Topics in the Calculus of Variations: Quasiconvexification of Distance Functions and Geometry in the Space of Matrices

by

Leila Yadollahi Farsani

A thesis submitted for the degree of *Doctor of Philosophy*

in the

University of Sussex

Department of Mathematics

April 2016

Declaration

I hereby declare that this thesis has not been and will not be submitted in whole or in part to another University for the award of any other degree.

Signature:

Leila Yadollahi Farsani

UNIVERSITY OF SUSSEX

Leila Yadollahi Farsani

Submitted for the degree of Doctor of Philosophy

April 2016

Quasiconvexification of distance functions and geometry in the space of matrices

Abstract

This thesis is concerned with C.B. Morrey's notion of quasiconvexity in the calculus of variations and the construction of various quasiconvex functions with desirable analytic and geometric properties. For variational integrals of the form

$$\mathbb{I}(u) = \int_\Omega f(Du(x))\, dx \tag{1}$$

the existence of minimizers is intimately linked with the *quasiconvexity* of the integrand f, namely, the condition

$$\int_\Omega f(\xi + D\varphi(x))\, dx \geq \int_\Omega f(\xi)\, dx, \tag{2}$$

for $\xi \in \mathbb{R}^{N\times n}$ and all $\phi \in C_c^\infty(\Omega, \mathbb{R}^N)$. In this thesis we study quasiconvex functions in the form of squared distance functions to various subsets $K \subset \mathbb{R}^{N\times n}$ or their relaxations when appropriate. We also extend this to the construction of various quasiconvex functions with specified growth and zero sets and investigate various connections with a number of old and longstanding problems in harmonic analysis and geometric functions theory, including, the Beurling-Ahlfors operator and the Burkholder functional. We also consider the p-Dirichlet energy over a space of vector-valued Sobolev maps from generalised annuli $\mathbb{X}$ into spheres and examine a geometric class of maps as solutions to the associated Euler-Lagrange equation (i.e., p-harmonic maps)

$$\Delta_p u + |\nabla u|^p u = 0, \qquad 1 < p < \infty, \tag{3}$$

where Δ_p is the so-called p-Laplacian. We prove that in even dimensions this system has infinitely many spherical twist solutions satisfying $u(x) = x|x|^{-1}$ on $\partial\mathbb{X}$.

Acknowledgements

I would like to firstly express my special appreciation to my supervisor Dr.Ali Taheri, whose help and academic support played an essential role in the completion of this thesis. I wish to also thank my other two supervisors namely Dr. Miroslav Chlebik and Dr. Qi Tang for their academic help and guidance, which made the completion of this thesis possible. I sincerely offer my gratitude to Professor Antonella De Santo and Dr. Bertram Duering for their invaluable help during times of particular hardships. My passion for mathematics is without doubt rooted back to my first exposure of mathematics during my secondary school, due to which I wish to express my gratitude to my secondary school mathematics teacher.

Undertaking this PhD has been a life-changing experience for me, with a number of easy, exciting and difficult memories, and I wish to express my appreciation to all those who have in one way of another contributed to the success of this academic achievement. I would like to also appreciate all types of academic and non-academic contributions I have received towards the completion of this thesis.

I am also indebted to all my friends who were always so helpful in many ways. Too many names to list in the scope of this acknowledgement.

A very special thanks goes out to my parents and my brothers for all their constant unconditional love and encouragement. I cant thank enough for all their help and supports in numerous ways.

Contents

Chapter 1

Introduction

Variational problems and their associated minimization or maximization principles form one of the most wide-ranging means of formulating mathematical models governing the equilibrium configurations and stable states of physical systems. In this thesis we consider variational problems that entail an integral functional in the form

$$\mathbb{I}(u) = \int_\Omega f(x, u(x), Du(x))\, dx, \tag{1.1}$$

where $\Omega \subset \mathbb{R}^n$ $(n \geq 1)$ is a bounded domain (open and connected set) with a point in Ω being denoted by $x = (x_1, x_2, \ldots, x_n)$, $u = (u_1, u_2, \ldots, u_N)$ is a map assigning the value $u(x) \in \mathbb{R}^N$ $(N \geq 1)$ to $x \in \Omega \subset \mathbb{R}^n$ and $f : \Omega \times \mathbb{R}^N \times \mathbb{R}^{N\times n} \mapsto \mathbb{R}$ the integrand or energy density. Associated with the energy functional $\mathbb{I}$ is the following minimization problem (**P**):

$$\alpha := \inf \Big\{ \mathbb{I}(u) : u \in \mathbb{X} \Big\}, \tag{1.2}$$

meaning that we wish to (hopefully) find a map $\bar{u} \in \mathbb{X}$ such that

$$\alpha = \mathbb{I}(\bar{u}) \leq \mathbb{I}(u), \text{ for all } u \in \mathbb{X}. \tag{1.3}$$

Here $\bar{u}$ is called a minimizer of $\mathbb{I}$ over $\mathbb{X}$ and $\mathbb{X}$ is a suitable space of admissible maps itself forming part of the mathematical model. For the sake of this thesis the space of admissible maps $\mathbb{X}$ is often a suitable subset of Sobolev spaces of weakly differentiable functions: $\mathbb{X} = u_0 + W_0^{1,p}(\Omega; \mathbb{R}^N)$, where u_0 is a given fixed map and therefore the notation $u \in \mathbb{X}$ is the shortcut for meaning that $u = u_0$ on $\partial\Omega$ and $u - u_0 \in W_0^{1,p}(\Omega; \mathbb{R}^N)$. (Here and throughout the exponent p is restricted to the range $1 \leq p \leq \infty$ often excluding the end points $p = 1$ or ∞ for technical reasons.)

The first crucial question that arises in connection with problem (**P**) is, of course, the question of existence of minimizers. This naturally depends on the choice of admissible functions $\mathbb{X}$ as the qualitative features and properties of the integrand or energy density f. A natural choice for the space of admissible maps would probably be a subspace of $C^1(\overline{\Omega};\mathbb{R}^N)$ or even $C^\infty(\overline{\Omega};\mathbb{R}^N)$ to have existence and regularity of minimizers in one strike and further to be able to write down the first order differential equation associated to the minimization problem better known as the *Euler-Lagrange equation.* This however turns out to be a strategy too hard too implement in most problems, particularly those dealing with partial derivatives (i.e $n \geq 2$). The essence of the so-called *direct method of the calculus of variations* is to split the problem into two main parts. First to enlarge the space of admissible maps, for instance, by considering spaces of weakly differentiable functions such as Sobolev spaces $W^{1,p}$ ($p \geq 1$), so as to secure a general existence theorem and second to prove some regularity or smoothness results that should be satisfied by any minimizer of (**P**).

The existence of minimizers in the above framework can now be tackled by the direct method of the calculus of variations and this is intimately connected with the fundamental property of sequential weak lower semicontinuity in $W^{1,p}$. As it happens this property is related to the convexity properties of the integrand, specifically, convexity of the function $\xi \mapsto f(x,u,\xi)$ in the scalar case and *quasiconvexity* in the vectorial case. In this thesis we are mainly interested in the vectorial case $n, N \geq 2$, and so the focus is entirely on the question and investigation of quasiconvexity. In the remainder of this introduction we proceed by giving a brief outline of the thesis and the topics covered in each chapter.

Chapter 2

In chapter 2, we give an overview and background on the tools, techniques and main results from the literature used in the thesis. In particular we give a brief account on the Direct method in the Calculus of Variations. We express and describe in more depth the reason why the concept of quasiconvexity is very important in the Calculus of Variations. There are several other closely related convexity notions that are introduced and considered. We investigate the relations amongst these notions and prove various statements and theorems to that effect. As expressed earlier the concept of quasiconvexity arises in conjunction with the sequential weak lower semicontinuity in Sobolev spaces $W^{1,p}$. When dealing with the Euler-Lagrange equations the natural concept is ellipticity or Legendre-Hadamard condition and this immediately leads to the notion of *rank-one convexity.* A more geometric notion involving the minors and subdeterminants is that of *polyconvexity.* We present the definition of these notions and the relations amongst them. As a very special case and for the sake of future applications we also discuss these notions in the so-called *quadratic* case where tools from Fourier analysis and specifically Plancherel's theorem lead to an interesting equivalence between quasiconvexity and rank-one convexity [for quadratic forms $\xi \mapsto f(\xi)$]. Let $N = n$, the function

$$f(\xi) = \phi(\det \xi)$$

with $\phi : \mathbb{R} \to \mathbb{R}$ convex is polyconvex hence quasiconvex and subsequently rank-one convex but not in general convex. When $n \geq 2$ and $N \geq 3$ *Šverák* produced an example of a function f that is rank-one convex but not quasiconvex hence answering a longstanding conjecture of Morrey to the effect that quasiconvexity is *not* implied by rank-one convexity. The reverse implication is true regardless of the range of n, N. However it is still an open problem if there are rank-one convex f that are not quasiconvex in the case $N = n = 2$ or more generally $n \geq N = 2$. As part of the discussion we look at the function studied by Alibert-Dacorogna-Marcellini given by $f_\gamma : \mathbb{R}^{2\times 2} \mapsto \mathbb{R}$ for $\gamma \in \mathbb{R}$, where

$$f_\gamma(\xi) = |\xi|^2(|\xi|^2 - 2\gamma \det \xi). \tag{1.4}$$

Further investigation of this interesting function continues in Chapter 4.

