The Generic Failure of Lower-semicontinuity for the Linear Distortion Functional

Seyed Mohsen Hashemi

Supervisor

Distinguished Professor Gaven J. Martin

Co-supervisor

Associate Professor Winston L. Sweatman

A Thesis Presented in Partial Fulfilment of the Requirments for the Degree of

Doctor of Philosophy (PhD)
in Mathematics.

Institute of Advanced Studies (NZIAS)
Massey University of Auckland
New Zealand

December 2022

ABSTRACT

My research is primarily concerned with the convexity properties of distortion functionals (particularly the linear distortion) defined for quasiconformal homeomorphisms of domains in Euclidean n-spaces, though we will mainly stick to three-dimensions. The principal application is in studying the lower semi-continuity of distortion on uniformly convergent limits of sequences of quasiconformal mappings. For example, given the curve family or analytic definitions of quasiconformality - discussed in Chapter 3 - it is known that if $\{\mathbf{f_n}\}_{n=1}^{\infty}$ is a sequence of K-quasiconformal mappings (and here K depends on the particular distortion but is the same for every element of the sequence) which converges to a function $\mathbf{f}$, then the limit function is also K-quasiconformal.

Despite a widespread belief that this was also true for the geometric definition of quasiconformality (via the linear distortion $\mathcal{H}(\mathbf{f})$ defined below) Tadeusz Iwaniec gave a specific surprising example to show that the linear distortion function is not lower semicontinuous. The main aim of this thesis is to show that this failure of lower semicontinuity is actually far more common, perhaps even generic in the sense that it is true that under mild restrictions on a quasiconformal $\mathbf{f}$, there may be a sequence $\{\mathbf{f_n}\}_{n=1}^{\infty}$ with $\mathbf{f_n} \to \mathbf{f}$ uniformly and with $\limsup_{n\to\infty} \mathcal{H}(\mathbf{f_n}) < \mathcal{H}(\mathbf{f})$. The main result of this thesis is to show this is true for a wide class of linear mappings and give bounds for the maximal jump down in the limit.

Acknowledgements

First and foremost I offer my sincerest gratitude and respect to my supervisor, **Distinguished Professor Gaven J. Martin** and my co-supervisor **Associate Professor Winston Sweatman** who have supported me throughout my education in PhD. They are outstanding mathematicians and working with them is so nice and enjoyable. It is my honour that I have been one of their students.

The greatest measure of appreciation goes to my dear wife **Leila** and my beautiful daughter **Haya** for their enduring support and patience. If I am in this position, it is because of my wife's unsparing support. Furthermore, I would like to thank and acknowledge my parents for everything that they have given me in my life. They are the best and I have been fortunate enough to have such deserving parents. Finally, I am grateful to the friendly staff of the New Zealand Institute for Advanced Study (NZIAS) and the School of Natural and Computational Sciences (SNCS) at the Massey University of Albany for their help and kindness.

Contents

List of Figures

List of Tables

Nomenclature

$B^n(x,r)$	The ball with centre x and radius r	10
$m^*(S)$	Lebesgue outer measure	6
C^∞	Class on functions that they have continuous derivatives of all orders.	12
$\Delta\mathbf{f}$	The incremental change of $\mathbf{f}$	13
$\ell(T)$	The minimal stretching of the map T	19
$\ell(\gamma)$	The length of a path γ	51
$\ell_{\mathbf{f}}(\mathbf{x},r)$	The minimal stretching of the function $\mathbf{f}$ at $\mathbf{x}$	23
$\ell_\rho(\gamma)$	ρ-length of a path γ	53
Γ	Curve family	53
γ^o	Parametrisation of γ	52
$\mathbb{S}^{n-1}$	The unit sphere in $\mathbb{R}^n$	23
$\mathcal{K}_{\alpha,\beta}(A)$	Sectional distortions of A	70
$\mathfrak{M}$	σ-algebra on a set X	2
$\operatorname{diam}(A)$	The diameter of set A in $\mathbb{R}^n$	10
$\operatorname{dist}(A,B)$	The distance between the sets A and B in $\mathbb{R}^n$	10
$\operatorname{dom}\mathbf{f}$	Domain of $\mathbf{f}$	37
$\operatorname{epi}\mathbf{f}$	Epigraph of $\mathbf{f}$	37
$\operatorname{sub}\mathbf{f}$	Subgraph of $\mathbf{f}$	37
μ	Measure function	3
$\overline{A}$	Conjugate of matrix A with complex entries	17
∂E	The boundary of set E	37
$\upsilon_\rho(A)$	the ρ-volume of a Lebesgue measurable subset A of Ω	53
A^*	Conjugate transpose of matrix A with complex entries	17

A^c	The complement of A relative to set X	2
C^k	Class on functions that they have continuous derivatives up to (and including) k order	12
coE	The convex hull of set E	37
$d\mathbf{f}(\mathbf{a}, \mathbf{h})$	Total differential of **f**	13
$intE$	The interior of set E	37
$J_{\mathbf{f}}(x)$	Jacobian determinant	10
$K(\mathbf{f})$	Maximal distortion.	59
$K_I(\mathbf{f})$	Inner distortion.	59
$K_O(\mathbf{f})$	Outer distortion.	59
$L(T)$	The maximal stretching of the map T	19
$L^p(X)$	The set of Lebesgue measurable functions f that pth power of absolute value of f is integrable in X for $p \geq 1$	27
L_α	Lower-level set of height α of **f**	37
$L_{\mathbf{f}}(\mathbf{x}, r)$	The maximal stretching of the function **f** at **x**	23
$L^p_{loc}(\Omega)$	The set of locally Lebesgue integrable functions f in Ω for $p \geq 1$	27
$M_p(\Gamma)$	Modulus of curve family Γ	54
$M_{m\times n}(\mathbb{R})$	Space of all $m \times n$ matrices over the field of real number $\mathbb{R}$	40
$s_\gamma(t)$	the length function of γ	52
U_α	Upper-level set of height α of **f**	38
$\bar{B}^n(\mathbf{x}, r)$	The closure of the ball B^n with centre **x** and radius r.	57
$\lVert f \rVert_p$	L^p-norm of function f	26
$\Delta(E, F : G)$	The family of all open paths with endpoints that connect E and G through G.	55
$\langle \mathbf{v}, \mathbf{u} \rangle$	Euclidean inner product of two vectors **v** and **u**	16
$\mathbb{S}^{n-1}(\mathbf{x}, r)$	An n-dimensional sphere with centre **x** and radius r.	57
$\mathrm{Adm}(\Gamma)$	The collection of all admissible densities	53
ω_{n-1}	The surface area of an n-dimensional sphere with centre **x** and radius r.	57
$\partial B^n(\mathbf{x}, r)$	The boundary of an n-dimensional ball with centre **x** and radius r.	57
ϕ	Empty set	2

$\mathrm{supp}(\mathbf{f})$	The closure of set of all $\mathbf{x}$ in domain that $\mathbf{f}(\mathbf{x}) \neq 0$ with respect to domain.	29
τ	A toplogy on a set for example X	2
$A = \mathrm{Sing}(1, a, b)$	The diagonal matrix A with distinct ordered singular values 1, a, and b	123
$ACL(\Omega, \mathbb{R}^m)$	Absolutely continuous on lines	66
$B^n(\mathbf{x}, r)$	An open n-dimensional ball with centre $\mathbf{x}$ and radius r.	57
$C^0(\Omega, \mathbb{R}^m) = C^k(\Omega)$	The space of all functions $\mathbf{f}$ that are continuous in Ω.	29
$C^k(\Omega, \mathbb{R}^m) = C^0(\Omega)$	The space of all functions $\mathbf{f}$ that are k times continuously differentiable in Ω.	29
$C^\infty(\Omega, \mathbb{R}^m) = C^\infty(\Omega)$	The space of all smooth functions $\mathbf{f}$ (infinitely differentiable real-valued functions).	29
$C_0^\infty(\Omega, \mathbb{R}^m) = C_0^\infty(\Omega)$	The space of all compactly supported smooth functions $\mathbf{f}$.	29
$GL(n)$	General linear group	20
$O(n)$	Orthogonal group	20
V_n	The volume of an n-dimensional sphere with centre $\mathbf{x}$ and radius r.	57
$W^{k,p}(\Omega)$	The Sobolev space.	32
$W_0^{k,p}(\Omega)$	The space of functions $\mathbf{f} \in W^{k,p}(\Omega)$ that $\mathbf{f}$ is zero on $\partial\Omega$.	32
$W_{loc}^{k,p}(\Omega)$	The local Sobolev space.	32
$\mathit{aff}\ E$	The affine hull of set E	37

Chapter 0

Introduction

0.1 Introduction

The principal objects of study in this thesis are quasiconformal mappings, and in particular, the convexity properties of linear distortion functionals are investigated. Quasiconformal mappings are generalizations of conformal mappings, the defining feature is that they are geometric homeomorphisms of "bounded distortion". The notion of distortion can be quantified in various ways and we will see some of them here, indeed the different distortions and their properties are of fundamental interest to us.

Quasiconformal mappings can be considered not only on domains in the complex plane and higher dimensional Euclidean spaces, but also on Riemann surfaces and on Riemannian manifolds in all dimensions, and even on arbitrary metric spaces. However, although our results have consequences in this framework, we will not discuss these generalisations here. The importance of quasiconformal mappings in complex analysis was realized by Ahlfors and Teichmüller in the 1930s. Ahlfors used quasiconformal mappings to Nevanlinna's value distribution theory in his geometric approach. He chose the term "quasiconformal" in his work in 1935. Teichmüller used quasiconformal mappings to find and measure a distance between two conformally inequivalent compact Riemann surfaces. For more details and historical progress, you can see "What is a quasiconformal mapping?" [35] or [42].

Quasiconformal mappings are generalizations of conformal mappings. There are three definitions for quasiconformal mappings in Euclidean spaces: metric, geometric and analytic. We can start with the metric definition, which is the easiest way to see and which makes sense in arbitrary metric spaces. It explains the property that infinitesimal balls are transformed to infinitesimal ellipsoids of bounded eccentricity (see [35] or [42]).

Let Ω be a domain in $\mathbb{R}^n$, let $\mathbf{f} : \Omega \longrightarrow \mathbb{R}^n$ be a continuous injection. The **linear distortion or dilatation** of $\mathbf{f}$ at the point x in A is the quantity $\mathcal{H}_{\mathbf{f}}(\mathbf{x})$ or $\mathcal{H}(\mathbf{x}, \mathbf{f})$ defined by

$$\mathcal{H}_{\mathbf{f}}(\mathbf{x}) = \mathcal{H}(\mathbf{x}, \mathbf{f}) = \limsup_{r \to 0} \frac{L_f(\mathbf{x}, r)}{\ell_f(\mathbf{x}, r)}, \tag{1}$$

where for $0 < r < \operatorname{dist}(\mathbf{x}, \partial\Omega)$ we set

$$L_f(\mathbf{x}, r) = \max_{\|h\|=r} \|\mathbf{f}(\mathbf{x}+h) - \mathbf{f}(\mathbf{x})\|, \qquad \ell_f(\mathbf{x}, r) = \min_{\|h\|=r} \|\mathbf{f}(\mathbf{x}+h) - \mathbf{f}(\mathbf{x})\|.$$

In other words,

$$\mathcal{H}_{\mathbf{f}}(\mathbf{x}) = \mathcal{H}(\mathbf{x}, \mathbf{f}) = \limsup_{r\to 0} \frac{\max_{\|h\|=r} \|\mathbf{f}(\mathbf{x}+h) - \mathbf{f}(\mathbf{x})\|}{\min_{\|h\|=r} \|\mathbf{f}(\mathbf{x}+h) - \mathbf{f}(\mathbf{x})\|}. \tag{2}$$

Let Ω and Ω' be domains in $\mathbb{R}^n$ with $n \geq 2$, and let $\mathbf{f}$ be a homeomorphism of Ω onto Ω'. The **linear distortion of function f** is defined

$$\mathcal{H}(\mathbf{f}) = \sup\{\mathcal{H}_{\mathbf{f}}(\mathbf{x}) \, : \, \mathbf{x} \in \Omega\}, \tag{3}$$

where $\mathcal{H}_{\mathbf{f}}(\mathbf{x}) \geq 1$ is the linear distortion of $\mathbf{f}$ at $\mathbf{x}$ as defined at (1). The homeomorphism $\mathbf{f}$ is a conformal mapping if and only if $\mathcal{H}(\mathbf{f}) = 1$ (see [30] p.77). A homeomorphism $\mathbf{f} : \Omega \longrightarrow \mathbb{R}^n$ is quasiconformal mappings if $\mathcal{H}(\mathbf{f})$ is bounded. The central problem discussed in this thesis concerns the semicontinuity properties of distortion functionals for quasiconformal mappings. In particular, if $\{\mathbf{f}_\nu\}_{\nu=1}^{\infty}$ is a sequence of quasiconformal mappings of $\hat{\mathbb{R}}^n = \mathbb{R}^n \bigcup \{\infty\}$ onto itself and that

$$\mathbf{f}_\nu \longrightarrow \mathbf{f}, \tag{4}$$

uniformly on $\hat{\mathbb{R}}^n$, in the spherical metric

$$d(x, y) = \frac{|\mathbf{x} - \mathbf{y}|}{\sqrt{1 + |\mathbf{x}|^2}\sqrt{1 + |\mathbf{y}|^2}}.$$

We ask for which distortion functionals $\mathcal{K} = \mathcal{K}(\mathbf{g})$, defined for a quasiconformal mapping $\mathbf{g}$ of $\hat{\mathbb{R}}^n$ do we have the lower semicontinuity property

$$\mathcal{K}(\mathbf{f}) \leq \liminf_{\nu\to\infty} \mathcal{K}(\mathbf{f}_\nu), \tag{5}$$

so that the limit $\mathbf{f}$ is itself a $\mathcal{K}$-quasiconformal mapping of $\hat{\mathbb{R}}^n$ onto itself. For instance, $\mathcal{K}$ could be the linear distortion, maximal distortion, outer and inner distortion defined below etc. This lower semicontinuity property is related to issues of convexity of the functional $\mathcal{K}$ defined on the space of mappings or more precisely the pointwise differentials of these mappings. We will provide references which show that the functionals determined by the distortion of modulus of curve families and also those determined by the pointwise differential inequality do in fact have the lower semicontinuity property given at (5). For the linear distortion the question of lower semicontinuity has been answered negatively by Tadeusz Iwaniec (see [24]). In his paper [39], he proved that there is a sequence of quasiconformal mappings as at (4) such that for the linear distortion

$$\mathcal{H}(\mathbf{f}) > \lim_{\nu\to\infty} \mathcal{H}(\mathbf{f}_\nu). \tag{6}$$

Notice that the limit here does exist and the limit mapping $\mathbf{f}$ above must be quasiconformal (and so $\mathcal{H}(\mathbf{f}) < \infty$) since the other two functional distortions are controlled by the linear distortion functional. Since the very beginning of the multidimensional quasiconformal mappings theory, it has been widely believed that the class of K-quasiconformal mappings in $\mathbb{R}^n$ is closed with respect to uniform convergence, where K stands for the linear distortion. The key element of this thesis is the linear distortion function fails to be rank-one convex in dimensions higher than 2 (see [39]).

Lemma 0.1.1. (Iwaniec's Lemma) Given $n \geq 3$ and $H > 1$, there is a matrix A and a rank-one matrix B and numbers $t, s > 0$ such that

$$\mathcal{H}(A - sB) = \mathcal{H}(A + tB) = H < \mathcal{H}(A). \tag{7}$$

The reason for this unusual and anomalous behaviour of the linear distortion function is that it fails to be rank-one convex in dimensions higher than 2. Recall that the determinant function det : $\mathbb{R}^{n\times n} \longrightarrow \mathbb{R}$, in spite of the non-linearity of this polynomial of n^2 variables, is, in fact, linear in the directions of rank-one matrices. More precisely, this means that the function of the real variable $t \mapsto \det(A + tB)$ is linear if $\operatorname{rank}(B) \leq 1$. The same is true for lower-order minors and consequently for null-Lagrangians, being linear combinations of the minors of $D\mathbf{f}$. We know that if B is a rank-one matrix then it can be written as the tensor product of two vectors (see [42], page 197). The previous lemma and the next theorem are existence theorems and Iwaniec found a striking counterexample with an amazing calculation that refutes the lower semicontinuity property. Iwaniec's example is for dimension 3 and he extended his example to higher dimensions (see [42] and [30]).

Theorem 0.1.2. (Iwaniec's Theorem) For each $n \geq 3$ and $H > 1$, there exists a sequence $\{\mathbf{f}_\nu\}_{\nu=1}^{\infty}$ of quasiconformal mappings $\mathbf{f}_\nu : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ converging uniformly to a linear quasiconformal map $\mathbf{f}_0 : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ such that

$$\mathcal{H}(\mathbf{x}, \mathbf{f}_\nu) \equiv H < \mathcal{H}(\mathbf{x}, \mathbf{f}_0), \qquad \text{almost everywhere in } \mathbb{R}^n \quad \nu = 1, 2, \dots$$

These results show that there exists at least one linear mapping T and a sequence $T_\nu \longrightarrow T$ such that the linear distortion is not lower semicontinuous. The key idea in Iwaniec's work is that the linear distortion function fails to be rank-one convex in dimension $n \geq 3$. The linear distortion functional is defined pointwise from the differential matrix when it is nonsingular. At the points $x_0 \in \mathbb{R}^3$ of differentiability of a quasiconformal mapping a local analysis of the convexity properties of the linear distortion can be given if we only consider the nonsingular linear mappings $x \mapsto Tx$, $T = D\mathbf{f}(x_0) \in GL(3, \mathbb{R})$. To compute the linear distortion we study the eigenvalues of

$$A = T^t T \in Sym^+_{3\times 3}(\mathbb{R}),$$

the space of symmetric positive definite 3×3 matrices. Given such an A the spectral theorem tells us A is orthogonally diagonalisable and so we may as well suppose it is diagonal since we can diagonalise it by orthogonal (conformal) transformations (at a point). In this way, we reduce the problem of the convexity of the linear distortion functional to considering that functional defined on the space of 3×3 positive definite diagonal matrices $A = (a_{ij})$, $a_{ij} = 0$, if $i \neq j$. All the diagonal entries a_{ii} are positive and we can further suppose that $1 = a_{11} \leq a_{22} \leq a_{33}$, by scaling and a further conjugation by orthogonal matrices - neither of which affects the linear distortion.

We show that it is always possible to find a rank-one matrix B so that (7) holds for diagonal matrices A with distinct positive eigenvalues. An interesting question - possibly connected with some aspects of materials science - is to determine the "rank-one direction" for which the linear distortion function at A is **most** concave. This might identify the structure of the laminations for the minimisers of certain stored energy functionals occurring in the calculus of variations. An unanswered question, (partly posed by Gehring and Iwaniec) is to determine how big the difference between the left-hand side and the right-hand side of (7) can be? In Chapter 4, we answer this question, in the linear and general cases.

In the first chapter, we recall some background in analysis, linear algebra, topology and Sobolev space that we use in this thesis. The second chapter presents three different notions of convexity; named polyconvexity, quasiconvexity and rank-one convexity and their properties and relationships with each other. In Chapter 3, we define the linear distortion and three definitions of quasiconformal mappings and their relations. The main new results of my research can be found in Chapter 4. Finally, in the last chapter we describe our conclusion and potential applications in material science. In Chapter 4, we try to answer three problems which result in four new theorems. This means that not only do we extend Iwaniec's Theorem to all matrices A with distinct singular values, but also that we can find the maximum jump

$$\mathcal{H}(A) - \mathcal{H}(A_j),$$

with $A_j \longrightarrow A$ unifomly on compact sets.

The main foci of the PhD thesis are the following three problems:

Problem 1. Let

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{bmatrix}$$

be diagonal with $1 < a < b$. Determine vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^3$ with $\|\mathbf{u}\| = \|\mathbf{v}\| = 1$ so that with $B_0 = \mathbf{u} \otimes \mathbf{v}$ we have

$$\frac{d}{dt}\Big|_{t=0} \mathcal{H}(A + t\,B_0) \;=\; 0, \tag{8}$$

$$\frac{d^2}{dt^2}\Big|_{t=0} \mathcal{H}(A + t\,B_0) \;\leq\; \frac{d^2}{dt^2}\Big|_{t=0} \mathcal{H}(A + t\,B), \tag{9}$$

for every rank one matrix $B \in \mathbb{R}^{3\times 3}$ with $\frac{d}{dt}\Big|_{t=0} \mathcal{H}(A + t\,B) = 0$.

Thus problem 1 asks us to find the “best” rank-one direction - that which maximises the negative of the second derivative. With regard to this problem we might also ask if this direction is unique (up to sign). In this direction we expect to find the minimum values of $\mathcal{H}(A + tB)$.

The second problem that we will discuss next, is finding an interval for variable t for which $\mathcal{H}(A+tB) < \mathcal{H}(A)$. There are conjectures by Gehring and Iwaniec concerning the gap between the linear distortion and the outer distortion with regard to limiting processes, basically trying to estimate the maximum size of the jump showing the failure of lower semi-continuity. The solution to problem 2 below should go some way to providing concrete examples of what can happen and possibly resolve their conjectures, or offer more insight and potentially new conjectures.

Problem 2. Let

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{bmatrix}$$

be diagonal with $1 < a < b$ and suppose $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^3$ have $\|\mathbf{u}\| = \|\mathbf{v}\| = 1$ so that $B_0 = \mathbf{u} \otimes \mathbf{v}^t$ is a solution to Problem 1.

1. Determine the largest magnitudes of real numbers $\mathbf{t}_+ > 0$ and $\mathbf{t}_- < 0$ so that $\mathcal{H}(A + t\, B_0)$ is a smooth function of t in the interval $\mathbf{t}_- < t < \mathbf{t}_+$.
2. Determine $\mathcal{H}(A + \mathbf{t}_- B_0)$ and $\mathcal{H}(A + \mathbf{t}_+ B_0)$.

We expect that the values $\mathbf{t}_-$ and $\mathbf{t}_+$ will be where the singular values of $A + tB_0$ cross. These might be determined from a (rather challenging) discriminant problem. We also conjecture:
For all rank-one matrices B with

$$\frac{\mathrm{d}}{\mathrm{dt}}\Big|_{t=0} \mathcal{H}(A + t\, B) = 0, \quad \text{and} \quad \frac{\mathrm{d}^2}{\mathrm{dt}^2}\Big|_{t=0} \mathcal{H}(A + t\, B) < 0,$$

we have for all $t > 0$

$$\mathcal{H}(A + tB) \geq \max\{\mathcal{H}(A + \mathbf{t}_- B_0), \mathcal{H}(A + \mathbf{t}_+ B_0)\}.$$

This conjecture expresses the hope that the best direction also leads to the biggest gap between $\mathcal{H}(A)$ and $\mathcal{H}(A + t\, B)$ and therefore gives us the approximation to A of least linear distortion.

As can be seen in Problem 2, on the interval $(\mathbf{t}_-,\ \mathbf{t}_+)$, the value of $\mathcal{H}(A + tB_0)$ is less than $\mathcal{H}(A)$ and $\mathcal{H}(A + tB)$, where B is a rank-one matrix with the above conditions. Finally problem 3 seeks to put the previous two problems in a more general framework.

Problem 3. Let $C \in GL(3, \mathbb{R})$ and suppose that C has distinct singular values (so $C^t C$ has distinct real eigenvalues). Then

1. Determine the best rank-one direction B^* (as per Problem 1) for C.
2. Determine $\mathbf{t}_-, \mathbf{t}_+$ as per Problem 2, and also $\mathcal{H}(C + \mathbf{t}_- B^*)$ and $\mathcal{H}(C + \mathbf{t}_+ B^*)$, that produce the largest interval of concavity for $\mathcal{H}(C + tB^*)$ and the smallest values for the linear distortion of an approximant.

Finally, we discuss new models for static nonlinear deformations via scale-invariant conformal energy functionals based on linear distortion. In particular, we give examples to show that, despite equicontinuity estimates giving compactness, minimising sequences will have strictly lower energy than their limit, and that this energy gap can be quite large.

Chapter 1

Topology, Linear Algebra, Analysis and Linear Distortion

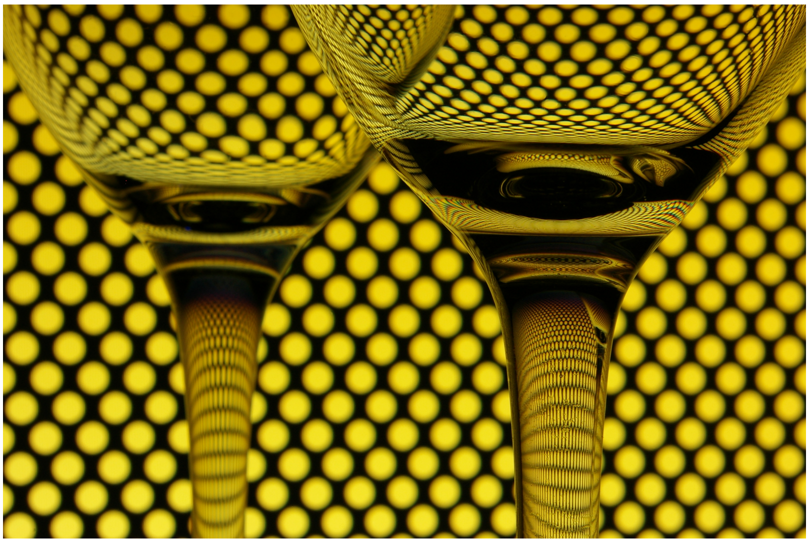

Figure 1.1: Distortion in Optics
Source: Images for linear distortion photography

Chapter 1

Linear Algebra

This chapter presents an introduction to linear algebra and some of its applications in geometry. It includes ten short sections that we need for the following chapters. In the first section we recall Euclidean n-space, some topology background and mathematical spaces.

1.1 Euclidean n-Space and Topology Background

This section is about mathematical structures called norms and spaces.

Definition 1.1.1. A collection τ of subsets of a set X is said be a **topology** in X if τ has the following properties:

1. $\phi \in \tau$ and $X \in \tau$.
2. If $V_i \in \tau$ for $i = 1, 2, \dots, n$, then $\bigcap\limits_{i=1}^{n} V_i \in \tau$.
3. If $\{V_\alpha\}$ is an arbitrary collection of members of τ (finite, countable or uncountable), then $\bigcup\limits_{\alpha} V_\alpha \in \tau$.

Properly speaking, a **topological space** is an ordered pair (X, τ). The members of τ are called the open sets in X. If X and Y are topological spaces and if f is a mapping of X into Y, then f is said to be continuous provided that $f^{-1}(V)$ is an open set in X for every open set V in Y (see [55], [51]).

Definition 1.1.2. A collection $\mathfrak{M}$ of subsets of a set X is said to be a σ-algebra in X if $\mathfrak{M}$ has the following properties (see [64], [4]):

1. $X \in \mathfrak{M}$.
2. If $A \in \mathfrak{M}$, then $A^c \in \mathfrak{M}$ where A^c is the complement of A relative to X.
3. If $A = \bigcup_{n=1}^{\infty} A_n$ and if $A_n \in \mathfrak{M}$ for $n = 1, 2, \dots$, then $A \in \mathfrak{M}$.

If $\mathfrak{M}$ is a σ-algebra in X, then X is called a **measurable space**, and the members of $\mathfrak{M}$ are called the **measurable sets** in X. The prefix σ refers to the fact that the property (3) is required to hold for all countable unions of members of $\mathfrak{M}$. If (3) is required for finite unions only, then $\mathfrak{M}$ is called the **algebra of sets**. If X is a measurable space, Y is a topological space and f is a mapping of X into Y, then f is said to be measurable provided that $f^{-1}(V)$ is a measurable set in X for every open set V in Y. Let X be a set and $\mathfrak{M}$ be a σ-algebra on X. A function μ from $\mathfrak{M}$ to the extended real number line is called the **measure** or **measure function** if it satisfies the following properties:

1. For all E in $\mathfrak{M}$, we have $\mu(E) \geq 0$.
2. $\mu(\phi) = 0$.

3. For all countable collections $\{E_k\}_{k=1}^{\infty}$ of pairwise disjoint sets in $\mathfrak{M}$, we have

$$\mu(\bigcup_{k=1}^{\infty} E_k) = \sum_{k=1}^{\infty} \mu(E_k).$$

Property (3) is called **countably additive**. A triple $(X, \mathfrak{M}, \mu)$ is called a **measure space**. A **positive measure** is a function μ defined on a σ-algebra $\mathfrak{M}$, whose range is in $[0, \infty]$ and which is countably additive (see [64], page 16). A measure space is a measurable space that has a positive measure μ defined on the σ-algebra of its measurable sets. A measure spaces might not be a topological spaces (see [64], [71], [76]). Let $(X, \mathfrak{M}_1)$ and $(Y, \mathfrak{M}_2)$ be two measurable spaces, then the function $f : X \longrightarrow Y$ is said to be a **measurable function** if the preimage of E under f is in $\mathfrak{M}_1$ for every $E \in \mathfrak{M}_2$.

Definition 1.1.3. A **metric space** is a set X in which a distance function or **metric** d is defined, with the following properties:

1. $0 \leq \mathrm{d}(x, y) \leq \infty$ for all x and y in X.
2. $\mathrm{d}(x, y) = 0$ if and only if $x = y$.
3. $\mathrm{d}(x, y) = \mathrm{d}(y, x)$ for all x and y in X.
4. $\mathrm{d}(x, y) \leq \mathrm{d}(x, z) + \mathrm{d}(z, y)$ for all x, y and z in X.

Theorem 1.1.1. If $\mathcal{F}$ is any collection of subsets of X, there exists a smallest σ-algebra $\mathfrak{M}^*$ in X such that $\mathcal{F} \subset \mathfrak{M}^*$.

This $\mathfrak{M}^*$ is sometimes called the σ-algebra generated by $\mathcal{F}$. We do not want to prove this theorem. There exists one in [64] on page 12.

Definition 1.1.4. Let X be a topological space. By theorem (1.1.1), there exists a smallest σ-algebra $\mathcal{B}$ in X such that every open set in X belongs to $\mathcal{B}$. The members of $\mathcal{B}$ are called the **Borel sets of** X [27], [58].

In particular, closed sets are Borel sets, and so are all countable unions of closed sets and all countable intersections of open sets. These last two are called $\mathbf{F}_\sigma$'s and $\mathbf{G}_\delta$'s, respectively, and play a considerable role. Since $\mathcal{B}$ is a σ-algebra, we may now regard X as a measurable space, with the Borel sets playing the role of the measurable sets; more concisely, we consider the measurable space $(X, \mathcal{B})$. If $f : X \longrightarrow Y$ is a continuous mapping of X, where Y is any topological space, then it is evident from the definitions that $f^{-1}(V) \in \mathcal{B}$ for every open set V in Y. In the other hand, every continuous mapping of X is Borel measurable. **Borel measurable mappings** are often called **Borel mappings** or **Borel functions** (see [26], [66] and [27]). We explain some other definitions that are useful for this thesis.

Definition 1.1.5. We need some spaces and their norms that are useful to understand the rest of this thesis.

1. The **closure** $\overline{E}$ of a set $E \subset X$ is the smallest closed set in X which contains E.
2. A set $K \subset X$ is **compact** if every open cover of K contains a finite subcover. In particular, if X is itself compact, then X is called a **compact space**.
3. X is a **Hausdorff space** if the following is true. If $p \in X$, $q \in X$ and $p \neq q$, then p has a neighbourhood U and q has a neighbourhood V such that $U \bigcap V = \phi$.
4. X is **locally compact space** if every point of X has a neighbourhood whose closure is compact.

5. **An inner product space** is a vector space V over the field F (F is either the field of real numbers $\mathbb{R}$ or the field of complex numbers $\mathbb{C}$) with an inner product, i.e., with a map
$$\langle \cdot, \cdot \rangle : V \times V \longrightarrow F$$
that satisfies the following axioms for all vectors $x, y, z \in V$ and all scalars $a \in F$:
(a) $\langle x, y \rangle = \overline{\langle y, x \rangle}$.
(b) $\langle ax, y \rangle = a\langle x, y \rangle$.
(c) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$.
(d) $\langle x, x \rangle \geq 0$.
(e) $\langle x, x \rangle = 0 \Leftrightarrow x = 0$.
The following norm that will be defined by
$$\|x\| = \langle x, x \rangle^{\frac{1}{2}} = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2},$$
where $x = (x_1, x_2, \ldots, x_n) \in V$, is called the **inner product norm or Euclidean norm** in vector space V.
6. A **normed vector space** is a pair $(V, \|\cdot\|)$ where V is a vector space and $\|\cdot\|$ a inner product norm on V.
7. The vector space $\mathbb{R}^n$ with inner product norm is called a **Euclidean n-space.**
8. A metric space M is called a **complete** or **Cauchy space** if every Cauchy sequence of points in M has a limit that is also in M. In the other words, if every Cauchy sequence in M converges in M.
9. A **Hilbert space** H is a real or complex inner product space that is also a complete metric space with respect to the distance function induced by the inner product.
10. A **function space** is a set of functions of a given kind from a set X to a set Y. It can be a topological space, vector space or both of them.
11. A **Banach space** is a complete normed vector space. Thus, a Banach space is a vector space with a metric that allows the computation of vector length and distance between vectors and is complete in the sense that a Cauchy sequence of vectors always converges to the well defined limit that is within the space.
12. If $f : V \longrightarrow W$ is a function between two vector spaces V and W on field F, and k is an integer, then f is said to be homogeneous of degree k if
$$f(\alpha \mathbf{v}) = \alpha^k f(\mathbf{v}) \tag{1.1}$$
for all nonzero $\alpha \in F$ and $\mathbf{v} \in V$. When the vector spaces involved are over the real numbers, a slightly less general form of homogeneity is often used, requiring only that the equation (1.1) holds for all $\alpha > 0$.
13. For a real number $p \geq 1$, the **p-norm** or $\mathbf{L^p} -$ **norm** of vector $\mathbf{x}$ in a real vector space is defined by
$$\|\mathbf{x}\|_p = (|x_1|^p + |x_2|^p + \ldots + |x_n|^p)^{\frac{1}{p}}.$$
The Euclidean norm from above falls into this class and is the 2-norm, and the 1-norm is the norm that corresponds to the rectilinear distance. The $\mathbf{L^\infty}$**-norm** is the limit of L^p-norm for $p \longrightarrow \infty$. So,
$$\|\mathbf{x}\|_\infty = \max\{|x_1|,\ |x_2|, \ldots,\ |x_n|\}.$$

We can summarise some of these definition in the diagram below. An arrow from space A to space B implies that space A is also a kind of space B. That means for instance, that a normed vector space is also a metric space.

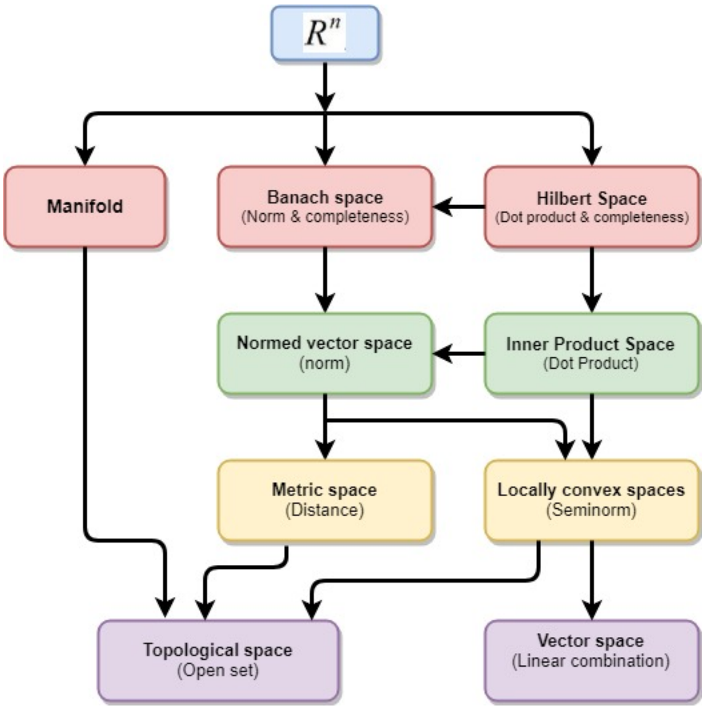

Figure 1.2: Abstract mathematical spaces

More information and theorems and examples can be found in [63], [64], [4].

1.2 Möbius n-Space

A **Möbius transformation** is a mapping $\mathbf{f} : \hat{\mathbb{R}}^n \longrightarrow \hat{\mathbb{R}}^n$ such that $\mathbf{f}$ is a composition of a finite number of the following transformations (see [74] p.14)

- Translation: $\mathbf{f}(\mathbf{x}) = \mathbf{x} + \mathbf{a}$, for a fixed $a \in \mathbb{R}^n$.
- Stretching: $\mathbf{f}(\mathbf{x}) = r\mathbf{x}$, where r is a positive real number.
- Orthogonal mapping: $\mathbf{f}$ is linear and $\|\mathbf{f}(\mathbf{x})\| = \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathbb{R}^n$.
- Inversion in a sphere $\mathbb{S}^{n-1}(\mathbf{a}, r)$: $\mathbf{f}(\mathbf{x}) = \mathbf{a} + \frac{r^2(\mathbf{x}-\mathbf{a})}{\|\mathbf{x}-\mathbf{a}\|^2}$, where $r > 0$ is the radius and $\mathbf{a} \in \mathbb{R}^n$ is the centre of the sphere.

The function π is called **stereographic projection** of $\hat{\mathbb{R}}^n$ to $\mathbb{S}^n$, defined as follows : $\pi(\infty) = e_{n+1}$, and for $\mathbf{x} \in \mathbb{R}^n$, $\pi(\mathbf{x})$ is the point of intersection with $\mathbb{S}^n$ of the Euclidean ray that issues from e_{n+1} and passes through $\mathbf{x}$. We can find the equation of function $\pi(\mathbf{x})$ by using elementary analytic geometry as

$$\pi(\mathbf{x}) = \Big(\frac{2x_1}{\|\mathbf{x}\|^2 + 1}, \dots, \frac{2x_n}{\|\mathbf{x}\|^2 + 1}, \frac{\|\mathbf{x}\|^2 - 1}{\|\mathbf{x}\|^2 + 1}\Big).$$

It is easy to see that $\pi(x)$ is an injective function. So, its inverse is defined

$$\pi^{-1}(\mathbf{y}) = \Big(\frac{y_1}{1-y_{n+1}}, \frac{y_2}{1-y_{n+1}} \dots, \frac{y_n}{1-y_{n+1}}\Big),$$

for $y \neq e_{n+1}$, while $\pi^{-1}(e_{n+1}) = \infty$. Stereographic projection provides a new metric structure on $\hat{\mathbb{R}}^n$ (see [30]). If $\mathbf{x}$ and $\mathbf{y}$ are two points in $\hat{\mathbb{R}}^n$, then the **chordal metric** on $\hat{\mathbb{R}}^n$ is defined by

$$q(x,y) = |\pi(\mathbf{x}) - \pi(\mathbf{x})| = \begin{cases} \dfrac{2\|\mathbf{x}-\mathbf{y}\|}{\sqrt{1+\|\mathbf{x}\|^2}\sqrt{1+\|\mathbf{y}\|^2}} & \text{if } \ \mathbf{x},\ \mathbf{y} \in \mathbb{R}^n \\ \dfrac{2}{\sqrt{1+\|\mathbf{x}\|^2}} & \text{if } \ \mathbf{x} \in \mathbb{R}^n,\ \mathbf{y} = \infty. \end{cases}.$$

1.3 Analysis and Measure Theory Background

In this section we introduce the **Lebesgue measure** and **conformal maps**. Let X be a measurable space and Y a topological space. If F is a symbol that denotes either $\mathbb{R}$ or $\mathbb{C}$, then let X be a measurable space then any constant function from X into F is measurable. The sum and product of two measurable functions from X into F are measurable. The complex conjugate of any measurable function from X into $\mathbb{C}$ is measurable. A function from X into $\mathbb{C}$ is measurable if and only if its real and imaginary parts are both measurable. All proofs of these statements exist in [76].

Definition 1.3.1. Let $(X, \mathfrak{M}, \mu)$ be a measure space. A measurable subset A of X is called a **null set** if $\mu(A) = 0$. A statement $P(x)$ is true for almost every x if there is a null set A such that $P(x)$ is true for every x not belonging to A. The measure μ is complete if every subset of every null set is measurable.

The **extended non-negative real axis** $[0, +\infty]$ is the non-negative real axis $[0, +\infty) = \{x \in \mathbb{R} : x \geq 0\}$with the additional element adjoined to it, which we label $+\infty$. We begin by defining the **Lebesgue outer measure** which assigns to each subset S of $\mathbb{R}$ an outer measure $m^*(S)$. Thus m^* will be a function

$$m^* : P(\mathbb{R}) \longrightarrow [0, +\infty]$$

where $P(\mathbb{R})$ denotes the power set of $\mathbb{R}$. m^* is not countably additive. Instead, it has the weaker property of countable subadditivity, meaning that,

$$m^*(\bigcup_{n\in\mathbf{N}} S_n) \leq \sum_{n\in\mathbf{N}} m^*(S_n)$$

for any sequence $\{S_n\}$ of subsets of $\mathbb{R}$.

Definition 1.3.2. If $S \subseteq \mathbb{R}$, the Lebesgue outer measure of S is defined by

$$m^*(S) = \inf\{\sum_{k=1}^{\infty} \ell(I_k) \mid \quad S \subseteq \bigcup_{k=1}^{\infty} I_k \qquad \textit{where } I_k \textit{ are open intervals}\}.$$

The Lebesgue outer measure has these following properties:

1. m^* is defined for every set of real numbers.
2. $0 \le m^*(S) \le \infty$.
3. $m^*(A) \le m^*(B)$ provided $A \subseteq B$ (monotonic).
4. $m^*(\phi) = 0$.
5. $m^*(\{a\}) = 0$.
6. $m^*(I) = \ell(I)$. I is an interval. The Lebesgue outer measure of an interval is its length.
7. $m^*(A + c) = m^*(A)$ (translation invariant).
8. $m^*(\bigcup_{n\in\mathbf{N}} S_n) \le \sum_{n\in\mathbf{N}} m^*(S_n)$ for every sequence of sets of real numbers (countable subadditivity).

Proofs of these properties exist in pages (96) and (97) of [17] or in the book [71].

Definition 1.3.3. A subset E of $\mathbb{R}$ is said to be **Lebesgue measurable** if

$$m^*(S \cap E) + m^*(S \cap E^c) = m^*(S)$$

for every S of $\mathbb{R}$. In this case, the outer measure $m^*(E)$ of E is called the **Lebesgue measure** of E, and is denoted $m(E)$.

Definition 1.3.3 is symmetric between E and E^c. Thus a set E is measurable if and only if its complement E^c is measurable. Lebesgue measure and Lebesgue measurable sets have the following properties:

1. If E and F are measurable subsets of $\mathbb{R}$, then $E \cup F$ is also measurable.
2. If E and F are measurable subsets of $\mathbb{R}$, then $E \cap F$ is also measurable.
3. Let $\{E_k\}$ be a sequence of pairwise disjoint measurable subsets of $\mathbb{R}$. Then the union $\bigcup_{k\in\mathbf{N}} E_k$ is measurable, and
$$m(\bigcup_{k\in\mathbf{N}} E_k) = \sum_{k\in\mathbf{N}} m(E_k).$$
4. Every interval I in $\mathbb{R}$ is Lebesgue measurable.
5. (**Heine-Borel Theorem**) Let $[a, b]$ be a closed interval in $\mathbb{R}$, and $\mathcal{C}$ be a family of open intervals that covers $[a, b]$. Then there exists a finite subcollection of $\mathcal{C}$ that covers $[a, b]$.
6. For every interval I in $\mathbb{R}$, $m(I) = \ell(I)$.

Proofs of all parts exist in (http://math.bard.edu/belk/math461/LebesgueMeasure.pdf) or you can see [17] and [71]. Assume $(X, \mathfrak{M})$ is a measurable space, then the function $f : X \longrightarrow \mathbb{R}$ is said to be a **Lebesgue measurable function** if and only if $\mathfrak{M}$ is the σ-algebra of Lebesgue measurable sets and for all $\alpha \in \mathbb{R}$ the set $\{\, x \in X \, : \, f(x) < \alpha\}$ is Lebesgue measurable. This definition can also be described by $\{\, x \in X \, : \, f(x) > \alpha\}$, $\{\, x \in X \, : \, f(x) \ge \alpha\}$ and $\{\, x \in X \, : \, f(x) \le \alpha\}$. All definitions are equivalent. Lebesgue measurable functions are important in mathematical analysis because they can be integrated.
Let $(X, \mathfrak{M}, \mu)$ be a measure space. If $s : X \longrightarrow [0, \infty)$ is a measurable simple function, of the form

$$s = \sum_{i=1}^{n} \alpha_i \chi_{A_i},$$

Where $\alpha_1, \alpha_2, \ldots, \alpha_n$ are the distinct values of s and A_i are disjoint members of $\mathfrak{M}$. If $E \in \mathfrak{M}$, we define

$$\int_E s d\mu = \sum_{i=1}^{n} \alpha_i \mu(A_i \cap E).$$

The convention $0 \cdot \infty = 0$ is used here; it may happen that $\alpha_i = 0$ foe some i and that $\mu(A_i \cap E) = \infty$. If $f : X \longrightarrow [0, \infty)$ is measurable, and $E \in \mathfrak{M}$, we define

$$\int_E f \mathrm{d}\mu = \sup \int_E s \mathrm{d}\mu, \tag{1.2}$$

the supremum being taken over all simple measurable function s such that $0 \leq s \leq f$. The left side of (1.2) is called the **Lebesgue integral** of f over E, with respect to the measure μ. It is a number in $[0, \infty]$ (see [64] p.19).
In this section, we start to define some useful ideas that help to define a new concept that is called linear distortion. Let A be a subset of $\mathbb{R}^n$ and B be a subset of $\mathbb{R}^m$. Then we can prove that for a function $\mathbf{f} : A \longrightarrow B$, the following are equivalent:

- $\mathbf{f}$ is continuous.
- For every set U that is open in B, then $\mathbf{f}^{-1}(U)$ is open in A.
- For every set U that is closed in B, then $\mathbf{f}^{-1}(U)$ is closed in A.

Proof of this proposition exists in [55] or [72]. But you should note that $\mathbb{R}^n$ and $\mathbb{R}^m$ are Euclidean normed and also topological spaces (see Figure 1.2). We next wish to describe what it means for A and B to be topologically the same. There should be a bijection between them that pairs open sets with open sets (see [72]).

Definition 1.3.4. A function $\mathbf{f} : A \longrightarrow B$ is called a **homeomorphism** if $\mathbf{f}$ is bijective and continuous and $\mathbf{f}^{-1}$ is continuous. If such a function exists, then A and B are said to be **homeomorphic**.

So, another way to define a homeomorphism is to say that it is a bijective correspondence $\mathbf{f} : A \longrightarrow B$ such that $\mathbf{f}(U)$ is open in B if and only if U in open in A. This remark shows that a homeomorphism $\mathbf{f} : A \longrightarrow B$ gives us a bijective correspondence not only between A and B but also between the collections of open sets of A and of Y (see [55]). A subset $U \subset \mathbb{R}^n$ is called **path-connected** if for every pair $\mathbf{p}, \mathbf{q} \in U$, there exists a continuous function $\mathbf{g} : [0, 1] \longrightarrow U$ with $\mathbf{g}(0) = \mathbf{p}$ and $\mathbf{g}(1) = \mathbf{q}$. A set $U \subset \mathbb{R}^n$ is called **connected** if there is no subset of U that is both open in U and closed in U. We can prove that every path-connected set $U \subset \mathbb{R}^n$ is connected (see [72] and [55]). Let (X, d) be a metric space and $f : X \longrightarrow \mathbb{R}$ be a real valued function. We recall that U_y and L_y denote upper contour set of f and lower contour set of f at y. They are defined as

$$U_y = f^{-1}([y, \infty)) = \{x \in X \mid f(x) \geq y\}$$

$$L_y = f^{-1}((-\infty, y]) = \{x \in X \mid f(x) \leq y\}.$$

Let $f : X \longrightarrow \mathbb{R}$ be a function. Then we can prove that the following are equivalent.

- For any $y \in \mathbb{R}$, U_y is closed.
- For any $y \in \mathbb{R}$, $f^{-1}((-\infty, y)) = [U_y]^c$ is open.
- For any $x \in X$, if the sequence $\{x_n\}$ in X converges to x, then for any $\epsilon > 0$ there is a natural number N such that for all $n > N$, $f(x) > f(x_n) - \epsilon$.

Also, we can prove that the following are equivalent.

- For any $y \in \mathbb{R}$, L_y is closed.
- For any $y \in \mathbb{R}$, $f^{-1}((y,\infty)) = [L_y]^c$ is open.
- For any $x \in X$, if the sequence $\{x_n\}$ in X converges to x, then for any $\epsilon > 0$ there is a natural number N such that for all $n > N$, $f(x) < f(x_n) + \epsilon$.

We do not prove these theorems and skip on their proofs that can be found all of the proofs in most of the topology and mathematical analysis books.

Definition 1.3.5. Let $f : X \longrightarrow \mathbb{R}$ be a function.

1. f is **upper semicontinuous** if and only if for any $y \in \mathbb{R}$, $f^{-1}((-\infty, y))$ is open.
2. f is **lower semicontinuous** if and only if for any $y \in \mathbb{R}$, $f^{-1}((y,\infty))$ is open.

Informally, a function is upper semicontinuous if it is continuous or, if not, it only jumps up; a function is lower semicontinuous if it is continuous or, if not, it only jumps down. We can show that f is continuous if and only if it is both upper and lower semicontinuous. It is obvious that f is lower semicontinuous if and only if $-f$ is upper continuous and vice versa. There is another definition for lower and upper semicontinuous. Let $f : X \longrightarrow \mathbb{R}$ be a function. f is called lower semicontinuous at $a \in X$ if

$$\liminf_{x \to a} f(x) \geq f(a),$$

and is called upper semicontinuous at $a \in X$ if

$$\limsup_{x \to a} f(x) \leq f(a).$$

Theorem 1.3.1. Let $f : X \longrightarrow \mathbb{R}$ be a function. Then the following conditions are equivalent:

1. f is lower semicontinuous on X.
2. The lower contour set of f at $y \in \mathbb{R}$, L_y, is closed.
3. The $\operatorname{epi} f$ is closed in $X \times \mathbb{R}$ ($\operatorname{epi} f$ is defined in Equation (2.7)).

We can express this theorem for upper semicontinuous functions on X. For more details and the proof you can see [14] and [42]. An important property of lower semicontinuous functions is given by the following well-known Weierstrass theorem. A lower semicontinuous function f on a compact topological space X takes a minimum value on X (see [14]). We need to next theorem to define another definition.

Theorem 1.3.2. Let A be a subset of $\mathbb{R}^n$, and let $\mathbf{f} : A \longrightarrow \mathbb{R}^n$ be an injective function. Suppose that $\mathbf{f}$ is continuous in some neighbourhood of a point x, that $\mathbf{f}$ is differentiable at x itself, and that $D\mathbf{f}(x)$ is nonsingular. Then $y = \mathbf{f}(x)$ is an interior point of $\mathbf{f}(A)$, the function $\mathbf{g} = \mathbf{f}^{-1} : \mathbf{f}(A) \longrightarrow A$ is differentiable at y, and

$$D\mathbf{g}(y) = [D\mathbf{f}(x)]^{-1}$$

Consider an open set U in $\mathbb{R}^n$ and injective function $\mathbf{f}$ belonging to C^k for some $k \geq 1$. In the case $k \geq 1$ assume additionally that $J_{\mathbf{f}}(x) \neq 0$ for every x in U, where $J_{\mathbf{f}}$ denotes the

Jacobian determinant of $\mathbf{f}$, i.e., $J_{\mathbf{f}}(x) = \det[D\mathbf{f}(x)]$. (see [42], [30] and [6]).
Let A be a subset of $\mathbb{R}^n$ and B be a subset of $\mathbb{R}^m$. In this case, A and B are called **diffeomorphic**, if there exists a smooth bijective function $\mathbf{f} : A \longrightarrow B$ whose inverse is also smooth. Then $\mathbf{f}$ is called a **diffeomorphism**. In general, we can say, for k a positive integer or ∞, a C^k-diffeomorphism $\mathbf{f}$ between open sets A and B in $\mathbb{R}^n$ is a bijection $\mathbf{f} : A \longrightarrow B$ such that both $\mathbf{f}$ and $\mathbf{f}^{-1}$ are of class C^k. Unless otherwise stated, the word "diffeomorphism" will mean C^1-diffeomorphism. If $k \geq 1$ and the set A is connected, then either $J_{\mathbf{f}}(x) > 0$ for every x in A or $J_{\mathbf{f}}(x) < 0$ for all such x. Now we can explain a criterion that enables us to determine whether a function is, at least on a local level, a diffeomorphism.

Theorem 1.3.3. Let U be an open set in $\mathbb{R}^n$, and let $\mathbf{f}$ be a member of the class C^k with $1 \leq k \leq \infty$. If x is a point of U for which $J_{\mathbf{f}}(x) \neq 0$, then there exists an $r > 0$ such that the restriction of $\mathbf{f}$ to the ball $B = B^n(x, r)$ is a C^k-diffeomorphism of B onto $\mathbf{f}(B)$.

If an embedding $\mathbf{f} : U \longrightarrow \mathbb{R}^n$ is differentiable at $a \in U$, then the Jacobian determinant $J_{\mathbf{f}}(x) = \det[D\mathbf{f}(x)]$ represents the infinitesimal change of volume under the mapping $\mathbf{f}$. To be precise, we have

$$|J_{\mathbf{f}}(a)| = \lim_{r \to 0} \frac{|\mathbf{f}(B(a,r))|}{|B(a,r)|}, \tag{1.3}$$

where $|B|$ denotes the n-dimensional measure of B. For more details you can see [30] and [42].
If A and B are nonempty subsets of $\mathbb{R}^n$, then $\operatorname{diam}(A)$ and $\operatorname{dist}(A, B)$ denote the **chordal diameter** of A and the **chordal distance** between A and B, respectively. Thus

$$\operatorname{diam}(A) = \sup\big\{q(x,y) \ : \ x, y \in A\big\}, \qquad \operatorname{dist}(A,B) = \inf\big\{q(x,y) \ : \ x \in A\,,\ y \in B\big\}.$$

We shorten $\operatorname{dist}(\{x\}, B)$ to the more compact $\operatorname{dist}(x, B)$ (see [30]). Now, we define the main subject of this thesis that is called linear distortion or dilatation. We know that the **domain in topology** means an open connected set. So, in the rest of this thesis, whenever we use the domain term, we mean a connected open set in $\mathbb{R}^n$.

1.4 Differentiation in Higher-Dimensions

Definition 1.4.1. Suppose $\mathbf{f} : \mathbb{R}^n \longrightarrow \mathbb{R}$ is a scalar field of n variables. The **partial derivative of f** with respect to x_i is the ordinary derivative of the partial function with respect to x_i. Symbolically, we have

$$\frac{\partial \mathbf{f}}{\partial x_i} = \lim_{h \to 0} \frac{\mathbf{f}(x_1, \ldots, x_i + h, \ldots, x_n) - \mathbf{f}(x_1, \ldots, x_n)}{h}. \tag{1.4}$$

Sometimes, in some functions the partial derivatives exist, but the existence of partial derivatives alone is not enough for differentiablity of the function $\mathbf{f}$. Because of that, we need to define the differentiability in the general case. Let X be an open set in $\mathbb{R}^n$ and $\mathbf{f} : X \longrightarrow \mathbb{R}$ be a scalar field; let $\mathbf{a} = (a_1, a_2, \ldots, a_n) \in X$. We say that $\mathbf{f}$ is **differentiable at a** if all the partial derivatives $\mathbf{f}_{x_i}(\mathbf{a})$, $i = 1, 2, \ldots, n$, exist and if the function $\mathbf{h} : \mathbb{R}^n \longrightarrow \mathbb{R}$ defined by

$$\mathbf{h}(\mathbf{x}) = \mathbf{f}(\mathbf{a}) + \mathbf{f}_{x_1}(\mathbf{a})(x_1 - a_1) + \ldots + \mathbf{f}_{x_n}(\mathbf{a})(x_n - a_n) \tag{1.5}$$

is a good linear approximation to **f** near **a**, meaning that

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{\mathbf{f}(\mathbf{x})-\mathbf{h}(\mathbf{x})}{\|\mathbf{x}-\mathbf{a}\|} = \mathbf{0}. \tag{1.6}$$

We can use vector and matrix notation to rewrite things a bit. Define the **gradient** of a scalar field $\mathbf{f} : X \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ to be a vector

$$\nabla\mathbf{f} = (\frac{\partial \mathbf{f}}{\partial x_1}, \dots, \frac{\partial \mathbf{f}}{\partial x_n}).$$

Let X be an open set in $\mathbb{R}^n$ and let $\mathbf{f} : X \longrightarrow \mathbb{R}^m$ be a vector field of n variables. We define the **matrix of partial derivatives of f**, denoted D**f**, to be the $m \times n$ matrix whose ij-th entry is $\frac{\partial f_i}{\partial x_j}$, where $f_i : X \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ is the i-th component function of **f**. That is,

$$\mathbf{f}'(\mathbf{x}) = D\mathbf{f}(x_1, x_2, \dots, x_n) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{bmatrix} \tag{1.7}$$

The i-th row of D**f** is nothing more than Df_i, and the entries of Df_i are precisely the components of the gradient vector ∇f_i. (Indeed, in the case where $m = 1$, ∇f and $D\mathbf{f}$ mean exactly the same thing.) Now we can define derivative for vector fields.

Definition 1.4.2. Let X be an open set in $\mathbb{R}^n$, let $\mathbf{f} : X \longrightarrow \mathbb{R}^m$, and let $\mathbf{a} \in X$. We say that **f** is **differentiable at a** if D**f**(**a**) exists and if the function $\mathbf{h} : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ defined by

$$\mathbf{h}(\mathbf{x}) = \mathbf{f}(\mathbf{a}) + D\mathbf{f}(\mathbf{a})(\mathbf{x}-\mathbf{a})$$

is a good linear approximation to **f** near **a**. That is, we require

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{\|\mathbf{f}(\mathbf{x})-\mathbf{h}(\mathbf{x})\|}{\|\mathbf{x}-\mathbf{a}\|} = \lim_{\mathbf{x}\to\mathbf{a}} \frac{\|\mathbf{f}(\mathbf{x})-[\mathbf{f}(\mathbf{a})+D\mathbf{f}(\mathbf{a})(\mathbf{x}-\mathbf{a})]\|}{\|\mathbf{x}-\mathbf{a}\|} = \mathbf{0}. \tag{1.8}$$

In fact, we could have approached our discussion of differentiability much more abstractly right from the beginning. We could have defined a function $\mathbf{f} : X \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ to be differentiable at a point $\mathbf{a} \in X$ to mean that there exists some linear mapping $A : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ such that

$$\lim_{\mathbf{x}\to\mathbf{a}} \frac{\|\mathbf{f}(\mathbf{x})-[\mathbf{f}(\mathbf{a})+A(\mathbf{x}-\mathbf{a})]\|}{\|\mathbf{x}-\mathbf{a}\|} = \mathbf{0}.$$

We can say if $\mathbf{f} : X \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is differentiable at **a**, then it is continuous at **a**. Also, we can prove that if $\mathbf{f} : X \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is such that, for $i = 1, \dots, m$ and $j = 1, \dots, n$ all $\frac{\partial f_i}{\partial x_j}$ exist and are continuous in a neighborhood of **a** in X, then **f** is differentiable at **a**. In general we can show that a function $\mathbf{f} : X \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is differentiable at $\mathbf{a} \in X$ if and only if each of its component functions $f_i : X \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$, is differentiable at **a** for $i = 1, 2, \dots, m$. If you want to see the proofs of these statements you can see [5], [63] or [18]. In general, if $\mathbf{f} : X \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ is a scalar field of n variables, then the **kth-order partial derivative** with respect to the variables $x_{i_1}, x_{i_2}, \dots, x_{i_k}$ (in that order), where $i_1, i_2, \dots, i_k$ are integers in the set $\{1, 2, \dots, n\}$ is the iterated derivative

$$\frac{\partial^k \mathbf{f}}{\partial x_{i_k} \dots \partial x_{i_2} \partial x_{i_1}} = \frac{\partial}{\partial x_{i_k}} \dots \frac{\partial}{\partial x_{i_2}} \frac{\partial}{\partial x_{i_1}} \mathbf{f}(x_1, x_2, \dots, x_n).$$

Equivalent notation for this kth-order partial is

$$\mathbf{f}_{x_{i_1} x_{i_2} \dots x_{i_k}}(x_1, x_2, \dots, x_n).$$

Definition 1.4.3. Assume X is an open set in $\mathbb{R}^n$. A scalar field $\mathbf{f} : X \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ whose partial derivatives up to (and including) order at least k exist and are continuous on X is said to be of **class** C^k. If $\mathbf{f}$ has continuous partial derivatives of all orders on X, then $\mathbf{f}$ is said of **class** C^∞ or **smooth**. A vector field $\mathbf{f} : X \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is of class C^k (or of class C^∞) if and only if each of its component functions is of class C^k (or of class C^∞).

Theorem 1.4.1. Let $\mathbf{f} : X \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a scalar field of class C^k. Then the order in which we calculate any k*th*-order partial derivative does not matter: If $(i_1, \dots, i_k)$ are any k integers (not necessarily distinct) between 1 and n, and if $(j_1, \dots, j_k)$ is any permutation of these integers, then

$$\frac{\partial^k \mathbf{f}}{\partial x_{i_1} \dots \partial x_{i_k}} = \frac{\partial^k \mathbf{f}}{\partial x_{j_1} \dots \partial x_{j_k}}$$

We can also prove the chain rule is true for a scalar field and a vector field. This theorem has a long proof (see [18] or [63]).

1.5 Higher-Order Taylor's Series

Our goal in this section is to provide a means of approximating any scalar field by a polynomial of given degree, known as the **Taylor polynomial**.

Theorem 1.5.1. (First-order Taylor's formula in several variables) Let X be an open set in $\mathbb{R}^n$ and suppose that $\mathbf{f} : X \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ is differentiable at the point $\mathbf{a}$ in X. Let

$$\mathbf{p_1}(\mathbf{x}) = \mathbf{f}(\mathbf{a}) + D\mathbf{f}(\mathbf{a})(\mathbf{x} - \mathbf{a}). \tag{1.9}$$

Then

$$\mathbf{f}(\mathbf{x}) = \mathbf{p_1}(\mathbf{x}) + R_1(\mathbf{x}, \mathbf{a}),$$

where $R_1(\mathbf{x}, \mathbf{a}) / \left\| \mathbf{x} - \mathbf{a} \right\| \longrightarrow 0$ as $\mathbf{x} \longrightarrow \mathbf{a}$.

We may also express the first-order Taylor polynomial using the gradient in place of formula (1.9), so,

$$\mathbf{p_1}(\mathbf{x}) = \mathbf{f}(\mathbf{a}) + \nabla \mathbf{f}(\mathbf{a})(\mathbf{x} - \mathbf{a}).$$

Before we explore higher-orders of Taylor's theorem in several variables, we consider the linear approximation in further detail.

Definition 1.5.1. Let $\mathbf{f} : X \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ and let $\mathbf{a} \in X$. The **incremental change of f**, denoted $\Delta\mathbf{f}$, is

$$\Delta\mathbf{f} = \mathbf{f}(\mathbf{a}+\mathbf{h}) - \mathbf{f}(\mathbf{a}),$$

where $\mathbf{h} = \mathbf{x} - \mathbf{a}$. The **total differential of f**, denoted $\mathrm{d}\mathbf{f}(\mathbf{a},\mathbf{h})$, is

$$\mathrm{d}\mathbf{f}(\mathbf{a},\mathbf{h}) = \frac{\partial \mathbf{f}}{\partial x_1}(\mathbf{a})h_1 + \frac{\partial \mathbf{f}}{\partial x_2}(\mathbf{a})h_2 + \ldots + \frac{\partial \mathbf{f}}{\partial x_n}(\mathbf{a})h_n.$$

The significance of the differential is that for $\mathbf{h} \approx \mathbf{0}$,

$$\Delta\mathbf{f} \approx \mathrm{d}\mathbf{f}.$$

We have abbreviated $\mathrm{d}\mathbf{f}(\mathbf{a},\mathbf{h})$ by $\mathrm{d}\mathbf{f}$. The incremental change $\Delta\mathbf{f}$ equals the change in z-coordinate of the graph of $z = \mathbf{f}(x,y)$ as a point in $\mathbb{R}^2$ changes from $\mathbf{a} = (a,b)$ to $\mathbf{a}+\mathbf{h} = (a+h, b+k)$. The differential $\mathrm{d}\mathbf{f}$ equals the change in z-coordinate of the graph of the tangent plane at (a,b).

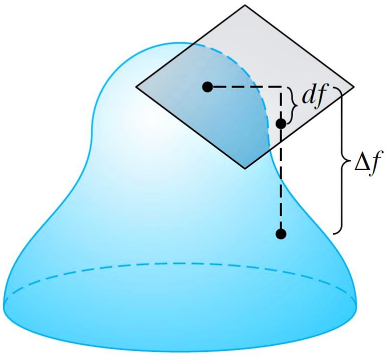

Figure 1.3: The total differential of $\mathbf{f}$
Source: Page 250 of book [18]

Now we state the second-order version of Taylor's theorem precisely. Taylor's theorem can be explained for every order but it requires that $\mathbf{f}$ be a C^k class function.

Theorem 1.5.2. (Second-order Taylor's formula) Let X be an open set in $\mathbb{R}^n$, and suppose that $\mathbf{f} : X \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ is of class C^2. Let

$$\mathbf{p_2}(\mathbf{x}) = \mathbf{f}(\mathbf{a}) + \sum_{i=1}^{n} \mathbf{f}_{x_i}(x_i - a_i) + \frac{1}{2}\sum_{i,j=1}^{n} \mathbf{f}_{x_i x_j}(\mathbf{a})(x_i - a_i)(x_j - a_j). \tag{1.10}$$

Then

$$\mathbf{f}(\mathbf{x}) = \mathbf{p_2}(\mathbf{x}) + R_2(\mathbf{x},\mathbf{a}),$$

where $\left|R_2(\mathbf{x},\mathbf{a})\right| / \left\|\mathbf{x}-\mathbf{a}\right\|^2 \longrightarrow 0$ as $\mathbf{x} \longrightarrow \mathbf{a}$.

