i}}\Big|^{\frac{1}{4}}\cos\bar\theta_{ij}\ ,\ \Big|\frac{\lambda_{2,i}}{\lambda_{1,i}}\Big|^{\frac{1}{4}}\sin\bar\theta_{ij}\Big]^t,$$

from which it follows that

$$|D_i^{-1}\bar{\mathbf{e}}_j^*| = |\bar{\mathbf{e}}_j|\Big(\Big|\frac{\lambda_{1,i}}{\lambda_{2,i}}\Big|^{\frac{1}{2}}\cos^2\bar\theta_{ij} + 2\cos\bar\theta_{ij}\sin\bar\theta_{ij} + \Big|\frac{\lambda_{2,i}}{\lambda_{1,i}}\Big|^{\frac{1}{2}}\sin^2\bar\theta_{ij}\Big)^{\frac{1}{2}},$$

where $|\bar{\mathbf{e}}_j| = \frac{h_0}{2}\Lambda_j\Big|\frac{\lambda_{1,j}}{\lambda_{2,j}}\Big|^{\frac{1}{4}}$. Hence

$$|D_i^{-1}\bar{\mathbf{e}}_j^*| \le \frac{h_0}{2}\Lambda_j\Big(\Big|\frac{\lambda_{1,i}}{\lambda_{2,i}}\frac{\lambda_{1,j}}{\lambda_{2,j}}\Big|^{\frac{1}{2}}\cos^2\bar\theta_{ij} + 2\cos\bar\theta_{ij}\sin\bar\theta_{ij}\Big|\frac{\lambda_{1,j}}{\lambda_{2,j}}\Big|^{\frac{1}{2}} + \frac{\lambda_{2,i}}{\lambda_{1,i}}\frac{\lambda_{2,j}}{\lambda_{1,j}}\Big|^{\frac{1}{2}}\sin^2\bar\theta_{ij}\Big)^{\frac{1}{2}}.$$

Using a similar approach as to obtain (3.71), we obtain

$$|D_i^{-1}\bar{\mathbf{e}}_j^*| \le (1+\xi)(1+\frac{1}{10^4(\frac{3}{2})^{\frac{1}{2}}})\frac{h_0}{2}\Lambda_j. \tag{3.73}$$

The length of the third edge-vector of $D_i^{-1} \circ R_{\bar\theta_{ij}}(\Lambda_j D_j)(T_0)$ is estimated in a similar way as in (3.71). Combining this with (3.71) and (3.73) yields the result. □

The triangle $D_i^{-1} \circ R_{\bar\theta_{ij}}(\Lambda_j D_j)(T_0)$ is isotropic with its sides being comparable to those of $\Lambda_j T_0$ which is an equilateral triangle: From (3.43), with $\Lambda_i = s^\eta(K_p(\pi_i) + 2Ch_{\pi_i}^2\omega(r))^{-\frac{q}{2}}$, let us estimate the quotient

$$\frac{\Lambda_j}{\Lambda_i} = \Big(\frac{K_p(\pi_i) + 2Ch_{\pi_i}^2\omega(r)}{K_p(\pi_j) + 2Ch_{\pi_j}^2\omega(r)}\Big)^{\frac{q}{2}}, \tag{3.74}$$

where $h_{\pi_i}$ is the diameter of an optimal triangle for $\pi_i$ by using the construction in Algorithm 3.1. First, from (2.54) we have $K_p(\pi_i) = K_p(\pi_0)|\lambda_{1,i}\lambda_{2,i}|^{\frac{1}{2}}$ where $\pi_0(x,y) = x^2 + y^2$. Assuming without loss of generality that $\lambda_{1,j}$ is the closest to $\lambda_{1,i}$ of the eigenvalues of $Q_{\pi_j}$, and also that $\lambda_{2,j}$ is the closest to $\lambda_{2,i}$, we deduce from (3.19) that

$$\begin{aligned}|\lambda_{1,i}\lambda_{2,i} - \lambda_{1,j}\lambda_{2,j}| &= |(\lambda_{1,i} - \lambda_{1,j})\lambda_{2,i} + \lambda_{1,j}(\lambda_{2,i} - \lambda_{2,j})|\\ &\le 2\|Q_{\pi_i} - Q_{\pi_j}\|_2\lambda_{2,i}\\ &\le 3\omega(\sqrt{2}r)\lambda_{2,i},\end{aligned} \tag{3.75}$$

by virtue of (2.4) and and (3.45). It follows from (3.75) that

$$|K_p(\pi_i) - K_p(\pi_j)| \le 3^{\frac{1}{2}}K_p(\pi_0)\omega(\sqrt{2}r)^{\frac{1}{2}}\lambda_{2,i}^{\frac{1}{2}}. \tag{3.76}$$

On the other hand, by using (3.66) and (3.67), we obtain

$$\begin{aligned}h_{\pi_i}^2 &= \rho_0^2\Big|\frac{\lambda_{2,i}}{\lambda_{1,i}}\Big|^{\frac{1}{2}} + \frac{h_0^2}{4}\Big|\frac{\lambda_{1,i}}{\lambda_{2,i}}\Big|^{\frac{1}{2}}\\ &\le \rho_0^2\Big(\Big|\frac{\lambda_{2,j}}{\lambda_{1,j}}\Big|^{\frac{1}{2}} + \xi^2\Big) + \frac{h_0^2}{4}\Big(\Big|\frac{\lambda_{1,j}}{\lambda_{2,j}}\Big|^{\frac{1}{2}} + \xi^2\Big)\\ &= h_{\pi_j}^2 + h_0^2\xi^2,\end{aligned}$$

by virtue of the fact that $\rho_0^2 + \frac{h_0^2}{4} = h_0^2$. Combining this with (3.76) yields

$$\begin{aligned}K_p(\pi_i) + 2Ch_{\pi_i}^2\omega(r) \le{}& K_p(\pi_j) + 2Ch_{\pi_j}^2\omega(r)\\ &+ 3^{\frac{1}{2}}K_p(\pi_0)\omega(\sqrt{2}r)^{\frac{1}{2}}\lambda_{2,i}^{\frac{1}{2}} + 2Ch_0^2\xi^2\omega(r).\end{aligned}$$

Since $\varepsilon = \frac{3}{2}\omega(\sqrt{2}r)$, clearly from (3.63) we have

$\xi^2 = 3\omega(\sqrt{2}r)^{\frac{1}{2}}\delta_f^{-1}|f|^{\frac{1}{2}}_{W^2_\infty(\Omega)}$. Thus the inequalities

$$\begin{aligned}\frac{3^{\frac{1}{2}}K_p(\pi_0)\omega(\sqrt{2}r)^{\frac{1}{2}}\lambda_{2,i}^{\frac{1}{2}}}{K_p(\pi_j)+2Ch^2_{\pi_j}\omega(r)} &\leq \frac{3^{\frac{1}{2}}K_p(\pi_0)\omega(\sqrt{2}r)^{\frac{1}{2}}\lambda_{2,i}^{\frac{1}{2}}}{K_p(\pi_0)|\lambda_{1,j}\lambda_{2,j}|^{\frac{1}{2}}} \\ &\leq \frac{3^{\frac{1}{2}}\omega(\sqrt{2}r)^{\frac{1}{2}}\left(\frac{3}{2}\right)^{\frac{1}{2}}|f|^{\frac{1}{2}}_{W^2_\infty(\Omega)}}{\delta_f} \\ &\leq \xi^2.\end{aligned}$$

From the fact that $h_0^2 = \frac{h_0^2}{4} + \rho_0^2$ where $\rho_0^2 = \sqrt{3} > \frac{h_0^2}{4} = \frac{1}{\sqrt{3}}$, we can easily prove that for any $A \geq 1$, we have $h_0^2 \leq \frac{h_0^2}{4}\frac{1}{A} + \rho_0^2 A$. Hence $h_0^2 \leq h^2_{\pi_j}$ holds for any $j$. Also, since $h_{\pi_j} \geq 1$, clearly

$$\frac{2Ch_0^2\xi^2\omega(r)}{K_p(\pi_j)+2Ch^2_{\pi_j}\omega(r)} \leq \frac{2Ch_0^2\xi^2\omega(r)}{2Ch^2_{\pi_j}\omega(r)} = \frac{h_0^2}{h^2_{\pi_j}}\xi^2 \leq \xi^2,$$

after assuming that $\omega(r) \neq 0$. Obviously in the case $\omega(r) = 0$, we conclude that

$$K_p(\pi_i) + 2Ch^2_{\pi_i}\omega(r) \leq (1+2\xi^2)\Big(K_p(\pi_j)+2Ch^2_{\pi_j}\omega(r)\Big), \tag{3.77}$$

and hence, from (3.74),

$$\frac{\Lambda_j}{\Lambda_i} \leq (1+2\xi^2)^{\frac{q}{2}}. \tag{3.78}$$

Note that the inverse inequality can be proved by using a similar argument. Since $\varepsilon = \frac{3}{2}\omega(\sqrt{2}r)$, we deduce from (3.63) and (3.17) that

$$\begin{aligned}\xi &= 3^{\frac{1}{2}}\omega(\sqrt{2}r)^{\frac{1}{4}}\delta_f^{-\frac{1}{2}}|f|^{\frac{1}{4}}_{W^2_\infty(\Omega)} \\ &\leq \frac{1}{10^{\frac{5}{2}}\delta_f^{\frac{1}{2}}C_{\delta_f}} \\ &\leq \frac{1}{2}10^{-\frac{5}{2}},\end{aligned} \tag{3.79}$$

by virtue of the fact that $\frac{1}{\delta_f^{\frac{1}{2}}C_{\delta_f}} \leq \frac{1}{2}$ which can be easily proved. Hence, with $d_i' = |\bar{\mathbf{e}}_i^*| = h_0\Lambda_i$ and $d_j' = |\bar{\mathbf{e}}_j^*|$, combining (3.71) and (3.73) with (3.78) yields the inequalities below with $\xi_1 = \frac{1}{10^4(\frac{3}{2})^{\frac{1}{2}}}$,

$$\begin{aligned}
d'_j &\leq (1+\xi)(1+\xi_1)h_0\Lambda_j \leq (1+\xi)(1+2\xi^2)^{\frac{q}{2}}(1+\xi_1)h_0\Lambda_i \\
&\leq (1+\xi)(1+\sqrt{2}\xi)(1+\xi_1)h_0\Lambda_i \leq (1+4\xi)(1+\xi_1)d'_i \\
&\leq (1+8\xi+\xi_1)d'_i \leq \left(1+\frac{9}{2}10^{-\frac{5}{2}}\right)d'_i \\
&\leq (1+10^{-\frac{3}{2}})d'_i, \qquad (3.80)
\end{aligned}$$

since $\xi_1$ is less than the upper bound of $\xi$ in (3.79). The inverse inequality, that is $d'_i \leq (1+10^{-\frac{3}{2}})d'_j$, can be proved by using a similar argument.

**Lemma 3.3.8.** *Let $r$ be sufficiently small so that (3.17) holds. Suppose that $S_i, S_j$ are two neighboring sub-squares. Then the angle $\theta^*_{ij}$ (resp. $\bar{\theta}^*_{ij}$) formed from the segments in $\psi_i^{-1}(\mathcal{L}_i)$ and $\psi_i^{-1}(\mathcal{L}_j)$ (resp. $\psi_i^{-1}(\bar{\mathcal{L}}_i)$ and $\psi_i^{-1}(\bar{\mathcal{L}}_j)$) is bounded by $C'_1\varepsilon^{\frac{1}{4}}$, where $\varepsilon = \frac{3}{2}\omega(\sqrt{2}r)$ and $C'_1$ a constant.*

*Proof.* After a back transformation (see Figure 3.18), the directional vectors $\mathbf{e}^*_i$ and $\mathbf{e}^*_j$, respectively associated with $\psi_i^{-1}(\mathcal{L}_i)$ and $\psi_i^{-1}(\mathcal{L}_j)$ are given by

$$\mathbf{e}^*_i = \Lambda_i[\rho_0 \ \ -\frac{h_0}{2}]^t, \qquad (3.81)$$

$$\mathbf{e}^*_j = |\mathbf{e}_j|\Big[\Big|\frac{\lambda_{1,i}}{\lambda_{2,i}}\Big|^{\frac{1}{4}}\cos(\alpha_j-\bar{\theta}_{ij}),\ -\Big|\frac{\lambda_{2,i}}{\lambda_{1,i}}\Big|^{\frac{1}{4}}\sin(\alpha_j-\bar{\theta}_{ij})\Big]^t. \qquad (3.82)$$

In order to estimate the angle $\theta^*_{ij}$, we shall use the scalar product of $\mathbf{e}^*_i$ and $\mathbf{e}^*_j$. We have that $|\mathbf{e}_j| = \frac{\rho_0\Lambda_j}{\cos\alpha_j}\Big|\frac{\lambda_{2,j}}{\lambda_{1,j}}\Big|^{\frac{1}{4}}$, and also

$$\begin{aligned}
A_{ij} &:= \cos(\alpha_j-\bar{\theta}_{ij}) = \cos\alpha_j\cos\bar{\theta}_{ij} + \sin\alpha_j\sin\bar{\theta}_{ij}, \\
B_{ij} &:= \sin(\alpha_j-\bar{\theta}_{ij}) = \sin\alpha_j\cos\bar{\theta}_{ij} - \cos\alpha_j\sin\bar{\theta}_{ij}.
\end{aligned}$$

We thus have

$$\begin{aligned}
\mathbf{e}^*_j &= \frac{\rho_0\Lambda_j}{\cos\alpha_j}\Big|\frac{\lambda_{2,j}}{\lambda_{1,j}}\Big|^{\frac{1}{4}}\Big[\Big|\frac{\lambda_{1,i}}{\lambda_{2,i}}\Big|^{\frac{1}{4}}A_{ij}\,,\ -\Big|\frac{\lambda_{2,i}}{\lambda_{1,i}}\Big|^{\frac{1}{4}}B_{ij}\Big]^t \\
&= \rho_0\Lambda_j\Big[\Big|\frac{\lambda_{2,j}}{\lambda_{1,j}}\frac{\lambda_{1,i}}{\lambda_{2,i}}\Big|^{\frac{1}{4}}A_{ij}\,,\ -\Big|\frac{\lambda_{2,j}}{\lambda_{1,j}}\frac{\lambda_{2,i}}{\lambda_{1,i}}\Big|^{\frac{1}{4}}B_{ij}\Big]^t.
\end{aligned}$$

The square of the scalar product of $\mathbf{e}_i^*$ and $\mathbf{e}_j^*$ satifies

$$\begin{aligned}(\mathbf{e}_i^*\mathbf{e}_j^*)^2 =&\rho_0^2\Lambda_i^2\Lambda_j^2\Big(\rho_0\Big|\frac{\lambda_{2,j}}{\lambda_{1,j}}\frac{\lambda_{1,i}}{\lambda_{2,i}}\Big|^{\frac{1}{4}}A_{ij}+\frac{h_0}{2}\Big|\frac{\lambda_{2,j}}{\lambda_{1,j}}\frac{\lambda_{2,i}}{\lambda_{1,i}}\Big|^{\frac{1}{4}}B_{ij}\Big)^2\\ =&\rho_0^2\Lambda_i^2\Lambda_j^2\Big(\rho_0^2\Big|\frac{\lambda_{2,j}}{\lambda_{1,j}}\frac{\lambda_{1,i}}{\lambda_{2,i}}\Big|^{\frac{1}{2}}A_{ij}^2+\rho_0h_0\Big|\frac{\lambda_{2,j}}{\lambda_{1,j}}\Big|^{\frac{1}{2}}A_{ij}B_{ij}+\frac{h_0^2}{4}\Big|\frac{\lambda_{2,j}}{\lambda_{1,j}}\frac{\lambda_{2,i}}{\lambda_{1,i}}\Big|^{\frac{1}{2}}B_{ij}^2\Big).\end{aligned} \tag{3.83}$$

We also have that

$$(|\mathbf{e}_i^*||\mathbf{e}_j^*|)^2 = h_0^2\rho_0^2\Lambda_i^2\Lambda_j^2\Big(\Big|\frac{\lambda_{2,j}}{\lambda_{1,j}}\frac{\lambda_{1,i}}{\lambda_{2,i}}\Big|^{\frac{1}{2}}A_{ij}^2+\Big|\frac{\lambda_{2,j}}{\lambda_{1,j}}\frac{\lambda_{2,i}}{\lambda_{1,i}}\Big|^{\frac{1}{2}}B_{ij}^2\Big). \tag{3.84}$$

Hence, since $h_0^2-\rho_0^2=\frac{h_0^2}{4}$, we have

$$\begin{aligned}\sin^2\theta_{ij}^* =&1-\cos^2\theta_{ij}^*=\frac{(|\mathbf{e}_i^*||\mathbf{e}_j^*|)^2-(\mathbf{e}_i^*\mathbf{e}_j^*)^2}{(|\mathbf{e}_i^*||\mathbf{e}_j^*|)^2}\\ =&\frac{\rho_0^2\Lambda_i^2\Lambda_j^2}{(|\mathbf{e}_i^*||\mathbf{e}_j^*|)^2}\Big(\frac{h_0^2}{4}\Big|\frac{\lambda_{2,j}}{\lambda_{1,j}}\frac{\lambda_{1,i}}{\lambda_{2,i}}\Big|^{\frac{1}{2}}A_{ij}^2-\rho_0h_0\Big|\frac{\lambda_{2,j}}{\lambda_{1,j}}\Big|^{\frac{1}{2}}A_{ij}B_{ij}+\rho_0^2\Big|\frac{\lambda_{2,j}}{\lambda_{1,j}}\frac{\lambda_{2,i}}{\lambda_{1,i}}\Big|^{\frac{1}{2}}B_{ij}^2\Big)\\ =&\frac{\rho_0^2\Lambda_i^2\Lambda_j^2}{(|\mathbf{e}_i^*||\mathbf{e}_j^*|)^2}\Big(\frac{h_0}{2}\Big|\frac{\lambda_{2,j}}{\lambda_{1,j}}\frac{\lambda_{1,i}}{\lambda_{2,i}}\Big|^{\frac{1}{4}}A_{ij}-\rho_0\Big|\frac{\lambda_{2,j}}{\lambda_{1,j}}\frac{\lambda_{2,i}}{\lambda_{1,i}}\Big|^{\frac{1}{4}}B_{ij}\Big)^2.\end{aligned}$$

Recall from (3.48) that $\tan\alpha_j=\frac{h_0}{2\rho_0}\Big|\frac{\lambda_{1,j}}{\lambda_{2,j}}\Big|^{\frac{1}{2}}$. Thus

$$\rho_0\Big|\frac{\lambda_{2,j}}{\lambda_{1,j}}\frac{\lambda_{2,i}}{\lambda_{1,i}}\Big|^{\frac{1}{4}}\tan\alpha_j\cos\bar\theta_{ij}=\frac{h_0}{2}\Big|\frac{\lambda_{1,j}}{\lambda_{2,j}}\frac{\lambda_{2,i}}{\lambda_{1,i}}\Big|^{\frac{1}{4}}\cos\bar\theta_{ij}.$$

Hence we have

$$\begin{aligned}\sin^2\theta_{ij}^* =&\frac{\rho_0^2\Lambda_i^2\Lambda_j^2}{(|\mathbf{e}_i^*||\mathbf{e}_j^*|)^2}\cos^2\alpha_j\Big(\frac{h_0}{2}\Big(C-\frac{1}{C}\Big)\cos\bar\theta_{ij}+\frac{h_0}{2}C\tan\alpha_j\sin\bar\theta_{ij}\\ &+\rho_0\Big|\frac{\lambda_{2,j}}{\lambda_{1,j}}\frac{\lambda_{2,i}}{\lambda_{1,i}}\Big|^{\frac{1}{4}}\sin\bar\theta_{ij}\Big)^2,\end{aligned} \tag{3.85}$$

where $C=\Big|\frac{\lambda_{2,j}}{\lambda_{1,j}}\frac{\lambda_{1,i}}{\lambda_{2,i}}\Big|^{\frac{1}{4}}\leq 1+\xi$ with $\xi$ defined in (3.63), and estimated in (3.79). Without loss of generality, we assume that $\frac{1}{C}\leq C$ (otherwise we use $C'=\frac{1}{C}$), and thus

$$|C-\frac{1}{C}|=\Big|\frac{C^2-1}{C}\Big|\leq\Big|\frac{(1+\xi)^2-1}{C}\Big|=\frac{2\xi+\xi^2}{C}\leq 2\xi+\xi^2.$$

Now recall from Proposition 3.3.2 that $\sin\bar{\theta}_{ij}\leq C_{\delta_f}\varepsilon$, and that $\tan\alpha_j\leq 1$ since $\left|\frac{\lambda_{1,j}}{\lambda_{2,j}}\right|\leq 1$ and $h_0=2/3^{\frac{1}{4}},\rho_0=3^{\frac{1}{4}}$. This, coupled with (3.68), imply that

$$|\sin^2\theta^*_{ij}|\leq\frac{\rho_0^2\Lambda_i^2\Lambda_j^2}{(|\mathbf{e}_i^*||\mathbf{e}_j^*|)^2}\Big(\frac{h_0}{2}(2\xi+\xi^2)+\frac{h_0}{2}(1+\xi)C_{\delta_f}\varepsilon+\frac{\rho_0C_1^{\frac{1}{2}}\varepsilon^{\frac{1}{4}}(1+\xi)}{10^5C_{\delta_f}^{\frac{1}{2}}|f|^{\frac{1}{4}}_{W^2_\infty(\Omega)}}\Big)^2.$$

Combining this with (3.63) yields

$$\begin{aligned}|\sin^2\theta^*_{ij}|\leq&\frac{\rho_0^2\Lambda_i^2\Lambda_j^2}{(|\mathbf{e}_i^*||\mathbf{e}_j^*|)^2}\Big(\frac{3h_0}{2}\sqrt{2}\varepsilon^{\frac{1}{4}}\delta_f^{-\frac{1}{2}}|f|^{\frac{1}{4}}_{W^2_\infty(\Omega)}+h_0C_{\delta_f}\varepsilon+h_0c_0\varepsilon^{\frac{1}{4}}\Big)^2\\ \leq&\frac{c_0'}{\left|\frac{\lambda_{2,j}}{\lambda_{1,j}}\frac{\lambda_{1,i}}{\lambda_{2,i}}\right|^{\frac{1}{2}}A_{ij}^2+\left|\frac{\lambda_{2,j}}{\lambda_{1,j}}\frac{\lambda_{2,i}}{\lambda_{1,i}}\right|^{\frac{1}{2}}B_{ij}^2}\varepsilon^{\frac{1}{2}}\\ \leq&C_1'\varepsilon^{\frac{1}{2}},\end{aligned}\tag{3.86}$$

by virtue of (3.84), thereby proving the result. □

We are now ready to prove the following.

**Proposition 3.3.9.** *Let $r$ be sufficiently small so that (3.17) holds. For a fixed $i\in\{1,\ldots,m^2\}$, consider the back transformation $\Delta_i$ of the triangulation $\Delta_{s,r}$ defined in (3.44). Then:*

*i. For any segment $\ell_0\in\mathcal{L}_i$ extended according to Setting 1, the angle between the segments $\varphi_i^{-1}(\ell_0)$ and $\varphi_i^{-1}(\ell_0^{ext})$ is either $\pi/6$ or $\pi/3$;*

*ii. For any segment $\ell_0\in\mathcal{L}_i\cup\bar{\mathcal{L}}_i$ extended according to Setting 2-a, 2-b or 3 up to a neighboring sub-square $S_j$ of $S_i$, the angle between the segments $\varphi_i^{-1}(\ell_0)$ and $\varphi_i^{-1}(\ell_0^{ext})$ is less than $\frac{\pi}{7}$.*

*Proof.* Denote by $\theta$ the angle between the segments $\psi_i^{-1}(\ell_0)$ and $\psi_i^{-1}(\ell_0^{\text{ext}})$. Given a regular triangle $T$ in $R_i$, we denote by $d_i',h_i'$ the lengths of the shortest and

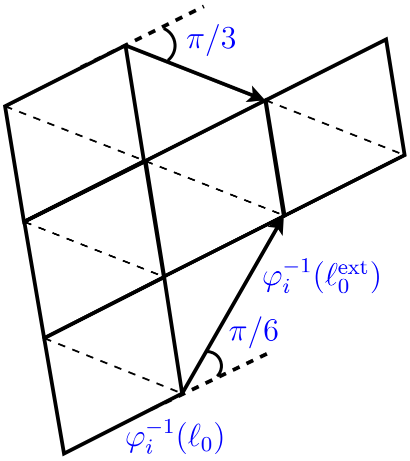


Figure 3.19: Connection according to Setting 1 for regular regions after a back transformation.

largest edges of $\psi_i^{-1}(T)$.

i. The back transformation of $\Delta_{s,r}$ by using $\psi_i$ transforms the regular triangles contained in $S_i$ into equilateral triangles. Given a segment $\ell_0 \in \mathcal{L}_i$ and the extended segment $\ell_0^{\text{ext}}$ obtained by using Setting 1, as shown in Figure 3.19, the segment $\psi_i^{-1}(\ell_0^{\text{ext}})$ is necessarily one of the diagonals of a parallelogram defined by two equilateral triangles of $\psi_i^{-1}(S_i)$. Thus, the angle $\theta$ is either $\pi/6$ or $\pi/3$;

The proof of the second part is divided into two parts ii-a and ii-b (on page 147) depending on which Setting is used to extend a given segment $\ell_0$.

ii-a. Consider a segment $\ell_0 \in \mathcal{L}_i$ which is extended by using Setting 2-a following the direction $\mathbf{e}_i$. Given the segment $\bar{\ell}_1 \in \bar{\mathcal{L}}_j$ (described in Setting 2-a) which intersects the line $\ell_0^{\text{line}}$, let $z_0$ denote the common vertex of $\psi_i^{-1}(\ell_0)$ and $\psi_i^{-1}(\ell_0^{\text{ext}})$, and let $z$ denote the intersection of the line extending $\psi_i^{-1}(\ell_0)$ with $\psi_i^{-1}(\bar{\ell}_1)$.

We first recall that the angle $\theta_{ij}^*$ (resp. $\bar{\theta}_{ij}^*$) between the segments in $\psi_i^{-1}(\mathcal{L}_i)$ and in $\psi_i^{-1}(\mathcal{L}_j)$ (reps. $\psi_i^{-1}(\bar{\mathcal{L}}_i)$ and $\psi_i^{-1}(\bar{\mathcal{L}}_j)$) is estimated according to Lemma 3.3.8. Also, after a back transformation the angles in a parallelogram of $\psi_i^{-1}(R_i)$ are $\frac{\pi}{3}$ and $\frac{2\pi}{3}$. These angles are perturbed by $\theta_{ij}^*$ and $\bar{\theta}_{ij}^*$ for the parallelograms in

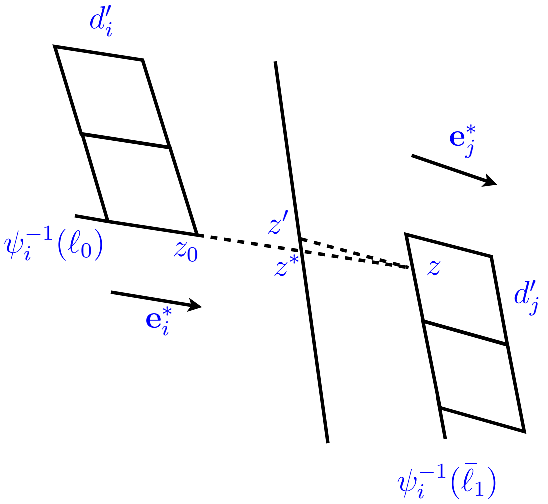


Figure 3.20: After a back transformation, the distance $|z_0 - z|$ is nearly equal to $d_i + d_j'$.

$\psi_i^{-1}(R_j)$. The perturbed angles are denoted by $\beta_{ij}$ and $\beta_{ij}'$. More precisely,

$$\begin{aligned}(\beta_{ij},\beta_{ij}') \in &\Big\{ \Big(\frac{\pi}{3}+\theta_{ij}^*+\bar\theta_{ij}^*, \frac{2\pi}{3}-\theta_{ij}^*-\bar\theta_{ij}^*\Big), \Big(\frac{\pi}{3}+\theta_{ij}^*-\bar\theta_{ij}^*, \frac{2\pi}{3}-\theta_{ij}^*+\bar\theta_{ij}^*\Big),\\ &\Big(\frac{\pi}{3}-\theta_{ij}^*+\bar\theta_{ij}^*, \frac{2\pi}{3}+\theta_{ij}^*-\bar\theta_{ij}^*\Big), \Big(\frac{\pi}{3}-\theta_{ij}^*-\bar\theta_{ij}^*, \frac{2\pi}{3}+\theta_{ij}^*+\bar\theta_{ij}^*\Big)\Big\}.\end{aligned} \tag{3.87}$$

Our first task consists in estimating the minimum distance between $z_0$ and $z$ (see (3.93)). Let $z^*$ denote the intersection of the line extending $\psi_i^{-1}(\ell_0)$ with the common edge of $\psi_i^{-1}(S_i)$ and $\psi_i^{-1}(S_j)$. Then by construction (see Step (iii) of Section 3.2.2), we have $|z_0 - z^*| \geq d_i' = |\bar{\mathbf{e}}_i^*|$.

Consider the intersection point $z'$ of the common edge $E_{ij}$ of $\psi_i^{-1}(S_i)$ and $\psi_i^{-1}(S_j)$ with the line parallel to $\mathbf{e}_j^*$ and passing through $z$, see Figure 3.20. By considering the triangle formed by $z, z^*$ and $z'$, denote by $\vartheta, \gamma$ the interior angles at $z^*, z'$, respectively. By using the sine rule, we have $\frac{|z-z^*|}{\sin\gamma} = \frac{|z-z'|}{\sin\vartheta}$ which yields

$$|z - z^*| = \frac{\sin\gamma}{\sin\vartheta}|z-z'| \geq \frac{\sin\gamma}{\sin\vartheta} d_j', \tag{3.88}$$

by virtue of the fact that by construction $|z - z'| \geq d_j'$, where $d_j'$ is the length of the shortest edge of a parallelogram in $\psi_i^{-1}(R_j)$. The right hand side of the above inequality is an increasing function of $\gamma \in [0, \frac{\pi}{2}]$. Clearly $\gamma \geq 0$, and $\gamma > \frac{\pi}{2}$

implies that $|z-z^*| \geq |z-z'|$, which does not provide the minimum lower bound possible for $|z-z^*|$. We will show below that there is an angle $\gamma_0$ for which $\gamma \geq \gamma_0$ lead to the lower bound for $|z-z^*|$.