Chapter 3

In this chapter we discuss squared distance functions to sets K in the space of matrices with particular emphasis and interest in their convexity properties and relaxation. We start by setting the general problem and giving some useful and illustrative examples including when $K = \{A, B\}$ with $\operatorname{rank}(A - B) = 1$ as well as K=subspace with rank-one direction, K=subspace with no rank-one direction, $K = \mathbf{SO}(n)$ and more. We then specialise exclusively to the special orthogonal group $\mathbf{SO}(n)$ $n \geq 2$ by first showing that the squared distance to $\mathbf{SO}(n)$ is not quasiconvex and then finding an explicit formula for this distance function. Then we specialise further to the case $n = 2$ and discuss matters further by connecting to the rich and well-developed topic of geometric function theory, plane quasiconformal maps and subspaces of conformal and anti-conformal matrices.

Chapter 4

In this chapter, we use exclusively the technique of Zhang in construction of Quasiconvex functions with linear growth via maximal function methods. Then we introduce Müller's improvement and variant of Zhang's lemma. It also includes the examples of Dacorogna-Marcellini. The parameter dependent function which is convex, quasiconvex, polyconvex, rank-one convex for different ranges of parameters. This chapter then will end with some open problems.

Chapter 5

In this chapter we look at the p-Dirichlet energy for a space of maps from a generalised annulus and taking values in spheres. We take a close look at the associated Euler-Lagrange equation, specifically, the p-harmonic map equation, and examine a class of geometric maps, called spherical twists, as solutions. Spherical twists are mappings y from $\mathbb{X} = \{x \in \mathbb{R}^n : a < |x| < b\}$ to $\mathbb{S}^{n-1}$ taking the explicit form

$$y(x) = Q(r)\theta, \qquad a < r = |x| < b, \quad \theta = x|x|^{-1} \in \mathbb{S}^{n-1}, \tag{1.5}$$

where $Q = Q(r)$ lies in $\mathbf{SO}(n)$ and we require maps y to satisfy the boundary conditions $y(x) = x|x|^{-1}$ on $\partial\mathbb{X}$. So it is evident that the compact Lie group $\mathbf{SO}(n)$ still plays a significant role in this chapter. After some general discussion and standard calculations we consider the p-Dirichlet energy as being restricted to the space of spherical twists. This transforms the problem into one posed over closed curves in the pointed space $(\mathbf{SO}(n), \mathrm{I}_n)$. After a careful analysis of the stationary solutions to this variational problem (closely related to a rescaled p-geodesic problem posed over the Lie group) we proceed by extracting out of all such stationary curves, those that furnish a solution to the p-harmonic map equation. It turns out that in even dimensions the probelm has infinitely many solutions where as in odd dimensions there is only one spherical twist solution, namely, the map $y = x|x|^{-1}$. This chapter continues earlier works by Shahrokhi and Taheri [42, 44, 45, 43] and Taheri [51, 52, 54, 53] and is part of joint work with these authors.

Chapter 2

Preliminaries and Background Material

2.1 The Direct Method of the Calculus of Variations

The direct method of the calculus of variations is a classical and fundamental method for proving the existence of minimizers and maximizers. It generalises the principle that a real-valued continuous function on a compact subset of a topological space attains both its infimum and supremum. In most variational problems one is interested solely in minimization of a functional (which typically takes the form of some energy or entropy) and attainment of infimum is related to existence of ground states or equilibirium states. The direct method of the calculus of variations gives sufficient conditions for the existence of minimizers of the energy or entropy functional over a suitable space of admissible states and as one naturally expects convexity and coercivity here play essential roles. Let us proceed by formulating this and introducing the necessary terminology and concepts in a form most suitable and convenient for future applications. (We point out that the extent that topics are dealt with in this preliminary chapter is proportional and directed towards later needs only and readers wishing a more in depth exposition should consult the references in the bibliography.)

Definition 2.1.1. (Sequential lower semicontinuity)
A functional $\mathbb{I} : \mathbb{X} = W^{1,p}(\Omega, \mathbb{R}^N) \to \mathbb{R} \cup \{+\infty\}$ (with $1 \leq p < \infty$) is said to be sequentially weakly lower semicontinuous (often *swlsc* for short) if and only if, for every sequence $u_j \rightharpoonup u$ in $\mathbb{X}$,

$$\liminf_{j \nearrow \infty} \mathbb{I}(u_j) \geq \mathbb{I}(u). \tag{2.1}$$

If $p = \infty$, then, $\mathbb{I}$ is said to be sequentially weak* lower semicontinuous in $W^{1,\infty}(\Omega, \mathbb{R}^N)$, if and only if (2.1) holds for every sequence $u_j \overset{*}{\rightharpoonup} u$ in $\mathbb{X} = W^{1,\infty}(\Omega; \mathbb{R}^N)$.

One of the main themes in the calculus of variations and this thesis is to consider the problem of minimizing the integral functional $\mathbb{I}$ given by

$$\mathbb{I}(u) = \int_\Omega f(Du(x))\, dx. \tag{1}$$

Here $f : \mathbb{R}^{n\times N} \to \mathbb{R}$ (the integrand) is a continuous real-valued function on the space of real $N \times n$ matrices denoted $\mathbb{R}^{N\times n}$, and Ω is a bounded open subset of $\mathbb{R}^n$. Furthermore $Du(x)$ denotes the gradient matrix of $u : \Omega \to \mathbb{R}^N$ at the point $x \in \Omega$. The question to begin with is when and under what conditions on f is the integral $\mathbb{I}$ sequentially weakly lower semicontinuous on the Sobolev space $\mathbb{X} = W^{1,p}(\Omega; \mathbb{R}^N)$? It is a classical and fundamental result of C.B. Morrey that a necessary condition for this is that f is quasiconvex. Furthermore under some additional conditions this quasiconvexity condition can be shown to be sufficient. (See [39][38] as well as [1]). Motivated by this discussion we now give a brief description and summary of the various convexity notions in the space of matrices (specifically $N \times n$ real matrices) used and needed in this thesis and discuss their relationship to one-another with particular emphasis on the vectorial case, that is, when $\min(N, n) \geq 2$, where these concepts become substantially different.

2.2 On Various Convexity Notions in the Space of $N \times n$ Matrices

A major interest in this thesis is on Morrey's quasiconvexity which as described above arises from the characterization of integrands or energy densities $f : \mathbb{R}^{N\times n} \to \mathbb{R}$ which give rise to lower semicontinuous energy functionals $\mathbb{I}(u) = \int f(Du)dx$. Towards this end let us proceed by introducing a number of related convexity notions associated with integrands f as above.

First we recall that a function f is said to be convex if and only if for every pair of matrices $A, B \in \mathbb{R}^{N\times n}$ and scalar $0 \leq t \leq 1$ we have

$$f(tA + (1-t)B) \leq tf(A) + (1-t)f(B). \tag{2.2}$$

Whilst for many applications convexity is an important and crucial property for the sake of multi-dimensional calculus of variations it is far too strong and we need to look at *weaker* convexity notions and conditions.

Definition 2.2.1. (Quasiconvexity)
A function $f : \mathbb{R}^{N\times n} \to \mathbb{R}$ is said to be quasiconvex at $A \in M^{N\times n}$ if and only if and for every smooth compactly supported $\varphi \in C_c(\Omega; \mathbb{R}^N)$ the following inequality holds:

$$\int_\Omega f(A + D\varphi)\, dx \geq \int_\Omega f(A)\, dx. \tag{2.3}$$

If the above inequality holds for every A then f is said to be quasiconvex (everywhere).

It can be shown that this defition does not depend on the choice of the domain $\Omega \subset \mathbb{R}^n$ in the sense that if it is true for one domain then it also holds for any other domain. (This can be done by using a Vitali's type covering argument. See, e.g., J.M. Ball [10].)