Proof of this theorem exists in [5] and [18]. Recall that the formula for the first-order Taylor polynomial $\mathbf{p_1}$ was written as

$$\mathbf{p_1}(\mathbf{x}) = \mathbf{f}(\mathbf{a}) + \nabla\mathbf{f}(\mathbf{a})(\mathbf{x}-\mathbf{a}).$$

If we consider $\mathbf{h} = \mathbf{x} - \mathbf{a}$. Then the formula that was written in the above line becomes

$$\mathbf{p_1}(\mathbf{x}) = \mathbf{f}(\mathbf{a}) + D\mathbf{f}(\mathbf{a})\mathbf{h} = \mathbf{f}(\mathbf{a}) + \sum_{i=1}^{n} \frac{\partial \mathbf{f}}{\partial x_i}(\mathbf{a})h_i. \tag{1.11}$$

We can show this formula by using vector and matrix notation. It turns out it is possible to do something similar for the second-order polynomial $\mathbf{p_2}$ (see [18]).

Definition 1.5.2. The **Hessian matrix** of a function $\mathbf{f} : X \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ is the matrix whose ijth entry is $\partial^2\mathbf{f}/\partial x_j \partial x_i$. That is,

$$\mathrm{H}(\mathbf{f}) = \begin{bmatrix} \mathbf{f}_{x_1x_1} & \mathbf{f}_{x_1x_2} & \cdots & \mathbf{f}_{x_1x_n} \\ \mathbf{f}_{x_2x_1} & \mathbf{f}_{x_2x_2} & \cdots & \mathbf{f}_{x_2x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{f}_{x_nx_1} & \mathbf{f}_{x_nx_2} & \cdots & \mathbf{f}_{x_nx_n} \end{bmatrix}. \tag{1.12}$$

Now let's look again at the formula for $\mathbf{p_2}$ in Theorem 1.5.2:

$$\mathbf{p_2}(\mathbf{x}) = \mathbf{f}(\mathbf{a}) + \sum_{i=1}^{n} \mathbf{f}_{x_i}(\mathbf{a})(x_i - a_i) + \frac{1}{2}\sum_{i,j=1}^{n} \mathbf{f}_{x_ix_j}(\mathbf{a})(x_i - a_i)(x_j - a_j).$$

If we consider $h_i = (x_i - a_i)$ and $h_j = (x_j - a_j)$, then we can write

$$\mathbf{p_2}(\mathbf{x}) = \mathbf{f}(\mathbf{a}) + \sum_{i=1}^{n} \mathbf{f}_{x_i}(\mathbf{a})h_i + \frac{1}{2}\sum_{i,j=1}^{n} \mathbf{f}_{x_ix_j}(\mathbf{a})h_ih_j.$$

We have let $\mathbf{h} = \mathbf{x} - \mathbf{a}$ (see [18]). This can be written as

$$\mathbf{p_2}(\mathbf{x}) = \mathbf{f}(\mathbf{a}) + \begin{bmatrix} \mathbf{f}_{x_1}(\mathbf{a}) & \mathbf{f}_{x_2}(\mathbf{a}) & \cdots & \mathbf{f}_{x_n}(\mathbf{a}) \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix} + \frac{1}{2} \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix}^t \begin{bmatrix} \mathbf{f}_{x_1x_1} & \mathbf{f}_{x_1x_2} & \cdots & \mathbf{f}_{x_1x_n} \\ \mathbf{f}_{x_2x_1} & \mathbf{f}_{x_2x_2} & \cdots & \mathbf{f}_{x_2x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{f}_{x_nx_1} & \mathbf{f}_{x_nx_2} & \cdots & \mathbf{f}_{x_nx_n} \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix}$$

Thus, we see that

$$\mathbf{p_2}(\mathbf{x}) = \mathbf{f}(\mathbf{a}) + \sum_{i=1}^{n} \mathbf{f}_{x_i}(\mathbf{a})(h_i) + \frac{1}{2}\mathbf{h}^t\mathrm{H}\mathbf{f}(\mathbf{a})\mathbf{h}. \tag{1.13}$$

For more information you can see [18]. If we want to write the second-order Taylor's formula for function $\mathbf{f}$, then

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{a}) + \sum_{i=1}^{n} \mathbf{f}_{x_i}(\mathbf{a})(h_i) + \frac{1}{2}\mathbf{h}^t\mathrm{H}\mathbf{f}(\mathbf{a})\mathbf{h} + R_2(\mathbf{x}, \mathbf{a}),$$

where $\left|R_2(\mathbf{x}, \mathbf{a})\right| / \left\|\mathbf{h}\right\|^2 \longrightarrow 0$ as $\mathbf{x} \longrightarrow \mathbf{a}$.

A scalar field $\mathbf{f}$ is said to have an **absolute maximum** at a point $\mathbf{a}$ of a set X in $\mathbb{R}^n$ if

$$\mathbf{f}(\mathbf{x}) \leq \mathbf{f}(\mathbf{a})$$

for all $\mathbf{x} \in X$. The number $\mathbf{f}(\mathbf{a})$ is called the **absolute maximum value** of $\mathbf{f}$ on X. The function $\mathbf{f}$ is said to have **relative maximum** at $\mathbf{a}$ if the top inequality is satisfied for every $\mathbf{x}$ in some n-ball $\mathrm{B}(\mathbf{a})$ lying in X. The terms **absolute minimum** and **relative minimum** are defined in an analogous fashion, using the opposite direction for the top inequality. A number which is either a relative maximum or a relative minimum of $\mathbf{f}$ is called an **extremum of f.** Assume $\mathbf{f}$ is differentiable at $\mathbf{a}$. If $\nabla\mathbf{f}(\mathbf{a}) = D\mathbf{f}(\mathbf{a})$ is either zero or undefined, then the point $\mathbf{a}$ is called a **critical point of f**. A critical point is called a **saddle point** if every n-ball $\mathrm{B}(\mathbf{a})$ contains points $\mathbf{x}$ such that $\mathbf{f}(\mathbf{x}) < \mathbf{f}(\mathbf{a})$ and other points such that $\mathbf{f}(\mathbf{x}) > \mathbf{f}(\mathbf{a})$.

Theorem 1.5.3. Let X be an open set in $\mathbb{R}^n$ and let $\mathbf{f} : X \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be differentiable. If $\mathbf{f}$ has a relative extremum at $\mathbf{a} \in X$, then $D\mathbf{f}(\mathbf{a}) = \mathbf{0}$.

We do not want to prove this theorem, if you want to see the proof of this theorem, see [18]. The next theorem describes the kind and nature of the critical points in terms of the algebraic sign of the quadratic form of the Hessian matrix.

Theorem 1.5.4. Let X be an open set in $\mathbb{R}^n$ and $\mathbf{f} : X \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}$ be a function of class C^2. Suppose that $\mathbf{a} \in X$ is a critical point of $\mathbf{f}$

1. If the Hessian $H\mathbf{f}(\mathbf{a})$ is positive definite, then $\mathbf{f}$ has a relative minimum at $\mathbf{a}$.
2. If the Hessian $H\mathbf{f}(\mathbf{a})$ is negative definite, then $\mathbf{f}$ has a relative maximum at $\mathbf{a}$.
3. If $\det H\mathbf{f}(\mathbf{a}) \neq 0$ but $H\mathbf{f}(\mathbf{a})$ is neither positive nor negative definite, then $\mathbf{f}$ has a saddle point at $\mathbf{a}$.

1.6 Linear Algebra Background

Theorem 1.6.1. If A and B are symmetric matrices with the same size, and if k is any scalar, then:
(a) A^t is symmetric.
(b) $A + B$ and $A - B$ are symmetric.
(c) kA is symmetric.

Definition 1.6.1. If A and B are square matrices with the same size such that $AB = BA$, then A and B are called **commutative**

It is not true, in general, that the product of symmetric matrices is symmetric. But it is easy to see, the product of two symmetric matrices is symmetric if and only if the matrices commute.

Theorem 1.6.2. If A is an invertible symmetric matrix, then A^{-1} is symmetric.

Theorem 1.6.3. If A is an invertible matrix, then AA^t and A^tA are also invertible.

Definition 1.6.2. A square matrix A with the property

$$A^{-1} = A^t$$

is said to be an **orthogonal matrix**

It follows from this definition that a square matrix A is orthogonal if and only if

$$AA^t = A^tA = I.$$

Theorem 1.6.4. For orthogonal matrices, we can prove the following statements:
(a) The inverse of an orthogonal matrix is orthogonal.
(b) A product of orthogonal matrices is orthogonal.
(c) If A is orthogonal, then $\det(A) = 1$ or $\det(A) = -1$.

Definition 1.6.3. Let A be an $n \times n$ matrix. If there is an orthogonal matrix P such that the matrix $P^{-1}AP$ is an orthogonal, then A is said to be **orthogonally diagonalizable** and P is said to **orthogonally diagonalize** A.

Definition 1.6.4. A **quadratic form** in the n real variables $x_1, x_2, \ldots, x_n$ is an expression that can be written as

$$\begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} A \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \tag{1.14}$$

where A is a symmetric $n \times n$ matrix. If we let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ then formula (1.4) can be written more compactly as

$$\mathbf{x^t} A\mathbf{x}. \tag{1.15}$$

If we use the fact that A is symmetric, that is, $A = A^t$, then the formula (1.15) can be expressed in terms of the Euclidean inner product by writing

$$\mathbf{x^t} A\mathbf{x} = \mathbf{x^t}(A\mathbf{x}) = \langle A\mathbf{x}, \mathbf{x} \rangle = \langle \mathbf{x}, A\mathbf{x} \rangle.$$

For more information you can see [3], [47] or [38].

Theorem 1.6.5. Let A be a symmetric $n \times n$ matrix whose eigenvalues in decreasing size order are $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$. If $\mathbf{x}$ is constrained so that $\|\mathbf{x}\| = 1$ relative to the Euclidean inner product on $\mathbb{R}^{\ltimes}$, then :
(a) $\lambda_1 \geq \mathbf{x^t} A\mathbf{x} \geq \lambda_n$.
(b) $\mathbf{x^t} A\mathbf{x} = \lambda_n$ if $\mathbf{x}$ is an eigenvector of A corresponding to λ_n and $\mathbf{x^t} A\mathbf{x} = \lambda_1$ if $\mathbf{x}$ is an eigenvector of A corresponding to λ_1.

We do not want to prove this theorem. Its proof exists in the most linear algebra book. For more information see [3] page 480.

Definition 1.6.5. A quadratic form $\mathbf{x^t} A\mathbf{x}$ is called **positive definite** if $\mathbf{x^t} A\mathbf{x} > 0$ for all $\mathbf{x} \neq 0$, and the symmetric matrix A is called a **positive definite matrix**.

A symmetric matrix A is positive definite if and only if all the eigenvalues of A are positive. Also, we can prove a symmetric matrix A is positive definite if and only if the determinant of every principle submatrix is positive. For matrices with real entries, the orthogonal matrices and symmetric matrices play an important role in the orthogonal diagonalization problem. But we need something for matrices with complex entries. We review some definitions of matrices with complex entries. If A is a matrix with complex entries, then the **conjugate transpose** of A, denoted by A^*, is defined by

$$A^* = \overline{A}^t. \tag{1.16}$$

If A and B are matrices with complex entries and k is any complex number, then we can prove some properties like $(A^*)^* = A$, $(A+B)^* = A^* + B^*$, $(kA)^* = \overline{k}A^*$ or $(AB)^* = B^*A^*$. By Definition 1.6.2 a matrix with real entries is called orthogonal if $A^{-1} = A^t$. This sort of matrix is called unitary matrix for complex entries.

Definition 1.6.6. A square matrix A with complex entries is called **unitary** if

$$A^{-1} = A^*. \tag{1.17}$$

Recall that a square matrix A with real entries is called orthogonally diagonalizable if there is an orthogonal matrix P such that $P^{-1}AP$ is diagonal. For complex matrices we have an analogous concept. A square matrix A with complex entries is called **unitarily diagonalizable** if there is a unitary P such that $P^{-1}AP = P^*AP$ is diagonal; the matrix P is said to **unitarily diagonalize** A.

Definition 1.6.7. A square matrix A with complex entries is called **Hermitian** or **self-adjoint** if

$$A = A^*.$$

It is easy to recognise Hermitian matrices by inspection: The entries on the main diagonal are real numbers, and the mirror of each entry across the main diagonal is its complex conjugate. Also, a matrix A with complex entries is called **normal** if

$$AA^* = A^*A.$$

We have two results with these definitions that we will use in the next chapters (you can see [3] or [47]).

Theorem 1.6.6. The eigenvalues of a Hermitian matrix are real numbers.

Theorem 1.6.7. The eigenvalues of a symmetric matrix with real entries are real numbers.

Now, we should review some properties of diagonalizable matrices. An $n \times n$ matrix A is called diagonalizable if it can be written as $A = P^{-1}DP$, where D is a diagonal matrix. This is possible if and only if there is a basis $B = \{\mathbf{b_1}, \mathbf{b_2}, \dots, \mathbf{b_n}\}$ for $\mathbb{R}^n$ where $\mathbf{b_i}$ are eigenvectors of A. The corresponding eigenvalues sit along the diagonal of D, and the matrix $P = \begin{bmatrix}\mathbf{b_1} & \mathbf{b_2}, \dots, \mathbf{b_n}\end{bmatrix}$. Thus $P = P_B$, the "change of coordinate" matrix $P_B[\mathbf{x}]_B = \mathbf{x}$ and $(P_B)^{-1}(\mathbf{x}) = [\mathbf{x}]_B$. An orthogonal matrix is a square matrix for which $A^{-1} = A^t$; equivalently, an orthogonal matrix is a square matrix with orthonormal columns.

Theorem 1.6.8. (The Spectral Theorem) A (real) $n \times n$ matrix A is orthogonally diagonalizable if and only if A is symmetric.

Theorem 1.6.9. Let A be an $n \times n$ real symmetric matrix, and let

$$Q(\mathbf{y}) = \mathbf{y}A\mathbf{y}^t = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}y_iy_j.$$

Then we have
(a) $Q(\mathbf{y}) > 0$ for all vectors $\mathbf{y} \neq 0$ (A is positive definite) if and only if all eigenvalues of A are positive real numbers.
(b) $Q(\mathbf{y}) < 0$ for all vectors $\mathbf{y} \neq 0$ (A is negative definite) if and only if all eigenvalues of A are negative real numbers.

The last concept we will describe here is the singular values of a matrix that is used in Chapters 4 and 5. If A is an $m \times n$ matrix, then $n \times n$ matrix A^tA is a symmetric matrix and hence can be orthogonally diagonalized, by the spectral theorem. Not only are the eigenvalues of A^tA all real (Theorem 1.6.7), they are all nonnegative.

Definition 1.6.8. If A is an $m \times n$ matrix, the **singular values** of A are the square roots of the eigenvalues of A^tA and are denoted by $\sigma_1, \sigma_2, \dots, \sigma_n$. It is conventional to arrange the singular values so that $\sigma_1 \leq \sigma_2 \leq \dots \leq \sigma_n$.

We explain the following extremely important theorems and properties of singular values of a matrix A without proofs. For more details and proofs, you can see [60] pages 590 to 599.

Theorem 1.6.10. If A is an $m \times n$ matrix with singular values $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ and $\sigma_{r+1} = \sigma_{r+2} = \dots = \sigma_n = 0$. Then there exist an $m \times m$ orthogonal matrix U, an $n \times n$ orthogonal matrix V, and an $m \times n$ matrix

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}$$

such that

$$A = U\Sigma V^t,$$

where $1 \leq r \leq n$.

The factorisation of A as in Theorem 1.6.10 is called a **singular value decomposition** or **SVD** of A. The columns of U are called **left singular vectors** of A, and the columns of V are called **right singular vectors** of A (see [60]). The SVD provides new geometric insight into the effect of transformations. We have said several times that an $m \times n$ matrix transforms the unit sphere in $\mathbb{R}^n$ into an ellipsoid in $\mathbb{R}^m$. For this propose, we can express the following theorem which is proved in [60] on page 598.

Theorem 1.6.11. If A is an $m \times n$ matrix with rank r. Then the image of the unit sphere in $\mathbb{R}^n$ under the matrix transformation that maps $\mathbf{x}$ to $A\mathbf{x}$ is

- the surface of an ellipsoid in $\mathbb{R}^m$ if $r = n$.
- a solid ellipsoid in $\mathbb{R}^m$ if $r < n$ (The dimension of this ellipsoid is r).

In general, we can describe the effect of an $m \times n$ matrix A on the unit sphere in $\mathbb{R}^n$ in terms of the effect of each factor in its SVD, $A = U\Sigma V^t$, from right to left. Since V^t is an orthogonal matrix, it maps the unit sphere to itself. The $m \times n$ matrix Σ does two things:

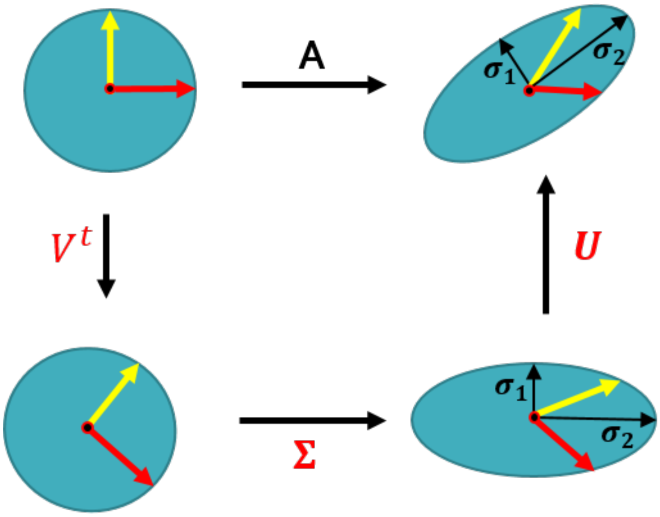

Figure 1.4: The singular value decomposition
$A = U\Sigma V^t$

- The diagonal entries $\sigma_{r+1} = \sigma_{r+2} = \ldots = \sigma_n = 0$ collapse $n-r$ of the dimensions of the unit sphere, leaving an r-dimensional unit sphere, which the nonzero diagonal entries $\sigma_1, \sigma_2, \ldots, \sigma_r$ then distort into an ellipsoid.
- The orthogonal matrix U then aligns the axes of this ellipsoid with the orthonormal basis vectors $\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_r}$ in $\mathbb{R}^m$

For more information about theorems and applications of singular values of the matrix A and its geometric approach, we refer you to [60] chapter 7 section 4.

1.7 Dilatation and Distortion of Linear Maps

Let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be a linear transformation. We define the **maximal stretching L(T)** and **minimal stretching** $\boldsymbol{\ell}(\mathbf{T})$ of T by

$$L(T) = \max_{|\mathbf{x}|=1} |T(\mathbf{x})|, \qquad \ell(T) = \min_{|\mathbf{x}|=1} |T(\mathbf{x})|. \tag{1.18}$$

The quantity $L(T)$ frequently goes by a different name, The **operator norm of T**, under the alternate notation $|T|$. For the composition ST of linear transformations $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ and $S : \mathbb{R}^m \longrightarrow \mathbb{R}^p$, it is true that

$$L(ST) \leq L(S)L(T), \qquad \ell(ST) \geq \ell(S)\ell(T). \tag{1.19}$$

A linear transformation $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is nonsingular if and only if $\ell(T) > 0$ in which event

$$L(T^{-1}) = \ell(T)^{-1}, \qquad \ell(T^{-1}) = L(T)^{-1}. \tag{1.20}$$

Recall that the nonsingular linear transformations of $\mathbb{R}^n$ form a group under composition, the **general linear group GL(n)**. A linear transformation $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is called an **orthogonal transformation** if $|T(\mathbf{x})| = |\mathbf{x}|$ for every $\mathbf{x}$ in $\mathbb{R}^n$ or, equivalently, that

$\langle T(\mathbf{x}), T(\mathbf{y})\rangle = \langle \mathbf{x}, \mathbf{y}\rangle$ for all $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$. The orthogonal transformations of $\mathbb{R}^n$ constitute a subgrup of GL(n), the **orthogonal group O(n)**. An element T of GL(n) belongs to $O(n)$ if and only if $T^{-1} = T^*$, where T^* denotes the adjoint of T, the unique linear transformation $T^* : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ that satisfies

$$\langle T(\mathbf{x}), \mathbf{y}\rangle = \langle \mathbf{x}, T^*(\mathbf{y})\rangle$$

for all $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$. If U is an element of $O(n)$ (we represent orthogonal transformations by the letters U and V) then $L(U) = \ell(U) = 1$, while the determinant of U is either 1 or -1. If an orthogonal transformation U has $\det(U) = 1$, we call U a **rotation** of $\mathbb{R}^n$.
The **special orthogonal group SO(n)** is the subgroup of $O(n)$ consisting of all such rotations. It is not hard to see that

$$L(VTU) = L(T), \qquad \ell(VTU) = \ell(T), \tag{1.21}$$

for any linear transformation $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ whenever $U \in O(n)$ and $V \in O(m)$. Let U be an open set in $\mathbb{R}^n$, and $\mathbf{x} \in U$. We define the maximal stretching $L_{\mathbf{f}}(\mathbf{x})$ and the minimal stretching $\ell_{\mathbf{f}}(\mathbf{x})$ at $\mathbf{x}$ of a function $\mathbf{f} : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n$ by :

$$L_{\mathbf{f}}(\mathbf{x}) = \limsup_{\mathbf{h}\to 0} \frac{\|\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x})\|}{\|\mathbf{h}\|},$$

$$\ell_{\mathbf{f}}(\mathbf{x}) = \liminf_{\mathbf{h}\to 0} \frac{\|\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x})\|}{\|\mathbf{h}\|}.$$

Theorem 1.7.1. If $\mathbf{f} : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is differentiable at $\mathbf{x} \in U$, then $L_{\mathbf{f}}(\mathbf{x}) = L(D\mathbf{f}(\mathbf{x})) = L(\mathbf{f}'(\mathbf{x}))$ and $\ell_{\mathbf{f}}(\mathbf{x}) = \ell(D\mathbf{f}(\mathbf{x})) = \ell(\mathbf{f}'(\mathbf{x}))$.

Proof. We prove the first part. Recall that if $T : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is a linear transformation then

$$L(T) = \max_{|\mathbf{x}|=1} |T(\mathbf{x})|, \qquad \ell(T) = \min_{|\mathbf{x}|=1} |T(\mathbf{x})|.$$

We can chose a small $\mathbf{h}$ such that $\mathbf{x} + \mathbf{h} \in U$. By Theorem 1.5.1, we can write

$$|\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x})| = |\mathbf{f}'(\mathbf{x})\mathbf{h} + \mathbf{h}\epsilon(\mathbf{h})| \leq |\mathbf{h}| L(\mathbf{f}'(\mathbf{x})) + |\mathbf{h}||\epsilon(\mathbf{h})|.$$

So, we have

$$L_{\mathbf{f}}(\mathbf{x}) \leq \limsup_{\mathbf{h}\to 0} (L(\mathbf{f}'(\mathbf{x})) + |\epsilon(\mathbf{h})|) = L(\mathbf{f}'(\mathbf{x})).$$

Also, we can chose $\mathbf{h} \in \mathbb{S}^{n-1}$ such that $L(\mathbf{f}'(\mathbf{x})) = |\mathbf{f}'(\mathbf{x})\mathbf{h}|$. Then by definition of derivative we have

$$L_{\mathbf{f}}(\mathbf{x}) \geq \lim_{t\to 0} \frac{|\mathbf{f}(\mathbf{x}+t\mathbf{h}) - \mathbf{f}(\mathbf{x})|}{t} = |\mathbf{f}'(\mathbf{x})\mathbf{h}| = L(\mathbf{f}'(\mathbf{x})),$$

and the proof is completed. The second part of proof is the same as the first part. ♠

Corollary 1.7.1.1. If $\mathbf{f} : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is differentiable at $\mathbf{x} \in U$, then $L_{\mathbf{f}}(\mathbf{x}) = |\mathbf{f}'(\mathbf{x})|$.

Let $n \geq 2$. A linear transformation $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is **conformal** if T is nonsingular and preserves Euclidean angles, in the sense that

$$\theta[T(\mathbf{x}), T(\mathbf{y})] = \theta(\mathbf{x}, \mathbf{y})$$

for all nonzero vectors $\mathbf{x}$ and $\mathbf{y}$ in $\mathbb{R}^n$. For more information you can see [30].

Theorem 1.7.2. Let $n \geq 0$. The following statements concerning a linear transformation $T \in GL(n)$ are equivalent :

1. T is conformal;
2. $T = \lambda U$, where λ is a positive number and U belongs to O(n);
3. $\|T\|^n = |\det(T)|$;
4. $T^*T = |\det(T)|^{2/n} I$.

where I denotes the identity matrix.

The proof of this theorem exists on pages (19) and (20) of the book [30].
Now we want to introduce properties of distortion and two new definitions of inner and outer distortion. We recall some definitions of linear algebra. When a linear transformation $S : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ enjoys the property that $S^* = S$, we say that S is symmetric or self-adjoint. For example, if $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is an arbitrary linear transformation, then both T^*T and TT^* are symmetric or self-adjoint. By the Spectral Theorem 1.6.8, we can say, S is a symmetric linear transformation of $\mathbb{R}^n$, then there exists $U \in O(n)$ such that $D = U^{-1}SU$ has the form

$$D(\mathbf{x}) = (\lambda_1 x_1, \lambda_2 x_2, \ldots, \lambda_n x_n).$$

In particular, by Theorem 1.6.7 the eigenvalues of S (that is, $\lambda_1, \lambda_2, \ldots, \lambda_n$) are real. The previous transformation D is called a **diagonal transformation**. We observe that

$$\min_{1\leq i\leq n} |\lambda_i| \leq |D(\mathbf{x})| = \sqrt{\lambda_1^2 x_1^2 + \lambda_2^2 x_2^2 + \ldots + \lambda_n^2 x_n^2} \leq \max_{1\leq i\leq n} |\lambda_i| \tag{1.22}$$

whenever $|\mathbf{x}| = 1$, so $L(D) = \max_{1\leq i\leq n} |\lambda_i|$ and $\ell(D) = \min_{1\leq i\leq n} |\lambda_i|$. Also, we have $\det(D) = \lambda_1\lambda_2 \ldots \lambda_n$. To say that a linear transformation $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is positive definite (respectively, positive semidefinite) means that $\langle \mathbf{x}, T(\mathbf{x})\rangle > 0$ (respectively, $\langle \mathbf{x}, T(\mathbf{x})\rangle \geq 0$) for every nonzero vector $\mathbf{x}$ in $\mathbb{R}^n$. For instance, if $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is an arbitrary linear transformation, then T^*T and TT^* are positive semidefinite. If, in addition, T is nonsingular, then T^*T and TT^* are actually positive definite. By Theorem 1.6.9, any eigenvalue of a positive definite (respectively, positive semidefinite) linear transformation is positive (respectively, nonnegative). You can see [42] or [30].
Suppose that a linear transformation $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is both symmetric and positive semidefinite. We can list the eigenvalues of T in the manner $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_n \geq 0$ and assert the existence of U in $O(n)$ for which $U^{-1}TU = D$, where $D(\mathbf{x}) = (\lambda_1 x_1, \lambda_2 x_2, \ldots, \lambda_n x_n)$. According to formula (1.21), we can write that

$$L(T) = L(U^{-1}TU) = L(D) = \lambda_1.$$

Similarly, $\ell(T) = \lambda_n$ and $\det(T) = \lambda_1\lambda_2 \ldots \lambda_n$. We can explain the next theorem.

Theorem 1.7.3. Any linear transformation $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ can be factored as $T = PU$, where U belongs to O(n) and P is both symmetric and positive semidefinite.

The linear transformation P in Theorem 1.7.3 is uniquely determined by T: denoting by I the identity transformation of $\mathbb{R}^n$, we compute

$$TT^* = (PU)(PU)^* = PUU^*P^* = PIP = P^2,$$

and learn that P is the (known to be unique) symmetric, positive semidefinite square root of TT^* (see [30], page 10). This observation entitles us to list the eigenvalues of P as $\lambda_1^{1/2}, \lambda_2^{1/2}, \ldots, \lambda_n^{1/2}$, where $\lambda_1 \geq \lambda_2 \geq \ldots \lambda_n \geq 0$ are the eigenvalues of TT^*. Therefore, there is an orthogonal transformation V such that $V^{-1}PV = D$, with $y = D(\mathbf{x})$ given as follows:

$$y_1 = \lambda_1^{1/2} x_1 \,, \ldots, \; y_n = \lambda_n^{1/2} x_n,$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. If T is nonsingular then all eigenvalues of TT^* are positive, which makes it apparent that D maps the sphere $\mathbb{S}^{n-1}$ bijectively to the ellipsoid E whose equation is $(y_1^2/\lambda_1) + (y_2^2/\lambda_2) + \ldots + (y_n^2/\lambda_n) = 1$. Because $U(\mathbb{S}^{n-1}) = V(\mathbb{S}^{n-1}) = \mathbb{S}^{n-1}$, we can write

$$T(\mathbb{S}^{n-1}) = PU(\mathbb{S}^{n-1}) = P(\mathbb{S}^{n-1}) = PV(\mathbb{S}^{n-1}) = VD(\mathbb{S}^{n-1}) = V(E).$$

So, we can say

$$L(T) = L(PU) = L(P) = \sqrt{\lambda_1} = \sigma_1$$

and

$$\ell(T) = \ell(PU) = \ell(P) = \sqrt{\lambda_n} = \sigma_n,$$

where σ_1 and σ_n are singular values of T. A further result of Theorem 1.7.3 is the useful inequality

$$\ell(T)^n \leq |\det(T)| \leq L(T)^n, \tag{1.23}$$

valid for any linear transformation $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$. We know that

$$|\det(T)| = \det(P) = \sqrt{\lambda_1 \lambda_2 \ldots \lambda_n},$$

in which $\lambda_1, \lambda_2, \ldots, \lambda_n$ are eigenvalues of TT^* (see [30]). Now, we can express a theorem such that it summarizes these above discussions.

Theorem 1.7.4. Let $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be a linear transformation, and list the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ of TT^* in nonincreasing order. Then there exist orthogonal transformations $U, V \in O(n)$ such that $VTU = D$, where

$$D(\mathbf{x}) = (\lambda_1^{1/2} x_1, \lambda_2^{1/2} x_2, \ldots, \lambda_n^{1/2} x_n).$$

Definition 1.7.1. We can define the **linear dilatation** $\mathcal{H}(T)$, the **inner dilatation** $\mathcal{H}_I(T)$, and the **outer dilatation** $\mathcal{H}_O(T)$ of a nonsingular linear transformation $T : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ as

$$\mathcal{H}(T) = \frac{L(T)}{\ell(T)}, \qquad \mathcal{H}_I(T) = \frac{|\det(T)|}{\ell(T)^n}, \qquad \mathcal{H}_O(T) = \frac{L(T)^n}{|\det(T)|}. \tag{1.24}$$

If T is singular it is customary to set $\mathcal{H}(T) = \mathcal{H}_I(T) = \mathcal{H}_O(T) = \infty$ (see [30]).

In geometric terms, $\mathcal{H}(T)$ measures the eccentricity of the ellipsoid $T(\mathbb{S}^{n-1})$ while $\mathcal{H}_I(T)$ and $\mathcal{H}_O(T)$ relate the volume of $T(\mathbf{B^n})$ to the volumes of the balls centered at the origin that are, respectively, inscribed in and circumscribed about $T(\mathbb{S}^{n-1})$. We remark that

$$\mathcal{H}(T^{-1}) = \mathcal{H}(T), \qquad \mathcal{H}_I(T^{-1}) = \mathcal{H}_O(T), \qquad \mathcal{H}_O(T^{-1}) = \mathcal{H}_I(T) \tag{1.25}$$

and that

$$\mathcal{H}(VTU) = \mathcal{H}(T), \qquad \mathcal{H}_I(VTU) = \mathcal{H}_I(T), \qquad \mathcal{H}_O(VTU) = \mathcal{H}_O(T) \tag{1.26}$$

if V and U are members of $O(n)$. Because of (1.26) and Theorem 1.7.4, we can show that

$$1 \leq \mathcal{H}_O(T) \leq \mathcal{H}_I(T)^{n-1}, \qquad 1 \leq \mathcal{H}_I(T) \leq \mathcal{H}_O(T)^{n-1} \tag{1.27}$$

and also that

$$1 \leq \ \min\{\mathcal{H}_I(T), \mathcal{H}_O(T)\} \ \leq \ \mathcal{H}(T)^{n/2} \ \leq \ \max\{\mathcal{H}_I(T), \mathcal{H}_O(T)\} \ \leq \ \mathcal{H}(T)^{n-1} \tag{1.28}$$

We can see that as a result of (1.28) $\mathcal{H}(T) = \mathcal{H}_I(T) = \mathcal{H}_O(T)$ when $n = 2$ (see [42] and [30]).

Definition 1.7.2. Let Ω be a domain in $\mathbb{R}^n$, let $\mathbf{f} : \Omega \longrightarrow \mathbb{R}^n$ be a continuous injection. The **linear distortion or dilatation** of $\mathbf{f}$ at the point $\mathbf{x}$ in A is the quantity $\mathcal{H}_{\mathbf{f}}(\mathbf{x})$ or $\mathcal{H}(\mathbf{x}, \mathbf{f})$ defined by

$$\mathcal{H}_{\mathbf{f}}(\mathbf{x}) = \mathcal{H}(\mathbf{x}, \mathbf{f}) = \limsup_{r \to 0} \frac{L_f(\mathbf{x}, r)}{\ell_f(\mathbf{x}, r)}, \tag{1.29}$$

where for $0 < r < \operatorname{dist}(\mathbf{x}, \partial\Omega)$ we set

$$L_f(\mathbf{x}, r) = \max_{\|h\|=r} \|\mathbf{f}(\mathbf{x}+h) - \mathbf{f}(\mathbf{x})\|, \qquad \ell_f(\mathbf{x}, r) = \min_{\|h\|=r} \|\mathbf{f}(\mathbf{x}+h) - \mathbf{f}(\mathbf{x})\|.$$

In the other words,

$$\mathcal{H}_{\mathbf{f}}(\mathbf{x}) = \mathcal{H}(\mathbf{x}, \mathbf{f}) = \limsup_{r \to 0} \frac{\max_{\|h\|=r} \|\mathbf{f}(\mathbf{x}+h) - \mathbf{f}(\mathbf{x})\|}{\min_{\|h\|=r} \|\mathbf{f}(\mathbf{x}+h) - \mathbf{f}(\mathbf{x})\|}. \tag{1.30}$$

Let Ω and Ω' be domains in $\mathbb{R}^n$ with $n \geq 2$, and let $\mathbf{f}$ be a homeomorphism of Ω onto Ω'. The **linear distortion of function f** is defined

$$\mathcal{H}(\mathbf{f}) = \sup\{\mathcal{H}_{\mathbf{f}}(\mathbf{x}) : \ \mathbf{x} \in \Omega\}, \tag{1.31}$$

where $\mathcal{H}_{\mathbf{f}}(\mathbf{x}) \geq 1$ is the linear distortion of $\mathbf{f}$ at $\mathbf{x}$ as defined at (1.29). $\mathbf{f}$ is a conformal mapping if and only if $\mathcal{H}(\mathbf{f}) = 1$ (see [30] p.77).
At points where the differential $D\mathbf{f}(a)$ exists and is nonsingular we clearly have the limit in (1.30) existing. This shows the following equation is correct.

$$\max_{|h|=1} |D\mathbf{f}(\mathbf{x})h| \leq \mathcal{H}(\mathbf{x}, \mathbf{f}) \min_{|h|=1} |D\mathbf{f}(\mathbf{x})h| \tag{1.32}$$

Geometrically this means that the differential $D\mathbf{f}(a) : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ maps the unit sphere to an ellipsoid for which the ratio of the lengths of the largest and smallest semiaxes, the **eccentricity**, is equal to $\mathcal{H}(a, \mathbf{f})$. If we define $C\mathbf{f}(\mathbf{x}) = D^t\mathbf{f}(\mathbf{x}).D\mathbf{f}(\mathbf{x})$, this matrix is referred to as the **right Cauchy-Green strain tensor**. The positive square roots of its eigenvalues,

the singular values of $D\mathbf{f}(a)$, are the **principal stretchings**. The corresponding eigenvectors are called the **principal directions** of the deformation at $a \in \Omega$ (see [42]).
Our **distortion tensor** will be represented by the positive symmetric matrix defined by

$$G(\mathbf{x}) = J_{\mathbf{f}}(\mathbf{x})^{-\frac{2}{n}} D^t\mathbf{f}(\mathbf{x}) D\mathbf{f}(\mathbf{x}). \tag{1.33}$$

Let

$$0 < \mu_1(\mathbf{x}) \leq \mu_2(\mathbf{x}) \leq \ldots \leq \mu_n(\mathbf{x}) \tag{1.34}$$

denote the positive square roots of the eigenvalues of the direction tensor $G(\mathbf{x})$. Now returning to equation (1.32) one easily obtains the identity

$$\mathcal{H}_{\mathbf{f}}(\mathbf{x}) = \mathcal{H}(\mathbf{x}, \mathbf{f}) = \frac{\mu_n(\mathbf{x})}{\mu_1(\mathbf{x})} \tag{1.35}$$

Thus $\mathcal{H}_{\mathbf{f}}(\mathbf{x}) = \mathcal{H}(\mathbf{x}, \mathbf{f})$ measures the maximum possible relative distortion of linear objects (e.g. length of curves) at the infinitesimal level at $a \in \Omega$ (see [42]). Now we recall the following result of Rademacher-Stepanoff.

Theorem 1.7.5. (Rademacher-Stepanoff) Let $\mathbf{f} : \Omega \longrightarrow \mathbb{R}^n$ be an arbitrary mapping defined on an open set $\Omega \subset \mathbb{R}^n$. Consider the set

$$E = \Big\{ a \in \Omega \; : \; \limsup_{x \to a} \frac{|\mathbf{f}(\mathbf{x}) - \mathbf{f}(a)|}{|\mathbf{x} - a|} < \infty \Big\}.$$

Then E is a Lebesgue measurable set and $\mathbf{f}$ is differentiable at almost every point of E.

Proof. Proof of this theorem exists in Federer [26], page 218. ♠

In this thesis, we use the term **embedding** to mean a continuous injection. Also, we remark that in the Rademacher-Stepanoff theorem the mapping $\mathbf{f}$ need not even be assumed measurable. Let Ω be an open subset of $\mathbb{R}^n$, $n \geq 2$, and let $\mathbf{f} : \Omega \longrightarrow \mathbb{R}^n$ be an embedding. Then the linear distortion function $\mathcal{H}_{\mathbf{f}}(\mathbf{x}) = \mathcal{H}(\mathbf{x}, \mathbf{f})$ defined in (1.30) is a Borel function (see [42] and [30]). The first step towards establishing the regularity properties of mapping of finite distortion is given by the following differentiability theorem.

Theorem 1.7.6. Every embedding $\mathbf{f} : \Omega \longrightarrow \mathbb{R}^n$ is differentiable at almost every point of the set

$$E = \{ \mathbf{x} \in \Omega \; : \; \mathcal{H}(\mathbf{x}, \mathbf{f}) < \infty \}.$$

Moreover, for almost every point $\mathbf{x} \in E$ we have

$$|D\mathbf{f}(\mathbf{x})| \leq \mathcal{H}(\mathbf{x}, \mathbf{f}) |J_{\mathbf{f}}(\mathbf{x})|^{\frac{1}{n}} \tag{1.36}$$

and

$$\max_{|h|=1} |D\mathbf{f}(\mathbf{x})h| \leq \mathcal{H}(\mathbf{x}, \mathbf{f}) \min_{|h|=1} |D\mathbf{f}(\mathbf{x})h| \tag{1.37}$$

A proof can be found in [42], page 103 and 104. We can explain the following corollary that will be useful.

Corollary 1.7.6.1. Every continuous locally injective mapping $\mathbf{f} : \Omega \longrightarrow \mathbb{R}^n$ whose linear distortion function is finite almost everywhere is differentiable almost everywhere in Ω.

1.8 Tensor Product

In this section, we explain definitions and properties of another method of the product of vectors. We are familiar with the inner and cross products, but we introduce another type of product for the other concepts.

Definition 1.8.1. (tensor product of vectors) If $\mathbf{u}$ and $\mathbf{v}$ are vectors on length m and n, respectively, their tensor product $\mathbf{u}\otimes\mathbf{v}$ is defined as an $m\times n$ matrix defined by $(\mathbf{u}\otimes\mathbf{v})_{ij} = u_i v_j$. In other words,

$$\mathbf{u}\otimes\mathbf{v} = \mathbf{u}\mathbf{v}^{\mathbf{t}}.$$

This is called the **outer product** of vectors $\mathbf{u}$ and $\mathbf{v}$. The outer product contrasts with the dot product, which takes as input a pair of coordinate vectors and products a scalar. For complex vectors, it is customary to use the conjugate transpose of $\mathbf{v}$, so in this case we can write

$$\mathbf{u}\otimes\mathbf{v} = \mathbf{u}\mathbf{v}^{*}.$$

Theorem 1.8.1. If A is an $m\times n$ matrix then A is a rank-one matrix if and only if there are two vectors $\mathbf{u}$ and $\mathbf{v}$ with m and n components, respectively such that

$$A = \mathbf{u}\otimes\mathbf{v}. \tag{1.38}$$

It is easy to show that, if the rank of a real symmetric matrix is equal to 1, then the diagonal elements of the matrix can not be all zero. We leave this proof for readers. Now, we want to explain another result for rank-one matrices. This theorem will be proved in 3-dimensions, the general proof for n-dimensions will be similar. It is used in the analysis of rank-one convexity for the linear distortion later.

Theorem 1.8.2. Let A be a 3×3 rank-one symmetric matrix. Then A can be written as

$$A = \lambda(1,x,y)\otimes(1,x,y). \tag{1.39}$$

In the other words, the matrix A can be expressed by

$$A = \lambda\mathbf{u}\otimes\mathbf{u} = \lambda\mathbf{u}\mathbf{u}^{t},$$

where $\mathbf{u} = (1,x,y)$ is a vector in $\mathbb{R}^3$.

Proof. Since A is a rank-one matrix, by Theorem 1.8.1, there are two vectors $\mathbf{v}$ and $\mathbf{w}$ in $\mathbb{R}^3$ with

$$A = \mathbf{v}\otimes\mathbf{w} = \mathbf{v}\mathbf{w}^{t}.$$

The vectors $\mathbf{v_1} = \frac{\mathbf{v}}{\|\mathbf{v}\|}$ and $\mathbf{w_1} = \frac{\mathbf{w}}{\|\mathbf{w}\|}$ are unit vectors. So, $\mathbf{v} = \|\mathbf{v}\|\mathbf{v_1}$ and $\mathbf{w} = \|\mathbf{w}\|\mathbf{w_1}$. Thus,

$$A = \mathbf{v}\mathbf{w}^{t} = \|\mathbf{v}\|\|\mathbf{w}\|\mathbf{v_1}\mathbf{w_1}^{t}.$$

If $\|\mathbf{v}\|\|\mathbf{w}\| = c\in\mathbb{R}$, then $A = c\mathbf{v_1}\mathbf{w_1}^{t}$. Since A is symmetric, we get

$$\mathbf{v_1}\mathbf{w_1}^{t} = \mathbf{w_1}\mathbf{v_1}^{t}, \tag{1.40}$$

where $\mathbf{v_1}$ and $\mathbf{w_1}$ are unit vectors. If (1.40) is multiplied on the left by $\mathbf{w_1}^t$ and on the right by $\mathbf{v_1}$,

$$\begin{aligned} &\Rightarrow \quad \mathbf{w_1}^t(\mathbf{v_1}\mathbf{w_1}^t)\mathbf{v_1} = \mathbf{w_1}^t(\mathbf{w_1}\mathbf{v_1}^t)\mathbf{v_1} \\ &\Rightarrow \quad (\mathbf{w_1}^t\mathbf{v_1})(\mathbf{w_1}^t\mathbf{v_1}) = (\mathbf{w_1}^t\mathbf{w_1})(\mathbf{v_1}^t\mathbf{v_1}) \\ &\Rightarrow \quad (\mathbf{w_1}^t\mathbf{v_1})^2 = \|\mathbf{v}\|^2\|\mathbf{w}\|^2 \\ &\Rightarrow \quad (\mathbf{w_1}^t\mathbf{v_1})^2 = 1 \times 1 = 1, \end{aligned}$$

which implies

$$|\mathbf{v_1}\mathbf{w_1}| = 1 = \|\mathbf{v_1}\|\|\mathbf{w_1}\|.$$

This is the equality case of Cauchy-Schwarz which holds if and only if $\mathbf{v_1}$ and $\mathbf{w_1}$ are linearly dependent. The two are both unit vectors so it follows that $\mathbf{v_1} = \pm\mathbf{w_1}$. Therefore, we have

$$A = \pm c\ (\mathbf{v_1} \otimes \mathbf{v_1}) = \pm c\ (\mathbf{v_1}\mathbf{v_1}^t).$$

Let $\mathbf{v_1} = (c_1, c_2, c_3)$, then $\mathbf{v_1} = c_1(1, \frac{c_2}{c_1}, \frac{c_3}{c_1})$. So,

$$A = \pm c\ (c_1c_2)(1, \frac{c_2}{c_1}, \frac{c_3}{c_1}) \otimes (1, \frac{c_2}{c_1}, \frac{c_3}{c_1}).$$

If $\pm c\ (c_1c_2) = \lambda$, $\frac{c_2}{c_1} = x$ and $\frac{c_3}{c_1} = y$, then

$$A = \lambda(1, x, y) \otimes (1, x, y),$$

as required. These arguments complete the proof of Theorem 1.8.2. ♠

1.9 $\mathbf{L^p}$ Spaces vs $\ell^{\mathbf{p}}$ Spaces

In section (1.1), we explained the L^p and $\ell^{\mathbf{p}}$ norms. Now, we can extend these concepts to functions and their spaces.

Definition 1.9.1. Let $(X, \mathfrak{M}, \mu)$ be a measure space and $1 \leq p < \infty$. The $\mathbf{L^p(X)}$**-space** is a set of equivalence classes of measurable functions $f : X \longrightarrow \mathbb{R}$ such that

$$\|f\|_p = \Big(\int_X |f|^p \mathrm{d}\mu\Big)^{\frac{1}{p}} < \infty. \tag{1.41}$$

The $\|f\|_p$ is called the L^p-norm of f. The notation $L^p(X)$ assumes that the measure μ is understood (see [33] and [56]).

Definition 1.9.2. Let Ω be an open set in the Euclidean space $\mathbb{R}^n$ and $f : \Omega \longrightarrow \mathbb{R}$ be a Lebesgue measurable function. If $1 \leq p \leq \infty$ the function f satisfies

$$\int_A |f|^p \mathrm{dx} < \infty. \tag{1.42}$$

This means $f \in L^p(A)$ for all compact subsets $A \in \Omega$. f is called locally integrable and the set of all functions is denoted by $L^p_{loc}(\Omega)$ (see [33] and [56]).

The L^p functions are allowed to take values of $\pm\infty$, like any measurable functions. However, it follows from the definition of an L^p function that it takes finite values almost everywhere, so there is no harm in restricting to L^p functions $X \longrightarrow \mathbb{R}$.
It is obvious that any scalar multiple of an L^p function is again L^p. Moreover, if f and g are L^p functions, then by Minkowski's inequality

$$\|f+g\|_p \leq \|f\|_p + \|g\|_p < \infty, \tag{1.43}$$

so $f+g$ is an L^p function. Thus the set of L^p functions forms a vector space. For $1 \leq p < \infty$, the $L^p(X)$-space is a Banach space. So, we can say, $L^p(X)$ for $1 \leq p < \infty$ is the Banach space of measurable functions f with $|f|^p$ integrable in X (see [33], [23] and [16]).

Example 1.9.1. Let $f : [0,1] \longrightarrow \mathbb{R}$ be a function

$$f(x) = \frac{1}{x},$$

where the value of $f(0)$ is immaterial. Then by the monotone convergence theorem,

$$\int_{[0,1]} |\ f\ |\ dm = \lim_{t\to 0^+} \int_{[t,1]} \frac{1}{x} dm(x) = \lim_{t\to 0^+} \Big[\log x\Big]_t^1 = \infty.$$

Thus, the function f is not L^1. It is easy to see that f is not L^p function for any $p \in [1,\infty)$. If

$$f(x) = \frac{1}{\sqrt{x}},$$

then,

$$\int_{[0,1]} |\ f\ |\ dm = \lim_{t\to 0^+} \int_{[t,1]} \frac{1}{\sqrt{x}} dm(x) = \lim_{t\to 0^+} \Big[2\sqrt{x}\Big]_t^1 = 2.$$

In fact,

$$\int_{[0,1]} \frac{1}{x^r} dm(x) = \begin{cases} \infty & \text{if } \ r \geq 1 \\ \dfrac{1}{1-r} & \text{if } \ r < 1. \end{cases}$$

It follows that the function $f(x) = \frac{1}{x^r}$ is L^p function if and only if $pr < 1$. For instance, $f(x) = \frac{1}{\sqrt{x}}$ is L^p for all $p \in [1,2)$, but is not L^p for any $p \in [2,\infty)$.

♣

It is not difficult to see that if $(X, \mathfrak{M}, \mu)$ is a measure space and $1 \leq p \leq q < \infty$ and $\mu(X) = 1$, then

$$\|f\|_p \leq \|f\|_q, \tag{1.44}$$

for every measurable function f. More generally, if $0 < \mu(X) < \infty$, then

$$\|f\|_p \leq \mu(X)^r \|f\|_q, \tag{1.45}$$

for every measurable function f, where $r = (1/p) - (1/q)$, and hence every L^q function is also L^p.

Definition 1.9.3. If $p \in [1, \infty)$, the $\ell^{\mathbf{p}}$**-norm** of a sequence $\{a_n\}$ of real numbers is defined by the formula

$$\|\{a_n\}\|_p = \left(\sum_{n \in \mathbb{N}} |a_n|^p \right)^{\frac{1}{p}}. \tag{1.46}$$

An $\ell^{\mathbf{p}}$ **sequence** is a sequence $\{a_n\}$ of real number for which

$$\sum_{n \in \mathbb{N}} |a_n|^p < \infty. \tag{1.47}$$

Example 1.9.2. Recall from calculus and analysis that the p-series

$$\sum_{n=1}^{\infty} \frac{1}{n^p},$$

converges if and only if $p > 1$. It follows that sequence $\{1/n^p\}$ is ℓ^1 if and only if $p > 1$. For instance,

$$\{\frac{1}{n^2}\} \text{ is } \ell^1, \text{ but } \{\frac{1}{\mathrm{n}}\} \text{ and } \{\frac{1}{\sqrt{\mathrm{n}}}\} \text{ are not.}$$

Moreover, since $(1/n^r)^p = 1/n^{rp}$, we find that $\{1/n^r\}$ is ℓ^p if and only if $p > 1/r$. Therefore,

$$\{\frac{1}{n^2}\} \text{ is } \ell^2, \text{ but not } \ell^1 \text{ and } \{\frac{1}{\sqrt{\mathrm{n}}}\} \text{ is } \ell^3, \text{ but not } \ell^2.$$

♣

It is easy to see that, for each $1 \leq p < \infty$ the set ℓ^p is a vector space and $\|.\|_p$ is a norm on it. Also, we can prove that if $1 \leq p < q < \infty$, then every ℓ^p sequence is also ℓ^q [33].

Definition 1.9.4. Let $(X, \mathfrak{M}, \mu)$ be a measure space, let $\{f_n\}$ be a sequence of measurable functions on X and let $P \in [1, \infty)$ (see [23] and [16]).

1. A sequence $\{f_n\}$ is an $\mathbf{L}^{\mathbf{p}}$ **Cauchy sequence** if for every $\epsilon > 0$ there exists an $N \in \mathbb{N}$, so that
$$i, j \geq N \Longrightarrow \|f_i - f_j\| < \epsilon. \tag{1.48}$$
2. A sequence $\{f_n\}$ has **bounded** $\mathbf{L}^{\mathbf{p}}$**-variation** if
$$\sum_{n \in \mathbb{N}} \|f_{n+1} - f_n\| < \infty. \tag{1.49}$$
3. A sequence $\{f_n\}$ **converges in** $\mathbf{L}^{\mathbf{p}}$ to a measurable function f, if
$$\lim_{n \longrightarrow \infty} \|f_n - f\|_p = 0. \tag{1.50}$$

Theorem 1.9.1. Let $(X, \mathfrak{M}, \mu)$ be a measure space, let $\{f_n\}$ be a sequence of measurable functions on X with bounded L^p-variation. Then $\{f_n\}$ converges pointwise almost everywhere to the measurable function f, and $f_n \longrightarrow f$ in L^p.

Definition 1.9.5. Let $(X, \mathfrak{M}, \mu)$ be a measure space, and f be a measurable function on X. The $\mathbf{L}^\infty$**-norm** of f is defined as follows

$$\|f\|_\infty := \inf\{c > 0 \ : \ |f| \le c \text{ almost everywhere}\}. \tag{1.51}$$

Then $\|f\|_{L^\infty} = \infty$ if the set of the right-hand side is empty (see [33]). The function f is called an $\mathbf{L}^\infty$ **function** if $\|f\|_\infty < \infty$.

Not only does the set

$$\{c > 0 \ : \ |f| \le c \text{ almost everywhere}\}, \tag{1.52}$$

have an infimum, but it also has a minimum. If $|f| \le c+1/n$ almost everywhere for all $n \in \mathbb{N}$, then it follows that $|f| \le c$ almost everywhere. The L^∞-norm $\|f\|_\infty$ is sometimes called the **essential supremum** of $|f|$, and L^∞ functions are sometimes said to be **essentially bounded** or **bounded almost everywhere**. Note that a continuous function on $\mathbb{R}$ is an L^∞ function if and only if it is bounded, in which case $\|f\|_\infty$ is equal to the supremum of $|f|$.

Definition 1.9.6. Let $\{a_n\}$ be a sequence of real numbers. The ℓ^∞**-norm** of $\{a_n\}$ is defined as follows:

$$\|\{a_n\}\|_\infty = \sup_{n\in\mathbb{N}} |a_n|. \tag{1.53}$$

Therefore, an ℓ^∞ sequence is bounded. Note that if $p \in [1, \infty)$, then any ℓ^p sequence must be ℓ^∞, because any ℓ^p sequence must converge to zero (see [33] and [56]).

1.10 Weak Derivative and Sobolev Spaces

We recall Definition 1.4.3. Let Ω be a nonempty subset of $\mathbb{R}^n$. We define the function spaces $C^k(\Omega, \mathbb{R}^m) = C^k(\Omega)$ for $k = 0, 1, 2, \dots$ by: $C^0(\Omega, \mathbb{R}^m) = C^0(\Omega)$ is the class of all continuous functions $\mathbf{f} : \Omega \longrightarrow \mathbb{R}^m$. For $k \ge 1$, $C^k(\Omega)$ consists of the functions $\mathbf{f}$ in $C^0(\Omega)$ such that all partial derivatives

$$\frac{\partial^\ell \mathbf{f}}{\partial x_{i_\ell} \dots \partial x_{i_2} \partial x_{i_1}} = \frac{\partial}{\partial x_{i_\ell}} \dots \frac{\partial}{\partial x_{i_2}} \frac{\partial}{\partial x_{i_1}} \mathbf{f}(x_1, x_2, \dots, x_n), \tag{1.54}$$

with $1 \le \ell \le k$ and $1 \le i_1, i_2, \dots, i_\ell \le n$ are defined and continuous in Ω (see [30]). The class $C^\infty(\Omega, \mathbb{R}^m) = C^\infty(\Omega)$ is the intersection of the classes $C^k(\Omega)$, $k = 0, 1, 2, \dots$,

$$C^\infty(\Omega) = \bigcap_{k=0}^{\infty} C^k(\Omega). \tag{1.55}$$

Assume that a function $\mathbf{f}$ belomgs to $C^0(\Omega)$, we employ **supp(f)** to designate its support. This means the closure relative to Ω of the set of all $\mathbf{x} \in \Omega$ such that $\mathbf{f}(\mathbf{x}) \ne 0$ (see [30]). So,

$$\operatorname{supp}(\mathbf{f}) = \overline{\{\mathbf{x} \in \Omega : \ \mathbf{f}(\mathbf{x}) \ne 0\}}. \tag{1.56}$$

For $0 \le k \le \infty$, we use $C_0^k(\Omega, \mathbb{R}^m) = C_0^k(\Omega)$ to denote the subclass to $C^k(\Omega)$ comprising those functions $\mathbf{f}$ for which the support of $\mathbf{f}$ is a compact subset of Ω. Those functions in $C^\infty(\Omega)$ whose support is compactly contained in Ω are called the **test functions**, and

$$C_0^\infty(\Omega, \mathbb{R}^m) = C_0^\infty(\Omega) = C^\infty(\Omega) \cap C_0^k(\Omega) = \{\mathbf{f} : \ \mathbf{f} \text{ is a compactly supported smooth function}\}. \tag{1.57}$$

Example 1.10.1. Let Ω be the unit ball $B(0,1)$ and $\mathbf{f} : \Omega \longrightarrow \mathbb{R}$ be a function given by

$$\mathbf{f}(\mathbf{x}) = 1 - |\mathbf{x}|.$$

It is obvious that for all $\mathbf{x} \in B(0,1)$, $\mathbf{f}(\mathbf{x}) > 0$, so, the supp($\mathbf{f}$)= $\overline{B(0,1)}$ [46].

♣

Example 1.10.2. Let $f : \mathbb{R} \longrightarrow \mathbb{R}$ be a function given by

$$f(x) = \begin{cases} x^2 & \text{if } x \geq 0 \\ -x^2 & \text{if } x < 0. \end{cases}$$

The function $\mathbf{f}$ is $C^1(\mathbb{R})$ but not $C^2(\mathbb{R})$, (see [46]).

♣

Let us start with the motivation for the definition of weak derivatives. Let $\Omega \in \mathbb{R}^n$ be an open set, $u \in C^1(\Omega)$ and $\varphi \in C_0^\infty(\Omega)$. Integration by parts gives

$$\int_\Omega u.\, \frac{\partial \varphi}{\partial x_j}\, \mathrm{dx} = u(\mathbf{x})\varphi(\mathbf{x}) - \int_\Omega \frac{\partial u}{\partial x_j}\, \varphi \, \mathrm{dx}, \tag{1.58}$$

There is no boundary term because φ has a compact support in Ω and thus vanishes near $\partial\Omega$ [46]. The basic intuition is that weakly differentiable functions look differentiable except for a set of zero measure. This allows functions that are not normally considered differentiable to have a weak derivative that is defined everywhere on the original domain of definition. The reason why weak derivatives ignore sets of zero measure is precisely because weak derivatives are defined by integrals, and integrals can not see the behaviour on sets of zero measure. Let $u \in C^k(\Omega)$, $k = 1, 2, 3, \ldots$, and let $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n) \in \mathbb{W}^n$ ($\mathbb{W} = \mathbb{N} \cup \{0\}$ is whole numbers) be a **multi-index** such that the order of multi-index $|\alpha| = \alpha_1 + \alpha_2 + \ldots + \alpha_n$ is at most k. We denote

$$D^\alpha u = \frac{\partial^{|\alpha|} u}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}} = \frac{\partial^{\alpha_1}}{\partial x_1^{\alpha_1}} \cdots \frac{\partial^{\alpha_n}}{\partial x_n^{\alpha_n}} u. \tag{1.59}$$

Definition 1.10.1. Let Ω be an open set in $\mathbb{R}^n$ ($n = 1$ is allowed), and let $u, v : \Omega \longrightarrow \mathbb{R}$ be functions that are locally integrable in Ω, $u, v \in L^1_{loc}$. Let $\alpha \in \mathbb{W}^n$ be a multi-index. We say that $v : \Omega \longrightarrow \mathbb{R}$ is a **weak partial derivative** or **distributional** of u of order α in Ω with respect to the variable $\mathbf{x}$ if it is true that (see [30])

$$\int_\Omega u(\mathbf{x})\, D^\alpha \varphi(\mathbf{x})\, \mathrm{dm_n} = (-1)^{|\alpha|} \int_\Omega D^\alpha u(\mathbf{x})\, \varphi(\mathbf{x})\, \mathrm{dm_n} = (-1)^{|\alpha|} \int_\Omega v(\mathbf{x})\, \varphi(\mathbf{x})\, \mathrm{dm_n}, \tag{1.60}$$

where $\varphi \in C_0^\infty(\Omega)$. Recall that to define the weak derivative $\partial^\alpha u$, we do not need the existence of derivatives of smaller order (like in the classical definition). The weak derivative is defined as an element of L^1_{loc}, so we can change it on some set of measure zero. If $\tilde{v}$ is a second function satisfying these conditions, then it is not difficult to show that $\tilde{v} = v$ almost everywhere in Ω, so v is essentially unique (see [30], and [46]).We can also prove that if $\mathbf{f} \in L^1_{loc}$ satisfies

$$\int_\Omega \mathbf{f}\, \varphi \, \mathrm{dx} = 0, \tag{1.61}$$

for every $\varphi \in C_0^\infty(\Omega)$, then $\mathbf{f} = 0$ almost everywhere in Ω. This is an integral way to say that a function is zero almost everywhere.

Example 1.10.3. Let $n = 1$ and $\Omega = (0, 2)$. Consider

$$u(x) = \begin{cases} x & \text{if } \ 0 < x < 1 \\ 1 & \text{if } \ 1 \leq x < 2, \end{cases}$$

and

$$v(x) = \begin{cases} 1 & \text{if } \ 0 < x < 1 \\ 0 & \text{if } \ 1 \leq x < 2. \end{cases}$$

We want to prove that $u' = v$ in the weak sense. We must show

$$\int_0^2 u\,\varphi' \mathrm{dx} = -\int_0^2 v\,\varphi\,\mathrm{dx},$$

for every $\varphi \in C_0^\infty((0, 2))$. By the fundamental theorem of calculus and integration by parts, we get

$$\begin{aligned} \int_0^2 u(x)\,\varphi'(x)\mathrm{dx} &= \int_0^1 x\,\varphi'(x)\mathrm{dx} + \int_1^2 \varphi'(x)\mathrm{dx} \\ &= \Big[x\,\varphi(x)\Big]_0^1 - \int_0^1 \varphi(x)\mathrm{dx} + (\varphi(2) - \varphi(1)) \\ &= -\int_0^1 \varphi(x)\mathrm{dx} \\ &= -\int_0^1 1\,\varphi(x)\mathrm{dx} - \int_1^2 0\,\varphi(x)\mathrm{dx} \\ &= -\int_0^2 v\,\varphi(x)\mathrm{dx}, \end{aligned}$$

for every $\varphi \in C_0^\infty((0, 2))$. Recall that φ is zero on $\partial\Omega$, so $\varphi(2) = \varphi(0) = 0$ (see [46] and [48]).

♣

Definition 1.10.2. Suppose that Ω is an open set in $\mathbb{R}^n$. The **Sobolev space** $W^{k,p}(\Omega)$ consists of functions $u \in L^p(\Omega)$ such that for every multi-index α with $|\alpha| \leq k$, the weak derivative $D^\alpha u$ exists and $D^\alpha u \in L^p(\Omega)$. Therefore

$$W^{k,p}(\Omega) = \big\{u \in L^p(\Omega) : \ D^\alpha u \in L^p(\Omega), \ |\alpha| \leq k\big\}. \tag{1.62}$$

If $u \in W^{k,p}(\Omega)$ we define its norm (see [46] and [42])

$$\|u\|_{W^{k,p}(\Omega)} = \left(\sum_{|\alpha| \leq k} \int_\Omega |D^\alpha u|^p \mathrm{dx}\right)^{\frac{1}{p}}, \qquad 1 \leq p < \infty, \tag{1.63}$$

and

$$\|u\|_{W^{k,\infty}(\Omega)} = \sum_{|\alpha| \leq k} \operatorname*{ess\,sup}_{\Omega} |D^\alpha u|. \tag{1.64}$$

So, the Sobolev space $W^{k,p}(\Omega)$ consists of functions in $L^p(\Omega)$ that have weak derivatives up to order k and they belong to $L^p(\Omega)$. The Sobolev space $W^{k,p}(\Omega)$, with $1 \leq p \leq \infty$ and $k = 1, 2, 3, \ldots$, is a Banach space. The **local Sobolev space** $W^{k,p}_{loc}(\Omega)$ consists of those functions whose partials of order k or less are in $L^p_{loc}(\Omega)$. We say that the sequence $\{\mathbf{f_i}\}$ converges to $\mathbf{f}$ in $W^{k,p}_{loc}(\Omega)$ if

$$\lim_{i\to\infty} \|\partial^\alpha \mathbf{f_i} - \partial^\alpha \mathbf{f}\|_{L^p(X)} = 0, \tag{1.65}$$

for every compact set $X \subset \Omega$ and every $|\alpha| \leq k$, see [42]. The space $W^{k,p}_0(\Omega)$ is defined as the closure of $C^\infty_0(\Omega)$, the compactly supported smooth functions, with respect to the norms defined at (1.63) and (1.64). The space $W^{k,p}_0(\Omega)$ is the set of all functions u which are an element of the space $W^{k,p}(\Omega)$ where u is zero on the boundary of Ω ($\partial\Omega$).

$$W^{k,p}_0(\Omega) = \{u \in W^{k,p}(\Omega) : \text{ } u \text{ is zero on } \partial\Omega\}. \tag{1.66}$$

For $k = 1$, the norms on $W^{k,p}(\Omega)$ in (1.63) and (1.64) can be expressed by

$$\|u\|_{W^{k,p}(\Omega)} = \left(\|u\|^p_{L^p(\Omega)} + \|Du\|^p_{L^p(\Omega)}\right)^{\frac{1}{p}}, \tag{1.67}$$

for $1 \leq p < \infty$, and for $p = \infty$ we have

$$\|u\|_{W^{k,\infty}(\Omega)} = \operatorname*{ess\,sup}_{\Omega} |u| + \operatorname*{ess\,sup}_{\Omega} |Du|. \tag{1.68}$$

Example 1.10.4. Let $\Omega = [0, 2\pi] \times [0, 2\pi]$ and $u : \Omega \longrightarrow \mathbb{R}$ be a function given by

$$u(\mathbf{x}) = u(x_1, x_2) = \sin x_1 + \cos x_2.$$

We see that $\frac{\partial u}{\partial x_1} = \cos x_1$ and $\frac{\partial u}{\partial x_2} = -\sin x_2$.

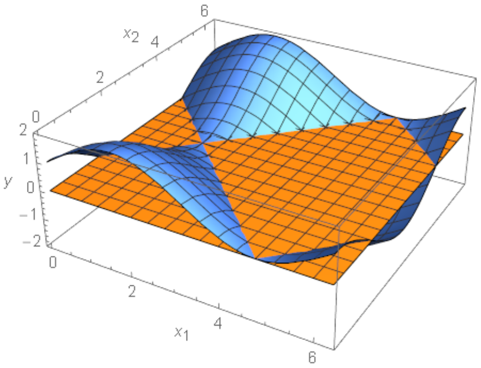

Figure 1.5: The Sobolev function in $W^{1,2}(\Omega)$ space given by $y = \sin(x_1) + \cos(x_2)$.