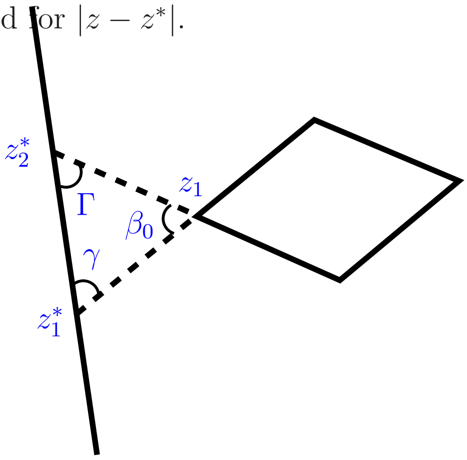


Figure 3.21: Nearest that $z_1$ can be to an edge of $\psi_i^{-1}(S_j)$

In fact, $|z-z^*|$ is longer than $|z_1-z_1^*|$ where $z_1$ is a vertex of $\psi_i^{-1}(\ell_1)$ which is nearer to $E_{ij}$ than $z$ (there is always such a vertex since $z$ belongs to a parallelogram having an edge on $\psi_i^{-1}(\ell_1)$), and $z_1^*$ the intersection point of $E_{ij}$ with the line parallel to $\mathbf{e}_j^*$ and passing through $z_1$. Clearly, $d_j'$ is a lower bound for $|z_1-z_1^*|$. We are looking for the lowest value of $\gamma$ for which the minimum $|z_1-z_1^*| = d_j'$ is attained. Such a value is determined through the following settings, as illustrated in Figure 3.21: The line parallel to $\bar{\mathbf{e}}_j^*$ and passing through $z_1$ intersects the edge $E_{ij}$ at a point $z_2^*$. By construction we also have $|z_1-z_2^*| \geq d_j'$. For the triangle formed by $z_1, z_1^*, z_2^*$, the interior angle at $z_2^*$ is denoted $\Gamma$ and the one at $z_1$ which we denote by $\beta_0$ is either $\beta_{ij}$ or $\beta_{ij}'$ (the one at $z_1^*$ is obviously $\gamma$).

Both lower bounds $|z_1-z_1^*| = d_j'$ and $|z_1-z_2^*| = d_j'$ (see Figure 3.21) are attained if $\gamma = \Gamma = \gamma_0 := \frac{\pi-\beta_0}{2}$ and if the distance of $z_1$ to the edge $E_{ij}$ is $d_j' \sin\gamma_0$. Indeed, by using the sine rule clearly $\frac{|z_1-z_1^*|}{\sin\Gamma} = \frac{|z_1-z_2^*|}{\sin\gamma}$ which yields

$$|z_1-z_1^*| = \frac{\sin\Gamma}{\sin\gamma}.$$

Thus, if $\gamma < \Gamma$ then $\frac{\sin\Gamma}{\sin\gamma} > 1$ and $|z_1-z_1^*| > |z_1-z_2^*| = d_j'$. Hence $\gamma = \gamma_0$ is the minimum angle that makes $|z_1-z_1^*|$ as short as possible. Moreover, from (3.87) the largest angle that $\beta_0$ can be is $\frac{2\pi}{3} + \theta_{ij}^* + \bar{\theta}_{ij}^*$, which yields $\gamma_0 = \frac{\pi}{6} - \frac{\theta_{ij}^*+\bar{\theta}_{ij}^*}{2}$.

Let us now come back to our previous setting with the points $z, z^*$ and $z'$ (see Figure 3.20). Since $\gamma \geq \gamma_0$, we deduce from (3.88) that

$$|z - z^*| \geq \frac{\sin \gamma_0}{\sin \vartheta} d'_j. \tag{3.89}$$

From the fact that

$$\vartheta = \pi - \gamma - \theta^*_{ij} \leq \pi - \gamma_0 - \theta^*_{ij} = \frac{5\pi}{6} + \frac{\bar{\theta}^*_{ij} - \theta^*_{ij}}{2},$$

we obtain

$$\begin{aligned}\frac{\sin \gamma_0}{\sin \vartheta} =& \frac{\cos(\bar{\theta}^*_{ij} + \theta^*_{ij}) - \sqrt{3}\sin(\bar{\theta}^*_{ij} + \theta^*_{ij})}{\cos(\bar{\theta}^*_{ij} - \theta^*_{ij}) - \sqrt{3}\sin(\bar{\theta}^*_{ij} - \theta^*_{ij})} \\ =& \frac{1 - \tan \bar{\theta}^*_{ij} \tan \theta^*_{ij} - \sqrt{3}\tan \bar{\theta}^*_{ij} - \sqrt{3}\tan \theta^*_{ij}}{1 + \tan \bar{\theta}^*_{ij} \tan \theta^*_{ij} - \sqrt{3}\tan \bar{\theta}^*_{ij} + \sqrt{3}\tan \theta^*_{ij}} \\ =& 1 - \frac{2\tan \frac{\theta^*_{ij}}{2}(\sqrt{3} + \tan \frac{\bar{\theta}^*_{ij}}{2})}{1 - \sqrt{3}\tan \frac{\bar{\theta}^*_{ij}}{2} + \tan \frac{\theta^*_{ij}}{2}(\sqrt{3} + \tan \frac{\bar{\theta}^*_{ij}}{2})},\end{aligned} \tag{3.90}$$

after expansions and simplifications of the cosines and sines. Observe that $\theta^*_{ij}$ and $\bar{\theta}^*_{ij}$ are small (Lemma 3.3.8) enough so that $-\sqrt{3}\tan \frac{\bar{\theta}^*_{ij}}{2} + \tan \frac{\theta^*_{ij}}{2}(\sqrt{3} + \tan \frac{\bar{\theta}^*_{ij}}{2}) \leq \frac{1}{2}$, from which the above inequality reads

$$\frac{\sin \gamma_0}{\sin \vartheta} \geq 1 - \varrho_{ij}, \tag{3.91}$$

where $\varrho_{ij} = 4\tan \frac{\theta^*_{ij}}{2}(\sqrt{3} + \tan \frac{\bar{\theta}^*_{ij}}{2})$ which can be very small too. We deduce from (3.89) that

$$|z - z^*| \geq (1 - \varrho_{ij})d'_j. \tag{3.92}$$

Now, observe that

$$|z_0 - z| = |z_0 - z^*| + |z - z^*| \geq d'_i + (1 - \varrho_{ij})d'_j. \tag{3.93}$$

The above estimation will be useful for each of the following two cases, whether

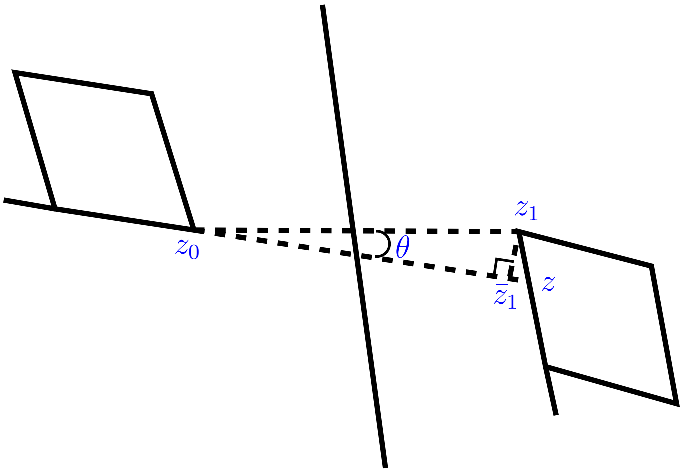


Figure 3.22: Positions of the points $z, z_0, z_1, \bar{z}_1$ when $z$ is not near a corner and $\widehat{z} = \frac{\pi}{3} \pm \bar{\theta}^*_{ij}$

or not $z$ is near to a re-entrant corner.

$*$ Let $z_1$ be the nearest vertex of $\psi_i^{-1}(\ell_1)$ such that $z_1$ is not a re-entrant corner and such that $|z-z_1| \leq \frac{d'_j}{2}$. Consider the triangle formed by $z_0, z, z_1$. Recall that the segments in $\psi_i^{-1}(\mathcal{L}_i)$ and $\psi_i^{-1}(\mathcal{L}_j)$ are almost parallel in the sense that they make small angles (and similarly for the segments in $\psi_i^{-1}(\bar{\mathcal{L}}_i)$ and $\psi_i^{-1}(\bar{\mathcal{L}}_j)$). Thus the interior angle $\widehat{z}$ at $z$ belongs to the set

$$\left\{\frac{\pi}{3} \pm \bar{\theta}^*_{ij}, \frac{2\pi}{3} \pm \bar{\theta}^*_{ij}\right\}. \tag{3.94}$$

Observe that $\theta$ is the angle between $[z_0, z]$ and $[z_0, z_1]$. We have two cases.

- Suppose that $\widehat{z} = \frac{\pi}{3} \pm \bar{\theta}^*_{ij}$, as shown in Figure 3.22. Denoting by $\bar{z}_1$ the orthogonal projection of $z_1$ onto the the segment $[z, z_0]$, we have

$$\begin{aligned} \tan\theta =& \frac{|z_1 - \bar{z}_1|}{|z_0 - \bar{z}_1|} = \frac{|z - z_1| \sin \widehat{z}}{|z_0 - z| - |z - z_1| \cos \widehat{z}} \\ \leq& \frac{\frac{d'_j}{2} \sin \widehat{z}}{d'_i + (1 - \varrho_{ij}) d'_j - \frac{d'_j}{2} \cos \widehat{z}}, \end{aligned} \tag{3.95}$$

by virtue of (3.93). Since from (3.80) we have that $d'_i = (1 \pm \delta) d'_j$ for some

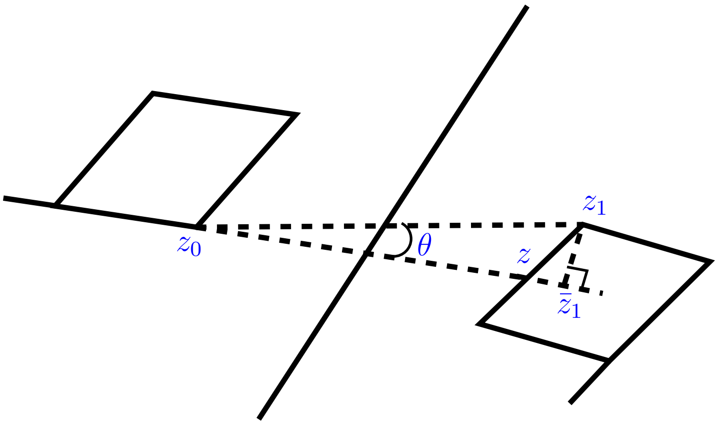


Figure 3.23: Positions of the points $z, z_0, z_1, \bar{z}_1$ when $z$ is not near a corner and $\widehat{z} = \frac{2\pi}{3} \pm \bar{\theta}^*_{ij}$

$0 \leq \delta \leq \frac{1}{10^{\frac{3}{2}}}$, after simplification by $d'_j$ we obtain

$$\tan\theta \leq \frac{\frac{1}{2}}{\frac{3}{2} - \delta - \varrho_{ij}} = \frac{1}{3 - 2(\delta + \varrho_{ij})} \leq \frac{1}{3 - \frac{1}{5}} \leq \tan\frac{\pi}{7}, \tag{3.96}$$

by choosing $\varrho_{ij}$ small enough (see (3.91)) so that $2(\delta + \varrho_{ij}) \leq \frac{1}{5}$.

●● Suppose that $\widehat{z} = \frac{2\pi}{3} \pm \bar{\theta}^*_{ij}$, as shown in Figure 3.23. Then by denoting $\bar{z}_1$ the projection of $z_1$ onto the line extending the segment $[z_0, z]$, we have

$$\begin{aligned} \tan\theta &= \frac{|z_1 - \bar{z}_1|}{|z_0 - \bar{z}_1|} \leq \frac{|z_1 - \bar{z}_1|}{|z - z_0|} \leq \frac{|z - z_1|}{|z - z_0|} \\ &\leq \frac{\frac{d'_j}{2}}{d'_i + (1 - \varrho_{ij})d'_j} \leq \frac{1}{4 - 2(\delta + \varrho_{ij})} \leq \tan\frac{\pi}{7}. \end{aligned} \tag{3.97}$$

$**$ Suppose now that $z_1$ is the nearest vertex of $\psi_i^{-1}(\ell_1)$ such that $z_1$ is not a re-entrant corner but $\frac{d'_j}{2} < |z - z_1| \leq d'_j$. Let $z^*$ denote the re-entrant corner near to $z$ and let $\tilde{z}$ be the middle point of $z^*$ and $z_1$ (see Figure 3.24). Consider the point $\tilde{z}_1$ obtained from the intersection of the segment $[z_1, \tilde{z}_2]$ with the line parallel to $\mathbf{e}^*_j$ and passing through $\tilde{z}$, where $\tilde{z}_2 \in \psi_i^{-1}(\bar{L}_j)$ is the closest vertex to $z^*$ such that the triangle defined by $z_1, z^*, \tilde{z}_2$ is not included in $\psi_i^{-1}(R_j)$. Clearly the two triangles $T, \tilde{T}$ formed by $z_1, \tilde{z}, \tilde{z}_1$ and $z_1, z^*, \tilde{z}_2$ have the same shape. Thus

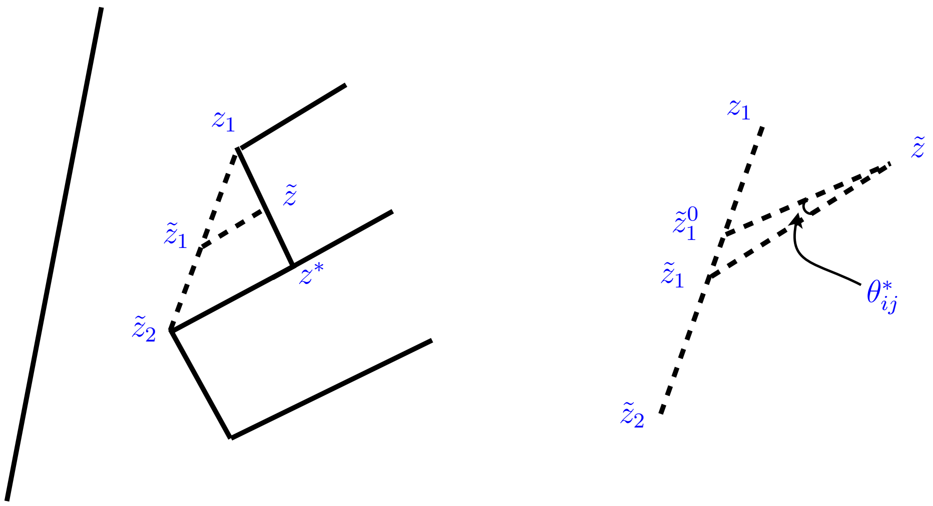


Figure 3.24: Different points in the neighborhood of a re-entrant corner.

necessarily

$$\frac{|\tilde{z}-\tilde{z}_1|}{|\tilde{z}-z_1|}=\frac{|z^*-\tilde{z}_2|}{|z^*-z_1|}.$$

The numerator and denominator in the right hand side of the above equality are equal to $d_j'$ up to a perturbation by a small number (see (3.71) and (3.73)). Since $|\tilde{z}-z_1|=\frac{|z^*-z_1|}{2}$, the above equality yields that

$$|\tilde{z}-\tilde{z}_1|=(1+\delta')\frac{d_j'}{2}, \tag{3.98}$$

where $|\delta|\leq 3\xi$ where $\xi$ is given in (3.63).

The interior angle $\widehat{z^*}$ at $z^*$ of the triangle $\tilde{T}$ belongs to the set

$$\left\{\frac{\pi}{3}\pm\theta_{ij}^*\pm\bar{\theta}_{ij}^*,\frac{2\pi}{3}\pm\theta_{ij}^*\pm\bar{\theta}_{ij}^*\right\}.$$

We again have two cases.

- Suppose that $\widehat{z^*}=\frac{\pi}{3}\pm\theta_{ij}^*\pm\bar{\theta}_{ij}^*$ (similar to Figure 3.22). Since the edge lengths of $\tilde{T}$ are approximately $d_j'$, the interior angle at $\tilde{z}_2$ is also approximately $\frac{\pi}{3}$. Thus, for the triangle $T$ the interior angle at $\tilde{z}_1$ is approximately $\frac{\pi}{3}$. Consider the triangle $\tilde{T}^0$ formed by the points $\tilde{z}$, $\tilde{z}_1$ and $\tilde{z}_1^0$ such that the point $\tilde{z}^0$ belongs to the segment $[\tilde{z}_1, z_1]$, the interior angle at $\tilde{z}$ is $\theta^*_{ij}$ and such that $|\tilde{z} - \tilde{z}^0_1| \leq |\tilde{z} - \tilde{z}_1|$. By the sine rule, we have

$$\frac{|\tilde{z} - \tilde{z}^0_1|}{\sin \widehat{\tilde{z}_1}} = \frac{|\tilde{z}_1 - \tilde{z}^0_1|}{\sin \theta^*_{ij}}. \tag{3.99}$$

On the other hand, we have

$$|\tilde{z}_1 - \tilde{z}^0_1| \cos \widehat{\tilde{z}_1} + |\tilde{z} - \tilde{z}^0_1| = |\tilde{z} - \tilde{z}_1|. \tag{3.100}$$

From (3.99), we obtain $|\tilde{z}_1 - \tilde{z}^0_1| = \frac{\sin \theta^*_{ij}}{\sin \widehat{\tilde{z}_1}} |\tilde{z} - \tilde{z}^0_1|$ which, together with (3.100) and in view of (3.98), yields that $|\tilde{z} - \tilde{z}^0_1| = \frac{|\tilde{z} - \tilde{z}_1|}{\cos \theta^*_{ij}} \frac{\tan \widehat{\tilde{z}_1}}{\tan \theta^*_{ij} + \tan \widehat{\tilde{z}_1}}$, and thus

$$|\tilde{z} - \tilde{z}^0_1| \geq |\tilde{z} - \tilde{z}_1| \Big( 1 - \frac{\tan \theta^*_{ij}}{\tan \theta^*_{ij} + \tan \widehat{\tilde{z}_1}} \Big) \geq \frac{d'_j}{2} (1 - \varrho'_{ij}), \tag{3.101}$$

where $\varrho'_{ij} = O(\tan \theta^*_{ij})$. Recall from (3.98) that $|\tilde{z} - \tilde{z}_1|$ is approximately equal to $\frac{d'_j}{2}$. Thus in a similar way that we obtain (3.93), we have that

$$\begin{aligned} |z_0 - \tilde{z}| =& |z_0 - z^*| + |z^* - \tilde{z}^0_1| + |\tilde{z}^0_1 - \tilde{z}| \\ \geq& d'_i + (1 - \varrho_{ij}) d'_j + (1 - \varrho'_{ij}) \frac{d'_j}{2} \\ =& (\frac{5}{2} - \delta - \varrho_{ij} - \frac{\varrho'_{ij}}{2}) d'_j. \end{aligned}$$

Denote by $\bar{z}_1$ the projection of $z_1$ onto the segment $[z_0, z]$. Recalling that $\theta$ is the angle between $[z_0, z]$ and $[z_0, z_1]$, we have that

$$\tan \theta = \frac{|z_1 - \bar{z}_1|}{|z_0 - \bar{z}_1|} = \frac{|z - z_1| \sin \widehat{z}}{|z_0 - z| - |z - z_1| \cos \widehat{z}},$$

where $\widehat{z} = \frac{\pi}{3} \pm \bar{\theta}^*_{ij}$ and $|z_1 - \bar{z}_1| \leq d'_j$. Hence we obtain

$$\tan \theta \leq \frac{d'_j \sin \widehat{z}}{(\frac{5}{2} - \delta - \varrho_{ij} - \frac{\varrho'_{ij}}{2}) d'_j - d'_j \cos \widehat{z}},$$

with $\sin\widehat{z} \leq \frac{\sqrt{3}}{2} + \frac{1}{2}\sin\bar{\theta}^*_{ij}$ and $\cos\widehat{z} \leq \frac{1}{2} + \frac{\sqrt{3}}{2}\sin\bar{\theta}^*_{ij}$. Thus

$$\begin{aligned}\tan\theta \leq & \frac{\frac{\sqrt{3}}{2} + \frac{1}{2}\sin\bar{\theta}^*_{ij}}{(\frac{5}{2} - \delta - \varrho_{ij} - \frac{\varrho'_{ij}}{2}) - (\frac{1}{2} + \frac{\sqrt{3}}{2}\sin\bar{\theta}^*_{ij})} \\ = & \frac{\sqrt{3} + \bar{\theta}^*_{ij}}{4 - 2\delta - 2\varrho_{ij} - \varrho'_{ij} - \sqrt{3}\bar{\theta}^*_{ij}} \\ \leq & \tan\frac{\pi}{7},\end{aligned} \tag{3.102}$$

with $\bar{\theta}^*_{ij}$ small enough (e.g. $\leq \frac{1}{100}$) so that $2\delta + 2\varrho_{ij} + \varrho'_{ij} + \sqrt{3}\bar{\theta}^*_{ij} \leq \frac{1}{5}$.

•• Suppose that $\widehat{z^*} = \frac{2\pi}{3} \pm \theta^*_{ij} \pm \bar{\theta}^*_{ij}$ (similar to Figure 3.22). We use the same notation as in the previous case. In order to bound $\tan\theta$, we shall estimate the lowest possible value for $|z_0 - z|$. This is achieved with the following setting: The point $z$ coincides with the middle point $\tilde{z}$ and the segment $[\tilde{z}, \tilde{z}^0_1]$ is shorter than $[\tilde{z}, \tilde{z}_1]$ whose length is approximately $\frac{d'_j}{2}$ (see (3.98)).

The difference here is that $\bar{z}_1$ is the projection of $z_1$ onto the line extending $[z_0, z]$. In this case we have

$$\tan\theta = \frac{|z_1 - \bar{z}_1|}{|z_0 - \bar{z}_1|} \leq \frac{|z_1 - z|}{|z_0 - \bar{z}_1|}.$$

The numerator is clearly less than $d'_j$. For the denominator, using a similar method as in the previous case yields that

$$|\tilde{z} - \tilde{z}^0_1| \geq |\tilde{z} - \tilde{z}_1|\left(1 - \frac{\tan\theta^*_{ij}}{\tan\theta^*_{ij} + \tan\widehat{\tilde{z}_1}}\right). \tag{3.103}$$

Since $\widehat{z^*} = \frac{2\pi}{3} \pm \theta^*_{ij} \pm \bar{\theta}^*_{ij}$ and $|z^* - z_1|, |z^* - \tilde{z}_2|$ have lengths approximately $d'_j$, the triangle $T$ is nearly isosceles and we have $\widehat{\tilde{z}_1} \approx \frac{\pi}{6}$. Thus

$$\begin{aligned}|z_0 - z| = & |z_0 - \bar{z}^*| + |\bar{z}^* - \tilde{z}^0_1| + |\tilde{z}^0_1 - \tilde{z}| \\ \geq & d'_i + (1 - \varrho_{ij})d'_j + (1 - \varrho'_{ij})\frac{d'_j}{2},\end{aligned}$$

which, by using a similar argument as when proving (3.102), yields

$$\begin{aligned}\tan\theta \leq & \frac{\frac{1}{2}+\frac{\sqrt{3}}{2}\sin\bar{\theta}^*_{ij}}{(\frac{5}{2}-\delta-\varrho_{ij}-\frac{\varrho'_{ij}}{2})-(\frac{\sqrt{3}}{2}+\frac{1}{2}\sin\bar{\theta}^*_{ij})} \\ = & \frac{1+\sqrt{3}\bar{\theta}^*_{ij}}{5-\sqrt{3}-2\delta-2\varrho_{ij}-\varrho'_{ij}-\bar{\theta}^*_{ij}} \\ \leq & \tan\frac{\pi}{7}. \end{aligned} \tag{3.104}$$

The second part of the proof is given below (the first part starts on page 138).

ii-b. Suppose that $\ell_0$ is extended according to Setting 2-b. We adopt the same notation as in the previous case, and illustrations are similar to those already shown in Figures 3.22-3.23-3.24. We consider the triangle formed by $z_0, z_1, z$ and denote by $\widehat{z}$ the interior angle at $z$. The angle $\theta$ is the angle between the segments $[z_0, z]$ and $[z_0, z_1]$. We again have two cases.

- Suppose that $\widehat{z} = \frac{\pi}{3} \pm \bar{\theta}^*_{ij}$. Denote by $\bar{z}_1$ the projection of $z_1$ onto the segment $[z_0, z]$. Then, we have

$$\tan\theta = \frac{|z_1-\bar{z}_1|}{|z_0-\bar{z}_1|} = \frac{|z_1-z|\sin\widehat{z}}{|z_0-z|-|z_1-z|\cos\widehat{z}}.$$

With $L' = |z_0 - z|$ and $|z_1 - z| \leq \frac{d'_j L'}{8d'_i}$, we obtain

$$\tan\theta \leq \frac{\frac{d'_jL'}{8d'_i}}{L'-\frac{d'_jL'}{8d'_i}} = \frac{d'_j}{8d'_i-d'_j} = \frac{1}{7-8\delta} \leq \tan\frac{\pi}{7}, \tag{3.105}$$

by virtue of the fact that $d'_i = (1-\delta)d'_j$.

●● Suppose that $\widehat{z} = \frac{2\pi}{3} \pm \bar{\theta}^*_{ij}$. Let $\bar{z}_1$ be the projection of $z_1$ onto the line extending the segment $[z_0, z]$. Then

$$\begin{aligned}\tan\theta = & \frac{|z_1-\bar{z}_1|}{|z_0-\bar{z}_1|} \leq \frac{|z_1-\bar{z}_1|}{|z_0-z|} \leq \frac{|z_1-z|}{|z_0-z|} \leq \frac{\frac{d'_jL'}{8d'_i}}{L'} \\ = & \frac{1}{8-8\delta} \leq \tan\frac{\pi}{7}. \end{aligned} \tag{3.106}$$

A similar argument is used if $\ell_0$ is extended by using Setting 3. This concludes our proof. □

In Proposition 3.3.9-(ii), the angle $\gamma^*$ will be different if we change some parameters in the algorithms in Section 3.2.2.

- The minimum distance between a vertex in a regular region to the edge of the sub-square that contains it is obtained from step (iii) of Section 3.2.2. Changing this distance results in different estimations of $|z_0 - z|$ in (3.93) of part ii-a;

- Changing the factor $\frac{1}{8}$ in the tolerance $\frac{d_j L}{8h_i}$ introduced in Setting 2-b also results in a different tolerance, and thus (3.105) and (3.106) will change accordingly;

- The parameter $r$ should satisfy (3.17) so that $d'_i = (1 \pm \delta) d'_j$, where $|\delta| < \frac{1}{10^{\frac{3}{2}}}$ allowing us to prove (3.96) and (3.97);

- The parameter $s$ needs to be small enough so that in each sub-square the area covered by irregular regions is sufficiently small (see (3.3.12) of Section 3.4), and so that one of the connections in Setting 1-4 occurs. For instance, $s$ should be small enough so that at times the tolerance (on some enlarged segment) is larger than twice the maximum distance between a vertex of a regular region and the edge of the sub-square that contains it, i.e there is $L'_0$ such that $\frac{d'_j L'_0}{8d'_i} \geq 2d'_j$ and for which

$$16d'_i \leq L' \leq 17d'_i. \tag{3.107}$$

  The above inequalities say that the system of parallel segments $\psi_i^{-1}\left(\bar{\mathcal{L}}_j\right)$ should have at least 16 segments. This can be easily achieved since $d'_i = h_0 \Lambda_i$, where $\Lambda_i$ has the factor $s^\eta$. Other conditions on $s$ are imposed in (3.112) in Section 3.4.

We now present the following result.

**Proposition 3.3.10.** *Let $r$ be sufficiently small so that (3.17) holds, and let $T \in \Delta_{s,r}$ be a triangle whose barycenter is inside a sub-square $R_i$. Then, the interior angles of $\psi_i^{-1}(T)$ are at most $\pi - \frac{\pi}{50} = \frac{49\pi}{50}$.*

*Proof.* We present several cases depending on how the triangle $T$ is obtained.

• If $T_0$ is a regular triangle of $R_i$, then $\psi_i^{-1}(T_0)$ is an equilateral triangle, its interior angles are exactly $\frac{\pi}{3}$;

• Suppose that $T_0$ was already a triangle before the triangulation performed in Section 3.2.4, and that it has no vertex on the boundary of $\Omega$. Then, it necessarily has one edge obtained from Setting 1. We have two exclusive cases:

- the other two edges of $T_0$ are part of the boundary of the regular region $R_i$ ($P_0$ is then necessarily contained in $S_i$). This implies that the interior angles of $\psi_i^{-1}(T_0)$ are necessarily less than the maximum interior angle of a parallelogram obtained after back transformation, that is $\frac{2\pi}{3}$;
- the other two edges are obtained by the intersection of two extended segments. By taking into account the result *ii* of Proposition 3.3.9, and knowing that the angles in a parallelogram after back transformation are $\frac{\pi}{3}$ and $\frac{2\pi}{3}$ which can be altered by at most $\theta^* < \frac{\pi}{7}$ (defined by (3.56)), the maximum interior angle of $T_0$ is bounded by $\frac{2\pi}{3} + \theta^* + \gamma^* < \frac{20\pi}{21}$, with $\gamma^* \leq \frac{\pi}{7}$ from *ii* of Proposition 3.3.9.

• Suppose that $T_0$ was already a triangle before the triangulation performed in Section 3.2.4 and that $T_0$ has a vertex on the boundary of $\Omega$. Then, it necessarily has an edge $\ell$ overlapping the boundary and the other two edges are obtained by the intersections of two extended segments from Setting 4 or 5. Note that after the back transformation by $\psi_i$ the interior angles of a regular parallelogram in $S_i$ are $\frac{\pi}{3}$ and $\frac{2\pi}{3}$. These angles are altered by at most $\theta^*$ (defined by (3.56)) in any neighboring sub-square. Thus the interior angle formed from the intersection of the two extended segments is at most $\frac{2\pi}{3} + \theta^*$, with obviously $\theta^* < \frac{\pi}{7}$. The interior angles $\theta_0$ and $\bar{\theta}_0$ at the end-points of $\ell$ respectively belong to some sets $\Theta_k$

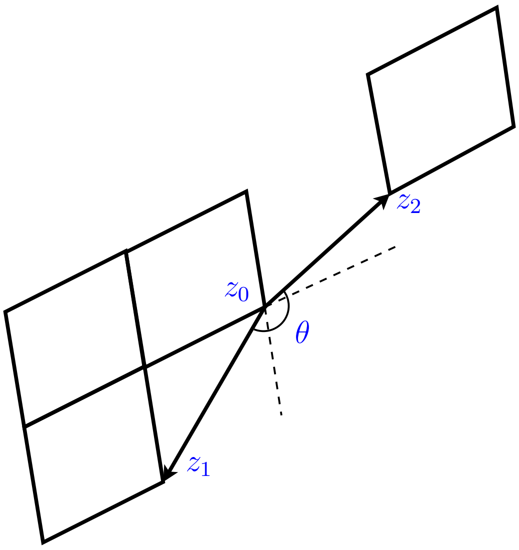


Figure 3.25: An example of big angle: $z_0$ is connected to $z_1$ by using Setting 1, and to $z_2$ by using either Setting 2 or Setting 3.

and $\bar{\Theta}_{k'}$ (both defined in (3.53) and (3.54)) for some $k, k' \in \{i, j\}$, with $j$ being the index of the neighboring sub-square $S_j$ which is the closest (apart from $S_i$) to the barycenter of $T_0$. We observe that $\theta_0$ necessarily belongs to one of the two smallest angles in $\Theta_k$, and similarly for $\bar{\theta}_0$ it belongs to one of the two smallest angles in $\bar{\Theta}_{k'}$. Thus both cannot be greater than $\frac{\pi}{2}$.