Definition 2.2.2. (Rank-one Convexity)
A function $f : \mathbb{R}^{N\times n} \to \mathbb{R}$ is said to be rank-one convex at $A \in \mathbb{R}^{N\times n}$ if and only if for every rank-one matrix $B = a \otimes b \in \mathbb{R}^{N\times n}$ the function

$$t \mapsto f(A + tB) \tag{2.4}$$

is convex. (Note that this means as a function of the variable t with $-\infty < t < \infty$). We say that f is rank-one convex if and only if f is convex at every $A \in \mathbb{R}^{N\times n}$. Thus here for every $A, B \in \mathbb{R}^{N\times n}$ with $\operatorname{rank}(A - B) \leq 1$ and every t we have

$$f(tA + (1 - t)B) \leq tf(A) + (1 - t)f(B), \qquad B - A = a \otimes b. \tag{2.5}$$

It is well-known that rank-one convexity is a necessary condition for quasiconvexity. It was a longstanding open problem from the work of Morrey in 1952 if the two notions are equivalent, however, in 1992 Šverâk produced an example to show that when $n \geq 2,\ N \geq 3$ this is not the case and indeed rank-one convexity does *not* imply quasiconvexity. The question to this date remains completely open in the case $n = 2,\ N = 2$ and there are conjecture asserting that here the two notions could be equivalent. We return to this discussion in a later chapter. The next and final convexity notion in our list is that of polyconvexity as introduced by Morrey and Ball as formulated below.

Definition 2.2.3. (Polyconvexity)
A function $f : \mathbb{R}^{N\times n} \to \mathbb{R}$ is said to be polyconvex if and only if the function $A \mapsto f(A)$ can be written as a convex function of the minors of A, that is, f is a convex function of all $p \times p$ subdeterminants of A where $1 \leq p \leq \min\{m, n\}$.

In order to clarify the concept let us give a basic example by considering the case $n = 2, N = 2$. Indeed here a function $f : \mathbb{R}^{2\times 2} \to \mathbb{R}$ is polyconvex if and only if there is a convex function, say, $g : \mathbb{R}^5 \to \mathbb{R}$ such that

$$f(X) = g(X, \det X), \qquad X \in \mathbb{R}^{2\times 2}. \tag{2.6}$$

Note that here we are using $X \mapsto (X, \det X)$ as a map from $\mathbb{R}^{2\times 2}$ into $\mathbb{R}^5$ and f is represented as a composition of this map with the convex function g. For higher dimensions the situation is essentially the same but naturally the notation is more cumbersome. Let us also remark before moving on, that polyconvexity is a weaker notion than convexity, that is, it follows from the latter but does not imply it. For instance, letting $h : \mathbb{R} \to \mathbb{R}$ be a convex function, we see the function $f : \mathbb{R}^{2\times 2} \to \mathbb{R}$ given by

$$f(X) = h(\det X), \qquad X \in \mathbb{R}^{2\times 2}, \tag{2.7}$$

is polyconvex but not necessarily convex.

Before moving on we also remark that the above convexity notions can be extended without much difficulty to the cases where the function f takes on the value $+\infty$ as well. For various applications, e.g., in nonlinear hyperelasticity this is a crucial assumption as, for example, it allows to incorporate the values $f(A) = +\infty$ when A satisfies $\det A \leq 0$. In the language of nonlinear elasticity this is precisely the formulation of the physical condition that material does not interpenetrate itself.

Let us now take a closer look at the convexity notion of polyconvexity and give a more detailed form of the definition that is more in line with the description in the 2×2 case outlined in the example above. Indeed a function f is polyconvex if and only if there exists $g : \mathbb{R}^{\tau(n,N)} \to \mathbb{R}$ convex such that

$$f(X) = [g \circ T](X) = g(T(X)), \qquad X \in \mathbb{R}^{N\times n}. \tag{2.8}$$

Here $T : \mathbb{R}^{N\times n} \to \mathbb{R}^{\tau(n,N)}$ is the map given by the array of subdeterminants

$$T(X) = (X,\ \mathrm{adj}_2 X, \ldots,\ \mathrm{adj}_{n\wedge N} X), \tag{2.9}$$

where $X \mapsto \mathrm{adj}_s X$ stands for the assignment of $s \times s$ minors of X to the matrix X (note that here $\mathbb{X} \in \mathbb{R}^{N\times n}, 2 \leq s \leq n \wedge N = \min\{n, N\}$) and

$$\tau(n, N) = \sum_{s=1}^{n\wedge N} \sigma(s), \tag{2.10}$$

with the quantity $\sigma(s)$ given explicitly by

$$\sigma(s) = \binom{N}{s} = \binom{n}{s} = \frac{N!n!}{(s!)^2(N-s)!(n-s)!}. \tag{2.11}$$

Having introduced all the necessary convexity notions, let us now proceed by saying a few words about the relationship between them. As before suppose $f : \mathbb{R}^{N\times n} \to \mathbb{R}$. Then we have the implications

$$f\, convex \Rightarrow f\, polyconvex \Rightarrow f\, quasiconvex \Rightarrow f\, rank\text{-}one\, convex. \tag{2.12}$$

Thus as the above implications suggests convexity is the strongest and rank-one convexity is the weakest among them all. We also point out that none of these implications work in the reverse direction. It is quite remarkable that the two notions at the start and the end of the arrows can be verified directly, e.g., when f is differentiable or sufficiently smooth by a certain positivity of the second derivatives whereas there is no such pointwise condition on quasiconvexity. (See J. Kristensen [33].) As a matter of fact subject to $f \in C^2(\mathbb{R}^{N\times n})$ it is seen that rank-one convexity is equivalent to the so-called Legendre-Hadamard condition (or sometimes ellipticity condition)

$$\sum_{i,j=1}^{N} \sum_{\alpha,\beta=1}^{n} \frac{\partial^2 f(\xi)}{\partial \xi_\alpha^i \partial_\beta^j} \lambda^i \lambda^j \mu_\alpha \mu_\beta \geq 0 \tag{2.13}$$

for every

$$\lambda \in \mathbb{R}^N, \mu \in \mathbb{R}^n, \xi = (\xi_\alpha^i)_{1\leq\alpha\leq n}^{1\leq i\leq N} \in \mathbb{R}^{N\times n}. \tag{2.14}$$

For a proof of these implications and further discussion, we refer the reader to the standard texts on the subject, e.g., B. Dacorogna [20]. For the sake of future reference and in line with what indicated earlier, let us also note that when $f : \mathbb{R}^{N\times n} \to \mathbb{R} \cup \{+\infty\}$ (that is f takes on the value $+\infty$ as well), then

$$f\, convex \Rightarrow f\, polyconvex \Rightarrow f\, rank\text{-}one\, convex. \tag{2.15}$$

We point out that in scalar case, namely when $\min(n, N) = 1$, all these convexity notions are equivalent and hence coincide with the usual convexity. It is in passing from scalar to vectorial case that the real difference between these convexity notions begin to present itself.

2.3 Relaxation, Quasiconvex Envelopes and Existence of Minimizers

The direct method of the calculus of variations is based on the observation that on a reflexive Banach space $\mathbb{X}$ a functional $\mathbb{I}$ that is bounded from below attains its infimum if it is firstly coercive and secondly it is sequentially weakly lower semicontinuous. For funcationals in the form

$$\mathbb{I}(u) = \int_\Omega f(Du(x))\,dx \tag{2.16}$$

defined over $\mathbb{X} = W^{1,p}(\Omega;\mathbb{R}^N)$ (with $1 < p$ to ensure reflexivity) the coercivity condition follows from a pointwise condition on the integrand f, namely,

$$f(X) \geq c_1|X|^p - c_2, \qquad c_1 > 0, c_2 \geq 0, \tag{2.17}$$

and the sequential weak lower semicontinuity follows from the growth condition

$$f(X) \leq c_0(1 + |X|^p), \qquad c_0 > 0, \tag{2.18}$$

and quasiconvexity. (See, e.g., Acerbi and Fusco [1].)