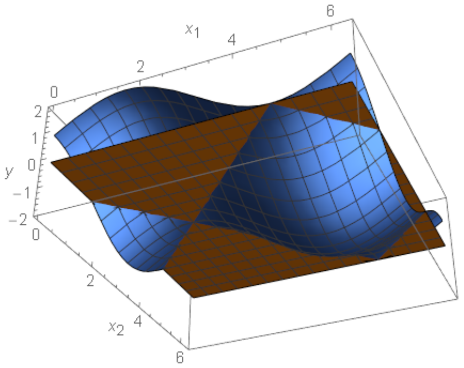

Figure 1.6: The Sobolev function in $W^{1,2}(\Omega)$ space given by $y = \sin(x_1) + \cos(x_2)$ from another perspective.

Moreover, we obtain

$$\int_\Omega |u|^2 \mathrm{dx} = \int_\Omega (\sin x_1 + \cos x_2)^2 \mathrm{dx} \leq \int_\Omega |1 + 1|^2 \mathrm{dx} = \int_\Omega 4 \mathrm{dx} = 4|\Omega| < \infty,$$

$$\int_\Omega \left|\frac{\partial u}{\partial x_1}\right|^2 \mathrm{dx} = \int_\Omega |\cos x_1|^2 \mathrm{dx} \le \int_\Omega |1|^2 \mathrm{dx} = |\Omega| < \infty,$$

and

$$\int_\Omega \left|\frac{\partial u}{\partial x_2}\right|^2 \mathrm{dx} = \int_\Omega |-\sin x_2|^2 \mathrm{dx} \le \int_\Omega |1|^2 \mathrm{dx} = |\Omega| < \infty.$$

Thus, $u \in W^{1,2}(\Omega)$, and this means the function u is a Sobolev-function (see [49]).

♣

Example 1.10.5. Let $\Omega = (0,1) \times (0,1)$ be the open unit square and let $u : \Omega \longrightarrow \mathbb{R}$ be a function given by

$$u(\mathbf{x}) = u(x_1, x_2) = \sqrt{x_1}.$$

The partial derivatives of u are $\frac{\partial u}{\partial x_1} = \frac{1}{2\sqrt{x_1}}$, and $\frac{\partial u}{\partial x_2} = 0$. So, we have

$$\int_\Omega |u|^2 \mathrm{dx} = \int_0^1 \int_0^1 \sqrt{x_1}^2 \, \mathrm{dx}_1 \mathrm{dx}_2 = \frac{1}{2} \int_0^1 \mathrm{dx}_2 = \frac{1}{2} < \infty,$$

$$\int_\Omega \left|\frac{\partial u}{\partial x_2}\right|^2 \mathrm{dx} = \int_0^1 \int_0^1 0 \, \mathrm{dx}_2 \, \mathrm{dx}_1 = 0. \int_0^1 \mathrm{dx}_1 = 0 < \infty,$$

and

$$\int_\Omega \left|\frac{\partial u}{\partial x_1}\right|^2 \mathrm{dx} = \int_0^1 \int_0^1 \left|\frac{1}{2} x^{-\frac{1}{2}}\right|^2 \mathrm{dx}_1 \, \mathrm{dx}_2 = \frac{1}{4} \int_0^1 \int_0^1 x^{-1} \mathrm{dx}_1 \, \mathrm{dx}_2$$

$$= \frac{1}{4} \int_0^1 (\ln 1 - \ln 0) \, \mathrm{dx}_2 = \frac{1}{4} \times \infty = \infty.$$

This shows that $u \notin W^{1,2}(\Omega)$, but $u \in W^{1,1}(\Omega)$. That means whether u is or is not a Sobelev function depends on the space (see [49], page 6).

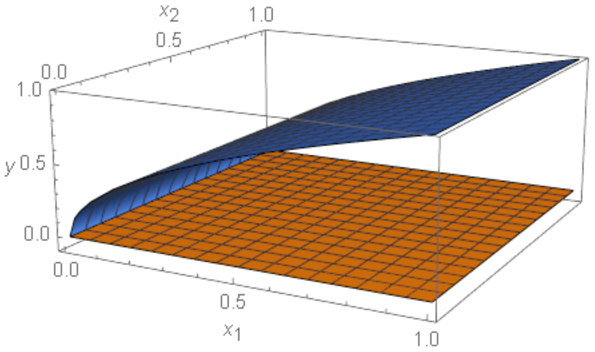

Figure 1.7: The Sobolev function in $W^{1,1}(\Omega)$, and the non-Sobolev function in $W^{1,2}(\Omega)$ given by $u(x_1, x_2) = \sqrt{x_1}$.

♣

Chapter 2

Polyconvex, Quasiconvex and Rank-One Convex Functions

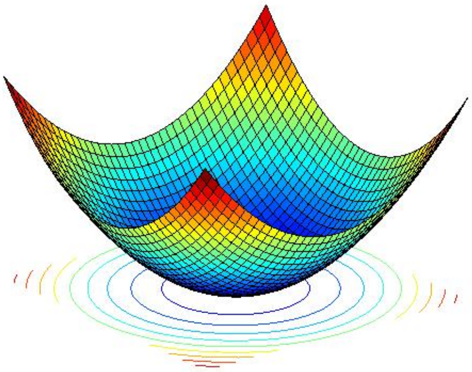

Figure 2.1: A graph of a simple convex function with the equation: $\mathbf{f}(x,y) = x^2 + y^2$.

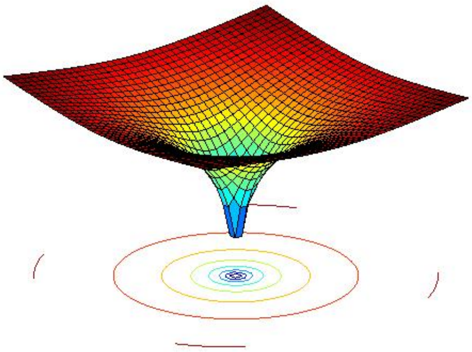

Figure 2.2: A quasiconvex function that is not convex with equation: $\mathbf{f}(x,y) = \log(x^2 + y^2)$.

Chapter 2

Polyconvex, Quasiconvex and Rank-One Convex Functions

We have assigned this chapter to an important subject in this thesis concerning convexity. We seek to generalise convexity for the space of all $m \times n$ matrices. In the first section, we describe the concept of convexity in higher dimensions. In the second section, we define polyconvex, quasiconvex, rank-one convex functions, and the relations between them.

2.1 Convex Functions in Higher Dimensions

In this section we define convex and concave function in all dimensions.

Definition 2.1.1. Let I be a nondegenerate interval of $\mathbb{R}$. A function $f : I \longrightarrow \mathbb{R}$ is called **convex** if

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda)f(x_2) \tag{2.1}$$

for all x_1 and x_2 in I and $\lambda \in [0,1]$. It is called **strictly convex** if the inequality (2.1) holds strictly whenever x_1 and x_2 are distinct points and $\lambda \in (0,1)$. If the function $-f$ is convex or strictly convex then we say that f is **concave** or **strictly concave**, respectively. If f is both convex and concave in I, then f is said to be **affine** (see [14], [52], [59]).

The convexity of a function $f : I \longrightarrow \mathbb{R}$ means geometrically that for all $u,\ v \in I$ and $u < v$, the points of the graph of f in the interval $[u,v]$ are under or on the chord joining the endpoints $(u, f(u))$ and $(v, f(v))$.

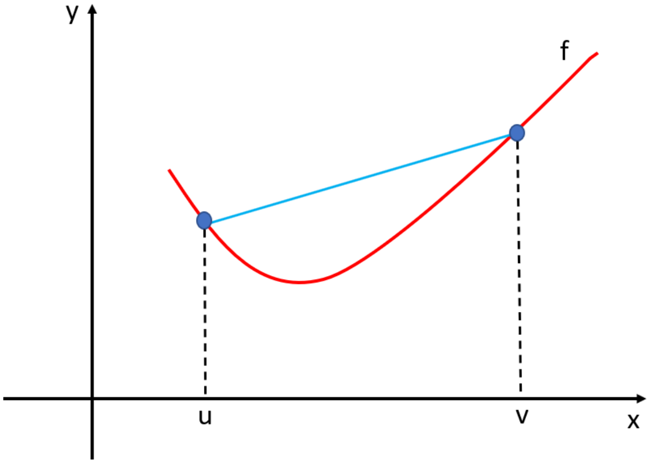

Figure 2.3: The graph of a convex function is under the chord.

Theorem 2.1.1. (**J.Jensen Midpoint Convex**) Let $f : I \longrightarrow \mathbb{R}$ be a continuous function. Then f is convex if and only if it is midpoint convex, that is,

$$f(\frac{x+y}{2}) \leq \frac{f(x)+f(y)}{2} \qquad for\ \ all\ \ x, y \in I. \tag{2.2}$$

Corollary 2.1.1.1. Let $f : I \longrightarrow \mathbb{R}$ be a continuous function. Then f is convex if and only if

$$f(x+h)+f(x-h)-2f(x) \geq 0 \tag{2.3}$$

for all $x \in I$ and all $h > 0$ such that $x+h$ and $x-h$ are in I.

Corollary 2.1.1.2. (**The second derivative test for convexity**) Suppose that $f : I \longrightarrow \mathbb{R}$ is a twice differentiable function. Then:

1. f is convex if and only if $f''(x) \geq 0$ for all x in I;
2. f is strictly convex if and only if $f''(x) \geq 0$ and the set of points where f'' vanishes does not include intervals of positive length.
3. f is concave if and only if $f''(x) \leq 0$ for all x in I;
4. f is strictly concave if and only if $f''(x) \leq 0$ and the set of points where f'' vanishes does not include intervals of positive length.

Theorem 2.1.2. (The operations with convex functions)

1. Adding two convex functions defined on the same interval we obtain a convex function; if one of them is strictly convex, then the sum is also strictly convex.
2. Multiplying a (strictly) convex function by a positive scalar, we obtain, also, a (strictly) convex function.
3. The restriction of every (strictly) convex function to a subinterval of its domain is also a (strictly) convex function.
4. If $f : I \longrightarrow \mathbb{R}$ is a convex (strictly convex) function and $g : \mathbb{R} \longrightarrow \mathbb{R}$ is a nondecreasing (an increasing) convex function, then $g \circ f$ is a convex (strictly convex) function.

For more information and proofs you can see chapter one of book [59]. Convex functions provide basic techniques in optimization theory, partial differential equations and geometric inequalities. The natural domain for a convex function is a convex set. That is why we shall start by recalling some basic facts on convex sets. A function f defined on an interval J is said to be **Lipschitz** if there exists a constant $L \geq 0$ such that

$$|f(x)-f(y)| \leq L|x-y| \qquad \text{for all } x, y \in J. \tag{2.4}$$

A subset C of a vector space (linear space are assumed henceforth to be real) E to be is said to be convex if it contains the line segment

$$[x, y] = \{(1-\lambda)x + \lambda y | \lambda \in [0,1]\}$$

connecting any of its points x and y.

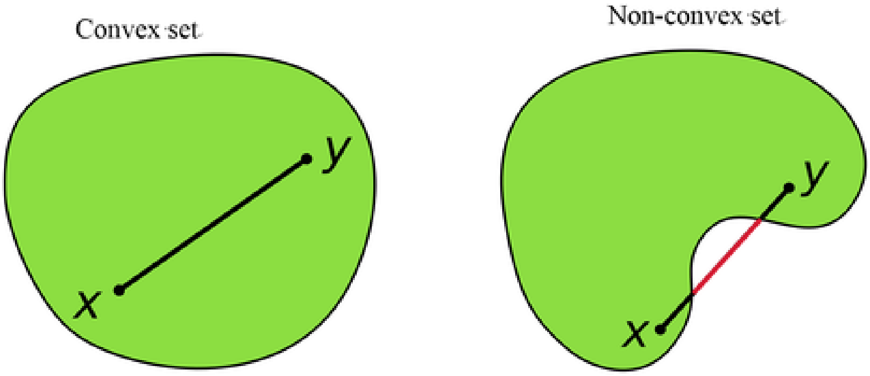

Figure 2.4: The convex and Non-convex (concave) sets.
Source: The convex function images

A subset A of E is said to be an **affine set** if it contains the whole line through any two of its points. Algebraically, this means

$$\text{for all } x, y \in A \text{ and } \lambda \in \mathbb{R} \qquad (1-\lambda)x + \lambda y \in A.$$

For the given set $E \subset \mathbb{R}^n$, $\overline{E}$, ∂E, $\text{int}E$ and E^c respectively stand for the closure, the boundary, the interior and the complement of E. The **affine hull** of a set $E \subset \mathbb{R}^n$ is the smallest affine set containing E. It is denoted by *aff* E . A **hyperplane** $H \subset \mathbb{R}^n$ is a set of the form

$$H = \{\mathbf{x} \in \mathbb{R}^n | \, \langle \mathbf{x}, a \rangle = \alpha\}, \tag{2.5}$$

where $a \in \mathbb{R}^n$, $a \neq 0$ and $\alpha \in \mathbb{R}$. Also, we can define the **convex hull** of a set $E \subset \mathbb{R}^n$, denoted by coE, is the smallest convex set containing E.

Definition 2.1.2. Let U be a convex subset of $\mathbb{R}^n$ (real linear space) and let $\mathbf{f} : U \longrightarrow \mathbb{R}$ be a function

1. The **subgraph of f**, denoted *sub* **f**, is the set

$$\text{sub}\,\mathbf{f} = \{(\mathbf{x}, y) \in U \times \mathbb{R} \mid \mathbf{f}(\mathbf{x}) \geq y\}. \tag{2.6}$$

2. The **epigraph of f**, denoted *epi* **f**, is the set

$$\text{epi}\,\mathbf{f} = \{(\mathbf{x}, y) \in U \times \mathbb{R} \mid \mathbf{f}(\mathbf{x}) \leq y\}. \tag{2.7}$$

3. The **domain** of **f** is defined as

$$\text{dom}\,\mathbf{f} = \{\mathbf{x} \in \mathbb{R}^n \mid \mathbf{f}(\mathbf{x}) < \infty\}. \tag{2.8}$$

4. The **lower contour set of f at** $\alpha \in \mathbb{R}$ **or lower-level set of f at** $\alpha \in \mathbb{R}$, is defined as

$$L_\alpha = \{\mathbf{x} \in U \mid \mathbf{f}(\mathbf{x}) \leq \alpha\}. \tag{2.9}$$

5. The **upper contour set of f at** $\alpha \in \mathbb{R}$ **or upper-level set of f at** $\alpha \in \mathbb{R}$, denoted U_α, is the set
$$U_\alpha = \{\mathbf{x} \in U \mid \mathbf{f}(\mathbf{x}) \geq \alpha\}. \tag{2.10}$$

The subgraph of a function is the region lying below the graph of the function, and the epigraph of a function is the region lying above the graph of the function. If U is a convex subset of $\mathbb{R}^n$ and $\mathbf{f} : U \longrightarrow \mathbb{R}$ is a function, we say that $\mathbf{f}$ is convex on U if *sub* $\mathbf{f}$ is a convex set. Also, we say that $\mathbf{f}$ is convex on U if *epi* $\mathbf{f}$ is a convex set. Note that concave and convex functions are required to have convex domains. For more information about these sets you can see [59] or [19]. Source of the figure below: (http://scottmccracken.weebly.com/uploads/9/0/6/6/9066859/convexity-print_version.pdf)

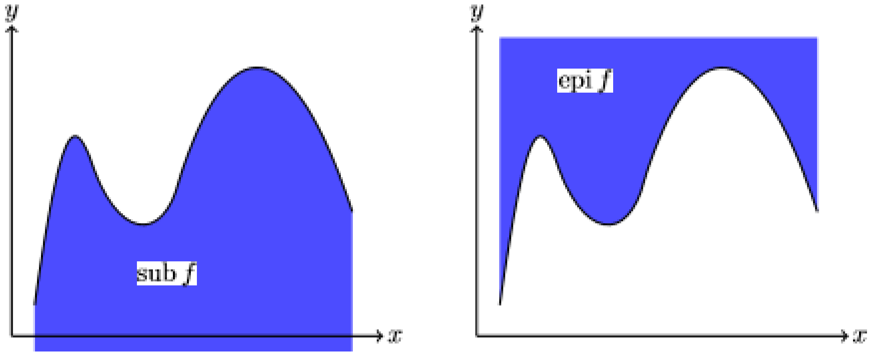

Figure 2.5: The subgraph and epigraph of f.

Definition 2.1.3. Let U be a convex subset of $\mathbb{R}^n$ and let $\mathbf{f} : U \longrightarrow \mathbb{R}$ be a function. Then

1. $\mathbf{f}$ is convex if and only if for all $\mathbf{x}, \mathbf{y} \in U$ and $\lambda \in [0, 1]$, we have
$$\mathbf{f}(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) \leq \lambda\mathbf{f}(\mathbf{x}) + (1-\lambda)\mathbf{f}(\mathbf{y}).$$
2. $\mathbf{f}$ is concave if and only if for all $\mathbf{x}, \mathbf{y} \in U$ and $\lambda \in [0, 1]$, we have
$$\mathbf{f}(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) \geq \lambda\mathbf{f}(\mathbf{x}) + (1-\lambda)\mathbf{f}(\mathbf{y}).$$
3. $\mathbf{f}$ is strictly convex if and only if for all $\mathbf{x}, \mathbf{y} \in U$ and $\lambda \in (0, 1)$, we have
$$\mathbf{f}(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) < \lambda\mathbf{f}(\mathbf{x}) + (1-\lambda)\mathbf{f}(\mathbf{y}).$$
4. $\mathbf{f}$ is strictly concave if and only if for all $\mathbf{x}, \mathbf{y} \in U$ and $\lambda \in (0, 1)$, we have
$$\mathbf{f}(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) > \lambda\mathbf{f}(\mathbf{x}) + (1-\lambda)\mathbf{f}(\mathbf{y}).$$

It is easy to prove that if U is a convex subset of $\mathbb{R}^n$ and $\mathbf{f} : U \longrightarrow \mathbb{R}$ is a function, then $\mathbf{f}$ is a convex function if and only if $-\mathbf{f}$ is concave and vice versa. If $\mathbf{f}_i : U \longrightarrow \mathbb{R}$ are convex (concave) functions and a_i are positive numbers $i = 1, \ldots, k$, then
$$a_1\mathbf{f}_1 + \cdots a_k\mathbf{f}_k$$

is a convex (concave) function. Convexity in the case of several variables is equivalent to convexity on each line segment included in the domain of definition. So we can say, a function $\mathbf{f} : U \longrightarrow \mathbb{R}$ is convex if and only if for every two points $\mathbf{x}$ and $\mathbf{y}$ in U the function

$$\varphi : [0, 1] \longrightarrow \mathbb{R}, \qquad \varphi(t) = \mathbf{f}((1 - t)\mathbf{x} + t\mathbf{y})$$

is convex. We can prove that each level set of height α of a convex function is a convex set. Convex functions exhibit a series of nice properties related to maximum and minimum of functions. We want to describe some theorems about convex and concave functions.

Theorem 2.1.3. Let $U \subseteq \mathbb{R}^n$ be an open and convex set and let $\mathbf{f} : U \longrightarrow \mathbb{R}$ be either a convex or a concave function. Then $\mathbf{f}$ is a continuous function.

Theorem 2.1.4. Let $U \subseteq \mathbb{R}^n$ be an open and convex set and let $\mathbf{f} : U \longrightarrow \mathbb{R}$ be differentiable.

1. A function $\mathbf{f}$ is convex if and only if for any $\mathbf{x}, \mathbf{y} \in U$ we have

$$\mathbf{f}(\mathbf{x}) \geq \nabla\mathbf{f}(\mathbf{y})(\mathbf{x} - \mathbf{y}) + \mathbf{f}(\mathbf{y}). \tag{2.11}$$

2. A function $\mathbf{f}$ is concave if and only if for any $\mathbf{x}, \mathbf{y} \in U$ we have

$$\mathbf{f}(\mathbf{x}) \leq \nabla\mathbf{f}(\mathbf{y})(\mathbf{x} - \mathbf{y}) + \mathbf{f}(\mathbf{y}). \tag{2.12}$$

The proofs of these theorems are long and we don't intend to describe them here. But you can see their proofs in books [59], [19] or article (https://pages.wustl.edu/files/pages/imce/nachbar/concavity.pdf). By Theorem 1.4.1 we know that the order in which we calculate any kth-order partial derivative does not matter. The matrix of second order partial derivatives is the Hessian denoted by $H\mathbf{f}(\mathbf{x}) = D^2\mathbf{f}(\mathbf{x}) = D(\nabla\mathbf{f})(\mathbf{x})$. By Theorem 1.4.1 if f is C^2 then $H\mathbf{f}(\mathbf{x}) = D^2\mathbf{f}(\mathbf{x})$ is symmetric.

Theorem 2.1.5. Let $U \subseteq \mathbb{R}^n$ be an open and convex set and let $\mathbf{f} : U \longrightarrow \mathbb{R}$ be C^2.

1. If $H\mathbf{f}(\mathbf{x}) = D^2\mathbf{f}(\mathbf{x})$ is negative definite for every $\mathbf{x} \in U$ then f is strictly concave.
2. If $H\mathbf{f}(\mathbf{x}) = D^2\mathbf{f}(\mathbf{x})$ is negative semi-definite for every $\mathbf{x} \in U$ then f is concave.
3. If f is concave and C^2 then $H\mathbf{f}(\mathbf{x}) = D^2\mathbf{f}(\mathbf{x})$ is negative semi-definite for every $\mathbf{x} \in U$.
4. If $H\mathbf{f}(\mathbf{x}) = D^2\mathbf{f}(\mathbf{x})$ is positive definite for every $\mathbf{x} \in U$ then f is strictly convex.
5. If $H\mathbf{f}(\mathbf{x}) = D^2\mathbf{f}(\mathbf{x})$ is positive semi-definite for every $\mathbf{x} \in U$ then f is convex.
6. If f is convex and C^2 then $H\mathbf{f}(\mathbf{x}) = D^2\mathbf{f}(\mathbf{x})$ is positive semi-definite for every $\mathbf{x} \in U$.

2.2 Polyconvex, Quasiconvex and Rank-One Convex Functions and Examples

We start this section with three important definitions and their relationships to each other. One of these definitions is related to the main question of this thesis and so we will find some equivalent conditions for it. One major application of convex and concave functions is optimization. As we will see, any relative maximum of a concave function is an absolute maximum. The notion of polyconvexity is a generalisation of the notion of convexity for functions defined on spaces of matrices.

Definition 2.2.1. A function $\mathcal{F} : M_{m\times n}(\mathbb{R}) \longrightarrow \mathbb{R} \cup \{+\infty\}$ is **polyconvex** if there exists a convex function $\mathcal{G} : \mathbb{R}^{\tau(m,n)} \longrightarrow \mathbb{R} \cup \{+\infty\}$ (in general non-unique) such that

$$\mathcal{F}(A) = \mathcal{G}(T(A)),$$

for every $A \in M_{m\times n}(\mathbb{R})$, where $T : \mathbb{R}^{m\times n} \longrightarrow \mathbb{R}^{\tau(m,n)}$ is such that

$$T(A) := (A,\ \mathrm{Minors}_2 A,\ \mathrm{Minors}_3 A, \dots,\ \mathrm{Minors}_{m\wedge n} A).$$

In this definition, $\mathrm{Minors}_r A$ stands for the matrix of all $r \times r$ minors of the matrix $A \in M_{m\times n}(\mathbb{R})$, $2 \le r \le m \wedge n = \min\{m, n\}$ and

$$\tau(m,n) := \sum_{r=1}^{m\wedge n} \sigma(r), \quad \text{where} \quad \sigma(r) := \frac{m!n!}{(r!)^2(m-r)!(n-r)!}.$$

In other words, $\mathcal{F}$ is polyconvex if $\mathcal{F}(A)$ is a convex function of subdeterminants of the matrix A or $\mathcal{F}$ is polyconvex if $\mathcal{F}(A)$ is a convex function of the minors of the matrix A (See [19], [69], [70], [68] or [21]). We will show that the function $\mathcal{G}$ is not unique.

Example 2.2.1. In the case $m = n = 2$ (see [19], [9]), we have

$$\sigma(1) = \frac{1! \times 1!}{(1!)^2(2-1)!(2-1)!} = 4, \qquad \sigma(2) = \frac{2! \times 2!}{(2!)^2(2-2)!(2-2)!} = 1,$$

$$\tau(m,n) := \sum_{r=1}^{2} \sigma(r) = \sigma(1) + \sigma(2) = 5, \qquad T(A) = (A, \det A).$$

Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

and

$$\mathcal{F}(A) = |A|^2 = a^2 + b^2 + c^2 + d^2 = (a-d)^2 + (c+d)^2 + 2\det A.$$

Suppose $\mathcal{G}_1,\ \mathcal{G}_2 : \mathbb{R}^{\tau(m,n)} = \mathbb{R}^5 \longrightarrow \mathbb{R}$ is defined by

$$\mathcal{G}_1(A,x) := |A|^2 \qquad \text{and} \qquad \mathcal{G}_2(A,x) := (a-d)^2 + (c+d)^2 + 2x.$$

Then it is easy to check $\mathcal{G}_1$ and $\mathcal{G}_2$ are convex, $\mathcal{G}_1 \neq \mathcal{G}_2$ and

$$\mathcal{F}(A) = \mathcal{G}_1(T(A)) = \mathcal{G}_1(A, \det A) = \mathcal{G}_2(T(A)) = \mathcal{G}_2(A, \det A).$$

♣

Polyconvexity is a more general property than convexity, because there are examples of polyconvex functions that are not convex. The following example shows this concept.

Example 2.2.2. Let $\mathcal{F} : M_{2\times 2}(\mathbb{R}) \longrightarrow \mathbb{R}\cup\{+\infty\}$ be a function (see [62], p 354) with equation

$$\mathcal{F}(A) = \begin{cases} \dfrac{1}{\det A} & \text{if } \det(A) > 0, \\ \\ +\infty & \text{if } \det(A) \leq 0. \end{cases}$$

There is a convex function $\mathcal{G} : \mathbb{R}^5 \longrightarrow \mathbb{R} \cup \{+\infty\}$ such that

$$\mathcal{F}(A) = \mathcal{G}(T(A)), \qquad \text{where} \quad \mathrm{T(A) = (A, \det A)}.$$

Let

$$\mathcal{G}(A, x) := \begin{cases} \dfrac{1}{x} & \text{if } x > 0, \\ \\ +\infty & \text{if } x \leq 0. \end{cases}$$

The function $\mathcal{G}$ is convex, and that means the function $\mathcal{F}$ is polyconvex. But the function $\mathcal{F}$ is not convex. Let

$$A = \begin{bmatrix} 1 & 9 \\ 0 & 1 \end{bmatrix} \qquad , \qquad B = \begin{bmatrix} 0 & -4 \\ 1 & 0 \end{bmatrix},$$

be two 2×2 matrices in $M_{2\times 2}(\mathbb{R})$. So, $\det(A) = 1$ and $\det(B) = 4$, then $\mathcal{F}(A) = 1$ and $\mathcal{F}(B) = 1/4$. Therefore

$$\lambda\, \mathcal{F}(A) + (1-\lambda)\, \mathcal{F}(B) = \lambda + (1-\lambda)\frac{1}{4} = \frac{1}{4} + \frac{3}{4}\lambda > 0 \qquad \forall \lambda \in [0,1].$$

But

$$\lambda\, A + (1-\lambda)B = \lambda \begin{bmatrix} 1 & 9 \\ 0 & 1 \end{bmatrix} + (1-\lambda) \begin{bmatrix} 0 & -4 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \lambda & 13\lambda - 4 \\ 1-\lambda & \lambda \end{bmatrix}.$$

Thus $\det(\lambda\, A + (1-\lambda)B) = 14\lambda^2 - 17\lambda + 4$, for all $\lambda \in [0,1]$. If we choose $\lambda = \frac{1}{2}$, then

$$\det(\lambda\, A + (1-\lambda)B) = 14(\frac{1}{2})^2 - 17(\frac{1}{2}) + 4 = -1.$$

Hence we have $\mathcal{F}(\lambda\, A + (1-\lambda)B) = +\infty$. So, this proves the function $\mathcal{F}$ is polyconvex but not convex.

♣

There is also another example that is a polyconvex function but not convex. Let $\mathcal{F} : M_{2\times 2}(\mathbb{R}) \longrightarrow \mathbb{R} \cup \{+\infty\}$ be a function with equation

$$\mathcal{F}(A) := \det(A).$$

It is not difficult to see this function is polyconvex but not convex (see [19]). Now we describe the concepts of quasiconvexity and rank-one convex. At first, we explain an easy definition for quasiconvex functions, and we extend the following definition to an advanced definition.

Definition 2.2.2. Let U be a convex subset of $\mathbb{R}^n$ and let $\mathcal{F} : U \longrightarrow \mathbb{R}$ be a function. Then

1. A function $\mathcal{F}$ is **quasiconvex** if all its lower-level sets are convex.

2. A function $\mathcal{F}$ is **quasiconcave** if all its upper-level sets are concave.

Suppose that L_α is the lower-level set of $\mathcal{F}$ at $\alpha \in \mathbb{R}$, then $x \in L_\alpha$ is equivalent to $\mathcal{F}(x) \leq \alpha$. So:

$$\begin{aligned}
&\mathcal{F} \text{ is quasiconvex}\\
&\Leftrightarrow \forall \alpha \in \mathbb{R},\ \forall \mathbf{x}, \mathbf{y},\ \forall \lambda \in [0,1] : \Big[\mathbf{x}, \mathbf{y} \in L_\alpha \Rightarrow \lambda\mathbf{x} + (1-\lambda)\mathbf{y} \in L_\alpha\Big]\\
&\Leftrightarrow \forall \alpha \in \mathbb{R},\ \forall \mathbf{x}, \mathbf{y},\ \forall \lambda \in [0,1] : \Big[\mathcal{F}(\mathbf{x}) \leq \alpha, \mathcal{F}(\mathbf{y}) \leq \alpha \Rightarrow \mathcal{F}(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) \leq \alpha\Big]\\
&\Leftrightarrow \forall \alpha \in \mathbb{R},\ \forall \mathbf{x}, \mathbf{y},\ \forall \lambda \in [0,1] : \Big[\max\big\{\mathcal{F}(\mathbf{x}), \mathcal{F}(\mathbf{y})\big\} \leq \alpha \Rightarrow \mathcal{F}(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) \leq \alpha\Big]\\
&\Leftrightarrow \forall \mathbf{x}, \mathbf{y},\ \forall \lambda \in [0,1] : \Big[\mathcal{F}(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) \leq \max\big\{\mathcal{F}(\mathbf{x}), \mathcal{F}(\mathbf{y})\big\}\Big]
\end{aligned}$$

So,

1. A function $\mathcal{F}$ is quasiconvex if and only if
$$\forall \mathbf{x}, \mathbf{y},\ \forall \lambda \in [0,1] : \Big[\mathcal{F}(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) \leq \max\big\{\mathcal{F}(\mathbf{x}), \mathcal{F}(\mathbf{y})\big\}\Big]. \tag{2.13}$$

2. A function $\mathcal{F}$ is quasiconcave if and only if
$$\forall \mathbf{x}, \mathbf{y},\ \forall \lambda \in [0,1] : \Big[\mathcal{F}(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) \geq \min\big\{\mathcal{F}(\mathbf{x}), \mathcal{F}(\mathbf{y})\big\}\Big]. \tag{2.14}$$

Definition 2.2.3. Let U be a convex subset of $\mathbb{R}^n$ and let $\mathcal{F} : U \longrightarrow \mathbb{R}$ be a function.

1. A function $\mathcal{F}$ is **strictly quasiconvex** if for all $\mathbf{x}$, $\mathbf{y} \in U$ with $\mathbf{x} \neq \mathbf{y}$ and all $\lambda \in (0,1)$, we have
$$\mathcal{F}(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) < \max\big\{\mathcal{F}(\mathbf{x}), \mathcal{F}(\mathbf{y})\big\}. \tag{2.15}$$

2. A function $\mathcal{F}$ is **strictly quasiconcave** if for all $\mathbf{x}$, $\mathbf{y} \in U$ with $\mathbf{x} \neq \mathbf{y}$ and all $\lambda \in (0,1)$, we have
$$\mathcal{F}(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) > \min\big\{\mathcal{F}(\mathbf{x}), \mathcal{F}(\mathbf{y})\big\}. \tag{2.16}$$

Example 2.2.3. Let $\mathcal{F} : U \subset \mathbb{R}^2 \longrightarrow \mathbb{R}$ be a function with equation $\mathcal{F}(\mathbf{x}) = \log(x^2 + y^2)$, where $\mathbf{x} = (x, y) \in \mathbb{R}^2$. The lower-level set of function $\mathcal{F}$ is $\mathcal{F}(\mathbf{x}) \leq \alpha$. So, the lower-level set of $\mathcal{F}$ is

$$\mathcal{F}(\mathbf{x}) \leq \alpha \Leftrightarrow \log(x^2 + y^2) \leq \alpha \Leftrightarrow x^2 + y^2 \leq 10^\alpha.$$

It is obvious that the lower-level set of the function $\mathcal{F}$ is convex, therefore, the function $\mathcal{F}$ is quasiconvex. With a little calculation, we can see that the function $\mathcal{F}$ is not convex, because with $\mathbf{x} = (1,1)$ and $\mathbf{y} = (3,3)$ in $U \subset \mathbb{R}^2$, we have

$$\mathcal{F}(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) = \mathcal{F}(3 - 2\lambda, 3 - 2\lambda) = \log((3-2\lambda)^2 + (3-2\lambda)^2) = \log 2 + 2\log(3 - 2\lambda).$$

If $\lambda = 1/2$ than $\mathcal{F}(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) = 0.903$, But the value of $\lambda\mathcal{F}(1,1) + (1-\lambda)\mathcal{F}(3,3) = 0.7781$. So, $\mathcal{F}$ is not convex.

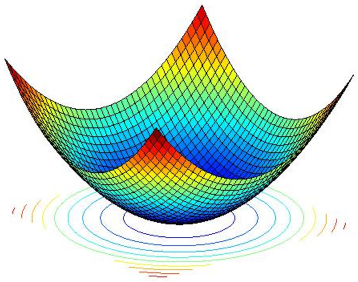

Figure 2.6: The graph of a simple convex function given by: $\mathcal{F}(x, y) = x^2 + y^2$.

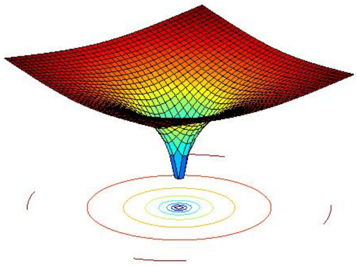

Figure 2.7: The quasiconvex function that is not convex given by: $\mathcal{F}(x, y) = \log(x^2 + y^2)$.

♣

Let U be a convex subset of $\mathbb{R}^n$ and let $\mathcal{F} : U \longrightarrow \mathbb{R}$ be a function. Then $\mathcal{F}$ is quasiconvex (or quasiconcave) if and only if $-\mathcal{F}$ is quasiconcave (or quasiconvex). Also, $\mathbf{f}$ is strictly quasiconvex (or strictly quasiconcave) if and only if the function $-\mathcal{F}$ is strictly quasiconcave (or strictly quasiconvex). In addition, let $\mathcal{F} : U \longrightarrow \mathbb{R}$ be a quasiconvex function. If $\mathcal{G} : \mathbb{R} \longrightarrow \mathbb{R}$ is an increasing function, then $\mathcal{G} \circ \mathcal{F}$ is quasiconvex. If $\mathcal{F} : U \longrightarrow \mathbb{R}$ is a quasiconcave and $\mathcal{G} : \mathbb{R} \longrightarrow \mathbb{R}$ is an increasing function, then $\mathcal{G} \circ \mathcal{F}$ is quasiconcave.

Theorem 2.2.1. Let U be a convex subset of $\mathbb{R}^n$ and let $\mathcal{F} : U \longrightarrow \mathbb{R}$ be a differentiable function on U. Then

1. $\mathcal{F}$ is quasiconvex if and only if

$$\forall\, \mathbf{x}, \mathbf{y} \in U, \qquad \mathcal{F}(\mathbf{x}) \geq \mathcal{F}(\mathbf{y}) \Rightarrow D\mathcal{F}(\mathbf{x}).(\mathbf{y} - \mathbf{x}) \leq 0. \tag{2.17}$$

2. $\mathcal{F}$ is quasiconcave if and only if

$$\forall\, \mathbf{x}, \mathbf{y} \in U, \qquad \mathcal{F}(\mathbf{y}) \geq \mathcal{F}(\mathbf{x}) \Rightarrow D\mathcal{F}(\mathbf{x}).(\mathbf{y} - \mathbf{x}) \geq 0. \tag{2.18}$$

Proof. **(Proof of part 2)** Suppose first that $\mathcal{F}$ is quasiconcave on U. Fix $\mathbf{x}, \mathbf{y} \in U$ and assume $\mathcal{F}(\mathbf{y}) \geq \mathcal{F}(\mathbf{x})$. For any $\lambda \in [0, 1]$, we have

$$\mathcal{F}(\mathbf{x} + \lambda(\mathbf{y} - \mathbf{x})) = \mathcal{F}(\lambda\mathbf{y} + (1 - \lambda)\mathbf{x}) \geq \min\left\{\mathcal{F}(\mathbf{x}), \mathcal{F}(\mathbf{y})\right\} = \mathcal{F}(\mathbf{x}).$$

So for all $\lambda \in (0, 1]$ we can write

$$\frac{\mathcal{F}(\mathbf{x} + \lambda(\mathbf{y} - \mathbf{x})) - \mathcal{F}(\mathbf{x})}{\lambda} \geq 0.$$

We know that if $\lambda \longrightarrow 0^+$, then the following limit gives us the directional derivative concerning the vector $(\mathbf{y} - \mathbf{x})$. Therefore,

$$D_{(\mathbf{y}-\mathbf{x})}\mathcal{F}(\mathbf{x}) = \lim_{\lambda \longrightarrow 0^+} \frac{\mathcal{F}(\mathbf{x} + \lambda(\mathbf{y} - \mathbf{x})) - \mathcal{F}(\mathbf{x})}{\lambda} \geq 0.$$

So, by calculus we have $D_{(\mathbf{y}-\mathbf{x})}\mathcal{F}(\mathbf{x}) = D\mathcal{F}(\mathbf{x}).(\mathbf{y}-\mathbf{x}) \geq 0$. Conversely, assume that for all $\mathbf{x}, \mathbf{y} \in U$ such that $\mathcal{F}(\mathbf{y}) \geq \mathcal{F}(\mathbf{x})$, we have $D\mathcal{F}(\mathbf{x}).(\mathbf{y}-\mathbf{x}) \geq 0$. Choose any $\mathbf{x}, \mathbf{y} \in U$, and suppose without loss of generality that

$$\min\big\{\mathcal{F}(\mathbf{x}), \mathcal{F}(\mathbf{y})\big\} = \mathcal{F}(\mathbf{x}).$$

Now, we must prove that for any $\mathbf{x}, \mathbf{y} \in U$ and $\lambda \in [0,1]$, we have

$$\mathcal{F}(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) \geq \min\big\{\mathcal{F}(\mathbf{x}), \mathcal{F}(\mathbf{y})\big\} = \mathcal{F}(\mathbf{x}).$$

Suppose that is not true. This means there exist $\mathbf{x}, \mathbf{y} \in U$ and $\lambda \in [0,1]$, we have

$$\mathcal{F}(\lambda\mathbf{x} + (1-\lambda)\mathbf{y}) < \min\big\{\mathcal{F}(\mathbf{x}), \mathcal{F}(\mathbf{y})\big\} = \mathcal{F}(\mathbf{x}) \leq \mathcal{F}(\mathbf{y}).$$

For each $t \in [0,1]$, we define $\mathbf{z}_t = (1-t)\mathbf{x} + t\mathbf{y} = \mathbf{x} + t(\mathbf{y}-\mathbf{x})$, and $g(t) = \mathcal{F}((1-t)\mathbf{x} + t\mathbf{y})$. Since U is convex, then $\mathbf{z}_t \in U$. Since $g(\lambda) < g(0) = \mathcal{F}(\mathbf{x})$, there exists $\lambda_1 \in (0,\lambda)$ such that $g'(\lambda_1) = (g(\lambda) - g(0))/(\lambda - 0)$. Similarly, since $g(\lambda) < g(1) = \mathcal{F}(\mathbf{y})$, there exists $\lambda_2 \in (\lambda, 1)$ such that $g'(\lambda_2) = (g(\lambda) - g(1))/(\lambda - 1)$. Thus we have $0 < \lambda_1 < \lambda < \lambda_2 < 1$ and $g'(\lambda_2) > 0 > g(\lambda_1)$. This implies

$$g'(\lambda_1)(\lambda_2 - \lambda_1) < 0 \qquad \text{and} \qquad g'(\lambda_2)(\lambda_1 - \lambda_2) < 0.$$

If $g(\lambda_1) \geq g(\lambda_2)$, then $\mathcal{F}((1-\lambda_1)\mathbf{x} + \lambda_1\mathbf{y}) \geq \mathcal{F}((1-\lambda_2)\mathbf{x} + \lambda_2\mathbf{y})$. So, by the condition we get $g'(\lambda_2)(\lambda_1 - \lambda_2) \geq 0$. If $g(\lambda_1) \leq g(\lambda_2)$, by the condition we get $g'(\lambda_1)(\lambda_2 - \lambda_1) \geq 0$. Both cases give us a contradiction and the proof is completed. For more information you can see [19], [59]. ♠

This proof says that the angle between the gradient and the vector $(\mathbf{y}-\mathbf{x})$ is acute or right. Part (2) is illustrated in the following Figure 2.8 (Source of Figure 2.8: http://scottmccracken.weebly.com/uploads/9/0/6/6/9066859/convexity-print_version.pdf). We now explain another definition for quasiconvex functions. In this definition, the domain of the function $\mathcal{F}$ is the set $M_{m\times n}(\mathbb{R})$, the space of all $m \times n$ matrices. First, it is necessary to state a few concepts.

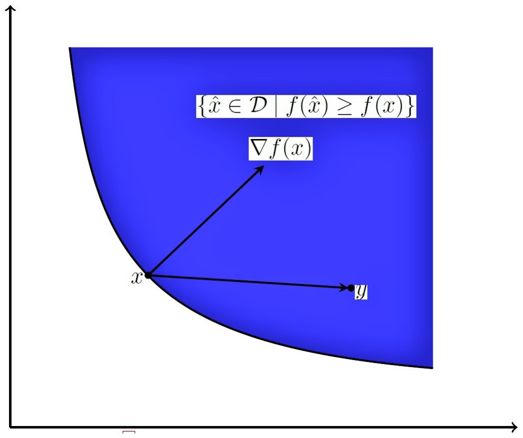

Figure 2.8: The angle between vectors $\nabla\mathbf{f}(\mathbf{x})$ and $(\mathbf{y}-\mathbf{x})$ is acute.

Definition 2.2.4. Let $U \subset \mathbb{R}^n$ be a bounded open set, $\varphi : U \subset \mathbb{R}^n \longrightarrow \mathbb{R}^m$ is continuous, has uniformly bounded gradient and $\varphi = 0$ on ∂U. Let $M_{m\times n}(\mathbb{R})$ be the space of all $m \times n$ matrices over the field of real numbers $\mathbb{R}$. Then a continuous function $\mathcal{F} : M_{m\times n}(\mathbb{R}) \longrightarrow \mathbb{R}$ is quasiconvex if

$$\int_U \big[\mathcal{F}(A + \nabla\varphi(\mathbf{x})) - \mathcal{F}(A)\big]\mathrm{dx} \geq 0 \Longleftrightarrow \frac{1}{\text{meas } U}\int_U \mathcal{F}(A + D\varphi(\mathbf{x}))\mathrm{dx} \geq \mathcal{F}(A), \tag{2.19}$$

for every $A \in M_{m\times n}(\mathbb{R})$, for every bounded open set $U \subset \mathbb{R}^n$ and every function φ with those conditions (See [19], [21], [69] and [67]).

In this definition $\nabla\varphi(\mathbf{x})$ stands for the gradient matrix of partial derivatives of the function φ. So,

$$\nabla\varphi(\mathbf{x}) = D\varphi(\mathbf{x}) = \begin{bmatrix} \frac{\partial\varphi_1}{\partial x_1} & \frac{\partial\varphi_1}{\partial x_2} & \cdots & \frac{\partial\varphi_1}{\partial x_n} \\ \frac{\partial\varphi_2}{\partial x_1} & \frac{\partial\varphi_2}{\partial x_2} & \cdots & \frac{\partial\varphi_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial\varphi_m}{\partial x_1} & \frac{\partial\varphi_m}{\partial x_2} & \cdots & \frac{\partial\varphi_m}{\partial x_n} \end{bmatrix},$$

where $\varphi(x) \in W_0^{1,\infty}(U, \mathbb{R}^n)$ (Sobolev Space); see [19]. In the next theorem (Theorem 2.2.2), we will explain that every polyconvex function is a quasiconvex function. That means the Examples 2.2.2 and $\mathcal{F}(A) := \det A$ where $A \in M_{m\times n}(\mathbb{R})$ are quasiconvex. But there are several examples with $m, n \geq 2$ of quasiconvex functions that are not polyconvex. We will see two of them in this chapter. The last main definition in this chapter is the concept of rank-one convex function that Charles B. Morrey introduced for the first time with another terminology [53], but the term rank-one was used by John M. Ball in [9].

Definition 2.2.5. Let $M_{m\times n}(\mathbb{R})$ denote the space of all $m \times n$ matrices over the field of real numbers $\mathbb{R}$ and U be an open subset of $M_{m\times n}(\mathbb{R})$. A function $\mathcal{F} : U \subset M_{m\times n}(\mathbb{R}) \longrightarrow \mathbb{R}$ is said to be **rank-one convex** if for every $A \in U$ the function of real variable $t \mapsto \mathcal{F}(A + tB)$ is convex near $t = 0$, whenever the rank of B is less than or equal to one [42].

Rank-one convexity expresses the convexity of a function along all line segments in $M_{m\times n}$ ($M_{m\times n}$ denotes the set all of real $m \times n$ matrices) whose the initial and end points differ by a matrix of rank one (see [67], [25]). This leads us to define another definition for rank-one convex function, as a result of the definition (2.2.5). It is easy to prove that these two definitions are equivalent (see [9], pages 352 and 353).

Definition 2.2.6. A function $\mathcal{F} : U \subset M_{m\times n}(\mathbb{R}) \longrightarrow \mathbb{R}$ is a rank-one convex function if $\mathcal{F}(tA + (1-t)B) \leq t\mathcal{F}(A) + (1-t)\mathcal{F}(B)$, whenever $t \in (0, 1)$ and $A,\ B \in M_{m\times n}(\mathbb{R})$ with $\operatorname{rank}(A - B) \leq 1$ (See [19], [20], [21], [42] and [67]).

It is not difficult to see that every rank-one convex function is continuous [67].

Theorem 2.2.2. Let $\mathcal{F} : M_{m\times n}(\mathbb{R}) \longrightarrow \mathbb{R}$. Then

1. $\mathcal{F}$ convex $\Rightarrow$ $\mathcal{F}$ polyconvex $\Rightarrow$ $\mathcal{F}$ quasiconvex $\Rightarrow$ $\mathcal{F}$ rank-one convex.
2. If $m = 1$ or $n = 1$, then all these notions are equivalent.

3. If $\mathcal{F} : M_{m\times n}(\mathbb{R}) \longrightarrow \mathbb{R}$ is convex, polyconvex, quasiconvex or rank-one convex, then $\mathcal{F}$ is locally Lipschitz.

Proof. There are long and complete proofs for this theorem in [11], [9], [54] and [19]. The converse implications of all parts of result (1) above are false, when $n \geq 3$. ♠

For instance, when $m = n = 2$, Example 2.2.2 or the function $\mathcal{F}(A) = \det(A)$ is polyconvex but not convex (See [21]). Also, we can find quasiconvex functions such that they are not polyconvex. The following two examples show quasiconvex functions which are not polyconvex. One example has been given by Sverak (see [68]). You can find the other given example by Alibert and Dacorogna in [2].

Example 2.2.4. (The example of Alibert-Dacorronga-Marcellini): Let $M_{2\times 2}(\mathbb{R})$ be the set of 2×2 real matrices endowed with Euclidean norm

$$|A|^2 = \sum_{i,j=1}^{2} a_{ij}^2 \qquad\qquad \text{for all} \;\; A \in M_{2\times 2}(\mathbb{R}).$$

Let $k \in \mathbb{R}$ and let $\mathcal{F}_\gamma : M_{m\times n}(\mathbb{R}) \longrightarrow \mathbb{R}$ be defined as

$$\mathcal{F}_\gamma(A) = |A|^2(|A|^2 - 2\gamma \det A).$$

Then

$$\begin{aligned}
\mathcal{F}_\gamma \;\text{ is convex} &\iff |\gamma| \leq \gamma_c = \frac{2}{3}\sqrt{2} \approx 0.942809,\\
\mathcal{F}_\gamma \;\text{ is polyconvex} &\iff |\gamma| \leq \gamma_p = 1,\\
\mathcal{F}_\gamma \;\text{ is quasiconvex} &\iff |\gamma| \leq \gamma_q = 1 + \epsilon, \qquad (\text{there exists } \;\epsilon > 0),\\
\mathcal{F}_\gamma \;\text{ is rank}-\text{one convex} &\iff |\gamma| \leq \gamma_r = \frac{2}{\sqrt{3}} \approx 1.1547.
\end{aligned}$$

The last result and the fact that if $\mathcal{F}_\gamma$ is polyconvex, then $|\gamma| \leq 1$, were established by Dacorogna-Marcellini [22]. All the other results were first proved in Alibert-Dacorogna [2]. The long and complete proof of this example exists in [19] (see pages 221 to 233). The example also provides a quasiconvex function that is not polyconvex. The problem of knowing if $\gamma_q = 2/\sqrt{3}$ is still open. if $\gamma_q < 2/\sqrt{3}$, then this would provide a rank-one convex function that is not quasiconvex (see [19]). The following picture shows the conditions in this example. You can see [75].

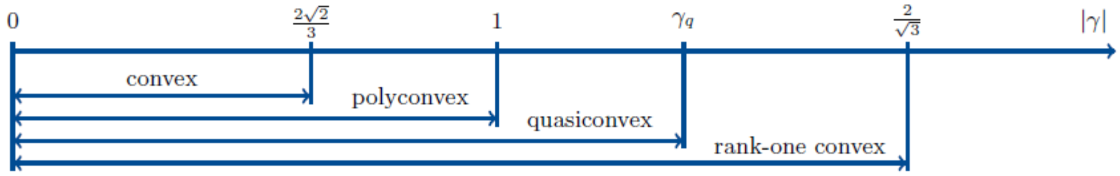

Figure 2.9: The example of Alibert-Dacorronga-Marcellini.

♣

The existence of a rank-one convex function that is not quasiconvex when $m \geq 3$, $n \geq 2$ was proved by Vladimir Sverak [69]. A classical result in this direction is that a quadratic function $\mathbf{f}$ is quasiconvex if and only if it is rank-one convex (see [11], [54] and [70]). This can be proved by using the Fourier transformation.
In the case $m = 2$ and $n \geq 2$, however, it is an **open problem** to determine whether

$$\mathcal{F} \text{ is rank-one convex} \nRightarrow \mathcal{F} \text{ is quasiconvex.}$$

Example 2.2.5. (Vladimir Sverak's example): We need to explain some lemmas and propositions that are useful for Sverak's example (see [69]). The proof of this example is divided into five steps.
Step 1: We do not prove the following lemmas, you can find their proofs in Sverak [69] or Dacorogna [19]. The following lemma can be found for example in [12].

Lemma 2.2.3. A continuous function $\mathcal{F}_\gamma : M_{m\times n}(\mathbb{R}) \longrightarrow \mathbb{R}$ is quasiconvex if and only if

$$\int_{[0,1]^n} \mathcal{F}(A + D\varphi(\mathbf{x}))\mathrm{dx} \geq \mathcal{F}(A), \tag{2.20}$$

for each $A \in M_{m\times n}(\mathbb{R})$ and each smooth function $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ which is periodic with respect to $\mathbb{Z}^n$. That means $\varphi(\mathbf{x} + \mathbf{k}) = \varphi(\mathbf{x})$ for each $\mathbf{x} \in \mathbb{R}^n$ and each $\mathbf{k} \in \mathbb{Z}^n$, and $\mathbf{k} = (k_1, k_2, \dots, k_n)$ where $k_i \in \mathbb{Z}$.

Step 2: Let $w : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ be a function given by

$$w(\mathbf{x}) = \frac{1}{2\pi}\begin{pmatrix} \sin(2\pi x_1) \\ \sin(2\pi x_2) \\ \sin(2\pi(x_1 + x_2)) \end{pmatrix}. \tag{2.21}$$

So,

$$\nabla w(\mathbf{x}) = Dw(\mathbf{x}) = \frac{1}{2\pi}\begin{pmatrix} \cos(2\pi x_1) & 0 \\ 0 & \cos(2\pi x_2) \\ \cos(2\pi(x_1 + x_2)) & \cos(2\pi(x_1 + x_2)) \end{pmatrix},$$

for each $\mathbf{x} \in \mathbb{R}^2$. For each $\mathbf{x} \in \mathbb{R}^2$ the gradient $Dw(\mathbf{x})$ of w lies in the linear subspace Ω of $M_{3\times 2}(\mathbb{R})$ given by

$$\Omega = \left\{ \begin{pmatrix} r & 0 \\ 0 & s \\ t & t \end{pmatrix} : \; r, s, t \in \mathbb{R} \right\}.$$

Let $\mathbf{f} : \Omega \longrightarrow \mathbb{R}$ be defined by

$$\mathbf{f}\left(\begin{pmatrix} r & 0 \\ 0 & s \\ t & t \end{pmatrix}\right) = -rst. \tag{2.22}$$

It is qbvious that the only rank-one directions in Ω are given by

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 1 \end{pmatrix}.$$

Thus the function $\mathbf{f}$ is convex (in fact linear) on each rank-one line contained in Ω. By expanding formula

$$\cos(a+b) = \cos a \cos b - \sin a \sin b,$$

we have

$$\int_{[0,1]^2} \mathbf{f}(Dw(\mathbf{x}))\mathrm{dx} = -\int_{[0,1]^2} (\cos 2\pi x_1)^2 (\cos 2\pi x_2)^2 \mathrm{dx}_1 \mathrm{dx}_2 < 0 = \mathbf{f}(0). \tag{2.23}$$

Now we can extend the function $\mathbf{f}$ to a rank-one convex function on $M_{3\times2}(\mathbb{R})$. This is not obvious and easy.

Step 3: We can consider $M_{3\times2}(\mathbb{R})$ as the six-dimensional Euclidean space. The norm of $A \in M_{3\times2}(\mathbb{R})$ is given by

$$|A|^2 = \sum_{\substack{1\le i\le 3\\ 1\le j\le 2}} (a_{ij})^2.$$

Now we can state the following lemma without its proof. This lemma was proved by Sverak in [69], pages 187 and 188. The keen reader can also find this proof in Dacorogna [19].

Lemma 2.2.4. Let Ω be the linear subspace of $M_{3\times2}(\mathbb{R})$ defined in Step 2 and let $P : M_{3\times2}(\mathbb{R}) \longrightarrow \Omega$ be the orthogonal projection such that

$$P(A) = \begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \\ \frac{(a_{31}+a_{32})}{2} & \frac{(a_{31}+a_{32})}{2} \end{bmatrix}, \qquad \text{for all } \mathrm{A} \in \mathrm{M}_{3\times2}(\mathbb{R}).$$

Let $\mathbf{f} : \Omega \longrightarrow \mathbb{R}$ be defined by Equation 2.22. Then for each $\epsilon > 0$, there exists $k = k(\epsilon) > 0$ such that the function $\mathcal{F}_\gamma : M_{m\times n}(\mathbb{R}) \longrightarrow \mathbb{R}$ given by

$$\mathcal{F}(A) = \mathbf{f}(P(A)) + \epsilon|A|^2 + \epsilon|A|^4 + k|A - P(A)|^2 \tag{2.24}$$

is rank-one convex on $M_{m\times n}(\mathbb{R})$.

Step 4: In this step, we are ready to explain a theorem that gives us a rank-one convex function that is not quasiconvex.

Theorem 2.2.5. There exists $\epsilon > 0$ and $k > 0$ such that the function $\mathcal{F}$ given by (2.24) is rank-one convex but is not quasiconvex.

Proof. Let w be the function defined (2.21). Since Dw is bounded, we see from (2.23) that we can choose $\epsilon > 0$ such that

$$\int_{[0,1]^2} \Big(\mathbf{f}(Dw(\mathbf{x})) + \epsilon|Dw(\mathbf{x})|^2 + \epsilon|Dw(\mathbf{x})|^4\Big)\mathrm{dx} < 0. \tag{2.25}$$

By Lemma 2.2.4 we can find $k = k(\epsilon)$ such that the function

$$\mathcal{F}(A) = \mathbf{f}(P(A)) + \epsilon|A|^2 + \epsilon|A|^4 + k|A - P(A)|^2$$

is rank-one convex. Since $|Dw(\mathbf{x}) - P(Dw(\mathbf{x}))| = 0$ for all $\mathbf{x} \in \mathbb{R}^2$, we have from (2.25) that

$$\int_{[0,1]^2} \mathcal{F}(Dw(\mathbf{x}))\mathrm{dx} < 0 = \mathcal{F}(0), \tag{2.26}$$

and using Lemma 2.2.3, we see that $\mathcal{F}$ is not quasiconvex. The proof of Theorem 2.2.5 is completed. ♠

Notice that the function $\mathcal{F}$ is a polynomial of degree four. We have finally reached the final step that can complete the proof of Sverak's example.
Step 5: In this step we prove a corollary of Theorem 2.2.5 that generalises the example to higher dimensions.

Corollary 2.2.5.1. Let $n \geq 2$ and $m \geq 3$ and let $T : M_{m\times n}(\mathbb{R}) \longrightarrow M_{3\times 2}(\mathbb{R})$ is defined by

$$T(X) = \begin{bmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \end{bmatrix}.$$

Let $\mathcal{F}_\circ : M_{m\times n}(\mathbb{R}) \longrightarrow \mathbb{R}$ be defined $\mathcal{F}_\circ(X) = \mathcal{F}(T(X))$, where $\mathcal{F}$ is the function from the theorem above. Then $\mathcal{F}_\circ$ is rank-one convex but is not quasiconvex.

It is easy to see, because T maps rank-one lines to rank-one lines, $\mathcal{F}_\circ$ is rank-one convex. Considering the periodic function

$$w_\circ(x_1, x_2, \ldots, x_n) = (w_1(x_1, x_2), w_2(x_1, x_2), w_3(x_1, x_2), 0, \ldots, 0),$$

where w is the function defined above. We see from (2.26) that $\mathcal{F}_\circ$ is not quasiconvex. The example is finished.

♣

If $m = n = 2$ the above implication would have far-reaching consequences. In the article [15], the authors claimed to have found a counterexample for the case $m = n = 2$, that $\mathcal{F}$ is a rank-one convex function but is not quasiconvex. But in the article [57], Patrizio Neff has shown that the example in [15] is not a counterexample and this case is still an open problem.

Chapter 3

Quasiconformal Mappings

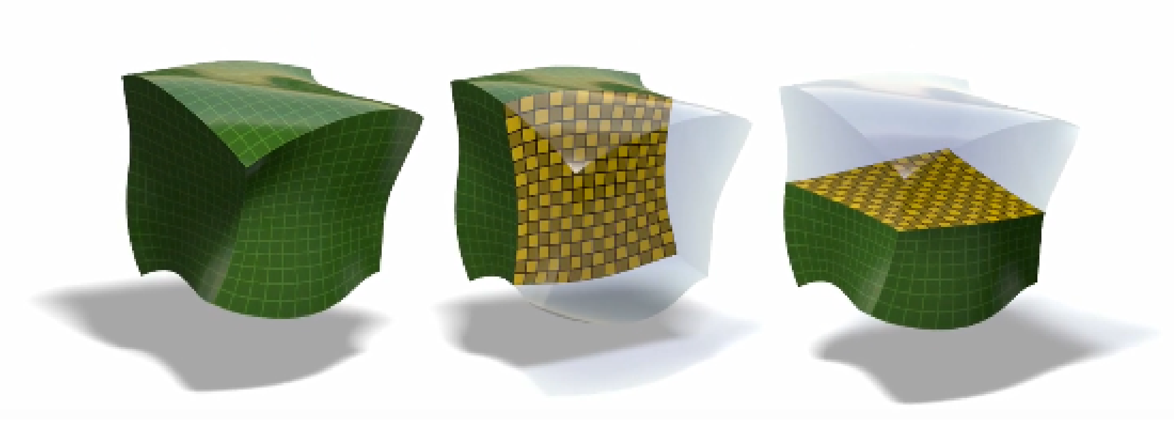

Figure 3.1: Quasiconformal Maps

Source: https://www.youtube.com/watch?v=iOwPGG5-54Q

Chapter 3

Quasiconformal Mappings

Quasiconformal mappings are generalizations of conformal mappings. They can be considered not only in Riemann surfaces but also on Riemannian manifolds in all domains, even on arbitrary metric spaces. Quasiconformal mappings occur naturally in various mathematical and often a priori unrelated contexts. In mathematics, the distortion is a measure of the amount by which a function from the Euclidean plane to itself distorts circles to ellipses. There are three main definitions for quasiconformal mappings in Euclidean space: geometric, analytic, and metric definitions. In Introduction 0.1 we defined the linear distortion and the history of the quasiconformal maps. For more details, you can see [28], [31], [29], [35], [74], [30], [6] and [42].

In this chapter, at first, we recall the curve family definition. In the two first sections of this chapter, we explain the concept and definitions of linear, outer, and inner distortion and their properties. The third section aims to show the lower semicontinuity property of the inner, outer and maximal distortion functions. We will finish this chapter with Iwaniec's Lemma and Theorem that show the linear distortion function, is not lower semicontinuous. Also we illustrate with some figures and examples (see [42] and [30]).

3.1 The Moduli of Curve Families

In this section, we will explain the notion of the moduli of curve families. This subject is used to describe quasiconformal mappings and their properties in this chapter. We are starting with some essential definitions and theorems.

Definition 3.1.1. A **path** in $\hat{\mathbb{R}}^n$ is a continuous mapping $\gamma : I \longrightarrow \hat{\mathbb{R}}^n$ where I is an interval in $\mathbb{R}$. The path is said to be closed or open if I is closed or open, respectively. The **locus** $|\gamma|$ of a path $\gamma : I \longrightarrow \hat{\mathbb{R}}^n$ is the point set $\gamma(I) \subset \hat{\mathbb{R}}^n$. The **subpath** of $\gamma : I \longrightarrow \hat{\mathbb{R}}^n$ is a restriction of γ to a subinterval of I. Suppose $a = t_0 \leq t_1 \leq \ldots \leq t_k = b$ is a partition of interval $[a, b]$. Then we define the **length** of $\gamma : [a, b] \longrightarrow \hat{\mathbb{R}}^n$ and denote by $\ell(\gamma)$ as

$$\ell(\gamma) = \sup \Big\{ \sum_{i=1}^{k} |\gamma(t_i) - \gamma(t_{i-1})| \Big\}. \tag{3.1}$$

Thus $0 \leq \ell(\gamma) \leq \infty$ and it is obvious to see $\ell(\gamma) = 0$ if and only if γ is a constant. If $\ell(\gamma) < \infty$ then γ is called **rectifiable path**. A path in $\hat{\mathbb{R}}^n$ such that $\infty \in |\gamma|$ is non-rectifiable, except for the constant path $\gamma(t) = \infty$ (see [74] p.1).

Theorem 3.1.1. Let $\gamma : [a,b] \longrightarrow \hat{\mathbb{R}}^n$ be a rectifiable path, and let $a = t_0 \leq t_1 \leq \ldots \leq t_k = b$ be a partition of interval $[a,b]$. Then every restriction $\gamma|[t_{i-1} - t_i]$ is rectifiable, and

$$\ell(\gamma) = \sum_{i=1}^{k} \ell(\gamma|[t_{i-1}, t_i]). \tag{3.2}$$

Proof. This proof is obvious (see [74], [30]). ♠

Definition 3.1.2. The path $\gamma^o : [0, \ell(\gamma)] \longrightarrow \hat{\mathbb{R}}^n$ is the **normal representation** of γ (see [30], [74]). It is also called the parametrisation of γ by means of its arc length.

Definition 3.1.3. A path $\gamma : [a,b] \longrightarrow \hat{\mathbb{R}}^n$ is **locally rectifiable** if each compact subpath of γ is rectifiable. We then denote

$$\ell(\gamma) = \sup \ell(\alpha), \qquad \text{over all compact subpaths } \alpha \text{ of } \gamma.$$

It is obvious that the concepts rectifiable and locally rectifiable coincide for closed paths. Furthermore, for a closed path γ, the two definitions of $\ell(\gamma)$ are equivalent. For instance, the path $\gamma : (-1,1) \longrightarrow \mathbb{R}^2$ defined by

$$\gamma(t) = (t, t\sin(\frac{1}{t})) \in \mathbb{R}^2,$$

is not locally rectifiable. Because its subpath $\gamma|(0,1)$ is locally rectifiable but not rectifiable [74]. We can prove that, if $\gamma : (a,b) \longrightarrow \hat{\mathbb{R}}^n$ is a rectifiable open path, then there exists a unique extension to closed path $\gamma_* : (a,b) \longrightarrow \hat{\mathbb{R}}^n$, Also $\ell(\gamma) = \ell(\gamma_*)$. Let $\gamma : [a,b] \longrightarrow \hat{\mathbb{R}}^n$ be a rectifiable path. For all $t \in [a,b]$, we denote $\ell(\gamma|[a,t])$ by $s_\gamma(t)$ or only by $s(t)$. The function $s_\gamma : [a,b] \longrightarrow \mathbb{R}$ is called the **length function** of γ.

Lemma 3.1.2. Let $\gamma : [a,b] \longrightarrow \hat{\mathbb{R}}^n$ be a rectifiable path of length ℓ, and let $s_\gamma = s$ be its length function. Then

1. $|\gamma(t) - \gamma(u)| \leq s(t) - s(u)$, where $a \leq u \leq t \leq b$;
2. s is nondecreasing and continuous;
3. s is absolutely continuous if and only if γ is as well;
4. $s'(t)$ and $\gamma'(t)$ exist and satisfy $s'(t) = |\gamma'(t)|$ for almost every t in $[a,b]$;
5. both s' and γ' belong to $L^1([a,b])$ and

$$\int_a^b |\gamma'(t)|\mathrm{dt} = \int_{\mathrm{a}}^{\mathrm{b}} |\mathrm{s}'(\mathrm{t})|\mathrm{dt} \leq \ell, \tag{3.3}$$

with equality holding precisely when γ is absolutely continuous.