- Suppose now that $T_0$ was contained in some irregular polygon $P_0$ before the triangulation performed in Section 3.2.4. Suppose that $P_0$ has no intersection with the boundary of $\Omega$. Then $P_0$ is a quadrilateral whose edges are obtained from the intersections of some extended segments from Setting 1-5. Recall the angle $\gamma^* \leq \frac{\pi}{7}$, as described in *ii* of Proposition 3.3.9. We have two cases:

- If none of the edges of $P_0$ are obtained by Setting 1: Recall from Lemma 3.3.5 that extended segments from the same type cannot intersect when using Setting 2-a, 2-b and thus Setting 3. Obviously, there are no intersections of extended segments of the same type when using Setting 4-5. Hence, the maximum interior angle of $\psi_i^{-1}(P_0)$ must be less than the maximum interior angle $\frac{2\pi}{3}$ of a parallelogram, altered by at most twice of $\gamma^*$. Indeed, at the vertex where the interior angle is $\frac{2\pi}{3}$, there are two extensions creating turning of segments, each by at most $\gamma^*$. Hence the bound $\frac{2\pi}{3} + \frac{2\pi}{7} = \frac{20\pi}{21}$

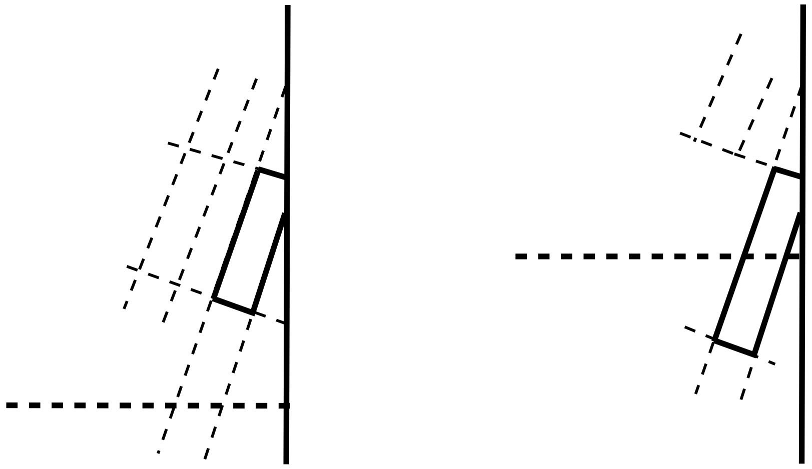


Figure 3.26: Interface regions in the neighborhood of the boundary of $\Omega$.

for the interior angles of $P_0$, and therefore of $T_0$;

- If one edge $\ell$ of $P_0$ is obtained by using Setting 1: Suppose that the maximum interior angle of $\psi_i^{-1}(P_0)$ is at its vertex $z_0$ (see Figure 3.25). Then necessarily $z_0$ is a corner of $\psi_i^{-1}(R_i)$. We deduce from Proposition 3.3.9 that the interior angle at $z_0$ is less than $\frac{2\pi}{3} + \frac{\pi}{6} + \gamma^* \leq \frac{41\pi}{42}$. This is also the bound for the interior angles of $T_0$.

• Suppose that $T_0$ was contained in some irregular polygon $P_0$ before the triangulation performed in Section 3.2.4. Suppose that $P_0$ intersects the boundary of $\Omega$. In fact, we can apply the above analysis in the case where the intersection is a point. Hence, we assume that $P_0$ possesses an edge $\ell$ overlapping the boundary of $\Omega$. Suppose that the maximum interior angle of $P_0$ is at its vertex $z_0 \in \ell_0$ (see Figure 3.26) which is the common end-point of the two edges $\ell$ and $\ell_0$ of $P_0$, where $\ell_0$ is part of some extended segment obtained from Setting 4. Consider the edge $\ell_1 = [z_2, z_1]$ of $P_0$ sharing a vertex $z_1 \neq z_0$ with $\ell_0$, and let $T_1$ be the triangle formed by the end-points of $\ell_0$ and $\ell_1$ (see Figure 3.27). Denote by $\vartheta_0$ the interior angle of that triangle at $z_0$, and by $\vartheta_1$ the one at $z_1$. Then either $\vartheta_1 \approx \frac{\pi}{3}$ or $\vartheta_1 \approx \frac{2\pi}{3}$ in the sense that either $|\vartheta_1 - \frac{\pi}{3}| \leq \theta^*$ or $|\vartheta_1 - \frac{2\pi}{3}| \leq \theta^*$, where $\theta^*$ is defined in (3.56).

As in the proof of Proposition 3.3.9, let the side length of a parallelogram in $\psi_i^{-1}(R_i)$ be denoted by $d_i'$. We have the following cases (see Figure 3.26 and

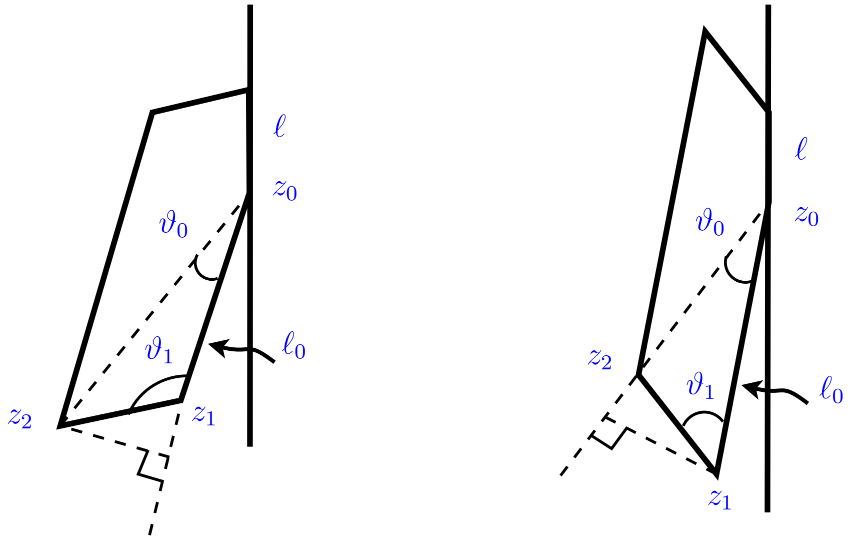


Figure 3.27: Interior angles of irregular regions.

Figure 3.27):

*a.* If $\ell_1$ is part of an extended segment from $R_i$, then $|\ell_1| = |z_2 - z_1| \geq \frac{1}{8}d'_i$ which is the minimum length of a tolerance. In this case, since either $\vartheta_1 \approx \frac{\pi}{3}$ or $\vartheta_1 \approx \frac{2\pi}{3}$, we have $|\ell_0| = |z_1 - z_0| \leq d'_i$ and $|z_2 - z_0| \leq d'_i$. On one hand, if $\vartheta_1$ is approximately $\frac{2\pi}{3}$, we have

$$\tan\vartheta_0 \geq \frac{d'_i}{8|z_2 - z_0|} \geq \frac{1}{8} \geq \tan\frac{\pi}{30}.$$

On the other hand, if $\vartheta_1$ is approximately $\frac{\pi}{3}$, then since $|z_1 - z_0| \leq d'_i$,

$$\tan\vartheta_0 \geq \frac{1}{8|z_1 - z_0|} \geq \frac{1}{8} \geq \tan\frac{\pi}{30};$$

*b.* If $\ell_1$ is not part of an extended segment from $R_i$, i.e part of an extended segment from a neighboring regular region $R_j$, then $|\ell_1| = |z_2 - z_1| = d'_j$ (see proof of Proposition 3.3.9 for the descriptions of $d'_i, d'_j$). Also, we have $|\ell_0| \leq 6\sqrt{2}(1+\xi)d'_j$, with $\xi$ as in (3.63). Moreover, since either $\vartheta_1 \approx \frac{\pi}{3}$ or $\vartheta_1 \approx \frac{2\pi}{3}$, we have that $|z_2 - z_0| \leq 6\sqrt{2}(1+\xi)d'_j$. Thus, if $\vartheta_1$ is approximately

$\frac{2\pi}{3}$, we have

$$\tan\vartheta_0 \geq \frac{d'_j}{|z_2 - z_0|} \geq \frac{1}{6\sqrt{2}(1+\xi)} \geq \frac{\sqrt{2}}{18} \geq \tan\frac{\pi}{50}.$$

Similarly, if $\vartheta_1$ is approximately $\frac{\pi}{3}$, then

$$\tan\vartheta_0 \geq \frac{d'_j}{|z_1 - z_0|} \geq \frac{1}{6\sqrt{2}(1+\xi)} \geq \frac{\sqrt{2}}{18} \geq \tan\frac{\pi}{50}.$$

From the above analysis, we conclude that $\vartheta_0 \geq \frac{\pi}{50}$. This implies that after a back transformation, the maximum interior angles in any irregular polygon decreases by at least $\frac{\pi}{50}$. □

### 3.3.4 On the area covered by irregular triangles

We respectively denote by $\Delta_{s,r}^{\text{reg}}$ and $\Delta_{s,r}^{\text{irr}}$ the sets of regular and irregular triangles in $\Delta_{s,r}$.

Let $h_M := \sup_{T\in\Delta_{s,r}^{\text{reg}}} h_T$ denote the longest edge of a regular triangle. Given a sub-square $S_i$, with $i \in \{1,\dots,m^2\}$, there are four *quarter-disks* $\mathcal{D}_i^{(1)}, \mathcal{D}_i^{(2)}, \mathcal{D}_i^{(3)}, \mathcal{D}_i^{(4)}$ of radius $4\sqrt{2}h_M$ and each centered at the four corners of $S_i$, see Figure 3.28.

A quarter-disk $\mathcal{D}_i^k$ is big enough to contain at least one parallelogram of the regular region. Also, from a geometric viewpoint, the following properties are observed:

$(i)$ If $\mathcal{D}_i^k$ has a side on the boundary of $\Omega$, then that side must possess the end-point of an extended segment obtained from Setting 4;

$(ii)$ If the center of the quarter-disk is a corner of $\Omega$ (two of its sides are on the boundary of $\Omega$), then $\mathcal{D}_i^k$ contains at least two extended segments from Setting 4 connecting vertices of $R_i$ with the the boundary of $\Omega$. It cannot contain an extended segment from Setting 2-3;

$(iii)$ A half-disk formed by two quarter-disks $\mathcal{D}_i^k, \mathcal{D}_j^{k'}$ must contain an extended segment obtained from Setting 2-a.

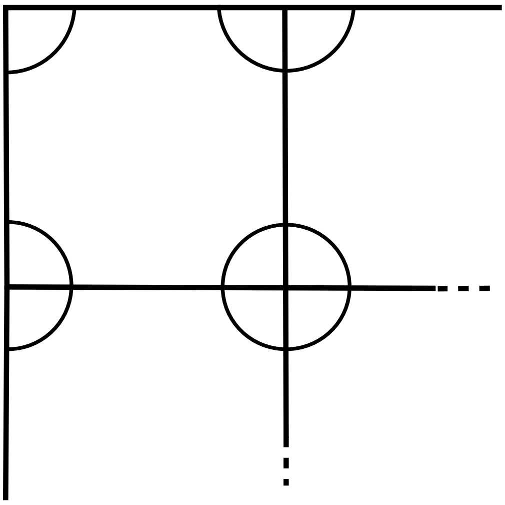


Figure 3.28: Quarter-disks of radius $4\sqrt{2}h_M$.

The first property is useful to ensure that for an extended segment intersecting a quarter-disk, the part of the intersected segment that is inside the quarter-disk has a length less than the diameter of the quarter-disk. The second property is a consequence of the first, whereas the third is by construction.

We are able prove the following result.

**Proposition 3.3.11.** *Let $r$ be sufficiently small so that (3.17) holds. There is an absolute constant $C_*$ such that for any irregular triangle $T_1 \in \Delta_{s,r}^{irr}$, we have $h_{T_1} \le C_* \sup_{T\in\Delta_{s,r}^{reg}} h_T$.*

*Proof.* To prove the above result, we first investigate the length of the longest edge that an irregular polygon $P$ has. Note that an edge $e$ of an irregular region is either an edge of a parallelogram or part of an extended segment, with the latter case detailed below.

Suppose that $e$ is part of an extended segment $\ell_0^{\text{ext}}$ obtained from the extension of $\ell_0 \in \mathcal{L}_i$ in the direction of $\mathbf{e}_i$, for some $i \in \{1, \dots, m^2\}$ (a similar argument is used if $\ell_0 \in \bar{\mathcal{L}}_i$ is extended in the direction $\bar{\mathbf{e}}_i$).

Setting 1: Obviously, if $\ell_0^{\text{ext}}$ is obtained from Setting 1 then its length is less than $\sqrt{2}h_M$;

Setting 4: Suppose that $\ell_0^{\text{ext}}$ is obtained from Setting 4 and denote its end-points by $z_0, z$, with $z_0$ belonging to the regular region $R_i$, and $z$ to the boundary of $\Omega$. Recall the families of segments described after the steps of Setting 5 in Section 3.2.3.

a. Suppose that $\ell_0^{\text{ext}}$ does not intersect any quarter-disk: If $\ell_0^{\text{ext}}$ is cut by segments from the other family of extended segments, then the length of the left over cut segment is less than $h_M$. Otherwise, by construction, the length of $\ell_0^{\text{ext}}$ is less than $2h_M$;

b. Suppose that $\ell_0^{\text{ext}}$ intersects a quarter-disk:

- If $z_0, z$ belong to the quarter-disk, then $|z_0 - z|$ is less than the diameter of that quarter-disk which is also less than twice the radius $8\sqrt{2}h_M$;
- If $z_0$ is in a quarter-disk and $z$ not, then similarly to Case a the length of $\ell_0^{\text{ext}}$ is less than $2h_M$;
- If $z$ is in a quarter-disk and $z_0$ not: Whether or not $\ell_0^{\text{ext}}$ is cut by some extended segment(s) from the other family, the length of the left over cut segment is less than the diameter of the quarter-disk. That diameter is also less than twice the radius $8\sqrt{2}h_M$ of the quarter-disk.

Setting 2-a: Suppose that $\ell_0$ is extended by using Setting 2-a and the direction $\mathbf{e}_i$. The resulting extended segment connects a vertex $z_0$ of $R_i$ with a vertex $z_1$ of $R_j$. We shall first estimate the length of the segment $[z_0, z]$ where $z$ is the intersection point of the line $\ell_0^{\text{line}}$ with a segment $\ell_1 \in \bar{\mathcal{L}}_j$, as described in Setting 2-a. The lengths $|z_0 - z_1|$ and $|z_0 - z|$ are comparable (i.e. one is less than $n$-times of the other, for some $n \in \mathbb{N}$) if the latter can be compared to $h_i$ (or $h_j$): The angle that $\ell_0^{\text{line}}$ makes with $\ell_1$ belongs to

$$\{\beta_i \pm \bar{\theta}_{ij},\, \pi - \beta_i \pm \bar{\theta}_{ij},\, \beta_i + 2\alpha_i \pm \bar{\theta}_{ij},\, \pi - \beta_i - 2\alpha_i \pm \bar{\theta}_{ij}\},$$

where $\beta_i$ and $\alpha_i + 2\alpha_i$ are the interior angles of a parallelogram in $R_i$,

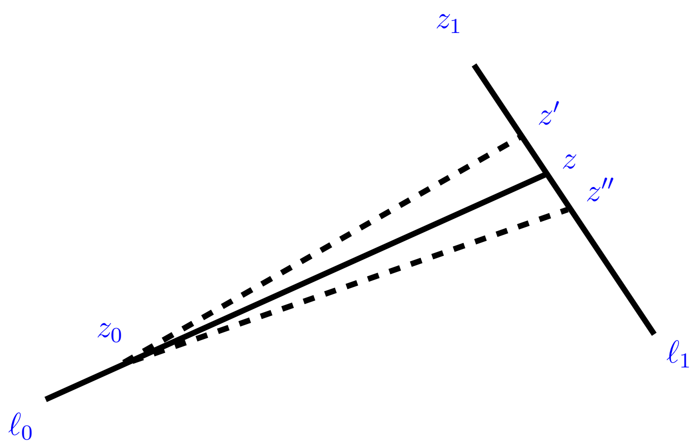


Figure 3.29: The points $z', z''$ close to $z$.

described in (3.48) and (3.55), and both are less than $\frac{\pi}{2}$. The length of $|z_0 - z|$ is less or equal to the diameter of the triangles formed by $z_0, z, z'$ and $z_0, z, z''$ where $[z', z'']$ is the part of $\ell_1$ with length $2h_j$ and midpoint $z$ (see Figure 3.29). Amongst the possible angles that $\ell_0^{\text{line}}$ makes with $\ell_1$, the angle $\beta_i - \bar{\theta}_{ij}$ (or equivalently $\pi - \beta_i + \bar{\theta}_{ij}$) creates the case for which one of the lengths $|z_0 - z'|, |z_0 - z''|$ is the longest possible. The angles at $z$ are $\beta_i - \bar{\theta}_{ij}$ and $\pi - \beta_i + \bar{\theta}_{ij}$, both being less than $\pi$. Hence $\max\{|z_0 - z'|, |z_0 - z''|\} \leq |z_0 - z| + h_j$. Clearly $|z_0 - z_1| \leq \max\{|z_0 - z'|, |z_0 - z''|\}$, and thus the lengths of $|z_0 - z_1|$ and $|z_0 - z|$ are comparable provided that the latter can be compared to $h_i$.

We now show that the length $|z_0 - z|$ is comparable to $h_i$: Consider the point of intersection $z^*$ of $\ell_0^{\text{line}}$ with the common edge $E_{ij}$ of $S_i$ and $S_j$. By using the argument for the case Setting 4 above, we have $|z_0 - z^*| \leq 8\sqrt{2}h_M$.

Let $z'''$ be the intersection point of $E_{ij}$ with the line parallel to $\mathbf{e}_j$ and passing through $z$. By using the argument for the case Setting 4 above, we have $|z - z'''| \leq 8\sqrt{2}h_M$. The angle between the segments $[z, z''']$ and $[z, z^*]$ is $\theta_{ij}$. In view of (3.17) $\theta_{ij}$ is small enough so that the lengths $|z - z'''|, |z - z^*|$ are comparable, (i.e. one is less than $k$-times of the other, for some $k \in \mathbb{N}$, say $k \leq 10$) even if the parallelograms are strongly anisotropic and have sides

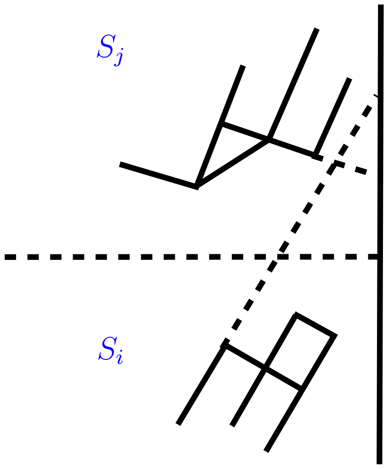


Figure 3.30: Extensions from Setting 4 near the boundary of $\Omega$.

nearly parallel to the sides of the sub-square. Thus,

$$|z^* - z| \leq (8\sqrt{2} + k)h_M, \quad k \leq 10. \tag{3.108}$$

Setting 2-b: Similarly to the previous cases, denote by $z$ the intersection point of the line extending $\ell_0$ with a line extending a segment of $\bar{\mathcal{L}}_j$ as described in Setting 2-b. We can apply the same argument as in the previous cases (above for Setting-4 and 2-a). Denote by $z^*$ the intersection point of $E_{ij}$ with the segment $[z_0, z]$. The length $|z_0 - z^*|$ is bounded by $8\sqrt{2}h_M$ as previously proved. Thus, although the point $z$ does not belong to $R_j$ but is at a distance of at most $h_M$, we can estimate $|z^* - z|$ by using a similar argument as in the case Setting 2-a above, and obtain an estimation of the form (3.108) with and additional term $h_M$ on the right hand side.

Setting 3: This case relies on the results in Setting 2-a and 2-b.

Setting 5: The length of any segments obtained from Setting 5 are less than the diameter of a quarter-disk, also less than the diameter of a half-disk, i.e. $8\sqrt{2}h_M$.

Hence, the length of an edge of an irregular region is at most $Ch_M$ where $C$ is an absolute constant. This is the upper bound for the lengths of an irregular triangle obtained by intersections of extended segments. If the polygon $P$ is a quadrilateral, then its diameter is less than twice its longest edge, thus the bound $2Ch_M$. If the polygon is a pentagon, its diameter is less than three times its longest edge, i.e. $3Ch_M$, and if the polygon is an hexagon, its diameter is less than four times the longest edge, i.e. $4Ch_M$. Hence, the diameter of an irregular triangle after the final splitting in Section 3.2.4 is less than the diameter of the polygon that contains it originally, thus the bound $4Ch_M$. □

In the result below, we estimate the space covered by irregular triangles in a sub-square.

**Proposition 3.3.12.** *Let $r$ be sufficiently small so that (3.17) holds. For each $i \in \{1, \dots, m^2\}$, the square $\omega_i$ centered at $b_i$ and with side lengths $r - 8\sqrt{2}\sup_{T\in\Delta_{s,r}^{reg}} h_T$ does not intersect any irregular triangle. Therefore, the area covered by irregular triangles satisfies*

$$\sum_{T\in\Delta_{s,r}^{irr}} |T| \leq 8rm^2 \sup_{T\in\Delta_{s,r}^{reg}} h_T. \tag{3.109}$$

*Proof.* Let $T_1 \in \Delta_{s,r}^{\mathrm{irr}}$ be an irregular region which intersects $S_i$. It is clear that a vertex $v \in T_1$ cannot be inside $\omega_i$, otherwise all of the vertices $v \pm \mathbf{e}_i$ and $v \pm \bar{\mathbf{e}}_i$ must be vertices of $R_i$, which means that $v$ is interior to the regular region $R_i$, thereby leading to a contradiction. Since irregular regions are obtained by extensions of segments and their intersections, clearly the vertices of $T_1$ are either outside of $\omega_i$, or coincide with some of the vertices of $R_i$ which are also outside of $\omega_i$.

In order to estimate the area covered by irregular regions in $S_i$, we simply observe that such an area is less than the area formed from four rectangles of sides lengths $r$ and $2h_T$, with $T$ being a regular triangle of $R_i$, that is,

$$4\Big(r \cdot 2h_T\Big) = 8rh_T, \tag{3.110}$$

with the factor 2 due to the removal of vertices as described in Section 3.2.2, and the factor 4 due to the four sides of $S_i$. The proof is obtained by replacing $h_T$ with $\sup_{T\in\Delta_{s,r}^{\mathrm{reg}}}$, then summing up over all the $m^2$ sub-squares. □

## 3.4 Asymptotic error estimations in $L_p$ and $W_p^1$

By using the triangulation $\Delta_{s,r}$ constructed in Section 3.2, we present the error bounds in $L_p$-norm and $W_p^1$-seminorm resulting from the approximation of $f$ on $\Omega$. We assume that $1 \leq p < \infty$. The analysis is marked by the separation of the errors on regular regions and irregular regions. Recall that regular triangles which define the regular regions are isosceles (see Section 3.2.2), so that the estimations on them are less tedious as compared to the estimations on irregular triangles which may have arbitrary shapes.

### 3.4.1 $L_p$-norm of the error on regular regions

From the construction of the triangulation in Section 3.2, a regular triangle $T \subset R_k$, $k \in \{1, \ldots, m^2\}$, is an isosceles triangle of the form

$$T = \pm\Lambda_k T_0 + \mathbf{t}, \tag{3.111}$$

with $T_0$ being the nearly-optimal triangle obtained by using Algorithm 3.1 with $\pi = \pi_k$, where $\Lambda_k = s^{\eta}\Big(K_p(\pi_k) + 2Ch_{\pi_k}^2\omega(r)\Big)^{-\frac{q}{2}}$ with $h_{\pi_k}$ denoting the diameter of an optimal triangle for $\pi_k = \pi_{b_k}$, and where $\mathbf{t}$ is a translation vector so that $T \subset R_k$ (see Step (i) and (iii) of Section 3.2.2), and where also $1+\frac{1}{2p} > \eta > 1+\frac{1-q}{8p}$.

We now assume that $s$ is sufficiently small so that, for all $k = 1, \ldots, m^2$,

$$s^{\frac{1}{2p}} \leq \Big|\frac{\lambda_{1,k}}{\lambda_{2,k}}\Big| \quad \text{and} \quad s^{\eta-1-\frac{1-q}{8p}} \leq |\lambda_{1,k}|^{\frac{q}{2}}. \tag{3.112}$$

Note that $\frac{1}{2p} \geq \frac{1-q}{8p}$ holds since $0 < q \leq 1$.

We prove Proposition 3.4.1 by using a similar argument as in [29].

**Proposition 3.4.1.** *Given $k \in \{1, \dots, m^2\}$, consider a regular triangle $T \subset R_k$ as given in (3.111). If $r$ is sufficiently small so that $K_p(\pi_z) \geq C_p\omega(r)$ holds for all $z \in T$, then*

$$h_T \leq C_1 s \quad \textit{and} \quad |T| \leq s^{2\eta}\Big(K_p(\pi_{b_k}) + Ch_{\pi_k}^2\omega(r)\Big)^{-q} \tag{3.113}$$

*hold for some constant $C_1$, with $C$ being the constant occurring in (3.116).*

*If, moreover, $\max_{z\in T}|z - b_k| \leq r$ and $C_1 s \leq r$ hold, then*

$$e_T(f)^p \leq s^{2p\eta}\int_T \left(\Big(K_p(\pi_z) + Ch_{\pi_k}^2\omega(r)\Big)^{1-q} + C_{p,\delta_f}\omega(\sqrt{2}r)^{\frac{1}{2}}\right)^p dz, \tag{3.114}$$

*where $C_{p,\delta_f}$ is a constant depending only on $\delta_f$ and $p$.*

*Proof.* First, assume that $T$ is a regular triangle obtained without adjustment of angle described in Proposition 3.2.4. Recall from Lemma 2.4.5 that $h_{T_0} \sim \left|\frac{\lambda_{2,k}}{\lambda_{1,k}}\right|^{\frac{1}{4}}$. By using simple inequalities, observe that $\Big(K_p(\pi_{b_k}) + 2Ch_{\pi_k}^2\omega(r)\Big)^{\frac{q}{2}}$ has as lower bounds

$$2^{\frac{q}{2}}\omega(r)^{\frac{q}{2}}h_{T_0}^{q} \sim \omega(r)^{\frac{q}{2}}\left|\frac{\lambda_{2,k}}{\lambda_{1,k}}\right|^{\frac{q}{4}} \quad \text{and} \quad K_p(\pi_{b_k})^{\frac{q}{2}} \sim |\lambda_1\lambda_2|^{\frac{q}{4}}.$$

Hence, from (3.111), there is a constant $C_1$ such that the diameter $h_T = \Lambda_k h_{T_0}$ of $T$ satisfies

$$\begin{aligned} h_T &\leq \frac{s^\eta h_{T_0}}{\Big(K_p(\pi_{b_k}) + 2Ch_{T_0}^2\omega(r)\Big)^{\frac{q}{2}}} \\ &\leq C_1 s^\eta \min\left\{\omega(r)^{-\frac{q}{2}}\left|\frac{\lambda_{2,k}}{\lambda_{1,k}}\right|^{\frac{1}{4}-\frac{q}{4}}, |\lambda_{1,k}\lambda_{2,k}|^{-\frac{q}{4}}\left|\frac{\lambda_{2,k}}{\lambda_{1,k}}\right|^{\frac{1}{4}}\right\} \\ &= C_1 s^\eta \left|\frac{\lambda_{2,k}}{\lambda_{1,k}}\right|^{\frac{1-q}{4}} \min\left\{\omega(r)^{-\frac{q}{2}}, |\lambda_{1,k}|^{-\frac{q}{2}}\right\}. \end{aligned} \tag{3.115}$$

By using (3.112), the inequalities $\left|\frac{\lambda_{2,k}}{\lambda_{1,k}}\right|^{\frac{1-q}{8p}} \leq s^{-\frac{1-q}{8p}}$ and $s^{\eta-1-\frac{1-q}{4}}|\lambda_{1,k}|^{-\frac{q}{2}} \leq 1$ hold. We then deduce from (3.115) that

$$h_T \leq C_1 s^{\eta-\frac{1-q}{4}}|\lambda_{1,k}|^{-\frac{q}{2}} = C_1 s\Big(s^{\eta-1-\frac{1-q}{4}}|\lambda_{1,k}|^{-\frac{q}{2}}\Big) \leq C_1 s.$$

Recall from the proof of Lemma 2.4.1 that there is a constant $C$ such that

$$K_p(\pi) - K_p(\pi') \le Ch_{\bar{T}}^2 \|\pi - \pi'\|, \tag{3.116}$$

whenever $\bar{T} \in \Delta_p(\pi')$, with $\pi, \pi' \in \mathbb{H}_2$. Let the point $z \in T$ be fixed. By using (3.116), we have $K_p(\pi_z) - K_p(\pi_{b_k}) \le Ch_{\pi_k}^2 \|\pi_z - \pi_{b_k}\| \le Ch_{\pi_k}^2 \omega(r)$ which yields $0 \le K_p(\pi_z) - Ch_{\pi_k}^2 \omega(r) \le K_p(\pi_{b_k})$ for all $z \in T$. The area $|T| = \Lambda_k^2 |T_0| = \Lambda_k^2$, with

$$|T| = s^{2\eta}\Big(K_p(\pi_{b_k}) + 2Ch_{\pi_k}^2 \omega(r)\Big)^{-q} \le s^{2\eta}\Big(K_p(\pi_z) + Ch_{\pi_k}^2 \omega(r)\Big)^{-q}.$$

Using (3.27) of Proposition 3.2.4 for $\varepsilon = \frac{3}{2}\omega(\sqrt{2}r)$, Proposition 2.5.4 with $c_1 = \Lambda_k$, and the fact that $h_T \le C_1 s \le r$, we find that, for any $z \in T$,

$$\begin{aligned} e_T(f) &\le \Big(K_p(\pi_z) + Ch_{\pi_k}^2 \omega\Big(\max\{|z-t|, h_T\}\Big)\Big)|T|^{1+\frac{1}{p}} + 9\mu_k^2 h_T^2 \varepsilon^{\frac{1}{2}} |T|^{\frac{1}{p}} \\ &\le \Big(K_p(\pi_z) + Ch_{\pi_k}^2 \omega(r)\Big)|T|^{1+\frac{1}{p}} + 9\mu_k^2 h_T^2 \varepsilon^{\frac{1}{2}} |T|^{\frac{1}{p}}, \end{aligned} \tag{3.117}$$

by virtue of (3.113).