When f fails to be quasiconvex the functional $\mathbb{I}$ is not sequentially weakly lower semicontinuous and the direct method of the calculus of variations as indicated above does not apply. One of the ways of getting around this difficulty is to consider the quasiconvex relaxation of the integrand (i.e., the quasiconvex envelope of the function f) and subsequently the relaxation of the functional $\mathbb{I}$. By the quasiconvex envelope of f we mean the largest quasiconvex function which is smaller than f. This will hereafter be denoted by Qf or f^{qc}. Towards this end let $f : R^{N\times n} \mapsto \mathbb{R}$ be a given continuous function. Then we say

$$Qf = \sup\{g \leq f : \ g \text{ quasiconvex}\}. \tag{1}$$

The quasiconvex envelope of f can be seen to be quasiconvex and hence replacing f with Qf in the integral functional will result in the so-called relaxed functional $\bar{\mathbb{I}}$, specifically,

$$\bar{\mathbb{I}}(u) = \int_\Omega Qf(Du(x))\,dx. \tag{2}$$

Although it is straightforward from the definition that $Qf \leq f$ and hence $\bar{\mathbb{I}}(u) \leq \mathbb{I}(u)$ for all $u \in \mathbb{X} = W^{1,p}(\Omega;\mathbb{R}^N)$ it can be shown that in the level of the minimization problem we have

$$\inf \mathbb{I}(u) = \inf \bar{\mathbb{I}}(u). \tag{2.19}$$

This means that for every $u \in W^{1,p}(\Omega,\mathbb{R}^n)$, there exists a sequence $(u_j) \subset u + W_0^{1,p}(\Omega,\mathbb{R}^n)$ such that,

$$u_j \rightharpoonup u, \tag{2.20}$$

in $W^{1,p}(\Omega,\mathbb{R}^n)$ as $j \nearrow \infty$ while

$$\int_\Omega f(D(u_j)(x))\,dx \to \int_\Omega Qf(Du(x))\,dx, \tag{2.21}$$

as $j \nearrow \infty$. Note however that in the process of passing from $\mathbb{I}$ to $\bar{\mathbb{I}}$ the relaxed functional will attain its minimum whereas in general the original functional $\mathbb{I}$ will not. Indeed what is happening here is that minimizing sequences of the original functional $\mathbb{I}$ will converge weakly to minimizers of the relaxed functional $\bar{\mathbb{I}}$ but in general a minimizer of the relaxed functional is *not* a minimizer of the original functional $\mathbb{I}$. The converse of the above is also true in the sense that all minimizers of $\mathbb{I}$ are weak limits of a minimizing sequence for $\bar{\mathbb{I}}$.

Let us note that one can define the various semiconvex envelopes of a given integrand f in a way similar to that of its quasiconvex envelope. These in turn will be the convex envelopes, the polyconvex envelopes and the rank-one convex envelopes of f and denoted in turn by f^c, f^{pc} and f^{rc} respectively. (Here one restricts in the definition of $Qf = f^{qc}$ the functions g with $g \le f$ to be convex, polyconvex or rank-one convex respectively.) As a result in view of the relations between the various convexity notions discussed earlier it is not difficult to see that one has the chain of inequalities

$$f^c \le f^{pc} \le f^{qc} \le f^{rc}. \tag{2.22}$$

2.4 Convexity Notions and Quadratic Forms

Of particular interest is the case when the function $f : \mathbb{R}^{N\times n} \to \mathbb{R}$ is quadratic (or a quadratic form). Here as one can easily see the Euler-Lagrange equations associated with $\mathbb{I}$ are linear and one can say a lot more about the relationship between the various convexity notions introduced earlier. To fix notation let us agree to write

$$f(\xi) = \langle \mathbb{M}\xi ; \xi \rangle, \qquad \xi \in \mathbb{R}^{N\times n}, \tag{2.23}$$

where without loss of generality $\mathbb{M}$ can be taken a symmetric matrix in $\mathbb{R}^{(N\times n)\times(N\times n)}$ and $\langle ; \rangle$ denotes the scalar product in $\mathbb{R}^{N\times n}$. We then have the following.

Theorem 2.4.1. *Suppose f is a quadratic form. Then the following statements hold.*
(i) *f is rank-one convex if and only if f is quasiconvex.*
(ii) *If $N = 2$ or $n = 2$, then*

$$f \, \textit{polyconvex} \Leftrightarrow f \, \textit{quasiconvex} \Leftrightarrow f \, \textit{rank-one convex}. \tag{2.24}$$

We shall go through the argument and proof of this theorem shortly however before attending to this let us note that if $N, n \ge 3$, then in general

$$f \, \textit{rank-one convex} \not\Rightarrow f \, \textit{polyconvex}, \tag{2.25}$$

and also that even if $N = n = 2$ and f is quadratic then in general

$$f\,polyconvex \not\Rightarrow f\,convex. \tag{2.26}$$

In preparation for the proof of the above theorem let us suppose as before that $\mathbb{M}$ is a real symmetric $(N \times n) \times (N \times n)$ matrix and that f is the quadratic form $f(\xi) = \langle\, \mathbb{M}\xi; \xi\rangle$. Then the following results hold.

(i) f is convex if and only if

$$f(\xi) \geq 0 \tag{2.27}$$

for every $\xi \in \mathbb{R}^{N\times n}$.

(ii) f is polyconvex if and only if there exists $\alpha \in \mathbb{R}^{\sigma(2)}$ such that

$$f(\xi) \geq \langle \alpha; \mathrm{adj}_2\xi\rangle$$

for every $\xi \in \mathbb{R}^{N\times n}$. Here $\langle .;.\rangle$ denotes the scalar product in the real vector space $\mathbb{R}^{\sigma(2)}$ and $\sigma(2) = \binom{N}{2}\binom{n}{2}$.

(iii) f is quasiconvex if and only if for one (and hence every) non-empty bounded open set $\Omega \subset \mathbb{R}^n$ and for every $\varphi \in W_0^{1,\infty}(\Omega; \mathbb{R}^N)$

$$\int_\Omega f(D\varphi(x))\,dx \geq 0. \tag{2.28}$$

(iv) f is rank-one convex if and only if

$$f(\xi) \geq 0 \tag{2.29}$$

for every rank-one matrix $\xi = a \otimes b$ where $a \in \mathbb{R}^N$ and $b \in \mathbb{R}^n$.

As we mentioned before quasiconvexity implies rank-one convexity and in general the reverse implication is not true. Interestingly however if the function f is a quadratic form it can be shown that rank-one convexity and quasiconvexity are equivalent notions and so in particular quasiconvexity is implied by rank-one convexity.

Proposition 2.4.1. *Let f be a quadratic function. Then*

$$f\,quasiconvex \Longleftrightarrow f\,rank\text{-}one\,convex. \tag{2.30}$$

Proof. This observation is due originally to Morrey himself and the proof is a direct consequence of the Plancherel formula. Indeed here we aim to show that for rank-one convex quadratic form f and compactly supported φ,

$$\int\limits_{\mathbb{R}^n} f(\nabla\varphi)\,dx \geq 0.$$

To this end let $f : \mathbb{R}^{N\times n} \to \mathbb{R}$ be a quadratic and rank-one convex function. Then

$$f(a \otimes b) \geq 0.$$

Since f is quadratic by assumption we can write

$$f(a \otimes b) = \langle \mathbb{M}\,(a \otimes b), (a \otimes b)\rangle \geq 0. \tag{2.31}$$

Now the aim is to show that

$$\int\limits_{\mathbb{R}^n} \langle \mathbb{M}\,\nabla\varphi, \nabla\varphi\rangle\, dx \geq 0, \tag{2.32}$$

for every $\varphi \in C_0^\infty(\mathbb{R}^n; \mathbb{R}^n)$.

Upon taking Fourier transform it is clear that

$$\widehat{\nabla\varphi}(\xi) = \widehat{\frac{\partial\varphi_i}{\partial x_j}} = 2\pi i \xi_j \hat{\varphi}_i.$$

Hence substitution in f gives

$$\int\limits_{\mathbb{R}^n} \langle \mathbb{M}\,\widehat{\nabla\varphi}, \widehat{\nabla\varphi})\rangle\, d\xi = \int\limits_{\mathbb{R}^n} \langle Q(-2\pi i \xi_j \hat{\varphi}_i), -2\pi i \xi_j \hat{\varphi}_i\rangle\, d\xi.$$

Now using Plancherel theorem we can write

$$\int\limits_{\mathbb{R}^n} \langle \mathbb{M}\,(2\pi i \xi_j \hat{\varphi}_i), 2\pi i \xi_j \hat{\varphi}_i\rangle = \int\limits_{\mathbb{R}^n} \langle \mathbb{M}\,(-2\pi i \xi_j \varphi_i), 2\pi i \xi_j \varphi_i\rangle.$$

As $\xi \otimes \varphi = (\xi_j \varphi_i)$ is a rank-one matrix, it follows that

$$\int\limits_{\mathbb{R}^n} \langle \mathbb{M}\,(2\pi i \xi_j \varphi_i), 2\pi i \xi_j \varphi_i\rangle = \int\limits_{\mathbb{R}^n} \langle \mathbb{M}\,2\pi i(\xi_j \otimes \varphi_i), 2\pi i(\xi_j \otimes \varphi_i)\rangle.$$

Finally since f is a rank-one convex function, then the above integral is ≥ 0, which means

$$\int\limits_{\mathbb{R}^n} f(\nabla\varphi) \geq 0,$$

and so the assertion follows. □

2.5 Squared Distance Functions; Subspaces with no Rank-one Directions

In this final section we briefly review some of the main properties of distance functions to subspaces and particularly those subspaces containing no rank-one matrices. Squared distance functions are of interest mainly because they are quadratic forms that capture and encode the convexity properties of the sets which they measure distance from. (Here we confine the discussion mostly to subspaces but in later chapters this assumption is relaxed.)