Proof. The proof of this lemma is long and we leave it to the keen reader (see [30], [74]). ♠

Definition 3.1.4. Assume that $\Omega \subset \mathbb{R}^n$ is a Borel set and that $\rho : A \longrightarrow \hat{\mathbb{R}}$ is a nonnegative Borel function. For each rectifiable closed path $\gamma : [a, b] \longrightarrow \Omega$, we define the line integral of ρ over γ as follows:

$$\int_\gamma \rho \,\mathrm{ds} = \int_0^{\ell(\gamma)} \rho(\gamma^{\mathrm{o}}(\mathrm{t}))\mathrm{dt}, \tag{3.4}$$

where γ^o is the normal representation of γ.

The integral on the right exists, because $\rho \circ \gamma^o$ is a nonnegative Borel function. Its value may be $+\infty$. We can show that if $\gamma : [a, b] \longrightarrow \Omega$ is absolutely continuous, then

$$\int_\gamma \rho \,\mathrm{ds} = \int_0^{\ell(\gamma)} \rho(\gamma^{\mathrm{o}}(\mathrm{t}))|\gamma'(\mathrm{t})|\mathrm{dt} \tag{3.5}$$

In section (1.5) we defined conformal mappings. Now we want to give another definition for it.

Definition 3.1.5. Let Ω and Ω' be domains in $\mathbb{R}^n$, and $\mathbf{f} : \Omega \longrightarrow \Omega'$ be a homeomorphism. Then $\mathbf{f}$ is conformal if $\mathbf{f} \in W^{1,2}_{loc}$ and $|D\mathbf{f}(\mathbf{x})h| = |D\mathbf{f}(\mathbf{x})||h|$ for almost all $\mathbf{x} \in \Omega$ and $h \in \mathbb{R}^n$.

We can also say a C^1 homeomorphism $\mathbf{f}$ is conformal if and only if $|D\mathbf{f}(\mathbf{x})|^n = J_{\mathbf{f}}(\mathbf{x})$ for all $\mathbf{x} \in \Omega$. By **Liouville's theorem**, if Ω is a domain in $\mathbb{R}^n$ with $n \geq 3$, then any conformal mapping $\mathbf{f} : \Omega \longrightarrow \mathbb{R}^n$ is a restriction to Ω of a Möbius transformation (see [42] and [30]). It is well known that a two-dimensional diffeomorphism with positive Jacobian is conformal if and only if it is complex analytic. We also know that, for $n \geq 3$, every conformal mapping is a Möbius transformation. Suppose that Γ is a **family of smooth paths in** Ω. Γ is called a **curve family** and it is simply a collection of curves. Let Ω and Ω' be domains in $\mathbb{R}^n$, with $n \geq 2$, and $\mathbf{f} : \Omega \longrightarrow \Omega'$ be a diffeomorphism. Assume that Γ is a family of smooth paths in Ω. We denote $\mathbf{f}(\Gamma)$ for $\{\mathbf{f} \circ \gamma \,:\, \gamma \in \Gamma\}$. So, $\mathbf{f}(\Gamma)$ is also a family of smooth paths in Ω'.
Given a curve family Γ, an **admissible density** is a Borel function $\rho : \Omega \longrightarrow [0, \infty)$, if a ρ-length $\ell_\rho(\gamma)$ to any smooth path $\gamma \in \Gamma$:

$$\ell_\rho(\gamma) = \int_\gamma \rho(x)|\mathrm{dx}| \geq 1 \qquad \text{for every } \gamma \in \Gamma. \tag{3.6}$$

In the other words, the continuous function $\rho : \Omega \longrightarrow [0, \infty)$ is an admissible density if for all $\gamma \in \Gamma$, we have $\ell_\rho(\gamma) \geq 1$. There is an analogous notion of ρ-volume: the ρ-volume $\upsilon_\rho(A)$ of a Lebesgue measurable subset A of Ω is given by

$$\upsilon_\rho(A) = \int_A \rho^n(x)\mathrm{dx}, \tag{3.7}$$

dx denoting integration with respect to Lebesgue measure in $\mathbb{R}^n$. The collection of all such admissible densities is denoted by Adm(Γ) (see [30] p.78).
The condition $\ell_\rho(\gamma) \geq 1$ prevents ρ from being, on average, excessively small along the trajectory of any $\gamma \in \Gamma$. So, if ρ is an admissible density for Γ and if Γ contains sufficiently many paths and their collective trajectories fill up a substantial portion of Ω, one might expect that the associated volume $\upsilon_\rho(\Omega)$ could not itself be small.

Definition 3.1.6. Let $p \geq 1$. The **p-modulus of curve family** Γ is

$$M_p(\Gamma) = \inf\{\upsilon_\rho(\Omega) \ : \ \rho \in \mathrm{Adm}(\Gamma)\} = \inf \int_{\mathbb{R}^n} \rho^p(x)\mathrm{dx}, \tag{3.8}$$

where the infimum is over all admissible densities for Γ.

If $\mathrm{Adm}(\Gamma) = \phi$ then we set $M_p(\Gamma) = +\infty$. This happens only if Γ contains a constant path, because if $\gamma \in \Gamma$ is a constant curve, the condition

$$\int_\gamma \rho(x)|\mathrm{dx}| \geq 1, \qquad \text{for every } \gamma \in \Gamma,$$

can not be satisfied since the integral is always zero. Therefore, there are no admissible densities for a curve family with a constant curve. In the definition of the modulus of a curve family we have infimum taken over an empty set. So, $\inf \phi = +\infty$ (recall that $\inf \phi = +\infty$ and $\sup \phi = -\infty$ on the extended real numbers $\hat{\mathbb{R}} = \mathbb{R} \cup \{+\infty, -\infty\}$). From the definition of the modulus of curve family, we have $0 \leq M_p(\Gamma) \leq \infty$ (see [30], [74]). The case $p = n$ is the most important. In this case we simply write $M(\Gamma)$ for the modulus. There are some properties of the modulus of a curve family and we explain and list them.

1. $\mathrm{Adm}(\Gamma) = \phi$ if and only if Γ contains a single constant path.
2. If $\Gamma_1 \subset \Gamma_2$, then $M(\Gamma_1) \leq M(\Gamma_2)$.
3. $M(\phi) = 0$.
4. $M(\bigcup_{i=1}^{\infty} \Gamma_i) \leq \sum_{i=1}^{\infty} M(\Gamma_i)$.
5. If Γ_1 and Γ_2 are curve families such that every curve in Γ_2 has a subcurve in Γ_1, then $M(\Gamma_1) \geq M(\Gamma_2)$.

All proofs exist in Väiälä's book [74] from page 16 to 20. One of the most important results is that modulus is a conformal invariant. This will be proved by Theorem 3.1.5.

Definition 3.1.7. Let Γ_1 and Γ_2 be curve families in $\mathbb{R}^n$. We say that Γ_2 is minimized by Γ_1 and denote $\Gamma_2 \succ \Gamma_1$, if every $\gamma \in \Gamma_2$ has a subcurve which belongs to Γ_1. A curve $\beta : J \longrightarrow \hat{\mathbb{R}}^n$ is called of a subcurve $\gamma : I \longrightarrow \hat{\mathbb{R}}^n$, if J is a subinterval of I and β is the restriction of γ to J and it is shown by $\gamma \succ \beta$.

Theorem 3.1.3. If $\Gamma_2 \succ \Gamma_1$, then $M_p(\Gamma_2) \leq M_p(\Gamma_1)$.

Proof. The assertion is trivial if $M_p(\Gamma_1) = +\infty$. If $M_p(\Gamma_1) < +\infty$, consider a function ρ from $\mathrm{Adm}(\Gamma_1)$. If a curve from Γ_1 is locally rectifiable and if α is chosen so that $\alpha \succ \gamma$, then

$$\int_\alpha \rho \,\mathrm{ds} \geq \int_\gamma \rho \,\mathrm{ds} \geq 1.$$

This means that ρ belongs to $\mathrm{Adm}(\Gamma_2)$. Thus

$$M_p(\Gamma_2) \leq \int_{\mathbb{R}^n} \rho^p \,\mathrm{dm_n}.$$

Since ρ was an arbitrary member of $\mathrm{Adm}(\Gamma_1)$, the inequality $M_p(\Gamma_2) \leq M_p(\Gamma_1)$ follows when the infimum ρ is taken over all such ρ (see [30], [74]). ♠

Theorem 3.1.4. Suppose that the curves of a family Γ lie in a Borel set $G \subset \mathbb{R}^n$, and for every locally rectifiable $\gamma \in \Gamma$, we have $\ell(\gamma) \geq r > 0$. Then

$$M_p(\Gamma) \leq \frac{m(G)}{r^p}.$$

Proof. Define the function $\rho : \mathbb{R}^n \longrightarrow \mathbb{R}$ by

$$\rho(\mathbf{x}) = \begin{cases} \dfrac{1}{r} & \text{if } \mathbf{x} \in G \\ \\ +\infty & \text{if } \mathbf{x} \notin G. \end{cases}$$

Then

$$\int_\gamma \rho \,\mathrm{ds} = \int_\gamma \frac{1}{\mathrm{r}} \,\mathrm{ds} = \ell(\gamma).\frac{1}{\mathrm{r}} \geq \mathrm{r}.\frac{1}{\mathrm{r}} = 1,$$

so ρ finds itself in $\mathrm{Adm}(\Gamma)$, and

$$M_p(\Gamma) \leq \int_{\mathbb{R}^n} \rho^p \,\mathrm{dm_n} = \int_{\mathbb{R}^\mathrm{n}} \frac{1}{\mathrm{r^p}} \,\mathrm{dm_n} = \frac{\mathrm{m(G)}}{\mathrm{r^p}}.$$

This completes the proof (see [30], [74] and [73]). ♠

Here is some notation that we will use in the following examples:
Let $E, F, G \subset \hat{\mathbb{R}}^n$, the notation $\Delta(E, F : G)$ stands for the family of open paths with endpoints that connect E and F through G. More precisely, a path $\gamma : [a, b] \longrightarrow \hat{\mathbb{R}}^n$ belongs to $\Delta(E, F : G)$ if and only if $\gamma(a) \in E$, $\gamma(b) \in F$ and for $a < t < b$ we have $\gamma(t) \in G$ (see [30], [74] and [73]).

Example 3.1.1. (The Cylinder) Let E be a Borel set in $\mathbb{R}^{n-1}$ and let $h > 0$. Set

$$G = \{\mathbf{x} \in \mathbb{R}^n : \ (x_1, x_2, \ldots, x_{n-1}) \in E \ \text{ and } \ 0 < \mathrm{x_n} < \mathrm{h}\}.$$

Then G is a cylinder with bases E and $F = E + he_n$ and with height h. Set $\Gamma = \Delta(E, F : G)$.

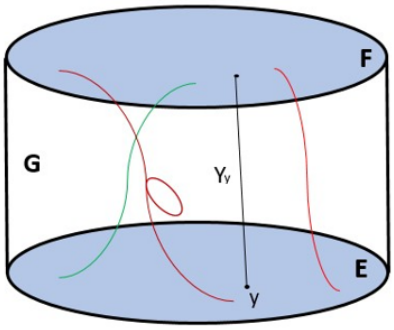

Figure 3.2: The cylinder of the family of open paths with endpoints that connect E and F through G.

We show that

$$M_p(\Gamma) = \frac{m(G)}{h^p} = \frac{h.m_{n-1}(E)}{h^p} = \frac{m_{n-1}(E)}{h^{p-1}}.$$

Since $\ell(\gamma) \geq h$, for every $\gamma \in \Gamma$, Theorem 3.1.4 implies

$$M_p(\Gamma) \leq \frac{m(G)}{h^p}.$$

Let ρ be an arbitrary function in $\mathrm{Adm}(\Gamma)$. For each $y \in E$, let $Y_y : [0, h] \longrightarrow \mathbb{R}^n$ be the vertical segment $Y_y(t) = y + te_n$. Then $Y_y \in \Gamma$. Suppose that $p > 1$, we obtain by Hölder's inequality

$$\begin{aligned} 1 &\leq \Big(\int_{Y_y} 1.\rho \,\mathrm{ds}\Big)^{\mathrm{p}} \leq \Big[\Big(\int_0^{\mathrm{h}} 1^{\mathrm{q}} \,\mathrm{ds}\Big)^{\frac{1}{\mathrm{q}}}\Big(\int_0^{\mathrm{h}} \rho^{\mathrm{p}} \,\mathrm{ds}\Big)^{\frac{1}{\mathrm{p}}}\Big]^{\mathrm{p}} \\ &= h^{\frac{p}{q}} \int_0^h \rho^p \,\mathrm{ds} = \mathrm{h}^{\mathrm{p}-1} \int_0^{\mathrm{h}} \rho^{\mathrm{p}}(\mathrm{y} + \mathrm{te_n}) \,\mathrm{dt}. \end{aligned}$$

So, we have

$$1 \leq h^{p-1} \int_0^h \rho^p(y + te_n) \,\mathrm{dt}.$$

Integration over $y \in E$ yields by Fubini's theorem

$$m_{n-1}(E) \leq h^{p-1} \int_E \mathrm{dm}_{\mathrm{n}-1} \int_0^{\mathrm{h}} \rho^{\mathrm{p}}(\mathrm{y} + \mathrm{te_n}) \,\mathrm{dt} = \mathrm{h}^{\mathrm{p}-1} \int_{\mathrm{G}} \rho^{\mathrm{p}} \,\mathrm{dm} = \mathrm{h}^{\mathrm{p}-1} \int \rho^{\mathrm{p}} \,\mathrm{dm}.$$

Since this holds for every $\rho \in \mathrm{Adm}(\Gamma)$, we get

$$M_p(\Gamma) \geq \frac{m_{n-1}(E)}{h^{p-1}}.$$

Thus, we proved that $M_p(\Gamma) = M_p(\Gamma_\circ)$, where $\Gamma_\circ$ is the subfamily of Γ consisting of the vertical segments Y_y (see [30], p.101 and [74], p.21).

♣

Remark. In Example 3.1.1, the modulus is invariant under similarity mappings if and only if $p = n$. This is the reason why the case $p = n$ is so important in the theory of quasiconformal mappings. Later in this section we will prove that $M(\Gamma)$ is a conformal invariant (see [74]). For the next example we use the following notation

$$\begin{aligned} B^n(\mathbf{x}, r) &= \{\mathbf{y} \in \mathbb{R}^n : |\mathbf{x} - \mathbf{y}| < r\}, \\ \bar{B}^n(\mathbf{x}, r) &= \{\mathbf{y} \in \mathbb{R}^n : |\mathbf{x} - \mathbf{y}| \leq r\}, \\ \mathbb{S}^{n-1}(\mathbf{x}, r) &= \{\mathbf{y} \in \mathbb{R}^n : |\mathbf{x} - \mathbf{y}| = r\}. \end{aligned}$$

If $r = 1$, then the unit ball and the unit sphere are shown by $B^n(1) = B^n$ and $\mathbb{S}^{n-1}(1) = \mathbb{S}^{n-1}$, respectively. The boundary of the n-dimensional ball is shown by $\partial B^n(\mathbf{x}, r)$. So, we can write $\mathbb{S}^{n-1}(\mathbf{x}, r) = \partial B^n(\mathbf{x}, r)$. The notation $\bar{B}^n(\mathbf{x}, r)$ is called the closure of $B^n(\mathbf{x}, r)$, and it is obvious to see

$$\bar{B}^n(\mathbf{x}, r) = B^n(\mathbf{x}, r) \cup \partial B^n(\mathbf{x}, r).$$

Recall that the volume of the n-dimensional unit ball can be expressed by

$$V_n = \frac{\pi^{\frac{n}{2}}}{\Gamma(\frac{n}{2} + 1)}, \tag{3.9}$$

where Γ is Euler's gamma function. The surface area of $\mathbb{S}^{n-1}$ is denoted by ω_{n-1}. It is well known that $\omega_{n-1} = nV_n$.

Example 3.1.2. Let x_0 be a point of $\mathbb{R}^n$, and $0 < a < b < \infty$. Let $\Gamma = \Delta(E, F : G)$, where $E = \mathbb{S}^{n-1}(\mathbf{x_0}, a)$, $F = \mathbb{S}^{n-1}(\mathbf{x_0}, b)$, and $G = B^n(\mathbf{x_0}, b) \setminus \bar{B}^n(\mathbf{x_0}, a)$. Then for $1 < p < \infty$, it is true that

$$M_p(\Gamma) = \omega_{n-1}\Big(\log\frac{b}{a}\Big)^{1-n},$$

where ω_{n-1} shows the surface area of $\mathbb{S}^{n-1}$. Let $\rho \in \mathrm{Adm}(\Gamma)$. For each unit vector $\mathbf{y} \in \mathbb{S}^{n-1}$, we define $Y_{\mathbf{y}} : [a, b] \longrightarrow \mathbb{R}^n$ to be the radial segment, given by $Y_{\mathbf{y}}(t) = t\,\mathbf{y}$.

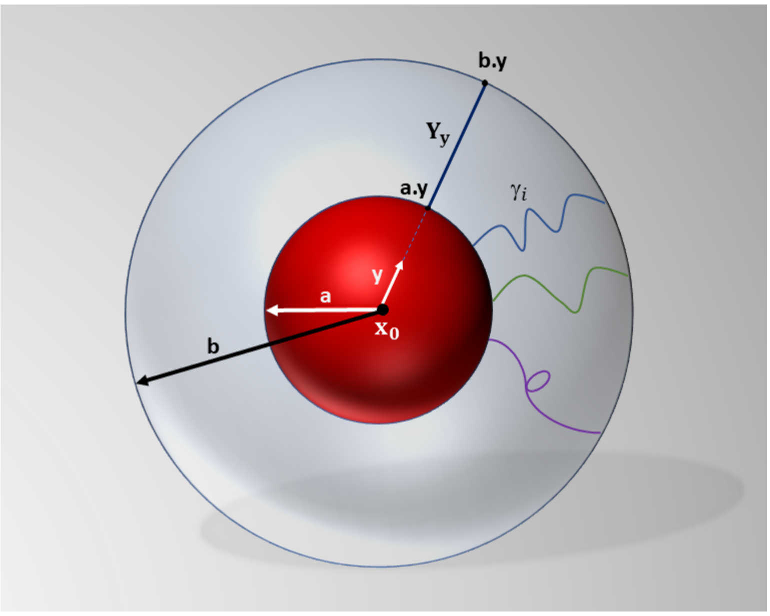

Figure 3.3: The spherical ring with $0 < a < b < \infty$

By Hölder's inequality (for $p = \frac{n}{n-1}$ and $q = n$), we obtain

$$\begin{aligned}
1 &\le \Big(\int_{Y_{\mathbf{y}}} 1\,\rho\,\mathrm{ds}\Big)^{\mathrm{n}} = \Big(\int_{\mathrm{a}}^{\mathrm{b}} \big(\rho(\mathrm{t}\mathbf{y})\,\mathrm{t}^{1-\frac{1}{\mathrm{n}}}\big)\mathrm{t}^{\frac{1}{\mathrm{n}}-1}\,\mathrm{dt}\Big)^{\mathrm{n}} \\
&\le \Bigg[\Big(\int_a^b \big(\rho(t\mathbf{y})\,t^{1-\frac{1}{n}}\big)^n\,\mathrm{dt}\Big)^{\frac{1}{\mathrm{n}}}\Big(\int_{\mathrm{a}}^{\mathrm{b}} \big(\mathrm{t}^{\frac{1}{\mathrm{n}}-1}\big)\,\mathrm{dt}\Big)^{\frac{\mathrm{n}-1}{\mathrm{n}}}\Bigg]^{\mathrm{n}} \\
&= \Big(\int_a^b \rho^n(t\mathbf{y})\,t^{n-1}\,\mathrm{dt}\Big)\Big(\int_{\mathrm{a}}^{\mathrm{b}} \mathrm{t}^{-1}\,\mathrm{dt}\Big)^{\mathrm{n}-1} \\
&= \big(\log\frac{b}{a}\big)^{n-1}\int_a^b \rho^n(t\mathbf{y})\,t^{n-1}\,\mathrm{dt}
\end{aligned}$$

So, by integrating over $\mathbf{y} \in \mathbb{S}^{n-1}$, we have

$$\omega_{n-1} \le \big(\log\frac{b}{a}\big)^{n-1}\int \rho^n\,\mathrm{dm}. \tag{3.10}$$

Taking the infimum over all $\rho \in \mathrm{Adm}(\Gamma)$, we obtain

$$\omega_{n-1} \le \big(\log\frac{b}{a}\big)^{n-1} M(\Gamma_G). \tag{3.11}$$

Next we define the function ρ given by

$$\rho(\mathbf{x}) = \begin{cases} \dfrac{1}{|\mathbf{x}| \log \frac{b}{a}}, & \text{if } \ \mathbf{x} \in G, \\ \\ 0, & \text{if } \mathbf{x} \notin G. \end{cases}$$

It is obvious that ρ is admissible for Γ_G, because

$$\ell_\rho(\gamma) \geq \int_a^b \frac{1}{|\mathbf{x}| \log \frac{b}{a}} \, \mathrm{dx} = \frac{1}{\log \frac{\mathrm{b}}{\mathrm{a}}} \int_{\mathrm{a}}^{\mathrm{b}} \frac{1}{|\mathbf{x}|} \, \mathrm{dx} = 1,$$

so, $\rho \in \mathrm{Adm}(\Gamma_{\mathrm{G}})$ and hence

$$M(\Gamma_G) \leq \int_{\mathbb{R}^n} \rho^n \, \mathrm{dm} = \omega_{\mathrm{n}-1} \big(\log \frac{\mathrm{b}}{\mathrm{a}} \big)^{-\mathrm{n}} \int_{\mathrm{a}}^{\mathrm{b}} \frac{1}{\mathrm{s}} \, \mathrm{ds} = \omega_{\mathrm{n}-1} \big(\log \frac{\mathrm{b}}{\mathrm{a}} \big)^{1-\mathrm{n}}. \tag{3.12}$$

By (3.11) and (3.12), we have

$$M(\Gamma_G) = \omega_{n-1} \Big(\log \frac{b}{a} \Big)^{1-n},$$

and this example is proved (see [30], [74], [61] and [73]).

♣

Let A be a subset of $\hat{\mathbb{R}}^n$ and that $\mathbf{f} : A \longrightarrow \hat{\mathbb{R}}^n$ is continuous. If Γ is a family of paths in A, then the family $\tilde{\Gamma} = \{\mathbf{f} \circ \gamma : \ \gamma \in \Gamma\}$ is called the image of Γ under $\mathbf{f}$.

Theorem 3.1.5. Let Ω and Ω' be domains in $\mathbb{R}^n$, $\mathbf{f} : \Omega \longrightarrow \Omega'$ is a conformal map and $\tilde{\Gamma} = \mathbf{f}(\Gamma)$, then $M(\Gamma) = M(\tilde{\Gamma})$, for all $\Gamma \subset \Omega$.

Proof. Let $\tilde{\rho} \in \mathrm{Adm}(\tilde{\Gamma})$, now we can define $\rho(x) = \tilde{\rho}(\mathbf{f}(x)) \, |D\mathbf{f}(x)|$. We have $\rho \in \mathrm{Adm}(\Gamma)$ since for all locally rectifiable $\gamma \in \Gamma$

$$\int_\gamma \rho \mathrm{ds} = \int_\gamma \tilde{\rho}(\mathbf{f}(x)) \, |D\mathbf{f}(x)| \mathrm{ds} = \int_{\mathbf{f} \circ \gamma} \tilde{\rho}(x) \mathrm{ds} \geq 1.$$

Therefore, we can write

$$\begin{aligned} M(\Gamma) &\leq \int_\Omega \rho^n \mathrm{dx} \\ &= \int_\Omega (\tilde{\rho}(\mathbf{f}(x)))^n \, |D\mathbf{f}(x)|^n \mathrm{dx} \\ &= \int_\Omega (\tilde{\rho}(\mathbf{f}(x)))^n \, J_{\mathbf{f}}(x) \mathrm{dx} \\ &= \int_\Omega (\tilde{\rho}(x))^n \mathrm{dx}. \end{aligned}$$

Now by taking the infimum over all $\tilde{\rho}$ we obtain $M(\Gamma) \leq M(\tilde{\Gamma})$. Since $\mathbf{f}$ is conformal, $\mathbf{f}^{-1}$ is conformal. So, with the same reasoning we have $M(\tilde{\Gamma}) \leq M(\Gamma)$, and so, $M(\Gamma) = M(\tilde{\Gamma})$ (see [42] or [74]). ♠

Suppose that $\mathbf{f} : \Omega \longrightarrow \Omega'$ is a diffeomorphism between domains Ω and Ω' in $\mathbb{R}^n$, where $n \geq 2$. In the proof of Theorem 3.1.5, we defined ρ with respect to $\tilde{\rho}$ that is called the **push forward**. Let $\rho : \Omega \longrightarrow [0, \infty)$ be defined by

$$\rho(x) = \tilde{\rho}\big(\mathbf{f}(x)\big)\, |D\mathbf{f}(x)|.$$

Let $\gamma : [a, b] \longrightarrow \Omega$ be a path in Γ. Then $\beta = \mathbf{f} \circ \gamma$ is a path and member of $\tilde{\Gamma}$, because it is easy to prove that

$$1 \leq \ell_{\tilde{\rho}}(\beta) = \int_{\beta} \tilde{\rho}(x)|\mathrm{dx}| = \int_a^b \tilde{\rho}[\beta(t)]|\beta'(t)|\mathrm{dt} = \int_a^b \tilde{\rho}[\beta(t)]|\mathbf{f}'[\gamma(t)]\,\gamma'(t)|\mathrm{dt}$$

$$\leq \int_a^b \tilde{\rho}[\beta(t)]|\mathbf{f}'[\gamma(t)]|\,|\gamma'(t)|\mathrm{dt} = \int_a^b \rho(\gamma(t))|\gamma'(t)|\mathrm{dt} = \int_{\gamma} \rho(x)|\mathrm{dx}| = \ell_{\rho}(\gamma).$$

3.2 Definitions of Quasiconformal Mappings

In this section, we give three main definitions for quasiconformal mappings in Euclidean space, geometric, metric, and analytic definitions. We also establish various results concerning their properties and differences.

Definition 3.2.1. Let Ω and Ω' be domains in $\mathbb{R}^n$, and $\mathbf{f} : \Omega \longrightarrow \Omega'$ be a homeomorphism. Let Γ be a curve family in Ω and its image $\tilde{\Gamma} = \big\{\mathbf{f} \circ \gamma \;:\; \gamma \in \Gamma\big\}$ is a curve family in Ω'. The **outer distortion** $K_O(\mathbf{f})$ and **inner distortion** $K_I(\mathbf{f})$ of a function $\mathbf{f}$ are defined as the following formulas

$$K_O(\mathbf{f}) = \sup_{\Gamma \in \Omega} \frac{M(\Gamma)}{M(\tilde{\Gamma})}, \qquad K_I(\mathbf{f}) = \sup_{\Gamma \in \Omega} \frac{M(\tilde{\Gamma})}{M(\Gamma)}, \tag{3.13}$$

where $M(\Gamma)$ and $M(\tilde{\Gamma})$ are not simultaneously 0 or ∞. The **maximal distortion** of $\mathbf{f}$ is defined by

$$K(\mathbf{f}) = \max\big\{K_O(\mathbf{f}), K_I(\mathbf{f})\big\}. \tag{3.14}$$

We note that $K_I \geq 1$ or $K_O \geq 1$, hence $K \geq 1$, since $M(\Gamma) = M(\tilde{\Gamma}) = \infty$, whenever one and hence the other of the families Γ and $\tilde{\Gamma}$ contains a constant path, and the constant paths have no influence on $K_I(\mathbf{f})$ and $K_O(\mathbf{f})$. In order to avoid technical difficulties, we shall assume from now on that every path family contains only non-constant paths (see [74] p.41 and 42).

Definition 3.2.2. (Geometric definition of quasiformality): Let Ω and Ω' be domains in $\mathbb{R}^n$, and $\mathbf{f} : \Omega \longrightarrow \Omega'$ be a homeomorphism. Let Γ be a curve family in Ω and its image $\tilde{\Gamma} = \big\{\mathbf{f} \circ \gamma \;:\; \gamma \in \Gamma\big\}$ is a curve family in Ω'. Then $\mathbf{f}$ is said to be a **K-quasiconformal mapping** of Ω to Ω' provided $K(\mathbf{f}) < \infty$. If, in fact, $K(\mathbf{f}) \leq K < \infty$, then we call $\mathbf{f}$ a K-quasiconformal mapping. For this to be true it is necessary and sufficient that

$$\frac{1}{K} M(\Gamma) \leq M(\tilde{\Gamma}) \leq K M(\Gamma) \tag{3.15}$$

holds for every family Γ of curves in Ω (see [30] p.206, [42] p.14 and [74]).

From the definition of the inner and outer distortion we obtain the following relations :

Theorem 3.2.1. Let Ω and Ω' be domains in $\mathbb{R}^n$, and $\mathbf{f}:\Omega\longrightarrow\Omega'$ be a homeomorphism. The following properties hold for every $x\in\Omega$:

$$\begin{array}{ll} 1)\ K_I(\mathbf{f})=K_O(\mathbf{f^{-1}}) & 2)\ K_O(\mathbf{f})=K_I(\mathbf{f^{-1}}) \\ 3)\ K(\mathbf{f^{-1}})=K(\mathbf{f}) & 4)\ K_I(\mathbf{f}\circ\mathbf{g})\le K_I(\mathbf{f})K_I(\mathbf{g}) \\ 5)\ K_O(\mathbf{f}\circ\mathbf{g})\le K_O(\mathbf{f})K_O(\mathbf{g}) & 6)\ K(\mathbf{f}\circ\mathbf{g})\le K(\mathbf{f})K(\mathbf{g}) \end{array} \tag{3.16}$$

Proof. (Proof of part 1) Suppose that Γ is a family of paths in Ω and $\tilde{\Gamma}=\mathbf{f}(\Gamma)$. By Definition 3.2.1, we have

$$\frac{M(\tilde{\Gamma})}{M(\Gamma)}=\frac{M(\tilde{\Gamma})}{M(\mathbf{f^{-1}}(\tilde{\Gamma}))}\le K_O(\mathbf{f^{-1}})$$

and by taking the supremum over all Γ we obtain $K_I(\mathbf{f})\le K_O(\mathbf{f^{-1}})$. By same process we can show $K_O(\mathbf{f^{-1}})\le K_I(\mathbf{f})$, so $K_I(\mathbf{f})=K_O(\mathbf{f^{-1}})$. Proofs of part (2) and (3) are similar to (1). (Proof of part 4) Consider $K_I(\mathbf{f}\circ\mathbf{g})$, by definition of inner distortion we have

$$\frac{M(\tilde{\Gamma})}{M(\Gamma)}=\frac{M((\mathbf{f}\circ\mathbf{g})\Gamma)}{M(\Gamma)}=\frac{M((\mathbf{f}\circ\mathbf{g})\Gamma)}{M(\mathbf{g}(\Gamma))}.\frac{M(\mathbf{g}(\Gamma))}{M(\Gamma)}\le K_I(\mathbf{f})K_I(\mathbf{g}).$$

Taking the supremum over all Γ from both sides gives us $K_I(\mathbf{f}\circ\mathbf{g})\le K_I(\mathbf{f})K_I(\mathbf{g})$. For parts (5) and (6) we have same processes (see [42] and [74]). ♠

Corollary 3.2.1.1. If $\mathbf{f}$ is a K-quasiconformal mapping, then $\mathbf{f^{-1}}$ is a K-quasiconformal mapping.

Corollary 3.2.1.2. If $\mathbf{h}=\mathbf{f}\circ\mathbf{g}$ and $\mathbf{f}$ is a K_1-quasiconformal mapping and $\mathbf{g}$ is a K_2-quasiconformal mapping, then $\mathbf{h}$ is a K_1K_2-quasiconformal mapping.

While modulus isn't easy to understand at first glance, once mastered it is a powerful tool in geometric function theory. In addition, the geometric definition (3.2.1) leads quickly to many strong properties of quasiconformal mappings; for instance, the inverse of a quasiconformal mapping is also quasiconformal, which is not apparent from the metric definition. In the Definition 1.7.1, we defined linear, inner and outer dilatations for linear transformation T. We know that the derivative of T is a linear transformation (see [30] and [42]). We can define the **linear distortion** $\mathcal{H}(T)$, **inner distortion** $\mathcal{H}_I(T)$, and the **outer distortion** $\mathcal{H}_O(T)$ of a nonsingular linear transformation $T:\mathbb{R}^n\longrightarrow\mathbb{R}^n$ as

$$\mathcal{H}(T)=\frac{L(T)}{\ell(T)},\qquad \mathcal{H}_I(T)=\frac{|\det(T)|}{\ell(T)^n},\qquad \mathcal{H}_O(T)=\frac{L(T)^n}{|\det(T)|}. \tag{3.17}$$

If T is singular it is customary to set $\mathcal{H}(T)=\mathcal{H}_I(T)=\mathcal{H}_O(T)=\infty$ (see [30]). Let the transformation T is presented by a $n\times n$ matrix A, and let $0<\mu_1\le\mu_2\le\ldots\le\mu_n$ be singular values of the matrix A (this means they are the eigenvalues of matrix A^tA and real positive numbers). Recall that $L(T)=L(A)=\mu_n$, $\ell(T)=\ell(A)=\mu_1$, and $|\det(T)|=|\det(A)|=\mu_1.\mu_2.\ldots.\mu_n$. Hence, by (3.17) we can write

$$\mathcal{H}(T)=\frac{\mu_n}{\mu_1},\qquad \mathcal{H}_I(T)=\frac{\mu_2.\mu_3.\ldots.\mu_n}{\mu_1^{n-1}},\qquad \mathcal{H}_O(T)=\frac{\mu_n^{n-1}}{\mu_1.\mu_2.\ldots.\mu_{n-1}}. \tag{3.18}$$

Theorem 3.2.2. Let Ω and Ω' be domains in $\mathbb{R}^n$, and $A : \Omega \longrightarrow \Omega'$ be a linear bijection. Then

1. $\mathcal{H}_I(A) \leq \mathcal{H}_O(A)^{n-1}$,
2. $\mathcal{H}_O(A) \leq \mathcal{H}_I(A)^{n-1}$,
3. $\mathcal{H}(A)^n = \mathcal{H}_I(A).\mathcal{H}_O(A)$,
4. $\mathcal{H}(A) \leq \min\{\mathcal{H}_I(A), \mathcal{H}_O(A)\} \leq \mathcal{H}(A)^{\frac{n}{2}} \leq \max\{\mathcal{H}_I(A), \mathcal{H}_O(A)\} \leq \mathcal{H}(A)^{n-1}$.

Proof. (Proof of part 3) Let $0 < \mu_1 \leq \mu_2 \leq \ldots \leq \mu_n$ be singular values of the matrix A. Thus

$$\begin{aligned}\mathcal{H}(A)^n &= \left(\frac{\mu_n}{\mu_1}\right)^n = \frac{\mu_n^n}{\mu_1^n} \\ &= \frac{\mu_n^n}{\mu_1^n}.\left(\frac{\mu_2.\mu_3.\cdots.\mu_{n-1}}{\mu_2.\mu_3.\cdots.\mu_{n-1}}\right) \\ &= \frac{\mu_n^{n-1}}{\mu_1\mu_2.\cdots.\mu_{n-1}}.\frac{\mu_2.\mu_3.\cdots.\mu_{n-1}\mu_n}{\mu_1^{n-1}} \\ &= \mathcal{H}_O(A).\mathcal{H}_I(A).\end{aligned}$$

This finishes the proof of part (3). The proofs of the other parts are the similar. ♠

It is obvious that all three linear, outer, and inner distortions are bigger than or equal to 1. They have the following geometric intuition. The image of the unit ball $B(0,1)$ under A is an ellipsoid $E(A)$. Let $B_I(A)$ and $B_O(A)$ be the inscribed and circumscribed balls of $E(A)$, respectively (see [42] and [74]). See Figure 3.4.

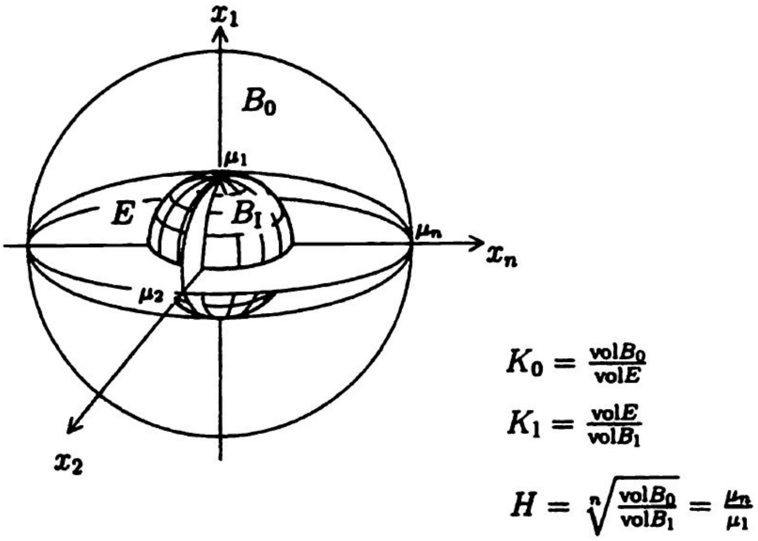

Figure 3.4: Eccentricity of the ellipsoid
Source: Page 110 of book [42]

We recall Definition 1.7.2 that defined the linear distortion of a function at a point or its domain. If A is a linear transformation given by its matrix. If $n = 2$, computing distortion can be done by a relatively simple formula (see [74]).

Example 3.2.1. Let $A : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ be a linear bijection given by (see [74], pp 44 and 45)

$$A(z) = ax + by + i(cx + dy),$$

where z is the complex notation in $\mathbb{R}^2$ with $z = x + iy$ and a, b, c, and d are real numbers such that $ad - bc \neq 0$. Since $n = 2$, then $\mathcal{H}(A) = \mathcal{H}_I(A) = \mathcal{H}_O(A)$. Let

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

We know that the singular values of matrix A is the square roots of the eigenvalues of matrix A^tA, thus

$$A^tA = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \cdot \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a^2 + c^2 & ab + cd \\ ab + cd & b^2 + d^2 \end{bmatrix}.$$

The charactristic equation of A^tA is

$$\det(A^tA - \lambda I) = \lambda^2 - (a^2 + b^2 + c^2 + d^2)\lambda + (ad - bc)^2 = 0,$$

where I is the 2×2 identity matrix. By solving the characteristic equation we get

$$\lambda_1 = \frac{1}{2}\Big(a^2 + b^2 + c^2 + d^2 - \sqrt{(a^2 + b^2 + c^2 + d^2)^2 - 4(ad - bc)^2}\Big),$$

$$\lambda_2 = \frac{1}{2}\Big(a^2 + b^2 + c^2 + d^2 + \sqrt{(a^2 + b^2 + c^2 + d^2)^2 - 4(ad - bc)^2}\Big).$$

So, the singular values of matrix A are

$$\mu_1 = \sqrt{\lambda_1} = \frac{\sqrt{a^2 + b^2 + c^2 + d^2 - \sqrt{(a^2 + b^2 + c^2 + d^2)^2 - 4(ad - bc)^2}}}{\sqrt{2}},$$

and

$$\mu_2 = \sqrt{\lambda_2} = \frac{\sqrt{a^2 + b^2 + c^2 + d^2 + \sqrt{(a^2 + b^2 + c^2 + d^2)^2 - 4(ad - bc)^2}}}{\sqrt{2}}.$$

It is easy to see that $\mu_1 \leq \mu_2$, hence the linear distortion of matrix A is

$$\mathcal{H}(A) = \mathcal{H}_I(A) = \mathcal{H}_O(A) = \frac{\sqrt{a^2 + b^2 + c^2 + d^2 + \sqrt{(a^2 + b^2 + c^2 + d^2)^2 - 4(ad - bc)^2}}}{\sqrt{a^2 + b^2 + c^2 + d^2 - \sqrt{(a^2 + b^2 + c^2 + d^2)^2 - 4(ad - bc)^2}}}.$$

♣

We now define the linear distortion of a function at a point x.

Definition 3.2.3. Let Ω be a domain in $\mathbb{R}^n$, let $\mathbf{f} : \Omega \longrightarrow \mathbb{R}^n$ be a continuous injection. The **linear distortion or dilatation** of $\mathbf{f}$ at the point $\mathbf{x}$ in A is the quantity $\mathcal{H}_{\mathbf{f}}(\mathbf{x})$ or $\mathcal{H}(\mathbf{x}, \mathbf{f})$ defined by

$$\mathcal{H}_{\mathbf{f}}(\mathbf{x}) = \mathcal{H}(\mathbf{x}, \mathbf{f}) = \limsup_{r \to 0} \frac{L_f(\mathbf{x}, r)}{\ell_f(\mathbf{x}, r)}, \tag{3.19}$$

where for $0 < r < \operatorname{dist}(\mathbf{x}, \partial\Omega)$ we set

$$L_f(\mathbf{x}, r) = \max_{\|h\|=r} \|\mathbf{f}(\mathbf{x}+h) - \mathbf{f}(\mathbf{x})\|, \qquad \ell_f(\mathbf{x}, r) = \min_{\|h\|=r} \|\mathbf{f}(\mathbf{x}+h) - \mathbf{f}(\mathbf{x})\|.$$

In other words,

$$\mathcal{H}_{\mathbf{f}}(\mathbf{x}) = \mathcal{H}(\mathbf{x}, \mathbf{f}) = \limsup_{r \to 0} \frac{\max_{\|h\|=r} \|\mathbf{f}(\mathbf{x}+h) - \mathbf{f}(\mathbf{x})\|}{\min_{\|h\|=r} \|\mathbf{f}(\mathbf{x}+h) - \mathbf{f}(\mathbf{x})\|}. \tag{3.20}$$

Let Ω and Ω' be domains in $\mathbb{R}^n$ with $n \geq 2$, and let $\mathbf{f}$ be a homeomorphism of Ω onto Ω'. The **linear distortion of function f** is defined

$$\mathcal{H}(\mathbf{f}) = \sup\{\mathcal{H}_{\mathbf{f}}(\mathbf{x}) : \ \mathbf{x} \in \Omega\}, \tag{3.21}$$

where $\mathcal{H}_{\mathbf{f}}(\mathbf{x}) \geq 1$ is the linear distortion of $\mathbf{f}$ at $\mathbf{x}$ as defined at (1.29). $\mathbf{f}$ is a conformal mapping if and only if $\mathcal{H}(\mathbf{f}) = 1$ (see [30] p.77).

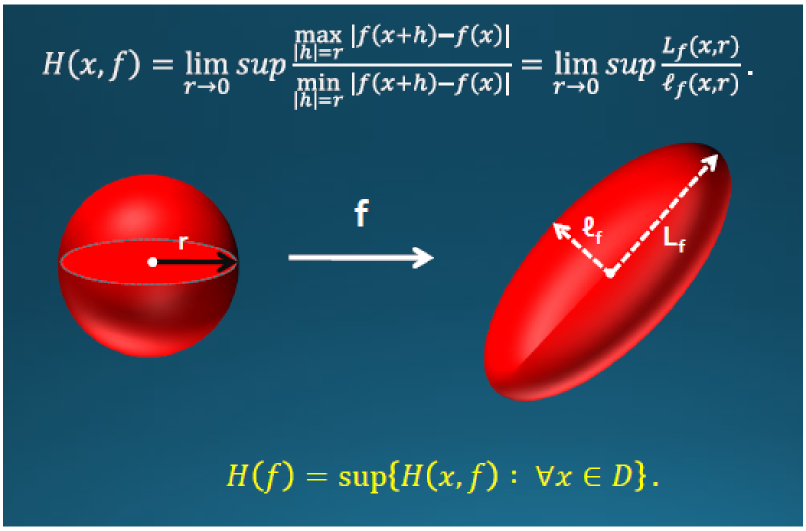

Figure 3.5: Linear distortion of the funcion $\mathbf{f}$ at a piont in the domain.

Let Ω and Ω' be domains in $\mathbb{R}^n$, and $\mathbf{f} : \Omega \longrightarrow \Omega'$ be a diffeomorphism. This means a diffeomorphism is a C^1-homeomorphism whose Jacobian $J_{\mathbf{f}}(\mathbf{x})$ does not vanish. If the function $\mathbf{f}$ is a diffeomorphism, then by (3.17), Theorem 1.7.1 and Corollary 1.7.1.1, we have $L_{\mathbf{f}}(\mathbf{x}) = |D\mathbf{f}(\mathbf{x})|$ and $\det(D\mathbf{f}(\mathbf{x})) = J_{\mathbf{f}}(\mathbf{x})$, so

$$\mathcal{H}_O(D\mathbf{f}(\mathbf{x})) = \frac{|D\mathbf{f}(\mathbf{x})|^n}{|J_{\mathbf{f}}(\mathbf{x})|}, \qquad \mathcal{H}_I(D\mathbf{f}(\mathbf{x})) = \frac{|J_{\mathbf{f}}(\mathbf{x})|}{\ell(D\mathbf{f}(\mathbf{x}))^n}. \tag{3.22}$$

We now introduce some theorems that lead us to other definitions of quasiconformal mappings and the relations between them (see [42] and [74]).

Theorem 3.2.3. Let Ω and Ω' be domains in $\mathbb{R}^n$, with $n \geq 2$, and $\mathbf{f} : \Omega \longrightarrow \Omega'$ be a homeomorphism with $\mathbf{f}$ differentiable at the point $\mathbf{x} \in \Omega$. If $K_O(\mathbf{f}) < \infty$, then

$$|D\mathbf{f}(x)|^n \leq K_O(\mathbf{f})|J_{\mathbf{f}}(\mathbf{x})|. \tag{3.23}$$

Proof. The reader can see [74] pages 47 and 48. ♠

Theorem 3.2.4. Suppose $\mathbf{f} : \Omega \longrightarrow \Omega'$ is a diffeomorphism, then

$$K_O(\mathbf{f}) = \sup_{\mathbf{x}\in\Omega} \mathcal{H}_O[D\mathbf{f}(\mathbf{x})], \qquad K_I(\mathbf{f}) = \sup_{\mathbf{x}\in\Omega} \mathcal{H}_I[D\mathbf{f}(\mathbf{x})]. \tag{3.24}$$

Proof. It suffices to prove the first equation, because the second one follows it by applying the inverse mapping $\mathbf{f}$. Suppose that $H_O[D\mathbf{f}(\mathbf{x})] = K < \infty$. We have to show

$$K_O(\mathbf{f}) = \frac{M(\Gamma)}{M(\tilde{\Gamma})} \leq K,$$

for an arbitrary path family Γ in Ω. The proof is similar to the proof of the Theorem 3.1.5 for the conformal invariance. In fact the Theorem 3.1.5 is a special case of this theorem. Let $\rho : \mathbb{R}^n \longrightarrow \hat{\mathbb{R}}^n$ be a function that

$$\rho(\mathbf{x}) = \begin{cases} \rho'(\mathbf{f}(\mathbf{x}))|D\mathbf{f}(\mathbf{x})| & \text{if } \mathbf{x} \in \Omega, \\ 0 & \text{if } \mathbf{x} \notin \Omega. \end{cases}$$

Suppose that γ is a locally rectifiable path in Γ. By Definition 3.1.4 we obtain

$$\int_\gamma \rho \,\mathrm{ds} \geq \int_{\mathbf{f}\circ\gamma} \rho' \,\mathrm{ds} \geq 1.$$

Therefore, $\rho \in \mathrm{Adm}(\Gamma)$, and $\rho' \in \mathrm{Adm}(\tilde{\Gamma})$, which implies

$$\begin{aligned} M(\Gamma) &\leq \int_\Omega \rho^n \mathrm{dm} = \int_\Omega \rho'(\mathbf{f}(\mathbf{x}))^n |D\mathbf{f}(\mathbf{x})|^n \,\mathrm{dm} \\ &\leq K \int_\Omega \rho'(\mathbf{f}(\mathbf{x}))^n |J_{\mathbf{f}}(\mathbf{x})| \,\mathrm{dm} = K \int_{\mathbf{f}(\Omega)} (\rho')^n \,\mathrm{dm} \\ &\leq K \int_{\mathbb{R}^n} (\rho')^n \,\mathrm{dm}. \end{aligned}$$

Since this holds for every $\rho' \in \mathrm{Adm}(\tilde{\Gamma})$, we get $M(\Gamma) \leq K\, M(\tilde{\Gamma})$. By Theorem 3.2.4 we can prove the reverse inequality. So

$$K_O(\mathbf{f}) = \sup_{\mathbf{x}\in\Omega} \mathcal{H}_O[D\mathbf{f}(\mathbf{x})].$$

By $\mathbf{f}^{-1} : \Omega' \longrightarrow \Omega$ and $K_I(\mathbf{f}) = K_O(\mathbf{f}^{-1})$, we can prove the second equation, thus

$$K_I(\mathbf{f}) = \sup_{\mathbf{x}\in\Omega} \mathcal{H}_I[D\mathbf{f}(\mathbf{x})].$$

♠

Corollary 3.2.4.1. Let $\mathbf{f} : \Omega \longrightarrow \Omega'$ be a diffeomorphism, then $\mathbf{f}$ is K-quasiconformal if and only if

$$\frac{|D\mathbf{f}(\mathbf{x})|^n}{K} \le |J_{\mathbf{f}}(\mathbf{x})| \le K\, \ell(D\mathbf{f}(\mathbf{x}))^n,$$

for all $x \in \Omega$. Moreover,

$$1 \le K_O(\mathbf{f}) \le K_I(\mathbf{f})^{n-1}, \qquad 1 \le K_I(\mathbf{f}) \le K_O(\mathbf{f})^{n-1}.$$

Corollary 3.2.4.2. If $\mathbf{f}$ is a differentiable quasiconformal mapping at point $\mathbf{x}$ in Ω, then either $D\mathbf{f}(\mathbf{x}) = 0$ or $J_{\mathbf{f}}(\mathbf{x}) \neq 0$ (see [74], page 48).

Example 3.2.2. Let $A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ be a linear bijection. Then $D(A(\mathbf{x})) = A$ for all $\mathbf{x} \in \mathbb{R}^n$. By Theorem 3.2.3 we obtain

$$K_O(A) = \mathcal{H}_O(A), \qquad K_I(A) = \mathcal{H}_I(A).$$

Therefore A is a quasiconformal mapping.

♣

The definition of linear distortion tells us $1 \le \mathcal{H}_{\mathbf{f}}(\mathbf{x}) \le \infty$. If $A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a bijective linear mapping, $\mathcal{H}(\mathbf{x}, A) = \mathcal{H}(A)$, for all $\mathbf{x} \in \mathbb{R}^n$. We omit the very technical proof of the following theorem. It can be found in [74] pages 78 and 79.

Theorem 3.2.5. Assume $\mathbf{f} : \Omega \longrightarrow \Omega'$ is a homeomorphism such that for some $K < \infty$ either $K_O(\mathbf{f}) \le K$ or $K_I(\mathbf{f}) \le K$. Then $\mathcal{H}(\mathbf{x}, \mathbf{f})$ is bounded by a constant which depends only on n and K.

Corollary 3.2.5.1. If $\mathbf{f} : \Omega \longrightarrow \Omega'$ is a quasiconformal mapping, then $\mathcal{H}(x, \mathbf{f})$ is bounded.

Definition 3.2.4. A function $f : [a, b] \longrightarrow \mathbb{R}^m$ is said to be **absolutely continuous** if for each $\epsilon > 0$ there exists a $\delta > 0$ such that

$$\sum_{i=1}^{k} |f(b_i) - f(a_i)| < \epsilon$$

whenever $a = a_1 < b_1 \le a_2 < b_2 \le \ldots \le a_k < b_k = b$ and

$$\sum_{i=1}^{k} |b_i - a_i| < \delta.$$

It is easy to see that every absolutely continuous function is uniformly continuous and, therefore continuous. Also, we can prove that every Lipschitz-continuous function is absolutely continuous (see [42] and [64]).

Definition 3.2.5. For $j = 1, 2, \ldots, n$, let $I_j = [a_j, b_j]$ be a closed interval and $Q = I_1 \times I_2 \times \ldots \times I_n$ be a rectangular box that lies in a domain Ω in $\mathbb{R}^n$. Consider the jth face

$$Q_j = I_1 \times \ldots I_{j-1} \times \{a_j\} \times I_{j+1} \times \ldots I_n.$$

Then we can say a continuous function $\mathbf{f} : \Omega \longrightarrow \mathbb{R}^m$ is $ACL(\Omega, \mathbb{R}^m)$, **absolutely continuous on lines** if for every $j = 1, 2, \ldots, n$ and almost every $a \in Q_j$, with respect to $(n-1)$-measure, the function

$$t \mapsto \mathbf{f}(a + te_j), \qquad a_j \leq t + \langle a, e_j \rangle \leq b_j,$$

is absolutely continuous. e_j denotes the jth unit basis vector (see [42] p.106). For $p \geq 1$, an ACL mapping $\mathbf{f} : \Omega \longrightarrow \mathbb{R}^m$ is said to be ACL^p if all partial derivatives of $\mathbf{f}$ are in $L^p_{loc}(\Omega)$. In other words,

$$\frac{\partial \mathbf{f}}{\partial x_i} \in L^p_{loc}(\Omega), \qquad 1 \leq i \leq n.$$

Sometimes, ACL^p is known as $W^{1,p}$ (see [77], Chapter 2). There is a fundamental result of Gehring and Väiälä, its proof exists in Väiälä's book [74]. This result links geometric and analytic concepts in the theory of quasiconformal mappings.

Theorem 3.2.6. Let $\mathbf{f} : \Omega \longrightarrow \mathbb{R}^n$ be a homeomorphism such that $\mathcal{H}(\mathbf{x}, \mathbf{f})$ is bounded, for all $\mathbf{x} \in \Omega$. Then $\mathbf{f}$ is $ACL(\Omega, \mathbb{R}^n)$.

Theorem 3.2.7. (**Gehring-Väiälä**) Every quasiconformal mapping $\mathbf{f} : \Omega \longrightarrow \mathbb{R}^n$ lies in the space $ACL(\Omega, \mathbb{R}^n)$.

There are others important results in Väiälä's book [74] that we now explain.

Theorem 3.2.8. Suppose that Ω and Ω' are domains in $\mathbb{R}^n$,with $n \geq 2$, and $\mathbf{f} : \Omega \longrightarrow \Omega'$ be a homeomorphism such that $\mathcal{H}(\mathbf{f})$ is bounded. Then $\mathbf{f}$ is differentiable almost everywhere.

Proof. The proof of this theorem can be found in [74], page 109. ♠

Corollary 3.2.8.1. Every quasiconformal mapping is differentiable almost everywhere.

Theorem 3.2.9. If Ω and Ω' are domains in $\mathbb{R}^n$, with $n \geq 2$, and $\mathbf{f} : \Omega \longrightarrow \Omega'$ is a homeomorphism and if $1 \leq K \leq \infty$, then the following are equivalent:

1. $K_O(\mathbf{f}) \leq K$
2. $\mathbf{f}$ is ACL, almost everywhere differentiable, and $|\mathbf{f}'(\mathbf{x})|^n \leq K|J_{\mathbf{f}}(\mathbf{x})|$ almost everywhere.

Moreover, each of these conditions implies that $\mathbf{f} \in ACL^n$ (see [74] p.110 and 111).

Corollary 3.2.9.1. Every quasiconformal map is ACL^n.

Lemma 3.2.10. If $\mathbf{f} : \Omega \longrightarrow \Omega'$ is a homeomorphism, such that $\mathcal{H}(\mathbf{x}, \mathbf{f})$ is bounded by a constant $\mathcal{H}$, for all $\mathbf{x} \in \Omega$. Then (see [74])

$$|\mathbf{f}'(\mathbf{x})|^n \leq \mathcal{H}^{n-1}|J_{\mathbf{f}}(\mathbf{x})|.$$

Proof. By Theorem 3.2.6, the homeomorphism $\mathbf{f}$ is ACL, and by Theorem 3.2.8, $\mathbf{f}$ is differentiable almost everywhere. Also, by Corollary 3.2.4.2, and Theorem 1.7.1, we can write that for $x \in \Omega$ either $\mathbf{f}'(\mathbf{x}) = 0$ or

$$0 < \mathcal{H}(\mathbf{f}'(\mathbf{x})) = \mathcal{H}(\mathbf{x}, \mathbf{f}) = \mathcal{H}.$$

If $\mathbf{f}'(\mathbf{x}) = 0$, then the result is obvious, therefore we choose $0 < \mathcal{H}(\mathbf{f}'(\mathbf{x})) = \mathcal{H}(\mathbf{x}, \mathbf{f}) = \mathcal{H}$. Accordingly,

$$\frac{|\mathbf{f}'(\mathbf{x})|^n}{|J_{\mathbf{f}}(\mathbf{x})|} = \mathcal{H}_O(\mathbf{f}'(\mathbf{x})) \leq \max\{\mathcal{H}_I(\mathbf{f}'), \mathcal{H}_O(\mathbf{f}')\} \leq \mathcal{H}(\mathbf{f}')^{n-1} = \mathcal{H}^{n-1},$$

or

$$|\mathbf{f}'(\mathbf{x})|^n \leq \mathcal{H}^{n-1}|J_{\mathbf{f}}(\mathbf{x})|.$$

♠

Theorem 3.2.11. Let Ω and Ω' be domains in $\mathbb{R}^n$,with $n \geq 2$. Then the homeomorphism $\mathbf{f} : \Omega \longrightarrow \Omega'$ is a quasiconformal mapping if and only if $\mathcal{H}(\mathbf{f})$ is bounded.

Proof. Let $\mathbf{f}$ be a quasiconformal mapping. Then by Corollary 3.2.5.1, $\mathcal{H}(\mathbf{x}, \mathbf{f})$ is bounded, for all $\mathbf{x} \in \Omega$. Conversely, suppose that $\mathcal{H}(\mathbf{x}, \mathbf{f}) \leq \mathcal{H} < \infty$, for all $\mathbf{x} \in \Omega$. By Theorem 3.2.6, the homeomorphism $\mathbf{f}$ is ACL, and by Theorem 3.2.8, $\mathbf{f}$ is differentiable almost everywhere. By Corollary 3.2.4.2 and Theorem 1.7.1, for all $\mathbf{x} \in \Omega$, we have $\mathbf{f}'(\mathbf{x}) = 0$ or $0 < \mathcal{H}(\mathbf{f}'(\mathbf{x})) = \mathcal{H}(\mathbf{x}, \mathbf{f}) \leq \mathcal{H}$. In both cases, $|\mathbf{f}'(\mathbf{x})|^n \leq \mathcal{H}^{n-1}|J_{\mathbf{f}}(x)|$. Lemma 3.2.10 tells us

$$K_O(\mathbf{f}) \leq \mathcal{H}^{n-1},$$

for $\mathcal{H}^{n-1} = K$. Since $K_O(\mathbf{f}) = K_I(\mathbf{f}^{-1})$, Theorem 3.2.5 leads to $\mathcal{H}(\mathbf{y}, \mathbf{f}^{-1})$ is bounded, for all $\mathbf{y} \in \Omega'$. By repeating the above process we get $K_I(\mathbf{f}) = K_O(\mathbf{f}^{-1}) < \infty$. Hence $\mathbf{f}$ is a quasiconformal mapping (see [42], [30] and [74]). ♠

Corollary 3.2.11.1. A homeomorphism $\mathbf{f} : \Omega \longrightarrow \Omega'$ is quasiconformal if and only if one of the distortions $K_O(\mathbf{f})$ or $K_I(\mathbf{f})$ is finite.

Proof. Let $\mathbf{f}$ be quasiconformal, then by Theorem 3.2.11 , $\mathcal{H}(\mathbf{x}, \mathbf{f})$ is bounded and at least one of the inner or outer distortion is finite. Conversely, assume $K_O(\mathbf{f})$ or $K_I(\mathbf{f})$ is finite. By Theorem 3.2.5, we can say, $\mathcal{H}(\mathbf{x}, \mathbf{f})$ is bounded and also by Theorem 3.2.11, $\mathbf{f}$ is a quasiconformal map (see [74] p.114). ♠

Now, by Theorem 3.2.11, we have another definition for quasiconformal mappings that is called the metric definition of quasiconformal mappings. For more details, you can see [42] or [30].

Definition 3.2.6. (Metric definition of quasiconformality) Let Ω and Ω' be domains in $\mathbb{R}^n$, with $n \geq 2$, and $\mathbf{f} : \Omega \longrightarrow \Omega'$ be a homeomorphism. Then $\mathbf{f}$ is called **$\mathcal{H}$-quasiconformal**, if there is a constant $1 \leq \mathcal{H} \leq \infty$ such that $\mathcal{H}(\mathbf{f}) \leq \mathcal{H}$.

This distortion function is a quantity that we shall meet repeatedly. If the function $\mathbf{f}$ is conformal, then $\mathcal{H}(\mathbf{x}, \mathbf{f}) = 1$. Indeed, the converse is also true. This metric definition is simple, direct, very general, and quite appealing from a geometric-aesthetic vantage point. Further, it requires no regularity of differentiability properties of the mapping to formulate. Unfortunately, this metric definition, while aesthetically pleasing, is difficult to work with. Nowadays, the following analytic definition of quasiconformality is more common (see [42] and [30]). The following theorems and definitions show the relationship between quasiconformality and bounded distortion for diffeomorphism. But before the analytic definition, we give some definitions that we need. At first, we present an important theorem.

Theorem 3.2.12. Let Ω and Ω' be domains in $\mathbb{R}^n$, with $n \geq 2$, and $\mathbf{f} : \Omega \longrightarrow \Omega'$ be a homeomorphism. Then $\mathbf{f}$ is K-quasiconformal if and only if the following conditions are satisfied:

1. $\mathbf{f}$ is ACL,
2. $\mathbf{f}$ is differentiable almost everywhere,
3. For almost every $x \in \Omega$ we have

$$\frac{|\mathbf{f}'(\mathbf{x})|^n}{K} \leq |J_{\mathbf{f}}(\mathbf{x})| \leq K\ \ell(\mathbf{f}'(\mathbf{x}))^n. \tag{3.25}$$

Proof. Let $\mathbf{f}$ be a K-quasiconformal, then by Theorem 3.2.7 (Gehring-Väiälä) $\mathbf{f}$ is ACL, and by Corollary 3.2.8.1 the homeomorphism $\mathbf{f}$ is differentiable almost everywhere. So, properties (1) and (2) are true. Since $\mathbf{f}$ is K-quasiconformal, we can write

$$\frac{|\mathbf{f}'(\mathbf{x})|^n}{|J_{\mathbf{f}}(\mathbf{x})|} \leq K$$

for all $\mathbf{x} \in \Omega$ that $\mathbf{f}$ is differentiable at $\mathbf{x}$. We may also choose $|J_{\mathbf{f}}(\mathbf{x})| \neq 0$. We know that the $\mathbf{f}^{-1}$ is also K-quasiconformal and differentiable at $\mathbf{y} = \mathbf{f}(\mathbf{x})$. Hence,

$$|J_{\mathbf{f}}(\mathbf{x})| = |J_{\mathbf{f}^{-1}}(\mathbf{y})| \leq K|\mathbf{f}'^{-1}(\mathbf{y})| = K\ \ell(\mathbf{f}'(\mathbf{x}))^n.$$

The above relation shows us the condition (3) is also true. Conversely, assume that the three conditions hold. That means $\mathbf{f}$ is ACL, differentiable almost everywhere, and for almost $\mathbf{x} \in \Omega$ the following inequality holds

$$\frac{|\mathbf{f}'(\mathbf{x})|^n}{K} \leq |J_{\mathbf{f}}(\mathbf{x})| \leq K\ \ell(\mathbf{f}'(\mathbf{x}))^n.$$

Note that the above inequality gives us $K_O(\mathbf{f}) \leq K$ and this inequality by Theorem 3.2.9 implies $K_O(\mathbf{f})$ is finite. Therefore, by Corollary 3.2.11.1, $\mathbf{f}$ is a quasiconformal map. Since $|J_{\mathbf{f}}(\mathbf{x})| \neq 0$ almost everywhere, we have $K_I(\mathbf{f}) \leq K$ (see [30], p 243 and [74], p 115). ♠

This theorem leads us to define a quasiconformal mapping in a third way. We had geometric and metric definitions before, but the next one is called the analytic definition of quasiconformal mapping.

Definition 3.2.7. (Analytic definition of quasiconformality) Let Ω and Ω' be domains in $\mathbb{R}^n$, with $n \geq 2$, and $\mathbf{f} : \Omega \longrightarrow \Omega'$ be a homeomorphism. Then $\mathbf{f}$ is called **K-quasiconformal** if the first partial derivatives of $\mathbf{f}$ exist as Lebesgue functions, the Jacobian determinant of $\mathbf{f}$ is locally integrable and does not change sign in Ω and there is a constant $K_O \geq 1$, such that $\mathbf{f}$ satisfies the distortion inequality

$$|\mathbf{f}'(x)|^n \leq K_O\,|J_{\mathbf{f}}(x)|, \tag{3.26}$$

for almost everywhere in Ω (see [42] p.106). In a simple way, we can say $\mathbf{f}$ is quasiconformal if the following conditions are satisfied:

1. **f** is ACL,
2. **f** is differentiable almost everywhere,
3. For almost every $x \in \Omega$, Equation 3.26 holds.

Now we can explain a theorem that shows the equivalence of these three definitions of quasiconformal mappings.