Note that, by using (3.115), we also have

$$h_T \le C_1 s^\eta \Big(\delta_f^{-1} |f|_{W^2_\infty(\Omega)}^{\frac{1-q}{4}}\Big) |\lambda_{1,k}|^{-\frac{q}{2}} \le C_1 s^\eta \delta_f^{-1-\frac{q}{2}} |f|_{W^2_\infty(\Omega)}^{\frac{1-q}{4}},$$

by virtue of (3.13). Since the angle $\mu_k$ associated with $\pi_{b_k}$ is less than $2\pi$, we have $9\mu_k^2 \le 36\pi^2$. Thus, denoting the constant $C_{p,\delta_f} = 24\pi^2(\frac{3}{2})^{\frac{3}{2}}\Big(C_1^2 \delta_f^{-2-q} |f|_{W^2_\infty(\Omega)}^{\frac{1-q}{2}}\Big)$, we deduce from (3.117) that

$$\begin{aligned} e_T(f) &\le s^{2\eta}\Big(K_p(\pi_z) + Ch_{\pi_k}^2 \omega(r)\Big)^{1-q} |T|^{\frac{1}{p}} + C_{p,\delta_f} s^{2\eta} \omega(\sqrt{2}r)^{\frac{1}{2}} |T|^{\frac{1}{p}} \\ &\le s^{2\eta}\Big(\Big(K_p(\pi_z) + Ch_{\pi_k}^2 \omega(r)\Big)^{1-q} + C_{p,\delta_f} \omega(\sqrt{2}r)^{\frac{1}{2}}\Big)|T|^{\frac{1}{p}}. \end{aligned}$$

Since $|T|$ depends only on $t$, integrating over $z \in T$ yields

$$|T|e_T(f)^p = \int_T e_T(f)^p \mathrm{d}z$$
$$\leq s^{2p\eta} \int_T \left( \left( K_p(\pi_z) + Ch_{\pi_k}^2 \omega(r) \right)^{1-q} + C_{p,\delta_f} \omega(\sqrt{2}r)^{\frac{1}{2}} \right)^p |T| \mathrm{d}z.$$

The result (3.114) is obtained by simplification by $|T|$ in both sides. □

We denote by $\Omega_{\mathrm{reg}}$ the space covered by the union of all regular triangles, and by $\Delta_{s,r}^{\mathrm{reg}}$ the set of regular triangles in $\Delta_{s,r}$.

**Proposition 3.4.2.** *There are two constants $C_{p,\delta_f}$ and $C_{\delta_f}$ such that the errors on regular regions satisfy*

$$\sum_{T \in \Delta_{s,r}^{reg}} e_T(f)^p \leq s^{2p\eta} \int_\Omega \left( \left( K_p(\pi_z) + C_{\delta_f} \omega(r) \right)^{1-q} + C_{p,\delta_f} \omega(\sqrt{2}r)^{\frac{1}{2}} \right)^p dz. \quad (3.118)$$

*Proof.* For any regular triangle $T \in \Delta_{s,r}^{\mathrm{reg}}$, the conditions in Proposition 3.4.1 for $s$ and $r$ so that $C_1 s \leq r$ and $\max_{z \in T} |z - t| \leq r$ hold are satisfied by construction of $\Delta_{s,r}$. With $h_{\pi_k}^2 \leq h_0^2 \delta_f^{-\frac{1}{2}} |f|_{W_\infty^2(\Omega)}^{\frac{1}{2}}$, we easily deduce from (3.114) that

$$\sum_{T \in \Delta_{s,r}^{\mathrm{reg}}} e_T(f)^p \leq s^{2p\eta} \int_{\Omega_{\mathrm{reg}}} \left( \left( K_p(\pi_z) + C'_{p,\delta_f} \omega(r) \right)^{1-q} + C_{p,\delta_f} \omega(r)^{\frac{1}{2}} \right)^p \mathrm{d}z,$$

where $C_{\delta_f} = 2Ch_0^2 \delta_f^{-\frac{1}{2}} |f|_{W_\infty^2(\Omega)}^{\frac{1}{2}}$, thereby proving (3.118). □

### 3.4.2 $L_p$-norm on irregular regions

Denote by $\Delta_{s,r}^{\mathrm{irr}}$ the set of irregular triangles in $\Delta_{s,r}$. Given an irregular triangle $T \in \Delta_{s,r}^{\mathrm{irr}}$, we use Lemma 2.2.1 to obtain

$$e_T(f) \leq Ch_T^2 |f|_{W_p^2(T)} \leq Ch_T^2 |T|^{\frac{1}{p}} |f|_{W_\infty^2(T)},$$

where $C$ is an absolute constant. Thus,

$$\sum_{T\in\Delta_{s,r}^{\mathrm{irr}}} e_T(f)^p \le C^p |f|^p_{W^2_\infty(\Omega)} \Big(\sup_{T\in\Delta_{s,r}^{\mathrm{irr}}} h_T\Big)^{2p} \Big(\sum_{T\in\Delta_{s,r}^{\mathrm{irr}}} |T|\Big). \tag{3.119}$$

**Proposition 3.4.3.** *The errors on irregular regions satisfies*

$$\sum_{T\in\Delta_{s,r}^{irr}} e_T(f)^p \le s^{2p+1}\Big(8(C_*C_1)^{2p+1}C^p rm^2 |f|^p_{W^2_\infty(\Omega)}\Big). \tag{3.120}$$

*Proof.* From Proposition 3.3.11, together with Proposition 3.4.1, for any irregular triangle $T\in\Delta_{s,r}^{\mathrm{irr}}$, we have $h_T \le C_*C_1 s$. Hence from (3.109), the area covered by all irregular triangles satisfies

$$\sum_{T\in\Delta_{s,r}^{\mathrm{irr}}} |T| \le 8C_*C_1 rm^2 s. \tag{3.121}$$

The result in (3.120) is obtained by combining the above inequality with (3.119), together with the result in Proposition 3.3.11 and (3.113) of Proposition 3.4.1. □

### 3.4.3 Sobolev seminorm on regular regions

Given a regular triangle $T$ contained in a sub-square $R_i$, $i\in\{1,\dots,m^2\}$, by using the estimation (3.113) in Proposition 3.4.1, we have

$$\begin{aligned} |T|^{\frac{1}{2}} &\le s^\eta \Big(K_p(\pi_i) + 2Ch^2_{\pi_i}\omega(r)\Big)^{-\frac{q}{2}} \\ &\le s^\eta K_p(\pi_0)^{-\frac{q}{2}} |\det \pi_i|^{-\frac{q}{4}}, \end{aligned}$$

by virtue of (2.54), where $\pi_0(x,y) = x^2+y^2$.

With $\pi = \pi_{b_i}$, for any $z\in T$ we have $|z-b_i| \le \sqrt{2}r$ and the condition (2.88) is satisfied with $\nu = \omega(\sqrt{2}r)$. By virtue of the choice of $r$ in (3.17), the inequality $\delta_f \le |\lambda_{1,i}\lambda_{2,i}|^{\frac{1}{2}}$ clearly holds for all $i\in\{1,\dots,m^2\}$, and we have

$$\omega(r) \le \min\Big\{1, |\lambda_{1,i}\lambda_{2,i}|^{\frac{1}{2}}\Big\}. \tag{3.122}$$

Hence, by using (2.93) we find that

$$|f - I_T f|^p_{W^1_p(T)} \lesssim |\det \pi_{b_i}|^{\frac{p}{4}} |\lambda_{2,i}|^{\frac{p}{2}} |T|^{1+\frac{p}{2}} \lesssim s^{p\eta} |T| |\det \pi_{b_i}|^{\frac{q}{4}} |\lambda_{2,i}|^{\frac{p}{2}}, \tag{3.123}$$

by virtue of the fact that $p - pq = q$.

In order to estimate (3.123) independently of $b_i$, we consider the function $g(z) := |\det H_f(z)|^{\frac{q}{4}} \|H_f(z)\|_2^{\frac{p}{2}}$ and prove the following result.

**Lemma 3.4.4.** *For any $z, t \in \Omega$ such that $|z - t| \leq \sqrt{2}r$, there is a constant $\tilde{C}_p$ depending only on $p$ such that*

$$|g(z) - g(t)| \leq \tilde{C}_p \omega(\sqrt{2}r)^{\frac{q}{4}}.$$

*Proof.* Consider each of the terms in the right hand side of

$$\begin{aligned} g(z) - g(t) =& \|H_f(z)\|_2^{\frac{p}{2}} (|\det H_f(z)|^{\frac{q}{4}} - |\det H_f(t)|^{\frac{q}{4}}) \\ &+ |\det H_f(t)|^{\frac{q}{4}} (\|H_f(z)\|_2^{\frac{p}{2}} - \|H_f(t)\|_2^{\frac{p}{2}}). \end{aligned} \tag{3.124}$$

By virtue of (2.4), clearly $\|H_f(z)\|_2^{\frac{p}{2}} \leq (\frac{3}{2})^{\frac{p}{2}} \|\pi_z\|^{\frac{p}{2}} \leq (\frac{3}{2})^{\frac{p}{2}} |f|^{\frac{p}{2}}_{W^2_\infty(\Omega)}$. Also, since $|\det \pi_z| \leq \|H_f(z)\|_2^2$, we also have that $|\det H_f(t)|^{\frac{q}{4}} \leq (\frac{3}{2})^{\frac{q}{2}} |f|^{\frac{q}{2}}_{W^2_\infty(\Omega)}$. In a similar way that (3.75) is proved, and taking into account (3.45) so that $\|Q_{\pi_z} - Q_{\pi_t}\|_2 \leq \frac{3}{2}\omega(\sqrt{2}r)$, we have

$$\Big| |\det H_f(z)| - |\det H_f(t)| \Big| \leq 2(\frac{3}{2})^2 \omega(\sqrt{2}r) |f|_{W^2_\infty(\Omega)}.$$

Since $q < 1$, we have that $||\det H_f(z)|^{\frac{q}{4}} - |\det H_f(t)|^{\frac{q}{4}}| \leq 2^{\frac{q}{4}} (\frac{3}{2})^{\frac{q}{2}} \omega(r)^{\frac{q}{4}} |f|^{\frac{q}{4}}_{W^2_\infty(\Omega)}$ (see (3.75)).

Next, since $|z - t| \leq \sqrt{2}r$, we obtain from (2.4) that

$$\Big| \|H_f(z)\|_2 - \|H_f(t)\|_2 \Big| \leq \frac{3}{2} \omega(\sqrt{2}r). \tag{3.125}$$

Considering the function $\ell(x) = x^{\frac{p}{2}}$ on an interval $[a, b] \subset \mathbb{R}$, by the mean value theorem, we have $|\ell(b)-\ell(a)| \leq |b-a|\ell'(c)$ for some $c \in (a,b)$. Thus,

- If $p \geq 2$, by using the mean value theorem and (3.125), we have

$$\begin{aligned}\Big| \|H_f(z)\|_2^{\frac{p}{2}} - \|H_f(t)\|_2^{\frac{p}{2}} \Big| &\leq \frac{p}{2}\Big| \|H_f(z)\|_2 - \|H_f(t)\|_2 \Big| \|H_f(z)\|_2^{\frac{p}{2}-1} \\ &\leq \frac{p}{2}(\frac{3}{2})^{\frac{p}{2}}\omega(\sqrt{2}r)|f|_{W^2_\infty(\Omega)}^{\frac{p}{2}-1};\end{aligned}$$

- If $p \leq 2$, (3.125) implies $\Big| \|H_f(z)\|_2^{\frac{p}{2}} - \|H_f(t)\|_2^{\frac{p}{2}} \Big| \leq (\frac{3}{2})^{\frac{p}{2}}\omega(\sqrt{2}r)^{\frac{p}{2}}$.

It follows that

$$\Big| \|H_f(z)\|_2^{\frac{p}{2}} - \|H_f(t)\|_2^{\frac{p}{2}} \Big| \leq (\frac{3}{2})^{\frac{p}{2}}\Big(\frac{p}{2}|f|_{W^2_\infty(\Omega)}^{\frac{p}{2}-1}+1\Big)\omega(\sqrt{2}r)^{\frac{p}{2}}.$$

We now deduce from (3.124) that $|g(z)-g(t)| \leq \tilde{C}_p\omega(\sqrt{2}r)^{\frac{q}{4}}$, where

$$\tilde{C}_p = 2^{\frac{q}{4}}(\frac{3}{2})^{\frac{2p+q}{4}}|f|_{W^2_\infty(\Omega)}^{\frac{p+q}{2}} + \frac{p}{2}(\frac{3}{2})^{\frac{p+q}{2}}|f|_{W^2_\infty(\Omega)}^{\frac{q}{2}}\Big(|f|_{W^2_\infty(\Omega)}^{\frac{p}{2}-1}+1\Big).$$

This concludes our proof. □

**Proposition 3.4.5.** *The $W^1_p$-seminorm on regular regions satisfies*

$$\sum_{T\in\Delta^{reg}_{s,r}} |f-I_Tf|^p_{W^1_p(T)} \lesssim s^{p\eta}\int_\Omega \Big(|\det \pi_z|^{\frac{q}{4}}\|H_f(z)\|_2^{\frac{p}{2}} + \tilde{C}_p\omega(\sqrt{2}r)^{\frac{q}{4}}\Big)dz, \qquad (3.126)$$

*where $\tilde{C}_p$ is a in Lemma 3.4.4.*

*Proof.* By using Lemma 3.4.4 and (3.123), for any $z \in T$, we have

$$\begin{aligned}|f-I_Tf|^p_{W^1_p(T)} \lesssim s^{p\eta}|T|g(b_i) &\leq s^{p\eta}|T|\Big(g(z)+\tilde{C}_p\omega(\sqrt{2}r)^{\frac{q}{4}}\Big) \\ &\lesssim s^{p\eta}|T|\Big(|\det \pi_z|^{\frac{q}{4}}\|H_f(z)\|_2^{\frac{p}{2}} + \tilde{C}_p\omega(\sqrt{2}r)^{\frac{q}{4}}\Big).\end{aligned}$$

By integrating over $z \in T$ and simplifying by $|T|$, we obtain

$$|f-I_Tf|^p_{W^1_p(T)} \lesssim s^{p\eta}\int_T \Big(|\det \pi_z|^{\frac{q}{4}}\|H_f(z)\|_2^{\frac{p}{2}} + \tilde{C}_p\omega(\sqrt{2}r)^{\frac{q}{4}}\Big)\mathrm{d}z. \qquad (3.127)$$

Summing up over all regular triangles gives (3.126). □

### 3.4.4 Sobolev seminorm on irregular regions

Let $T$ be an irregular triangle and $S_i$, with $i \in \{1, \dots, m^2\}$, one of the closest sub-squares to the barycenter $b_T$ of $T$. Let $\varphi_i$ be the invertible linear map as described in Section 3.2.4. According to Corollary 3.3.10, the interior angles of the triangle $\varphi_i^{-1}(T)$ are far from $\pi$. We then apply Lemma 2.6.1 to obtain

$$|f - I_T f|_{W_p^1(T)} \lesssim \operatorname{cond}(\varphi_i)^2 h_T |f|_{W_p^2(T)} = \left|\frac{\lambda_{2,i}}{\lambda_{1,i}}\right| h_T |f|_{W_p^2(T)}. \tag{3.128}$$

**Proposition 3.4.6.** *The $W_p^1$-seminorm on irregular regions satisfies*

$$\sum_{T \in \Delta_{s,r}^{irr}} |f - I_T f|^p_{W_p^1(T)} \lesssim s^{p+\frac{1}{2}} \left( r m^2 |f|^p_{W_\infty^2(\Omega)} \right). \tag{3.129}$$

*Proof.* As in the proof of (3.120), the diameter $h_T$ of an irregular triangle $T$ is less than $\leq C_* C_1 s$, where $C_*$ is the constant in Proposition 3.3.11 and $C_1$ the constant in Proposition 3.4.1 so that (3.121) holds. Using (3.112) we have $\left|\frac{\lambda_{2,i}}{\lambda_{1,i}}\right| \leq s^{-\frac{1}{2p}}$, and thus from (3.128), we have

$$|f - I_T f|_{W_p^1(T)} \lesssim s^{1-\frac{1}{2p}} |f|_{W_p^2(T)} \leq s^{1-\frac{1}{2p}} |T|^{\frac{1}{p}} |f|_{W_\infty^2(T)}.$$

By virtue of (3.121), we find that

$$\sum_{T \in \Delta_{s,r}^{\mathrm{irr}}} |f - I_T f|^p_{W_p^1(T)} \lesssim s^{p-\frac{1}{2}} |f|^p_{W_\infty^2(\Omega)} \left( \sum_{T \in \Delta_{s,r}^{\mathrm{irr}}} |T| \right) \lesssim s^{p+\frac{1}{2}} \left( r m^2 |f|^p_{W_\infty^2(\Omega)} \right),$$

which proves the result. □

In the asymptotic estimations which we prove in Section 3.4.6, we show that the errors on irregular triangles in both $L_p$-norm and $W_p^1$-seminorm are significantly small compared to the errors on regular regions.

### 3.4.5 Number of triangles

In the result below we estimate the number of triangles in $\Delta_{s,r}$. We assume that $r$ is sufficiently small so that (3.15) holds.

**Proposition 3.4.7.** *The number of triangles in $\Delta_{s,r}$ satisfies*

$$\#\Big(\Delta_{s,r}\Big) \le s^{-2\eta}\Big(\int_{\Omega_{reg}}\Big(K_p(\pi_z) + \tilde{C}_{\delta_f}\omega(\sqrt{2}r)\Big)^q dz + \tilde{C}'_{\delta_f}s\Big), \tag{3.130}$$

*where $\tilde{C}_{\delta_f}$ and $\tilde{C}'_{\delta_f}$ are constants depending only on $\delta_f$.*

*Proof.* We shall first estimate the number of triangles inside of, or intersecting, a sub-square $S_i$. Let $P_i$ denote a regular parallelogram of $R_i$ and $T_i$ a regular triangle. Clearly, from (3.43),

$$|P_i| = 2|T_i| = \frac{2s^{2\eta}}{\Big(K_p(\pi_{b_i}) + 2Ch^2_{\pi_i}\omega(r)\Big)^q}.$$

Denote by $P_i^h$ and $P_i^\rho$ the longest and shortest side lengths of $P_i$, with

$$P_i^h = \Lambda_i\Big(\rho_0^2\Big|\frac{\lambda_{2,i}}{\lambda_{1,i}}\Big|^{\frac{1}{2}} + \frac{h_0^2}{4}\Big|\frac{\lambda_{1,i}}{\lambda_{2,i}}\Big|^{\frac{1}{2}}\Big)^{\frac{1}{2}}, \tag{3.131}$$

$$P_i^\rho = h_0\Lambda_i\Big|\frac{\lambda_{1,i}}{\lambda_{2,i}}\Big|^{\frac{1}{4}}. \tag{3.132}$$

The number of regular parallelograms in $R_i$ is denoted by $N_i(s)$. By virtue of (2.71) and (2.4), there is a constant $c_2$ such that

$$\max_{z\in\Omega} h_{\pi_z} \le c_2(\frac{3}{2})^{\frac{1}{4}}\delta_f^{-\frac{1}{4}}|f|^{\frac{1}{4}}_{W^2_\infty(\Omega)} := C_f. \tag{3.133}$$

Observe from (3.116) that $K_p(\pi_z) \le K_p(\pi_{b_i}) + 2Ch^2_{\pi_i}\omega(r) \le K_p(\pi_z) + 4CC_f^2\omega(\sqrt{2}r)$ for any $z \in S_i$, where $C_f$ is defined in (3.133). Thus

$$|T| = \Lambda_i^2 = \frac{s^{2\eta}}{\Big(K_p(\pi_{b_i}) + 2Ch^2_{\pi_i}\omega(r)\Big)^q} \ge \frac{s^{2\eta}}{\Big(K_p(\pi_z) + 4CC_f^2\omega(\sqrt{2}r)\Big)^q}.$$

The number of regular parallelograms in $R_i$ is then bounded by

$$N_i(s) \leq \frac{|S_i|}{|2T_i|} \leq \frac{1}{2} s^{-2\eta} |S_i| \Big( K_p(\pi_z) + 4CC_f^2 \omega(\sqrt{2}r) \Big)^q,$$

for any $z \in S_i$. By integrating over $z \in S_i$, the number of regular triangles $2N_i(s)$ in $R_i$ satisfies

$$2N_i(s) \leq s^{-2\eta} \int_{S_i} \Big( K_p(\pi_z) + 4CC_f^2 \omega(\sqrt{2}r) \Big)^q \mathrm{d}z.$$

The total number of all regular triangles then satisfies

$$\#\Big(\Delta_{s,r}^{\mathrm{reg}}\Big) \leq s^{-2\eta} \int_{\Omega} \Big( K_p(\pi_z) + 4CC_f^2 \omega(\sqrt{2}r) \Big)^q \mathrm{d}z. \tag{3.134}$$

We shall now estimate the number of irregular polygons. Let $s$ be sufficiently small and fixed. Using the result in Proposition 3.3.11 and its proof, the length of an extended segment from Setting 4 is less than $C_* h_M$, with $h_M = \sup_{T \in \Delta_s^{\mathrm{reg}}} h_T$. Now, from Proposition 3.4.1, we have $h_M \leq C_1 s$. Hence, the maximum length of an extended segment $\ell^{\mathrm{ext}}$ from Setting 4 is bounded by $C_1 C_* s$.

Our next step is to extend the regular region $R_i$ to a bigger one that can cover $S_i$, instead of the procedure of segment extensions in Section 3.2.3. We define a stripe of regular parallelograms as a collection of parallelograms in $R_i$ which are glued by their shortest or longest edges. We then extend the stripe up to the side(s) of $S_i$ by gluing more additional parallelograms. The number of the additional parallelograms in the extended stripe satisfies

$$N_i'(s) = \left\lceil \frac{C_1 C_* s}{P_i^\rho} \right\rceil. \tag{3.135}$$

where $P_i^\rho$ is the length of the shortest edge of $P_i$, as given in (3.132).

Observe that the obtained extended regular region does not necessarily cover $S_i$: For instance, if the stripes parallel to $P_i^h$ or $P_i^\rho$ are parallel to the sides of $S_i$, then some spaces in the neighborhood of the corners of $S_i$ may not be covered.

Hence, we shall construct a new kind of stripes, which we call *additional* stripes which will cover the quarter-disks defined in Section 3.3.4. An *additional* stripe is defined by two lines where both are parallel to either $P_i^h$ or $P_i^\rho$, and from which one can obtain a stripe of parallelograms (from the translates of $P_i$) by gluing the parallelograms by their shortest or longest edges. Note that we allow the extended stripes and the additional ones to overlap.

Recall that each of the quarter-disks defined in Section 3.3.4 contains at least one regular parallelogram. Recall also that the radius of each quarter disk is $4\sqrt{2}h_M$, and its diameter $8h_M$ is less than the maximum length of an extended segment from Setting 4. Hence, the number of additional stripes in any direction of $\mathbf{e} \in \{\pm\mathbf{e}_i, \pm\bar{\mathbf{e}}_i\}$ is less than $N_i'(s)$. Observe also that $N_i'(s)$ is the minimum number of parallelograms that can be inserted inside each additional stripes so that they cover the quarter-disk corresponding to them, where each parallelograms in these additional stripes have a non-empty intersection with the quarter-disks.

Let $N_i^{(1)}$ be defined by

$$N_i^{(1)} = \lceil \frac{\sqrt{2}r}{P_i^\rho} \rceil. \tag{3.136}$$

Let also $N_i^{(2)}$ be defined by

$$N_i^{(2)} = \lceil \frac{\sqrt{2}r}{P_i^h} \rceil. \tag{3.137}$$

Clearly, $N_i^{(2)} \leq N_i^{(1)}$, and the maximum number of stripes that cover $S_i$ is less than $N_i^{(1)}$.

For each stripe, there are two directions of extensions to the sides of $S_i$, and there are two kinds of stripes (parallel to $P_i^h$ or $P_i^\rho$). The number of stripes in $R_i$ is less than $2N_i^{(1)}$. Hence, the total number of parallelograms from the extensions of stripes and from the additional stripes, is bounded by

$$\begin{aligned} N_i^{\mathrm{irr}}(s) &= (2N_i^{(1)})(2N_i'(s)) + (4N_i'(s))(N_i'(s)) \\ &\leq 8N_i^{(1)}N_i'(s), \end{aligned} \tag{3.138}$$

by virtue of $N_i'(s) \leq N_i^{(1)}$.

Observe now that, counting a pentagon or a hexagon obtained by the segment extension procedure of Section 3.2.3 can be replaced by counting twice the parallelogram associated with it: A pentagon (resp. a hexagon) can be associated with a translated version of a regular parallelogram which is cut by an edge (resp. two edges) of the sub-square $S_i$. Now, a pentagon or a hexagon is counted as one polygon, but it will be divided into three or four triangles during the final triangulation; whereas if we count twice the parallelogram associated with it, then we would have four triangles since a parallelogram will be divided into two triangles.

We can associate a parallelogram of the extended and additional stripes to each quadrilateral obtained by the segment extension procedure of Section 3.2.3, since for sufficiently small $r, s$ the number of quadrilaterals having a non-empty intersection with $S_i$ is less than the number of these parallelograms. Hence, the number of irregular triangles inside of, or intersecting $S_i$, is less than four times $N_i^{\mathrm{irr}}(s)$, that is,

$$4N_i^{\mathrm{irr}}(s) \leq 32N_i^{(1)}N_i'(s). \tag{3.139}$$

Let $K_f = \max_{z\in\Omega} K_p(\pi_z) + 4CC_f^2\omega(\sqrt{2}r)$. Then, from (3.132), there is an absolute constant $C'$ such that

$$\frac{1}{P_i^\rho} \leq s^{-\eta}C'\frac{K_f^{\frac{q}{2}}}{h_0}(\frac{3}{2})|f|^{\frac{1}{4}}\delta_f^{-\frac{1}{4}}.$$

Combining now (3.136) with (3.135), there exists a constant $K_{\delta_f}$ such that

$$4N_i^{\mathrm{irr}}(s) \leq K_{\delta_f}s^{1-2\eta}. \tag{3.140}$$

Summing up over all sub-squares, the total number of irregular triangles is less than $s^{1-2\eta}m^2K_{\delta_f}$. Hence the result with $\tilde{C}_{\delta_f} = 4CC_f^2$ and $\tilde{C}'_{\delta_f} = m^2K_{\delta_f}$. $\square$

### 3.4.6 Asymptotic error estimations

We combine the results obtained in the previous sections. For the asymptotic $L_p$-norm, we use (3.118), (3.120) and (3.130) to obtain

$$\begin{aligned}\big(\#\Delta_{s,r}\big)\|f - I_{\Delta_{s,r}}f\|_{L_p(\Omega)} \leq \Big(\int_\Omega \Big(K_p(\pi_z) + \tilde{C}_{\delta_f}\omega(\sqrt{2}r)\Big)^q \mathrm{d}z + \tilde{C}'_{\delta_f}s\Big) \\ \Big(\int_\Omega \Big(\Big(K_p(\pi_z) + C_{\delta_f}\omega(r)\Big)^{1-q} + C_{p,\delta_f}\omega(\sqrt{2}r)^{\frac{1}{2}}\Big)^p \mathrm{d}z + C's^{2p+1-2p\eta}\Big)^{\frac{1}{p}},\end{aligned} \tag{3.141}$$

where $C' = 8(C_*C_1)^{2p+1}C^prm^2|f|^p_{W^2_\infty(\Omega)}$. Note that $2p+1-2p\eta > 0$ since $\eta$ is chosen so that $1 + \frac{1-q}{8p} < \eta < 1 + \frac{1}{2p}$. In a similar way, for the $W^1_p$-seminorm of the error, we combine (3.126), (3.129) and (3.130) to obtain

$$\begin{aligned}\big(\#\Delta_{s,r}\big)^{\frac{1}{2}}|f - I_{\Delta_{s,r}}f|_{W^1_p(\Omega)} \lesssim \Big(\int_\Omega \Big(K_p(\pi_z) + \tilde{C}_{\delta_f}\omega(\sqrt{2}r)\Big)^q \mathrm{d}z + \tilde{C}'_{\delta_f}s\Big)^{\frac{1}{2}} \\ \Big(\int_\Omega \Big(|\det \pi_z|^{\frac{q}{4}}\|H_f(z)\|_2^{\frac{p}{2}} + \tilde{C}_p\omega(\sqrt{2}r)^{\frac{q}{4}}\Big)\mathrm{d}z + C_1's^{p+\frac{1}{2}-p\eta}\Big)^{\frac{1}{p}},\end{aligned} \tag{3.142}$$

where $C_1' = rm^2|f|^p_{W^2_\infty(\Omega)}$. Note that $p + \frac{1}{2} - p\eta > 0$.