Proposition 2.5.1. *Let $L \subset \mathbb{R}^{N\times n}$ be a subspace with no rank-one directions. Then there exists $\epsilon = \epsilon(L) > 0$ such that the squared distance function to L satisfies the following bound*

$$dist_L^2(u \otimes v) \geq \epsilon ||u \otimes v||^2. \tag{2.33}$$

Proof. Let f denotes the squared distance function to L. The idea is to minimize f over the compact set $K = \mathbb{S}^{N-1} \times \mathbb{S}^{n-1}$. Towards this end note that in virtue of f being quadratic we have $f(tA) = t^2 f(A)$ for every $t \in \mathbb{R}$. Thus in the case $\xi = (u \otimes v)$ for non-zero $u \in \mathbb{R}^N$ and $v \in \mathbb{R}^n$, upon setting $\bar{u} = u/||u||$ and $\bar{v} = v/||v||$, we have

$$\begin{aligned} f(u \otimes v) &= \|u\|^2 \|v\|^2 f(\bar{u} \otimes \bar{v}) \\ &= \|u \otimes v\|^2 f(\bar{u} \otimes \bar{v}). \end{aligned}$$

Now since $\bar{u}$ and $\bar{v}$ are unit vectors in $\mathbb{R}^N$ and $\mathbb{R}^n$ respectively to deduce the assertion it suffices to show that

$$\epsilon(L) := \inf_{\bar{u}\otimes\bar{v}\in K} f(\bar{u} \otimes \bar{v}) > 0. \tag{2.34}$$

However since the above infimum is attained over the compact set $K = \mathbb{S}^{N-1} \times \mathbb{S}^{n-1}$ the latter follows upon noting that L contains no rank-one matrix, that is, $f(\xi) > 0$ for every non-zero $\xi \in L$. The proof is finished. $\square$

The squared distance function to any subspace being quadratic and non-negative is evidently convex. As a result of the above proposition given $L \subset \mathbb{R}^{N\times n}$ a subspace with no rank-one directions it is seen that a slight *negative* perturbation of the squared distance function f to L, specifically, the function g defined by

$$g(\xi) = f(\xi) - \epsilon(L)\|\xi\|^2 \tag{2.35}$$

is rank-one convex (note that $g(\xi) \geq 0$ for all rank-one ξ) and in view of being quadratic is also quasiconvex yet g is not convex.

Let us now give another interpretation of the above. As before let $L \subset \mathbb{R}^{N\times n}$ be a subspace with no rank-one directions and denote by $L^\perp \subset \mathbb{R}^{N\times n}$ the orthogonal complement of L. Then it is clear that the squared distance function to L can be written as

$$f(\xi) = \mathrm{dist}_L^2(\xi) = \|P_{L^\perp}(\xi)\|^2. \tag{2.36}$$

Here and below we write P_L and $P_{L^\perp}$ for the orthogonal projections onto the subspaces L and $L^\perp$ respectively. Now consider minimizing the quotient

$$\lambda_L = \min_{|a|=|b|=1} \frac{|P_{L^\perp}(a \otimes b)|^2}{|P_L(a \otimes b)|^2}. \tag{2.37}$$

Then as before since L does not contain any rank-one matrices, we have $0 < \lambda_L < \infty$. Furthermore from the definition of λ_L, it is easy to see that λ_L is the largest positive number λ such that the quadratic function

$$q_\lambda(\xi) = \|P_{L^\perp}(\xi)\|^2 - \lambda \|P_L(\xi)\|^2 \tag{2.38}$$

is rank-one convex and hence quasiconvex. As a matter of fact relating to the above discussion we can write

$$\mathrm{dist}_L^2(\xi) = \|P_{L^\perp}(\xi)\|^2 \geq \epsilon(L)\|\xi\|^2 \tag{2.39}$$

and so

$$\|P_{L^\perp}(\xi)\|^2 - \epsilon(L)\left[\|P_{L^\perp}(\xi)\|^2 + \|P_L(\xi)\|^2\right] \geq 0. \tag{2.40}$$

Rearranging terms we can therefore write

$$\|P_{E^\perp}(A)\|^2 - \frac{\epsilon}{1-\epsilon}\|P_E(A)\|^2 \geq 0. \tag{2.41}$$

Hence comparison with the expression for q_λ we have the following relation between the quantities $\epsilon = \epsilon(L) > 0$ and λ_L

$$\lambda_L = \frac{\epsilon(L)}{1-\epsilon(L)}. \tag{2.42}$$

In the next chapter we give various examples of subspaces L where we can obtain explicitly via basic linear algebra the constant $\epsilon(L) > 0$ and hence $\lambda_L > 0$ for construction of quadratic quasiconvex functions. Let us end by giving a couple of basic examples of subspaces with no rank-one directions that will be used later on.

- Firstly in the 2×2 it can be easily seen that a non-zero matrix is rank-one if and only if it has a vanshing determinant. This simple test can be used to show that the two complimentary subspaces

$$\mathbb{R}^{2\times 2} = C_+ \oplus C_- \tag{2.43}$$

  that is the subspaces of conformal and anti-conformal matrices have no rank-one directions. (Note that each subspace is two dimensional.)

- As another example one can show that when $n = N$ the subspace $L \subset \mathbb{R}^{N\times n}$ of skew symmetric matrics (i.e., $\xi^t = -\xi$) contains no rank-one directions. Indeed assuming $u \otimes v \in L$ we must have

$$u \otimes v + v \otimes u = 0. \tag{2.44}$$

  Now if u and v are colinear vectors there is nothing to prove. Otherwise testing the above equation against any vector $u^\perp$ in the plane generated by u, v results in

$u = 0$. We point out that in contrast the space of symmetric matrices does contain rank-one directions however that is not to say that certain of its subspaces can not have no rank-one directions.

Chapter 3

Distance Functions and their Quasiconvexification

3.1 Introduction

The computation of the quasiconvex envelope of a function $\mathbb{I}(.)$ provides information on the asymptotic behavior of minimizing sequences for the corresponding functional. Quasiconvex relaxation of certain distance functions to a given set in the space of matrices is an important area in the study of optimal design problems. As obtaining of an explicit formula is hard, hence an estimate of the lower bound of the quasiconvex relaxation will provide us useful information on the set itself and on the relaxed function. The aim is now to compute some of these quasiconvex envelopes for certain functions $f : \mathbb{R}^{N\times n} \mapsto \mathbb{R}$ which are defined on the set of $N \times n$ matrices through a quadratic forms.

3.2 p-Distance Functions to Sets $K \subset \mathbb{R}^{N\times n}$ and their Quasiconvexification

Definition 3.2.1. Let $F(P)$ denotes the distance function from a point $P \in \mathbb{M}^{N\times n}$ to a set $K \subset \mathbb{M}^{N\times n}$. The p-distance function $\text{dist}^p(., K)$ characterizes the geometry of K and can be defined as

$$F(P) = \text{dist}^p(P, K) = \inf_{A\in K} |P - A|^p. \tag{3.1}$$

Let $\mathbb{M}^{N\times n}$ be the space of all $N \times n$ real matrices with $\mathbb{R}^{Nn}$ norm. If $E \subset \mathbb{M}^{N\times n}$ is a linear subspace, we denote P_E and $P_{E^\perp}$ as the orthogonal projection from $\mathbb{M}^{N\times n}$ to E and its orthogonal complement $E^\perp$ respectively.

Lemma 3.2.1. We can show the Quasiconvex envelope of $\text{dist}^p(Q, P)$ with

$$F(P) = Q\text{dist}^p(P, K)$$

for some closed subset $K \subset \mathbb{M}^{N\times n}$, where $\mathbb{M}^{N\times n}$ is the space of all $N \times n$ real matrices, without knowing the exact formula of $F(P)$, and where $Q\text{dist}^p(P, K)$ is the quasiconvexification of $\text{dist}^p(P, K)$, and $p > 1$.