Theorem 3.2.13. Let Ω and Ω' be domains in $\mathbb{R}^n$, with $n \geq 2$, and $\mathbf{f} : \Omega \longrightarrow \Omega'$ be a homeomorphism. For all Γ in Ω the following are equivalent:

1. $\frac{1}{K} M(\Gamma) \leq M(\tilde{\Gamma}) \leq K M(\Gamma)$.
2. $\mathcal{H}(x, \mathbf{f})$ is bounded.
3. **f** is ACL, **f** is differentiable almost everywhere and for almost every $x \in \Omega$ we have
$$|\mathbf{f}'(x)|^n \leq K\, |J_{\mathbf{f}}(x)|.$$

Proof. To get from (1) to (2) we can use Theorem 3.2.5 and Corollary 3.2.5.1. Also, we can prove by Theorem 3.2.12, the formula (1) implies (3). By Lemma (3.2.10), Theorem 3.2.11 and Theorem 3.2.12 we can show that (2) and (3) are equivalent. But the proofs that (2) or (3) implies (1) are technical and we omit those. For the complete proof, you can see [74] pages 121 to 123. ♠

We should note that in 2 dimensions $\mathcal{H} = K$. This means if $n = 2$ then the constant $\mathcal{H}$ and the constant K are equal. But if $n \geq 3$, $\mathcal{H}$ and K are different constants in the metric definition and analytic or geometric definitions. Because if $0 < \mu_1 \leq \mu_2 \leq \ldots \leq \mu_n$, then in the analytic definition we have

$$|\mathbf{f}'(\mathbf{x})|^n \leq K |J_{\mathbf{f}}(\mathbf{x})|.$$

Therefore,

$$(\mu_n)^n \leq K \mu_1 \mu_2 \cdots \mu_n.$$

On the other hand, we know that

$$\frac{\mu_n}{\mu_1} \leq \mathcal{H}.$$

So, it is obvious to see that if $n = 2$, the constants $\mathcal{H}$ and K are equal. Now, we can extend these functions for linear mappings. Suppose $A \in M_{n \times n}(\mathbb{R})$ is a nonsingular matrix, and $0 < \mu_1 \leq \mu_2 \leq \ldots \leq \mu_n$ are square roots of eigenvalues of the matrix $A^t A$ (this means that they are singular values of matrix A). Here and in what follows, all distortion functions of a singular matrix are assumed to be equal to ∞, except for the zero matrix where by convention they are all equal to 1. In two dimensions all these distortion functions coincide. This is not the case for $n \geq 3$. However, these distortion functions are coupled by the inequalities (see [42]).

$$\begin{array}{llll} \mathcal{H} \leq K_O & K \leq K_O^{n-1} & K_O \leq K_I^{n-1} & K_I \leq K_O^{n-1} \\ \mathcal{H} \leq K_I & K \leq K_I^{n-1} & K_O \leq \mathcal{H}^{n-1} & K_I \leq \mathcal{H}^{n-1} \\ \mathcal{H} \leq K^{\frac{2}{n}} & K \leq H^{n-1} & K_O \leq K & K_I \leq K \end{array} \tag{3.27}$$

Proof. Recall that $L(A) = \mu_n$, $\ell(A) = \mu_1$ and $|\det(A)| = \mu_1.\mu_2.\ldots.\mu_n$. We want to prove

$$K_O(A) \leq K_I(A)^{n-1}.$$

Let $0 < \mu_1 \leq \mu_2 \leq \ldots \leq \mu_n$ be square roots of eigenvalues of the matrix A^tA (this means that they are singular values of matrix A). So,

$$\Leftrightarrow \frac{\mu_n^{n-1}}{\mu_1.\mu_2\cdots\mu_{n-1}} \leq \left(\frac{\mu_2\cdots\mu_n}{\mu_1^{n-1}}\right)^{n-1},$$

$$\Leftrightarrow \frac{\mu_n^{n-1}}{\mu_1.\mu_2\cdots\mu_{n-1}} \leq \left(\frac{\mu_2^{n-1}\cdots\mu_n^{n-1}}{(\mu_1^{n-1})^{n-1}}\right),$$

$$\Leftrightarrow \frac{1}{\mu_1.\mu_2\cdots\mu_{n-1}} \leq \left(\frac{\mu_2^{n-1}\cdots\mu_{n-1}^{n-1}}{(\mu_1^{n-1})^{n-1}}\right),$$

$$\Leftrightarrow \frac{\mu_1^{n-1}}{\mu_1.\mu_2\cdots\mu_{n-1}} \leq \left(\frac{\mu_2^{n-1}\cdots\mu_{n-1}^{n-1}}{(\mu_1^{n-1})^{n-2}}\right),$$

$$\Leftrightarrow \frac{\mu_1}{\mu_2}.\frac{\mu_1}{\mu_3}\cdots\frac{\mu_1}{\mu_{n-1}} \leq \left(\frac{\mu_2}{\mu_1}\right)^{n-1}\left(\frac{\mu_3}{\mu_1}\right)^{n-1}\cdots\left(\frac{\mu_{n-1}}{\mu_1}\right)^{n-1}.$$

This is obviously true because the left-hand side is less than or equal to 1 and the right-hand side is bigger than or equal to 1. Thus the inequality is proved. ♠

In fact, there are more distortion functions which are most readily defined in terms of singular values of A. We call them the **sectional distortions of A** and they are defined for $1 \leq \alpha,\ \beta \leq n-1$ as

$$\mathcal{K}_{\alpha,\beta}(A) = \frac{(\mu_n\mu_{n-1}\cdots\mu_{n-\alpha+1})^\beta}{(\mu_1\mu_2\cdots\mu_\beta)^\alpha} \tag{3.28}$$

There is no repetition here. We can say $\mathcal{K}_{\alpha,\beta} = \mathcal{K}_{\alpha',\beta'}$ if and only if $\alpha = \alpha'$ and $\beta = \beta'$ (see [42]). Outer, inner and linear distortions can be described as particular cases of the $\mathcal{K}_{\alpha,\beta}$:

1. $\mathcal{K}_{1,n-1}(A) = (\mu_1\mu_2\cdots\mu_{n-1})^{-1}(\mu_n)^n = K_O(A)$,
2. $\mathcal{K}_{n-1,1}(A) = (\mu_1)^{1-n}(\mu_2\mu_3\cdots\mu_n) = K_I(A)$,
3. $\mathcal{K}_{1,1}(A) = \mu_n/\mu_1 = H(A)$.

Also note the following symmetry relation:

$$\mathcal{K}_{\alpha,\beta}(A^{-1}) = \mathcal{K}_{\alpha,\beta}(A),$$

when $\det(A) \neq 0$ (see [42] p.110). Also we can write that

$$\mathcal{K}_{\alpha,\beta}(AB) \leq \mathcal{K}_{\alpha,\beta}(A)\mathcal{K}_{\alpha,\beta}(B).$$

The distortion functions $\mathcal{K}_{\alpha,\beta}$ with $\alpha+\beta = n$ have the very nice property of being polyconvex. This implies lower semicontinuity on the space ACL^n. In even dimensions, the distortion functions are expressed by $\mathcal{K} = (\mathcal{K}_{\ell,\ell})^{1/\ell}$, when $\ell = n/2$. This definition has all desired features mentioned above, which the inner, outer and linear distortions do not. So, for this reason we denote

$$\mathcal{K}_\ell(A) = \mathcal{K}_{\ell,n-\ell}(A) = \frac{(\mu_n\mu_{n-1}\cdots\mu_{n-\ell+1})^n}{(\det(A))^\ell},$$

for $\ell = 1, 2, \ldots, n-1$, and in even dimensions we set

$$\mathcal{K}(A) = \frac{\mu_n\cdots\mu_{\ell+1}}{\mu_\ell\cdots\mu_1} \qquad\qquad \ell = \frac{n}{2}.$$

All the distortion functions $\mathcal{K}_{\alpha,\beta}$, $1 \le \alpha,\, \beta \le n-1$, can be expressed in terms of the distortion functions $\mathcal{K}_\ell$ by the formula

$$\mathcal{K}_{\alpha,\beta} = (\mathcal{K}_\alpha^\beta \mathcal{K}_{n-\beta}^\alpha)^{\frac{1}{n}}.$$

Having analysed these distortion functions for linear mappings, we define for a ACL^n mapping $\mathbf{f} : \Omega \longrightarrow \mathbb{R}^n$ the pointwise distortion functions

$$\mathcal{K}_{\alpha,\beta}(\mathbf{x}, \mathbf{f}) = \mathcal{K}_{\alpha,\beta}(\mathbf{f}'(\mathbf{x})) \tag{3.29}$$

and in even dimensions

$$\mathcal{K}(\mathbf{x}, \mathbf{f}) = \mathcal{K}(\mathbf{f}'(\mathbf{x})) \tag{3.30}$$

at the points $x \in \Omega$ where the differential exists (see [42] p.111). We will finish this section by recalling the definition of the distortion tensor

$$G(x) = \begin{cases} |J_{\mathbf{f}}(\mathbf{x})|^{-\frac{2}{n}}(\mathbf{f}'(\mathbf{x}))^t(\mathbf{f}(\mathbf{x})) & \text{if } J_{\mathbf{f}}(\mathbf{x}) \neq 0 \\ I & \text{if } J_{\mathbf{f}}(\mathbf{x}) = 0. \end{cases} \tag{3.31}$$

3.3 Convergence of Quasiconformal Mappings

In the previous two sections, we showed that there are three definitions of quasiconformality that are equivalent. They were curve family, analytic and geometric definitions (see [36]). We recall some definitions and properties of convergence sequences and after that, we start the convergence of quasiconformal mappings.

Let Ω and Ω' be domains in $\mathbb{R}^n$, and $\mathbf{f_n} : \Omega \longrightarrow \Omega'$ be a sequence of functions. We say that the sequence $\{\mathbf{f_n}\}_{n=1}^\infty$ is **uniformly convergent** with limit $\mathbf{f} : \Omega \longrightarrow \Omega'$ if and only if for every $\epsilon > 0$, there exists an integer N such that

$$\|\mathbf{f_n}(\mathbf{x}) - \mathbf{f}(\mathbf{x})\| < \epsilon,$$

for all $n \ge N$ and all $\mathbf{x} \in \Omega$. The sequence $\{\mathbf{f_n}\}_{n=1}^\infty$ is said to be **locally uniformly convergent** with limit $\mathbf{f}$, if Ω is a metric space and for every $\mathbf{x} \in \Omega$, there exists an $r > 0$ such that the sequence $\{\mathbf{f_n}\}_{n=1}^\infty$ is uniformly convergent on $B(\mathbf{x}, r) \cap \Omega$. The sequence $\{\mathbf{f_n}\}_{n=1}^\infty$ is **pointwise convergent** to $\mathbf{f}$ if and only if for every $\epsilon > 0$ and every $\mathbf{x} \in \Omega$, there exists an integer N, depending on ϵ and on $\mathbf{x}$, such that

$$\|\mathbf{f_n}(\mathbf{x}) - \mathbf{f}(\mathbf{x})\| < \epsilon,$$

for all $n > N$. It is clear that uniformly convergence implies locally uniformly convergence and easy to see that local uniformly convergence results in pointwise convergence (see [63]). Let (X, τ) be a topological space and (Y, d_Y) be a metric space. A sequence $\{\mathbf{f_n}\}_{n=1}^\infty$ of functions

$\mathbf{f_n} : X \longrightarrow Y$ is said to converge to $\mathbf{f} : X \longrightarrow Y$ uniformly on compact subsets of X if, for every compact set $K \subseteq X$ and every $\epsilon > 0$, there is the positive integer $N_{K,\epsilon}$ such that

$$\|\mathbf{f_n}(\mathbf{x}) - \mathbf{f}(\mathbf{x})\| < \epsilon,$$

for all $\mathbf{x} \in X$ and $n \leq N$ (see [64] and [55]). Now we want to investigate the convergence of quasiconformal mappings. For this purpose, we explain some theorems such that their proofs are omitted.

Definition 3.3.1. Let $\{D_n\}$ be a sequence of sets in $\hat{\mathbb{R}}^n$ with $n \leq 2$. The kernel of $\{D_n\}$, denoted by $ker_{n\to\infty}D_n$ is the set defined as follows:

$$ker_{n\to\infty}D_n = \bigcup_{n=1}^{\infty} int\Big(\bigcap_{i=n}^{\infty} D_i\Big).$$

Hence, a point x belongs to $ker_{n\to\infty}D_n$ if and only if x has a neighbourhood U that is included in D_n for all sufficiently large n. This implies that each compact subset of $ker_{n\to\infty}D_n$ is contained in D_n once n suitably large. It is obvious that $ker_{n\to\infty}D_n$ is an open set and this set need not be connected, however, even if each of the sets D_n is a domain ([30] p.284).

Theorem 3.3.1. Let $\{D_n\}$ be a sequence of domains in $\hat{\mathbb{R}}^n$ with $n \leq 2$, let $\mathbf{f_n}$ be a K-quasiconformal mapping of D_n onto a domain D'_n, and let D be a subdomain of $ker_{n\to\infty}D_n$. Assume that $\mathbf{f_n} \longrightarrow \mathbf{f}$ pointwise in D. There are the following three possibilities for the limit mapping $\mathbf{f}$:

1. $\mathbf{f}$ might take exactly two values in D (one of these at precisely one point of D) in which event the convergence of $\{\mathbf{f_n}\}_{n=1}^{\infty}$ is not locally uniform in D;
2. $\mathbf{f}$ might be a homeomorphism of D onto a subdomain of $ker_{n\to\infty}D'_n$, in which event the convergence of $\{\mathbf{f_n}\}_{n=1}^{\infty}$ is locally uniform in D;
3. $\mathbf{f}$ might be constant in D, in which event the convergence of $\{\mathbf{f_n}\}_{n=1}^{\infty}$ may or may not be locally uniform in D.

Proof. There is long and complete proof in [30] in pages 284 to 286. Also you can see [74] pages 69 and 70 or [28]. ♠

Theorem 3.3.2. Let $\{D_n\}$ be a sequence of domains in $\hat{\mathbb{R}}^n$ with $n \leq 2$, let $\mathbf{f_n}$ be a K-quasiconformal mapping of D_n into $\hat{\mathbb{R}}^n$, and let D be a subdomain of $ker_{n\to\infty}D_n$. Suppose that $\mathbf{f_n} \longrightarrow \mathbf{f}$ locally uniformly in D, where $\mathbf{f}$ is a homeomorphism. Then

$$K_I(\mathbf{f}) \leq \liminf_{n\to\infty} K_I(\mathbf{f_n}), \qquad K_O(\mathbf{f}) \leq \liminf_{n\to\infty} K_O(\mathbf{f_n}).$$

In fact, if each of the mappings $\mathbf{f_n}$ is a K-quasiconformal mapping, then $\mathbf{f}$ is a K-quasiconformal mapping as well.

Corollary 3.3.2.1. Let D be a domain in $\mathbb{R}^n$. The functions $K_I(\mathbf{f})$, $K_O(\mathbf{f})$, and $K(\mathbf{f})$ defined on the space of quasiconformal mappings of D into $\mathbb{R}^n$ with topology of local uniform convergence are lower semicontinuous (see [30] p.287).

Theorem 3.3.3. Assume that $\{\mathbf{f_n}\}_{n=1}^{\infty}$ is a sequence of K-quasiconformal mappings of $\hat{\mathbb{R}}^n$ onto itself and that $\mathbf{f_n} \longrightarrow \mathbf{f}$ uniformly on $\hat{\mathbb{R}}^n$. Then $\mathbf{f}$ is a K-quasiconformal mapping of $\hat{\mathbb{R}}^n$ onto itself. Moreover, $\mathbf{f_n^{-1}} \longrightarrow \mathbf{f^{-1}}$ uniformly on $\hat{\mathbb{R}}^n$ as well.

Theorem 3.3.4. If $\{\mathbf{f_n}\}_{n=1}^{\infty}$ is a sequence of homeomorphism of $\hat{\mathbb{R}}^n$ and if $\{\mathbf{f_n}\}_{n=1}^{\infty}$ converges uniformly on $\hat{\mathbb{R}}^n$ to a homeomorphism $\mathbf{f}$, then $\mathbf{f_n^{-1}} \longrightarrow \mathbf{f^{-1}}$ uniformly on $\hat{\mathbb{R}}^n$.

There exist technical proofs in [30] from page 283 to 293. More details can be found in [1], [74] and [32]. These theorems show us that if we have a sequence $\{\mathbf{f_n}\}_{n=1}^{\infty}$ of K-quasiconformal mappings (with curve family definition) on $\hat{\mathbb{R}}^n$, and if $\{\mathbf{f_n}\}_{n=1}^{\infty}$ converges uniformly on $\hat{\mathbb{R}}^n$ to a $\mathbf{f}$, Then $\mathbf{f}$ is a K-quasiconformal mapping of $\hat{\mathbb{R}}^n$ onto itself and the distortion of $\mathbf{f}$ is equal to or less than distortion of all functions $\mathbf{f_n}$. This fact is also true for the analytic definition for distortion.
These theorems show us that, with curve family and analytic definitions of quasiconformal mappings, if $\{\mathbf{f_n}\}_{n=1}^{\infty}$ is a sequence of K-quasiconformal mappings of $\hat{\mathbb{R}}^n$ onto itself and that $\mathbf{f_n} \longrightarrow \mathbf{f}$ uniformly on $\hat{\mathbb{R}}^n$, then $\mathbf{f}$ is a K-quasiconformal mapping of $\hat{\mathbb{R}}^n$ onto itself. This theorem says the sequence $\{\mathbf{f_n}\}_{n=1}^{\infty}$ and function $\mathbf{f}$ have the same constants and sometimes the constant of $\mathbf{f}$ is less than the constant of $\{\mathbf{f_n}\}_{n=1}^{\infty}$ (for more details and proofs you can see [42], [30], [74] and [40]).

3.4 The Failure of Lower Semicontinuity of Linear Distortion

Theorems 3.3.3 and 3.3.4 show the lower semicontinuity holds for geometric and analytic definitions of quasiconformality. But for the metric definition of quasiconformality, we have not proved the linear distortion function to be lower semicontinuous. If $\{\mathbf{f_n}\}_{n=1}^{\infty}$ is a sequence of $\mathcal{H}$-quasiconformal mappings (metric definition) of $\mathbb{R}^n$ onto itself and that $\mathbf{f_n} \longrightarrow \mathbf{f}$ uniformly on $\hat{\mathbb{R}}^n$, then $\mathbf{f}$ is a K-quasiconformal mapping of $\mathbb{R}^n$ but $\mathbf{f}$ might not be a $\mathcal{H}$-quasiconformal mapping. We can ask the following question regarding to lower semicontinuity of the linear distortion: In the above conditions for metric definition, is

$$\mathcal{H}(\mathbf{x}, \mathbf{f}) \leq \limsup_{n\to\infty} \mathcal{H}(\mathbf{x}, \mathbf{f_n})?$$

This question has been answered negatively by Tadeusz Iwaniec (see [24]). In the article [39], he proved for a certain sequence of quasiconformal mappings $\mathbf{f_n}$ converging to $\mathbf{f}$ locally uniformly that $\mathcal{H}(\mathbf{x}, \mathbf{f}) > \mathcal{H}(\mathbf{x}, \mathbf{f_n})$ almost everywhere in $\mathbb{R}^n$. The key idea is that the linear distortion function fails to be rank-one convex in dimensions $n \geq 3$. At first, we recall some definitions and after that explain the lemma such that it is useful for the next theorem. Recall that $M_{n\times n}(\mathbb{R})$ denotes the real vector space of all $n \times n$ matrices endowed with the norm

$$|A| = \max_{|h|=1} |Ah|.$$

By Theorem 1.8.1, a matrix B is rank one if and only if it can be written as the tensor product of two vectors. Therefore, there are two vectors $\mathbf{u} = (u_1, u_2, \ldots, u_n)$ and $\mathbf{v} = (v_1, v_2, \ldots, v_n)$ in $\mathbb{R}^n$ such that

$$B = \mathbf{u} \otimes \mathbf{v} = [b_{ij}] = [u_i v_j], \qquad i, j = 1, 2, \ldots, n.$$

By definition of linear distortion for matrix $A \in M_{n\times n}(\mathbb{R})$ (3.18) we have

$$\mathcal{H}(A) = \frac{\max_{|h|=1} |Ah|}{\min_{|h|=1} |Ah|}.$$

The linear distortion function $\mathcal{H}(A)$ is not rank-one convex. This content is described in the next lemma.

Lemma 3.4.1. (Iwaniec's Lemma) Given $n \geq 3$ and $H > 1$, there is a matrix A and a rank-one matrix B and numbers $t, s > 0$ such that

$$\mathcal{H}(A - sB) = \mathcal{H}(A + tB) = H < \mathcal{H}(A).$$

There is a long and technical proof with interesting computation in [39] in dimension 3. This proof can be extended to higher dimensions. Now we can explain the theorem that the linear distortion function fails to be rank-one convex in dimensions higher than 2 (see [39]).

Theorem 3.4.2. (Iwaniec's Theorem) For each $n \geq 3$ and $H > 1$, there exists a sequence $\{\mathbf{f_n}\}_{n=1}^{\infty}$ of quasiconformal mappings $\mathbf{f_n} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ converging uniformly to a linear quasiconformal map $\mathbf{f_0} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ such that

$$\mathcal{H}(\mathbf{x}, \mathbf{f_n}) \equiv H < \mathcal{H}(\mathbf{x}, \mathbf{f_0}), \qquad \text{almost everywhere in } \mathbb{R}^n \quad n = 1, 2, \ldots$$

Proof. The complete proof is in [42], page 198. ♠

In the next chapter, we try to determine the general cases for this question and answer them. We are interested in finding the best direction of vectors **u** and **v** and the interval for variable t so that the conditions are satisfied in that interval. Also, we solve these questions for general linear mapping.

Chapter 4

New Research

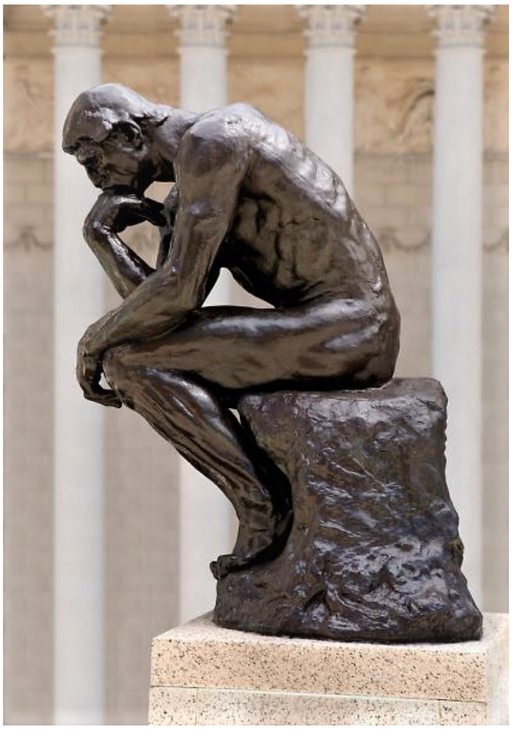

Figure 4.1: The Thinker is a bronze sculpture by Auguste Rodin
Source: Images for The Thinker

Chapter 4

New Research

Iwaniec showed in [39] that there exists at least one linear mapping T with the property described in Theorem 3.4.2. That is there is a sequence $\{\mathbf{f}_\nu\}$ of $\mathcal{H}$-quasiconformal mappings ($\mathcal{H}$ is the linear distortion) and $\mathbf{f}_\nu \longrightarrow T$ locally uniformly in $\mathbb{R}^n$ and

$$\lim \mathcal{H}(\mathbf{f}_\nu) = H_\infty < \mathcal{H}(T). \tag{4.1}$$

In order to get our generalisation, we note that, Iwaniec's construction will work (it is a common construction in the calculus of variations) provided we are able to solve the following three problems. Given a linear mapping $\mathbf{x} \mapsto A\mathbf{x} + b$, we must find a rank-one direction B_0 such that $t \mapsto \mathcal{H}(A + tB_0)$ has a local maximum at $t = 0$. We will identify the best rank-one direction in the sense that

$$\left.\frac{\mathrm{d}}{\mathrm{dt}}\right|_{t=0} \mathcal{H}(A + t\,B_0) = 0, \qquad \text{and} \qquad \left.\frac{\mathrm{d}^2}{\mathrm{dt}^2}\right|_{t=0} \mathcal{H}(A + t\,B_0), \tag{4.2}$$

is largest. This is a long calculation which occupies most of the chapter. Then, to estimate the jump between $\lim \mathcal{H}(\mathbf{f}_\nu)$ and $\mathcal{H}(T)$, as at (4.1), we must identify the largest interval on which $t \mapsto \mathcal{H}(A + tB_0)$ is concave. We will identify this interval for all linear mappings T with distinct singular values. Problem 1. is about finding the best rank-one direction B that which maximises the difference between $\mathcal{H}(A+tB)$ and $\mathcal{H}$ where A is a 3×3 diagonal matrix with distinct singular values.

Problem 1. Let

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{bmatrix}$$

be diagonal with $1 < a < b$. Determine vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^3$ with $\|\mathbf{u}\| = \|\mathbf{v}\| = 1$ so that with $B_0 = \mathbf{u} \otimes \mathbf{v}$, we have

$$\left.\frac{\mathrm{d}}{\mathrm{dt}}\right|_{t=0} \mathcal{H}(A + t\,B_0) \;=\; 0 \tag{4.3}$$

$$\left.\frac{\mathrm{d}^2}{\mathrm{dt}^2}\right|_{t=0} \mathcal{H}(A + t\,B_0) \;\leq\; \left.\frac{\mathrm{d}^2}{\mathrm{dt}^2}\right|_{t=0} \mathcal{H}(A + t\,B) \tag{4.4}$$

for every rank one matrix $B \in \mathbb{R}^{3\times 3}$ with $\left.\frac{\mathrm{d}}{\mathrm{dt}}\right|_{t=0} \mathcal{H}(A + t\,B) = 0$.

Thus Problem 1. asks us to find the **best rank one direction**. That which maximises the negative of the second derivative. With regard to this problem, we might also ask if this direction is unique (up to sign). In this direction, we expect to find the minimum values of $\mathcal{H}(A+tB)$ and this leads us to the second problem, to find an interval around 0, where this happens.

We know the linear distortion function $\mathcal{H}(A+tB)$ is a real-valued function with respect to t. Problem 2. investigates what would be the biggest interval for t in the direction identified by the solution to Problem 1.

Problem 2. Let

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{bmatrix}$$

be diagonal with $1 < a < b$ and suppose $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^3$ have $\|\mathbf{u}\| = \|\mathbf{v}\| = 1$ so that $B_0 = \mathbf{u} \otimes \mathbf{v}$ is a solution to Problem 1.

1. Determine the largest real numbers $\mathbf{t}_+ > 0$ and $\mathbf{t}_- < 0$ so that $\mathcal{H}(A + t\,B_0)$ is a smooth function of t in the interval $\mathbf{t}_- < t < \mathbf{t}_+$.
2. Determine $\mathcal{H}(A + \mathbf{t}_- B_0)$ and $\mathcal{H}(A + \mathbf{t}_+ B_0)$.

We expect that the values $\mathbf{t}_-$ and $\mathbf{t}_+$ to be where the singular values of $A + tB_0$ cross. These will be determined from a (rather challenging) discriminant problem.

Conjecture: We also conjecture for all rank-one matrices B with

$$\frac{\mathrm{d}}{\mathrm{dt}}\Big|_{t=0} \mathcal{H}(A + t\,B) = 0, \quad \text{and} \quad \frac{\mathrm{d}^2}{\mathrm{dt}^2}\Big|_{t=0} \mathcal{H}(A + t\,B) < 0,$$

we have for all $t > 0$

$$\mathcal{H}(A + tB) \geq \max\{\mathcal{H}(A + \mathbf{t}_- B_0), \mathcal{H}(A + \mathbf{t}_+ B_0)\}.$$

This conjecture expresses the hope that the best direction also leads to the biggest gap between $\mathcal{H}(A)$ and $\mathcal{H}(A + t\,B)$ and therefore gives us the approximation to A of least linear distortion.

The solution to Problem 2, through the construction, leads to the identification of the biggest jump between $\mathcal{H}(A)$ and $\mathcal{H}(A + t\,B_0)$ in the best direction. This requires another lengthy calculation examining the behaviour of the singular values of $A + tB_0$ to determine when they cross. We solve these first two problems in the case A is diagonal, then show how to reduce the general case to this case. Finally, we put all this together to get the generalisation we seek with an estimate of the largest possible jump in linear distortion.

Iwaniec's Lemma gives an example that fails lower semicontinuity. If

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & c^2 \end{bmatrix},$$

and $B = \mathbf{u} \otimes \mathbf{v}$, where $\mathbf{u} = (1, b, c)$ and $\mathbf{v} = (1, -b, c)$. Thus

$$\mathcal{H}(A + tB) = c^2 - 2t^2 + O(t^3) \leq \mathcal{H}(A) = c^2.$$

It follows that there is some some small t_0 so that

$$\mathcal{H}(A) > \mathcal{H}(A + \mathbf{t}_0 B).$$

So, the lower semicontinuity property for linear distortion functions does not hold (see [39]). As you see, there are only two variables in his example, b and c. Despite the simplicity of having two variables in solving this problem, it was still solved with a complex and extraordinary calculation, and this shows us that solving the problem, in general, is basically a very difficult discussion.

The natural question here would be what happens if matrix C, instead of being diagonal, is an arbitrary 3×3 matrix with distinct singular values. Next Problem 3. addresses this question. We know that every matrix C with distinct singular values can be factorised by Theorem 1.6.10 as

$$C = UAV^t,$$

where U and V^t are 3×3 orthogonal matrices and A is a diagonal matrix with entries that are singular values of the matrix C.

Problem 3. Let $C \in GL(3, \mathbb{R})$ and suppose that C has distinct singular values (so C^tC has distinct real eigenvalues). Then

1. Determine the best rank-one direction B^* (as per Problem 1) for C.
2. Determine $\mathbf{t}_-, \mathbf{t}_+$ as per Problem 2, and also $\mathcal{H}(C+\mathbf{t}_-B^*)$ and $\mathcal{H}(C+\mathbf{t}_+B^*)$, that is the largest interval of concavity for $\mathcal{H}(C+tB^*)$ and the smallest values for the linear distortion of an approximation.

Finally, we put the solutions to these three problems together to get our main result.

4.1 Solving Problem 1.

Theorem 4.1.1. Let

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{bmatrix}$$

be diagonal with $1 < a < b$. Then there are vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^3$ with $\|\mathbf{u}\| = \|\mathbf{v}\| = 1$ such that with $B_0 = \mathbf{u} \otimes \mathbf{v}$, we have

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{dt}}\Big|_{t=0} \mathcal{H}(A + t\,B_0) &= 0 \\ \frac{\mathrm{d}^2}{\mathrm{dt}^2}\Big|_{t=0} \mathcal{H}(A + t\,B_0) &\leq \frac{\mathrm{d}^2}{\mathrm{dt}^2}\Big|_{t=0} \mathcal{H}(A + t\,B) \end{aligned}$$

for every rank one matrix $B \in \mathbb{R}^{3\times3}$ with $\frac{\mathrm{d}}{\mathrm{dt}}\Big|_{t=0} \mathcal{H}(A + t\,B) = 0$.

Proof. Our problem now is to determine the best rank-one matrix $B_0 = \mathbf{u}_0.\mathbf{v}_0^t$, $\|\mathbf{u}_0\| = \|\mathbf{v}_0\| = 1$, so that

$$\frac{d}{dt}\Big|_{t=0} \mathcal{H}(A + tB_0) = 0, \tag{4.5}$$

and that for all other rank one matrices B with $\frac{d}{dt}\Big|_{t=0} \mathcal{H}(A + tB) = 0$ we have

$$\frac{d^2}{dt^2}\Big|_{t=0} \mathcal{H}(A + tB_0) \leq \frac{d^2}{dt^2}\Big|_{t=0} \mathcal{H}(A + tB). \tag{4.6}$$

That is the coefficient of the quadratic term in the series expansion of $\mathcal{H}(A+tB)$ is as negative as possible. Since $\mathbf{u}, \mathbf{v} \in \mathbb{S}^2$, a compact space, there is an absolute minimum. Let

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{bmatrix}$$

be a diagonal matrix and B be a rank-one matrix such that $B = \mathbf{u} \otimes \mathbf{v} = \mathbf{u} \cdot \mathbf{v}^t$, where

$$\mathbf{u}^t = (\sqrt{1-r^2}, r\cos(\theta_1), r\sin(\theta_1))$$

and

$$\mathbf{v}^t = (\sqrt{1-s^2}, s\cos(\theta_2), s\sin(\theta_2))$$

are unit vectors in $\mathbb{R}^3$, for $0 \leq r, s \leq 1$ and θ_1, θ_2 in $[0, 2\pi]$. The matrix A^tA has three eigenvalues 1, a^2 and b^2 where $1 < a < b$. By definition of $\mathcal{H}(A)$ we have

$$\mathcal{H}(A) = \sqrt{\frac{\text{The largest eigenvalue of } A^tA}{\text{The smallest eigenvalue of } A^tA}} = \sqrt{\frac{b^2}{1}} = b.$$

But $\mathcal{H}(A+tB)$ is a function with respect to t, r, s, θ_1 and θ_2. Since $(A+tB)^t(A+tB)$ is symmetric, then all eigenvalues of $X = (A+tB)^t(A+tB)$ are nonnegative real numbers. Assume that $\lambda_1(t)$, $\lambda_2(t)$ and $\lambda_3(t)$ are eigenvalues of X. The functions $\lambda_1(t)$, $\lambda_2(t)$ and $\lambda_3(t)$ are continuous and differentiable for degree two of differentiations with respect to t. By using the Taylor series for several variable functions, we get

$$\lim_{t\to 0} \lambda_1(t) = \lambda_1(0) = 1, \qquad \lim_{t\to 0} \lambda_2(t) = \lambda_2(0) = a^2, \qquad \lim_{t\to 0} \lambda_3(t) = \lambda_3(0) = b^2.$$

These limits show that for sufficiently small t in the neighbourhood of 0, we have

$$\lambda_1(t) < \lambda_2(t) < \lambda_3(t).$$

So, for t sufficiently small,

$$\mathcal{H}(A+tB) = \sqrt{\frac{\lambda_3(t)}{\lambda_1(t)}}.$$

If t tends to infinity then

$$\lim_{t\to+\infty} \mathcal{H}(A+tB) = \lim_{t\to+\infty} \mathcal{H}(\frac{A}{t} + B) = \lim_{t\to+\infty} \mathcal{H}(B) = \infty,$$

because B is a rank-one matrix. The two matrices $A+tB$ and $\frac{1}{t}A+B$ have the same ratio of singular values. That means $\mathcal{H}(A+tB)$ and $\mathcal{H}(\frac{A}{t}+B)$ are conformally equivalent. Note that

$$(\frac{A}{t} + B)^t(\frac{A}{t} + B) = \frac{1}{t^2}A^tA + \frac{1}{t}(A^tB + B^tA) + B^tB.$$

Eigenvalues are solutions to the an equation with coefficients from this matrix. Eigenvalues converge locally uniformly to B^tB which has only one non-zero eigenvalue. For various choices of B the following figures show what can happen.

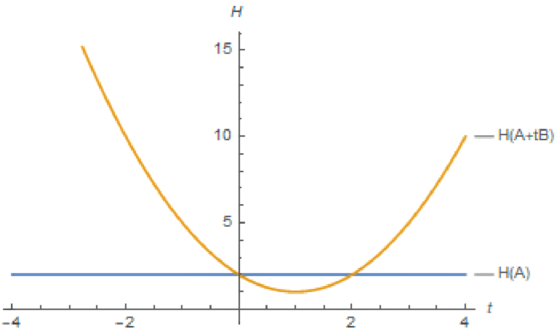

Figure 4.2: Linear distortion function $\frac{\mathrm{d}}{\mathrm{dt}}\mathcal{H}(A+tB)\Big|_{t=0} < 0$, $\frac{\mathrm{d}^2}{\mathrm{dt}^2}\mathcal{H}(A+tB)\Big|_{t=0} > 0$.

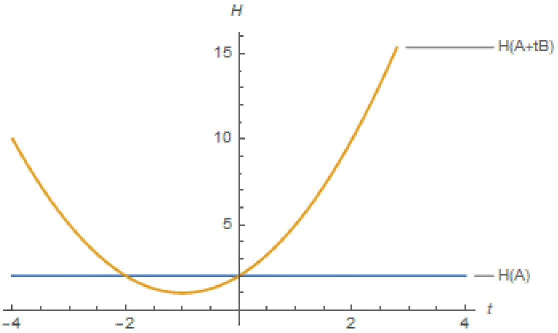

Figure 4.3: Linear distortion function $\frac{\mathrm{d}}{\mathrm{dt}}\mathcal{H}(A+tB)\Big|_{t=0} > 0$, $\frac{\mathrm{d}^2}{\mathrm{dt}^2}\mathcal{H}(A+tB)\Big|_{t=0} > 0$.

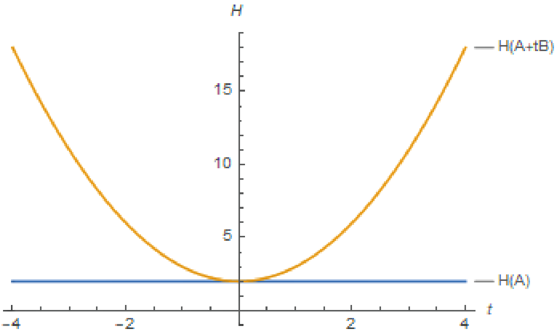

Figure 4.4: Linear distortion function $\frac{\mathrm{d}}{\mathrm{dt}}\mathcal{H}(A+tB)\Big|_{t=0} = 0$, $\frac{\mathrm{d}^2}{\mathrm{dt}^2}\mathcal{H}(A+tB)\Big|_{t=0} > 0$.

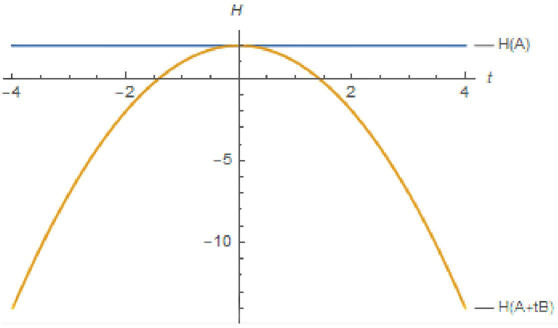

Figure 4.5: Linear distortion function $\frac{\mathrm{d}}{\mathrm{dt}}\mathcal{H}(A+tB)\Big|_{t=0} = 0$, $\frac{\mathrm{d}^2}{\mathrm{dt}^2}\mathcal{H}(A+tB)\Big|_{t=0} < 0$.

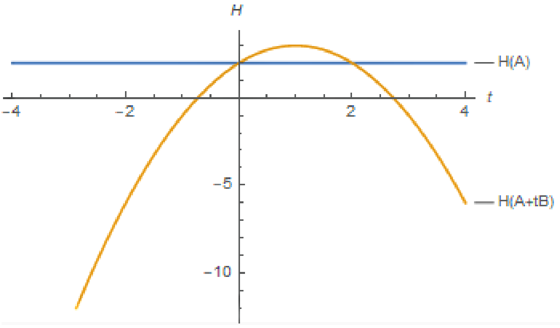

Figure 4.6: Linear distortion function $\frac{\mathrm{d}}{\mathrm{dt}}\mathcal{H}(A+tB)\Big|_{t=0} > 0$, $\frac{\mathrm{d}^2}{\mathrm{dt}^2}\mathcal{H}(A+tB)\Big|_{t=0} < 0$.

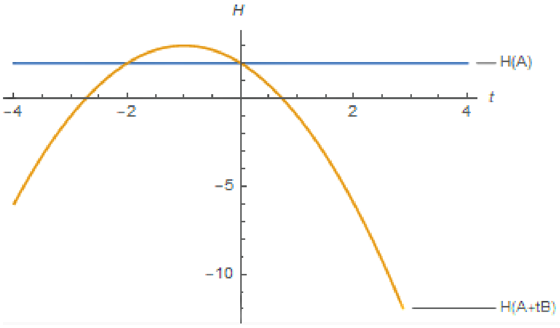

Figure 4.7: Linear distortion function $\frac{\mathrm{d}}{\mathrm{dt}}\mathcal{H}(A+tB)\Big|_{t=0} < 0$, $\frac{\mathrm{d}^2}{\mathrm{dt}^2}\mathcal{H}(A+tB)\Big|_{t=0} < 0$.

The functions $\mathcal{H}(A)$ and $\mathcal{H}(A+tB)$ may have one of the six previous graphs on a neighbourhood of 0. Those six cases depend on the signs of the first and the second derivatives of the function $\mathcal{H}(A+tB)$. In all cases the function $\mathcal{H}(A)$ is the constant b. In Figures 4.2 and 4.7 for positive small t near to 0, the function $\mathcal{H}(A+tB)$ is less than $\mathcal{H}(A)$ and for negative t near to 0 is bigger than $\mathcal{H}(A)$. In Figures 4.3 and 4.6 the function $\mathcal{H}(A+tB)$ for positive small t near to 0, is bigger and for negative t near to 0, is less than $\mathcal{H}(A)$. Also, in the Figure 4.4, we have

$$\frac{\mathrm{d}}{\mathrm{dt}}\mathcal{H}(A+tB)\Big|_{t=0} = 0, \qquad\qquad \frac{\mathrm{d}^2}{\mathrm{dt}^2}\mathcal{H}(A+tB)\Big|_{t=0} > 0.$$

These equations show us the function $\mathcal{H}(A+tB)$ on a neighbourhood of 0, is bigger than $\mathcal{H}(A)$ (see Figure 4.4). But in Figure 4.5, the first and the second derivative of $\mathcal{H}(A+tB)$ are

$$\frac{\mathrm{d}}{\mathrm{dt}}\mathcal{H}(A+tB)\Big|_{t=0} = 0, \qquad\qquad \frac{\mathrm{d}^2}{\mathrm{dt}^2}\mathcal{H}(A+tB)\Big|_{t=0} < 0.$$

These equations show that there is a neighbourhood of 0 such that for every $t \neq 0$ on this neighbourhood we have

$$\mathcal{H}(A+tB) < \mathcal{H}(A).$$

It is very obvious that in order to answer our main question, the state of Figure 4.5 is a desirable answer. In this case for every sufficiently small t near to 0 there are vectors $\mathbf{u}$ and $\mathbf{v}$ such that

$$\mathcal{H}(A+t\mathbf{u}\otimes\mathbf{v}) < \mathcal{H}(A).$$

In the rest of this section we try to find best vectors $\mathbf{u}_0$ and $\mathbf{v}_0$ and relations between their components and the entries of matrix A such that

$$\frac{\mathrm{d}}{\mathrm{dt}}\mathcal{H}(A+tB_0)\Big|_{t=0} = 0, \qquad\qquad \frac{\mathrm{d}^2}{\mathrm{dt}^2}\mathcal{H}(A+tB_0)\Big|_{t=0} < 0. \tag{4.7}$$

The matrix B can be written as

$$B = \begin{bmatrix} \sqrt{1-r^2}\sqrt{1-s^2} & \sqrt{1-r^2}s\cos(\theta_2) & \sqrt{1-r^2}s\sin(\theta_2) \\ r\sqrt{1-s^2}\cos(\theta_1) & rs\cos(\theta_1)\cos(\theta_2) & rs\cos(\theta_1)\sin(\theta_2) \\ r\sqrt{1-s^2}\sin(\theta_1) & rs\cos(\theta_2)\sin(\theta_1) & rs\sin(\theta_1)\sin(\theta_2) \end{bmatrix}.$$

Now, we consider the matrix $A+tB$.

$$A+tB = \begin{bmatrix} 1+t\sqrt{1-r^2}\sqrt{1-s^2} & ts\sqrt{1-r^2}\cos(\theta_2) & ts\sqrt{1-r^2}\sin(\theta_2) \\ tr\sqrt{1-s^2}\cos(\theta_1) & a+trs\cos(\theta_1)\cos(\theta_2) & trs\cos(\theta_1)\sin(\theta_2) \\ tr\sqrt{1-s^2}\sin(\theta_1) & trs\cos(\theta_2)\sin(\theta_1) & b+trs\sin(\theta_1)\sin(\theta_2) \end{bmatrix}.$$

Then set $X = (A+tB)^T(A+tB)$. We calculate for small t the smallest and the biggest eigenvalues of X. Let I be the identity 3×3 matrix, then the smallest eigenvalues of X can be found by

$$\det[X-\lambda_1 I] \approx \det[A+tB-(1+xt+yt^2)I] = 0,$$

where $\lambda_1 < \lambda_2 < \lambda_3$ are eigenvalues of the matrix X. The equation above shows:

$$x = 2\sqrt{1-r^2}\sqrt{1-s^2},$$

$$y = 1-s^2-\frac{\Big(ar\sqrt{1-s^2}\cos(\theta_1)+s\sqrt{1-r^2}\cos(\theta_2)\Big)^2}{a^2-1}-\frac{\Big(br\sqrt{1-s^2}\sin(\theta_1)+s\sqrt{1-r^2}\sin(\theta_2)\Big)^2}{b^2-1}.$$

So, we can write the first three terms of the smallest eigenvalue of the X. Therefore

$$\lambda_1(t) = 1+2\big(\sqrt{1-r^2}\sqrt{1-s^2}\big)t+\Bigg(1-s^2-\frac{\big(ar\sqrt{1-s^2}\cos(\theta_1)+s\sqrt{1-r^2}\cos(\theta_2)\big)^2}{a^2-1}-$$

$$\frac{\big(br\sqrt{1-s^2}\sin(\theta_1)+s\sqrt{1-r^2}\sin(\theta_2)\big)^2}{b^2-1}\Bigg)t^2+O(t^3).$$

With the same process, we can use the first three terms of the biggest eigenvalue of the matrix X and approximate the eigenvalue. The biggest eigenvalue is

$$\lambda_3(t) = b^2+2\,brs\sin(\theta_1)\sin(\theta_2)\;t+\Bigg(\frac{s^2}{2}-\frac{1}{2}s^2\cos(2\theta_2)+$$

$$\frac{\big(br\sqrt{1-s^2}\sin(\theta_1)+s\sqrt{1-r^2}\sin(\theta_2)\big)^2}{b^2-1}+\frac{s^2\big(rb\,\cos(\theta_2)\sin(\theta_1)+ra\,\cos(\theta_1)\sin(\theta_2)\big)^2}{b^2-a^2}\Bigg)t^2$$

$$+O(t^3).$$

Therefore for small enough t, the linear distortion function $\mathcal{H}$ is defined by

$$\mathcal{H}(A+tB)=\sqrt{\frac{\lambda_3(t)}{\lambda_1(t)}}=\frac{\mu_3}{\mu_1},$$

where $\lambda_3(t)$ and $\lambda_1(t)$ are the largest and smallest eigenvalues of the matrix X and μ_3 and μ_1 are the largest and the smallest singular values of X, respectively. By using Taylor series, $\mathcal{H}(A+tB)$ can be written as

$$\mathcal{H}(A+tB)=b+\omega_1 t+\omega_2 t^2+O(t^3),$$

where ω_1 and ω_2 can be described in terms of the four variables r, s, θ_1 and θ_2 and two parameters a and b. If we expand the linear distortion $\mathcal{H}(A+tB)$, then the exact formula takes the form:

$$H(A+tB)=b+\big(-b\sqrt{1-r^2}\sqrt{1-s^2}+rs\sin(\theta_1)\sin(\theta_2)\big)t+\frac{1}{2b}\Bigg(\frac{s^2}{2}-\frac{1}{2}s^2\cos(2\theta_2)$$

$$-4brs\sqrt{1-r^2}\sqrt{1-s^2}\;\sin(\theta_1)\sin(\theta_2)+\frac{\big(br\sqrt{1-s^2}\sin(\theta_1)+s\sqrt{1-r^2}\sin(\theta_2)\big)^2}{b^2-1}+$$

$$\frac{s^2\big(rb\,\cos(\theta_2)\sin(\theta_1)+ra\,\cos(\theta_1)\sin(\theta_2)\big)^2}{b^2-a^2}-\big(b\sqrt{1-r^2}\sqrt{1-s^2}-rs\sin(\theta_1)\sin(\theta_2)\big)^2+ \tag{4.8}$$

$$b^2\Bigg(-1+s^2+4(r^2-1)(s^2-1)+\frac{\Big(ar\sqrt{1-s^2}\cos(\theta_1)+s\sqrt{1-r^2}\cos(\theta_2)\Big)^2}{a^2-1}+$$

$$\frac{\big(br\sqrt{1-s^2}\sin(\theta_1)+s\sqrt{1-r^2}\sin(\theta_2)\big)^2}{b^2-1}\Bigg)\Bigg)t^2+O(t^3).$$

By equation (4.5), the first derivative of $\mathcal{H}(A+tB)$ must be zero. So

$$\big(-b\big(\sqrt{1-r^2}\sqrt{1-s^2}\,\big)+rs\sin(\theta_1)\sin(\theta_2)\big)=0.$$

After simplification, we get

$$\begin{aligned} b\big(\sqrt{1-r^2}\sqrt{1-s^2}\,\big)&=rs\sin(\theta_1)\sin(\theta_2),\\ b^2(1-r^2)(1-s^2)&=r^2s^2\sin^2(\theta_1)\sin^2(\theta_2) \end{aligned}\tag{4.9}$$

We want to minimise the quadratic coefficient of Equation (4.8). The quadratic coefficient is a function of the four variables r, s, θ_1 and θ_2 and the two parameters a and b. The quadratic coefficient is

$$\begin{aligned} Q(r,s,\theta_1,\theta_2)=\frac{1}{2b}\Bigg(&\frac{s^2}{2}-\frac{1}{2}s^2\cos(2\theta_2)-4brs\sqrt{1-r^2}\sqrt{1-s^2}\,\sin(\theta_1)\sin(\theta_2)+\\ &\frac{\big(br\sqrt{1-s^2}\sin(\theta_1)+s\sqrt{1-r^2}\sin(\theta_2)\big)^2}{b^2-1}+\frac{s^2\big(rb\,\cos(\theta_2)\sin(\theta_1)+ra\,\cos(\theta_1)\sin(\theta_2)\big)^2}{b^2-a^2}-\\ &\big(b\sqrt{1-r^2}\sqrt{1-s^2}-rs\sin(\theta_1)\sin(\theta_2)\big)^2+b^2\Bigg(-1+s^2+4(r^2-1)(s^2-1)+\\ &\frac{\Big(ar\sqrt{1-s^2}\cos(\theta_1)+s\sqrt{1-r^2}\cos(\theta_2)\Big)^2}{a^2-1}+\frac{\big(br\sqrt{1-s^2}\sin(\theta_1)+s\sqrt{1-r^2}\sin(\theta_2)\big)^2}{b^2-1}\Bigg)\Bigg). \end{aligned}\tag{4.10}$$

If we replace Equation (4.9) in Equation (4.10), we get

$$\begin{aligned} Q(r,s,\theta_1,\theta_2)=\frac{1}{8}\Bigg(&\frac{r^2\cos(2\theta_1)\big(-2(b^2+b^4)+s^2(-5+3b^2+2b^4)-(b^2-5)s^2\cos(2\theta_2)\big)}{b^3-b}+\\ &\frac{3r^2s^2+2b^4(5r^2-4)(s^2-1)+b^2\big(r^2(14-17s^2)+12s^2-8\big)+(-4b^2+3(b^2-1)r^2)s^2\cos(2\theta_2)}{b^3-b}+\\ &4bs^2\cos^2(\theta_2)\Big(\frac{1-r^2}{a^2-1}+\frac{r^2\sin^2(\theta_1)}{b^2-a^2}\Big)+4a^2r^2\cos^2(\theta_1)\Big(\frac{b(1-s^2)}{a^2-1}+\frac{s^2\sin^2(\theta_1)}{b^3-a^2b}\Big)-\\ &\frac{2a(b^2-1)r^2s^2\sin(2\theta_1)\sin(2\theta_2)}{(a^2-1)(a^2-b^2)}\Bigg). \end{aligned}\tag{4.11}$$

It is obvious that the function Q is π-periodic, so we assume the values of θ_1 and θ_2 are between 0 and π. By looking at (4.11), we have found the equation involves only r^2 and s^2. Assume that $\delta=r^2$ and $\eta=s^2$. The Equation 4.11 can be written as:

$$\begin{aligned}
Q(\delta,\eta,\theta_1,\theta_2) = \frac{1}{8}\Bigg(& \frac{\delta\cos(2\theta_1)\big(-2(b^2+b^4)+\eta(-5+3b^2+2b^4)-(b^2-5)\eta\cos(2\theta_2)\big)}{b^3-b} + \\
& \frac{3\delta\eta+2b^4(5\delta-4)(\eta-1)+b^2\big(\delta(14-17\eta)+12\eta-8\big)+(-4b^2+3(b^2-1)\delta)\eta\cos(2\theta_2)}{b^3-b} + \\
& 4b\eta\cos^2(\theta_2)\Big(\frac{1-\delta}{a^2-1}+\frac{\delta\sin^2(\theta_1)}{b^2-a^2}\Big)+4a^2\delta\cos^2(\theta_1)\Big(\frac{b(1-\eta)}{a^2-1}+\frac{\eta\sin^2(\theta_1)}{b^3-a^2b}\Big) - \\
& \frac{2a(b^2-1)\delta\eta\sin(2\theta_1)\sin(2\theta_2)}{(a^2-1)(a^2-b^2)}\Bigg).
\end{aligned} \tag{4.12}$$

The first derivative of $H(A+tB)$ is

$$\begin{aligned}
& b^2(1-r^2)(1-s^2) = r^2s^2\sin^2(\theta_1)\sin^2(\theta_2) \\
& \Rightarrow b^2(1-\delta)(1-\eta) = \delta\eta\sin^2(\theta_1)\sin^2(\theta_2) \\
& \Rightarrow \delta\eta = \frac{b^2\big(\delta+\eta-1\big)}{b^2-\sin^2(\theta_1)\sin^2(\theta_2)}
\end{aligned} \tag{4.13}$$

We can therefore eliminate the nonlinear term $\delta\eta$, using Equations 4.12 and 4.13, the function Q can be simplified to

$$\begin{aligned}
& Q(\delta,\eta,\theta_1,\theta_2) = -\frac{1}{32\,(a^2-1)\,(b^2-1)\,(b^2-a^2)\,(b^2-\sin^2(\theta_1)\sin^2(\theta_2))} \\
& b\Bigg(b^4(32-7\delta-7\eta)+8b^6(\delta+\eta-2)-7b^2(\delta+\eta)+2a^4\big(\delta+\eta-4-4b^2(\delta+\eta-3)\big) - \\
& a^2\big(\delta-8+b^2(40-21(\delta+\eta))+\eta+8b^4(\delta+\eta)\big)+(a^2-b^2)\Big[\Big(\eta-8-2a^2(-4+4b^2(1+\delta-\eta)+\eta)+ \\
& b^2(8\delta-8b^2(\eta-1)+\eta)\Big)\cos(2\theta_1)+(1-2a^2+b^2)\delta\cos(4\theta_1)\Big]+\Big[8(a^2-1)\big(-b^2(\delta+\eta-2)+ \\
& a^2(-1+b^2(\delta+\eta-1))\big)\cos(2\theta_1)+(a^2-b^2)\Big(-8+\delta+a^2(8-2\delta+8b^2(\delta-\eta-1)+b^2(-8b^2(\delta-1)+ \\
& +\delta+8\eta)+(2a^2-b^2-1)\delta\cos(4\theta_1)\Big)\Big]\cos(2\theta_2)-2(a^2-b^2)(2a^2-b^2-1)\eta\cos(4\theta_2)\sin^2(\theta_1) - \\
& 8ab(b^2-1)^2(\delta+\eta-1)\sin(2\theta_1)\sin(2\theta_2)\Bigg).
\end{aligned} \tag{4.14}$$

Since $1 \leq a \leq b$, then $\xi = 32(a^2-1)(b^2-1)(b^2-a^2)$ and $\mu = \big(b^2-\sin^2(\theta_1)\sin^2(\theta_2)\big)$ are positive. The function Q has four variables, we can rewrite it with respect to four variables δ, η, θ_1 and θ_2 as below,

$$Q(\delta,\eta,\theta_1,\theta_2) = \frac{\alpha\delta+\beta\eta+\gamma}{32(a^2-1)(b^2-1)(b^2-a^2)\big(b^2-\sin^2(\theta_1)\sin^2(\theta_2)\big)}. \tag{4.15}$$

So,

$$Q(\delta,\eta,\theta_1,\theta_2) = \frac{\alpha\delta+\beta\eta+\gamma}{\xi\mu}, \tag{4.16}$$

where α, β and γ are functions with respect to θ_1 and θ_2. Equation 4.14 leads us to the following equations:

$$\begin{aligned}\alpha(\theta_1,\theta_2) =& b\Bigg(a^2-2a^4+(7-21a^2+8a^4)b^2+(7+8a^2)b^4-8b^6-8\big(a^2-1\big)\\ & b^2\cos(2\theta_1)\Big(-a^2+b^2+\big(a^2-1\big)\cos(2\theta_2)\Big)-\big(a^2-b^2\big)\Big[\Big(1+b^2-8b^4\\ & +a^2\big(-2+8b^2\big)\Big)\cos(2\theta_2)+2\Big(1-2a^2+b^2\Big)\cos(4\theta_1)\sin^2(\theta_2)\Big]+\\ & 8ab\Big(b^2-1\Big)^2\sin(2\theta_1)\sin(2\theta_2)\Bigg),\end{aligned} \tag{4.17}$$

$$\begin{aligned}\beta(\theta_1,\theta_2) =& b\Bigg(a^2-2a^4+(7-21a^2+8a^4)b^2+(7+8a^2)b^4-8b^6+\cos(2\theta_1)\\ & \Big[-\big(a^2-b^2\big)\Big(1+b^2-8b^4+a^2\big(8b^2-2\big)\Big)-8\big(a^2-1\big)b^2\cos(2\theta_2)\Big]+\\ & 2\big(b^2-a^2\big)\Big[-4\big(a^2-1\big)b^2\cos(2\theta_2)+\Big(1-2a^2+b^2\Big)\cos(4\theta_2)\sin^2(\theta_1)\Big]\\ & +8ab\Big(b^2-1\Big)^2\sin(2\theta_1)\sin(2\theta_2)\Bigg),\end{aligned} \tag{4.18}$$

$$\begin{aligned}\gamma(\theta_1,\theta_2) =& -b\Bigg(32b^4-16b^6+8a^2\big(1-5b^2\big)+8a^4\big(3b^2-1\big)-8\big(a^2-b^2\big)\\ & \big(a^2-b^2-1\big)\cos(2\theta_1)+\cos(2\theta_2)\Big[8\big(b^2-1\big)\big(a^2-b^2\big)\big(1-a^2+b^2\big)-\\ & 8\big(a^2-1\big)\Big(a^2+\big(a^2-2\big)b^2\Big)\cos(2\theta_1)\Big]+8ab\Big(b^2-1\Big)^2\sin(2\theta_1)\sin(2\theta_2)\Bigg).\end{aligned} \tag{4.19}$$

The constraint is given at (4.13), and since the denominator at (4.14) does not vanish, we may multiply the constraint by this term and clear the multiplicative factor. We then consider Lagrange multipliers to examine the function.

$$F(\delta,\eta,\lambda) = \frac{\alpha\delta+\beta\eta+\gamma}{\xi\mu} - \lambda\big(b^2(1-\delta-\eta)+\delta\eta\mu\big). \tag{4.20}$$

We get the following three equations

1. $\frac{\partial F}{\partial \delta} = b^2\lambda - \eta\lambda\mu + \frac{\alpha}{\xi\mu} = 0,$
2. $\frac{\partial F}{\partial \eta} = b^2\lambda - \delta\lambda\mu + \frac{\beta}{\xi\mu} = 0,$
3. $\frac{\partial F}{\partial \lambda} = b^2(-1+\delta+\eta) - \delta\eta\mu = 0.$

Hence we have the two sets of solutions for δ and η.

$$\delta_1 = \frac{b\Big(b - \sqrt{\frac{\beta}{\alpha}}\sqrt{b^2-\mu}\Big)}{\mu}, \qquad \eta_1 = \frac{b\Big(b - \sqrt{\frac{\alpha}{\beta}}\sqrt{b^2-\mu}\Big)}{\mu}, \qquad \lambda_1 = -\frac{\sqrt{\alpha\beta}}{b\,\xi\,\mu\sqrt{b^2-\mu}}, \tag{4.21}$$

$$\delta_2 = \frac{b\Big(b + \sqrt{\frac{\beta}{\alpha}}\sqrt{b^2-\mu}\Big)}{\mu}, \qquad \eta_2 = \frac{b\Big(b + \sqrt{\frac{\alpha}{\beta}}\sqrt{b^2-\mu}\Big)}{\mu}, \qquad \lambda_2 = \frac{\sqrt{\alpha\beta}}{b\,\xi\,\mu\sqrt{b^2-\mu}}. \tag{4.22}$$

Since $0 \le \delta,\, \eta \le 1$, we must check which is the set of solutions between 0 and 1 that we want. Now let's show $0 \le \delta_1 \le 1$, $0 \le \eta_1 \le 1$, $\delta_2 \ge 1$ and $\eta_2 \ge 1$. We have $\mu = \big(b^2 - \sin^2(\theta_1)\sin^2(\theta_2)\big)$, so

$$\sqrt{b^2-\mu} = \sqrt{\sin^2(\theta_1)\sin^2(\theta_2)} = |\sin(\theta_1)\sin(\theta_2)| = \sin(\theta_1)\sin(\theta_2),$$

which is a positive real number. For δ_2, we have

$$\Big(b^2 + b\sqrt{\frac{\beta}{\alpha}}\,(\sin(\theta_1)\sin(\theta_2))\Big) \ge b^2 \ge \big(b^2 - \sin^2(\theta_1)\sin^2(\theta_2)\big).$$

So,

$$\delta_2 = \frac{\Big(b^2 + b\sqrt{\frac{\beta}{\alpha}}\,(\sin(\theta_1)\sin(\theta_2))\Big)}{\big(b^2 - \sin^2(\theta_1)\sin^2(\theta_2)\big)} \ge 1.$$

Also we can write

$$\Big(b^2 + b\sqrt{\frac{\alpha}{\beta}}\,(\sin(\theta_1)\sin(\theta_2))\Big) \ge b^2 \ge \big(b^2 - \sin^2(\theta_1)\sin^2(\theta_2)\big),$$

therefore,

$$\eta_2 = \frac{\Big(b^2 + b\sqrt{\frac{\alpha}{\beta}}\,(\sin(\theta_1)\sin(\theta_2))\Big)}{\big(b^2 - \sin^2(\theta_1)\sin^2(\theta_2)\big)} \ge 1.$$

We have proved that $\delta_2 > 1$ and $\eta_2 > 1$, this means that δ_2 and η_2 are not solutions for above system of equations. Now we must show that $0 \le \delta_1 \le 1$ and $0 \le \eta_1 \le 1$.
We claim, **the absolute minimum of the function Q is negative**. This can be proved by the following statements. Let $r = s$ and $\theta_2 = \pi - \theta_1$, then the vectors $\mathbf{u}$ and $\mathbf{v}$ are

$$\mathbf{u}^t = (\sqrt{1-r^2}, r\cos(\theta_1), r\sin(\theta_1)),$$

and

$$\mathbf{v}^t = (\sqrt{1-r^2}, -r\cos(\theta_1), r\sin(\theta_1)).$$

Then

$$A+tB=\begin{bmatrix} 1+t(1-r^2) & -tr\sqrt{1-r^2}\cos(\theta_1) & tr\sqrt{1-r^2}\sin(\theta_1)\\ tr\sqrt{1-r^2}\cos(\theta_1) & a-tr^2\cos^2(\theta_1) & tr^2\cos(\theta_1)\sin(\theta_1)\\ tr\sqrt{1-r^2}\sin(\theta_1) & -tr^2\cos(\theta_1)\sin(\theta_1) & b+tr^2\sin^2(\theta_1)\end{bmatrix}.$$

The smallest and the biggest eigenvalues of matrix $X=(A+tB)^t(A+tB)$ can be found

$$\lambda_1(t)=1+2(1-r^2)t+\left(\frac{(r^2-1)\big(1+a-b+r^2+ab\,(r^2-1)-(a+b)r^2\cos(2\theta_1)\big)}{(a+1)(b-1)}\right)t^2+O(t^3),$$

and

$$\lambda_3(t)=b^2+2\,br^2\sin^2(\theta_1)\,t-$$
$$\left(\frac{r^2\Big(b\big(3r^2+b(r^2-4)\big)+a\big(r^2+b(3r^2-4)\big)+(a-b)(b-1)\cos(2\theta_1)\Big)\sin^2(\theta_1)}{2(a+b)(b-1)}\right)t^2+O(t^3).$$

The linear distortion function is defined by

$$\mathcal{H}(A+tB)=\sqrt{\frac{\lambda_3(t)}{\lambda_1(t)}}=b+L(r,\theta_1)\,t+Q(r,\theta_1)\,t^2+O(t^3),$$

where $\lambda_3(t)$ and $\lambda_1(t)$ are the largest and smallest eigenvalues of the matrix X, and $L(r,\theta_1)$ and $Q(r,\theta_1)$ are linear and quadratic coefficients of the linear distortion function, respectively. The first derivative of the linear distortion function must be zero.

$$L(r,\theta_1)=b(r^2-1)+r^2\sin^2(\theta_1)=0,$$

so, we get

$$(r^2-1)=-\frac{r^2\sin^2(\theta_1)}{b}.$$

If we put the above equation in the quadratic coefficient function, it is simplified as

$$Q(r,\theta_1)=\frac{r^2\Big(4(a+1)(a+b)-\big(1+a+a^2+(a-1)b+b^2\big)r^2\Big)}{8(a+1)(b-1)(a+b)}$$
$$+\frac{r^2\Big(-4(a+1)(a+b)\cos(2\theta_1)+\big(1+a+a^2+(a-1)b+b^2\big)r^2\cos(4\theta_1)\Big)}{8(a+1)(b-1)(a+b)}.$$

We first find the critical points, these occur when

$$\frac{\partial Q}{\partial r}(r,\theta_1)=0,\qquad \frac{\partial Q}{\partial\theta_1}(r,\theta_1)=0.$$

The solutions of the partial derivatives above are

$$r=0,\ 1,\qquad \theta_1=0,\pi.$$

So, the critical points only happen on boundaries. If $r = 0$, then $Q = 0$. But if $r = 1$ then the function Q is

$$Q(1, \theta_1) = \frac{\Big(4(a+1)(a+b) - \big(1 + a + a^2 + (a-1)b + b^2\big)\Big)}{8(a+1)(b-1)(a+b)}$$

$$+ \frac{\Big(-4(a+1)(a+b)\cos(2\theta_1) + \big(1 + a + a^2 + (a-1)b + b^2\big)\cos(4\theta_1)\Big)}{8(a+1)(b-1)(a+b)}.$$

In this case the partial derivative of Q with respect to θ_1 is

$$\frac{\partial Q}{\partial \theta_1}(1, \theta_1) = \frac{8(a+1)(a+b)\sin(2\theta_1) - 4\big(1 + a + a^2 + (a-1)b + b^2\big)\sin(4\theta_1)}{8(a+1)(b-1)(a+b)} = 0.$$

Since $1 < a < b$, $0 \le r \le 1$, and $0 \le \theta_1 \le \pi$, the expression $8(a+b)(b-1)(a+1)$ can not be equal to zero. So, the function Q can be defined everywhere in $[0, \pi]$. In this case, there are five critical points on $[0, \pi]$.

$$\theta_1 = 0, \qquad \theta_1 = \frac{\pi}{2}, \qquad \theta_1 = \pi,$$

$$\theta_1 = \arctan\frac{y_1}{x_1}, \qquad \theta_1 = \pi - \arctan\frac{y_1}{x_1},$$

where

$$x_1 = \frac{(a+1)(a+b)}{1 + a + a^2 - b + ab + b^2},$$

and

$$y_1 = \frac{(b-1)\sqrt{1 + 2a + 2a^2 + 2ab + b^2}}{\sqrt{1 + 2a + 3a^2 + a^4 - 2b + 2a^3b + 3b^2 + 3a^2b^2 - 2b^3 + 2ab^3 + b^4}}.$$

The values of function Q_1 at the critical points are

$$Q_1(0) = 0, \qquad Q_1(\frac{\pi}{2}) = \frac{1}{b-1} > 0, \qquad Q_1(\pi) = 0,$$

$$Q_1(\arctan\frac{y_1}{x_1}) = Q_1(\pi - \arctan\frac{y_1}{x_1}) = -\frac{(b-1)^3}{4(a+1)(b+a)\big(1 + a + a^2 + b(a-1) + b^2\big)} < 0.$$

So, the absolute minimum of function Q is negative. **Thus, we have proved that the absolute minimum of the quadratic coefficient function is negative**. At this point, we know that the absolute negative minimum either occurs at δ_1, η_1 or on the boundary. In the next step, we will prove that the quadratic function Q is positive on the boundaries. The values of θ_1 and θ_2 are 0 or π on the boundaries. So,

$$\theta_1 = 0 \text{ and } \theta_2 = \theta, \qquad \theta_1 = \pi \text{ and } \theta_2 = \theta,$$

$$\theta_1 = \theta \;\text{ and }\; \theta_2 = 0, \qquad\qquad \theta_1 = \theta \;\text{ and }\; \theta_2 = \pi.$$

There are four cases here, and we will prove only the first case and leave the others for readers, as they are entirely similar. By the substitution $\theta_1 = 0$ and $\theta_2 = \theta$ at the very beginning, we obtain

$$\mathbf{u}^t = (\sqrt{1-r^2}, r, 0),$$

and

$$\mathbf{v}^t = (\sqrt{1-s^2}, s\cos(\theta), s\sin(\theta)).$$

The matrix $A + tB$ where $B = \mathbf{u} \otimes \mathbf{v}$ is

$$A + tB = \begin{bmatrix} 1 + t\sqrt{1-r^2}\sqrt{1-s^2} & ts\sqrt{1-r^2}\cos(\theta) & ts\sqrt{1-r^2}\sin(\theta) \\ tr\sqrt{1-s^2} & a + trs\cos(\theta) & trs\sin(\theta) \\ 0 & 0 & b \end{bmatrix}.$$

Consider $X = (A+tB)^t(A+tB)$, and $\lambda_3(t)$ and $\lambda_1(t)$ are the largest and smallest eigenvalues of the matrix X. The linear distortion function is

$$\mathcal{H}(A+tB) = \sqrt{\frac{\lambda_3(t)}{\lambda_1(t)}} = b + L(r,s,\theta)\,t + Q(r,s,\theta)\,t^2 + O(t^3),$$

where the linear ceofficient function is

$$L(r,s,\theta) = -b\sqrt{1-r^2}\sqrt{1-s^2},$$

and the quadratic coefficient function is

$$Q(r,s,\theta) = \frac{b}{2}\Bigg(2 + \frac{1}{2}\left(\frac{1}{b^2} - 4\right)s^2 + \frac{r^2(s^2-1)(2a^2-3)}{a^2-1} + \frac{2ars\sqrt{1-r^2}\sqrt{1-s^2}\cos(\theta)}{a^2-1} -$$

$$\frac{s^2(r^2-1)\cos^2(\theta)}{a^2-1} - \frac{s^2\cos(2\theta)}{2b^2} + \frac{\big((b^2-a^2)(b^2+1) - b^2(1-2a^2+b^2)r^2\big)s^2\sin^2(\theta)}{b^2(b^2-1)(b^2-a^2)}\Bigg).$$

We know the first derivative must be zero and this gives us

$$-b\sqrt{1-r^2}\sqrt{1-s^2} = 0.$$

So, $r = 1$ or $s = 1$. We should prove in both cases the function Q is positive on the boundaries.