Observing that $p(1-q) = q$ and taking the limit as $r \to 0$, we obtain

$$\begin{aligned}\lim_{r\to 0}\int_\Omega \Big(K_p(\pi_z) + \tilde{C}_{\delta_f}\omega(\sqrt{2}r)\Big)^q \mathrm{d}z &= \int_\Omega K_p(\pi_z)^q \mathrm{d}z, \\ \lim_{r\to 0}\int_\Omega \Big(\Big(K_p(\pi_z) + C_{\delta_f}\omega(r)\Big)^{1-q} + C_{p,\delta_f}\omega(\sqrt{2}r)^{\frac{1}{2}}\Big)^p \mathrm{d}z &= \int_\Omega K_p(\pi_z)^q \mathrm{d}z, \\ \lim_{r\to 0}\int_\Omega \Big(|\det \pi_z|^{\frac{q}{4}}\|H_f(z)\|_2^{\frac{p}{2}} + \tilde{C}_p\omega(\sqrt{2}r)^{\frac{q}{4}}\Big)\mathrm{d}z &= \int_\Omega |\det \pi_z|^{\frac{q}{4}}\|H_f(z)\|_2^{\frac{p}{2}}\mathrm{d}z.\end{aligned}$$

Thus, by virtue of (3.141) and (3.142), given a number $\varepsilon > 0$ and a sufficiently small $r_0 > 0$, for the triangulation $\Delta_s := \Delta_{s,r_0}$, we have

$$\limsup_{s\to 0}(\#\Delta_s)\|f - I_{\Delta_s}f\|_{L_p(\Omega)} \leq \Big(\int_\Omega \Big(K_p(\pi_z) + \varepsilon\Big)^q \mathrm{d}z\Big)^{\frac{1}{q}}, \tag{3.143}$$

$$\limsup_{s\to 0}(\#\Delta_s)^{\frac{1}{2}}|f - I_{\Delta_s}f|_{W_p^1(\Omega)} \lesssim \Big(\int_\Omega \Big(K_p(\pi_z) + \varepsilon\Big)^q \mathrm{d}z\Big)^{\frac{1}{2}} \Big(\int_\Omega |\det \pi_z|^{\frac{q}{4}}\|H_f(z)\|_2^{\frac{p}{2}}\mathrm{d}z + \varepsilon\Big)^{\frac{1}{p}}. \tag{3.144}$$

**Theorem 3.4.8.** *Let $f \in C^2(\Omega)$ be convex, and $1 \leq p < \infty$. There exists a sequence of triangulations $(\Delta_N)_{N\geq N_0}$, with $(\#\Delta_N) \leq N$, where the asymptotic estimations*

$$\limsup_{N\to\infty} N\|f - f_N\|_{L_p(\Omega)} \leq \Big(\int_\Omega \Big(K_p(\pi_z)\Big)^q dz\Big)^{\frac{1}{q}}, \tag{3.145}$$

$$\limsup_{N\to\infty} N^{\frac{1}{2}}|f - f_N|_{W_p^1(\Omega)} \lesssim \Big(\int_\Omega \Big(K_p(\pi_z)\Big)^q dz\Big)^{\frac{1}{2}} \Big(\int_\Omega |\det \pi_z|^{\frac{q}{4}}\|H_f(z)\|_2^{\frac{p}{2}}\, dz\Big)^{\frac{1}{p}}, \tag{3.146}$$

*hold, with $f_N := I_{\Delta_N}f$ being the approximant of $f$ on $\Omega$.*

*Proof.* To prove the above result, we shall first show that the number of regular triangles dominates the total number of triangles in the triangulation $\Delta_s$. Then, it will be sufficient to study the number of triangles in $\Delta_{s_1}^{\mathrm{reg}}$ and in $\Delta_{s_0}^{\mathrm{reg}}$, as shown in (3.153), resulting from a small perturbation $\varepsilon_0 = s_0^\eta - s_1^\eta > 0$.

As already shown in Proposition 3.4.7, for $s = s_0$, the number of irregular triangles is $o(s_0^{-2\eta})$ as $s_0 \to 0$. For $s_0$ small enough, we claim that $(\#\Delta_{s_0}^{\mathrm{reg}})$ dominates the number of irregular triangles. This is achieved if we derive a lower bound of order $s_0^{-2\eta}$ for the number of regular triangles $(\#\Delta_{s_0}^{\mathrm{reg}})$.

Recall that the number of regular parallelograms in $R_i$ is denoted by $N_i(s_0)$. Given a parallelogram $P_i$ of $R_i$, from Proposition 3.3.12 and Proposition 3.4.1, the area $N_i(s_0)|P_i|$ covered by regular triangles in $S_i$ satisfies

$$N_i(s_0)|P_i| \geq (r - 8\sqrt{2}h_M)^2 \geq (r - 8\sqrt{2}C_1 s_0)^2, \tag{3.147}$$

where $h_M := \sup_{T\in\Delta_{s,r}^{\text{reg}}} h_T$. Recall the fact that $K_p(\pi_z) \leq K_p(\pi_{b_i}) + 2Ch_{\pi_i}^2\omega(r) \leq K_p(\pi_z) + 4CC_f^2\omega(\sqrt{2}r)$, for any $z \in S_i$, where $C_f$ is defined in (3.133). After integrating over $z \in S_i$,

$$\frac{1}{2s_0^{2\eta}}\int_{S_i} K_p(\pi_z)^q \mathrm{d}z \leq \frac{r^2}{|P_i|} \leq \frac{1}{2s_0^{2\eta}}\int_{S_i}\Big(K_p(\pi_z) + 4CC_f^2\omega(r)\Big)^q \mathrm{d}z.$$

By using only the first inequality, we deduce from (3.147) that, for $s_0 \to 0$,

$$\begin{aligned} N_i(s_0) &\geq \frac{r^2}{|P_i|}\frac{(r-8\sqrt{2}C_1s_0)^2}{r^2} \geq \frac{(r-8\sqrt{2}C_1s_0)^2}{2r^2s_0^{2\eta}}\int_{S_i} K_p(\pi_z)^q \mathrm{d}z \\ &\geq \frac{1}{2}s_0^{-2\eta}\int_{S_i} K_p(\pi_z)^q \mathrm{d}z - o(s_0^{-2\eta}). \end{aligned} \tag{3.148}$$

Recall that the number of regular triangles is twice of $N_i(s_0)$. Now, summing up over all sub-squares implies that, for $s_0 \to 0$,

$$(\#\Delta_{s_0}^{\text{reg}}) \geq s_0^{-2\eta}\int_\Omega K_p(\pi_z)^q \mathrm{d}z - o(s_0^{-2\eta}), \tag{3.149}$$

thereby proving our claim that regular triangles dominate $(\#\Delta_{s_0})$.

Next, we shall study the number of triangles due to a small perturbation of $s_0^\eta$ (see (3.153) below). Let $A_0 = |R_i(s_0)|, A_1 = |R_i(s_1)|$ be the respective areas of the regular regions in $R_i(s_0)$ and $R_i(s_1)$. First, we claim that

$$A_1 \geq \frac{s_1^\eta}{s_0^\eta}A_0. \tag{3.150}$$

To prove this, we use the following observation: Suppose that a scaling of $\mathcal{P}_i$ by $s_0^\eta$ leads to the current partition of $S_i$. A scaling of $\mathcal{P}_i$ by $s_1^\eta = s_0^\eta\Big(\frac{s_1^\eta}{s_0^\eta}\Big)$ is equivalent to scaling $\mathcal{P}_i$ by $s_0^\eta$, then scaling it again by $t = \frac{s_1^\eta}{s_0^\eta}$. This is equivalent to scaling $S_i$ by $\frac{1}{t} = \frac{s_0^\eta}{s_1^\eta}$. Clearly, the area $A_1'$ of the regular region in $\frac{1}{t}S_i$ is greater than the area $A_0$ of the regular region in $S_i$. The fact that $tA_1' = A_1$ proves our claim in (3.150).

Denoting by $a_0 = |P_i(s_0)|, a_1 = |P_i(s_1)|$ the areas of regular parallelograms in $R_i(s_0), R_i(s_1)$, respectively. Note that $\frac{a_0}{a_1} = \frac{s_0^{2\eta}}{s_1^{2\eta}}$. Recalling that $N_i(s_0)$ denote the exact number of regular parallelograms of $R_i$, we have

$$N_0 := N_i(s_0) = \frac{A_0}{a_0} \quad \text{and} \quad N_1 := N_i(s_1) = \frac{A_1}{a_1}. \tag{3.151}$$

We have

$$N_1 = N_0 + N_0(bc + c - 1), \tag{3.152}$$

where

$$b = \frac{A_1 - \frac{s_1^\eta}{s_0^\eta} A_0}{\frac{s_1^\eta}{s_0^\eta} A_0} \quad \text{and} \quad c = \frac{s_0^\eta}{s_1^\eta}.$$

Denoting $A = |S_i|$ the area of $S_i$, observe that

$$b = \frac{\left(\frac{s_1^\eta}{s_0^\eta} A - \frac{s_1^\eta}{s_0^\eta} A_0\right) - \left(\frac{s_1^\eta}{s_0^\eta} A - A_1\right)}{\frac{s_1^\eta}{s_0^\eta} A_0} \leq \frac{A - A_0}{A_0}.$$

Note that $A - A_0$ is the area covered by irregular triangles in $S_i$, and it is negligible compared to the area $A_0$ of regular regions, so that $b \to 0$ as $s_0 \to 0$.

We also have, by setting $\varepsilon_0 = \varepsilon s_0^{\eta+\alpha}$, where $\alpha > 0$ and $\varepsilon \geq 0$,

$$c - 1 = \frac{\varepsilon_0}{s_0^\eta - \varepsilon_0} = \frac{\varepsilon s_0^\alpha}{1 - \varepsilon s_0^\alpha}.$$

Clearly, $c - 1 \to 0$ as $\varepsilon \to 0$ or $s_0 \to 0$.

It follows that $bc + c - 1 \to 0$ as $\varepsilon_0 \to 0$ and $s_0 \to 0$, and we deduce from (3.152) that $N_1 = N_0 + o(N_0)$ which, by considering all regular regions, yields

$$(\#\Delta_{s_1}^{\text{reg}}) \leq (\#\Delta_{s_0}^{\text{reg}}) + o\Big((\#\Delta_{s_0}^{\text{reg}})\Big), \tag{3.153}$$

as $\varepsilon_0 \to 0$ and $s_0 \to 0$.

Combining now the number of all triangles, we know that the total number of regular triangle dominates the number of irregular ones. As already mentioned before, it is therefore sufficient to study the jump (3.153) from $s_0^\eta$ to $s_1^\eta = s_0^\eta - \varepsilon_0$. This leads to the estimation below, for $\varepsilon_0 \to 0$,

$$(\#\Delta_{s_1}) \leq (\#\Delta_{s_0}) + o\Big((\#\Delta_{s_0})\Big). \tag{3.154}$$

We now proceed to define the triangulation $\Delta_N$ for a given $N \geq N_0$. For any $N \geq N_0$, consider $s_N$ defined by

$$s_N := \min\{s^\eta > 0 : \big(\#\Delta_s\big) \leq N\}. \tag{3.155}$$

Let $N$ be large enough so that $s_N$ is sufficiently small. For any $s^\eta = s_N^\eta - \varepsilon < s_N^\eta$, $\varepsilon > 0$, we have $N < (\#\Delta_s)$. Therefore, for $(\#\Delta_{s_N}) \to \infty$,

$$N < (\#\Delta_s) \leq (\#\Delta_{s_N}) + o\Big((\#\Delta_{s_N})\Big),$$

which implies that $N \to \infty$ is equivalent to $(\#\Delta_{s_N}) \to \infty$. It follows that

$$N - (\#\Delta_{s_N}) = o(N) \quad \text{and} \quad N^{\frac{1}{2}} - (\#\Delta_{s_N})^{\frac{1}{2}} = o(N^{\frac{1}{2}}), \tag{3.156}$$

as $N \to \infty$.

We use the extraction argument as in [29]: Since (3.143) and (3.144) hold, for any $n \geq 1$ there exists a sub-sequence of triangulations $\Big(\Delta_{s_N}^n\Big)_{N \geq N_0}$ such that

$$\begin{aligned}
\limsup_{N\to\infty} N\|f - I_{\Delta_{s_N}^n} f\|_{L_p(\Omega)} \leq& \Big(\int_\Omega \Big(K_p(\pi_z) + \frac{1}{n}\Big)^q \mathrm{d}z\Big)^{\frac{1}{q}}, \\
\limsup_{N\to\infty} N^{\frac{1}{2}}|f - I_{\Delta_{s_N}^n} f|_{W_p^1(\Omega)} \lesssim& \Big(\int_\Omega \Big(K_p(\pi_z) + \frac{1}{n}\Big)^q \mathrm{d}z\Big)^{\frac{1}{2}} \\
& \Big(\int_\Omega |\det \pi_z|^{\frac{q}{4}} \|H_f(z)\|_2^{\frac{p}{2}} \mathrm{d}z + \frac{1}{n}\Big)^{\frac{1}{p}}.
\end{aligned}$$

Define the triangulation $\Delta_N = \Delta_{s_N}^{n(N)}$ where

$$
\begin{aligned}
n(N) := \max\Big\{ n \leq N : N\|f - I_{\Delta_{s_N}^n} f\|_{L_p(\Omega)} \leq \Big(\int_\Omega \Big(K_p(\pi_z) + \frac{2}{n}\Big)^q \mathrm{d}z\Big)^{\frac{1}{q}}, \\
N^{\frac{1}{2}}|f - I_{\Delta_{s_N}^n} f|_{W_p^1(\Omega)} \lesssim \Big(\int_\Omega \Big(K_p(\pi_z) + \frac{2}{n}\Big)^q \mathrm{d}z\Big)^{\frac{1}{2}} \\
\Big(\int_\Omega |\det \pi_z|^{\frac{q}{4}} \|H_f(z)\|_2^{\frac{p}{2}} \mathrm{d}z + \frac{2}{n}\Big)^{\frac{1}{p}} \Big\}.
\end{aligned}
\tag{3.157}
$$

For $N$ large enough, the set is non-empty and finite, making $n(N)$ well-defined. Also, clearly $n(N)$ increases with $N$ so that $n(N) \to \infty$ as $N \to \infty$. Combining this with (3.156) and the two above inequalities yields the results. □

Observe that, by using the Cauchy-Schwarz inequality for integrals,

$$
\begin{aligned}
\int_\Omega |\det \pi_z|^{\frac{q}{4}} \|H_f(z)\|_2^{\frac{p}{2}} \mathrm{d}z &= K_p(\pi_0)^{-\frac{q}{2}} \int_\Omega K_p(\pi_z)^{\frac{q}{2}} \|H_f(z)\|_2^{\frac{p}{2}} \mathrm{d}z \\
&\leq K_p(\pi_0)^{-\frac{q}{2}} \Big(\int_\Omega K_p(\pi_z)^q \mathrm{d}z\Big)^{\frac{1}{2}} \Big(\int_\Omega \|H_f(z)\|_2^p \mathrm{d}z\Big)^{\frac{1}{2}} \\
&\leq (\frac{3}{2})^{\frac{p}{2}} K_p(\pi_0)^{-\frac{q}{2}} \Big(\int_\Omega K_p(\pi_z)^q \mathrm{d}z\Big)^{\frac{1}{2}} \Big(|f|^p_{W_p^2(\Omega)}\Big)^{\frac{1}{2}},
\end{aligned}
$$

since, by virtue of (2.4), $\|H_f(z)\|_2 \leq \frac{3}{2} \max\{D_{xx}^2 f(z), D_{xy}^2 f(z), D_{yy}^2 f(z)\}$. Then, we can also use the following asymptotic estimation which is only a bit coarser than (3.146),

$$
\begin{aligned}
\limsup_{N\to\infty} N^{\frac{1}{2}} |f - f_N|_{W_p^1(\Omega)} \lesssim & (\frac{3}{2})^{\frac{1}{2}} K_p(\pi_0)^{-\frac{q}{2p}} |f|^{\frac{1}{2}}_{W_p^2(\Omega)} \Big(\int_\Omega \Big(K_p(\pi_z)\Big)^q \mathrm{d}z\Big)^{\frac{1}{2}+\frac{1}{2p}} \\
\leq & C_p \Big[ |f|_{W_p^2(\Omega)} \Big(\int_\Omega \Big(K_p(\pi_z)\Big)^q \mathrm{d}z\Big)^{\frac{1}{q}} \Big]^{\frac{1}{2}},
\end{aligned}
\tag{3.158}
$$

since $\frac{1}{2} + \frac{1}{2p} = \frac{1}{2q}$, with $C_p = C(\frac{3}{2})^{\frac{1}{2}} K_p(\pi_0)^{-\frac{1}{2(p+1)}}$ where $C$ is an absolute constant.

As we already discussed in the introduction of this chapter (see also Section 1.2), the estimation in (3.145) is optimal in the sense that it cannot be further improved on certain triangulation, and that (3.158) is the first $W_p^1$-seminorm estimation derived from a triangulation which is optimal for the $L_p$-norm estimation.

## 3.5 A numerical illustration

In this section, we present a simple comparison of our method to the uniform method. The function $f$ considered in (3.160) is convex on $\Omega = (-1,1)^2$, its anisotropic behavior is shown in Figure 3.31. In order to analyze the sharpness of our error bounds and the quality of our mesh, the function $f$ is designed so that its Hessian matrix is near degenerate, with however well-separated eigenvalues.

An "anisotropic behavior" of a function is characterized by an "abrupt" variation of its graph at some points, some curves or some surfaces. In most cases, it occurs when there are two directions such that, in one direction the change of the derivative is maximal, whereas in the other one the derivative is minimal. In order to catch this anisotropic behavior, it is cheaper to use anisotropic triangles rather than numerous small triangles.

Note that in general, knowing the expressions of the Hessian whose determinant is positive does not necessarily lead to the expression of a convex function. However, by using any invertible linear map $\phi$, one can produce another convex function $f_0 \circ \phi$ from a given one $f_0$. Indeed, supposing that $\phi(x,y) = (\alpha x + \beta y, \gamma x + \delta y)$, we have

$$\begin{aligned}
D_x f_0(\alpha x + \beta y, \gamma x + \delta y) &= \alpha D_x f_0(\alpha x + \beta y, \gamma x + \delta y) + \gamma D_y f_0(\alpha x + \beta y, \gamma x + \delta y)\\
D_y f_0(\alpha x + \beta y, \gamma x + \delta y) &= \beta D_x f_0(\alpha x + \beta y, \gamma x + \delta y) + \delta D_y f_0(\alpha x + \beta y, \gamma x + \delta y)
\end{aligned}$$

from which we deduce that

$$\begin{aligned}
D_{xx} f_0(\alpha x + \beta y, \gamma x + \delta y) &= \alpha^2 D_{xx} f_0(\alpha x + \beta y, \gamma x + \delta y) + 2\alpha\gamma D_{xy} f_0(\alpha x + \beta y, \gamma x + \delta y)\\
&\quad + \gamma^2 D_{yy} f_0(\alpha x + \beta y, \gamma x + \delta y)\\
D_{yy} f_0(\alpha x + \beta y, \gamma x + \delta y) &= \beta^2 D_{xx} f_0(\alpha x + \beta y, \gamma x + \delta y) + 2\beta\delta D_{xy} f_0(\alpha x + \beta y, \gamma x + \delta y)\\
&\quad + \delta^2 D_{yy} f_0(\alpha x + \beta y, \gamma x + \delta y)\\
D_{xy} f_0(\alpha x + \beta y, \gamma x + \delta y) &= \alpha\beta D_{xx} f_0(\alpha x + \beta y, \gamma x + \delta y) + \gamma\delta D_{yy} f_0(\alpha x + \beta y, \gamma x + \delta y)\\
&\quad + (\alpha\delta + \gamma\beta) D_{xy} f_0(\alpha x + \beta y, \gamma x + \delta y),
\end{aligned}$$

or equivalently, the Hessian

$$H_{f_0\circ\phi} = \begin{bmatrix} \alpha & \gamma \\ \beta & \delta \end{bmatrix} \begin{bmatrix} (D_{xx}f_0)\circ\phi & (D_{xy}f_0)\circ\phi \\ (D_{xy}f_0)\circ\phi & (D_{yy}f_0)\circ\phi \end{bmatrix} \begin{bmatrix} \alpha & \beta \\ \gamma & \delta \end{bmatrix}. \tag{3.159}$$

This means that the sign of the determinant of $H_{f_0\circ\phi}$ is the same as the sign of the determinant $H_{f_0}$.

Convex functions with simple expressions can be easily found. Below we construct a convex function which possesses an anisotropic behavior. Let $f_0$ be the function defined by

$$f_0(x,y) = \frac{x^2}{400} - \frac{x^3y}{4800} - \frac{xy}{200} + \frac{y^2}{4}.$$

By taking the second derivatives of $f_0$, we can easily deduce that $\det H_{f_0}(x,y) = \frac{1}{2}\left(\frac{1}{200}-\frac{xy}{800}\right)-\left(\frac{x^2}{1600}+\frac{1}{200}\right)^2 > 0$, and thus $f_0$ is always convex on $\Omega$. By considering the linear map $\phi(x,y) = (x+y, x-y)$, we obtain the function $f = f_0\circ\phi$

$$f(x,y) = \frac{(x+y)^2}{400} - \frac{(x+y)^3(x-y)}{4800} - \frac{(x+y)(x-y)}{200} + \frac{(x-y)^2}{4}. \tag{3.160}$$

We have that

$$\begin{aligned} D_xf(x,y) =& \frac{x+y}{200} - \frac{(x+y)^3}{4800} - \frac{(x+y)^2(x-y)}{1600} - \frac{x}{100} + \frac{x-y}{2}, \\ D_yf(x,y) =& \frac{x+y}{200} + \frac{(x+y)^3}{4800} - \frac{(x+y)^2(x-y)}{1600} + \frac{y}{100} - \frac{x-y}{2}, \end{aligned}$$

thus the Hessian matrix of $f$ is given by

$$H_f(x,y) = \begin{bmatrix} \frac{1}{2} - \frac{1}{200} - \frac{x^2+xy}{400} & -\frac{1}{2} + \frac{1}{200} - \frac{x^2-y^2}{800} \\ -\frac{1}{2} + \frac{1}{200} - \frac{x^2-y^2}{800} & \frac{1}{2} + \frac{3}{200} + \frac{xy+y^2}{400} \end{bmatrix}.$$

It is clear that $H_f$ is convex on $\Omega$ by virtue of (3.159). By studying each of the entries of $H_f$, we can easily show that for any $r > 0$, $\omega(r) \leq \frac{1}{200}$ where $\omega$ is defined in (2.75).

First let $\Omega$ be divided into four sub-squares $S_1, \ldots, S_4$ of side length one and barycenters

$$b_1 = \Big(-\frac{1}{2}, \frac{1}{2}\Big), \quad b_2 = \Big(\frac{1}{2}, \frac{1}{2}\Big), \quad b_3 = \Big(-\frac{1}{2}, -\frac{1}{2}\Big) \quad \text{and} \quad b_4 = \Big(\frac{1}{2}, -\frac{1}{2}\Big).$$

The anisotropic behavior of $f$ at one of the barycenters is shown as follows: We have that $D_x f(b_1) = -\frac{99}{200}$ and $D_y f(b_1) = \frac{101}{200}$. We show that at $b_1$, the direction in which the derivative is maximal (resp. minimal) is the eigenvector of $H_f(b_1)$ corresponding to the largest (resp. smallest) eigenvalue. Since

$$H_f(b_1) = \begin{bmatrix} \frac{99}{200} & -\frac{99}{200} \\ -\frac{99}{200} & \frac{103}{200} \end{bmatrix},$$

its eigenvalues are 0.0099 and 1 with corresponding eigenvectors $\mathbf{v}_1 = [0.714 \;\; 0.7]^t$ and $\mathbf{v}_2 = [-0.7 \;\; 0.714]^t$. The derivatives in $\mathbf{v}_1$ and $\mathbf{v}_2$ directions satisfy

$$D_{\mathbf{v}_1} f(b_1) = 0.714 \times \frac{-99}{200} + 0.7 \times \frac{101}{200} = 7e - 05, \tag{3.161}$$

$$D_{\mathbf{v}_2} f(b_1) = -0.7 \times \frac{-99}{200} + 0.714 \times \frac{101}{200} = 0.707. \tag{3.162}$$

In the table below, we denote by $N$ the number of triangles which we obtain by triangulating $\Omega = (-1,1)^2$ by using the method described in Section 3.2.2 and Section 3.2.3, with tolerance $\frac{d_j L}{3h_i}$. To simplify the implementation, instead of using the back transformation procedure discussed in Section 3.2.4, we use the constraint Delaunay algorithm in order to triangulate the domain $\Omega$ after it is partitioned into polygons of at most six edges. The resulting triangulation, shown in Figure 3.31, does not significantly differ from the result expected by our algorithm. To show that regular triangles produce small errors compared to irregular triangles, triangles are colored according to their relative errors, i.e. $\dfrac{\text{error}}{\text{area}}$, where gray colors indicate high errors.

The overall error is denoted by $E = \|f - f_N\|_{L_2(\Omega)}$ and is plotted in red in Figure 3.33. In order to evaluate integrals on a given a triangle $T$, we use a Gaussian quadrature formula. Fix a right triangle $T_0$ with vertices $(0,1), (1,0), (0,0)$. Given a function $g$, using a Gaussian quadrature method from [34] shows that

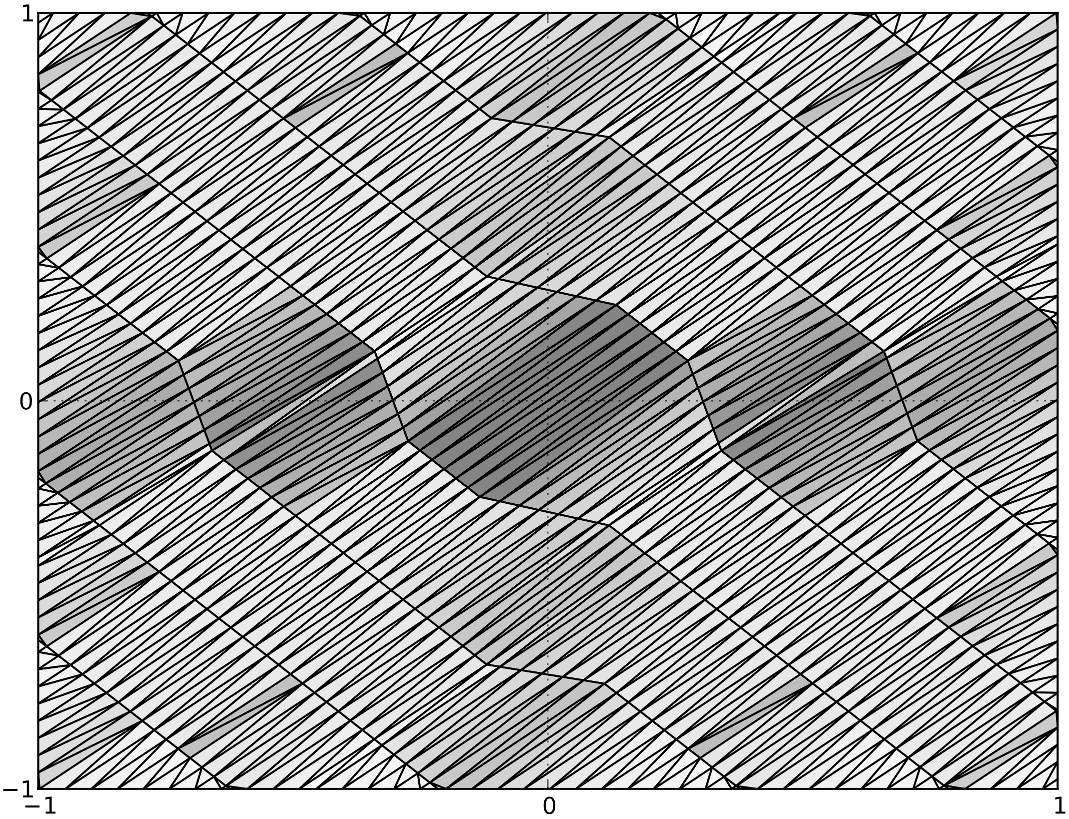


Figure 3.31: An anisotropic triangulation constructed by using our method, triangles are colored according to the relative error $\frac{\text{error}}{\text{area}}$, gray colored triangles indicate high errors.

the integral of $g$ on $T_0$ can be approximated as follows,

$$\int_{T_0} g(x)\,\mathrm{d}x \approx \frac{25}{96} g\Big(\frac{1}{5},\frac{1}{5}\Big) - \frac{9}{32} g\Big(\frac{1}{3},\frac{1}{3}\Big) + g\Big(\frac{1}{5},\frac{3}{5}\Big) + g\Big(\frac{3}{5},\frac{1}{5}\Big), \tag{3.163}$$

and equality occurs for polynomials of degree $\leq 3$.

Suppose now that $T$ is a non-degenerate triangle with vertices given by $p_1 = (x_1, y_1), p_2 = (x_2, y_2), p_3 = (x_3, y_3)$. Then, denoting by $a = x_2 - x_3$, $b = x_1 - x_3$, $c = y_2 - y_3$ and $d = y_1 - y_3$, we clearly have

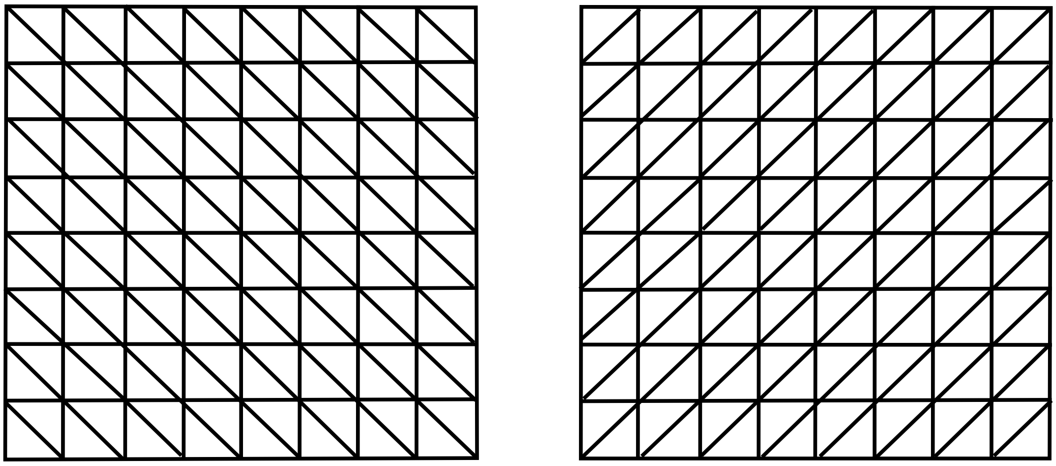
$$T = \phi(T_0), \quad \phi(x, y) := M[x\, y]^t + \mathbf{t}, \tag{3.164}$$

where $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $\mathbf{t} = [x_3\, y_3]^t$. Thus by a simple change of variables,

$$\|f\|^2_{L_2(T)} = \int_{\phi(T_0)} |f(x)|^2 \,\mathrm{d}x = \int_{T_0} |f \circ \phi(x)|^2 |M| \,\mathrm{d}x, \tag{3.165}$$

and we can apply (3.163) with $g = (f \circ \phi)^2$.

The regular triangles cover $\%A_{\text{reg}}$ of the area of $\Omega$, and they contribute $\%E^2_{\text{reg}}$ to the overall squared error $E^2$. Figure 3.33 shows that our slightly modified method can provide sharper estimations compared to uniform methods. We denote by $D$ the $W^1_p$-seminorm of the error, and by $\%R_{\text{reg}}$ the percentage of the error contributed by regular regions. The computation of the integral for the $W^1_p$-seminorm uses a similar method as for the $L_2$-norm (see (3.165)). In Figure 3.34 are shown the estimations for the derivatives.



(a) Uniform 1.

(b) Uniform 2.

Figure 3.32: Uniform triangulations to approximate $f$.

The number of triangles using a uniform triangulation is denoted by $N_{\text{uni}}$. We denote by $E_1$ and $E_2$ the $L_2$-norm of the errors which result from the approximation by using the uniform triangulations shown in Figure 3.32. In Figure 3.33, they are respectively plotted in green and blue lines. Between the two uniform triangulations 1 and 2, the latter is better adapted since the triangles are aligned in a direction which makes a small angle with the eigenvector $\mathbf{v}_1$, recalling that the derivatives of $f$ are minimal in $\mathbf{v}_1$ direction. We also see this from the $W_p^1$-seminorm of the derivatives, $D_1$ and $D_2$, which are plotted in Figure 3.34.