We consider the minimizing problem $\inf \mathbb{I}(u)$ subject to certain boundary conditions, where

$$\mathbb{I}(u) = \int_\Omega F_K(Du(x))dx. \tag{1}$$

We are interested in the case $p = 2$, where $F_K = \text{dist}^2(., K)$ denotes the squared Euclidean distance function to the compact set K in matrix space $\mathbb{M}^{N\times n}$. Here Ω is a bounded open subset of $\mathbb{R}^n$, $Du(x)$ denotes the gradient matrix of u of x. Function $f : \mathbb{R}^{n\times N} \to \mathbb{R}$ is a continuous real valued function on the space of real $N \times n$ matrices $\mathbb{M}^{N\times n}$ and $u : \Omega \to \mathbb{R}^N$ is a mapping in the Sobolev space $W^{1,2}(\Omega; \mathbb{R}^{\mathrm{N}})$. The motivation for the investigation of functional related to the distance function is, from purely mathematical viewpoint, that a class of functional is natural generalization of the Dirichlet integral (when K is a singleton), and also the potential applications to a class of differential inclusions, namely to find $u : \Omega \to \mathbb{R}^N$ satisfying a Dirichlet boundary condition on $\partial\Omega$ and $Du \in K$ almost everywhere in Ω. In addition, distance functions are frequently used in mathematical models for microstructure in solids where typical energy densities vanish a finite union of energy wells and are positive elsewhere. (see[12],[13].) If the set K is convex, then the metric projection $\pi : \mathbb{M}^{N\times n} \to K$ is uniquely defined, 1-Lipschitz and the squared distance function on the set K is convex and $C^{1,1}$. However, if K is not convex, then the squared distance function even fails to be quasiconvex [23] [63], and so we consider the relaxed functional instead,

$$\mathbb{I}^{qc}(u) = \int_\Omega F_K^{qc}(Du(x))dy \tag{2}$$

where $F_K^{qc}(.) = [\text{dist}^2(.)]^{qc}$ is the quasiconvex envelope of F_K. (see [20][36].)

Proposition 3.2.1. *Suppose $f : \mathbb{M}^{N\times n} \mapsto \mathbb{R}$ is continuous, then*

$$Qf(P) = \inf_{\phi\in C_0^\infty(\Omega,\mathbb{R}^N)} \frac{1}{\text{meas}(\Omega)} \int_\Omega f(P + D\phi(x))dx,$$

where $\Omega \subset \mathbb{R}^n$ is bounded domain. In particular, here infimum is independent of Ω.

As an example, we can see the set $K = SO(n)$ is not convex, then the squared distance function to it is not even quasiconvex.[60]. In the work [9] established Lipschitz regularity for the gradient of the relaxation F^{qc} and hence the Euler Lagrange equations

$$\int_{\Omega} \sigma(Du) : D\varphi dx = 0 \text{ for all } \varphi \in W_0^{1,2}(\Omega; \mathbb{R}^{\mathbb{N}})$$

with $\sigma = DF_K^{qc}$ can be used to drive regularity with minimizers (or extremal in general).

3.3 A Primer on Distance Functions and Metric Projections

For a closed subset K of $\mathbb{H}$ where it denotes a real and finite-dimensional of Hilbert space with inner product $\langle ., . \rangle$ and associated norm $\|.\|$. For a closed subset K of $\mathbb{H}$ we denote its squared distance function by F_K which is defined as

$$F_K(x) = \operatorname{dist}(x, K)^2 := \inf_{y \in K} \|x - y\|^2.$$

For each $x \in K$ the above infimum is easily seen to be attained. Hence the set

$$\pi_K(x) = \{y \in K : \|x - y\|^2 = F_K(x)\}$$

is a nonempty, closed subset of K. We consider $\pi K : H \to 2K$ as a multi-valued (set-valued) mapping, and refer to it as the metric projection onto K. For every closed subset K of C_1, C_2, (the space of conformal, anti-conformal respectively) matrices in $\mathbb{M}^{2\times 2}$, the quasiconvexification of the distance function $\operatorname{dist}(., K)$ is bounded below by itself that is,

$$c \operatorname{dist}(P, K) \leq Q\operatorname{dist}(P, K)$$

and the constant $c > 0$ is independent of K. From the definition of quasiconvex relaxation we have

$$Q\operatorname{dist}(P, K) \leq \operatorname{dist}(P, K).$$

As we see $Q\operatorname{dist}(P, K)$ is not convex, If $K \subset C_1$ (C_2, respectively) is closed and non-convex. It showed in [61] that $\operatorname{dist}(., K)$ is not rank-one convex in $\mathbb{M}^{2\times 2}$, justifying for any closed set $K \subset \mathbb{M}^{2\times 2}$ which is supported by C_1 (C_2, respectively). In the case where C_1 is the supporting space of K, we have that

$$c\operatorname{dist}(P, K) - C|P_{C_1}(P)| \leq Q\operatorname{dist}(P, K).$$

The connected subsets of $\mathbb{M}^{2\times 2}$, a closed connected set K does not have rank-one connections if and only if K is a Lipschitz graph of mapping f from a closed set of C_1 to C_2 or from a closed set of C_2 to C_1 respectively), such that

$$|f(A) - f(B)| < |A - B|, \qquad A \neq B$$

for any $p \in (1, \infty)$, there exist some $t(p) > 0$, if K is such a graph satisfying

$$|f(A) - f(B)| < k|A - B| \text{ and } k^p < t(p),$$

then the quasiconvexification $Q\text{dist}(., K)$ satisfies

$$\{P \in \mathbb{M}^{2\times 2}, Q\text{dist}(P, K) = 0\} = K.$$

Theorem 3.3.1. *The gradient of any extremal $u \in W^{1,2}(\Omega; \mathbb{R}^N)$ of the variational problem $\mathbb{I}^{qc}(u)$ belongs to $BMO_{loc}(\Omega; \mathbb{M}^{N\times n})$. Moreover, this regularity is optimal in the sense that for $n = N = 2$ and $K = SO(2)$ there exist minimizers of $\mathbb{I}^{qc}(u)$ that are not locally Lipschitz continuous on any open nonempty subset.[24]*

Distance Function to sets $K = \{A, B\}$

Let Ω be a bounded open subset of $\mathbb{R}^n$ and measure of Ω is its Lebesgue measure. We denote by $\mathbb{M}^{N\times n}$ the space of real $N \times n$ matrices with the standard $\mathbb{R}^{Nn}$ metric; hence the norm of $P \in \mathbb{M}^{N\times n}$ is defined by $|P| = (\text{tr}P^T P)^{1/2}$, where tr is the trace operator and P^T is the transpose of P. The inner product of two matrices in $\mathbb{M}^{N\times n}$ is $P : Q = \text{tr}P^T Q$. For a mapping $\phi : \Omega \subset \mathbb{R}^n \mapsto \mathbb{R}^N$, we denote by $D\phi(x)$ its gradient matrix in $\mathbb{M}^{N\times n}$ at $x \in \Omega$. $C_0^\infty(\Omega, \mathbb{R}^N)$ is the space of all $\mathbb{R}^N$-valued smooth functions with compact support in Ω. Let $K = \{A, B\}$,

$$F : \mathbb{M}^{N\times n} \mapsto \mathbb{R}, \quad F(P) \geq 0, \, F(P) = 0 \Leftrightarrow P \in K,$$

where $K = A, B$ with rank $(A - B) > 1$.

Theorem 3.3.2. *[32] Let $f : \mathbb{M}^{N\times n} \mapsto \mathbb{R}$ be a continuous function.*

$$f(P) = \min\{|P - A|^2, |P - B|^2\} = \text{dist}^2(P, \{A, B\}), \tag{3.2}$$

where $A, B \in \mathbb{M}^{N\times n}$ are fixed matrices. Then the quasiconvexification of f is given by

$$Qf(P) = \min_{0\leq\theta\leq 1}\{|P - \theta A - (1 - \theta)B|^2 + \theta(1 - \theta)[|A - B|^2 - \lambda_{\max}]\}, \tag{3.3}$$

where $\lambda_{\max}$ is the greatest eigenvalue of the matrix $(A - B)^T(A - B)$.

Then, from Theorem 3.3.2. and the construction in [32] we can present the following proposition.

Proposition 3.3.1. *Let*

$$f(P) = \min\{|P - A|^2, |P - B|^2\} = \text{dist}^2(P, \{A, B\}). \tag{3.4}$$

Then

$$Qf(.) = Rf(.). \tag{3.5}$$

Proof. (Proof of Proposition 3.3.2) We showed earlier that, $Qf \leq Rf \leq f$. Differentiating the right hand side of the equation we have

$$\theta = \frac{2(P-B).(A-B) - [|A-B|^2 - \lambda_{\max}]}{2\lambda_{\max}},$$

or

$$2(P-B).(A-B) = 2\theta\lambda_{\max} + |A-B|^2 - \lambda_{\max}.$$

When $\theta \leq 0$, this is,

$$(P-B).(A-B) \leq \frac{1}{2}[|A-B|^2 - \lambda_{\max}],$$

we should set $\theta = 0$, so that

$$Qf(P) = |P-B|^2 = \operatorname{dist}^2(P,K) = f(P).$$

Hence

$$Rf(P) = Qf(P) = f(P).$$

When $\theta \geq 1$, that is,

$$(P-B).(A-B) \geq \frac{1}{2}[|A-B|^2 - \lambda_{\max,}],$$

we should set $\theta = 1$, so that

$$Qf(P) = |P-B|^2 = \operatorname{dist}^2(P,K) = f(P).$$

Hence we still have

$$Rf(P) = Qf(P) = f(P).$$

When 0< $\theta < 1$, we consider

$$C_+ = P + (1-\theta)[(A-B)\eta] \otimes \eta \quad \text{and} \quad C_- = P - \theta[(A-B)\eta] \otimes \eta,$$