Case 1: Let $r = 1$, then

$$\begin{aligned} Q(1,s,\theta) &= \frac{b\big(s^2 + a^2(2-3s^2) + 2b^2(s^2-1) + (a^2-1)s^2(2\cos^2(\theta)-1)\big)}{4(a^2-1)(a^2-b^2)} \\ &= \frac{b\big(2s^2(a^2-1)(1-\cos^2(\theta)) + 2(b^2-a^2)(1-s^2)\big)}{4(a^2-1)(b^2-a^2)}. \end{aligned}$$

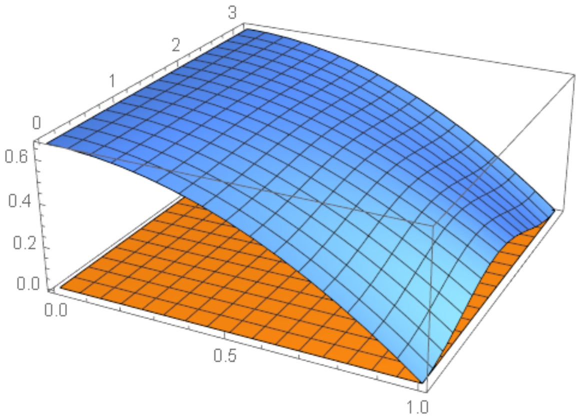

Figure 4.8: If $A = (1, 2, 4)$ is a diagonal matrix, then on the boundaries the function Q has positive values.

Since $1 < a < b$, $0 \leq s \leq 1$, and $0 \leq \theta \leq \pi$, then it is obvious to see

$$Q(1, s, \theta) \geq 0.$$

Case 2: Let $s = 1$, then

$$Q(r, 1, \theta) = \frac{b}{2}\left(\frac{(b^2-1)(b^2-a^2)(1-r^2)\cos^2(\theta) + (a^2-1)\big(r^2(a^2-1) + (2-r^2)(b^2-a^2)\big)}{(a^2-1)(b^2-1)(b^2-a^2)}\right)$$

Since $1 < a < b$, $0 \leq r \leq 1$, and $0 \leq \theta \leq \pi$, then it is obvious to see

$$Q(r, 1, \theta) \geq 0.$$

The other cases can be proved by the same process. **So, the quadratic function is positive on its boundaries**. The quadratic function Q is continuous and smooth on its compact domain, which means it has an absolute minimum. We have proved the absolute minimum is negative and the function Q is positive on its boundaries, thus the absolute minimum happens in an interior point of its domain. This shows us the absolute minimum occurs at δ_1 and η_1. Therefore

$$0 \leq \delta_1 \leq 1, \qquad 0 \leq \eta_1 \leq 1.$$

Now the equation of function Q can be described with respect to δ_1, η_1 and γ. But the equations δ_1, η_1 include the expressions

$$\sqrt{\frac{\beta}{\alpha}} \text{ and } \sqrt{\frac{\alpha}{\beta}},$$

respectively. That says, the ranges of functions α and β are important to correctly define the function Q, and the function Q can be defined by taking correct branches of square roots. The quadratic function is the Hessian matrix of a linear distortion function which is continuous and smooth in its domain. That means the quadratic function is well-defined.

Remark: There are computational issues associated with the choice of square root in our formula, that we resolved to present the graph of the function Q. For instance, if $A = (1, 2, 4)$ is a 3×3 diagonal matrix, then the graph of functions α and β are given below in Figure 4.9.

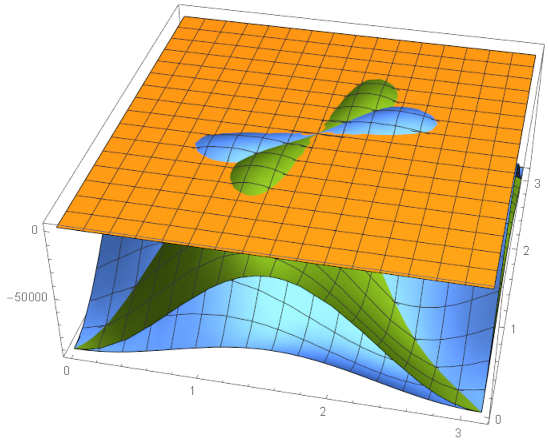

Figure 4.9: If $A = (1, 2, 4)$ is a diagonal matrix, then the functions α and β are the blue and the green graphs, respectively.

This issue came from the extra condition that happened in our algebraic solution for finding the quadratic function Q. This means in the actual quadratic function Q, we do not have the expressions

$$\sqrt{\frac{\beta}{\alpha}} \quad \text{and} \quad \sqrt{\frac{\alpha}{\beta}}.$$

So, the points of $[0, \pi] \times [0, \pi]$ belong to the domain of function Q if α and β have the same sign. By substitution δ_1, η_1 in Equation 4.16, we have

$$Q(\theta_1, \theta_2) = \frac{\alpha\left(\frac{b\left(b-\sqrt{\frac{\beta}{\alpha}}\sqrt{b^2-\mu}\right)}{\mu}\right) + \beta\left(\frac{b\left(b-\sqrt{\frac{\alpha}{\beta}}\sqrt{b^2-\mu}\right)}{\mu}\right) + \gamma}{\xi\mu}.$$

By simplification, the function Q can be written as

$$Q(\theta_1, \theta_2) = \frac{\alpha\left(b^2 - b\sqrt{\frac{\beta}{\alpha}}\sin(\theta_1)\sin(\theta_2)\right) + \beta\left(b^2 - b\sqrt{\frac{\alpha}{\beta}}\sin(\theta_1)\sin(\theta_2)\right) + \gamma\mu}{\xi\mu^2},$$

therefore,

$$Q(\theta_1, \theta_2) = \begin{cases} \dfrac{b^2(\alpha+\beta) - 2b\sqrt{\alpha}\sqrt{\beta}\sin(\theta_1)\sin(\theta_2) + \gamma\mu}{\mu^2\xi}, & \text{if } \alpha \geq 0 \text{ and } \beta \geq 0 \\[2ex] \dfrac{b^2(\alpha+\beta) + 2b\sqrt{\alpha\beta}\sin(\theta_1)\sin(\theta_2) + \gamma\mu}{\mu^2\xi}, & \text{if } \alpha < 0 \text{ and } \beta < 0. \end{cases} \tag{4.23}$$

But for the second case ($\alpha < 0$ and $\beta < 0$, the both case $\alpha < 0$ and $\beta > 0$ or $\alpha > 0$ and $\beta < 0$ are not acceptable), we can write

$$-2b\ \sqrt{\alpha}\sqrt{\beta} = -2b\ \sqrt{-1|\alpha|}\sqrt{-1|\beta|} = -2b\ i^2\ \sqrt{|\alpha|}\sqrt{|\beta|} = +2b\ \sqrt{|\alpha|.|\beta|} = +2b\ \sqrt{\alpha.\beta}\ ,$$

where $|\alpha|$ and $|\beta|$ are positive. Hence, $\alpha < 0$ and $\beta < 0$ conclude $\sqrt{\alpha}\sqrt{\beta}$ is a real number. Thus, in both cases ($\alpha \geq 0$ and $\beta \geq 0$) or ($\alpha < 0$ and $\beta < 0$), the function Q will be as follows

$$Q(\theta_1,\theta_2) = \frac{b^2(\alpha(\theta_1,\theta_2)+\beta(\theta_1,\theta_2)) - 2b\ \sqrt{\alpha(\theta_1,\theta_2)}\sqrt{\beta(\theta_1,\theta_2)}\sin(\theta_1)\sin(\theta_2) + \gamma(\theta_1,\theta_2)\mu(\theta_1,\theta_2)}{\mu^2(\theta_1,\theta_2)\xi} \tag{4.24}$$

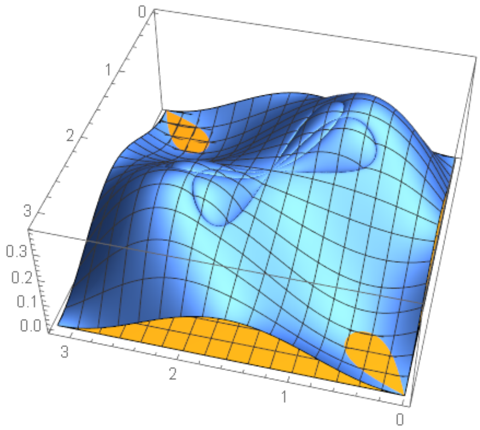

Figure 4.10: The function Q if $A = (1, 2, 4)$ is a diagonal matrix.

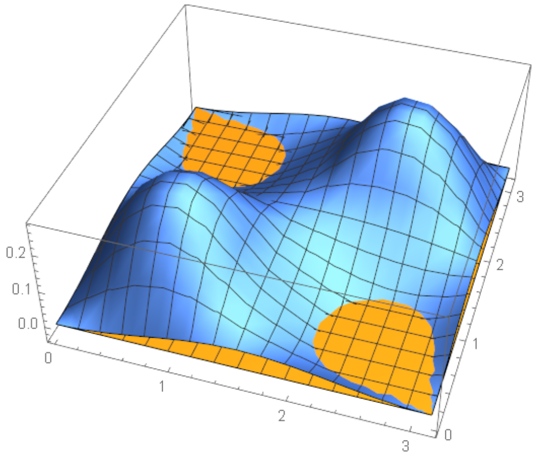

Figure 4.11: The function Q if $A = (1, 2, 10)$ is a diagonal matrix.

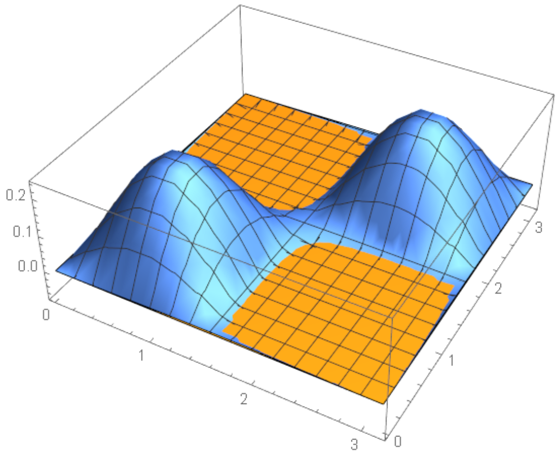

Figure 4.12: The function Q if $A = (1, 2, 105)$ is a diagonal matrix.

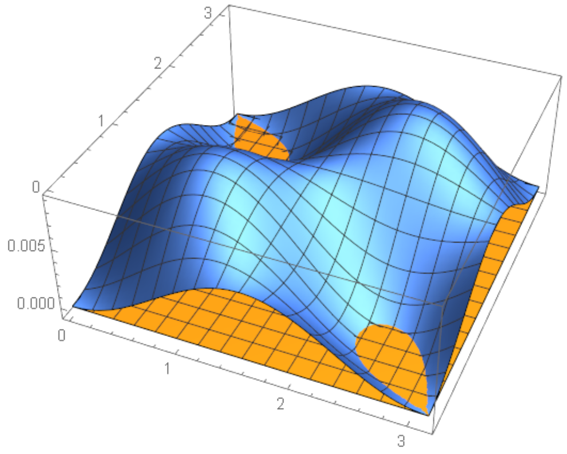

Figure 4.13: The function Q if $A = (1, 99, 154)$ is a diagonal matrix.

The function Q has various properties, for instance:

1. $Q(\theta_1, \theta_2) = Q(\theta_2, \theta_1)$,
2. $Q(\theta_1, \theta_2) = Q(-\theta_1, -\theta_2)$,
3. $Q(\theta_1, \theta_2) = Q(\pi - \theta_1, \pi - \theta_2)$,
4. $Q(\theta_1, \theta_2) = Q(2\pi - \theta_2, 2\pi - \theta_1)$,
5. $Q(\theta_1, \theta_2) = Q(\pi - \theta_2, \pi - \theta_1)$.

The proof of some of the above properties is easy, so we only prove the third and fifth properties. At first, we prove symmetry for functions α, β, γ and μ. Since $\cos(\pi - \theta) = -\cos(\theta)$, $\cos(2\pi - \theta) = \cos(\theta)$, $\sin(2\pi - \theta) = -\sin(\theta)$ and $\sin(\pi - \theta) = \sin(\theta)$, then

$$
\begin{aligned}
\alpha(\pi-\theta_1,\pi-\theta_2) =& b\Big(a^2-2a^4+(7-21a^2+8a^4)b^2+(7+8a^2)b^4-8b^6- \\
& 8(a^2-1)b^2\cos(2(\pi-\theta_1))\Big(-a^2+b^2+(a^2-1)\cos(2(\pi-\theta_2))\Big)- \\
& (a^2-b^2)\Big[\Big(1+b^2-8b^4+a^2(-2+8b^2)\Big)\cos(2(\pi-\theta_2))+ \\
& 2\Big(1-2a^2+b^2\Big)\cos(4(\pi-\theta_1))\sin^2(\pi-\theta_2)\Big]+ \\
& 8ab\Big(b^2-1\Big)^2\sin(2(\pi-\theta_1))\sin(2(\pi-\theta_2))\Big) \\
& = b\Big(a^2-2a^4+(7-21a^2+8a^4)b^2+(7+8a^2)b^4-8b^6-8(a^2-1) \\
& b^2\cos(2\theta_1)\Big(-a^2+b^2+(a^2-1)\cos(2\theta_2)\Big)-(a^2-b^2)\Big[\Big(1+b^2-8b^4 \\
& +a^2(-2+8b^2)\Big)\cos(2\theta_2)+2\Big(1-2a^2+b^2\Big)\cos(4\theta_1)(-\sin(\theta_2))^2\Big]+ \\
& 8ab\Big(b^2-1\Big)^2(-\sin(2\theta_1))(-\sin(2\theta_2))\Big) \\
& = b\Big(a^2-2a^4+(7-21a^2+8a^4)b^2+(7+8a^2)b^4-8b^6-8(a^2-1) \\
& b^2\cos(2\theta_1)\Big(-a^2+b^2+(a^2-1)\cos(2\theta_2)\Big)-(a^2-b^2)\Big[\Big(1+b^2-8b^4 \\
& +a^2(-2+8b^2)\Big)\cos(2\theta_2)+2\Big(1-2a^2+b^2\Big)\cos(4\theta_1)\sin^2(\theta_2)\Big]+ \\
& 8ab\Big(b^2-1\Big)^2\sin(2\theta_1)\sin(2\theta_2)\Big) \\
& = \alpha(\theta_1,\theta_2).
\end{aligned}
\tag{4.25}
$$

With the same method, we can prove

$$
\begin{aligned}
\beta(\pi-\theta_1,\pi-\theta_2) &= \beta(\theta_1,\theta_2), \\
\gamma(\pi-\theta_1,\pi-\theta_2) &= \gamma(\theta_1,\theta_2).
\end{aligned}
\tag{4.26}
$$

Also,

$$\begin{aligned}\mu(\pi-\theta_1,\pi-\theta_2) &= \big(b^2-\sin^2(\pi-\theta_1)\sin^2(\pi-\theta_2)\big)\\ &= \big(b^2-(-\sin(\theta_1))^2(-\sin(\theta_2))^2\big)\\ &= \big(b^2-\sin^2(\theta_1)\sin^2(\theta_2)\big)=\mu(\theta_1,\theta_2)\end{aligned} \tag{4.27}$$

Now by Equations 4.25, 4.26 and 4.27, we can prove the function Q is symmetric.

$$\begin{aligned}Q(\pi-\theta_1,\pi-\theta_2) &= \frac{b^2(\alpha(\pi-\theta_1,\pi-\theta_2)+\beta(\pi-\theta_1,\pi-\theta_2))}{\mu^2(\pi-\theta_1,\pi-\theta_2)\xi}\\ &- \frac{2b\sqrt{\alpha(\pi-\theta_1,\pi-\theta_2)}\sqrt{\beta(\pi-\theta_1,\pi-\theta_2)}\sin(\pi-\theta_1)\sin(\pi-\theta_2)}{\mu^2(\pi-\theta_1,\pi-\theta_2)\xi}\\ &+ \frac{\gamma(\pi-\theta_1,\pi-\theta_2)\mu(\pi-\theta_1,\pi-\theta_2)}{\mu^2(\pi-\theta_1,\pi-\theta_2)\xi}\\ &= \frac{b^2(\alpha(\theta_1,\theta_2)+\beta(\theta_1,\theta_2))-2b\sqrt{\alpha(\theta_1,\theta_2)}\sqrt{\beta(\theta_1,\theta_2)}\sin(\theta_1)\sin(\theta_2)+\gamma(\theta_1,\theta_2)\mu(\theta_1,\theta_2)}{\mu^2(\theta_1,\theta_2)\xi}\\ &= Q(\theta_1,\theta_2).\end{aligned} \tag{4.28}$$

Therefore, we showed the function Q is π-periodic. The proof of the fifth property is not easy because the functions α and β are not symmetric with respect to the line $\theta_1+\theta_2=\pi$, and the functions γ and μ are symmetric with respect to the same line. Although the α and β functions are not symmetric with respect to the line $\theta_1+\theta_2=\pi$, we prove that the function Q is symmetric with respect to that line. It can be calculated with some effort that,

$$\begin{aligned}&\alpha(\theta_1,\theta_2)-\alpha(\pi-\theta_2,\pi-\theta_1)=-[\beta(\theta_1,\theta_2)-\beta(\pi-\theta_2,\pi-\theta_1)]=\\ &b(a^2-b^2)(-1+2a^2-b^2)\big(\cos(2\theta_1)-\cos(2\theta_2)\big)\big(-1+4b^2+\cos(2\theta_1)+2\cos(2\theta_2)\sin^2(\theta_1)\big)\end{aligned} \tag{4.29}$$

and,

$$\gamma(\theta_1,\theta_2)-\gamma(\pi-\theta_2,\pi-\theta_1)=\mu(\theta_1,\theta_2)-\mu(\pi-\theta_2,\pi-\theta_1)=0. \tag{4.30}$$

Thus,

$$\begin{aligned}&Q(\theta_1,\theta_2)-Q(\pi-\theta_2,\pi-\theta_1)=\\ &=\frac{b^2(\alpha(\theta_1,\theta_2)+\beta(\theta_1,\theta_2))-2b\sqrt{\alpha(\theta_1,\theta_2)}\sqrt{\beta(\theta_1,\theta_2)}\sin(\theta_1)\sin(\theta_2)+\gamma(\theta_1,\theta_2)\mu(\theta_1,\theta_2)}{\mu^2(\theta_1,\theta_2)\xi}-\\ &\Bigg[\frac{b^2(\alpha(\pi-\theta_2,\pi-\theta_1)+\beta(\pi-\theta_2,\pi-\theta_1))}{\mu^2(\pi-\theta_2,\pi-\theta_1)\xi}-\\ &\frac{2b\sqrt{\alpha(\pi-\theta_2,\pi-\theta_1)}\sqrt{\beta(\pi-\theta_2,\pi-\theta_1)}\sin(\pi-\theta_2)\sin(\pi-\theta_1)}{\mu^2(\pi-\theta_2,\pi-\theta_1)\xi}+\\ &\frac{\gamma(\pi-\theta_2,\pi-\theta_1)\mu(\pi-\theta_2,\pi-\theta_1)}{\mu^2(\pi-\theta_2,\pi-\theta_1)\xi}\Bigg].\end{aligned} \tag{4.31}$$

Since $\sin(\pi-\theta_2)\sin(\pi-\theta_1)=\sin(\theta_1)\sin(\theta_2)$, $\gamma(\theta_1,\theta_2)=\gamma(\pi-\theta_2,\pi-\theta_1)$ and $\mu(\theta_1,\theta_2)=\mu(\pi-\theta_2,\pi-\theta_1)$ then the equation (4.28) can be simplified

$$
\begin{aligned}
&Q(\theta_1,\theta_2) - Q(\pi-\theta_2,\pi-\theta_1) = \\
&\frac{\Big(b^2(\alpha(\theta_1,\theta_2)+\beta(\theta_1,\theta_2))\Big) - \Big(b^2(\alpha(\pi-\theta_2,\pi-\theta_1)+\beta(\pi-\theta_2,\pi-\theta_1))\Big)}{\mu^2(\theta_1,\theta_2)\xi} - \\
&\frac{\Big(2b\sqrt{\alpha(\theta_1,\theta_2)}\sqrt{\beta(\theta_1,\theta_2)}\sin(\theta_1)\sin(\theta_2)\Big) + \Big(-2b\sqrt{\alpha(\pi-\theta_2,\pi-\theta_1)}\sqrt{\beta(\pi-\theta_2,\pi-\theta_1)}\sin(\theta_2)\sin(\theta_1)\Big)}{\mu^2(\theta_1,\theta_2)\xi} \\
&+\frac{\Big(\gamma(\theta_1,\theta_2)\mu(\theta_1,\theta_2)\Big) - \Big(\gamma(\pi-\theta_2,\pi-\theta_1)\mu(\pi-\theta_2,\pi-\theta_1)\Big)}{\mu^2(\pi-\theta_2,\pi-\theta_1)\xi}
\end{aligned}
\tag{4.32}
$$

The third fraction is zero. So,

$$
\begin{aligned}
&Q(\theta_1,\theta_2) - Q(\pi-\theta_2,\pi-\theta_1) = \\
&\frac{b^2\Big[\Big(\alpha(\theta_1,\theta_2)-\alpha(\pi-\theta_2,\pi-\theta_1)\Big)+\Big(\beta(\theta_1,\theta_2)-\beta(\pi-\theta_2,\pi-\theta_1)\Big)\Big]}{\mu^2(\theta_1,\theta_2)\xi} - \\
&\frac{2b\sin(\theta_1)\sin(\theta_2)\Big(\sqrt{\alpha(\theta_1,\theta_2)}\sqrt{\beta(\theta_1,\theta_2)}-\sqrt{\alpha(\pi-\theta_2,\pi-\theta_1)}\sqrt{\beta(\pi-\theta_2,\pi-\theta_1)}\Big)}{\mu^2(\theta_1,\theta_2)\xi}
\end{aligned}
\tag{4.33}
$$

The first fraction in Equation 4.33 is also zero by Equation 4.29. For the second fraction, we multiply and divide the conjugate of the numerator to the fraction, then

$$
\frac{2b\sin(\theta_1)\sin(\theta_2)\Big(\alpha(\theta_1,\theta_2)\beta(\theta_1,\theta_2)-\alpha(\pi-\theta_2,\pi-\theta_1)\beta(\pi-\theta_2,\pi-\theta_1)\Big)}{\mu^2(\theta_1,\theta_2)\ \xi\ \Big(\sqrt{\alpha(\theta_1,\theta_2)}\sqrt{\beta(\theta_1,\theta_2)}+\sqrt{\alpha(\pi-\theta_2,\pi-\theta_1)}\sqrt{\beta(\pi-\theta_2,\pi-\theta_1)}\Big)}
$$

But it is easy to check that the function $\alpha(\theta_1,\theta_2)\beta(\theta_1,\theta_2)$ is symmetric with respect to the line $\theta_1+\theta_2=\pi$. That means

$$
\Big(\alpha(\theta_1,\theta_2)\beta(\theta_1,\theta_2)-\alpha(\pi-\theta_2,\pi-\theta_1)\beta(\pi-\theta_2,\pi-\theta_1)\Big)=0.
$$

The equation above tells us the second fraction in (4.33) is equal to zero and we get

$$
Q(\theta_1,\theta_2)-Q(\pi-\theta_2,\pi-\theta_1)=0,
$$

so,

$$
Q(\theta_1,\theta_2)=Q(\pi-\theta_2,\pi-\theta_1).
$$

Therefore, the Function Q is symmetric with respect to the line $\theta_1+\theta_2=\pi$. These symmetries suggest that the absolute minimum value is taken along the line $\theta_2=\pi-\theta_1$. The function Q is positive on the boundary because,

$$
Q(0,\theta_2)=Q(\pi,\theta_2)=\frac{b\sin^2(\theta_2)}{2(b^2-a^2)}, \qquad Q(\theta_1,0)=Q(\theta_1,\pi)=\frac{b\sin^2(\theta_1)}{2(b^2-a^2)}.
$$

That means the absolute negative minimum must occur in interior points of the domain $[0, \pi] \times [0, \pi]$. But if we look at the symmetry line $\theta_1 + \theta_2 = \pi$, then there are two negative minima of the function Q. Now, we must show the function extremes are located on the symmetric lines $\theta_1 = \theta_2$ and $\theta_1 + \theta_2 = \pi$ (see figures (4.10) to (4.13)).
To find the extremes of the function Q, the two following equations of the partial derivatives of the function Q with respect to the two variables θ_1 and θ_2 must be solved.

$$\frac{\partial Q}{\partial \theta_1}(\theta_1, \theta_2) = 0, \qquad \frac{\partial Q}{\partial \theta_2}(\theta_1, \theta_2) = 0. \tag{4.34}$$

These two equations are very difficult to solve analytically because they are of eighth degree and trigonometric. Also the equations are very long and can not be simplified. Although the equations can not be solved analytically, fortunately, it can be proved that the extremes of the function Q are on the lines of symmetry. For this purpose, we must prove that the contour maps of the partial derivatives of Q intersect each other on lines of symmetry. See the figures below.

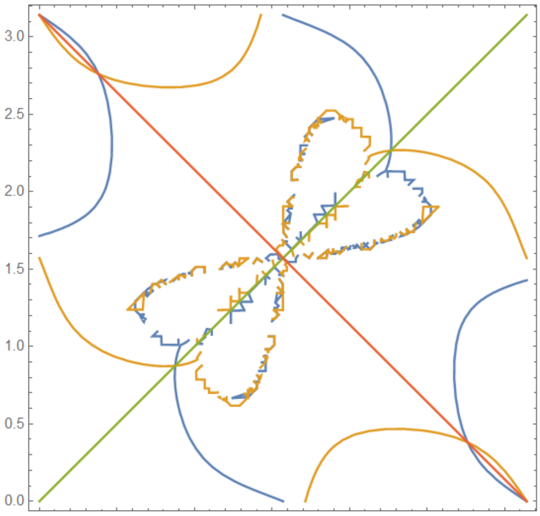

Figure 4.14: The contour maps of $\frac{\partial Q}{\partial \theta_1}(\theta_1, \theta_2)$ and $\frac{\partial Q}{\partial \theta_2}(\theta_1, \theta_2)$, if $A = (1, 2, 4)$ is a diagonal matrix.

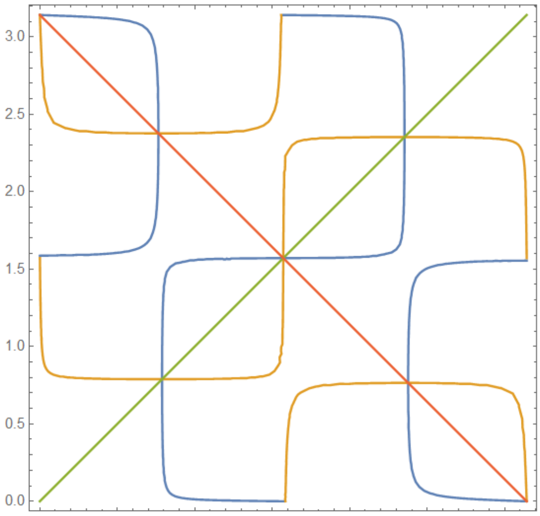

Figure 4.15: The contour maps of $\frac{\partial Q}{\partial \theta_1}(\theta_1, \theta_2)$ and $\frac{\partial Q}{\partial \theta_2}(\theta_1, \theta_2)$, if $A = (1, 2, 86)$ is a diagonal matrix.

At first, the partial derivatives of Q are symmetric with respect to the line $\theta_1 = \theta_2$. That means

$$\frac{\partial Q}{\partial \theta_1}(\theta_1, \theta_2) - \frac{\partial Q}{\partial \theta_2}(\theta_2, \theta_1) = 0,$$

or,

$$\frac{\partial Q}{\partial \theta_1}(\theta_1, \theta_2) = \frac{\partial Q}{\partial \theta_2}(\theta_2, \theta_1). \tag{4.35}$$

The point $(\frac{\pi}{2}, \frac{\pi}{2})$ is the intersection point of the lines of symmetry $\theta_1 = \theta_2$ and $\theta_1 + \theta_2 = \pi$. If the origin $(0, 0)$ is transferred to the point $(\frac{\pi}{2}, \frac{\pi}{2})$, then the two partial derivative of the function Q are symmetric with respect to the line $\theta_1 = \theta_2$. That means one of these two functions can be considered as f and the other one is f^{-1}. As a consequence, the solutions of the simultaneous equations on $[0,\pi]\times[0,\pi]$,

$$\begin{cases} \dfrac{\partial Q}{\partial \theta_1}(\theta_1,\theta_2)=0 & \text{if } \theta_1 \text{ and } \theta_2 \in [0,\pi]\times[0,\pi] \\ \\ \dfrac{\partial Q}{\partial \theta_2}(\theta_1,\theta_2)=0 & \text{if } \theta_1 \text{ and } \theta_2 \in [0,\pi]\times[0,\pi], \end{cases} \tag{4.36}$$

are located on the lines of symmetry $\theta_1=\theta_2$ and $\theta_1+\theta_2=\pi$, and the proof of this part is completed. Hence, the critical points of Q lie on lines of symmetry, and in this case, we need to find the cross-sections of Q concerning on lines $\theta_1=\theta_2$ and $\theta_1+\theta_2=\pi$. The following graphs show these two cross-sections of function Q.

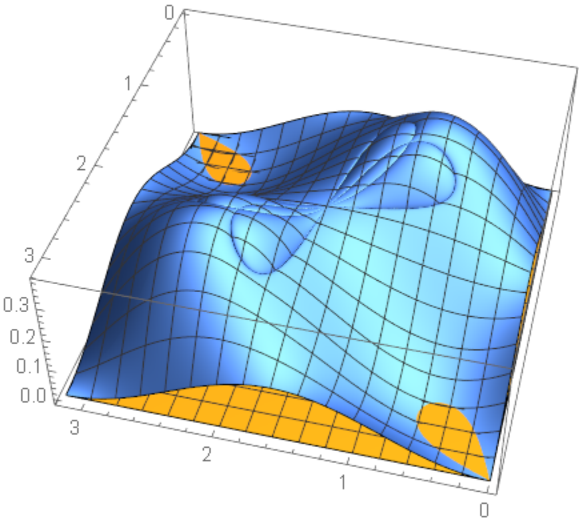

Figure 4.16: If $A=(1,2,4)$ is a diagonal matrix, then the function Q has two absolute negative minimums at interior points of the domain.

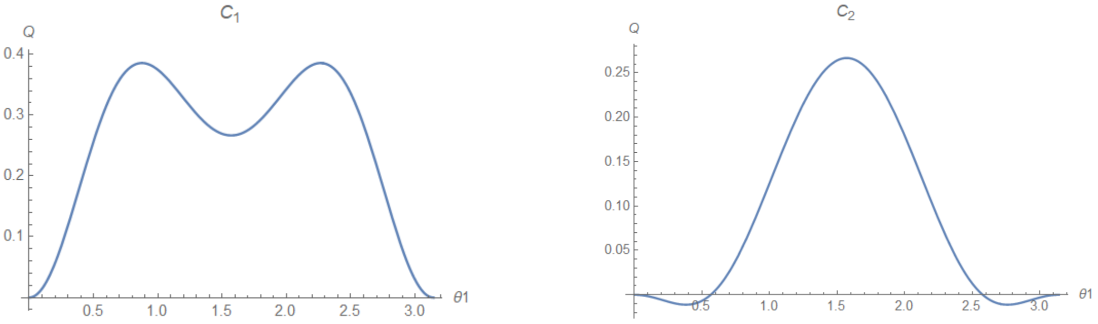

Figure 4.17: The cross-section of Q with respect to line $\theta_1=\theta_2$.

Figure 4.18: The cross-section of Q with respect to line $\theta_1+\theta_2=\pi$.

To find the absolute maximums and minimums of Q, we only need to consider these two cross-sections of Q. So, we have two cases.

Case 1: $(\theta_1 = \theta_2)$

In this case the functions α, β, γ and μ are one-variable functions and their equations can be written as follows:

$$\begin{aligned}\alpha(\theta_1) =& b\Big(a^2 - 2a^4 + \big(7 - 21a^2 + 8a^4\big)b^2 + \big(7 + 8a^2\big)b^4 - 8b^6 + \cos(2\theta_1)(a^2 - b^2)\\ & (-1 + 2a^2 - 9b^2 + 8b^4) - 8(a^2-1)^2 b^2 \cos^2(2\theta_1) + 2(a^2 - b^2)(2a^2 - b^2 - 1)\\ & \cos(4\theta_1)\sin^2(\theta_1) + 8ab(b^2-1)^2 \sin^2(2\theta_1)\Big),\end{aligned} \tag{4.37}$$

$$\begin{aligned}\beta(\theta_1) =& b\Big(a^2 - 2a^4 + \big(7 - 21a^2 + 8a^4\big)b^2 + \big(7 + 8a^2\big)b^4 - 8b^6 + \cos(2\theta_1)(a^2 - b^2)\\ & (-1 + 2a^2 - 9b^2 + 8b^4) - 8(a^2-1)^2 b^2 \cos^2(2\theta_1) + 2(a^2 - b^2)(2a^2 - b^2 - 1)\\ & \cos(4\theta_1)\sin^2(\theta_1) + 8ab(b^2-1)^2 \sin^2(2\theta_1)\Big),\end{aligned} \tag{4.38}$$

$$\begin{aligned}\gamma(\theta_1) =& -8b\Big(a^2(1 - 5b^2) - 2b^4(b^2 - 2) + a^4(3b^2 - 1) - 2(b^2 - 1)(b^2 - a^2)\\ & (1 - a^2 + b^2)\cos(2\theta_1) - (a^2 - 1)\big(a^2 + (a^2 - 2)b^2\big)\cos^2(2\theta_1) + \\ & ab(b^2 - 1)^2 \sin^2(2\theta_1)\Big).\end{aligned} \tag{4.39}$$

Also, for the function μ and constant ξ, we have

$$\mu(\theta_1) = (b^2 - \sin^4(\theta_1)) \quad , \quad \xi = 32(a^2-1)(b^2-1)(b^2-a^2). \tag{4.40}$$

These show that if $\theta_1 = \theta_2$ then $\alpha(\theta_1) = \beta(\theta_1)$. The function Q converts to a one-variable real-valued function and finding the maximum and minimum points is straight forward.

$$Q_1(\theta_1) = \frac{b^2(\alpha(\theta_1) + \beta(\theta_1)) - 2b\sqrt{\alpha(\theta_1)}\sqrt{\beta(\theta_1)}\sin^2(\theta_1) + \gamma(\theta_1)\mu(\theta_1)}{\mu^2(\theta_1)\xi} \tag{4.41}$$

where $\theta_1 \in [0, \pi]$ and Q_1 is the cross section of function Q with respect to $\theta_1 = \theta_2$. The first derivative of Q_1 with respect to θ_1 can be simplified as

$$\begin{aligned}\frac{\mathrm{dQ_1}}{\mathrm{d}\theta_1}(\theta_1) =& 2b^2\Bigg[\frac{\big(2 + 6b^2 + 4a^2(b+1) - 4a(b+1)^2\big)\sin(2\theta_1)}{(a-b)(b-1)(a-1)\big(1 + 2b - \cos(2\theta_1)\big)^3} + \\ & \frac{\big(-1 - 2a^2(b+1) + 2a(b+1)^2 + b(-2 + b - 2b^2)\big)\sin(4\theta_1)}{(a-b)(b-1)(a-1)\big(1 + 2b - \cos(2\theta_1)\big)^3}\Bigg] = 0\end{aligned} \tag{4.42}$$

Since $1 < a < b$ and $0 \leq \theta_1 \leq \pi$, the expression $(a-b)(b-1)(a-1)\big(1+2b-\cos(2\theta_1)\big)^3$ can not be equal to zero. So, the function Q_1 can be defined everywhere in $[0,\pi]$. In this case, there are five critical points on $[0,\pi]$.

$$\theta_1 = 0, \qquad \theta_1 = \frac{\pi}{2}, \qquad \theta_1 = \pi,$$

$$\theta_1 = \arctan\frac{y_1}{x_1}, \qquad \theta_1 = \pi - \arctan\frac{y_1}{x_1},$$

where

$$x_1 = \frac{(b-1)\sqrt{b}}{\sqrt{1-2a+2a^2+2b-4ab+2a^2b-b^2-2ab^2+2b^3}},$$

and

$$y_1 = \frac{\sqrt{b+1}\sqrt{1-2a+2a^2-2ab+b^2}}{\sqrt{1-2a+2a^2+2b-4ab+2a^2b-b^2-2ab^2+2b^3}}.$$

The values of function Q_1 at the critical points are

$$Q_1(0) = 0, \qquad Q_1(\frac{\pi}{2}) = \frac{b}{b^2-1} > 0, \qquad Q_1(\pi) = 0,$$

$$Q_1(\arctan\frac{y_1}{x_1}) = Q_1(\pi - \arctan\frac{y_1}{x_1}) = \frac{b(b-1)^3}{4(a-1)(b-a)(b+1)\big(1+a^2+b(b-1)-a(b+1)\big)}.$$

The end-points $\theta_1 = 0$ and $\theta_1 = \pi$ have value zero and $\theta_1 = \frac{\pi}{2}$ has positive value. If we apply the second derivative test for $\theta_1 = \frac{\pi}{2}$, then

$$\frac{\partial^2 Q_1}{\partial\theta_1^2}(\frac{\pi}{2}) = \frac{2b^2\big((a-1)^2+(b-a)^2\big)}{(a-1)(b-1)(b-a)(b+1)^2} > 0,$$

because $1 < a < b$. That means the point $\theta_1 = \frac{\pi}{2}$ is a local minimum of Q_1. So, the function Q_1 has boundary values zero at $\theta_1 = 0$ and $\theta_1 = \pi$ and two local and absolute maximums at

$$\theta_1 = \arctan\frac{y_1}{x_1}, \qquad \theta_1 = \pi - \arctan\frac{y_1}{x_1},$$

and finally, two absolute minimums at the end-points.
Since the function Q is symmetric with respect to the line $\theta_2 = \pi - \theta_1$, then the function Q_1 is symmetric with respect to the vertical line $\theta_1 = \frac{\pi}{2}$. Hence, these two absolute maximums have the same value.

$$Q_1(\arctan\frac{y_1}{x_1}) = Q_1(\pi - \arctan\frac{y_1}{x_1}) = \text{Absolute maximum value.}$$

This completes the proof of case 1.

Case 2: $(\theta_2 = \pi - \theta_1)$

In this case, the functions α, β, γ and μ are

$$\begin{aligned}\alpha(\theta_1) =& b\Big(a^2 - 2a^4 + \big(7 - 21a^2 + 8a^4\big)b^2 + \big(7 + 8a^2\big)b^4 - 8b^6 + \cos(2\theta_1)(a^2 - b^2)\\ & (-1 + 2a^2 - 9b^2 + 8b^4) - 8(a^2 - 1)^2 b^2 \cos^2(2\theta_1) + 2(a^2 - b^2)(2a^2 - b^2 - 1)\\ & \cos(4\theta_1)\sin^2(\theta_1) - 8ab(b^2 - 1)^2 \sin^2(2\theta_1)\Big),\end{aligned} \tag{4.43}$$

$$\begin{aligned}\beta(\theta_1) =& b\Big(a^2 - 2a^4 + \big(7 - 21a^2 + 8a^4\big)b^2 + \big(7 + 8a^2\big)b^4 - 8b^6 + \cos(2\theta_1)(a^2 - b^2)\\ & (-1 + 2a^2 - 9b^2 + 8b^4) - 8(a^2 - 1)^2 b^2 \cos^2(2\theta_1) + 2(a^2 - b^2)(2a^2 - b^2 - 1)\\ & \cos(4\theta_1)\sin^2(\theta_1) - 8ab(b^2 - 1)^2 \sin^2(2\theta_1)\Big),\end{aligned} \tag{4.44}$$

$$\begin{aligned}\gamma(\theta_1) =& -8b\Big(a^2(1 - 5b^2) - 2b^4(b^2 - 2) + a^4(3b^2 - 1) - 2(b^2 - 1)(b^2 - a^2)\\ & (1 - a^2 + b^2)\cos(2\theta_1) - (a^2 - 1)(a^2 + (a^2 - 2)b^2)\cos^2(2\theta_1) - \\ & ab(b^2 - 1)^2 \sin^2(2\theta_1)\Big),\end{aligned} \tag{4.45}$$

Also, for the function μ and constant ξ, we have

$$\mu(\theta_1) = (b^2 - \sin^4(\theta_1)), \qquad \xi = 32(a^2 - 1)(b^2 - 1)(b^2 - a^2). \tag{4.46}$$

These show that if $\theta_2 = \pi - \theta_1$ then $\alpha(\theta_1) = \beta(\theta_1)$. The function Q converts to a one-variable real-valued function and we can find the maximum and minimum points.

$$Q_2(\theta_1) = \frac{b^2(\alpha(\theta_1) + \beta(\theta_1)) - 2b\,\sqrt{\alpha(\theta_1)}\sqrt{\beta(\theta_1)}\sin^2(\theta_1) + \gamma(\theta_1)\mu(\theta_1)}{\mu^2(\theta_1)\xi} \tag{4.47}$$

where $\theta_1 \in [0, \pi]$ and Q_2 is the cross-section of function Q with respect to $\theta_2 = \pi - \theta_1$. The first derivative of Q_2 with respect to θ_1 can be simplified as follows

$$\begin{aligned}\frac{\mathrm{d}Q_2}{\mathrm{d}\theta_1}(\theta_1) =& \frac{-b^2\big(2b - 1 + \cos(2\theta_1)\big)}{8(a+1)(a+b)(b-1)(b - \sin^2(\theta_1))^2(b + \sin^2(\theta_1))^3}\Big[-1 - 3b^2 - 2a^2(b+1) - \\ & 2a(b+1)^2 + \big(1 + 2a^2(b+1) + 2a(b+1)^2 + b(2 - b + 2b^2)\big)\cos(2\theta_1)\Big]\sin(2\theta_1) = 0\end{aligned} \tag{4.48}$$

Since $1 < a < b$ and $0 \leq \theta_1 \leq \pi$, the expression $8(a+1)(a+b)(b-1)(b-\sin^2(\theta_1))^2(b+\sin^2(\theta_1))^3$ can not be equal to zero. So, the function Q_2 can be defined everywhere in $[0, \pi]$. In this case, there are five critical points on $[0, \pi]$.

$$\theta_1 = 0, \qquad \theta_1 = \frac{\pi}{2}, \qquad \theta_1 = \pi,$$

$$\theta_1 = \arctan\frac{y_2}{x_2}, \qquad \theta_1 = \pi - \arctan\frac{y_2}{x_2}$$

where

$$x_2 = \frac{(b-1)\sqrt{b}}{\sqrt{1+2a+2a^2+2b = 4ab+2a^2b-b^2+2ab^2+2b^3}}, \tag{4.49}$$

and

$$y_2 = \frac{\sqrt{b+1}\sqrt{1+2a+2a^2+2ab+b^2}}{\sqrt{1+2a+2a^2+2b+4ab+2a^2b-b^2+2ab^2+2b^3}}. \tag{4.50}$$

The values of function Q_2 at the critical points are

$$Q_2(0) = 0, \qquad Q_2(\frac{\pi}{2}) = \frac{b}{b^2-1} > 0, \qquad Q_2(\pi) = 0,$$

$$Q_2(\arctan\frac{y_2}{x_2}) = Q_2(\pi - \arctan\frac{y_2}{x_2}) = \frac{-b(b-1)^3}{4(a+1)(b+1)(a+b)(1+a+a^2+ab-b+b^2)}.$$

The end-points $\theta_1 = 0$ and $\theta_1 = \pi$ have value zero and $\theta_1 = \frac{\pi}{2}$ has the positive value. If we apply the second derivative test for $\theta_1 = \frac{\pi}{2}$, then

$$\frac{\mathrm{d}^2 Q_2}{\mathrm{d}\theta_1^2}(\frac{\pi}{2}) = -\frac{2b^2\big((a+b)^2+(a+1)^2\big)}{(a+1)(b-1)(a+b)(b+1)^2} < 0,$$

because $1 < a < b$. That means the point $\theta_1 = \frac{\pi}{2}$ is a local maximum of Q_2. So, the function Q_1 has boundary values zero at $\theta_1 = 0$ and $\theta_1 = \pi$ and two local and absolute minimums at

$$\theta_1 = \arctan\frac{y_2}{x_2}, \qquad \theta_1 = \pi - \arctan\frac{y_2}{x_2}.$$

Since the function Q is symmetric with respect to the line $\theta_1 = \theta_2$ then the function Q_2 is symmetric with respect to the vertical line $\theta_1 = \frac{\pi}{2}$.Also, the function Q_2 is a continuous and smooth function, therefore the values of the two minimums are equal and negative. Thus,

$$Q_2(\arctan\frac{y_2}{x_2}) = Q_2(\pi - \arctan\frac{y_2}{x_2}) < 0, \qquad \text{absolute negative minimun values.}$$

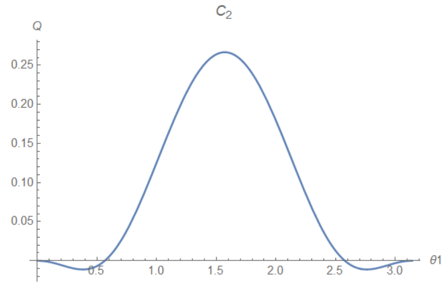

Figure 4.19: The cross section of Q with respect to the line $\theta_2 = \pi - \theta_1$.

As a result, we have found the best rank-one direction that which maximises the negative of the second derivative. Now, we are able to determine vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^3$ with $\|\mathbf{u}\| = \|\mathbf{v}\| = 1$. Recall

$$\mathbf{u}^t = (\sqrt{1-r^2}, r\cos(\theta_1), r\sin(\theta_1)),$$

and

$$\mathbf{v}^t = (\sqrt{1-s^2}, s\cos(\theta_2), s\sin(\theta_2)).$$

We know that in the best direction, $\theta_2 = \pi - \theta_1$ and $\alpha(\theta_1) = \beta(\theta_1)$. We chose $\delta = r^2$ and $\eta = s^2$. Then

$$\delta_1 = \frac{b\Big(b - \sqrt{\frac{\beta}{\alpha}}\sqrt{b^2-\mu}\Big)}{\mu}, \qquad \eta_1 = \frac{b\Big(b - \sqrt{\frac{\alpha}{\beta}}\sqrt{b^2-\mu}\Big)}{\mu}.$$

Since $\alpha(\theta_1) = \beta(\theta_1)$, then $r^2 = s^2$, or $r = \pm s$. Without loss of generality, assume that $r = s$, so,

$$r = s = \frac{b\Big(b - \sqrt{\frac{\beta}{\alpha}}\sqrt{b^2-\mu}\Big)}{\mu} = \frac{\big(b^2 - b\sin^2(\theta_1)\big)}{b^2 - \sin^4(\theta_1)} = \frac{b\big(b - \sin^2(\theta_1)\big)}{(b-\sin^2(\theta_1))(b+\sin^2(\theta_1))}.$$

Hence,

$$r = s = \frac{b}{(b+\sin^2(\theta_1))}. \tag{4.51}$$

Now, we have best rank-one direction

$$\mathbf{u}^t = \Big(\frac{\sqrt{2b\sin(\theta_1)+\sin^2(\theta_1)}}{(b+\sin^2(\theta_1))}, \frac{b\cos(\theta_1)}{(b+\sin^2(\theta_1))}, \frac{b\sin(\theta_1)}{(b+\sin^2(\theta_1))}\Big), \tag{4.52}$$

and,

$$\mathbf{v}^t = \Big(\frac{\sqrt{2b\sin(\theta_1)+\sin^2(\theta_1)}}{(b+\sin^2(\theta_1))}, -\frac{b\cos(\theta_1)}{(b+\sin^2(\theta_1))}, \frac{b\sin(\theta_1)}{(b+\sin^2(\theta_1))}\Big), \tag{4.53}$$

where

$$\theta_1 = \arctan\frac{y_2}{x_2}, \qquad\qquad \theta_1 = \pi - \arctan\frac{y_2}{x_2},$$

and x_2 and y_2 satisfy Equations (4.49) and (4.50). The vectors $\mathbf{u}$ and $\mathbf{v}$ can be written with respect to a and b as:

$$\mathbf{u} = \big(u_1, u_2, u_3\big),$$

$$\mathbf{v} = \big(u_1, -u_2, u_3\big),$$

where

$$u_1 = \frac{(b-1)}{\sqrt{2(b+1)(1+a+a^2+b(a-1)+b^2)}},$$

$$u_2 = \sqrt{\frac{1+2a^2+b^2+2a(1+b)}{2(1+a+a^2+b(a-1)+b^2}}, \tag{4.54}$$

$$u_3 = \frac{(b-1)\sqrt{b}}{\sqrt{2(b+1)(1+a+a^2+b(a-1)+b^2)}}.$$

The minimum of the function Q is a function with respect to two variables a and b. The equation of the minimum function with $1 < a < b$ is

$$\mathbf{Min(a,b)} = \frac{-b(b-1)^3}{4(a+1)(b+1)(a+b)(1+a+a^2+ab-b+b^2)}. \tag{4.55}$$

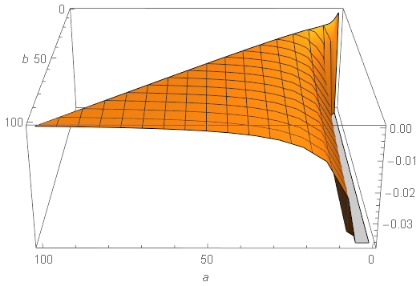

Figure 4.20: The minimum function of the function Q_2.

If $a \longrightarrow 1$ and $b \longrightarrow \infty$, then the limits of the minimum function are:

$$\lim_{a\longrightarrow 1} \mathrm{Min}(a,b) = -\frac{b(b-1)^3}{8(b+1)^2(b^2+3)},$$

$$\lim_{b\longrightarrow\infty} \mathrm{Min}(a,b) = -\frac{1}{4+4a}.$$

♠

We begin with these examples for Theorem 4.1.1:

Example 4.1.1. Let

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

be a diagonal 3×3 matrix, then by the equation (4.54) the matrix $B_0 = \mathbf{u}\otimes\mathbf{v}$ is the rank-one matrix, where

$$\mathbf{u} = \left(\frac{1}{\sqrt{30}}, \sqrt{\frac{5}{6}}, \sqrt{\frac{2}{15}}\right),$$

$$\mathbf{v} = \left(\frac{1}{\sqrt{30}}, -\sqrt{\frac{5}{6}}, \sqrt{\frac{2}{15}}\right).$$

So, the matrix $A + tB_0$ is

$$A + tB_0 = \begin{bmatrix} 1+\frac{t}{30} & -\frac{t}{6} & \frac{t}{15} \\ \frac{t}{6} & 2-\frac{5t}{6} & \frac{t}{3} \\ \frac{t}{15} & -\frac{t}{3} & 4+\frac{2t}{15} \end{bmatrix}.$$

If $K = (A+tB_0)^t(A+tB_0)$, then the characteristic equation is

$$\det\big(K-\lambda I\big) = 0,$$

where I is the 3×3 identity matrix, which is

$$-\lambda^3 + \Big(21 - \frac{11t}{5} + t^2\Big)\lambda^2 + \Big(-84 + 50t - \frac{1121t^2}{100}\Big)\lambda + \Big(64 - \frac{224t}{5} + \frac{196t^2}{25}\Big) = 0.$$

The equation above has three very long solutions but by Taylor's series the first three terms of the eigenvalues of the matrix K are equal to

$$\lambda_1(t) = 1 + \frac{1}{15}t + \frac{1}{60}t^2 + O(t^3),$$

$$\lambda_2(t) = 4 - \frac{10}{3}t + \frac{29}{36}t^2 + O(t^3),$$

$$\lambda_3(t) = 16 + \frac{16}{15}t + \frac{8}{45}t^2 + O(t^3),$$

where $\lambda_1(t) \leq \lambda_2(t) \leq \lambda_3(t)$, on an interval around 0. The linear distortion is

$$\mathcal{H}(A + tB_0) = \sqrt{\frac{\lambda_3(t)}{\lambda_1(t)}} = 4 + L(\theta_1, \theta_2)\, t + Q(\theta_1, \theta_2)\, t^2 + O(t^3).$$

The graph of quadratic coefficient function Q is

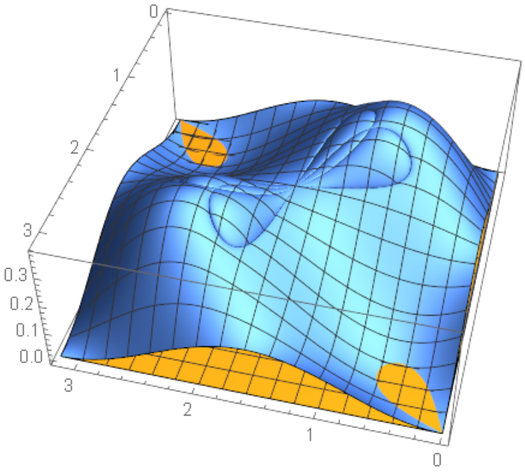

Figure 4.21: If $A = (1, 2, 4)$ is a diagonal matrix, the function Q has two absolute negative minimums at interior points of domain.

We have proved that all extreme points lie on the lines $\theta_2 = \theta_1$ and $\theta_2 = \pi - \theta_1$, therefore the planes $\theta_2 = \theta_1$ and $\theta_2 = \pi - \theta_1$ intersect the graph above at two curves C_1 and C_2. These two curves are shown below:

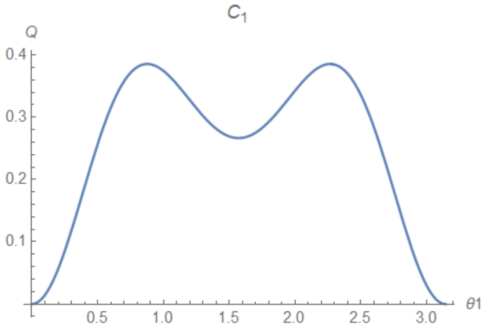

Figure 4.22: The cross section of Q with respect to line $\theta_1 = \theta_2$.

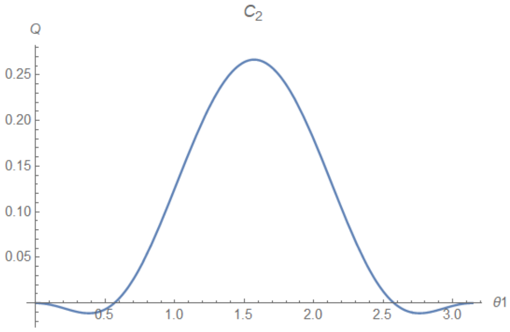

Figure 4.23: The cross section of Q with respect to line $\theta_1 + \theta_2 = \pi$.

$$\mathcal{H}(A + tB_0) = \sqrt{\frac{\lambda_3(t)}{\lambda_1(t)}} = \sqrt{\frac{16 + \frac{16t}{15} + \frac{8t^2}{45}}{1 + \frac{t}{15} + \frac{t^2}{60}}} = 4 - \frac{1}{90}t^2 + O(t^3).$$

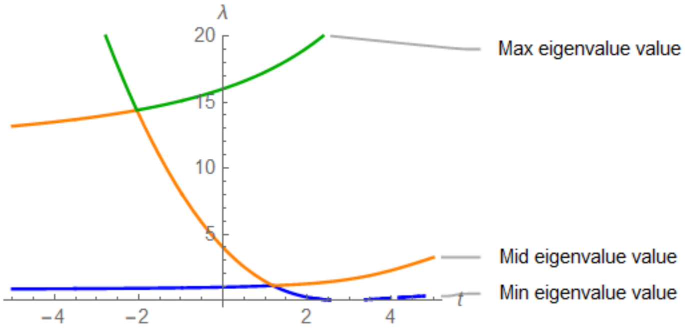

Figure 4.24: The eigenvalues of the matrix $K = (A + tB_0)^t(A + tB_0)$, when $A = (1, 2, 4)$ is a diagonal matrix and B_0 is the optimal direction.

It is easy to see the first derivative of $\mathcal{H}$ at $t = 0$ is 0 and the second derivative of $\mathcal{H}$ is $-\frac{1}{90}$. So,

$$\frac{\mathrm{d}}{\mathrm{dt}}\Big|_{t=0}\mathcal{H}(A + t\,B) = 0, \quad \text{and} \quad \frac{\mathrm{d}^2}{\mathrm{dt}^2}\Big|_{t=0}\mathcal{H}(A + t\,B) = -\frac{1}{90}.$$

This is equal to the minimum value of quadratic coefficient Q. So, there is a rank-one matrix B_0 such that

$$\frac{\mathrm{d}}{\mathrm{dt}}\Big|_{t=0}\mathcal{H}(A + t\,B_0) = 0,$$

$$\frac{\mathrm{d}^2}{\mathrm{dt}^2}\Big|_{t=0}\mathcal{H}(A + t\,B_0) \le \frac{\mathrm{d}^2}{\mathrm{dt}^2}\Big|_{t=0}\mathcal{H}(A + t\,B),$$

for every rank-one matrix $B \in \mathbb{R}^{3\times 3}$ with $\frac{\mathrm{d}}{\mathrm{dt}}\Big|_{t=0}\mathcal{H}(A + t\,B) = 0$.

♣

Example 4.1.2. Let

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{bmatrix},$$

and

$$\mathbf{u} = \left(\frac{7}{3\sqrt{158}}, \sqrt{\frac{109}{158}}, \frac{14}{3\sqrt{79}}\right),$$

$$\mathbf{v} = \left(\frac{7}{3\sqrt{158}}, -\sqrt{\frac{109}{158}}, \frac{14}{3\sqrt{79}}\right).$$

So, $B_0 = \mathbf{u} \otimes \mathbf{v}$ is the optimal rank-one direction. The first three terms of the eigenvalues of the matrix $K = (A + tB_0)^t(A + tB_0)$ are

$$\lambda_1(t) = 1 + \frac{49}{711}t + \frac{9653}{674028}t^2 + O(t^3),$$

$$\lambda_2(t) = 4 - \frac{218}{79}t + \frac{1967123}{3370140}t^2 + O(t^3),$$

$$\lambda_3(t) = 64 + \frac{3136}{711}t + \frac{12544}{31205}t^2 + O(t^3).$$

Therefore, by Taylor's series, the first three terms of the linear distortion of the matrix $A+tB_0$ can be written

$$\mathcal{H}(A + tB_0) = \sqrt{\frac{\lambda_3(t)}{\lambda_1(t)}} = \sqrt{\frac{64 + \frac{3136}{711}t + \frac{12544}{31205}t^2}{1 + \frac{49}{711}t + \frac{9653}{674028}t^2}} = 8 - \frac{343}{10665}t^2.$$

The graph of quadratic coefficient function Q is

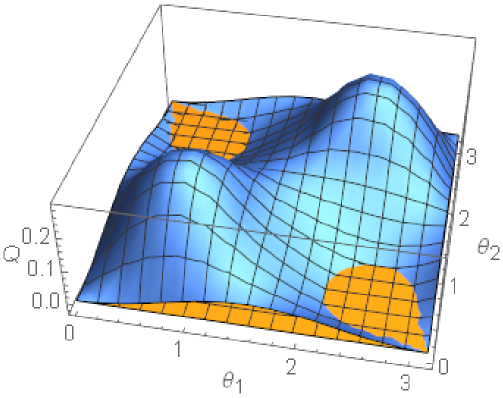

Figure 4.25: The function Q, where $A = (1, 2, 8)$ is a diagonal matrix.

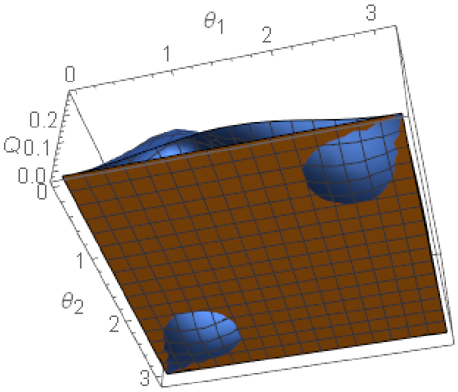

Figure 4.26: The function Q, where $A = (1, 2, 8)$ is a diagonal matrix.

It is easy to see the first derivative of $\mathcal{H}$ at $t = 0$ is 0 and the second derivative of $\mathcal{H}$ is $-\frac{343}{10665}$. So,

$$\left.\frac{\mathrm{d}}{\mathrm{dt}}\right|_{t=0} \mathcal{H}(A + t\,B) = 0, \quad \text{and} \quad \left.\frac{\mathrm{d}^2}{\mathrm{dt}^2}\right|_{t=0} \mathcal{H}(A + t\,B) = -\frac{343}{10665}.$$

We have proved that all extreme points lie on the lines $\theta_2 = \theta_1$ and $\theta_2 = \pi - \theta_1$, therefore the planes $\theta_2 = \theta_1$ and $\theta_2 = \pi - \theta_1$ intersect of the graph above at two curves C_1 and C_2. These two curves are shown below:

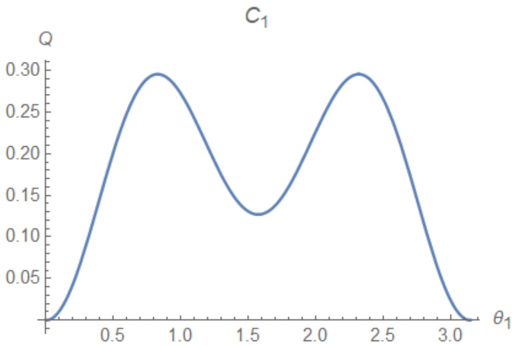

Figure 4.27: The cross section of Q with respect to line $\theta_1 = \theta_2$.

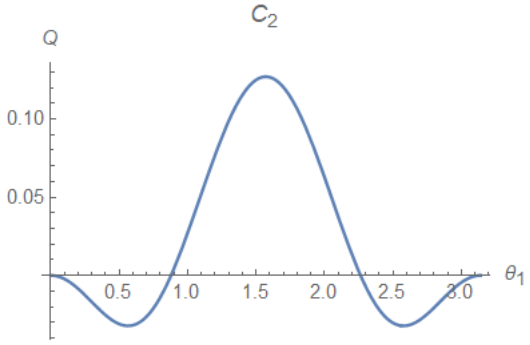

Figure 4.28: The cross section of Q with respect to line $\theta_1 + \theta_2 = \pi$.

So, $Q_{Min} = -\frac{343}{10665}$.

♣

Example 4.1.3. Let

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 120 \end{bmatrix},$$

and $B_0 = \mathbf{u} \otimes \mathbf{v}^t$, where

$$\mathbf{u} = \left(\frac{119}{11\sqrt{29054}}, \sqrt{\frac{14893}{29054}}, \frac{238\sqrt{\frac{15}{14527}}}{11} \right),$$

$$\mathbf{v} = \left(\frac{119}{11\sqrt{29054}}, -\sqrt{\frac{14893}{29054}}, \frac{238\sqrt{\frac{15}{14527}}}{11} \right).$$

The first three terms of the eigenvalues of the matrix $K = (A + tB_0)^t(A + tB_0)$ are

$$\lambda_1(t) = 1 + \frac{14161}{1757767}t + \frac{416744069}{306420974508}t^2 + O(t^3),$$

$$\lambda_2(t) = 4 - \frac{29786}{14527}t + \frac{5114656866379}{18691679444944}t^2 + O(t^3),$$

$$\lambda_3(t) = 14400 + \frac{203918400}{1757767}t + \frac{1129300099200}{1557639953749}t^2 + O(t^3).$$

Therefore, by Taylor's series, the first three terms of the linear distortion of the matrix $A+tB_0$ can be written

$$\mathcal{H}(A + tB_0) = \sqrt{\frac{\lambda_3(t)}{\lambda_1(t)}} = \sqrt{\frac{14400 + \frac{203918400}{1757767}t + \frac{1129300099200}{1557639953749}t^2}{1 + \frac{14161}{1757767}t + \frac{416744069}{306420974508}t^2}} = 120 - \frac{8425795}{107223787}t^2.$$

The graph of quadratic coefficient function Q on the domain $[0, \pi] \times [0, \pi]$ is

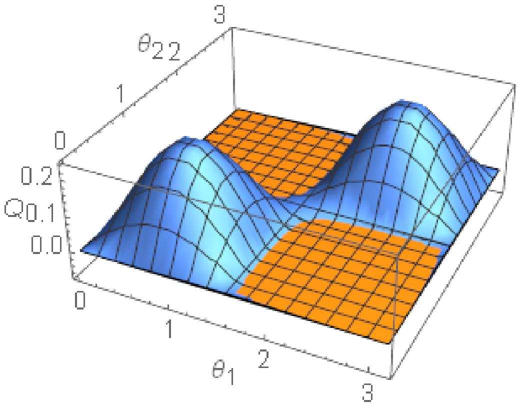

Figure 4.29: The function Q, where $A = (1, 2, 120)$ is a diagonal matrix.