The tables below summarizes the data, $L_2$-norms and $W_p^1$-seminorms of the errors, which enable us to plot Figure 3.33 and Figure 3.34.

Table 1

| $N$ | $E^2$ | $\%E^2_{\text{reg}}$ | $\%A_{\text{reg}}$ | $N_{\text{uni}}$ | $E^2_1$ | $E^2_2$ |
|---|---|---|---|---|---|---|
| 440 | $1.74e-5$ | 0.28 | 18.953 | 450 | $1.47e-5$ | $2.40e-6$ |
| 568 | $3.46e-6$ | 1.24 | 20.66 | 578 | $8.88e-6$ | $1.46e-6$ |
| 634 | $8.37e-7$ | 4.28 | 21.68 | 648 | $7.07e-6$ | $1.16e-6$ |
| 704 | $7.74e-7$ | 6.62 | 27.99 | 722 | $5.69e-6$ | $9.33e-7$ |
| 826 | $2.54e-7$ | 14.49 | 31.67 | 882 | $3.81e-6$ | $6.25e-7$ |
| 932 | $1.35e-7$ | 34.78 | 42.41 | 968 | $3.17e-6$ | $5.19e-7$ |
| 1004 | $1.35e-7$ | 42.92 | 49.17 | 1058 | $2.65e-6$ | $4.34e-7$ |
| 1260 | $6.75e-8$ | 46.80 | 47.63 | 1352 | $1.62e-6$ | $2.66e-7$ |
| 1440 | $6.11e-8$ | 64.58 | 52.59 | 1458 | $1.40e-6$ | $2.29e-7$ |
| 1868 | $3.66e-8$ | 60.72 | 64.46 | 1922 | $8.03e-7$ | $1.32e-7$ |
| 2142 | $2.79e-8$ | 68.10 | 64.35 | 2178 | $6.26e-7$ | $1.02e-7$ |

Table 2

| $N$ | $D^2$ | $\%D^2_{\text{reg}}$ | $N_{\text{uni}}$ | $D^2_1$ | $D^2_2$ |
|---|---|---|---|---|---|
| 440 | $3.97e-3$ | 1.66 | 450 | $1.45e-2$ | $2.91e-3$ |
| 568 | $1.57e-3$ | 4.12 | 578 | $1.13e-2$ | $2.26e-3$ |
| 634 | $7.84e-4$ | 7.72 | 648 | $1.01e-2$ | $2.02e-3$ |
| 704 | $7.29e-4$ | 12.14 | 722 | $9.05e-3$ | $1.81e-3$ |
| 826 | $4.60e-4$ | 17.72 | 882 | $7.40e-3$ | $1.48e-3$ |
| 932 | $3.64e-4$ | 26.26 | 968 | $6.75e-3$ | $1.35e-3$ |
| 1004 | $3.23e-4$ | 30.50 | 1058 | $6.17e-3$ | $1.24e-3$ |
| 1260 | $1.99e-4$ | 36.88 | 1352 | $4.83e-3$ | $9.67e-4$ |
| 1440 | $1.91e-4$ | 46.96 | 1458 | $4.48e-3$ | $8.97e-4$ |
| 1868 | $1.49e-4$ | 44.97 | 1922 | $3.40e-3$ | $6.81e-4$ |
| 2142 | $1.19e-4$ | 49.58 | 2178 | $3.00e-3$ | $6.01e-4$ |

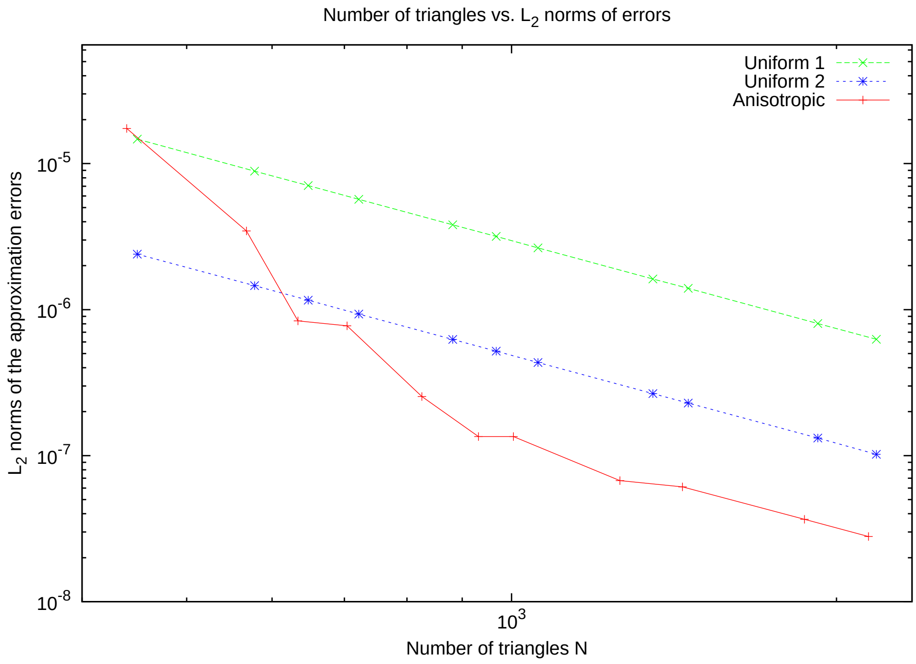


Figure 3.33: Squares of the $L_2$-norms of the approximation error $f - f_N$ by using our triangulation and the two uniform ones.

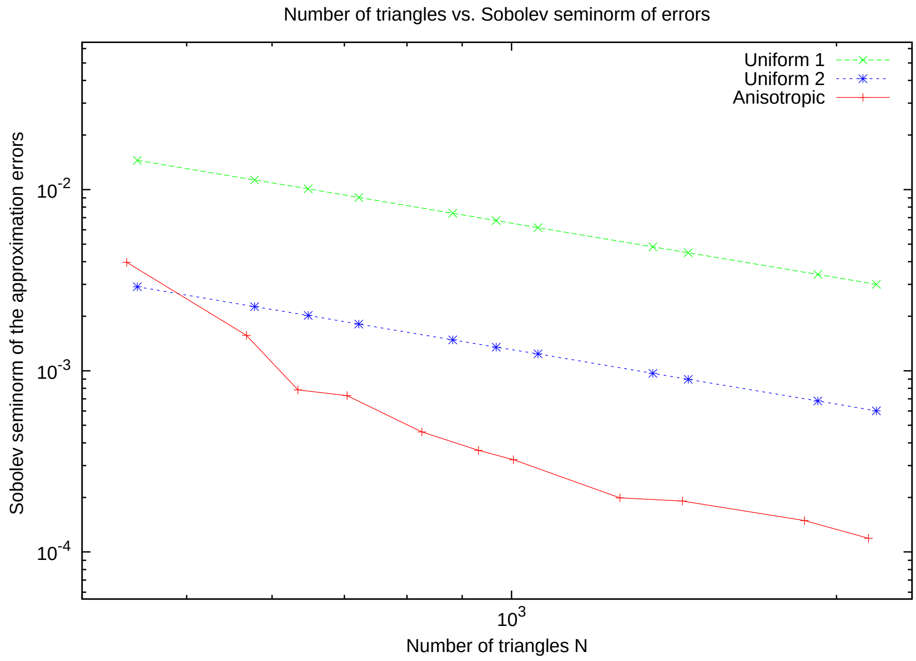


Figure 3.34: Squares of the $W_2^1$-seminorms of the approximation error $f - f_N$ by using our triangulation and the two uniform ones.

# Chapter 4

# Approximation by sums of piecewise polynomials

In this chapter, we present methods of approximation by using sums of piecewise polynomials of degree 0 or 1 in $d$ variables, extending the work in [16] on piecewise constants (see also [15]). The errors are measured in $L_p$-norm and $W_p^1$-seminorm. The basic idea consists in designing several overlaying partitions of the domain, responsible for the approximation of different components of the target function's gradient or Hessian. A key property of the new approximation method is that it achieves better approximation order in terms of the number $N$ of degrees of freedom comparing to standard methods on a single partition.

In the case of piecewise constants, with the errors measured in $L_p$-norm, the approximation order is $O(N^{-2/(d+1)})$ comparing to $O(N^{-1/d})$ achievable on a single isotropic partition. The order $O(N^{-2/(d+1)})$ has been shown in [16] for a single convex (anisotropic) partition. However, the construction of this partition requires the estimations of the average gradients of the target function $f$, whereas we use $d$ overlapping partitions independent of $f$.

In the case of piecewise linear approximations, with the errors measured in $L_\infty$-norm, it is known [16] that the order $O(N^{-2/d})$ cannot be improved on any single convex partition. We provide two methods of overlaying partitions, one with fewer partitions depending on $f$ and another with partitions independent of $f$, both with approximation order $O(N^{-6/(2d+1)})$ for the function and $O(N^{-3/(2d+1)})$

for the gradient, and the errors are measured in $L_p$-norm and $W_p^1$-seminorm.

Note that in contrast to Chapter 3, the approximant is not required to be continuous which makes partitioning easier by eliminating the need for irregular regions.

After providing in Section 4.1 the general concept of best approximation in the multivariate setting, we discuss in Section 4.2 the approximation method by using piecewise constants on a single partition. In Section 4.3, we present our approximation method by using sums of piecewise constant polynomials. The case for sums of piecewise linear polynomials is divided into two, in Section 4.4 and in Section 4.5, in the latter the directions of splitting in the partitions are fixed.

## 4.1 Generalities and notation

We shall start by introducing the general concept of best approximation. Given a bounded convex domain $\omega$ of $\mathbb{R}^d$, there is a constant $\rho_d$ which depends only on $d$ such that, for any $f \in W_p^1(\omega)$, we have the *Poincaré* inequality

$$\|f - f_\omega\|_{L_p(\omega)} \leq \rho_d \operatorname{diam}(\omega) \|\nabla f\|_{L_p(\omega)}, \tag{4.1}$$

where $f_\omega$ is the average of $f$ on $\omega$, that is

$$f_\omega = |\omega|^{-1} \int_\omega f(x)\,\mathrm{d}x, \tag{4.2}$$

with $|\omega|$ being the Lebesgue measure ($d$-dimensional volume) of $\omega$, and recalling from (1.22),

$$\|\nabla f\|_{L_p(\omega)} = \left( \int_\omega \left( \sum_{\nu=1}^{d} |D_{x_\nu} f(x)|^2 \right)^{\frac{p}{2}} \mathrm{d}x \right)^{\frac{1}{p}}. \tag{4.3}$$

Let $\Omega$ be a bounded domain in $\mathbb{R}^d$, $d \geq 2$. Given a partition $\Delta$ of $\Omega$ and a function $f : \Omega \to \mathbb{R}$, we are interested in estimating the error bounds resulting from its approximation by piecewise polynomials in the space

$$S_k(\Delta) = \Big\{ \sum_{\omega\in\Delta} q_\omega \chi_\omega : q_\omega \in \Pi_k^d \Big\}, \qquad \chi_\omega(x) := \begin{cases} 1, & \text{if } x \in \omega, \\ 0, & \text{otherwise,} \end{cases} \tag{4.4}$$

where $\Pi_k^d$, $k \geq 1$, denotes the space of polynomials of total degree $< k$ in $d$ variables[1]. For $k = 1, 2$, $S_1(\Delta)$ and $S_2(\Delta)$ are respectively the spaces of piecewise constant and piecewise linear polynomials on $\Omega$. Note that these functions are not necessarily continuous.

The best approximation error is measured in the $L_p$-norm $\| \cdot \|_{L_p(\Omega)}$,

$$E_k(f, \Delta)_p := \inf_{s\in S_k(\Delta)} \|f - s\|_{L_p(\Omega)}, \qquad 1 \leq p \leq \infty. \tag{4.5}$$

It is easy to check that,

$$E_k(f, \Delta)_p = \begin{cases} \Big( \sum_{\omega\in\Delta} E_k(f)^p_{L_p(\omega)} \Big)^{1/p} & \text{if } \ p < \infty, \\ \max_{\omega\in\Delta} E_k(f)_{L_\infty(\omega)} & \text{if } \ p = \infty, \end{cases} \tag{4.6}$$

where $E_k(f)_{L_p(\omega)}$ denotes the error of the best polynomial approximation on $\omega$,

$$E_k(f)_{L_p(\omega)} := \inf_{q\in\Pi_k^d} \|f - q\|_{L_p(\omega)}.$$

Indeed, suppose that the infimum in (4.5) is attained at $s_0 = \sum_{\omega\in\Delta} q_\omega^0 \chi_\omega$. Then

$$\begin{aligned} E_k(f, \Delta)_p =& \Big( \int_\Omega |(f - s_0)(x)|^p \mathrm{d}x \Big)^{\frac{1}{p}} = \Big( \sum_{\omega\in\Delta} \int_\omega |(f - s_0)(x)|^p \mathrm{d}x \Big)^{\frac{1}{p}} \\ =& \Big( \sum_{\omega\in\Delta} \int_\omega |(f - q_\omega^0)(x)|^p \mathrm{d}x \Big)^{\frac{1}{p}}. \end{aligned} \tag{4.7}$$

The right hand side of (4.6), for $p < \infty$, is clearly smaller or equal to the right hand side of (4.7). Conversely, the latter is smaller or equal to the right hand

[1] This is different of the standard definition where the total degree is $\leq k$.

side of (4.6) since $E_k(f, \Delta)_p$ is attained at $s_0$. The case $p = \infty$ is straightforward, and thus the (4.6) holds.

The Bramble-Hilbert lemma on convex domains [21] is stated as follows: If $\omega$ is a bounded convex domain and its restriction on $\omega$ belongs to the Sobolev space $W_p^k(\omega)$, then there is a polynomial $q \in \Pi_k^d$ such that

$$|f - q|_{W_p^r(\omega)} \leq \rho_{d,k} \operatorname{diam}^{k-r}(\omega) |f|_{W_p^k(\omega)}, \quad r = 0, \ldots, k, \tag{4.8}$$

where $\rho_{d,k}$ denotes a positive constant depending only on $d$ and $k$. In particular,

$$E_k(f)_{L_p(\omega)} \leq \rho_{d,k} \operatorname{diam}^k(\omega) \, |f|_{W_p^k(\omega)}. \tag{4.9}$$

Therefore, for any convex partition $\Delta$ of $\Omega$ and any $f \in W_p^k(\Omega)$,

$$E_k(f, \Delta)_p \leq \rho_{d,k} \max_{\omega \in \Delta} \operatorname{diam}^k(\omega) \, |f|_{W_p^k(\Omega)}. \tag{4.10}$$

Recall that $|\Delta|$ denotes the number of cells $\omega$ in $\Delta$. From the fact that

$$|\Omega| = \sum_{\omega \in \Delta} |\omega| \leq |\Delta| \max_{\omega \in \Delta} \operatorname{diam}(\omega)^d,$$

we have $\max_{\omega \in \Delta} \operatorname{diam}(\omega) \geq C |\Delta|^{-1/d}$, where $C$ depends only on $|\Omega|$ and $d$. Hence, in terms of $|\Delta|$, the approximation order that can be obtained from (4.10) is not better than

$$E_k(f, \Delta)_p = O(|\Delta|^{-k/d}). \tag{4.11}$$

This order is achieved for example for $\Omega = (0, 1)^d$ on convex partitions $\Delta_m$, $m = 1, 2, \ldots$, defined by splitting the cube $(0, 1)^d$ uniformly into $|\Delta_m| = m^d$ equal sub-cubes of edge length $1/m$.

Although the saturation order in (4.11) cannot be improved on isotropic partitions, it is shown in Theorem 4.2.1 that it can be improved on anisotropic partitions. For a system $\mathcal{P} = \{\Delta^{(1)}, \ldots, \Delta^{(n)}\}$ of several overlaying partitions of

$\Omega$, we consider the space of sums of piecewise polynomials

$$S_k(\mathcal{P}) = \left\{ \sum_{\nu=1}^{n} \sum_{\omega\in\Delta^{(\nu)}} q_{\nu,\omega}\chi_\omega \,:\, q_{\nu,\omega} \in \Pi_k^d \right\}. \tag{4.12}$$

A function in $S_k(\mathcal{P})$ is the sum of $n$ piecewise polynomials respectively belonging to $S_k\big(\Delta^{(\nu)}\big)$, $\nu = 1,\dots,n$. The corresponding best approximation error is measured with respect to the $L_p$-norm

$$E_k(f,\mathcal{P})_p := \inf_{s\in S_k(\mathcal{P})} \|f-s\|_p, \qquad 1 \le p \le \infty.$$

We set $|\mathcal{P}| = \sum\limits_{\nu=1}^{n} |\Delta^{(\nu)}|$.

## 4.2 Piecewise constant approximation

In [15, Theorem 2] it is shown that in (4.11) for $k = 1$, assuming higher smoothness of $f$ does not help to improve the order $E_1(f,\Delta_N)_\infty = O(|\Delta_N|^{-1/d})$ if the sequence of partitions $(\Delta_N)$ is *isotropic*, that is there is a constant $c > 0$ such that $\operatorname{diam}(\omega) \le c\rho(\omega)$ for all $\omega \in \bigcup_N \Delta_N$, where $\rho(\omega)$ is the maximum diameter of $d$-dimensional balls contained in $\omega$. More precisely, if $E_1(f,\Delta_N)_\infty = o(|\Delta_N|^{-1/d})$, $N \to \infty$, for a function $f \in C^1(\Omega)$ and some isotropic sequence of partitions $(\Delta_N)_{N\ge N_0}$ with $\lim\limits_{N\to\infty} \operatorname{diam}(\Delta_N) = 0$, then $f$ is a constant. Thus, $|\Delta|^{-1/d}$ is the *saturation order* of the piecewise constant approximation on isotropic partitions.

By using anisotropic partitions of $\Omega$, in [16] it has been shown that the approximation order of piecewise constants can be improved to $E_1(f,\Delta)_p = O(|\Delta|^{-2/(d+1)})$ on suitable anisotropic convex partitions obtained by a simple algorithm if $f \in W_p^2(\Omega)$.

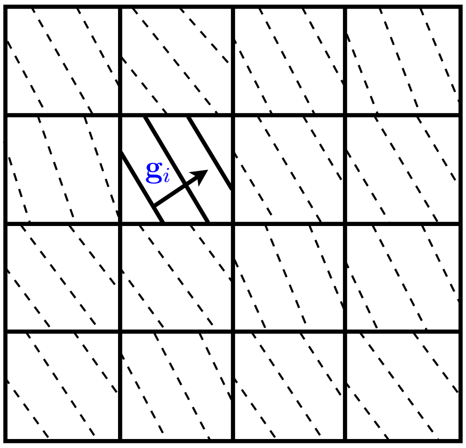


Figure 4.1: Example of a partition into $N_1$ cubes, with a cell partitioned into $N_2$ slices, with $m = 4$.

**Algorithm 4.1** ([16]): **A partition $\Delta_m$ of $\Omega$.**

Input $m \in \mathbb{N}$;

Assume that $f \in W_1^1(\Omega)$ where $\Omega = (0,1)^d$.

1. Split $\Omega$ into $N_1 = m^d$ cubes $\omega_1, \ldots, \omega_{N_1}$ of edge length $h = 1/m$;

2. Split each $\omega_i$ into $N_2$ slices $\omega_{ij}$, $j = 1, \ldots, N_2$, by equidistant hyperplanes orthogonal to the average gradient

$$\mathbf{g}_i := |\omega_i|^{-1} \int_{\omega_i} \nabla f(x)\, dx; \tag{4.13}$$

3. Set the partition $\Delta_m = \{\omega_{ij} : i = 1, \ldots, N_1,\ j = 1, \ldots, N_2\}$;

Then $|\Delta_m| = N_1 N_2$ and each $\omega_{ij}$ is a convex polyhedron with at most $2(d+1)$ facets.

In Figure 4.1 we show an example of partition into $N_1$ cubes for $d = 2$, where a cell is partitioned into $N_2$ slices according to Algorithm 4.1.

**Theorem 4.2.1** ([16]). *Assume that $f \in W_p^2(\Omega)$, $\Omega = (0,1)^d$, for some $1 \leq p \leq \infty$. For any $m = 1, 2, \ldots$, generate the partition $\Delta_m$ by using Algorithm 4.1 with $N_1 = m^d$ and $N_2 = m$. Then*

$$E_1(f, \Delta_m)_p \leq C_d |\Delta_m|^{-2/(d+1)} (|f|_{W_p^1(\Omega)} + |f|_{W_p^2(\Omega)}), \tag{4.14}$$

*where $C_d$ is a constant depending only on $d$.*

According to [16, Theorem 2], the saturation order of piecewise constant approximations on convex partitions is $|\Delta|^{-2/(d+1)}$, since for any $f \in C^2(\Omega)$ it cannot be improved any further. It is also shown in [16, Theorem 3] that the saturation order of piecewise linear approximations on convex partitions is $|\Delta|^{-2/d}$, which is the same as on isotropic partitions. The case $d = 2$ was proved by a different method in [26] for sequences of partitions $\Delta_N$ of $\Omega = (0,1)^2$.

## 4.3 Sums of piecewise constants on $\Omega \subset \mathbb{R}^d$

In contrast to the method using a single partition, we shall use a new algorithm which will involve a system of $d$ convex polyhedral partitions *independent* of $f$, that is, unlike in (4.13) where the splitting directions are fixed. On each partition, an approximant of $f$ can be determined (see (4.20) below). The error bounds are obtained by using triangular inequalities and the Poincaré inequality (4.1).

In the algorithm below, we produce a system of overlaying partitions $\mathcal{P}$ where the splitting directions are fixed.

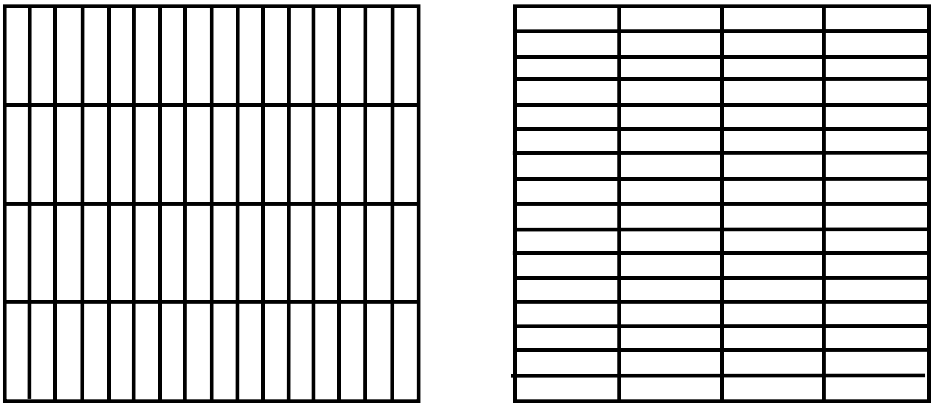


Figure 4.2: Partitions $\Delta^{(1)}, \Delta^{(2)}$ obtained from Algorithm 4.2 for piecewise constant approximation.

**Algorithm 4.2: Overlaying partitions $\Delta^{(\nu)}$, $\nu = 1, \ldots, d$, of $\Omega$**

Let $\Omega = (0,1)^d$ and let $m \in \mathbb{N}$.

1. Split $\Omega$ into $N_1 = m^d$ cubes $\omega_1, \ldots, \omega_{N_1}$ of edge length $h = 1/m$, whose edges are parallel to coordinate axes;
2. For each $\nu = 1, \ldots, d$, define $\Delta^{(\nu)}$ by splitting each $\omega_i$ into $N_2$ slices $\omega_{ij}^{(\nu)}$, $j = 1, \ldots, N_2$, by equidistant hyperplanes parallel to the subspace $x_\nu = 0$;
3. Set $\mathcal{P}_m = \{\Delta^{(1)}, \ldots, \Delta^{(d)}\}$;

Then $|\Delta^{(\nu)}| = N_1 N_2$ and each $\omega_{ij}^{(\nu)}$ is a $d$-dimensional box with its $\nu$-th dimension $\frac{h}{N_2}$ and all other dimensions $h$. We have $|\mathcal{P}_m| = dN_1N_2$.

Partitions $\Delta^{(1)}, \Delta^{(2)}$ in the case $d = 2$ are illustrated in Figure 4.2.

The following result is proved.

**Theorem 4.3.1.** *Assume that $f \in W_p^2(\Omega)$, $\Omega = (0,1)^d$, for some $1 \leq p \leq \infty$. For any $m = 1, 2, \ldots$, generate the system of partitions $\mathcal{P}_m$ by using Algorithm 4.2 with $N_1 = m^d$ and $N_2 = m$. Then*

$$E_1(f, \mathcal{P}_m)_p \leq C_d |\mathcal{P}_m|^{-2/(d+1)} (|f|_{W_p^1(\Omega)} + |f|_{W_p^2(\Omega)}), \tag{4.15}$$

*where $C_d$ is a constant depending only on $d$.*

*Proof.* For each $i = 1, \ldots, N_1$, let $\ell_i = c_i + \sum_{\nu=1}^{d} \ell_{i,\nu}$ denote a linear polynomial given by

$$\ell_{i,\nu} = a_{i,\nu}(x_\nu - x_{0,\nu}), \quad a_{i,\nu} = |\omega_i|^{-1} \int_{\omega_i} D_{x_\nu} f(x)\, dx, \quad \nu = 1, \ldots, d,$$

with $(x_{0,1}, \ldots, x_{0,\nu})$ being the barycenter of $\omega_i$ and the constant $c_i$ defined by the average of $f$ on $\omega_i$, i.e

$$c_i = |\omega_i|^{-1} \int_{\omega_i} f(x)\, dx.$$

Since $\int_{\omega_i} \ell_{i,\nu}(x)\, dx = 0$, we have by Poincaré inequality,

$$\begin{aligned} \|\Big(f - \sum_{\nu=1}^{d} \ell_{i,\nu}\Big) - c_i\|_{L_p(\omega_i)} &\leq \rho_d \operatorname{diam}(\omega_i) \|\nabla f - \sum_{\nu=1}^{d} \nabla \ell_{i,\nu}\|_{L_p(\omega_i)} \\ &\leq \frac{\sqrt{d}\rho_d}{m} \Big( \int_{\omega_i} \Big( \sum_{\nu=1}^{d} (D_{x_\nu} f(x) - a_{i,\nu})^2 \Big)^{p/2} dx \Big)^{1/p}. \end{aligned}$$

Poincaré inequality also implies

$$\begin{aligned} \|f - \sum_{i=1}^{N_1} \ell_i \chi_{\omega_i}\|_p &\leq \frac{\sqrt{d}\rho_d}{m} \Big( \sum_{i=1}^{N_1} \sum_{\nu=1}^{d} \int_{\omega_i} |D_{x_\nu} f(x) - a_{i,\nu}|^p\, dx \Big)^{\frac{1}{p}} \\ &\leq \frac{\sqrt{d}\rho_d}{m} \Big( \sum_{i=1}^{N_1} \sum_{\nu=1}^{d} \rho_d^p \operatorname{diam}(\omega_i)^p \|\nabla (D_{x_\nu} f)\|_{L_p(\omega_i)}^p \Big)^{\frac{1}{p}} \\ &\leq \frac{d\rho_d^2}{m^2} |f|_{W_p^2(\Omega)}. \end{aligned} \tag{4.16}$$

For fixed $i, \nu$, consider the sum of piecewise constant polynomials

$$s_{i,\nu} = \sum_{j=1}^{N_2} q_j \chi_{\omega_{ij}^{(\nu)}}, \tag{4.17}$$

where $q_j$ is a constant, $j = 1, \ldots, N_2$, which we shall make precise later. The $\nu$-th side of $\omega_{ij}^{(\nu)}$ is the straight line segment $[x_j^0, x_j^0 + \frac{1}{mN_2}]$ on the $x_\nu$-axis, and from a direct computation

$$\|\ell_{i,\nu} - s_{i,\nu}\|_{L_p(\omega_i)}^p = \sum_{j=1}^{N_2} \int_{\omega_{ij}^{(\nu)}} |a_{i,\nu}(x_\nu - x_{0,\nu}) - q_j|^p\, dx$$

$$= \frac{1}{m^{d-1}} \sum_{j=1}^{N_2} \int_0^{\frac{1}{mN_2}} |a_{i,\nu} x_\nu + a_{i,\nu}(x_j^0 - x_{0,\nu}) - q_j|^p \, dx_\nu$$
$$= \Big(\frac{1}{mN_2}\Big)^p \frac{|a_{i,\nu}|^p}{m^d(p+1)},$$

after choosing $q_j := a_{i,\nu}(x_j^0 - x_{0,\nu})$. Observe that

$$\frac{|a_{i,\nu}|^p}{m^d} = |\omega_i| \Big| \frac{1}{|\omega_i|} \int_{\omega_i} D_{x_\nu} f(x) \, dx \Big|^p \leq \int_{\omega_i} |D_{x_\nu} f(x)|^p \, dx.$$

Define $s$ as the sum of piecewise constant polynomials by

$$s = \sum_{i=1}^{N_1} (s_i + c_i) \chi_{\omega_i}, \tag{4.18}$$

where $s_i = \sum_{\nu=1}^d s_{i,\nu}$, we find that

$$\begin{aligned} \| \sum_{i=1}^{N_1} \ell_i \chi_{\omega_i} - s \|_p &= \Big( \sum_{i=1}^{N_1} \int_{\omega_i} \Big| \sum_{\nu=1}^d (\ell_{i,\nu}(x) - s_{i,\nu}(x)) \Big|^p \, dx \Big)^{\frac{1}{p}} \\ &\leq \Big( \sum_{i=1}^{N_1} d^{p-1} \sum_{\nu=1}^d \int_{\omega_i} |\ell_{i,\nu}(x) - s_{i,\nu}(x)|^p \, dx \Big)^{\frac{1}{p}} \\ &\leq \frac{d}{mN_2} \Big( \sum_{i=1}^{N_1} \sum_{\nu=1}^d \int_{\omega_i} |D_{x_\nu} f(x)|^p \, dx \Big)^{\frac{1}{p}} \\ &= \frac{d}{mN_2} |f|_{W_p^1(\Omega)}. \end{aligned} \tag{4.19}$$

Since $N_2 = m$ and $m^{-2} = \Big(\frac{|\mathcal{P}_m|}{d}\Big)^{\frac{-2}{d+1}}$, the bound (4.15) with $C_d = d^{1+\frac{2}{d+1}}(\rho_d^2 + 1)$ is obtained by combining (4.16) and (4.19). □

Observe that the sub-cubes $\omega_i$, $i = 1, \ldots, N_1$ are fixed on each partition $\Delta^{(\nu)}$, $\nu = 1, \ldots, d$, as described by step 1 of Algorithm 4.2. Hence the expression of $s$ in (4.17) is legitimate. In fact, $s$ can be expressed by $d$ piecewise polynomials $s = \sum_{\nu=1}^d f_\nu$ where

$$f_\nu = \sum_{i=1}^{N_1} (s_{i,\nu} + \frac{1}{d} c_i) \chi_{\omega_i}. \tag{4.20}$$

## 4.4 Sums of piecewise linears on $\Omega \subset \mathbb{R}^d$

In this section we approximate the function by using a sum of piecewise linear polynomials. As in the previous method, we design several overlaying partitions of $\Omega$ which are initially divided into sub-squares $\omega_i$, $i = 1, \ldots, N_1$.