where $\eta \in \mathbb{R}^n$ is a unit eigenvector of $(A-B)^T(A-B)$ corresponding to the biggest eigenvalue. We easily see that rank $(C_+ - C_-) = 1$ and $\theta C_+ + (1-\theta)C_- = P$. Sincce Rf is rank one convex, we have

$$R(f(P)) = Rf(\theta C_+ + (1-\theta)C_-) \leq \theta Rf(C_+) + Rf(C_-).$$

Since,

$$
\begin{aligned}
&(C_+ - B).(A-B) = (P + (1-\theta)([(A-B)\eta] \otimes \eta - B).(A-B) \\
&= (P-B).(A-B) + (1-\theta)([(A-B)\eta] \otimes \eta).(A-B) \\
&= \theta\lambda_{\max} + \frac{1}{2}(|A-B|^2 - \lambda_{\max}) + (1-\theta)\lambda_{\max} \\
&= \frac{1}{2}(|A-B|^2 + \lambda_{\max}),
\end{aligned}
$$

we have

$$Rf(C_+) = |C_+ - A|^2 = Qf(C_+) = f(C_+).$$

So, we can prove that

$$(C_- - B).(A-B) = \frac{1}{2}(|A-B|^2 + \lambda_{\max}),$$

so that

$$Rf(C_-) = |C_- - B|^2 = Qf(C_-) = f(C_-).$$

Consequently,

$$
\begin{aligned}
Rf(P) &\leq \theta|C_+ - A|^2 + (1-\theta)|C - B|^2 \\
&= \theta|P-A|^2 + (1-\theta)|P-B|^2 - \theta(1-\theta) - \lambda_{\max} \\
&= |P - \theta A + (1-\theta)B|^2 + \theta(1-\theta)(|A-B|^2 \lambda_{\max}) \\
&= Qf(P).
\end{aligned}
$$

So $Rf(P) = Qf(P)$ for every $P \in \mathbb{M}^{N\times n}$. □

Definition 3.3.1. *(supporting space of set)*
A non-empty, closed subset K of $\mathbb{M}^{2\times 2}$ is supported by C_1(C_2, respectively), if there exists an orthonormal basis of $\{c_1, c_2\}$ such that $c_1.P \geq 0$ for all $P \in K$ and $i = 1, 2$ and dot denotes the inner product of matrices when $n = 2$. We call C_1(C_2, respectively), the supporting space of K.

Lemma 3.3.1. Suppose that $K \subset C_1$ (C_2, respectively) is closed and non-convex. Then $\mathrm{dist}(., K)$ is not rank-one convex.

Theorem 3.3.3. *Suppose that $K \subset C_1$ (C_2, respectively) is closed (possibly unbounded). If we denote by*

$$K_\epsilon = \{P \in \mathbb{M}^{2\times 2},\ dist(P, K)\} \tag{3.6}$$

the ϵ-neighbourhood of K, there exists a constant $c > 0$ independent of K, such that

$$c\,\mathrm{dist}(P, K) \leq Q\mathrm{dist}(P, K) \leq \mathrm{dist}(P, K),$$

for every $P \in \mathbb{M}^{2\times 2}$.

We have then the following consequences of above Theorem.

Corollary 3.3.1. *Under the assumption of Theorem 3.3.3, we have*

$$Q_1(K_\epsilon) \subset K_\epsilon \tag{3.7}$$

for every $\epsilon > 0$, $K \subset \mathbb{M}^{N\times n}$ has a rank-one connection if there exist $A, B \in K$ such that rank $(A - B) = 1$.

We are interested in the situation when $\inf \mathbb{I}(u) = 0$ with F satisfying $F(P) = 0$, $P \in K$ and $F(P) > 0$, $P \neq K$. For the existence of such minimizers, quasiconvexity of the function f is the necessary and sufficient condition.

Theorem 3.3.4. *Let B be a compact subset of $\mathbb{R}^p$ and g a Carathéodory function of $U \times B$. There exists, a measurable mapping $\tilde{u} : \Omega \to B$ such that, for all $x \in \Omega$, we have*

$$g(x, \tilde{u}(x)) = \min_{a\in B}\{g(x, a)\}. \tag{3.8}$$

Proposition 3.3.2. *Let $B \subset \mathbb{R}^p$ be compact and let $u : \Omega \to \mathbb{R}^p$ be an integrable mapping. Then there exists a measurable mapping $\tilde{u} : \Omega \to B$ such that, for all $x \in \Omega$, we have*

$$\|u(x) - \tilde{u}(x)\| = \mathrm{dist}(u(x), B). \tag{3.9}$$

Theorem 3.3.5. *Let $A \in \mathbb{M}^{N\times n}$ be a matrix with rank $(A) > 1$ and consider the variational integral*

$$\mathbb{I}(u) = \int_\Omega \mathrm{dist}^2_{qc}(Du, \{A, -A\})dy. \tag{3.10}$$

(1) If A^TA is not proportional to the identity matrix, then $\mathbb{I}[.]$ has a unique minimizer $u \in W^{1,2}_{u_0}(\Omega; \mathbb{R}^N)$ for all $f \in L^2(\Omega; \mathbb{R}^N)$ and all $u_0 \in W^{1,2}(\Omega; \mathbb{R}^N)$.
(2) If A^TA is proportional to the identity matrix, then there exist infinitely many minimizers of $\mathbb{I}[.]$ on $W^{1,2}_0(\Omega; \mathbb{R}^N)$ and there exists at least one minimizer that is not locally Lipschitz continuous.

This result has to be contrasted with the importance of convexity of the set K in connection with regularity results. Note that $K = \{\mathbb{I}; -\mathbb{I}\} \subset \mathbb{M}^{2\times 2}$ is a non-convex set contained in the space of conformal matrices.

Theorem 3.3.6. *If $K \subset \mathbb{M}^{2\times 2}$ is a compact and convex subset in the two-dimensional subspace of all conformal matrices, then all extremals of the $\mathbb{I}(u)$ are locally Lipschitz continuous and even of a class $C^{1,\alpha}_{loc}(\Omega;\mathbb{R}^2)$ for some $\alpha \in (0,1]$.*

The regularity statements are optimal, are all closely connected to an explicit relaxation formula for functions that depend only on the conformal part of a 2×2 matrix. We use the orthogonal decomposition (with respect to the standard inner product $\langle .,. \rangle$ in $\mathbb{M}^{2\times 2}$ of a 2×2 matrix X into its conformal and anti-conformal part, $X = X^+ + X^-$. We include a short proof of this fact in the more general setting of functions that depend only on the conformal part, of the matrix. The key observation is that the conformal part of Du has constant length whereas the anti-conformal part has a logarithmic singularity at the origin. u is in fact a minimizer of the functional.

Theorem 3.3.7. *Suppose $K = \{P \in \mathbb{R}^{N\times n}, f(P) \le \alpha\}$ is compact, with $f : \mathbb{M}^{N\times n} \mapsto \mathbb{R}$ quasiconvex, Then,*

$$\{P \in \mathbb{M}^{N\times n}, Q\mathit{dist}(P,K) = 0\} = K. \tag{3.11}$$

Which means there exists a non-negative quasiconvex function with its zero set exactly K.