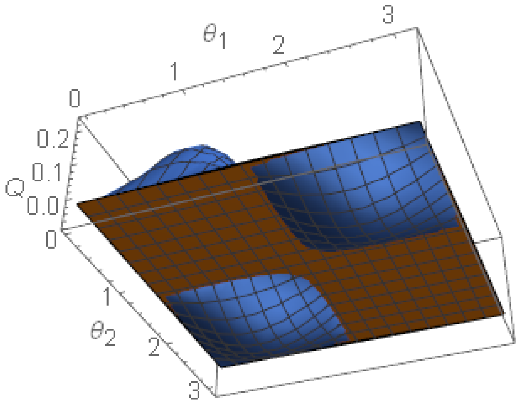

Figure 4.30: The function Q, where $A = (1, 2, 120)$ is a diagonal matrix.

It is easy to see the first derivative of $\mathcal{H}$ at $t = 0$ is 0 and the second derivative of $\mathcal{H}$ is -0.0785813972. So,

$$\frac{\mathrm{d}}{\mathrm{dt}}\bigg|_{t=0} \mathcal{H}(A + t\,B) = 0, \quad \text{and} \quad \frac{\mathrm{d}^2}{\mathrm{dt}^2}\bigg|_{t=0} \mathcal{H}(A + t\,B) = -\frac{8425795}{107223787}.$$

We have proved that all extreme points lie on the lines $\theta_2 = \theta_1$ and $\theta_2 = \pi - \theta_1$, therefore the planes $\theta_2 = \theta_1$ and $\theta_2 = \pi - \theta_1$ intersect of the graph above at two curves C_1 and C_2. These two curves are shown below:

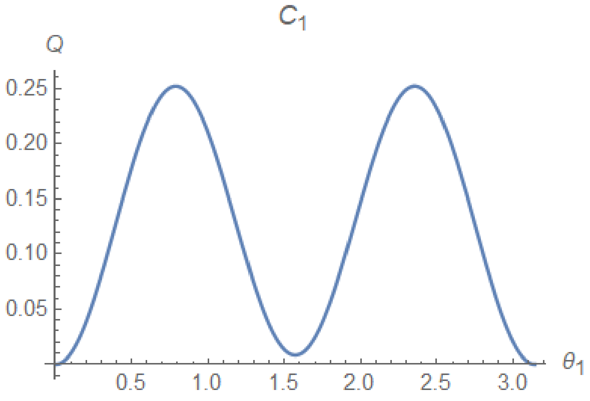

Figure 4.31: The cross section of Q with respect to line $\theta_1 = \theta_2$.

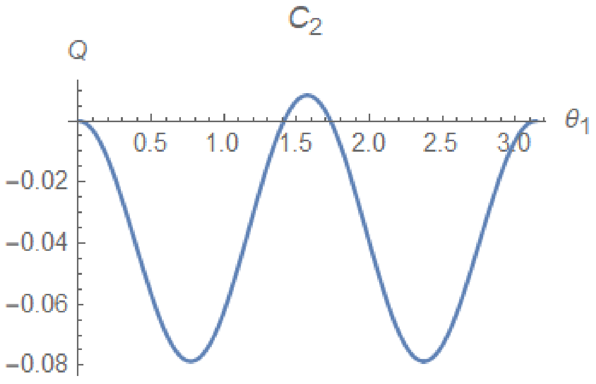

Figure 4.32: The cross section of Q with respect to line $\theta_1 + \theta_2 = \pi$.

So, the negative minimum of the quadratic coefficient function Q is equal to

$$Q_{Min} = -\frac{8425795}{107223787} = -0.0785813972.$$

♣

The second problem that we will address, is finding a neighbourhood of 0 for variable t for which $\mathcal{H}(A+tB_0) < \mathcal{H}(A)$. The regular branches of the maximum and minimum eigenvalues do not cross, but they have intersection points with the middle eigenvalue (See [34]). The middle eigenvalue intersects each of the largest and the smallest eigenvalues just once at typically asymmetric points ($\mathbf{t}_- < 0$ for the mid and max eigenvalues, and $\mathbf{t}_+ > 0$ for the mid and min eigenvalues).

These crossings typically yield different values of the linear distortion. The limiting jump is determined by the largest of these two values, [34]. The next section is devoted to this topic and some related examples.

4.2 Solving Problem 2.

There are conjectures by Gehring and Iwaniec concerning the gap between the linear distortion and the outer distortion with regard to limiting processes, basically trying to estimate the maximum size of the jump showing the failure of lower semi-continuity. In three dimensions as a consequence of Gehring and Iwaniec (see [32], and [28]) if $\mathbf{f_j} \longrightarrow \mathbf{f}$ and $\mathcal{H}(\mathbf{x}, \mathbf{f_j}) \leq H$, then

$$\mathcal{H}(\mathbf{x}, \mathbf{f}) \leq \frac{1}{2}(H + H^2)^{\frac{2}{3}}.$$

The solution to problem 2 should go some way to providing concrete examples of what can happen and possibly resolve their conjectures, or offer more insight and potentially new conjectures.

Theorem 4.2.1. Let

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{bmatrix}$$

be diagonal with $1 < a < b$ and suppose $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^3$ have $\|\mathbf{u}\| = \|\mathbf{v}\| = 1$ so that $B_0 = \mathbf{u} \otimes \mathbf{v}^t$ is a solution to Theorem 4.1.1.

1. There are the largest magnitude of two real numbers $\mathbf{t}_+ > 0$ and $\mathbf{t}_- < 0$ so that $\mathcal{H}(A + t\,B_0)$ is a smooth function of t in the interval $\mathbf{t}_- < t < \mathbf{t}_+$.
2. There are two real numbers $\mathcal{H}(A+\mathbf{t}_- B_0)$ and $\mathcal{H}(A+\mathbf{t}_+ B_0)$ such that for all rank-one matrices B with

   $$\frac{\mathrm{d}}{\mathrm{dt}}\Big|_{t=0} \mathcal{H}(A + t\,B) = 0, \quad \text{and} \quad \frac{\mathrm{d}^2}{\mathrm{dt}^2}\Big|_{t=0} \mathcal{H}(A + t\,B) < 0,$$

   we have for all $t > 0$

   $$\mathcal{H}(A + tB) \geq \max\{\mathcal{H}(A + \mathbf{t}_- B_0), \mathcal{H}(A + \mathbf{t}_+ B_0)\}.$$

   .

We expect that the values $\mathbf{t}_-$ and $\mathbf{t}_+$ are where the singular values of $A+tB_0$ cross. These are determined from a (rather challenging) discriminant problem. This conjecture expresses the hope that the best direction also leads to the biggest gap between $\mathcal{H}(A)$ and $\mathcal{H}(A+t\,B)$ and therefore gives us the approximation to A of least linear distortion via Iwaniec's construction.

We need to find the maximum interval on which $\mathcal{H}(A+tB_0)$ is smooth and concave. Thus we need to identify the discriminant of the eigenvalue equation for $\mathcal{H}(A+tB_0)$ as it is the transverse crossing of the eigenvalues which implies $\mathcal{H}$ loses smoothness.

Proof. The two vectors

$$\mathbf{u}=\Big(\frac{(b-1)}{\sqrt{2(b+1)(1+a+a^2+b(a-1)+b^2)}},\sqrt{\frac{1+2a^2+b^2+2a(1+b)}{2(1+a+a^2+b(a-1)+b^2}},\frac{(b-1)\sqrt{b}}{\sqrt{2(b+1)(1+a+a^2+b(a-1)+b^2)}}\Big),$$

and

$$\mathbf{v}=\Big(\frac{(b-1)}{\sqrt{2(b+1)(1+a+a^2+b(a-1)+b^2)}},-\sqrt{\frac{1+2a^2+b^2+2a(1+b)}{2(1+a+a^2+b(a-1)+b^2}},\frac{(b-1)\sqrt{b}}{\sqrt{2(b+1)(1+a+a^2+b(a-1)+b^2)}}\Big)$$

give the optimal rank-one matrix $B_0=\mathbf{u}\otimes\mathbf{v}$ as at (4.54) before. Suppose that the eigenvalues of the matrix $(A+tB_0)^t(A+tB_0)$ are $\lambda_1<\lambda_2<\lambda_3$. Using Taylor's series the first three terms of the eigenvalues can be calculated up to order 2 as

$$\lambda_1=1+\frac{(b-1)^2}{(b+1)(1+a+a^2-b+ab+b^2)}\,t+\frac{(b-1)^2(3+4a^2+2b^2-b+5a+3ab)}{4(a+1)(b+1)(1+a+a^2+b(a-1)+b^2)^2}\,t^2,$$

$$\lambda_2=a^2-\frac{a(1+2a+2a^2+2ab+b^2)}{(1+a+a^2-b+ab+b^2)}\,t+\frac{\big(1+2a(1+a)+2ab+b^2\big)\Big[a+2a^4(1+b)+}{4(a+1)(b+1)(a+b)(1+a+a^2+b(a-1)+b^2)^2}$$

$$\frac{4a^3(1+b)^2+b(1+b)(1+b^2)+a^2(1+b)(5+2b+5b^2)+ab(8+b(-2+b(8+b)))\Big]}{4(a+1)(b+1)(a+b)(1+a+a^2+b(a-1)+b^2)^2}\,t^2,$$

$$\lambda_3=b^2+\frac{(b-1)^2b^2}{(b+1)(1+a+a^2-b+ab+b^2)}\,t+\frac{b^2(b-1)^2(2+4a^2-b+3b^2+3a+5ab)}{4(a+b)(b+1)(1+a+a^2+b(a-1)+b^2)^2}\,t^2.$$

Therefore, the first four terms of the Taylor's series for the linear distortion function around 0 are

$$\mathcal{H}(A+t\,B_0)=\sqrt{\frac{\lambda_3(t)}{\lambda_1(t)}}=b-\frac{b(b-1)^3}{4(a+1)(b+1)(a+b)(1+a+a^2+ab-b+b^2)}\,t^2.$$

It is easy to see that $\frac{\mathrm{d}}{\mathrm{dt}}\Big|_{t=0}\mathcal{H}(A+t\,B_0)=0$, and the coefficient of the quadratic term is

$$Q(a,b)=-\frac{b(b-1)^3}{4(a+1)(b+1)(a+b)(1+a+a^2+ab-b+b^2)}=\mathrm{Min}(a,b).$$

This coefficient of the quadratic term is equal to the minimum function in the Equation 4.55, because we use the best rank-one direction matrix B_0. However, Taylor's series can not be used here and we need the actual eigenvalues of $(A+tB_0)^t(A+tB_0)$ to find the intersection points of the eigenvalues.
Let K be the 3×3 matrix $K=(A+tB_0)^t(A+tB_0)$. So K is symmetric and hence can

be orthogonally diagonalised, by the Spectral Theorem (Theorem 1.6.8). Not only are the eigenvalues of K all real (Theorem (1.6.7)), they are all nonnegative. To show this, let λ be an eigenvalue of K with corresponding unit vector $\mathbf{x}$. Then

$$0 \leq \|K\mathbf{x}\|^2 = (K\mathbf{x}) \cdot (K\mathbf{x}) = (K\mathbf{x})^t K\mathbf{x} = \mathbf{x}^t K^t K\mathbf{x} = \mathbf{x}^t \lambda \mathbf{x} = \lambda(\mathbf{x} \cdot \mathbf{x}) = \lambda \|\mathbf{x}\|^2 = \lambda,$$

so, all eigenvalues of matrix K are nonnegative. The characteristic equation $\det(K - \lambda I) = 0$ is:

$$\begin{aligned}
\det(K - \lambda I) = & \frac{b^2\left(2a^3(b+1) + 2a^2(b+1)(b-t+1) + 2a(b^3-4bt+1) - (b^3+b^2+b+1)t\right)^2}{4(b+1)^2\left(a^2+ab+a+b^2-b+1\right)^2} - \\
& \frac{1}{4(b+1)^2\left(a^2+ab+a+b^2-b+1\right)^2}\Big[4a^6(b+1)^2(1+b^2) + 8a^5(b+1)^2(1+b^2)(1+b-t) + \\
& 8b^6(t-2)t + t^2 + 4bt^2 + 4b^7t(t+2) - 4b^3t(3t+4) + 2b^4t(15t+4) + 4b^2(1+2t+2t^2) + \\
& b^8(t^2+4) - 4b^5(3t^2-2t-2) + 4a^4(b+1)\Big(3 + 3b^5 - 3b^4(t-1) - 3t + t^2 + b(t^2-10t+3) + \\
& b^3(t^2-10t+7) + b^2(t^2-6t+7)\Big) - 2a(b+1)\Big(2b^7t - (t-2)t - 3b^4(t-2) - 8bt^2 + b^3(-8t^2 + \\
& 6t-4) - b^6(t^2+4) - 4b^5(2t^2+1) - b^2(3t^2+4)\Big) + 4a^3(b+1)\Big(2 + 2b^6 + b^5(2-4t) - 4t + t^2 + \\
& b^4(t^2-8t+4) + 2b^2(t^2-6t+2) + 2b^3(3t^2-6t+4) + b(6t^2-8t+2)\Big) + a^2(b+1)\Big(4 + 4b^7 - \\
& 8t + 5t^2 - 4b^6(2t+1) + b^5(5t^2-28t+20) + 2b^3(11t^2-36t+6) + b(21t^2-28t-4) + b^4(12 + \\
& 8t + 21t^2) + b^2(20+8t+22t^2)\Big)\Big]\lambda + \frac{1}{(b+1)(1+a+a^2-b+ab+b^2)}\Big[1 + b^5 + a^4(b+1) + \\
& a^3(b+1)(1+b-2t) + b^3(t-1)^2 + t - 2bt + b^4t + t^2 + b^2(1+2b) + a(1+b)(1+b+b^2 + \\
& b^3 - t - b^2t + t^2 + bt^2) + a^2(b+1)(2+2b^2-2t+t^2-b(1+2t))\Big]\lambda^2 - \lambda^3 = 0.
\end{aligned} \tag{4.56}$$

This characteristic equation (4.56) has three nonnegative real roots. These three eigenvalues have long and very complicated equations. That means we can not find their intersection points by the following equations:

$$\lambda_3(t) = \lambda_1(t), \qquad \lambda_3(t) = \lambda_2(t), \qquad \lambda_2(t) = \lambda_1(t).$$

To avoid this complexity, we use an algebraic trick. A root to the above equations means that there is a repeated root of the characteristic equation (4.56). It seems quite remarkable that the discriminant admits a simple quadratic factor, but this is a consequence of the fact only two eigenvalues are crossing. This quadratic factor of Equation 4.58, given below, determines the crossing points of the singular values [34].

$$\text{Discriminant factorisation} = \frac{1}{16(\mathrm{b}+1)^5(1+\mathrm{a}+\mathrm{a}^2-\mathrm{b}+\mathrm{ab}+\mathrm{b}^2)^5}\big(\mathrm{P}(\mathrm{t})\big)^2\mathrm{R}(\mathrm{t}), \tag{4.57}$$

where $P(t)$ is equal to

$$\begin{aligned} P(t) =&\big(-1-a-2a^2-4b-7ab-4a^2b+2b^2-7ab^2-2a^2b^2-4b^3-ab^3-b^4\big)\, t^2+\\ &\big(-1+2a+a^2+4a^3-6b-4ab+7a^2b+8a^3b-b^2-12ab^2+7a^2b^2+4a^3b^2-\\ &b^3-4ab^3+a^2b^3-6b^4+2ab^4-b^5\big)\, t-2\big(-a+a^4+b-ab-2a^2b+2a^4b+b^2+\\ &2ab^2-4a^2b^2+a^4b^2+2ab^3-2a^2b^3+b^4-ab^4+b^5-ab^5\big)=0. \end{aligned} \tag{4.58}$$

The discriminant of this quadratic equation is

$$\begin{aligned} \triangle =&(b^2-1)^2\Big[1+4b+6b^3+4b^5+b^6+4a(b+1)\big(b^4+3b^3+3b+1\big)+4a^3(b+1)\\ &\big(3+(2+3b)\big)+2a^2(1+b+b^2)\big(5+b(6+5b)\big)+a^4\big(9+b(9b-2)\big)\Big], \end{aligned} \tag{4.59}$$

which is obviously positive. That means the quadratic equation has two real roots that are called $\mathbf{t}_+$ and $\mathbf{t}_-$. If

$$a_\circ t^2+b_\circ t+c_\circ=0,$$

is a quadratic equation, and $\triangle = b_\circ^2-4a_\circ c_\circ>0$. If $t_1\cdot t_2=\frac{c_\circ}{a_\circ}<0$, then the two real roots have different signs. In the quadratic equation (4.58), we have

$$\frac{c_\circ}{a_\circ}=\frac{2(a-1)(a-b)(b+1)^2(1+a+a^2-b+ab+b^2)}{1+a+2a^2+4b+7ab+4a^2b+2b^2(a^2-1)+7ab^2+4b^3+ab^3+b^4},$$

which is obviously negative, because $1<a<b$. Thus the two real roots of the quadratic equation (4.58) have different signs. Let $\mathbf{t}_+>0$ and $\mathbf{t}_-<0$. We write

$$\begin{aligned} \mathbf{t}_+ &= \frac{G_1(a,b)-(b^2-1)\sqrt{J_1(a,b)}}{2(-1-a-2a^2-4b-7ab-4a^2b+2b^2-7ab^2-2a^2b^2-4b^3-ab^3-b^4)},\\ \mathbf{t}_- &= \frac{G_1(a,b)+(b^2-1)\sqrt{J_1(a,b)}}{2(-1-a-2a^2-4b-7ab-4a^2b+2b^2-7ab^2-2a^2b^2-4b^3-ab^3-b^4)}, \end{aligned} \tag{4.60}$$

where,

$$\begin{aligned} G_1(a,b)=&1-2a-a^2-4a^3+6b+4ab-7a^2b-8a^3b+b^2+12ab^2-7a^2b^2-\\ &4a^3b^2+b^3+4ab^3-a^2b^3+6b^4-2ab^4+b^5, \end{aligned} \tag{4.61}$$

and

$$\begin{aligned} J_1(a,b) =&1 + 4a + 10a^2 + 12a^3 + 9a^4 + 4b + 16ab + 22a^2b + 20a^3b - 2a^4b+ \\ &12ab^2 + 32a^2b^2 + 20a^3b^2 + 9a^4b^2 + 6b^3 + 12ab^3 + 22a^2b^3 + 12a^3b^3+ \\ &16ab^4 + 10a^2b^4 + 4b^5 + 4ab^5 + b^6. \end{aligned} \tag{4.62}$$

If the Equation 4.58 is written as:

$$D_1\lambda^3 + C_1\lambda^2 + B_1\lambda + A_1 = 0, \tag{4.63}$$

where

$$\begin{aligned} D_1 =& -4(b+1)^2\big(1 + a + a^2 + (a-1)b + b^2\big)^2, \\ C_1 =& 4(b+1)\big(1 + a + a^2 + (a-1)b + b^2\big)\Big(1 + b^2 + b^3 + b^5 + a^4(b+1) + a^3(b+1) \\ &(1 + b - 2t) + (b-1)^2(b^2+1)t + (b^3+1)t^2 + a^2(b+1)\big(2 - b + 2b^2 - 2(b+1)t+ \\ &t^2\big) + a(b+1)\big(1 + b + b^2 + b^3 - (b^2+1)t + (b+1)t^2\big)\Big), \end{aligned}$$

$$\begin{aligned} B_1 =& -4a^6(b+1)^2(b^2+1) - 4(b^4+b)^2 - 8a^5(1+b)^2(b^2+1)(1+b-t) - 8b^2(b-1)^2 \\ &(b^3+1)t - \Big[1 + b\Big(4 + b\big[8 + b\big(-12 + b(30 + b(-12 + b(8 + 4(4+b)))))\big]\Big)\Big]+ \\ &4a^4(b+1)\Big(-(b+1)(3 + 7b^2 + 3b^4) + (b+3)(3b+1)(b^2+1)t - (b+1)(b^2+1)t^2\Big)+ \\ &4a^3(b+1)\Big(-2\big(1 + b + b^2(b+1)(2 + 2b + b^3)\big) + 4(b+1)(b^2+1)(1 + b + b^2)t- \\ &(b^2+1)(1 + b(6+b))t^2\Big) + a^2(b+1)\Bigg[-4(b+1)(b^2+1)\Bigg(1 + b\big(-2 + b(6+ \\ &(-2+b)b)\big)\Bigg) + 4\Bigg(2 + b\Bigg(7 + b\Big[-2 + b\Big(18 + b\big(-2 + b(7+2b)\big)\Big)\Big]\Bigg)\Bigg)t - (1+b) \\ &\Bigg(5 + b\Big(16 + b\big(9 + b(16 + 5b)\big)\Big)\Bigg)t^2\Bigg] + 2a(b+1)\Bigg[-4(b^2 + b^3 + b^5 + b^6) + 2t \\ &(1 + b^3(3 + 3b + b^4)) - \Bigg(1 + b\Bigg(8 + b\Big(3 + b\big(8 + b(3 + b(8+b))\big)\Big)\Bigg)\Bigg)t^2\Bigg], \end{aligned}$$

$$A_1 = b^2\Big(2a^3(b+1) + 2a^2(b+1)(1+b-t) - (b+1)(b^2+1)t + 2a(1 + b^3 - 4bt)\Big)^2.$$

The roots of the Equation 4.63 are

$$\begin{aligned}
\lambda_1 &= Z(a,b,t) - \frac{2^{\frac{1}{3}} \times Y(a,b,t)}{3 \times X(a,b,t)} + \frac{X(a,b,t)}{3 \times 2^{\frac{1}{3}} \times D_1}, \\
\lambda_2 &= Z(a,b,t) + \frac{\frac{1-i\sqrt{3}}{2} \times 2^{\frac{1}{3}} \times Y(a,b,t)}{3 \times X(a,b,t)} - \frac{\frac{1+i\sqrt{3}}{2} \times X(a,b,t)}{3 \times 2^{\frac{1}{3}} \times D_1}, \\
\lambda_3 &= Z(a,b,t) + \frac{\frac{1+i\sqrt{3}}{2} \times 2^{\frac{1}{3}} \times Y(a,b,t)}{3 \times X(a,b,t)} - \frac{\frac{1-i\sqrt{3}}{2} \times X(a,b,t)}{3 \times 2^{\frac{1}{3}} \times D_1}
\end{aligned} \tag{4.64}$$

where

$$X(a,b,t) = \left(-2C_1^3 + 9B_1C_1D_1 - 27A_1D_1 + \sqrt{-4\left(C_1^2 - 3B_1D_1\right)^3 + \left(2C_1^3 - 9B_1C_1D_1 + 27A_1D_1^2\right)^2} \right)^{\frac{1}{3}},$$

$$Y(a,b,t) = \frac{(-C_1^2 + 3B_1D_1)}{D_1} \qquad \text{and} \qquad \mathrm{Z(a,b,t)} = -\frac{\mathrm{C_1}}{\mathrm{3D_1}}.$$

We know that $0 \leq \lambda_1 \leq \lambda_2 \leq \lambda_3$, so, $0 \leq \sigma_1 = \sqrt{\lambda_1} \leq \sigma_2 = \sqrt{\lambda_2} \leq \sigma_3 = \sqrt{\lambda_3}$, are singular values of matrix $(A + tB_0)$. The linear distortion of matrix $A + tB_0$ is:

$$\mathcal{H}(A + tB_0) = \sqrt{\frac{\lambda_3}{\lambda_1}} = \frac{\sigma_3}{\sigma_1}. \tag{4.65}$$

These three eigenvalues are nonnegative real numbers, but they are described by three variables a, b, and t. The eigenvalues and the linear distortion are one-variable functions concerning t.

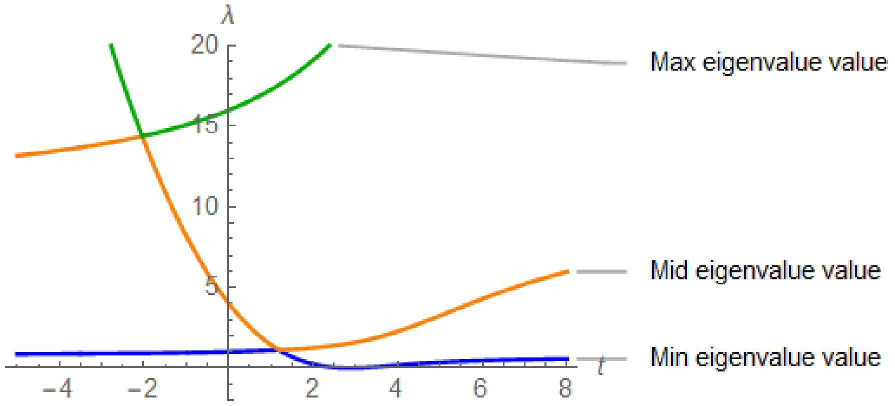

Figure 4.33: The eigenvalues of matrix $K = (A + tB_0)^t(A + tB_0)$, when $A = (1, 2, 4)$ is a diagonal matrix.

By using Equations 4.60 and 4.65, we can find the values of the following

$$\mathbf{t}_+, \quad \mathbf{t}_-, \quad \mathcal{H}(A + \mathbf{t}_+B_0), \quad \mathcal{H}(A + \mathbf{t}_-B_0).$$

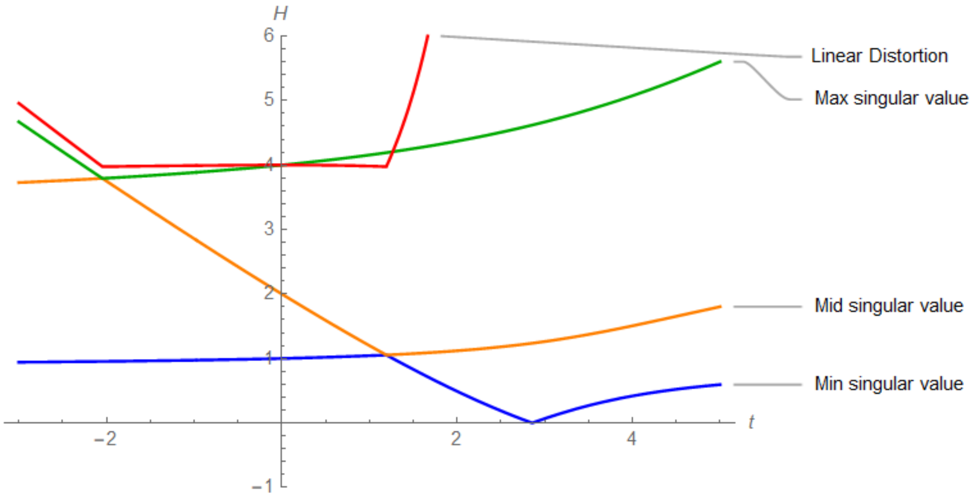

Figure 4.34: The eigenvalues and the linear distortion of matrix $K = (A + tB_0)^t(A + tB_0)$, when $A = (1, 2, 4)$ is a diagonal matrix.

So, the maximum jump is

$$\textbf{Maximum jump} = \min\{|\mathcal{H}(\mathrm{A} + \mathrm{t}_{+}\mathrm{B}_0) - \mathcal{H}(\mathrm{A})|, |\mathcal{H}(\mathrm{A} + \mathrm{t}_{-}\mathrm{B}_0) - \mathcal{H}(\mathrm{A})|\}.$$

Remember that the value of $\mathcal{H}(A)$ is b. ♠

Example 4.2.1. Following on from the Example 4.1.1, we compute $\mathcal{H}(A + tB)$, its graph, $\mathbf{t}_{+}$ and $\mathbf{t}_{-}$. The graph of eigenvalues of matrix K is shown in Figure 4.35.

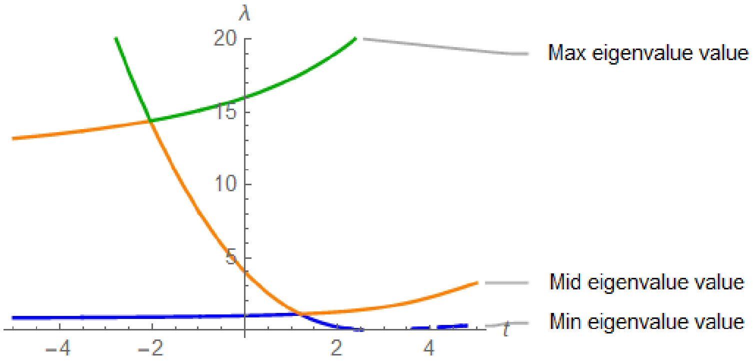

Figure 4.35: The intersection points of the eigenvalues of K.

The graphs of singular values and the linear distortion show two intersection points. At this point, we have shown how to find the best direction, and given this direction, how to find the maximal interval on which the linear distortion is smooth and concave (see [34]).

By using (4.60), (4.61), (4.62) and (4.65), those two points are (for $A = \mathrm{diag}(1, 2, 4)$)

$$\mathbf{t}_{+} = 1.19219, \quad \mathbf{t}_{-} = -2.04584, \quad \mathcal{H}(A + \mathbf{t}_{+}B_0) = 3.97539, \quad \mathcal{H}(A + \mathbf{t}_{-}B_0) = 3.97539.$$

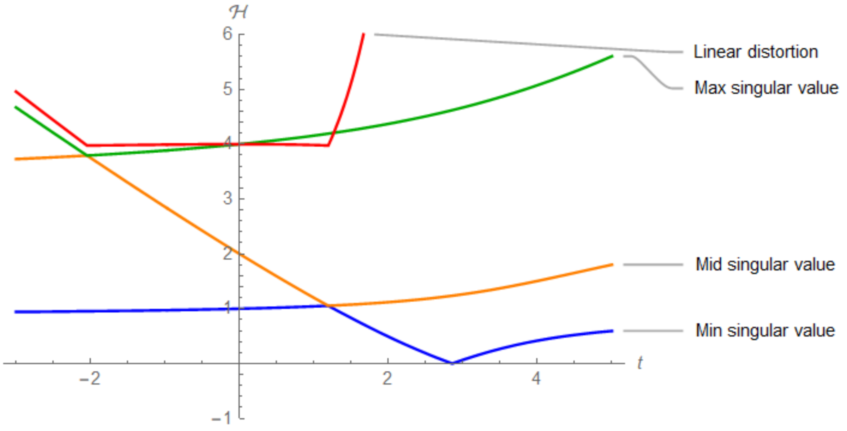

Figure 4.36: The singular values and the linear distortion of matrix $(A + tB_0)$, when $A = (1, 2, 4)$ is a diagonal matrix.

So, for all $t \in (-2.04484,\ 1.19219)$, we have $\mathcal{H}(A + tB_0) < \mathcal{H}(A)$. The the biggest jump is

$$\text{Maximum jump} = \min\{|\mathcal{H}(\mathrm{A} + \mathrm{t}_{+}\mathrm{B}_0) - \mathcal{H}(\mathrm{A})|, |\mathcal{H}(\mathrm{A} + \mathrm{t}_{-}\mathrm{B}_0) - \mathcal{H}(\mathrm{A})|\} = 0.02461,$$

since the value of $\mathcal{H}(A)$ is 4.

♣

Example 4.2.2. Let

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 8 \end{bmatrix}$$

be a diagonal 3×3 matrix, then by the Equation 4.54 the matrix $B_0 = \mathbf{u} \otimes \mathbf{v}$ is the best rank-one matrix direction, where

$$\mathbf{u} = \left(\frac{7}{3\sqrt{158}}, \sqrt{\frac{109}{158}}, \frac{14}{3\sqrt{79}} \right),$$

$$\mathbf{v} = \left(\frac{7}{3\sqrt{158}}, -\sqrt{\frac{109}{158}}, \frac{14}{3\sqrt{79}} \right).$$

If $K = (A + tB_0)^t(A + tB_0)$, then the graphs of the eigenvalues of the matrix K are shown in Figure 4.37

The graphs of singular values and the linear distortion show two intersection points. At these points we have sketched how to find the best direction, and given this direction, how to find the maximal interval on which the linear distortion is smooth the concave. By using the Equations 4.60, 4.61, 4.62 and 4.65, those two points are

$$\mathbf{t}_{+} = 1.4429628, \qquad \mathbf{t}_{-} = -6.095007, \qquad \mathcal{H}(A + \mathbf{t}_{+}B_0) = 7.90656, \qquad \mathcal{H}(A + \mathbf{t}_{-}B_0) = 7.60938.$$

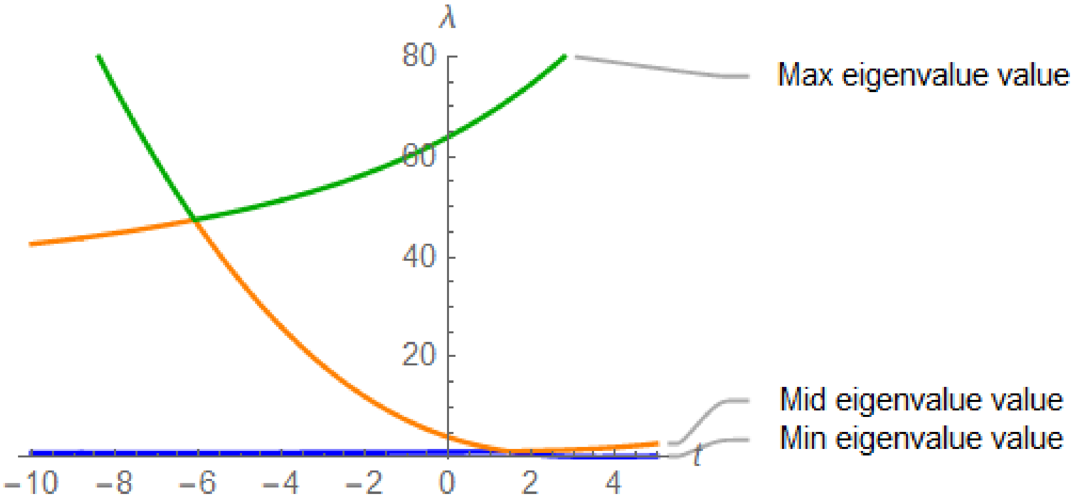

Figure 4.37: The intersection points of the eigenvalues of K.

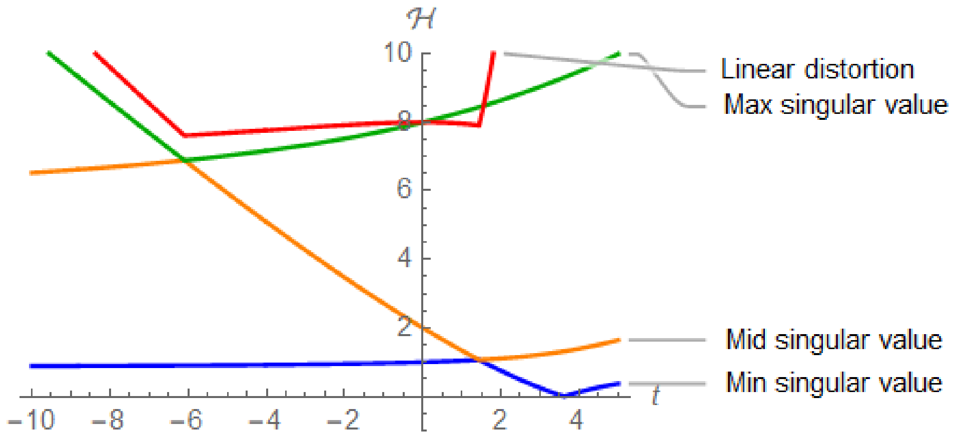

Figure 4.38: The singular values and the linear distortion of matrix $(A + tB_0)$, when $A = (1, 2, 8)$ is a diagonal matrix.

So, for all $t \in (-6.09501,\ 1.442963)$, we have $\mathcal{H}(A + tB_0) < \mathcal{H}(A)$. The biggest jump is

$$\text{Maximum jump} = \min\{|\mathcal{H}(\mathrm{A} + \mathrm{t}_{+}\mathrm{B}_0) - \mathcal{H}(\mathrm{A})|, |\mathcal{H}(\mathrm{A} + \mathrm{t}_{-}\mathrm{B}_0) - \mathcal{H}(\mathrm{A})|\} = 0.09344,$$

since the value of $\mathcal{H}(A)$ is 8.

♣

Now, we can extend the Iwaniec's Lemma and Theorem to (Lemma 3.4.1 and Theorem 3.4.2). Both his lemma and theorem are existence results. But we expand them to a wide range of matrices. The results of Problem (1) and Problem (2) in the previous section and this section show it is always possible to find a rank-one matrix B_0 so that

$$\mathcal{H}(A + t\ B_0) = \mathcal{H}(A + t\ \mathbf{u} \otimes \mathbf{v}) < \mathcal{H}(A), \tag{4.66}$$

holds for a diagonal matrix A with distinct positive singular values. The rank-one matrix B_0 is the best rank-one direction for which the linear distortion function at A is most concave. An unsolved and unanswered question (Partly posed by Frederick W. Gehring and T. Iwaniec) was to determine how big the difference between the left-hand and the right-hand side of inequality (4.66) can be?

In Problems 1 and 2, not only do we find the best direction for which the inequality (4.66) holds but we considered the Gehring and Iwaniec question and found the biggest jump and difference between the left-hand and the right-hand side of inequality (4.66) for a specific diagonal matrix A with distinct singular values. Iwaniec's lemma gives an example (actually counterexample) for the inequality (4.66). With the results we have established so far, we have

Theorem 4.2.2. Let

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{bmatrix}$$

be diagonal with $1 < a < b$. Then there are two unit vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^3$ described at (4.54), and a rank-one matrix $B_0 = \mathbf{u} \otimes \mathbf{v}$ (the best rank-one direction) and two real numbers $\mathbf{t}_- < 0$ and $\mathbf{t}_+ > 0$, the roots of $P(t)$ as in (4.58), such that

$$\frac{\mathrm{d}}{\mathrm{dt}}\Big|_{t=0} \mathcal{H}(A + t\,B_0) \;=\; 0 \tag{4.67}$$

$$\frac{\mathrm{d}^2}{\mathrm{dt}^2}\Big|_{t=0} \mathcal{H}(A + t\,B_0) \;\leq\; \frac{\mathrm{d}^2}{\mathrm{dt}^2}\Big|_{t=0} \mathcal{H}(A + t\,B) \leq 0 \tag{4.68}$$

for every rank-one matrix $B \in \mathbb{R}^{3\times 3}$ with $\frac{\mathrm{d}}{\mathrm{dt}}\Big|_{t=0} \mathcal{H}(A + t\,B) = 0$. Also, there is a largest interval $(\mathbf{t}_-, \mathbf{t}_+)$ so that $\mathcal{H}(A + t\,B_0)$ is a smooth function of t in this interval and

$$\max\{\mathcal{H}(A + \mathbf{t}_- B_0),\ \mathcal{H}(A + \mathbf{t}_+ B_0)\} < \mathcal{H}(A).$$

The jump and difference is defined to be

$$\textbf{Maximum jump} = \min\{|\mathcal{H}(\mathrm{A} + \mathbf{t}_+\mathrm{B}_0) - \mathcal{H}(\mathrm{A})|, |\mathcal{H}(\mathrm{A} + \mathbf{t}_-\mathrm{B}_0) - \mathcal{H}(\mathrm{A})|\},$$

Recall that the value of $\mathcal{H}(A)$ is b.

It is hard or maybe impossible to prove the Theorem 4.2.2 in n-dimensions, because the characteristic equations of $A + tB$ can not be solved and we have a lot of variables and parameters. But in a special case, the above arguments work in higher dimensions when we extend the matrices A and B by the rules (see [39])

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & a & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & b & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & a & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & a & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & a \end{bmatrix}$$

and

$$B = \begin{bmatrix} \sqrt{1-r^2}\sqrt{1-s^2} & \sqrt{1-r^2}s\cos(\theta_2) & \sqrt{1-r^2}s\sin(\theta_2) & 0 & 0 & \cdots & 0 \\ r\sqrt{1-s^2}\cos(\theta_1) & rs\cos(\theta_1)\cos(\theta_2) & rs\cos(\theta_1)\sin(\theta_2) & 0 & 0 & \cdots & 0 \\ r\sqrt{1-s^2}\sin(\theta_1) & rs\cos(\theta_2)\sin(\theta_1) & rs\sin(\theta_1)\sin(\theta_2) & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 \end{bmatrix}.$$

Remark: Remember that these two extended matrices A and B in $\mathbb{R}^{n\times n}$ do not give us the best rank-one direction and the maximum jump. They only show the inequality (4.66) can be true in n dimensions. More precisely, this means that the $n \times n$ matrices A and B can be a counterexample for the lower semicontinuity property of linear distortion functions. Note that we only need three distinct singular values for this to happen. Now we can extend Theorem 3.4.2 to the general case in $\mathbb{R}^3$. In this theorem, we construct a sequence $\{\mathbf{T}_\nu\}_{\nu=1}^{\infty}$ of uniformly bi-Lipschitz maps $\mathbf{T}_\nu : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$, that is,

$$\frac{1}{C}|\mathbf{x}-\mathbf{y}| \le |\mathbf{T}_\nu(\mathbf{x}) - \mathbf{T}_\nu(\mathbf{y})| \le C|\mathbf{x}-\mathbf{y}|$$

for all $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^3$ and $\nu = 1,\ 2,\ 3, \ldots$with a constant C independent of $\mathbf{x}$, $\mathbf{y}$ and ν. The linear distortion function of each $\mathbf{T}_\nu$ turns out to be constant and equal to H almost everywhere:

$$\mathcal{H}(\mathbf{x}, \mathbf{T}_\nu) = H, \qquad\qquad \nu = 1,\ 2, \ldots.$$

The sequence $\{\mathbf{T}_\nu\}$ converges uniformly to a linear map $\mathbf{T_0} : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ whose distortion is bigger than H (see [39], [42], [30]).

Theorem 4.2.3. (Generalisation of Iwaniec's Theorem) If $n = 3$ and $H > 1$ and $\mathbf{T_0} : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ with $\mathbf{T_0}(\mathbf{x}) = A\mathbf{x}$ such that A is a 3×3 diagonal matrix with distinct singular values, then there **always exists** a sequence $\{\mathbf{T}_\nu\}_{\nu=1}^{\infty}$ of quasiconformal mappings $\mathbf{T}_\nu : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ converging uniformly to a linear quasiconformal map $\mathbf{T_0}$ such that

$$\mathcal{H}(\mathbf{x}, \mathbf{T}_\nu) \equiv H < H_0 \equiv \mathcal{H}(\mathbf{x}, \mathbf{T_0}),$$

almost everywhere in $\mathbb{R}^3$.

Proof. This follows with the aid of Theorem 4.2.2. We have $B = \mathbf{u} \otimes \mathbf{v}$ for vectors $\mathbf{u}, \mathbf{v}$ in $\mathbb{R}^3$. Consider $\mathbf{T_0}(\mathbf{x}) = A\mathbf{x}$, and for $\nu = 1, 2, \ldots$, we define a sequence $\{\mathbf{T}_\nu\}_{\nu=1}^{\infty}$ by equation

$$\mathbf{T}_\nu(\mathbf{x}) = A\mathbf{x} + \frac{1}{\nu}h\,(\nu\mathbf{u}\cdot\mathbf{x})\mathbf{v}$$

where h is a periodic piecewise linear function on the real line as follows: Given $\mathbf{t}_- < 0$ and $\mathbf{t}_+ > 0$ of the lemma

$$h(r) = \begin{cases} \mathbf{t}_- r & \frac{i}{\mathbf{t}_+} - \frac{i-1}{\mathbf{t}_-} \le r \le \frac{i}{\mathbf{t}_+} - \frac{i}{\mathbf{t}_-}, \\ \mathbf{t}_+ r & \frac{i}{\mathbf{t}_+} - \frac{i}{\mathbf{t}_-} \le r \le \frac{i+1}{\mathbf{t}_+} - \frac{i}{\mathbf{t}_-}, \end{cases}$$

for any integer i. Then we can extend h to the entire line (saw-tooth function). the function h is a bounded Lipschitz function whose derivative assumes only the two values $\mathbf{t}_-$ and $\mathbf{t}_+$. The sequence $\{\mathbf{T}_\nu\}_{\nu=1}^\infty$ converges uniformly to $\mathbf{T_0}$. The derivative of $\mathbf{T}_\nu$ also assumes only two values, which are independent of ν, apart from countable set of points where it is not defined.

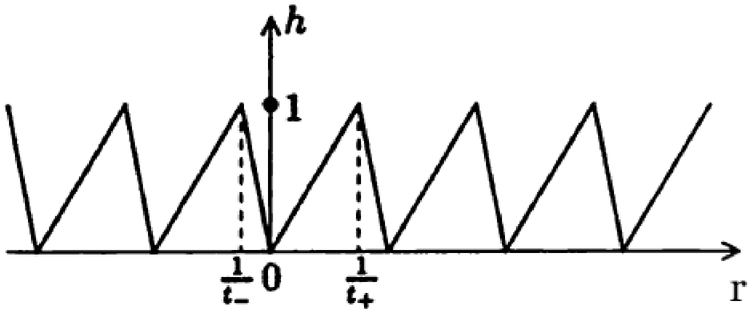

Figure 4.39: The saw-tooth function

$$D\mathbf{T}_\nu = A + h'(\nu\mathbf{u}\cdot\mathbf{x})\mathbf{u}\otimes\mathbf{v} = A + h'(\nu\langle\mathbf{u},\mathbf{x}\rangle)B \in \{A - \mathbf{t}_-B, A + \mathbf{t}_+B\}.$$

In either case, the linear distortion of $D\mathbf{T}_\nu(x)$ is equal to H:

$$\mathcal{H}(\mathbf{x},\mathbf{T}_\nu) = \mathcal{H}(D\mathbf{T}_\nu(\mathbf{x})) \equiv H < \mathcal{H}(A) = \mathcal{H}(\mathbf{x},\mathbf{T_0}).$$

Iwaniec's argument now proves the result. We remark that the sequence $\{h'(\nu\langle\mathbf{u},\mathbf{x}\rangle)\}_{\nu=1}^\infty$ converges weakly in $L^\infty(\mathbb{R}^3)$ to 0 as $\nu \longrightarrow \infty$, but not pointwise almost everywhere. ♠

4.3 Limit and the Jump Ratio

In this section we consider various limit regimes modelling the situation in one or two dimensions by letting $a, b \longrightarrow \infty$ in various ways. This has the same effect as letting the smallest singular value tend to 0 and effectively removing a variable from the deformation. Let $A = \mathrm{diag}(1,a,b)$ be a diagonal matrix with $1 < a < b$, and suppose $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^3$ have $\|\mathbf{u}\| = \|\mathbf{v}\| = 1$ so that $B_0 = \mathbf{u}\otimes\mathbf{v}$ is a solution to Problem 1. We have outlined above the closed form solutions for the two values where the singular values cross $\mathbf{t}_\pm$ and therefore have the following limits. If $b \longrightarrow \pm\infty$, then the limits of t_+ and t_- are

$$\lim_{b\to+\infty} \mathbf{t}_+(a,b) = 2(a-1), \qquad \lim_{b\to+\infty} \mathbf{t}_-(a,b) = -\infty.$$

The Figure 4.40 illustrates that the limit of $\mathbf{t}_-$ is $-\infty$ when $b \longrightarrow +\infty$. This is obvious, because

$$\lim_{b\to+\infty} \mathbf{t}_- = \lim_{b\to+\infty} \frac{(1+\cdots+b^5)+(b^2-1)\sqrt{(1+\cdots+b^6)}}{2(-1-\cdots-b^4)} = \lim_{b\to+\infty} \frac{2b^5+\cdots}{-2b^4+\cdots} = \lim_{b\to+\infty}(-b+\cdots) = -\infty.$$

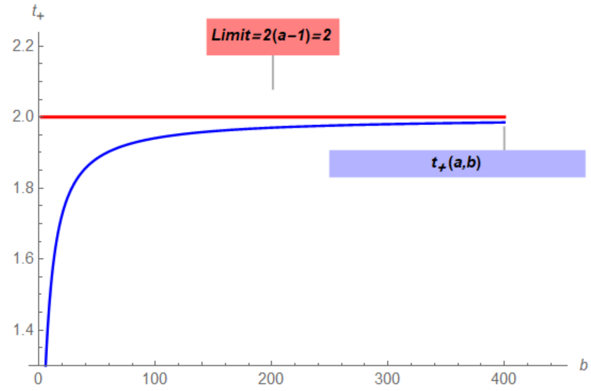

Figure 4.40: The limit of the function $\mathbf{t}_+(a,b)$ when $a=2$ and $b \longrightarrow +\infty$.

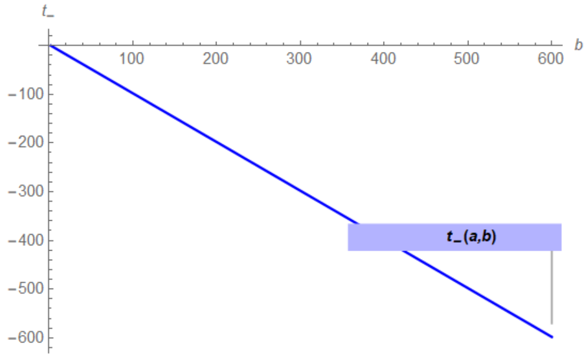

Figure 4.41: The limit of the function $\mathbf{t}_-(a,b)$ when $a=2$ and $b \longrightarrow +\infty$.

In this case, the graph of $\mathbf{t}_-$ is linear and tends to $-\infty$. Our explicit formulas allow computational investigation to help understand the size of the jump (see [34]). Assume the matrix

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{bmatrix}$$

is given, we define the jump ratio as the value of

$$\text{The jump ratio} = \frac{\max\{\mathcal{H}(A+t_-B_0),\ \mathcal{H}(A+t_+B_0)\}}{H(A)} = \frac{\max\{\mathcal{H}(A+t_-B_0),\ \mathcal{H}(A+t_+B_0)\}}{b},$$

since $\mathcal{H}(A)=b$. Recall the best rank-one direction matrix B_0 and the values $\mathbf{t}_-$ and $\mathbf{t}_+$ can be determined by numbers a, and b, where the singular values of $A+tB_0$ cross. We know that if $t \in [\mathbf{t}_-, \mathbf{t}_+]$, then $\mathcal{H}(A+tB_0)$ is concave and has absolute maximum b at $t=0$, and the absolute minimum happens at one of the two points $\mathbf{t}_-$ or $\mathbf{t}_+$. So,

$$\text{The jump ratio} = \max\{\frac{\mathcal{H}(\mathrm{A}+\mathbf{t}_-\mathrm{B}_0)}{\mathrm{b}}, \frac{\mathcal{H}(\mathrm{A}+\mathbf{t}_+\mathrm{B}_0)}{\mathrm{b}}\}. \tag{4.69}$$

If the matrix

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & f(x) \end{bmatrix},$$

where $1 < x < f(x)$, then the jump ratio will be different. Suppose that $A = \mathrm{Sing}(1, a, b)$ shows a digonal matrix A with distinct ordered singular values 1, a, and b. It is easy to see that the two trivial cases $A = \mathrm{Sing}(1, 1, x)$ or $A = \mathrm{Sing}(1, x, x)$ represent the constant graphs. These differences depend on the type of function f and its coefficients and powers. For instance, if $f(x) = x^s$ then the jump ratios for different $s > 1$ are

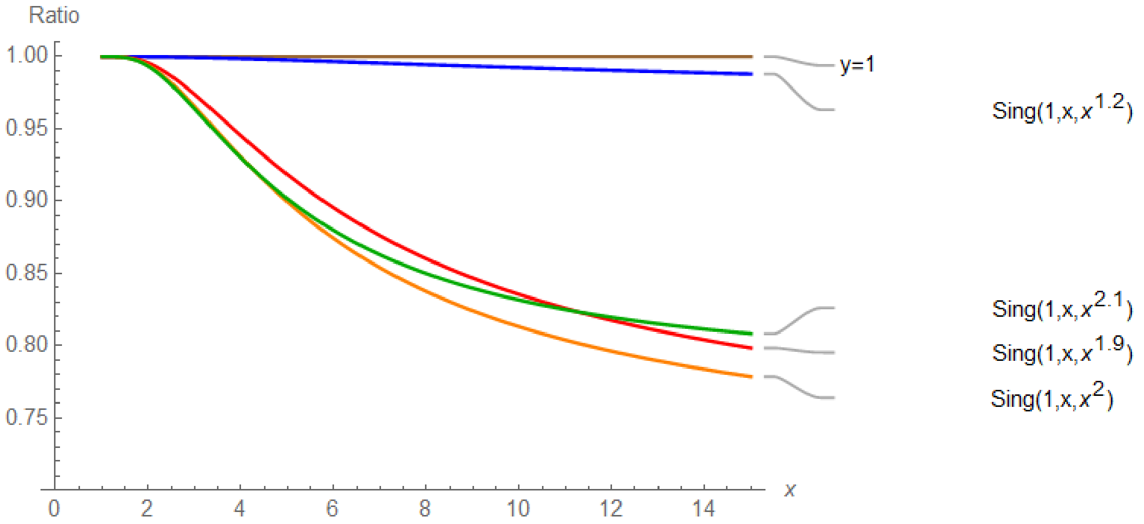

Figure 4.42: The relative jump ratio for $A \in \mathrm{Sing}(1, x, x^s)$, where $s > 1$. The function f is indicated by the labelling.

The following graph shows us the different curves of $f(x) = x^s$ for more values of s and $\mathrm{Sing}(1, x, x^2)$ has the biggest jump ratio. If $s > 2$ or $s < 2$ the jump ratio function tends to a number bigger than $1/\sqrt{2}$, so, the jump ratio in these cases are smaller than the case $s = 2$.

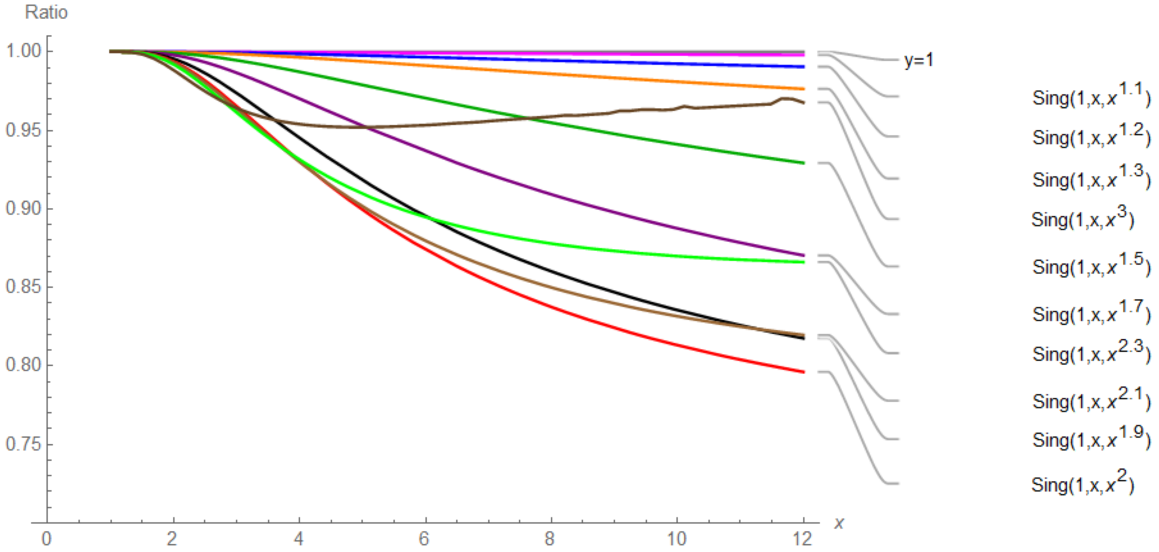

Figure 4.43: The relative jump ratio for $A \in \mathrm{Sing}(1, x, x^s)$, where $s > 1$. The function f is indicated by the labelling.

Now, we wnat to prove that the limit of the jump function for $A = \mathrm{Sing}(1, x, x^2)$ is equal to $1/\sqrt{2}$ as $a \longrightarrow \infty$. First, we substitute $b = a^2$ in Theorems 4.1.1 and 4.2.1 and find the best rank-one direction as $B = \mathbf{u} \otimes \mathbf{v}$ where

$$\mathbf{u} = (\frac{a-1}{\sqrt{2}\sqrt{(a^2+1)(a^2-a+1)}}, \frac{1}{\sqrt{2-\frac{2a}{a^2+1}}}, \frac{a(a^2-1)}{\sqrt{2}\sqrt{(a^2+1)(a^4+a^3+a+1)}}),$$

$$\mathbf{v} = (\frac{a-1}{\sqrt{2}\sqrt{(a^2+1)(a^2-a+1)}}, -\frac{1}{\sqrt{2-\frac{2a}{a^2+1}}}, \frac{a(a^2-1)}{\sqrt{2}\sqrt{(a^2+1)(a^4+a^3+a+1)}}).$$

Hence the linear distortion is

$$\mathcal{H}(A+tB) = a^2 - \frac{a(a-1)^3}{4(a^5+a^3+a^2+1)}t^2 + O(t^3). \tag{4.70}$$

However, we need to compute the actual eigenvalues and the points $\mathbf{t}_+$ and $\mathbf{t}_-$. These are

$$\mathbf{t}_+ = \frac{(a-1)(a^2+1)\Big(1-a\big(2-5a+a^3(5+(a-2)a)\big)+\sqrt{(a+1)^4(a^2+1)\big(1+a^2(7+a(a^3+7a-8))\big)}\Big)}{2+2a^2(a^5+6a^3+a^2+a+6)},$$

$$\mathbf{t}_- = \frac{(a-1)(a^2+1)\Big(1-a\big(2-5a+a^3(5+(a-2)a)\big)-\sqrt{(a+1)^4(a^2+1)\big(1+a^2(7+a(a^3+7a-8))\big)}\Big)}{2+2a^2(a^5+6a^3+a^2+a+6)}.$$

Notice that

$$\mathbf{t}_+ = 2a - 2 - \frac{8}{a} + \frac{6}{a^2} + O(\frac{1}{a^3}) \qquad \text{as } a \longrightarrow \infty,$$

and

$$\mathbf{t}_- = -a^2 + a - \frac{1}{a} + \frac{2}{a^2} + O(\frac{1}{a^3}) \qquad \text{as } a \longrightarrow \infty.$$

By substitution $b = a^2$ in Equations (4.63) and (4.64), we have

$$\lambda_1 = Z(a,t) - \frac{2^{\frac{1}{3}}Y(a,t)}{3 \times X(a,t)} + \frac{X(a,t)}{3 \times 2^{\frac{1}{3}} \times D(a)},$$

$$\lambda_2 = Z(a,t) + \frac{(1-i\sqrt{3}) \times Y(a,t)}{3 \times 2^{\frac{2}{3}} \times X(a,t)} - \frac{(1+i\sqrt{3}) \times X(a,t)}{3 \times 2^{\frac{4}{3}} \times D(a)},$$

$$\lambda_3 = Z(a,t) + \frac{(1+i\sqrt{3}) \times Y(a,t)}{3 \times 2^{\frac{2}{3}} \times X(a,t)} - \frac{(1-i\sqrt{3}) \times X(a,t)}{3 \times 2^{\frac{4}{3}} \times D(a)},$$

where $D(a) = -4(a^2+1)^2\big(1+a+a^2+(a-1)a^2+a^4\big)^2$ and $\lambda_1 \leq \lambda_2 \leq \lambda_3$. Actually note that all λ_i, terms are real with the same branch choices-solutions of a cubic equation with real roots. Each of the functions X, Y, Z, and D is a polynomial in a and t (both assumed large)or the square root of such. Now we have the actual eigenvalues and singular values. By Equation (4.69) the jump ratio is

$$\text{The jump ratio} = \frac{\max\{\mathcal{H}(A+t_{-}B), \mathcal{H}(A+t_{+}B)\}}{\mathcal{H}(A)}.$$

But we know that $H(A) = a^2$, so,

$$\text{The jump ratio} = \frac{\max\{\mathcal{H}(A+t_{-}B), \mathcal{H}(A+t_{+}B)\}}{a^2}.$$

Since t_+ and t_- are described by a, thus the jump ratio is a function with respect to a. If we calculate the Taylor series for each part of the jump ratio function as $a \longrightarrow \infty$, then

$$\text{The jump ratio} = \max\{\lim_{a\to+\infty} \frac{\mathcal{H}(A+t_{-}B)}{a^2}, \lim_{a\to+\infty} \frac{\mathcal{H}(A+t_{+}B)}{a^2}\}.$$

Let

$$t_{+}(a) = \frac{\mathcal{H}(A+t_{+}B)}{a^2}, \qquad t_{-}(a) = \frac{\mathcal{H}(A+t_{-}B)}{a^2}.$$

With regard to the variable a, both functions t_+ and t_- have very long equations. The limit of these functions can be calculated by using the following Taylor series

$$\sqrt{a^8 + 8a^6 - 8a^5 + 14a^4 - 8a^3 + 8a^2 + 1} = a^4 + 4a^2 - 4a - 1 + \frac{12}{a} + O(\frac{1}{a})^2, \tag{4.71}$$

and repeatedly removing the terms that are ineffective as $a \longrightarrow \infty$. Therefore

$$\lim_{a\to+\infty} t_{+}(a) = \lim_{a\to+\infty} \frac{\mathcal{H}(A+t_{+}B)}{a^2} = \lim_{a\to+\infty} \frac{\sqrt{\operatorname{Re}(\eta_1(a) + \eta_2(a))}}{a^2\sqrt{\operatorname{Re}(\eta_3(a) - \eta_4(a) - \eta_5(a))}}, \tag{4.72}$$

where (Re) is the real part of the expressions, and

$$\begin{aligned}
\eta_1(a) &= \frac{8a^{16} + 2^{2/3}(1 - i\sqrt{3})(a^7 + a^8)(-141a^{22} + 48a^{23} - 16a^{24})^{1/3}}{360a^{10} - 48a^{11} + 24a^{12}} \\
\eta_2(a) &= \frac{(1 + i\sqrt{3})\big(94a^{20}(a^2+1)^{2/3} - 8a^{21}(a^2+1)^{2/3} + 8a^{22}(a^2+1)^{2/3}\big)}{(-48a^{11} + 24a^{12})(-136a^{20} + 24a^{21} - 8a^{22})^{1/3}} \\
\eta_3(a) &= \frac{2a^{16} + 20a^{14} - 24a^{13} + 90a^{12} - 154a^{11}}{6a^{12} - 12a^{11} + 90a^{10} - 156a^9 + 450a^8} \\
\eta_4(a) &= \frac{808a^{18}(a^2+1)^{2/3} - 232a^{19}(a^2+1)^{2/3} + 112a^{20}(a^2+1)^{2/3} - 8a^{21}(a^2+1)^{2/3} + 8a^{22}(a^2+1)^{2/3}}{(900a^8 - 312a^9 + 180a^{10} - 24a^{11} + 12a^{12})(-1304a^{18} + 416a^{19} - 136a^{20} + 24a^{21} - 8a^{22})^{1/3}} \\
\eta_5(a) &= \frac{(a^8 + a^7 + 6a^6 - 3a^5)(-8a^{24} + 24a^{23} - 144a^{22} + 440a^{21} - 1440a^{20})^{1/3}}{6a^{12} - 12a^{11} + 90a^{10} - 156a^9 + 450a^8}.
\end{aligned} \tag{4.73}$$

Using the common denominator, simplifying algebraic expressions, the Taylor series, equation (4.72) can be written as follows

$$\lim_{a\to+\infty} t_+(a) = \lim_{a\to+\infty} \frac{\mathcal{H}(A+t_+B)}{a^2} = \lim_{a\to+\infty} \frac{\sqrt{\mathrm{Re}(\phi_1(a))}}{a^2\sqrt{\mathrm{Re}(\phi_2(a)-\phi_3(a)-\phi_4(a))}}, \tag{4.74}$$

where

$$\begin{aligned} \phi_1(a) &= \frac{16(-1)^{1/3}a^{52/3} + 4(1+i\sqrt{3})a^{16}(1+a^2)^{2/3}}{24(-1)^{1/3}a^{40/3}} \\ \phi_2(a) &= \frac{1}{3}(a^4+2a^3-a^2-18a+1) \\ \phi_3(a) &= \frac{3(1+a^2)^{2/3}(2a^{10}+2a^9+2a^8-32a^7+10a^6)}{(-1)^{1/3}(18a^{22/3}-18a^{19/3}+84a^{16/3}-138a^{13/3}+226a^{10/3})} \\ \phi_4(a) &= \frac{(-1)^{1/3}}{6}(a^2+3a-3-28a^{-1}-8a^{-2})(2a^2-2a+10-16a^{-1}+28a^{-2}). \end{aligned} \tag{4.75}$$

Now, It is easy to calculate the limit of Equation (4.74) as $a \longrightarrow \infty$:

$$\lim_{a\to+\infty} t_+(a) = \lim_{a\to+\infty} \frac{\mathcal{H}(A+t_+B)}{a^2} = \lim_{a\to+\infty} \frac{\sqrt{\mathrm{Re}(\phi_1(a))}}{a^2\sqrt{\mathrm{Re}(\phi_2(a)-\phi_3(a)-\phi_4(a))}} = \frac{1}{\sqrt{2}}. \tag{4.76}$$

By the same computation, we can prove that

$$\lim_{a\to+\infty} t_-(a) = \lim_{a\to+\infty} \frac{\mathcal{H}(A+t_-B)}{a^2} = \frac{1}{\sqrt{2}}. \tag{4.77}$$

So, the limit of the jump ratio function (as $a \longrightarrow \infty$) is

$$\text{The jump ratio} = \max\{\lim_{a\to+\infty} \frac{\mathcal{H}(A+t_-B)}{a^2}, \lim_{a\to+\infty} \frac{\mathcal{H}(A+t_+B)}{a^2}\} = \frac{1}{\sqrt{2}}.$$

This completes the proof of the limit.
These examples suggest if $A = (1, x, x^s)$ is a diagonal matrix then the largest jump can happen where $s = 2$. It seems the relative jump is

$$\text{The biggest relative jump} = \frac{\sqrt{2}}{2}.$$

when $x \longrightarrow \infty$. This means

$$\mathrm{Sing(1, x, x^2)} < \mathrm{Sing(1, x, x^s)},$$

for all $s > 1$ and $s \neq 2$. The graph of the jump ratio for diagonal matrix $A = (1, x, x^2)$ shows that the line $y = \frac{\sqrt{2}}{2}$ is a single horizontal asymptote and a lower bound. The function $f(x) = x^s$ can be replaced by other positive-valued functions.