Given a function $f \in W_p^2(\omega)$ where $\omega \subset \mathbb{R}^d$ is a bounded convex domain, recall that the average Hessian matrix of $f$ over $\omega$ is a $d \times d$ matrix whose entry $H_{\nu\mu}$ at the $\nu$-th row and $\mu$-th column is given by

$$H_{\nu\mu} = |\omega|^{-1} \int_\omega D^2_{x_\nu x_\mu} f(x)\,\mathrm{d}x.$$

**Algorithm 4.3: Overlaying partitions of $\Omega$**

Assume $f \in W_p^2(\Omega)$, $\Omega = (0,1)^d$ and $m \in \mathbb{N}$.

1. Split $\Omega$ into $N_1 = m^d$ cubes $\omega_1, \ldots, \omega_{N_1}$ of edge length $h = 1/m$, whose edges are parallel to coordinate axes;

2. For each $i = 1, \ldots, N_1$, compute the average Hessian matrix $H_i$ of $f$ over $\omega_i$, and let $\boldsymbol{\sigma}_i^{(\nu)}$, $\nu = 1, \ldots, d$, be the unit eigenvectors of $H_i$;

3. For each $\nu = 1, \ldots, d$, define $\Delta^{(\nu)}$ by splitting each $\omega_i$ into $N_2$ slices $\omega_{ij}^{(\nu)}$, $j = 1, \ldots, N_2$, by equidistant hyperplanes orthogonal to the eigenvector $\boldsymbol{\sigma}_i^{(\nu)}$;

Set $\mathcal{P}_m = \{\Delta^{(1)}, \ldots, \Delta^{(d)}\}$, where $|\Delta^{(\nu)}| = N_1 N_2$ and $|\mathcal{P}_m| = dN_1N_2$.

Partitions $\Delta^{(1)}, \Delta^{(2)}$ for Algorithm 4.3 in the case $d = 2$ are illustrated in Figure 4.3. On the first figure the splitting directions are orthogonal to the first eigenvectors, whereas in the second figure the directions of splittings are orthogonal to the second eigenvectors.

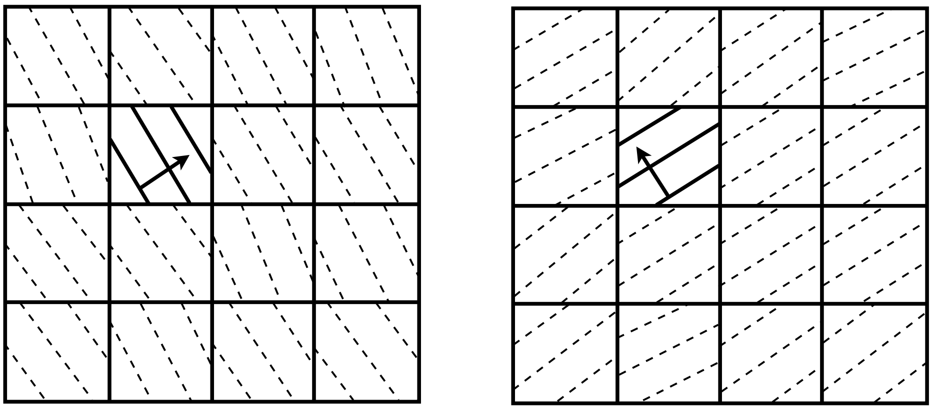


Figure 4.3: Partitions $\Delta^{(1)}, \Delta^{(2)}$ obtained from Algorithm 4.3 for piecewise linear approximation.

By using the above algorithm, we prove the following result.

**Theorem 4.4.1.** *Let $f \in W_p^3(\Omega)$, $\Omega = (0,1)^d$, for some $1 \le p < \infty$. For any $m = 1, 2, \dots$, generate the system of partitions $\mathcal{P}_m$ by using Algorithm 4.3 with $N_1 = m^d$ and $N_2 = \lceil m^{\frac{1}{2}} \rceil$. Then there exists a sum of piecewise linear functions $s_m \in S_2(\mathcal{P}_m)$ such that*

$$\|f - s_m\|_p \le C_1 |\mathcal{P}_m|^{-6/(2d+1)} (|f|_{W_p^2(\Omega)} + |f|_{W_p^3(\Omega)}), \tag{4.21}$$

$$|f - s_m|_{W_p^1(\Omega)} \le C_2 |\mathcal{P}_m|^{-3/(2d+1)} (|f|_{W_p^2(\Omega)} + |f|_{W_p^3(\Omega)}), \tag{4.22}$$

*where $C_1, C_2$ are constants depending only on $d$.*

*Proof.* Denote by $\Delta$ the partition of $\Omega$ into $N_1$ cubes $\omega_1, \dots, \omega_{N_1}$ of edge length $h = 1/m$. It follows from (4.8) that, for each $i = 1, \dots, N_1$, there exists a quadratic polynomial $q_i$ such that

$$\|f - q_i\|_{L_p(\omega_i)} \le \rho_{d,3} \operatorname{diam}(\omega_i)^3 |f|_{W_p^3(\omega_i)} \le \frac{d^{\frac{3}{2}} \rho_{d,3}}{m^3} |f|_{W_p^3(\omega_i)}, \tag{4.23}$$

$$|f - q_i|_{W_p^1(\omega_i)} \le \rho_{d,3} \operatorname{diam}(\omega_i)^2 |f|_{W_p^3(\omega_i)} \le \frac{d \rho_{d,3}}{m^2} |f|_{W_p^3(\omega_i)}, \tag{4.24}$$

$$|f - q_i|_{W_p^2(\omega_i)} \le \rho_{d,3} \operatorname{diam}(\omega_i) |f|_{W_p^3(\omega_i)} \le \frac{\sqrt{d} \rho_{d,3}}{m} |f|_{W_p^3(\omega_i)}. \tag{4.25}$$

It is clear from (4.23) that

$$\|f - \sum_{i=1}^{N_1} q_i \chi_{\omega_i}\|_p \leq \frac{d^{\frac{3}{2}} \rho_{d,3}}{m^3} |f|_{W_p^3(\Omega)}. \tag{4.26}$$

With $H_i$ denoting the average Hessian matrix of $f$ over $\omega_i$, let $\tilde{q}_i$ be the quadratic homogeneous polynomial whose coefficients are the entries of $H_i$, i.e

$$D^2_{x_\nu x_\mu} \tilde{q}_i = |\omega_i|^{-1} \int_{\omega_i} D^2_{x_\nu x_\mu} f(x)\, \mathrm{d}x, \quad \nu, \mu = 1, \ldots, d.$$

By using the Poincaré inequality, together with (4.25), clearly

$$\begin{aligned} \|D^2_{x_\nu x_\mu}(\tilde{q}_i - q_i)\|_{L_p(\omega_i)} &\leq \|D^2_{x_\nu x_\mu}(\tilde{q}_i - f)\|_{L_p(\omega_i)} + \|D^2_{x_\nu x_\mu}(f - q_i)\|_{L_p(\omega_i)} \\ &\leq \rho_d \operatorname{diam}(\omega_i) \|\nabla(D^2_{x_\nu x_\mu} f)\|_{L_p(\omega_i)} + \rho_{d,3} \operatorname{diam}(\omega_i) |f|_{W_p^3(\omega_i)}. \end{aligned} \tag{4.27}$$

From (4.8), there exists a linear polynomial $\tilde{\ell}_i$ such that

$$\|(q_i - \tilde{q}_i) - \tilde{\ell}_i\|_{L_p(\omega_i)} \leq \rho_{d,2} \operatorname{diam}(\omega_i)^2 |q_i - \tilde{q}_i|_{W_p^2(\omega_i)}, \tag{4.28}$$

$$|(q_i - \tilde{q}_i) - \tilde{\ell}_i|_{W_p^1(\omega_i)} \leq \rho_{d,2} \operatorname{diam}(\omega_i) |q_i - \tilde{q}_i|_{W_p^2(\omega_i)}. \tag{4.29}$$

The Hessian matrix $H_i$ can be diagonalized into $H_i = A^t D A$ where $A$ is an orthogonal matrix and $D$ a diagonal matrix with entries $\lambda_1, \ldots, \lambda_d$. With a slight abuse of notation we also denote by $A$ the linear mapping generated by the matrix $A$. Considering the linear transform $(X_1, \ldots, X_d)^t = A(x_1, \ldots, x_d)^t$, we use the notation

$$\bar{q}_i(X_1, \ldots, X_d) = \lambda_1 X_1^2 + \cdots + \lambda_d X_d^2,$$

where

$$(\tilde{q}_i + \tilde{\ell}_i) \circ A^{-1} = \bar{q}_i + \ell_i,$$

with $\ell_i$ being a linear polynomial in the variables $X_1, \ldots, X_d$. For $\nu = 1, \ldots, d$, each $X_\nu$ is a linear function of $x_1, \ldots, x_d$, and the eigenvector $\boldsymbol{\sigma}_i^{(\nu)}$ of $H_i$ is parallel to the $X_\nu$-axis.

Given $i, \nu$ and $j$, the set $A\omega_{ij}^{(\nu)}$ is contained between the hyperplanes $X_\nu = c_j$ and $X_\nu = c_j + \frac{\gamma_i^{(\nu)}}{mN_2}$, where $1 \leq \gamma_i^{(\nu)} \leq \sqrt{d}$ is the width of the unit cube in the direction of $\boldsymbol{\sigma}_i^{(\nu)}$. We set $\bar{s}_{i,\nu} = \sum_{j=1}^{N_2} \bar{\ell}_j \chi_{A(\omega_{ij}^{(\nu)})}$, where $\bar{\ell}_j = a_j X_\nu - b_j$, $a_j = 2\lambda_\nu c_j$ and $b_j = \lambda_\nu c_j^2$. Then

$$\begin{aligned}
\|\lambda_\nu X_\nu^2 - \bar{s}_{i,\nu}\|^p_{L_p(A(\omega_{ij}^{(\nu)}))} &= \int_{A\omega_{ij}^{(\nu)}} |\lambda_\nu X_\nu^2 - a_j X_\nu + b_j|^p \, dX \\
&\leq \frac{(\sqrt{d})^{d-1}}{m^{d-1}} \int_0^{\frac{\gamma_i^{(\nu)}}{mN_2}} |\lambda_\nu (X_\nu + c_j)^2 - a_j(X_\nu + c_j) + b_j|^p \, dX_\nu \\
&= \frac{|\lambda_\nu|^p}{(2p+1)m^d} \frac{(\sqrt{d})^{d-1}\big(\gamma_i^{(\nu)}\big)^{2p+1}}{N_2} \Big(\frac{1}{m^2 N_2^2}\Big)^p \\
&\leq \frac{|\lambda_\nu|^p}{m^d} \frac{(\sqrt{d})^{d+2p}}{N_2} \Big(\frac{1}{m^2 N_2^2}\Big)^p
\end{aligned} \tag{4.30}$$

by virtue of the fact that $\sqrt{d}$ is the diameter of the unit cube.

For each $\nu = 1, \ldots, d$, let $D_{X_\nu} g$ denote the partial derivative of a function $g$ with respect to the variable $X_\nu$. It is clear that $D_{X_\nu}(g \circ A^{-1})(X_1, \ldots, X_d) = (D_{\boldsymbol{\sigma}_\nu} g) \circ A^{-1}(X_1, \ldots, X_d)$, with $\boldsymbol{\sigma}_\nu = \boldsymbol{\sigma}_i^{(\nu)}$ being the $\nu$-th column vector of $A^{-1}$. It follows that

$$D^2_{X_\mu X_\nu}(g \circ A^{-1})(X_1, \ldots, X_d) = (D^2_{\sigma_\mu \sigma_\nu} g) \circ A^{-1}(X_1, \ldots, X_d).$$

For each $i, \nu$, since $\lambda_\nu = \frac{1}{2} D^2_{X_\nu X_\nu} \bar{q}_i(X)$ and $m^{-d} = |A\omega_i|$, we have

$$\frac{|\lambda_\nu|^p}{m^d} = \frac{1}{2^p} \int_{A\omega_i} |D^2_{X_\nu X_\nu}\big((\tilde{q}_i + \tilde{\ell}_i) \circ A^{-1}\big)(X)|^p \, dX = \frac{1}{2^p} \int_{\omega_i} |D^2_{\boldsymbol{\sigma}_\nu \boldsymbol{\sigma}_\nu}(\tilde{q}_i + \tilde{\ell}_i)(x)|^p \, \mathrm{d}x.$$

With $\bar{s}_i$ being the linear polynomial in the variables $X_1, \ldots, X_d$, defined by

$$\bar{s}_i = \ell_i + \sum_{\nu=1}^d \bar{s}_{i,\nu},$$

we have $(\tilde{q}_i + \tilde{\ell}_i) \circ A^{-1} - \bar{s}_i = \sum_{\nu=1}^{d} (\lambda_\nu X_\nu^2 - \bar{s}_{i,\nu})$ where, from (4.30),

$$\begin{aligned}
\| \sum_{\nu=1}^{d} (\lambda_\nu X_\nu^2 - \bar{s}_{i,\nu}) \|_{L_p(A(\omega_i))} &\leq \sum_{\nu=1}^{d} \| \lambda_\nu X_\nu^2 - \bar{s}_{i,\nu} \|_{L_p(A(\omega_i))} \\
&\leq d^{1-\frac{1}{p}} \Big( \sum_{\nu=1}^{d} \| \lambda_\nu X_\nu^2 - \bar{s}_{i,\nu} \|_{L_p(A(\omega_i))}^p \Big)^{\frac{1}{p}} \\
&\leq d^{1-\frac{1}{p}} \Big( \sum_{\nu=1}^{d} \sum_{j=1}^{N_2} \| \lambda_\nu X_\nu^2 - \bar{s}_{i,\nu} \|_{L_p(A(\omega_{ij}^{(\nu)}))}^p \Big)^{\frac{1}{p}} \\
&\leq \frac{d^{1-\frac{1}{p}} (\sqrt{d})^{2+\frac{d}{p}}}{2m^2 N_2^2} \Big( \sum_{\nu=1}^{d} \int_{\omega_i} |D^2_{\boldsymbol{\sigma}_\nu \boldsymbol{\sigma}_\nu} (\tilde{q}_i + \tilde{\ell}_i)(x)|^p \, dx \Big)^{\frac{1}{p}}, \qquad (4.31)
\end{aligned}$$

by virtue of the Hölder inequality (1.21).

Now denote by $a_{\mu,\nu}$ the entry of $A^{-1}$ at the $\mu$-th row and $\nu$-th column, with $|a_{\mu,\nu}| \leq 1$. Using the definition of directional derivatives, clearly

$$D^2_{\boldsymbol{\sigma}_\nu \boldsymbol{\sigma}_\nu} (\tilde{q}_i + \tilde{\ell}_i) = \sum_{j,k=1}^{d} a_{j,\nu} a_{k,\nu} D^2_{x_k x_j} (\tilde{q}_i + \tilde{\ell}_i).$$

We thus have the following inequalities,

$$\begin{aligned}
\int_{\omega_i} |D^2_{\boldsymbol{\sigma}_\nu \boldsymbol{\sigma}_\nu} (\tilde{q}_i + \tilde{\ell}_i)(x)|^p \, \mathrm{d}x &\leq d^{p-1} \sum_{j,k=1}^{d} \int_{\omega_i} |D^2_{x_k x_j} (\tilde{q}_i + \tilde{\ell}_i)(x)|^p \, \mathrm{d}x \\
&\leq d^{p-1} \sum_{j,k=1}^{d} \Big( \|D^2_{x_k x_j} (\tilde{q}_i + \tilde{\ell}_i) - D^2_{x_k x_j} f \|_{L_p(\omega_i)} + |f|_{W_p^2(\omega_i)} \Big)^p \\
&\leq d^{p+1} \Big( \rho_d \operatorname{diam}(\omega_i) |f|_{W_p^3(\omega_i)} + |f|_{W_p^2(\omega_i)} \Big)^p \\
&\leq d^{2p} \Big( \frac{\sqrt{d} \rho_d}{m} + 1 \Big)^p \Big( |f|^p_{W_p^3(\omega_i)} + |f|^p_{W_p^2(\omega_i)} \Big), \qquad (4.32)
\end{aligned}$$

by virtue of (4.1) and the Hölder inequality. For each $i = 1, \ldots, N_1$, denoting by $s_i$ the linear polynomial in the variables $x_1, \ldots, x_d$, given by

$$s_i = \bar{s}_i \circ A,$$

we have that $\|\tilde{q}_i + \tilde{\ell}_i - s_i\|_{L_p(\omega_i)} = \|\bar{q}_i - \bar{s}_i\|_{L_p(A\omega_i)}$. Hence (4.31) and (4.32) yield

$$\begin{aligned}\|\sum_{i=1}^{N_1}(\tilde{q}_i + \tilde{\ell}_i - s_i)\chi_{\omega_i}\|_p &= \Big(\sum_{i=1}^{N_1}\|\tilde{q}_i + \tilde{\ell}_i - s_i\|^p_{L_p(\omega_i)}\Big)^{\frac{1}{p}} \\ &\leq \frac{d^{4+\frac{d-1}{p}}}{m^2N_2^2}\Big(\frac{\sqrt{d}\rho_d}{m} + 1\Big)\Big(|f|_{W_p^3(\Omega)} + |f|_{W_p^2(\Omega)}\Big). \qquad (4.33)\end{aligned}$$

Consider the linear polynomial $s$ given by

$$s = \sum_{i=1}^{N_1} s_i\chi_{\omega_i}.$$

Since $m^{-3} \leq (mN_2)^{-2} \leq \Big(\frac{|\mathcal{P}_m|}{d}\Big)^{-6/(2d+1)}$, combining (4.26), (4.28) and (4.33) yields

$$\begin{aligned}\|f - s\|_p \leq \|f - &\sum_{i=1}^{N_1} q_i\chi_{\omega_i}\|_p \\ &+ \|\sum_{i=1}^{N_1}(q_i - \tilde{q}_i - \tilde{\ell}_i)\chi_{\omega_i}\|_p + \|\sum_{i=1}^{N_1}(\tilde{q}_i + \tilde{\ell}_i - s_i)\chi_{\omega_i}\|_p \\ &\leq C_1|\mathcal{P}_m|^{-6/(2d+1)}\Big(|f|_{W_p^3(\Omega)} + |f|_{W_p^2(\Omega)}\Big),\end{aligned}$$

where $C_1 = d^{\frac{6}{2d+1}}\Big(d^{\frac{3}{2}}\rho_{d,3} + d\rho_{d,2}(\rho_d + \rho_{d,3}) + \frac{d^{4+\frac{d-1}{p}}}{2}(\sqrt{d}\rho_d + 1)\Big)$, and (4.21) holds.

For each $i = 1, \ldots, N_1$ we observe that

$$\begin{aligned}|f - s_i|_{W_p^1(\omega_i)} \leq 3^{1-\frac{1}{p}}\Big(\sum_{\nu=1}^{d}\int_{\omega_i}&\Big(|D_{x_\nu}(f - q_i)(x)|^p \\ + |D_{x_\nu}(q_i - \tilde{q}_i - \tilde{\ell}_i)(x)|^p &+ |D_{x_\nu}(\tilde{q}_i + \tilde{\ell}_i - s_i)(x)|^p\Big)\,dx\Big)^{\frac{1}{p}}. \qquad (4.34)\end{aligned}$$

Note also that $A^{-1} = A^t$, so that the $\nu$-th row and $\mu$-th column of $A$ are exactly the $\mu$-th row and $\nu$-th $a_{\nu,\mu}$ column of $A^{-1}$. For each $\nu = 1, \ldots, d$, from the equality $\tilde{q}_i + \tilde{\ell}_i = (\bar{q}_i + \ell_i) \circ A$, we deduce that

$$D_{x_\nu}(\tilde{q}_i + \tilde{\ell}_i) = D_{x_\nu}((\bar{q}_i + \ell_i) \circ A) = (D_{\boldsymbol{\tau}_\nu}(\bar{q}_i + \ell_i)) \circ A = \sum_{\mu=1}^{d} a_{\nu,\mu}(D_{X_\mu}(\bar{q}_i + \ell_i)) \circ A,$$

where $\boldsymbol{\tau}_\nu$ denotes the $\nu$-th column vector of $A$. Similarly, $D_{x_\nu}s_i = D_{x_\nu}(\bar{s}_i \circ A) =$

$\sum_{\mu=1}^{d} a_{\mu,\nu}(D_{X_\mu}\bar{s}_i)\circ A$. It follows that

$$\begin{aligned}
\int_{\omega_i} |D_{x_\nu}(\tilde{q}_i+\tilde{\ell}_i-s_i)(x)|^p\,dx &\leq d^{p-1}\sum_{\mu=1}^{d}\int_{A\omega_i}|D_{X_\mu}(\bar{q}_i+\ell_i-\bar{s}_i)(X)|^p\,dX \\
&= d^{p-1}\sum_{\mu=1}^{d}\sum_{j=1}^{N_2}\int_{A\omega_{ij}^{(\mu)}}|2\lambda_\mu X_\mu - 2\lambda_\mu c_j|^p\,dX \\
&\leq \frac{d^{p-1}}{m^{d-1}}\sum_{\mu=1}^{d}\sum_{j=1}^{N_2}|2\lambda_\mu|^p\frac{1}{p+1}\Big(\frac{\sqrt{d}}{mN_2}\Big)^{p+1} \\
&= d^{\frac{3p-1}{2}}\Big(\frac{1}{mN_2}\Big)^p\sum_{\mu=1}^{d}\int_{\omega_i}|D^2_{\boldsymbol{\sigma}_\mu\boldsymbol{\sigma}_\mu}(\tilde{q}_i+\tilde{\ell}_i)(x)|^p\,dx. \qquad (4.35)
\end{aligned}$$

Combining (4.24) and (4.29), together with (4.35) and (4.32), we obtain

$$\begin{aligned}
|f-s_i|^p_{W^1_p(\omega_i)} \leq & 3^{p-1}\Big(\frac{d\rho_{d,3}}{m^2}\Big)^p|f|^p_{W^3_p(\omega_i)} + 3^{p-1}\Big(\frac{\sqrt{d}\rho_{d,2}}{m^2}\Big)^p|f|^p_{W^3_p(\omega_i)} \\
& + 3^p\Big(\frac{d^{\frac{7p-1}{2}}}{mN_2}\Big)^p\Big(\frac{\sqrt{d}\rho_d}{m}+1\Big)^p\Big(|f|^p_{W^3_p(\omega_i)}+|f|^p_{W^2_p(\omega_i)}\Big),
\end{aligned}$$

where, since $m^{-2}\leq (mN_2)^{-1}\leq \Big(\frac{|\mathcal{P}_m|}{d}\Big)^{-3/(2d+1)}$,

$$|f-s|_{W^1_p(\Omega)} \leq C_2|\mathcal{P}_m|^{-3/(2d+1)}\Big(|f|_{W^3_p(\Omega)}+|f|_{W^2_p(\Omega)}\Big),$$

with $C_2 = d^{\frac{3}{2d+1}}\Big(3d\rho_{d,3}+3\sqrt{d}\rho_{d,2}+3d^{\frac{7p-1}{2}}(\sqrt{d}\rho_d+1)\Big)$, and (4.22) is proved. □

## 4.5 Sums of piecewise linears with fixed splitting directions

In the previous section, the splitting directions in step 3 of Algorithm 4.3 depend on the eigenvectors of the average Hessian matrices of $f$. In this section, we present another method where the splitting directions are independent of the function.

The following result is needed in the proof of Theorem 4.5.2 below.

**Lemma 4.5.1.** *Any homogeneous quadratic polynomial $q$ can be represented as* *a linear combination of $\binom{d+1}{2}$ quadratic ridge functions*

$$q = \sum_{\nu=1}^{d} a_\nu x_\nu^2 + \sum_{\nu=1}^{d} \sum_{\mu=\nu+1}^{d} b_{\nu\mu}(x_\nu + x_\mu)^2, \tag{4.36}$$

*where*

$$a_\nu = \frac{1}{2} D_{x_\nu x_\nu} q - \frac{1}{2} \sum_{\mu \neq \nu} D_{x_\nu x_\mu} q, \quad b_{\nu\mu} = \frac{1}{2} D_{x_\nu x_\mu} q. \tag{4.37}$$

*Proof.* It is clear that $q$ in (4.36) is a quadratic polynomial which is indeed a combination of $d + (d-1) + \cdots + 1 = \binom{d+1}{2}$ ridge functions. Since second derivatives are linear operators, we just need to find the representation (4.36) for all quadratic monomials. For $q = x_\nu^2$, we simply take $a_\nu = 1$, and equate all other coefficients to zero. Moreover, for $\nu \neq \mu$,

$$2x_\nu x_\mu = (x_\nu + x_\mu)^2 - x_\nu^2 - x_\mu^2,$$

so that for $q = x_\nu x_\mu$ we can use $b_{\nu\mu} = \frac{1}{2}$, $a_\nu = a_\mu = -\frac{1}{2}$ which satisfy

$$a_\nu = a_\mu = \frac{1}{2} D_{x_\nu x_\mu} q - \frac{1}{2} \sum_{k \neq j} D_{x_k x_j} \quad \text{and} \quad b_{\nu\mu} = \frac{1}{2} D_{x_\nu x_\mu} q.$$

Note that if $q$ is of the form given in (4.36), then the formulas (4.37) follow directly from the fact that, for any $\nu$,

$$D_{x_\nu} q = 2a_\nu x_\nu + 2 \sum_{\nu \neq \mu} b_{x_\nu x_\mu}(x_\nu + x_\mu).$$

This concludes our proof. □

In the algorithm below, the initial partition in step 1 is the same as in Algorithm 4.3, however the splitting directions are fixed.

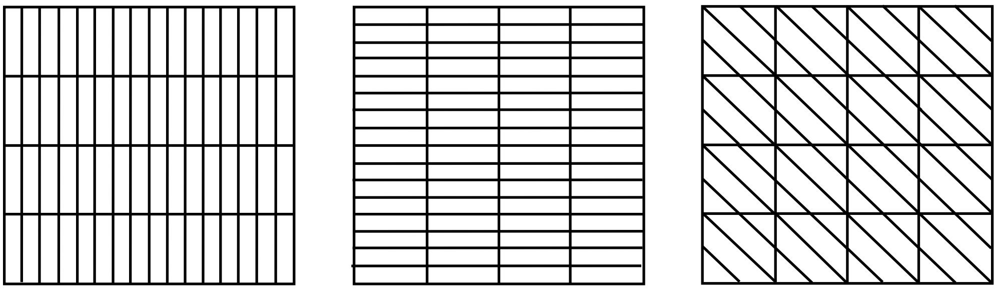


Figure 4.4: Partitions $\Delta^{(1)}, \Delta^{(2)}$ and $\Delta^{(1,2)}$ obtained from Algorithm 4.4.

**Algorithm 4.4: Partition of $\Omega$ with fixed directions**

Assume $f \in W_p^2(\Omega)$, $\Omega = (0,1)^d$ and let $m \in \mathbb{N}$.

1. Split $\Omega$ into $N_1 = m^d$ cubes $\omega_1, \dots, \omega_{N_1}$ of edge length $h = 1/m$, whose edges are parallel to coordinate axes;

2. For each $\nu, \mu = 1, \dots, d$, define $\Delta^{(\nu)}$ and $\Delta^{(\nu,\mu)}$ respectively by splitting each $\omega_i$, $i = 1, \dots, N_1$, into $N_2$ slices $\omega_{ij}^{(\nu)}$ and $\omega_{ij}^{(\nu,\mu)}$ $j = 1, \dots, N_2$, by equidistant hyperplanes parallel to $x_\nu = 0$ and $x_\nu + x_\mu = 0$, respectively.

Set $\mathcal{P}_m = \{\Delta^{(\nu)}, \Delta^{(\nu,\mu)}, \nu = 1, \dots, d,\ \mu = \nu + 1, \dots, d\}$ where for $\nu, \mu \in \{1, \dots, d\}$, $|\Delta^{(\nu)}| = |\Delta^{(\nu,\mu)}| = N_1 N_2$ and $|\mathcal{P}_m| = \binom{d+1}{2} N_1 N_2$.

Partitions $\Delta^{(1)}, \Delta^{(2)}$ and $\Delta^{(1,2)}$ in the case $d = 2$ are illustrated in Figure 4.4.

We prove the following result.