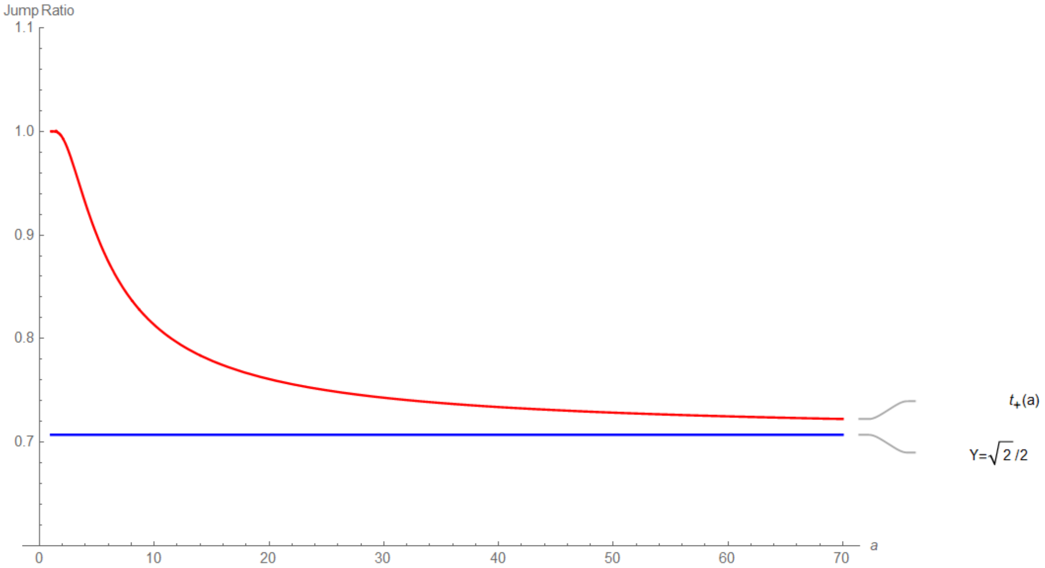

Figure 4.44: The relative jump ratio for $t_+(a)$ and its horizontal asymptote line $y = \frac{\sqrt{2}}{2}$ if $A = \text{Sing}(1, x, x^2)$.

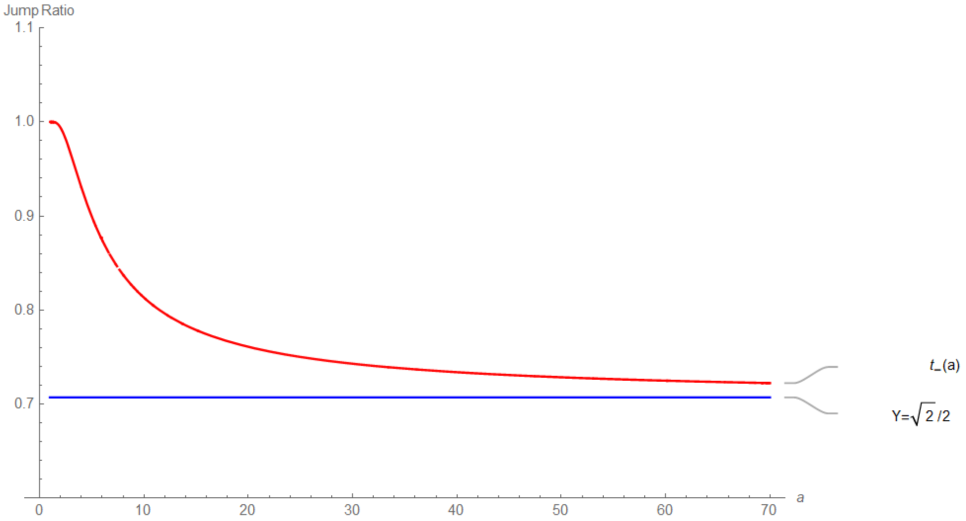

Figure 4.45: The relative jump ratio for $t_-(a)$ and its horizontal asymptote line $y = \frac{\sqrt{2}}{2}$ if $A = \text{Sing}(1, x, x^2)$.

For instance, f can be a linear function like $f(x) = x + q$ where q is a constant positive real number. For linear function $f(x) = x + q$, the graphs of the jump ratios show that with the bigger numbers of q the jump ratio tends to $y = \frac{\sqrt{2}}{2}$. So, if $q_1 < q_2$ then

$$\text{Sing}(1, x, x + q_1) > \text{Sing}(1, x, x + q_2),$$

when $x \longrightarrow \infty$. By increasing the number q, the relative jumps are smaller than before and the largest jump tends to $y = \frac{\sqrt{2}}{2}$. So, for different numbers q, the function f has different graphs as below:

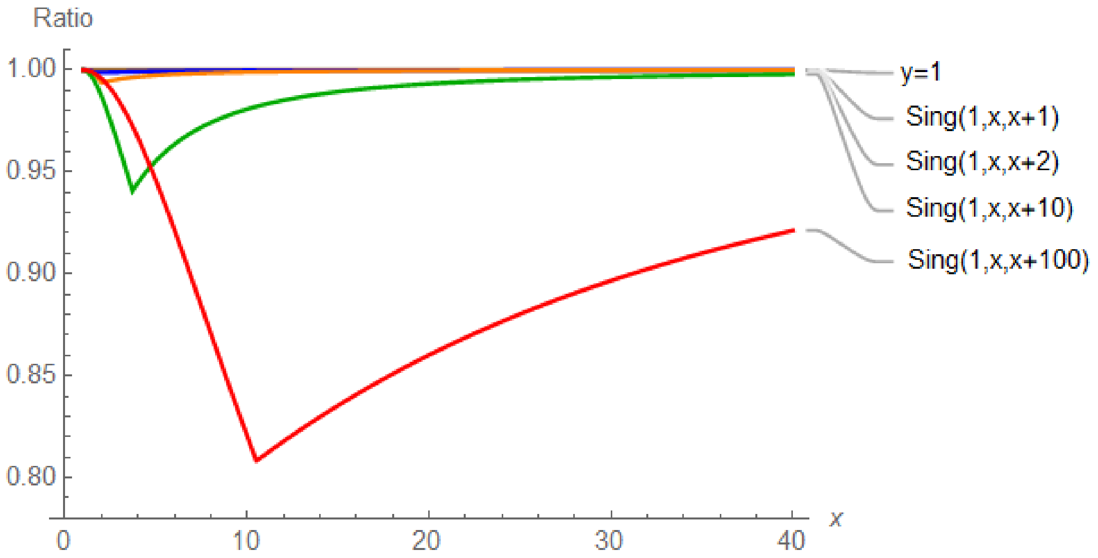

Figure 4.46: The relative jump ratio for $A \in \text{Sing}(1, x, x+q)$, where $q > 0$. The function f is indicated by the labelling.

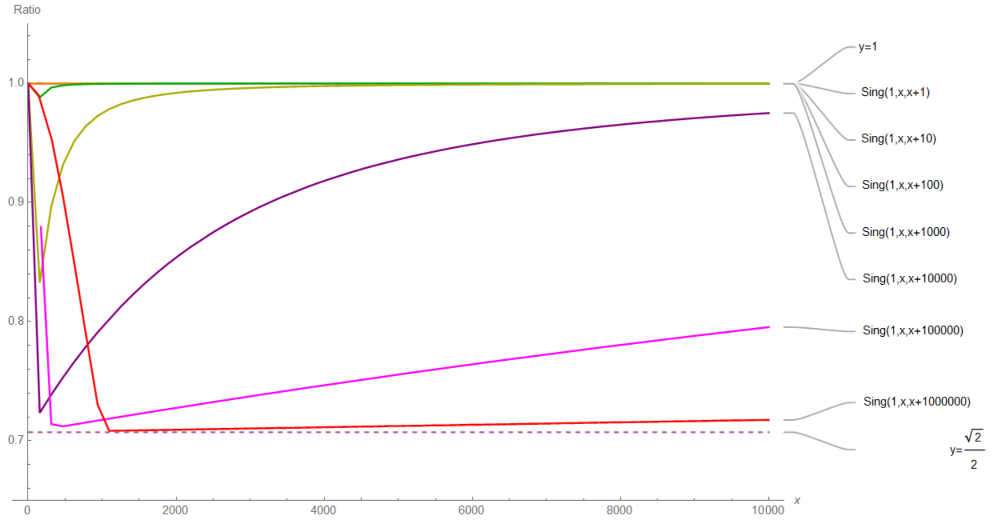

Figure 4.47: The relative jump ratio for $A \in \text{Sing}(1, x, x+q)$, where $q > 0$. The function f is indicated by the labelling.

These graphs also indicate the largest jump ratio for linear functions $f(x) = x + q$ with $q > 0$, tends to asymptote line $y = \frac{\sqrt{2}}{2}$.

If $f(x) = a^x$, then some relative jump functions are shown in Figure 4.48. If $a > 4$ then the graph of each relative jump has a minimum at $x = 2$, that gives the largest jump. If $a \longrightarrow +\infty$, the minimums of the relative jumps tends to $y = \frac{\sqrt{2}}{2}$ at $x = 2$.
These calculations show in the function $f(x) = a^x$, the minimum relative jump decreases faster than functions $f(x) = x^s$ and $f(x) = x + q$, but $f(x) \longrightarrow 1$, where $x \longrightarrow +\infty$.

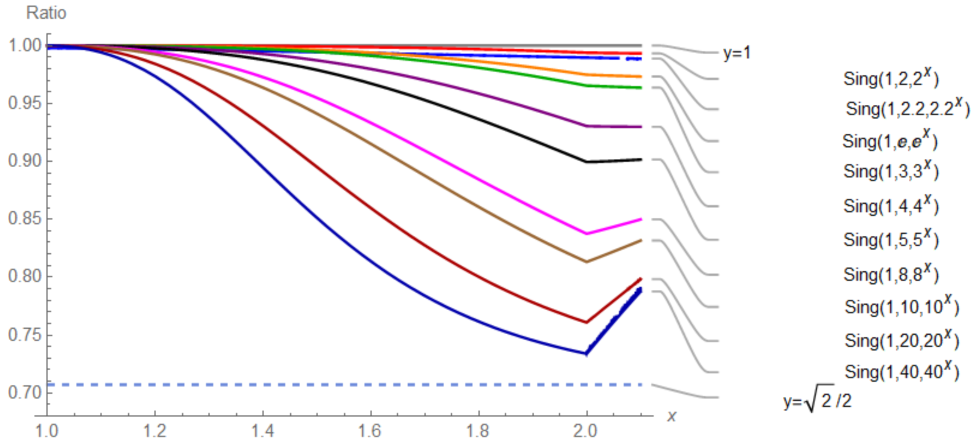

Figure 4.48: The relative jump ratio for $A \in \text{Sing}(1, a, a^x)$, where $x > 1$. The function f is indicated by the labelling.

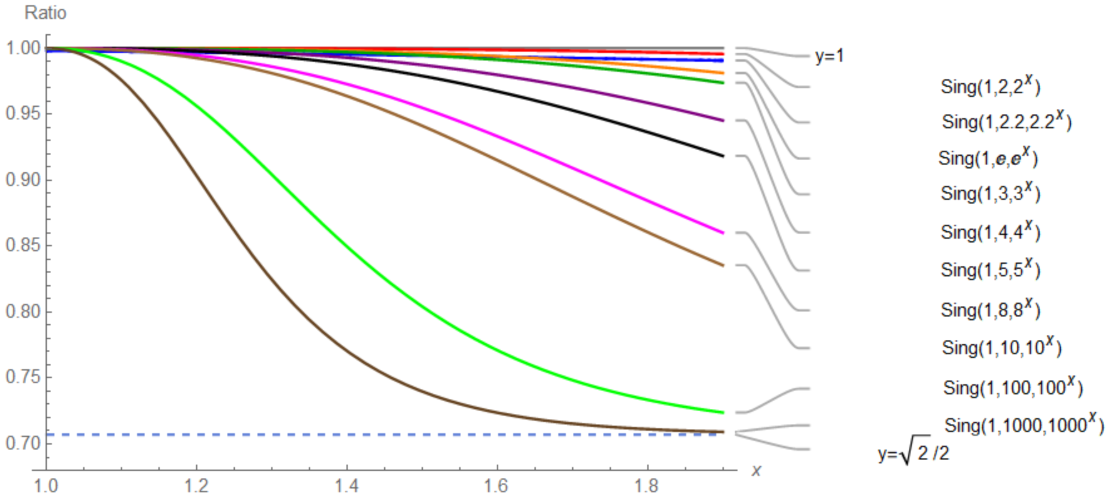

Figure 4.49: The relative jump ratio for $A \in \text{Sing}(1, a, a^x)$, where $x > 1$. The function f is indicated by the labelling.

A brilliant mathematician, Cauchy worked in many areas of mathematics. However, he is primarily known for his work on calculus and is recognized as one of the founders of the modern theory of mathematical analysis (see[65]). The functional equations that are particularly associated with Cauchy are the four equations in the table below.

Table 4.1: Cauchy's Functional Equations and Their Solutions.

No.	Functional Equations	Domains	Answers
1	$f(x+y) = f(x) + f(y)$	$x, y \in \mathbb{R}$	$f(x) = mx$, m is a constant.
2	$f(x+y) = f(x).f(y)$	$x, y \in \mathbb{R}$	$f(x) \equiv 0$, or $f(x) = a^x$, $a > 0$ is a constant.
3	$f(x.y) = f(x) + f(y)$	$x, y \in \mathbb{R}^+$	$f(x) = a \log \lvert x \rvert$, a is a constant.
4	$f(x.y) = f(x).f(y)$	$x, y \in \mathbb{R}^+$	$f(x) \equiv 0$, or $f(x) = \lvert x \rvert^a$, a is a constant.

We checked out the jump ratios and their minimums for all types of Cauchy's functional equations except the function $f(x) = \log|x|$. Note that the other three type were

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x^s \end{bmatrix}, \qquad A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x+q \end{bmatrix}, \qquad A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & a^x \end{bmatrix},$$

where $1 < x < f(x)$, $1 < x$, $a > 0$ and $a \neq 1$. But the last case is also one of these three types. Because by choosing

$$f(x) = \log|x| = g \qquad (\text{for all } \mathrm{x} > 1) \Longleftrightarrow \mathrm{x} = 10^{\mathrm{g}},$$

then the matrix A can be written as

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & f(x) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & \log x \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 10^g & 0 \\ 0 & 0 & g \end{bmatrix},$$

where $g > 1$. The two following matrices have the same eigenvalues and singular values.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 10^g & 0 \\ 0 & 0 & g \end{bmatrix}, \qquad \begin{bmatrix} 1 & 0 & 0 \\ 0 & g & 0 \\ 0 & 0 & 10^g \end{bmatrix}.$$

It is known that interchanging two rows in a diagonal matrix A does not change the linear distortion because elementary row operations in a matrix preserve the eigenvalues and of course, the singular values of A and linear distortion will be the same as before. So, we can say that the behaviour of Cauchy's functions for their relative jumps are identified as $x \longrightarrow +\infty$. Now, in the following graphs, we can compare all these types of Cauchy's functions.

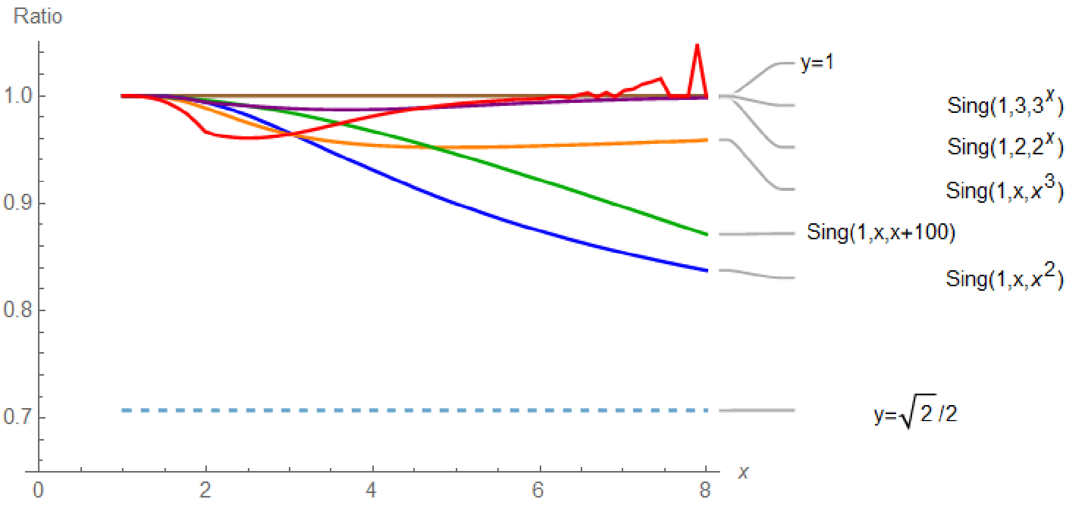

Figure 4.50: Comparison of the relative jumps for Cauchy's functions. The function f is indicated by the labelling.

The graphs above show if $f(x) = x^2$ then the relative jump tends to horizontal asymptote $y = \sqrt{2}/2$ as $x \longrightarrow +\infty$, in all types of Cauchy's functions.
However, we can not yet prove that this is best possible and that $A \in \mathrm{Sing}(1, x, x^2)$ is a critical case (see [34]). However we do have the following theorem.

Theorem 4.3.1. Let $m > \sqrt{2}$. Then there are a sequence of $\mathcal{H}$-quasiconformal self-mappings of $\mathbb{R}^3$ converging locally uniformly to an $\mathcal{H}_\infty$-quasiconformal self-mapping of $\mathbb{R}^3$ (and no smaller) with

$$m\,\mathcal{H} > \mathcal{H}_\infty$$

Proof. Suppose that

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a^2 \end{bmatrix},$$

and we define a linear map $\mathbf{f}_0 = A(\mathbf{x})$ that is a $\mathcal{H}_\infty$-quasiconformal self-mappings of $\mathbb{R}^3$. We define a sequence $\{\mathbf{f}_\nu\}_{\nu=1}^{\infty}$ by equation

$$\mathbf{f}_\nu(\mathbf{x}) = A\mathbf{x} + \frac{1}{\nu} h(\nu\langle \mathbf{u} \cdot \mathbf{x}\rangle)\mathbf{v},$$

where h is a periodic piecewise linear function on the real line as in Theorem (4.2.2). By the same process in the proof of Theorem (4.2.2), the following inequality holds,

$$\mathcal{H}(\mathbf{x}, \mathbf{f}_\nu) = \mathcal{H}(D\mathbf{f}_\nu(\mathbf{x})) \equiv H < \mathcal{H}_\infty(A) = \mathcal{H}_\infty(\mathbf{x}, \mathbf{f_0}).$$

So, the relative jump is

$$\text{The jump ratio} = \frac{\mathcal{H}(\mathbf{x}, \mathbf{f}_\nu)}{\mathcal{H}_\infty(\mathbf{x}, \mathbf{f_0})} > \frac{1}{\sqrt{2}} > \frac{1}{m}.$$

The inequality above can be written as

$$m\,\mathcal{H} > \mathcal{H}_\infty,$$

and this then proves the theorem. ♠

Theorem 4.3.1 is an existence theorem and we only showed that there is a sequence of $\mathcal{H}$-quasiconformal self-mappings of $\mathbb{R}^3$ converging locally uniformly to an $\mathcal{H}_\infty$-quasiconformal self-mapping of $\mathbb{R}^3$ so that the relative jump is bigger than or equal to $1/\sqrt{2} = \sqrt{2}/2$.
This leads us to a more general conjecture. In Theorem 4.2.1, we found the best rank-one direction matrix B_0 and the interval $[\mathbf{t}_-, \mathbf{t}_+]$ such that for all t in this interval we have

$$\max\{\mathcal{H}(A + \mathbf{t}_- B_0),\ \mathcal{H}(A + \mathbf{t}_+ B_0)\} < \mathcal{H}(A).$$

and for all individual linear distortion functions, the biggest jump was found. Theorem 4.2.2 shows that there always exists a sequence for $\mathcal{H}$-quasiconformal self-mappings of $\mathbb{R}^3$ converging locally uniformly to an $\mathcal{H}_\infty$-quasiconformal self-mapping of $\mathbb{R}^3$ so that

$$\mathcal{H}(\mathbf{x},\mathbf{T}_\nu) \equiv H < \mathcal{H}_\infty(\mathbf{x},\mathbf{T_0}) \qquad \text{for all } \nu = 1,2,3,\ldots.$$

Conjecture: Let $\{\mathbf{f}_\nu\}_{\nu=1}^{\infty}$ be a sequence of quasiconformal self-mappings of $\mathbb{R}^n$ converging locally uniformly to a quasiconformal self-mapping $\mathbf{f_0}$ of $\mathbb{R}^n$. Suppose that $\mathcal{H}(\mathbf{f}_\nu) \longrightarrow H$, as $\nu \longrightarrow +\infty$, then for the biggest jump or relative jump, the following inequality holds:

$$\frac{\mathcal{H}(\mathbf{f}_\nu)}{\mathcal{H}(\mathbf{f_0})} \longrightarrow \frac{H}{\mathcal{H}(\mathbf{f_0})} \geq \frac{1}{\sqrt{2}}.$$

In other words, $mH \geq \mathcal{H}(\mathbf{f_0})$, for all $m > \sqrt{2}$.

Gehring and Iwaniec showed in [32], that the limit mapping $\mathbf{f}$ of a weakly convergent sequence of mappings $\mathbf{f}_\nu$ with finite distortion also has finite distortion and give several dimension-free estimates for the distortion of $\mathbf{f}$.

Their arguments are based on the weak continuity of the Jacobian determinants and the concept of polyconvexity (see [32]). They presented the following theorem.

Theorem 4.3.2. Suppose that $\mathbf{f}_\nu : \Omega \subset \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a sequence of mappings of finite distortion which converges weakly in $W^{1,n}_{\text{loc}}$ to $\mathbf{f}$ and suppose that

$$\mathcal{H}(\mathbf{x},\mathbf{f}_\nu) \leq H < \infty \qquad \text{for } \nu = 1,2,3,...$$

almost everywhere in Ω. Then $\mathbf{f}$ has finite distortion and

$$\mathcal{H}(\mathbf{x},\mathbf{f}) \leq \frac{1}{2}(H + H^{n-1})^{\frac{2}{n}} \leq H^{2-\frac{2}{n}},$$

almost everywhere in Ω.

If $n = 3$, then as we said

$$\mathcal{H}(\mathbf{x},\mathbf{f}) \leq \frac{1}{2}(H + H^2)^{\frac{2}{3}}.$$

Notice that the ratio of the jump here is

$$\frac{\frac{1}{2}(H + H^2)^{\frac{2}{3}}}{H} \sim \frac{1}{2}H^{\frac{1}{3}} \qquad \text{as} \quad H \longrightarrow \infty.$$

Our results in the conjecture above and Theorem 4.3.1 suggest that

$$\mathcal{H}(\mathbf{x},\mathbf{f}) \leq \sqrt{2}\,H,$$

might be the best possible limit for large H.

4.4 Solving Problem 3.

Finally problem 3 seeks to put the previous two problems in a more general framework.

Theorem 4.4.1. Let $C \in GL(3,\mathbb{R})$ and suppose that C has distinct singular values (so C^tC has distinct real eigenvalues). Then

1. there is a best rank-one direction B^* (as per Problem 1) for C,
2. there are two numbers $\mathbf{t}_-, \mathbf{t}_+$ as per Problem 2, and also $\mathcal{H}(C+\mathbf{t}_-B^*)$ and $\mathcal{H}(C+\mathbf{t}_+B^*)$, that is the largest interval of concavity for $\mathcal{H}(C+tB^*)$ and the smallest values for the linear distortion of an approximation.

Proof. Let $C \in GL(3,\mathbb{R})$ be a 3×3 matrix with three distinct singular values $\sigma_1 < \sigma_2 < \sigma_3$. By the **singular value decomposition (SVD)**, Theorem 1.6.10, the matrix C can be written

$$C = UAV^t,$$

where U and V^t are 3×3 orthogonal matrices and

$$A = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \sigma_2 & 0 \\ 0 & 0 & \sigma_3 \end{bmatrix},$$

where σ_1, σ_2 and σ_3 are singular values of the matrix C. We want to determine the best rank-one direction matrix B^* for C such that

$$\left.\frac{d}{dt}\right|_{t=0}\mathcal{H}(C+t\,B^*) \;=\; 0 \tag{4.78}$$

$$\left.\frac{d^2}{dt^2}\right|_{t=0}\mathcal{H}(C+t\,B^*) \;\le\; \left.\frac{d^2}{dt^2}\right|_{t=0}\mathcal{H}(C+t\,B) \tag{4.79}$$

We claim that the best rank-one direction B^* is

$$B^* = UB_0V^t,$$

where U and V^t are two orthogonal matrices from singular value decomposition for C and B_0 is best rank-one direction for the diagonal matrix A from Problem 1. Because

$$\mathcal{H}(C+tB^*) = \mathcal{H}(UAV^t + t\,UB_0V^t) = \mathcal{H}(U(A+tB_0)V^t) \tag{4.80}$$

But by Equations (1.21) and (1.26) we have

$$\mathcal{H}(C+tB^*) = \mathcal{H}(A+tB_0), \tag{4.81}$$

where B_0 is the best rank-one direction matrix that is found in problem (1) and $t \in [\mathbf{t}_-, \mathbf{t}_+]$ that is found in problem (2). So, the best rank-one direction matrix for matrix C is the matrix $B^* = UB_0V^t$ and the real value t belongs to the interval $[\mathbf{t}_-, \mathbf{t}_+]$ and then

$$\mathcal{H}(C+tB^*) = H < \mathcal{H}(C) \qquad \text{for all t} \in [\mathbf{t}_-, \mathbf{t}_+].$$

We remark that the matrices U and V^t in the Singular Value Decomposition Theorem (SVD) for matrix $C = UAV^t$ are orthogonal matrices. They rotate the domain and do not have any affect on linear distortion (see Figure 1.4). ♠

Example 4.4.1. Let

$$C = \begin{bmatrix} 2 & -1 & 1 \\ -1 & \frac{5}{2} & -\frac{1}{2} \\ 1 & -\frac{1}{2} & \frac{5}{2} \end{bmatrix} \qquad \Longrightarrow \qquad C^tC = \begin{bmatrix} 6 & -5 & 5 \\ -5 & \frac{15}{2} & -\frac{7}{2} \\ 5 & -\frac{7}{2} & \frac{15}{2} \end{bmatrix}.$$

The 3×3 matrix C^tC is symmetric and hence can be orthogonally diagonalised by the Spectral Theorem. Not only are the eigenvalues of C^tC all real, they are all nonnegative. The singular values of C are the square roots of the eigenvalues of C^tC. By the Singular Value Decomposition (Theorem 1.6.10) the matrix C can be factorised as $C = UAV^t$, where the matrices U and V^t are orthogonal matrices and A is a diagonal matrix with singular values of C. So,

$$C = \begin{bmatrix} 2 & -1 & 1 \\ -1 & \frac{5}{2} & -\frac{1}{2} \\ 1 & -\frac{1}{2} & \frac{5}{2} \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{3}}{3} & 0 & -\frac{\sqrt{6}}{3} \\ -\frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} & -\frac{\sqrt{6}}{6} \\ \frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} & \frac{\sqrt{6}}{6} \end{bmatrix} \cdot \begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \frac{\sqrt{3}}{3} & -\frac{\sqrt{3}}{3} & \frac{\sqrt{3}}{3} \\ 0 & \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{6}}{3} & -\frac{\sqrt{6}}{6} & \frac{\sqrt{6}}{6} \end{bmatrix} = UAV^t.$$

Since the matrices U and V^t are orthogonal matrices, then

$$\mathcal{H}(C) = \mathcal{H}(UAV^t) = \mathcal{H}(A).$$

The best rank-one direction matrix B^* for the matrix C can be calculated by

$$\mathcal{H}(C + tB^*) = \mathcal{H}(UAV^t + t\, UB_0V^t) = \mathcal{H}(U(A + tB_0)V^t) = \mathcal{H}(A + tB_0).$$

But by Examples 4.1.1 and 4.2.1 the best rank-one direction matrix B_0 is

$$B_0 = \begin{bmatrix} \frac{1}{30} & -\frac{1}{6} & \frac{1}{15} \\ \frac{1}{6} & -\frac{5}{6} & \frac{1}{3} \\ \frac{1}{15} & -\frac{1}{3} & \frac{2}{15} \end{bmatrix}.$$

For all t in the interval $[\mathbf{t}_-, \mathbf{t}_+] = [-2.04484, 1.19219]$, we have

$$\mathcal{H}(C + tB^*) = \mathcal{H}(A + tB_0) = 4 - \frac{1}{90}t^2 < \mathcal{H}(C) = \mathcal{H}(A) = 4,$$

where

$$B^* = UB_0V^t = \begin{bmatrix} -\frac{4\sqrt{2}-9}{90} & \frac{6+2\sqrt{2}+20\sqrt{3}-5\sqrt{6}}{180} & \frac{-6-2\sqrt{2}+20\sqrt{3}-5\sqrt{6}}{180} \\ \frac{6+2\sqrt{2}-20\sqrt{3}+5\sqrt{6}}{180} & \frac{4\sqrt{2}-69}{180} & \frac{-81-4\sqrt{2}+20\sqrt{3}+10\sqrt{6}}{180} \\ \frac{-6-2\sqrt{2}-20\sqrt{3}+5\sqrt{6}}{180} & -\frac{81+4\sqrt{2}+20\sqrt{3}+10\sqrt{6}}{180} & \frac{4\sqrt{2}-69}{180} \end{bmatrix}.$$

The biggest jump in this example is

$$\min\{|\mathcal{H}(C + \mathbf{t}_\pm B^*) - \mathcal{H}(C)|\} = \min\{|\mathcal{H}(A + \mathbf{t}_\pm B_0) - \mathcal{H}(A)|\} = 0.02461.$$

♣

Chapter 5

Conclusion and Summary

Figure 5.1: Conclusion and Summary
Source: https://pixabay.com/en/photos/result/

Chapter 5

Conclusion and Summary

The central problem discussed in this thesis concerns the semicontinuity properties of the distortion functional for quasiconformal mappings. In particular, if $\{\mathbf{f_n}\}_{n=1}^{\infty}$ is a sequence of quasiconformal mappings of $\hat{\mathbb{R}}^n$ onto itself and

$$\mathbf{f_n} \longrightarrow \mathbf{f}, \tag{5.1}$$

uniformly on $\hat{\mathbb{R}}^n$, we ask for which distortion functionals $\mathcal{K} = \mathcal{K}(\mathbf{g})$, defined for a quasiconformal mappings $\mathbf{g}$ of $\hat{\mathbb{R}}^n$ do we have the lower semicontinuity property

$$\mathcal{K}(\mathbf{f}) \leq \liminf_{n\to\infty} \mathcal{K}(\mathbf{f_n}), \tag{5.2}$$

so that the limit $\mathbf{f}$ is itself a $\mathcal{K}$-quasiconformal mapping of $\hat{\mathbb{R}}^n$ onto itself. For instance, $\mathcal{K}$ could be the linear distortion, maximal distortion, outer or inner distortion and etc. This lower semicontinuity property is related to issues of convexity of the functional $\mathcal{K}$ defined on the space of mappings or more precisely the pointwise differentials of these mappings.

5.1 Summary of Thesis Problems

In this thesis, we discussed three different notions of convexity, namely polyconvexity, quasiconvexity, and rank-one convexity and their known (and unknown) relationships with one other. We then discussed three different distortion functions, which give rise to the same class of mappings - namely the quasiconformal mappings - but with possibly different associated constants. Each such functional gave rise to a different definition of quasiconformality. These three different definitions were:

1. the geometric definition via the distortion of the modulus of curve families
2. the metric definition via the linear distortion, and
3. the analytic definition defined via a differential inequality.

We provided references which show that the functional determined by the distortion of modulus of curve families and also that determined by the pointwise differential inequality do in fact have the lower semicontinuity property given at (5.2). You can see the Theorems 3.3.2 to 3.4.2. For the linear distortion the question of lower semicontinuity has been answered negatively by Tadeusz Iwaniec (see [24]). Theorem 3.4.2 is proved in his paper (see [39]). This asserts that there is a sequence of quasiconformal mappings as at (5.1) such that for the linear distortion

$$\mathcal{H}(\mathbf{f}) > \lim_{n\to\infty} \mathcal{H}(\mathbf{f_n}). \tag{5.3}$$

Notice that the limit here does exist and the limit mapping $\mathbf{f}$ above must be quasiconformal (and so $H(\mathbf{f}) < \infty$) since the other two distortion functionals are controlled by the linear distortion functional. Since the very beginning of the multidimensional theory of quasiconformal mappings, it has been widely believed that the class of K-quasiconformal mappings in $\mathbb{R}^n$ is closed concerning uniform convergence, where K stands for the linear distortion. In article [39], T. Iwaniec proved one lemma and a theorem that refutes this belief. The key element of his construction is that the linear distortion function fails to be rank-one convex in dimensions $n \geq 3$.

Lemma 5.1.1. (Iwaniec's Lemma) Given $n \geq 3$ and $H > 1$, there is a matrix A and a rank-one matrix B and numbers $t, s > 0$ such that

$$\mathcal{H}(A - sB) = \mathcal{H}(A + tB) = H < \mathcal{H}(A).$$

Theorem 5.1.2. (Iwaniec's Theorem) For each $n \geq 3$ and $H > 1$, there exists a sequence $\{\mathbf{f_n}\}_{n=1}^{\infty}$ of quasiconformal mappings $\mathbf{f_n} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ converging uniformly to a linear quasiconformal map $\mathbf{f_0} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ such that

$$\mathcal{H}(x, \mathbf{f_n}) \equiv H < \mathcal{H}(x, \mathbf{f_0}), \qquad \text{almost everywhere in } \mathbb{R}^n, \quad n = 1, 2, \ldots.$$

The linear distortion functional is defined pointwise from the differential matrix when it is nonsingular. At the points $x_0 \in \mathbb{R}^3$ of differentiability of a quasiconformal mapping a local analysis of the convexity properties of the linear distortion can be given if we only consider the nonsingular linear mappings $x \mapsto Tx$, $T = D\mathbf{f}(x_0) \in GL(3, \mathbb{R})$. To compute the linear distortion we study the singular values of

$$A = T^t T \in M_{3\times 3}(\mathbb{R}),$$

the space of symmetric positive definite 3×3 matrices. Given such an A the Spectral Theorem tells us A is orthogonally diagonalisable and so we may as well suppose it is diagonal since we can diagonalise it by orthogonal (conformal) transformations (at a point). In this way, we reduce the problem of the convexity of the linear distortion functional to considering that functionally defined on the space of 3×3 positive definite diagonal matrices $A = (a_{ij})$, $a_{ij} = 0$ if $i \neq j$, all the diagonal entries a_{ii} are positive and we can further suppose that $1 = a_{11} \leq a_{22} \leq a_{33}$, by scaling and a further conjugation by orthogonal matrices - neither of which affects the linear distortion. Now, we can extend the Iwaniec's Lemma and Theorem to (Lemma 3.4.1 and Theorem 3.4.2). Both of his lemma and theorem are existence theorems. But we expand them to a wide range of matrices. The results of Problem (1) and Problem (2) in the previous section and this section show it is always possible to find a rank-one matrix B_0 so that

$$\mathcal{H}(A + t\, B_0) = \mathcal{H}(A + t\, \mathbf{u} \otimes \mathbf{v}) < \mathcal{H}(A), \tag{5.4}$$

holds for a diagonal matrix A with distinct positive singular values. The rank-one matrix B_0 is the best rank-one direction for which the linear distortional function at A is most concave. An unsolved and unanswered question (Partly posed by Frederick W. Gehring and T. Iwaniec) was to determine how big the difference between the left-hand and the right-hand side of inequality (5.4) can be?

In problems (1) and (2), not only do we find best direction that the inequality (5.4) holds but we partly answered the Gehring and Iwaniec question and found the biggest jump and difference between the left-hand and the right-hand side of inequality (5.4) for a specific diagonal matrix A with distinct singular values. Iwaniec's lemma gives an example (actually counterexample) that the inequality (5.4) happened.

Theorem 5.1.3. Let

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & b \end{bmatrix}$$

be diagonal with $1 < a < b$. Then there are two unit vectors $\mathbf{u}$ and $\mathbf{v}$ in $\mathbb{R}^3$ described at (4.54), and a rank-one matrix $B_0 = \mathbf{u} \otimes \mathbf{v}$ (the best rank-one direction) and two real numbers $\mathbf{t}_- < 0$ and $\mathbf{t}_+ > 0$, the roots of $P(t)$ as a (4.58), such that

$$\frac{\mathrm{d}}{\mathrm{dt}}\Big|_{t=0} \mathcal{H}(A + t\,B_0) = 0 \tag{5.5}$$

$$\frac{\mathrm{d}^2}{\mathrm{dt}^2}\Big|_{t=0} \mathcal{H}(A + t\,B_0) \leq \frac{\mathrm{d}^2}{\mathrm{dt}^2}\Big|_{t=0} \mathcal{H}(A + t\,B) \leq 0 \tag{5.6}$$

for every rank-one matrix $B \in \mathbb{R}^{3\times 3}$ with $\frac{\mathrm{d}}{\mathrm{dt}}\Big|_{t=0} \mathcal{H}(A + t\,B) = 0$. Also, there is a largest interval $(\mathbf{t}_-, \mathbf{t}_+)$ so that $\mathcal{H}(A + t\,B_0)$ is a smooth function of t in this interval and

$$\max\{\mathcal{H}(A + \mathbf{t}_- B_0),\ \mathcal{H}(A + \mathbf{t}_+ B_0)\} < \mathcal{H}(A).$$

The jump and difference is defined to be

$$\textbf{Maximum jump} = \min\{|\mathcal{H}(\mathrm{A} + \mathbf{t}_+ \mathrm{B}_0) - \mathcal{H}(\mathrm{A})|, |\mathcal{H}(\mathrm{A} + \mathbf{t}_- \mathrm{B}_0) - \mathcal{H}(\mathrm{A})|\},$$

Recall that the value of $\mathcal{H}(A)$ is b.

Now we can extend the Theorem 5.1.2 to the general case in $\mathbb{R}^3$. In this theorem, we construct a sequence $\{\mathbf{T}_\nu\}_{\nu=1}^{\infty}$ of uniformly bi-Lipschitz maps $\mathbf{T}_\nu : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$, that is,

$$\frac{1}{C}|\mathbf{x} - \mathbf{y}| \leq |\mathbf{T}_\nu(\mathbf{x}) - \mathbf{T}_\nu(\mathbf{y})| \leq C|\mathbf{x} - \mathbf{y}|$$

for all $\mathbf{x}$, $\mathbf{y} \in \mathbb{R}^3$ and $\nu = 1,\ 2,\ 3, \ldots$with a constant C independent of $\mathbf{x}$, $\mathbf{y}$ and ν. The linear distortion function of each $\mathbf{T}_\nu$ turns out to be constant and equal to H almost everywhere:

$$\mathcal{H}(\mathbf{x}, \mathbf{T}_\nu) = H, \qquad\qquad \nu = 1,\ 2, \ldots.$$

The sequence $\{\mathbf{T}_\nu\}$ converges uniformly to a linear map $\mathbf{T_0} : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ whose distortion is bigger than H (see [39], [42], [30]).

Theorem 5.1.4. (Generalisation of Iwaniec's Theorem): If $n = 3$ and $H > 1$ and $\mathbf{T_0} : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ with $\mathbf{T_0}(\mathbf{x}) = A\mathbf{x}$ such that A is a 3×3 diagonal matrix with distinct singular values, then there **always exists** a sequence $\{\mathbf{T}_\nu\}_{\nu=1}^{\infty}$ of quasiconformal mappings $\mathbf{T}_\nu : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ converging uniformly to a linear quasiconformal map $\mathbf{T_0}$ such that

$$\mathcal{H}(\mathbf{x}, \mathbf{T}_\nu) \equiv H < H_0 \equiv \mathcal{H}(\mathbf{x}, \mathbf{T_0}),$$

almost everywhere in $\mathbb{R}^3$ and $\nu = 1, 2, \ldots$.

We showed if

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & f(x) \end{bmatrix},$$

and $f(x) = x^2$, then the relative jump tends to horizontal asymptote $y = \sqrt{2}/2$ as $x \longrightarrow +\infty$, in all types of Cauchy's functions. However we can not yet prove that this is best possible and that $A \in Sing(1, x, x^2)$ is a cirtical case (see [34]). However we do have the following theorem.

Theorem 5.1.5. Let $m > \sqrt{2}$. Then there are a sequence of $\mathcal{H}$-quasiconformal self-mappings of $\mathbb{R}^3$ converging locally uniformly to an $\mathcal{H}_\infty$-quasiconformal self-mapping of $\mathbb{R}^3$ (and no smaller) with

$$m\,\mathcal{H} > \mathcal{H}_\infty$$

Finally the last theorem that we have proved, seeks to put the previous two Theorems 5.1.3 and 5.1.4 in a more general framework.

Theorem 5.1.6. Let $C \in GL(3, \mathbb{R})$ and suppose that C has distinct singular values (so C^tC has distinct real eigenvalues). Then

1. there is a best rank-one direction B^* (as per Problem 1) for C,
2. there are two numbers $\mathbf{t}_-, \mathbf{t}_+$ as per Problem 2, and also $\mathcal{H}(C + \mathbf{t}_- B^*)$ and $\mathcal{H}(C + \mathbf{t}_+ B^*)$, that is the largest interval of concavity for $\mathcal{H}(C + tB^*)$ and the smallest values for the linear distortion of an approximation.

5.2 Potential Applications

Aspects of material science: The use of nonlinear elasticity to describe martensitic transformations and their microstructure starts with J.M. Ball and R.D. James (1987) [13], following the work of many authors applying nonlinear elasticity to crystals, especially J.L. Ericksen. There is a "linearized" version of the theory due to Khachaturyan and Roitburd. The lecture notes of Ball [8] gives a very good overview of the subject at an introductory level. In fact, one of the most studied phase transitions in nonlinear materials science is the Martensitic transition, the diffusionless structural transition from a high symmetry to a lower symmetry crystallographic phase. This transition is often induced by changing temperature or stress and shows athermal character and proceeds intermittently as a sequence of avalanches ultimately producing a complex multiscale microstructure.
Recently the theory of mappings of finite distortion has emerged and shown promise in modelling various aspects of nonlinear materials science with interesting associated extremal problems, see for example [6], [7], [37], [41], [42], [43], [45], [44] and [50]. Many of the distortion functionals studied in these cases are lower semicontinuous and extremals are quite regular - often diffeomorphisms. However, the images below suggest that this cannot be the case for the sort of phase transitions occurring in these structural transformations.

Thus we analyse a slightly different (but still well known) distortion functional where we lose polyconvexity and hope to gain models of non-smooth extrema. Here we are only concerned

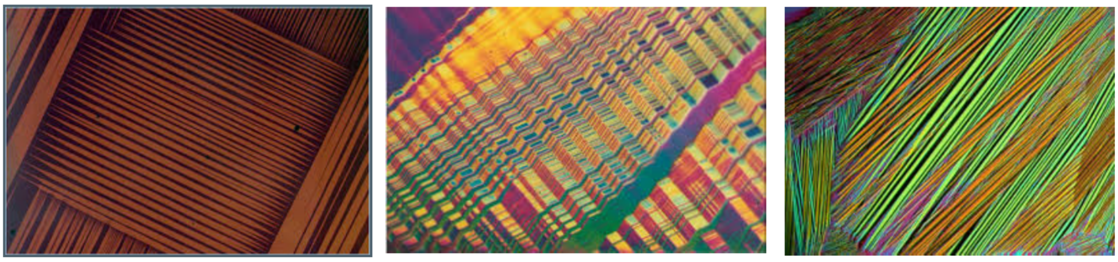

Figure 5.2: These images are from Chu and James's experiments in Cu-Al-Ni single crystals, and M. Morin

with static models but are well aware that the real problem is dynamical. We hope to address this elsewhere. Distortion functionals are polyconvex in two dimensions, so it is unlikely we will see these complicated micro-scale structures with purely two-dimensional models using distortion. However, an interesting feature of our models is that although fundamentally three-dimensional, limiting processes yield two-dimensional and even one-dimensional models that at least suggest connections. Distortion functionals are scale-invariant measures of the anisotropic nature of a deformation. Consider deforming a body $\Omega \subset \mathbb{R}^3$ by a homeomorphism $\mathbf{f} : \Omega \to f(\Omega) \subset \mathbb{R}^3$. Typically the assumption that deformation is a homeomorphism is given to us by the principle of inter penetrability of matter, see [10]. It is a remarkable feature of mappings of bounded distortion, coming from the modulus of continuity estimates on mappings and their inverses, that this topological condition is retained under limits (see [34]). Let $\Omega \subset \mathbb{R}^n$, $n \geq 3$, be a domain and $\mathbf{f} : \Omega \longrightarrow \mathbb{R}^n$ be a homeomorphism. Typically one assumes some regularity of the deformations in question, namely

$$\mathbf{f} \in W^{1,3}_{loc}(\Omega),$$

the Sobolev space of mappings whose first derivatives are in L^3. This condition is basically the minimal assumption one can make to ensure that the Jacobian determinant is locally integrable and that something like the change of variables formula might hold [42]. The distortion inequality assumes that the full differential of mapping is controlled by the Jacobian; there is a constant $K < \infty$ such that

$$\|\mathbf{f}'(\mathbf{x})\|^2 \leq K.J(\mathbf{x}, \mathbf{f}), \qquad \text{for almost every } \mathbf{x} \in \Omega. \tag{5.7}$$

Such mappings are called quasiconformal. Deformations satisfying the distortion inequality enjoy many properties, including higher regularity

$$f \in W^{1,3+\epsilon(K)}_{loc},$$

positive Jacobian $J(\mathbf{x}, \mathbf{f}) > 0$ almost everywhere and the change of variables formula [42]. Thus one can define the distortion function

$$\mathbf{K}(\mathbf{x}, \mathbf{f}) = \frac{\|\mathbf{f}'(\mathbf{x})\|^2}{J(\mathbf{x}, \mathbf{f})} \tag{5.8}$$

It is not obvious, but not too hard to see that $\mathbf{K}(\mathbf{x}, \mathbf{f})$ is a convex function of the minors of $Df(\mathbf{x})$. Thus one can consider various extremal problems, for instance, minimising the L^p-norms of $\mathbf{K}(\mathbf{x}, \mathbf{f})$ among deformations with prescribed boundary values linking with the calculus of variations. Some novel phenomena arise in such problems including the Nitsche phenomena where [7] geometric obstructions preclude the existence of a minimiser. However, it is expected that minimisers are smooth when they exist (and they should exist for all $p > 1$) if there is a candidate "barrier". For $\mathbf{x} \in \Omega$, we set

$$\mathcal{H}_{\mathbf{f}}(\mathbf{x}) = \mathcal{H}(\mathbf{x}, \mathbf{f}) = \limsup_{r \to 0} \frac{\max_{\|h\|=r} \|\mathbf{f}(\mathbf{x}+h) - \mathbf{f}(\mathbf{x})\|}{\min_{\|h\|=r} \|\mathbf{f}(\mathbf{x}+h) - \mathbf{f}(\mathbf{x})\|}. \tag{5.9}$$

Let Ω and Ω' be domains in $\mathbb{R}^3$, and let $\mathbf{f}$ be a homeomorphism of Ω onto Ω'. The linear distortion of function $\mathbf{f}$ is defined

$$\mathcal{H}(\mathbf{f}) = \sup\{\mathcal{H}_{\mathbf{f}}(\mathbf{x}) : \ \mathbf{x} \in \Omega\}, \tag{5.10}$$

We know that if $\mathcal{H}(\mathbf{x}, \mathbf{f})$ is bounded in Ω, then $\mathbf{f}$ is K-quasiconformal for some K bounded by a function of $\mathcal{H}$ (See [34]). In particular $\mathbf{f}$ lies in the Sobolev space $W^{1,3}_{loc}(\Omega)$ and satisfies the **distortion inequality**,

$$\max_{|\zeta|=1} |D\mathbf{f}(\mathbf{x})\zeta| \le \mathcal{H}(\mathbf{f}) \min_{|\zeta|=1} |D\mathbf{f}(\mathbf{x})\zeta|.$$

Geometrically this means that the differential $D\mathbf{f}(\mathbf{x}) : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ maps the unit sphere to an ellipsoid. The converse is also true, but a remarkable theorem of Heinonen and Koskela (1995) shows that one only requires the lim inf in (5.9) to gain quasiconformality (with the same constants).

A natural question is: Can these scale-invariant problems and Martensite transitions be modelled by distortion functionals? This relationship (if it exists at all) is predicated that at molecular scales we see piecewise linear mappings of uniformly bounded distortion as illustrated below. There are also some conjectural ideas regarding boundary values and questions such as are local scales determined by **"interfacial energy"** ?

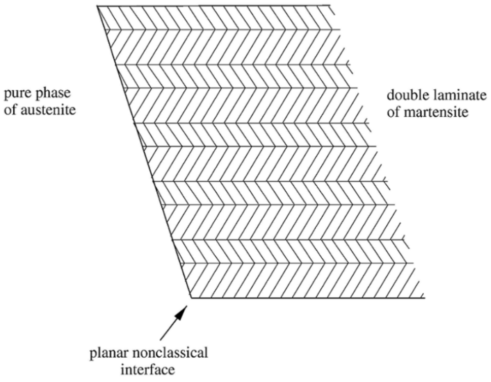

Figure 5.3: Nonclassical interface with double laminate [13].

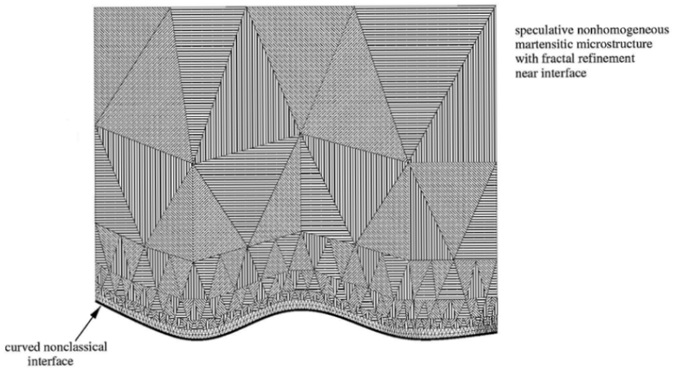

Figure 5.4: Nonclassical austeite-martensite interface [13].

Recall that a function $\mathcal{F} : U \subset \mathbb{R}^{n\times n} \longrightarrow \mathbb{R}$, defined on an open set U of $n \times n$ matrices, is a rank-one convex at $A_0 \in U$ if for every rank-one matrix $X \in \mathbb{R}^{n\times n}$ the function

$$t \mapsto \mathcal{F}(A_0 + tX)$$

is convex near $t = 0$. By Theorem 1.8.1, a matrix X has rank one if and only if it is the tensor product of two vectors $X = u \otimes v$, where $u, v \in \mathbb{R}^n$ (See [34]).
The Hadamard jump condition asserts that the piecewise linear function $F = \{A, B\}$ is continuous across the interface if and only if $A - B = u \otimes v$, and the normal to the interface is v.

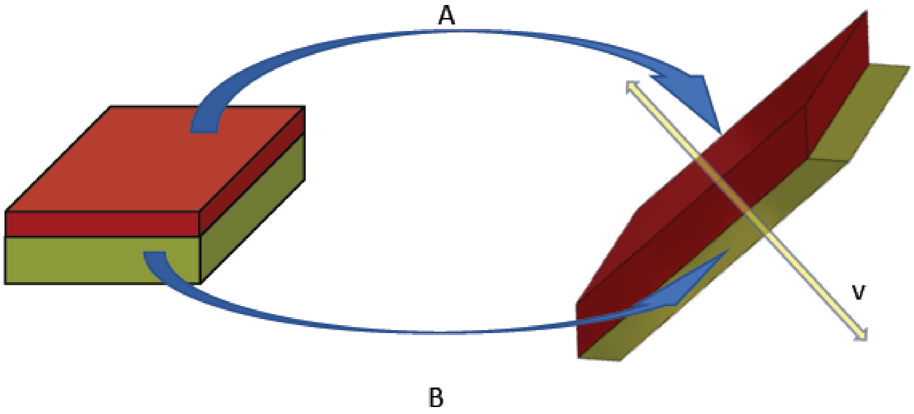

Figure 5.5: Rank-one connected planar interface

5.3 Further Research

The results above show that it is always possible to find a rank-one matrix B so that (5.4) holds for a diagonal matrix A with distinct positive eigenvalues. An interesting question - possibly connected with some aspects of materials science - is to determine the "rank-one direction" for which the linear distortion function at A is most concave. This might identify the structure of the laminations for the minimisers of certain stored energy functionals occurring in the calculus of variations. This problem is basically what we solved in this thesis. An unanswered question, (partly posed by Gehring and Iwaniec) is to determine how big the difference between the left-hand side and the right-hand side of (5.3) can be?
The Theorem 5.1.5 is an existence theorem and in this theorem we only showed that there is a sequence of $\mathcal{H}$-quasiconformal self-mappings of $\mathbb{R}^3$ converging locally uniformly to an $\mathcal{H}_\infty$-quasiconformal self-mapping of $\mathbb{R}^3$ so that the relative jump is bigger than or equal to $1/\sqrt{2} = \sqrt{2}/2$. This result leads us to a more general conjecture. In Theorem 5.1.3, we found the best rank-one direction matrix B_0 and the interval $[\mathbf{t}_-, \mathbf{t}_+]$ such that for all t in this interval we have

$$\max\{\mathcal{H}(A + \mathbf{t}_- B_0),\ \mathcal{H}(A + \mathbf{t}_+ B_0)\} < \mathcal{H}(A).$$

and for all individual linear distortion functions, the biggest jump was found. Theorem 4.2.2 shows that there always exists a sequence of $\mathcal{H}$-quasiconformal self-mappings of $\mathbb{R}^3$ converging locally uniformly to an $\mathcal{H}_\infty$-quasiconformal self-mapping of $\mathbb{R}^3$ so that

$$\mathcal{H}(\mathbf{x}, \mathbf{f}_\nu) \equiv H < \mathcal{H}_\infty(\mathbf{x}, \mathbf{f}_0) \qquad \text{for all } \nu = 1, 2, 3, \ldots.$$

Conjecture: Let $\{\mathbf{f}_\nu\}_{\nu=1}^{\infty}$ be a sequence of quasiconformal self-mappings of $\mathbb{R}^n$ converging locally uniformly to a quasiconformal self-mappings $\mathbf{f_0}$ of $\mathbb{R}^n$. Suppose that $\mathcal{H}(\mathbf{f}_\nu) \longrightarrow H$, as $\nu \longrightarrow +\infty$, then for the biggest jump or relative jump, the following inequality holds:

$$\frac{\mathcal{H}(\mathbf{f}_\nu)}{\mathcal{H}(\mathbf{f_0})} \longrightarrow \frac{H}{\mathcal{H}(\mathbf{f_0})} \geq \frac{1}{\sqrt{2}}.$$

In other words, $mH \geq \mathcal{H}(\mathbf{f_0})$, for all $m > \sqrt{2}$.

Bibliography

[1] Lars Valerian Ahlfors. *Lectures on quasiconformal mappings*, volume 3. van Nostrand Princeton, 1966.

[2] Jean-Jacques Alibert and Bernard Dacorogna. An example of a quasiconvex function that is not polyconvex in two dimensions. *Archive for rational mechanics and analysis*, 117(2):155–166, 1992.

[3] Howard Anton and Chris Rorres. Elementary linear algebra, john wiley & sons. *Inc., Newyork*, 1994.

[4] Tom M Apostol. *Mathematical Analysis*. Addison-Wesley Reading, MA, 1974.

[5] Tom M Apostol and Charles H Holbrow. Calculus, volume ii. *American Journal of Physics*, 31(5):401–402, 1963.

[6] Kari Astala, Tadeusz Iwaniec, and Gaven Martin. *Elliptic Partial Differential Equations and Quasiconformal Mappings in the Plane (PMS-48)*. Princeton University Press, 2008.

[7] Kari Astala, Tadeusz Iwaniec, Gaven J Martin, and Jani Onninen. Extremal mappings of finite distortion. *Proceedings of the London Mathematical Society*, 91(3):655–702, 2005.

[8] John M Ball. Défauts dans les cristaux et dans les cristaux liquides (paris lecture slides).

[9] John M Ball. Convexity conditions and existence theorems in nonlinear elasticity. *Archive for rational mechanics and Analysis*, 63(4):337–403, 1976.

[10] John M Ball. Global invertibility of sobolev functions and the interpenetration of matter. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 88(3-4):315–328, 1981.

[11] John M Ball. Does rank-one convexity imply quasiconvexity? In *Metastability and incompletely posed problems*, pages 17–32. Springer, 1987.

[12] John M Ball, JC Currie, and Peter J Olver. Null lagrangians, weak continuity, and variational problems of arbitrary order. *Journal of Functional Analysis*, 41(2):135–174, 1981.

[13] John M Ball and Richard D James. Fine phase mixtures as minimizers of energy. In *Analysis and Continuum Mechanics*, pages 647–686. Springer, 1989.

[14] Viorel Barbu and Teodor Precupanu. *Convexity and optimization in Banach spaces*. Springer Science & Business Media, 2012.

[15] MK Benaouda and JJ Telega. Two-dimensional examples of rank-one convex functions that are not quasiconvex. In *Annales Polonici Mathematici*, volume 73, pages 291–295. L'institut De Mathematiques De L'academie Polonaise, 2000.

[16] Vladimir I Bogachev and Oleg G Smolyanov. *Real and functional analysis*, volume 4. Springer Nature, 2020.

[17] Frank Burk. *Lebesgue measure and integration: an introduction*, volume 32. John Wiley & Sons, 2011.

[18] Susan Jane Colley. *Vector calculus.* Pearson,, 2012.

[19] Bernard Dacorogna. *Direct methods in the calculus of variations*, volume 78. Springer Science & Business Media, 2007.

[20] Bernard Dacorogna and Jean-Pierre Haeberly. Remarks on a numerical study of convexity, quasi-convexity, and rank-one convexity. *Progress in Nonlinear Differential Equations and Their Applications*, 25:143–154, 1996.

[21] Bernard Dacorogna and Jean-Pierre Haeberly. Some numerical methods for the study of the convexity notions arising in the calculus of variations. *ESAIM: Mathematical Modelling and Numerical Analysis*, 32(2):153–175, 1998.

[22] Bernard Dacorogna and Paolo Marcellini. A counterexample in the vectorial calculus of variations. *Material instabilities in continuum mechanics*, pages 77–83, 1988.

[23] Lokenath Debnath and Piotr Mikusinski. *Introduction to Hilbert spaces with applications.* Academic press, 2005.

[24] Peter Duren, Juha Heinonen, Brad Osgood, and Bruce Palka. *Quasiconformal mappings and analysis: a collection of papers honoring FW Gehring.* Springer Science & Business Media, 2012.

[25] Vera Ebbing. *Design of polyconvex energy functions for all anisotropy classes.* Inst. für Mechanik, Abt. Bauwissenschaften, 2010.

[26] Herbert Federer. *Geometric measure theory.* Springer, 2014.

[27] John M Franks. *A (terse) introduction to Lebesgue integration*, volume 48. American Mathematical Soc., 2009.

[28] Frederick W Gehring. *The Carathéodory convergence theorem for quasiconformal mappings in space.* Suomalainen Tiedeakatemia, 1963.

[29] Frederick W Gehring. Thel p-integrability of the partial derivatives of a quasiconformal mapping. *Acta Mathematica*, 130(1):265–277, 1973.

[30] Frederick W Gehring, Gaven J Martin, and Bruce P Palka. *An Introduction to the Theory of Higher-Dimensional Quasiconformal Mappings*, volume 216. American Mathematical Soc., 2017.

[31] FW Gehring. Quasiconformal mappings in space. *Bulletin of the American Mathematical Society*, 69(2):146–164, 1963.

[32] FW Gehring and T Iwaniec. The limit of mappings with finite distortion. In *Annales-Academiae Scientiarum Fennicae Series A1 Mathematica*, volume 24, pages 253–264. Academia Scientiarum Fennica, 1999.

[33] Markus Haase. *Functional analysis: an elementary introduction*, volume 156. American Mathematical Society, 2014.

[34] Seyed Mohsen Hashemi and Gaven J Martin. New models for deformations: Linear distortion and the failure of rank-one convexity. *arXiv e-prints*, pages arXiv–2004, 2020.

[35] Juha Heinonen. What is a quasiconformal mapping. *Notices of the AMS*, 53(11):1334–1335, 2006.

[36] Juha Heinonen and Pekka Koskela. Definitions of quasiconformality. *Inventiones mathematicae*, 120(1):61–79, 1995.

[37] Stanislav Hencl and Pekka Koskela. *Lectures on mappings of finite distortion*, volume 2096. Springer, 2014.

[38] Kenneth Hoffman and Ray Kunze. Linear algebra. 1971. *Englewood Cliffs, New Jersey.*

[39] Tadeusz Iwaniec. The failure of lower semicontinuity for the linear dilatation. *Bulletin of the London Mathematical Society*, 30(1):55–61, 1998.

[40] Tadeusz Iwaniec. Nonlinear analysis and quasiconformal mappings from the perspective of pdes. *Banach Center Publications*, 48(1):119–140, 1999.

[41] Tadeusz Iwaniec, Leonid V Kovalev, and Jani Onninen. Lipschitz regularity for inner-variational equations. *Duke Mathematical Journal*, 162(4):643–672, 2013.

[42] Tadeusz Iwaniec and Gaven Martin. *Geometric function theory and non-linear analysis.* Clarendon Press, 2001.

[43] Tadeusz Iwaniec, Gaven Martin, and Jani Onninen. On minimisers of l^p -mean distortion. *Computational Methods and Function Theory*, 14(2-3):399–416, 2014.

[44] Tadeusz Iwaniec and Jani Onninen. Neohookean deformations of annuli, existence, uniqueness and radial symmetry. *Mathematische Annalen*, 348(1):35–55, 2010.

[45] Tadeusz Iwaniec and Jani Onninen. Smoothing defected welds and hairline cracks. *SIAM Journal on Mathematical Analysis*, 48(1):281–301, 2016.

[46] Juha Kinnunen. Sobolev spaces. *Department of Mathematics and Systems Analysis, Aalto University*, 2017.

[47] Bernard Kolman and D Hill. *Introductory linear algebra.* Pearson Prentice Hall, 2005.

[48] Giovanni Leoni. *A first course in Sobolev spaces.* American Mathematical Soc., 2017.

[49] Dag Lukkassen. A short introduction to sobolev-spaces and applications for engineering students. 2004.

[50] Gaven J. Martin and Cong Yao. Exponential teichmüller theory. *to appear.*

[51] Bert Mendelson. *Introduction to topology.* Courier Corporation, 1990.

[52] Boris S Mordukhovich and Nguyen Mau Nam. An easy path to convex analysis and applications. *Synthesis Lectures on Mathematics and Statistics*, 6(2):1–218, 2013.

[53] Charles B Morrey Jr. Quasi-convexity and the lower semicontinuity of multiple integrals. *Pacific journal of mathematics*, 2(1):25–53, 1952.

[54] Charles Bradfield Morrey Jr. *Multiple integrals in the calculus of variations*. Springer Science & Business Media, 2009.

[55] James R Munkres. *Topology*. Prentice Hall, 2000.

[56] Joseph Muscat. *Functional analysis: an introduction to metric spaces, Hilbert spaces, and Banach algebras*. Springer, 2014.

[57] Patrizio Neff. Critique of" two-dimensional examples of rank-one convex functions that are not quasiconvex" by m.k. benaouda and j.j. telega. In *Annales Polonici Mathematici*, volume 86, page 193. L'institut De Mathematiques De L'academie Polonaise, 2005.

[58] Gail S Nelson. *A user-friendly introduction to Lebesgue measure and integration*, volume 78. American Mathematical Soc., 2015.

[59] Constantin Niculescu and Lars-Erik Persson. *Convex functions and their applications: a contemporary approach*. Springer Science & Business Media, 2006.

[60] David Poole. *Linear algebra: A modern introduction*. Cengage Learning, 2014.

[61] Antti Rasila. Introduction to quasiconformal mappings in n-space. In *Proceedings of the International Workshop on Quasiconformal*, volume 19, 2006.

[62] Michael Renardy and Robert C Rogers. *An introduction to partial differential equations*, volume 13. Springer Science & Business Media, 2006.

[63] Walter Rudin. Principles of mathematical analysis (international series in pure & applied mathematics). 1976.

[64] Walter Rudin. *Real and Complex Analysis*. McGraw-Hill, 1987.

[65] Christopher G Small. *Functional equations and how to solve them*. Springer, 2007.

[66] Dorairaj Somasundaram. *A Second Course in Mathematical Analysis*. Alpha Science International, 2010.

[67] Vladimír Šverák. Examples of rank-one convex functions. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 114(3-4):237–242, 1990.

[68] Vladimir Sverak. Quasiconvex functions with subquadratic growth. *Proceedings: Mathematical and Physical Sciences*, pages 723–725, 1991.

[69] Vladimír Šverák. Rank-one convexity does not imply quasiconvexity. *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, 120(1-2):185–189, 1992.

[70] Vladimír Šverák. On lower-semicontinuity of variational integrals. *Equadiff 8*, pages 217–220, 1994.

[71] Terence Tao. *An introduction to measure theory*, volume 126. American Mathematical Soc., 2011.

[72] Kristopher Tapp. *Differential Geometry of Curves and Surfaces.* Springer, 2016.

[73] Vesna Todorčević. *Harmonic quasiconformal mappings and hyperbolic type metrics.* Springer, 2019.

[74] Jussi Väisälä. *Lectures on n-dimensional quasiconformal mappings*, volume 229. Springer, 2006.

[75] Jendrik Voss, Ionel-Dumitrel Ghiba, Robert J Martin, and Patrizio Neff. A rank-one convex, non-polyconvex isotropic function on $GL^+(2)$ with compact connected sublevel sets. *arXiv preprint arXiv:2009.09690*, 2020.

[76] Nik Weaver. *Measure theory and functional analysis.* World Scientific Publishing Co Inc, 2013.

[77] William P Ziemer. *Weakly differentiable functions: Sobolev spaces and functions of bounded variation*, volume 120. Springer Science & Business Media, 2012.

Index