**Theorem 4.5.2.** *Let $f \in W_p^3(\Omega)$, $\Omega = (0,1)^d$, for some $1 \le p \le \infty$. For any $m = 1, 2, \dots$, generate the system of partitions $\mathcal{P}_m$ by using Algorithm 4.4 with $N_1 = m^d$ and $N_2 = \lfloor m^{\frac{1}{2}} \rfloor$. Then there is a sum of piecewise linear polynomials $s_m \in S_2(\mathcal{P}_m)$ such that*

$$\|f - s_m\|_p \le C_1 |\mathcal{P}_m|^{-6/(2d+1)} (|f|_{W_p^2(\Omega)} + |f|_{W_p^3(\Omega)}), \tag{4.38}$$

$$|f - s_m|_{W_p^1(\Omega)} \le C_2 |\mathcal{P}_m|^{-3/(2d+1)} (|f|_{W_p^2(\Omega)} + |f|_{W_p^3(\Omega)}), \tag{4.39}$$

*where $C_1, C_2$ are constants depending only on $d$.*

*Proof.* Denote by $\Delta$ the partition of $\Omega$ into $N_1$ cubes $\omega_1, \ldots, \omega_{N_1}$ of edge length $h = 1/m$. For each $i = 1, \ldots, N_1$, by (4.8) there is a quadratic polynomial $q_i$ such that

$$\|f - q_i\|_{L_p(\omega_i)} \leq \rho_{d,3} \operatorname{diam}(\omega_i)^3 |f|_{W_p^3(\omega_i)} \leq \frac{d^{\frac{3}{2}}\rho_{d,3}}{m^3} |f|_{W_p^3(\omega_i)}, \tag{4.40}$$

$$|f - q_i|_{W_p^1(\omega_i)} \leq \rho_{d,3} \operatorname{diam}(\omega_i)^2 |f|_{W_p^3(\omega_i)} \leq \frac{d\rho_{d,3}}{m^2} |f|_{W_p^3(\omega_i)}, \tag{4.41}$$

$$|f - q_i|_{W_p^2(\omega_i)} \leq \rho_{d,3} \operatorname{diam}(\omega_i) |f|_{W_p^3(\omega_i)} \leq \frac{\sqrt{d}\rho_{d,3}}{m} |f|_{W_p^3(\omega_i)}. \tag{4.42}$$

By using (4.36) and the notation therein, let $q_i = q_i^{(1)} + q_i^{(2)}$ where

$$q_i^{(1)} = \sum_{\nu=1}^{d} a_\nu x_\nu^2, \quad \text{and} \quad q_i^{(2)} = \sum_{\nu=1}^{d} \sum_{\mu=\nu+1}^{d} b_{\nu\mu}(x_\nu + x_\mu)^2.$$

For fixed $\nu = 1, \ldots, d$ and $j = 1, \ldots, N_2$, there exists $c_j$ such that the $\nu$-th side of $\omega_{ij}^{(\nu)}$ is given by $[c_i, c_i + \frac{1}{mN_2}]$. Considering the piecewise linear function

$$s_i^{(1)} = \sum_{j=1}^{N_2} \sum_{\nu=1}^{d} (2a_\nu c_j x_\nu - a_\nu c_j^2) \chi_{\omega_{ij}^{(\nu)}},$$

we have that

$$\begin{aligned} \|q_i^{(1)} - s_i^{(1)}\|_{L_p(\omega_i)}^p &= \int_{\omega_i} \left| \sum_{\nu=1}^{d} a_\nu x_\nu^2 - \sum_{j=1}^{N_2} \sum_{\nu=1}^{d} (2a_\nu c_j x_\nu - a_\nu c_j^2) \chi_{\omega_{ij}^{(\nu)}} \right|^p \mathrm{d}x \\ &\leq d^{p-1} \sum_{\nu=1}^{d} \int_{\omega_i} \left| a_\nu x_\nu^2 - \sum_{j=1}^{N_2} (2a_\nu c_j x_\nu - a_\nu c_j^2) \chi_{\omega_{ij}^{(\nu)}} \right|^p, \end{aligned}$$

by virtue of the Hölder inequality. Recalling that for each fixed $\nu$ the cube $\omega_i$ is split into $N_2$ slices $\omega_{ij}^{(\nu)}$, we use a linear change of variables to obtain

$$\begin{aligned} \|q_i^{(1)} - s_i^{(1)}\|_{L_p(\omega_i)}^p &\leq d^{p-1} \sum_{j=1}^{N_2} \sum_{\nu=1}^{d} \int_{\omega_{ij}^{(\nu)}} |a_\nu|^p |x_\nu - c_j|^{2p} dx \\ &= d^{p-1} \sum_{j=1}^{N_2} \sum_{\nu=1}^{d} \frac{1}{m^{d-1}} \int_{c_j}^{c_j + \frac{1}{mN_2}} |a_\nu|^p |x_\nu - c_j|^{2p} dx_\nu \end{aligned}$$

$$
\begin{aligned}
&= d^{p-1} \sum_{j=1}^{N_2} \sum_{\nu=1}^{d} \frac{1}{m^{d-1}} \int_0^{\frac{1}{mN_2}} |a_\nu|^p |x_\nu|^{2p} dx_\nu \\
&= d^{p-1} \sum_{\nu=1}^{d} \frac{|a_\nu|^p}{(2p+1)m^d} \Big(\frac{1}{mN_2}\Big)^{2p}. \qquad (4.43)
\end{aligned}
$$

By (4.37) and (4.42), for each $i = 1, \dots, N_1$, by adding and removing the terms $D^2_{x_\nu x_\nu} f$ and $D^2_{x_\nu x_\mu} f$, we get

$$
\begin{aligned}
\sum_{\nu=1}^{d} \frac{|a_\nu|^p}{m^d} =& \frac{1}{2^p} \int_{\omega_i} \sum_{\nu=1}^{d} \Big| D^2_{x_\nu x_\nu} q_i(x) - \sum_{\mu \neq \nu} D^2_{x_\nu x_\mu} q_i(x) \Big|^p dx \\
=& \frac{1}{2^p} \int_{\omega_i} \sum_{\nu=1}^{d} \Big( \Big| D^2_{x_\nu x_\nu} (q_i - f)(x) + D^2_{x_\nu x_\nu} f(x) \\
& + \sum_{\mu \neq \nu} \Big( D^2_{x_\nu x_\mu} (q_i - f)(x) + D^2_{x_\nu x_\mu} f(x) \Big) \Big|^p \Big) dx \\
\leq& \frac{4^{p-1}}{2^p} \Big( 2|q_i - f|^p_{W^2_p(\omega_i)} + 2|f|^p_{W^2_p(\omega)} \Big) \\
\leq& 2^{p-1} \Big( \frac{\sqrt{d}\rho_{d,3}}{m} \Big)^p |f|^p_{W^3_p(\omega_i)} + 2^{p-1} |f|^p_{W^2_p(\omega_i)}, \qquad (4.44)
\end{aligned}
$$

and (4.43) implies

$$
\|q_i^{(1)} - s_i^{(1)}\|^p_{L_p(\omega_i)} \leq \Big( \frac{2d}{m^2 N_2^2} \Big)^p \Big( \Big( \frac{\sqrt{d}\rho_{d,3}}{m} \Big)^p |f|^p_{W^3_p(\omega_i)} + |f|^p_{W^2_p(\omega_i)} \Big). \qquad (4.45)
$$

Considering the piecewise linear polynomial

$$
s_i^{(2)} = \sum_{j=1}^{N_2} \sum_{\nu=1}^{d} \sum_{\mu=\nu+1}^{d} \Big( 2b_{\nu\mu} b_j (x_\nu + x_\mu) - b_{\nu\mu} b_j^2 \Big) \chi_{\omega_{ij}^{(\nu,\mu)}},
$$

we obtain.

$$
\begin{aligned}
\|q_i^{(2)} - s_i^{(2)}\|^p_{L_p(\omega_i)} = \int_{\omega_i} \Big| & \sum_{\nu=1}^{d} \sum_{\mu=\nu+1}^{d} b_{\nu\mu} (x_\nu + x_\mu)^2 \\
& - \sum_{j=1}^{N_2} \sum_{\nu=1}^{d} \sum_{\mu=\nu+1}^{d} \Big( 2b_{\nu\mu} b_j (x_\nu + x_\mu) - b_{\nu\mu} b_j^2 \Big) \chi_{\omega_{ij}^{(\nu,\mu)}} \Big|^p \mathrm{d}x
\end{aligned}
$$

$$\begin{aligned}\leq & d^{2p-2}\sum_{\nu=1}^{d}\sum_{\mu=\nu+1}^{d}\int_{\omega_i}\Big|b_{\nu\mu}(x_\nu+x_\mu)^2\\ &-\sum_{j=1}^{N_2}\Big(2b_{\nu\mu}b_j(x_\nu+x_\mu)-b_{\nu\mu}b_j^2\Big)\chi_{\omega_{ij}^{(\nu,\mu)}}\Big|^p\,\mathrm{d}x.\end{aligned}$$

Recalling that for $\nu\neq\mu$ the cube $\omega_i$ is split into $N_2$ slices $\omega_{ij}^{(\nu,\mu)}$, we find that

$$\|q_i^{(2)}-s_i^{(2)}\|_{L_p(\omega_i)}^p\leq d^{2p-2}\sum_{j=1}^{N_2}\sum_{\nu=1}^{d}\sum_{\mu=\nu+1}^{d}\int_{\omega_{ij}^{(\nu,\mu)}}|b_{\nu\mu}|^p|x_\nu+x_\mu-b_j|^{2p}\,dx.$$

Given $\nu=1,\dots,d$ and $\mu=\nu+1,\dots,d$, there exists $b_j$ such that the $\nu$-th side of $\omega_{ij}^{(\nu,\mu)}$ lies between the hyperplanes $x_\nu+x_\mu=b_j$ and $x_\nu+x_\mu=b_j+\frac{\sqrt{2}}{mN_2}$. Consider the change of variable $X=x_\nu+x_\mu$ and $Y=x_\nu-x_\mu$ where $b_j\leq X\leq b_j+\frac{\sqrt{2}}{mN_2}$ and the range of $Y$ is at most $\frac{\sqrt{2}}{m}$. It follows that

$$\begin{aligned}\|q_i^{(2)}-s_i^{(2)}\|_{L_p(\omega_i)}^p &\leq d^{2p-2}\sum_{j=1}^{N_2}\sum_{\nu=1}^{d}\sum_{\mu=\nu+1}^{d}\frac{|b_{\nu\mu}|^p}{m^{d-2}}\Big(\frac{\sqrt{2}}{m}\int_{b_j}^{b_j+\frac{\sqrt{2}}{mN_2}}|X-b_j|^{2p}\,\mathrm{d}X\Big)\\ &\leq d^{2p-2}\sum_{j=1}^{N_2}\sum_{\nu=1}^{d}\sum_{\mu=\nu+1}^{d}\frac{\sqrt{2}|b_{\nu\mu}|^p}{m^{d-1}}\frac{1}{2p+1}\Big(\frac{\sqrt{2}}{mN_2}\Big)^{2p+1}. \end{aligned}\tag{4.46}$$

By (4.37) and (4.42), for each $i=1,\dots,N_1$, we deduce that

$$\begin{aligned}\sum_{\nu=1}^{d}\sum_{\mu=1}^{d}\frac{|b_{\nu\mu}|^p}{m^d}=&\sum_{\nu=1}^{d}\sum_{\mu=1}^{d}\int_{\omega_i}|\frac{1}{2}D_{x_\nu x_\mu}q_i(x)|^p dx\\ \leq&2^{p-1}\sum_{\nu=1}^{d}\sum_{\mu=1}^{d}\int_{\omega_i}\Big|\frac{1}{2}D_{x_\nu x_\mu}(q_i-f)\Big|^p dx+\frac{1}{2}|f|_{W_p^2(\omega_i)}^p\\ \leq&\frac{1}{2}\Big(\frac{\sqrt{d}\rho_{d,3}}{m}\Big)^p|f|_{W_p^3(\omega_i)}^p+\frac{1}{2}|f|_{W_p^2(\omega_i)}^p.\end{aligned}\tag{4.47}$$

Combining (4.47) and (4.46) yields

$$\|q_i^{(2)}-s_i^{(2)}\|_{L_p(\omega_i)}^p\leq d^{2p-2}\Big(\frac{2}{m^2N_2^2}\Big)^p\Big(\Big(\frac{\sqrt{d}\rho_{d,3}}{m}\Big)^p|f|_{W_p^3(\omega_i)}^p+|f|_{W_p^2(\omega_i)}^p\Big).\tag{4.48}$$

With $s_i = s_i^{(1)} + s_i^{(2)}$, combining (4.45) and (4.48) gives

$$\begin{aligned}\|q_i - s_i\|^p_{L_p(\omega_i)} \leq& 2^{p-1}\|q_i^{(1)} - s_i^{(1)}\|^p_{L_p(\omega_i)} + 2^{p-1}\|q_i^{(2)} - s_i^{(2)}\|^p_{L_p(\omega_i)} \\ \leq& \Big(d^p + d^{2p-2}\Big)\Big(\frac{4}{m^2N_2^2}\Big)^p\Big(\Big(\frac{\sqrt{d}\rho_{d,3}}{m}\Big)^p|f|^p_{W_p^3(\omega_i)} + |f|^p_{W_p^2(\omega_i)}\Big) \\ \leq& \Big(\frac{2d}{mN_2}\Big)^{2p}\Big(\Big(\frac{\sqrt{d}\rho_{d,3}}{m}\Big)^p + 1\Big)\Big(|f|^p_{W_p^3(\omega_i)} + |f|^p_{W_p^2(\omega_i)}\Big). \end{aligned} \tag{4.49}$$

The inequality $\max\{m^{-3}, (mN_2)^{-2}\} \leq 4\binom{d+1}{2}^{6/(2d+1)}|\mathcal{P}_m|^{-6/(2d+1)}$ is easily provable:

- Since $|\mathcal{P}_m| = \binom{d+1}{2}m^dN_2 \leq \binom{d+1}{2}m^{d+\frac{1}{2}}$, we have that $|\mathcal{P}_m|^{\frac{6}{2d+1}} \leq \binom{d+1}{2}^{6/(2d+1)}m^3$;
- Clearly $m \leq \lceil m^{\frac{1}{2}}\rceil^2 \leq (N_2+1)^2 \leq 4N_2^2$. Thus $m^3 \leq 4m^2N_2^2$ and hence $|\mathcal{P}_m|^{\frac{6}{2d+1}} \leq 4\binom{d+1}{2}^{6/(2d+1)}m^2N_2^2$.

Considering the piecewise polynomials $s = \sum_{i=1}^{N_1} s_i\chi_{\omega_i}$ and $q = \sum_{i=1}^{N_1} q_i\chi_{\omega_i}$, we now deduce from (4.40) and (4.49) that

$$\begin{aligned}\|f - s\|_p \leq& \Big(\sum_{i=1}^{N_1}\|f - q_i\|^p_{L_p(\omega_i)}\Big)^{\frac{1}{p}} + \Big(\sum_{i=1}^{N_1}\|q_i - s_i\|^p_{L_p(\omega_i)}\Big)^{\frac{1}{p}} \\ \leq& \frac{d^{\frac{3}{2}}\rho_{d,3}}{m^3}|f|_{W_p^3(\Omega)} + \frac{4d^2+2d}{m^2N_2^2}\Big(\sqrt{d}\rho_{d,3} + 1\Big)\Big(|f|_{W_p^3(\Omega)} + |f|_{W_p^2(\Omega)}\Big) \\ \leq& C_1|\mathcal{P}_m|^{-6/(2d+1)}\Big(|f|_{W_p^3(\Omega)} + |f|_{W_p^2(\Omega)}\Big), \end{aligned} \tag{4.50}$$

where $C_1 = 4\binom{d+1}{2}^{6/(2d+1)}\Big(d^{\frac{3}{2}}\rho_{d,3} + 6d^2\Big(\sqrt{d}\rho_{d,3} + 1\Big)\Big)$, thereby proving the result (4.38).

For each $i = 1, \ldots, N_1$, using the Hölder inequality yields

$$|q_i - s_i|^p_{W_p^1(\omega_i)} \leq 2^{p-1}\sum_{\nu=1}^{d}\Big(\|D_{x_\nu}(q_i^{(1)} - s_i^{(1)})\|^p_{L_p(\omega_i)} + \|D_{x_\nu}(q_i^{(2)} - s_i^{(2)})\|^p_{L_p(\omega_i)}\Big).$$

On one hand, a direct computation shows that, for each $\nu = 1, \ldots, d$,

$$\|D_{x_\nu}(q_i^{(1)} - s_i^{(1)})\|^p_{L_p(\omega_i)} = \frac{2^p}{p+1}\frac{|a_\nu|^p}{m^d}\Big(\frac{1}{mN_2}\Big)^p. \tag{4.51}$$

On the other hand, since for each $k = 1, \dots, d$,

$$D_{x_k} q_i^{(2)} = D_{x_k}\Big(\sum_{\nu=1}^{d}\sum_{\mu=\nu+1}^{d} b_{\nu\mu}(x_\nu + x_\mu)^2\Big) = \sum_{\mu\neq k}^{d} 2b_{k\mu}(x_k + x_\mu),$$

and

$$D_{x_k} s_i^{(2)} = D_{x_k}\sum_{j=1}^{N_2}\sum_{\nu=1}^{d}\sum_{\mu=1}^{d}\Big(2b_{\nu\mu}b_j(x_\nu + x_\mu) - b_{\nu\mu}b_j^2\Big) = \sum_{j=1}^{N_2}\sum_{\mu\neq k}^{d} 2b_{k\mu}b_j.$$

By using the Hölder inequality, we deduce that

$$\begin{aligned} \|D_{x_k}(q_i^{(2)} - s_i^{(2)})\|_{L_p(\omega_i)}^p &\le d^{p-1}\sum_{j=1}^{N_2}\sum_{\mu\neq k}^{d}|2b_{k\mu}|^p\int_{\omega_{ij}^{(k,\mu)}}|x_k + x_\mu - b_j|^p\,\mathrm{d}x \\ &\le d^{p-1}\sum_{j=1}^{N_2}\sum_{\mu\neq k}^{d}|2b_{k\mu}|^p\Big(\frac{\sqrt{2}}{m^{d-1}}\int_{b_j}^{b_j+\frac{\sqrt{2}}{mN_2}}|X - b_j|^p\,\mathrm{d}X\Big) \\ &= d^{p-1}\sum_{j=1}^{N_2}\sum_{\mu\neq k}^{d}|2b_{k\mu}|^p\frac{\sqrt{2}}{m^{d-1}}\frac{1}{p+1}\Big(\frac{\sqrt{2}}{mN_2}\Big)^{p+1} \\ &= \frac{d^{p-1}2^{\frac{3p}{2}+1}}{p+1}\sum_{\mu\neq k}^{d}\frac{|b_{k\mu}|^p}{m^d}\Big(\frac{1}{mN_2}\Big)^p, \end{aligned} \tag{4.52}$$

by virtue of a change of variable $X = x_\nu + x_\mu$, $Y = x_\nu - x_\mu$ where $b_j \le X \le b_j + \frac{\sqrt{2}}{mN_2}$ and the range of $Y$ not more that $\frac{\sqrt{2}}{m}$. From (4.51) and (4.52), together with (4.44) and (4.47), we find that

$$|q_i - s_i|_{W_p^1(\omega_i)}^p \le \Big(\frac{2}{mN_2}\Big)^p\Big(\frac{d^{p-1}2^{\frac{3p}{2}} + 2^{2p}}{p+1}\Big)\Big(\Big(\frac{\sqrt{d}\rho_{d,3}}{m}\Big)^p|f|_{W_p^3(\omega_i)}^p + |f|_{W_p^2(\omega_i)}^p\Big). \tag{4.53}$$

It is easy to show that $m^{-2} \le (mN_2)^{-1} \le 2\binom{d+1}{2}^{3/(2d+1)}|\mathcal{P}_m|^{-3/(2d+1)}$: The first inequality is obvious since $N_2 = \lfloor m^{\frac{1}{2}}\rfloor \le m$. Also, since $m^{\frac{1}{2}} \le \lceil m^{\frac{1}{2}}\rceil \le N_2 + 1 \le 2N_2$ we have that $m^{\frac{3}{2}} \le 2mN_2$. Combining this with the fact that $|\mathcal{P}_m|^{\frac{2}{2d+1}} \le \binom{d+1}{2}^{2/(2d+1)}m$ yields $|\mathcal{P}_m|^{\frac{3}{2d+1}} \le 2\binom{d+1}{2}^{3/(2d+1)}mN_2$.

Now combining (4.41) and (4.53), together with the Hölder inequality, we obtain

$$
\begin{aligned}
|f-s|_{W_p^1(\Omega)} \leq & \Bigg(2^{p-1}\sum_{i=1}^{N_1}\Big(|f-q_i|^p_{W_p^1(\omega_i)}+|q_i-s_i|^p_{W_p^1(\omega_i)}\Big)\Bigg)^{\frac{1}{p}} \\
\leq & C_2|\mathcal{P}_m|^{-3/(2d+1)}\Big(|f|_{W_p^3(\Omega)}+|f|_{W_p^2(\Omega)}\Big), \qquad (4.54)
\end{aligned}
$$

where $C_2 = 4\binom{d+1}{2}^{3/(2d+1)}\Big(d^2\rho_{d,3}+32d\Big(\sqrt{d}\rho_{d,3}+1\Big)\Big)$, thereby proving (4.39). □

The choice of $N_2$ in Theorem 4.5.2 is justified from the following argument: In order to estimate (4.50), we want that $\frac{1}{m^2N_2^2} \leq C\frac{1}{m^3}$ for some constant $C$, that is $m \leq CN_2^2$. If $N_2 = \lfloor m^{\frac{1}{k}}\rfloor$ for some $k \geq 1$, then $m \leq (N_2+1)^k \leq 2^kN_2^k$. The clearly obvious choice is $k=2$.

# Chapter 5

# Conclusion

The objective of this thesis was to develop new partitioning methods for the approximation of a function $f$ on a domain $\Omega \subset \mathbb{R}^d$, $d \geq 2$, by piecewise linear functions. By using our partitions, we estimate the approximation error in both $L_p$-norm and $W_p^1$-seminorm. In the two-dimensional case, we design conforming triangulations so that the approximant is continuous, whereas in the general multidimensional case, we do not impose continuity on the approximant.

In the first part, we start by investigating local errors resulting from the interpolation of a quadratic polynomial by a linear polynomial on a reference triangle $\hat{T}$. We show that, if the measure of non-degeneracy of the triangle $T$ on which we approximate the quadratic polynomial is bounded, then we can estimate the derivatives without the maximum angle condition necessarily met. We also discuss how an optimal triangle for a quadratic polynomial can be obtained. In the case where the determinant of the quadratic polynomial is positive, the optimal triangle is obtained by mapping an equilateral triangle by a linear map associated with the spectrum of the matrix associated with the quadratic polynomial. We provide a discussion on the characterization of optimal triangles for quadratic polynomials with a negative determinant, which still remains an open problem.

We carry on by studying the local errors from the interpolation of a twice differentiable function $f$ on triangles. We use a quadratic polynomial associated with the Hessian matrix $H_f$ as intermediate approximation. With the help of the shape function $K_p$, we are able to provide sharp error estimates when approximating on nearly optimal triangles which are scaled and shifted versions of optimal triangles. On a non-optimal triangle $T$ where we have no knowledge of its interior angles, it suffices to use a linear map $\varphi$ whose conditional number is bounded, and such that interior angles of the inverse image triangle $\varphi^{-1}(T)$ are far from the flat angle. If the measure of non-degeneracy of the triangle is bounded, then another alternative is to use again an intermediate quadratic polynomial.

The second part of this dissertation is devoted to the construction of a sequence of anisotropic triangulations $(\Delta_N)_{N \geq N_0}$, where an $W_p^1$-seminorm estimation is derived while maintaining the optimality of the asymptotic $L_p$-norm estimation. Our construction method shares some basic features with the constructions described in [2, 3, 29], namely using the Hessian $H_f$ for the initial step where we design the regular regions. The other step in our construction, on which lies the originality of this work, consists in obtaining the irregular regions by extending the segments which define the regular regions. Each extension is described by specific maneuvers in such a way that local error estimations in $L_p$-norm and $W_p^1$-seminorm can be derived from the approximations on the irregular triangles that are generated by the irregular regions.

In general, describing the shapes of the irregular triangles is problematic. However, using a "back transformation", we manage to show that the interior angles of the back transformed triangles are far from the flat angle. We show that these triangles cover only small parts of the domain $\Omega$, and that the error that they contribute to is negligible, as compared to the error coming from the regular triangles that are generated by the regular regions. We derive our asymptotic estimations in $L_p$-norm and $W_p^1$-seminorm by combining the local errors on all triangles.

In the third and final part, we use several overlaying partitions $\mathcal{P} = \{\Delta_1, \ldots, \Delta_d\}$ to approximate the target function. The splitting directions of the partitions are either fixed, or related to the properties of the Hessian. Also, each partition contributes to the design of the approximant which is discontinuous, and consists of a sum of piecewise linear functions. By using the Poincaré inequality and the

Bramble-Hilbert lemma for convex domains, our local error analysis addresses the best approximation problem on each cell of the partitions. We obtain error bounds in $L_p$-norm and $W_p^1$-seminorm, where our approximation orders improve on the ones obtained when using a single partition. Namely, we achieve the approximation order $O(N^{-6/(2d+1)})$ for the $L_p$-norm estimation, and the approximation order $O(N^{-3/(2d+1)})$ for the $W_p^1$-seminorm estimation, with $N$ being the number of degrees of freedom.

## Future work

We want to extend the results of Chapter 3 to functions which are not necessarily convex, and also to investigate whether our triangulations can be used for $p = \infty$. More properties of our triangulations are still to be found, for instance checking whether the triangles are regular (or quasi-uniform) with respect to certain metrics.

Another open question is to improve our $W_p^1$-seminorm estimation to the optimal result obtained in [30] and, moreover, construct triangulations where the piecewise linear interpolant is asymptotically optimal in $\limsup$ sense for both $L_p$-norm and $W_p^1$-seminorm errors.

Finally, we want to extend our results on sums of piecewise polynomials to higher order approximation. Also, we want to provide lower bounds in order to show that the obtained approximation orders cannot be improved when using any sums of piecewise polynomials on convex domains.

# Bibliography

[1] T. Apel. Anisotropic finite elements: Local estimates and applications. preprint SFB393/99-03, SFB 393, Technische Universität Chemnitz, 1999. 2, 6, 7, 12, 24, 36

[2] V. Babenko, Yu. Babenko, A. Ligun, and A. Shumeiko. On asymptotical behavior of the optimal linear spline interpolation error of $C^2$ functions. *East Journal on Approximations*, 12(1):71–101, 2006. 2, 3, 9, 10, 14, 40, 45, 56, 58, 64, 78, 79, 80, 81, 85, 86, 211

[3] V. Babenko, Yu. Babenko, and D. Skorokhodov. Exact asymptotics of the optimal $L_{p,\Omega}$-error of linear spline interpolation. *East Journal on Approximations*, 14(3):285–317, 2008. 2, 3, 9, 10, 14, 43, 45, 46, 56, 64, 78, 79, 80, 81, 83, 85, 86, 211

[4] I. Babuska and A.K. Aziz. On the angle condition in the finite element method. *SIAM Journal on Numerical Analysis*, 13(2):214–226, 1976. 6

[5] M. Bern and D. Eppstein. Mesh generation and optimal triangulation. *Computing in Euclidean Geometry*, pages 23–90, 1992. World Scientific, New York. 79

[6] K. Böröczky. Approximation of general smooth convex bodies. *Adv. in Math.*, 153:325–341, 2000. 79

[7] K. Böröczky and M. Ludwig. Approximation of convex bodies and a momementum lemma for power diagrams. *Monatshefte für Mathematik*, 127(2):101–110, 1999. 79

[8] J. Brandts, A. Hannukainen, S. Korotov, and M. Kří žek. On angle conditions in the finite element method. *SeMA Journal*, 56:81–95, 2011. 7

[9] S.C. Brenner and L.R. Scott. *The Mathematical Theory of Finite Element Methods*, volume 15 of *Texts in Applied Mathematics*. Springer, 2008. 2

[10] W. Cao. On the error of linear interpolation and the orientation, aspect ratio, and internal angles of a triangle. *Siam J. Numer. Anal.*, 43(1):19–40, 2005. 25

[11] W. Cao. An interpolation error estimate on anisotropic meshes in $\mathbb{R}^n$ and optimal metrics for mesh refinement. *SIAM J. Numer. Anal.*, 45(6):2368–2391, 2007. 80

[12] L. Chen, P. Sun, and J. Xu. Optimal anisotropic meshes for minimizing interpolation errors in $L^p$-norm. *Mathematics of Computation*, 76(257):179–204, 2007. 77, 80

[13] P.G. Ciarlet. *The finite element method for elliptic problems*. Classics in Applied Mathematics. Society for Industrial and Applied mathematics (SIAM), Philadelphia, 2002. 2, 5, 6, 23, 36

[14] A. Cohen, N. Dyn, F.Hecht, and J.-M. Mirebeau. Adaptive multiresolution analysis based on anisotropic triangulations. *Mathematics of Computation*, 81(278):789–810, 2012. 33, 79, 80

[15] O. Davydov. *Algorithms and error bounds for multivariate piecewise constant approximation*, volume 3, pages 27–45. Springer-Verlag, 2011. *in* “Approximation Algorithms for Complex Systems,” by E. H. Georgoulis, A. Iske and J. Levesley, Eds. 185, 189

[16] O. Davydov. Approximation by piecewise constants on convex partitions. *J. Approx. Theory*, 164:346–352, 2012. 3, 11, 13, 77, 185, 189, 190, 191

[17] E.F. D’Azevedo. Optimal triangular mesh generation by coordinate transformation. *SIAM J. Sci. Statist. Comput.*, 6:755–786, 1991. 6

[18] E.F. D’Azevedo. On adaptive mesh generation in two-dimensions. In *Proceedings, 8th International Meshing Roundtable*, pages 109–117, 1999. South Lake Tahoe, CA, U.S.A. 6

[19] E.F. D'Azevedo and R.B. Simpson. On optimal triangular meshes for minimizing the gradient error. *Numerische Mathematik*, 59:321–348, 1991. 6

[20] A.S. Deif. Rigorous perturbation bounds for eigenvalues and eigenvectors of a matrix. *Journal of Computational and Applied Mathematics*, 57:403–412, 1995. 87

[21] S. Dekel and D. Leviatan. The Bramble-Hilbert Lemma for convex domain. *Siam J. Math. Anal.*, 35:1203–1212, 2004. 11, 188

[22] A. Ern and J.-L. Guermond. *Theory and Practice of Finite Elements*, volume 159 of *Applied Mathematical Sciences*. Springer, 2004. 2

[23] L. Formaggia and S. Perotto. New anisotropic a priori error estimates. *Numerische Mathematik*, 89:641–667, 2001. 2, 6, 25, 29

[24] J.A. Gregory. *Error bounds for linear interpolation on triangles.* in Mathematics of Finite Elements and Applications II. Academic Press, London, 1976. 163-170. 6, 7

[25] P. Gruber. Volume approximation of convex bodies by inscribed polytopes. *Mathematische Annalen*, 281(2):229–245, 1988. 79

[26] A.S. Kochurov. Approximation by piecewise constant functions on the square. *East J. Approx*, (1):463–478, 1995. 191

[27] G. Kunert. Towards anisotropic mesh construction and error estimation in the finite element method. *Numerical methods for partial differential equations*, 18:625–648, 2002. 56

[28] M.J. Lai and L.L. Schumaker. *Spline Functions on triangulations.* Encyclopedia of Mathematics. Cambridge University Press, New York, 2007. 37

[29] J.-M. Mirebeau. Optimal meshes for finite elements of arbitrary order. *Constr. Approx.*, 32(2):339–383, 2010. 2, 3, 6, 9, 10, 14, 15, 21, 40, 41, 42, 54, 57, 58, 59, 60, 61, 64, 78, 80, 81, 82, 83, 85, 86, 160, 175, 211

[30] J.-M. Mirebeau. Optimally adapted meshes for finite elements of arbitrary order and $W^{1,p}$ norms. *Numerische Mathematik*, 120(12):271–305, 2012. 2, 3, 9, 10, 79, 80, 85, 86, 212

[31] J.-M. Mirebeau and A. Cohen. Greedy bisection generates optimally adapted triangulations. *Mathematics of Computation*, 81(278):811–837, 2012. 5, 18, 32, 33, 36, 79, 80, 86

[32] H. Pottmann, R. Krasauskas, B. Hamann, K. Joy, and W. Seibold. On piecewise linear approximation of quadratic functions. *Geometry of Graphics*, 4(1):31–53, 2000. 2, 43, 44, 45, 48, 80

[33] R.B. Simpson. Anisotropic mesh transformations and optimal error control. *Applied Numerical Mathematics*, 14:183–198, 1994. 6

[34] A. H. Stroud and D. Secrest. *Gaussian quadrature formulas*. Englewood Cliffs, Prentice-Hall, 1966. 180

[35] David S. Watkins. *Fundamentals of matrix computations*. John Wiley & Sons, Inc., New York. 1991. 87, 89