Proof. As the necessity is obvious, we need to prove the sufficiency of the conditions. In fact, we just need to prove the claim for the case that $\mathrm{dist}^p(.,K)$ is rank-one convex, because other types of semiconvexity imply rank-one convexity. Firstly we consider the condition (i), $p = 2$. Suppose the claim is not true. Then there exists a closed non-convex subset K of $\mathbb{M}^{N\times n}$ with $n, N \ge 2$, such that $\mathrm{dist}^2(.,K)$ is rank-one convex (if $n = 1$ or $N = 1$, all the semiconvex relaxations are the same and equal the convex relaxation.) There is some $P_0 \in \mathbb{M}^{N\times n}$ and $A, B \in K$, $A \ne B$, such that

$$|P_0 - A|^2 = |P_0 - B|^2 = \mathrm{dist}^2(P_0,K).$$

As $\{A,B\} \subset K$,

$$\mathrm{dist}^2(P,K) \le \ \mathrm{dist}^2(P,\{A,B\}) = \min\{|P-A|^2, |P-B|^2\} := f(P). \tag{3.12}$$

Since $\mathrm{dist}(P,K)$ is rank-one convex, we have

$$\mathrm{dist}^2(P,K) \le Rf(P) = Qf(P). \tag{3.13}$$

set $l = \mathrm{dist}(P_0,K)$. We will evaluate both functions in (3.13) at P_0 to reach a contradiction. The left-hand side of (3.13) at P_0 gives l^2. To evaluate the right-hand of (3.13) at P_0, we denote the angle between $A - P_0$ and $B - P_0$, which is less than or equal to π as 2α. From Theorem 3.3.2,

$$Rf(P_0) = Qf(P_0) = \min_{0\le 1\le \theta}\{|P_0 - \theta A - (1-\theta)B|^2 + \theta(1-\theta)[|A-B|^2 - \lambda_{\max}]\}, \quad (3.14)$$

where $\lambda_{\max}$ is the greatest eigenvalue of the matrix $(A-B)^T(A-B)$. We claim that

$$0 < \lambda_{\max} < |A-B|^2.$$

Otherwise, let $C = \frac{1}{2}(A+B)$, then $Qf(C) = 0$, so that $C \in K$, while

$$|P_0 - C|^2 < l^2 = \text{dist}^2(P_0, K)^2,$$

which leads to a contradiction. It is obvious that $\lambda_{\max} > 0$. Now, we decompose $P_0 - A$ and $P_0 - B$ as

$$P_0 - A = (C - A) + (P_0 - C), \qquad P_0 - B = (C - B) + (P_0 - C).$$

Notice that $P_0 - C$ is orthogonal to $C - A$ and $C - B$, while $C - A = -(C - B)$

$$|P_0 - C| = l\cos\alpha \qquad |C - A| = l\sin\alpha.$$

We then have

$$\begin{aligned} Rf(P_0) &= \min_{0\le\theta\ge 1}\{|P_0 - C + \theta(C-A) + (1-\theta)(C-B)|^2 + \theta(1-\theta)[|A-B|^2 - \lambda_{\max}]\} \\ &= \min_{0\le\theta\ge 1}\{|P_0 - C|^2 + (1-2\theta)(C-B)^|2 + \theta(1-\theta)[|A-B|^2 - \lambda_{\max}]\} \end{aligned}$$

It is easy to say the function reaches its minimum at $\theta = \frac{1}{2}$ by employing a simple calculus argument. Therefore,

$$\begin{aligned} Rf(P_0) &= |P_0 - C|^2 + \frac{1}{4}(|A-B|^2 - \lambda_{\max}) \\ &= l^2\cos^2\alpha + l^2\cos^2\alpha - \frac{1}{4}\lambda_{\max} \\ &= l^2 - \frac{1}{4}\lambda_{\max} < l^2, \end{aligned}$$

which contradicts (3.13). The proof for case (i) is complete.
For case (ii), $1 \le p < 2$, if $\text{dist}^p(., K)$ is rank-one convex, we have, from the fact that $t^{\frac{2}{p}}$ is a monotone increasing and convex function for $t \ge 0$, that $\text{dist}^2(., K) = [\text{dist}^p(., K)]^{2/p}$ is rank-one convex (see [11] for the quasiconvex set). The conclusion for case (i) implies that K is convex. □

Remark 3.3.1. *If $p \neq 2$. We do not have explicit formula for the quasiconvexification of*

$$Q\,dist^p(., \{A, B\}).$$

Remark 3.3.2. *An interesting problem in the study of $Qdist^p(.,K)$ is to find those points P, where*

$$Qdist^p(P,K) < dist^p(P,K).$$

Corollary 3.3.2. *Suppose K is a closed non-convex subset of $\mathbb{M}^{N\times n}$. For $P \in \mathbb{M}^{N\times n}$ such that the nearest-point property is not satisfied with respect to K, then*

$$(i) P dist^2(P,K) \le Qdist^2(P,K) \le Rdist^2(P,K) \le dist^2(P,K) - \lambda_P,$$

and

$$(ii) when 1 \le p < 2,$$

$$Qdist^2(P,K) \le Rdist^2(P,K) \le (dist^2(P,K) - \lambda_p)^{p/2},$$

where

$$\lambda_p = \sup\{\frac{1}{4n}|A-B|^2,\ A,B \in K, A \neq B, |P-A| = |P-B| = dist(P,K)\}.$$

Proof. We only need to consider the rank-one convex relaxation. For case (i), that is $p = 2$, we have

$$R\mathrm{dist}^2(P_0,K) \le Rf(P_0) = Qf(P_0) = l^2 - \frac{1}{4}\lambda_{\max} = \mathrm{dist}^2(P_0,K) - \frac{1}{4}\lambda_{\max}.$$

From the definition of $\lambda_{\max}$, we see that

$$\lambda_{\max} \ge \frac{1}{n}|A-B|^2.$$

Consequently,

$$R\mathrm{dist}^2(P_0,K) \le \mathrm{dist}^2(P_0,K) - \frac{1}{4n}|A-B|^2.$$

Hence

$$\begin{aligned} R\mathrm{dist}^2(P_0,K) \le \inf\{&\mathrm{dist}^2(P_0,K) - \frac{1}{4n}|A-B|^2, \\ &A,B \in K, |P-A| = |P-B| = \mathrm{dist}(P_0,K)\} \\ &= \mathrm{dist}^2(P_0,K) - \lambda_{P_0}. \end{aligned}$$

For case (ii), notice that

$$(R\mathrm{dist}^p(.,K))^{2/p} \le (\mathrm{dist}^p(.,K))^{2/p} = \mathrm{dist}^2(.,K).$$

Since $(Q\mathrm{dist}^p(.,K)])^{2/p}$ is a rank-one convex function when $1 \le p \le 2$, we see that

$$(R\mathrm{dist}^p(.,K))^{2/p} \le R\mathrm{dist}^2(.,K).$$

At the point P_0, which does not satisfy the nearest-point property with respect to K, we have

$$(R\mathrm{dist}^2(P_0, K))^{2/p} \le R\mathrm{dist}^2(P_0, K) \le \mathrm{dist}^2(P_0, K) - \lambda P_0.$$

Thus the conclusion follows. □

Proof. The necessity part is easy: when K is convex, $\mathrm{dist}^2(., K)$ is convex, so all the semiconvex relaxations equal $\mathrm{dist}^2(., K)$ itself;

$$C\mathrm{dist}^2(P, K) = P\mathrm{dist}^2(P, K) = Q\mathrm{dist}^2(P, K) = R\mathrm{dist}^2(P, K) = \mathrm{dist}^2(P, K)$$
$$\text{for all } P \in \mathbb{M}^{N\times n}.$$

Next, we prove that the condition is sufficient. In fact, we need to prove that if K is not convex, then for every $R > 0$, there exists $P_R \in E_k \subset \mathbb{M}^{N\times n}$, with $\mathrm{dist}(P_R, K) \ge R$, such that $R\mathrm{dist}^2(P_R, K) \le \mathrm{dist}^2(P_R, K)$. The idea of this proof is to find a point which, approximately, does not satisfy the nearest-point property with respect to K and is far away from K. If $\dim C(K) \le 1$, $C(K)$ must be a closed line segment from a single point. Since K is contractible, we see that $K = C(K)$. The conclusion follows. Therefore we have we may assume that $\dim K = C(K) \ge 2$, which is contained in k-dimensional plane E_k of $\mathbb{M}^{N\times n}$. We claim that there exists a supporting plane L of $C(K)$ in $\dim L = K - 1$, such that $K \cap L$ is not convex. Otherwise, since it is easy to check that $C(K \cap L) = C(K) \cap L$, we have

$$C(K) \cap L = \partial C(K) \cap L = K \cap L.$$

Since for every relative boundary point P of $C(K)$ in E_k, there exists at least one support of plane of $C(K)$ which contains P, we see that $\partial C(K) \subset K$. As K is a convex set so we see contradiction. Let L be a supporting plane of $C(K)$ such that $K \cap L$ is not convex. Therefore, there is a point $C \in L$ such that C does not satisfy the nearest point property with respect to $K \cap L$. We may assume that $C = 0$, the zero matrix in $\mathbb{M}^{N\times n}$, so that $L \subset E \subset \mathbb{M}^{N\times n}$ are all subspaces. Hence we can find two points $A, B \in K \cap L$, $A \ne B$, such that $|A| = |B| = \mathrm{dist}(0, K \cap L)$. Let $D \in E_k$ be the unit vector in E_k which is orthogonal to the subspace L, and L separates $C(K)$ and D. Now let $P = sD$, with $s > 0$, and let us give some estimates of the value of $Q\mathrm{dist}^2(P, K)$ and $\mathrm{dist}^2(P, K)$. Since

$$|P - A| - |P - A| = l = \sqrt{s^2 + |A|^2} := R,$$

we see that

$$R\mathrm{dist}^2(P, K) \le Rf(P) = Qf(P) = R^2 - \frac{1}{4}\lambda_{\max}.$$

Now, having f, Qf and Rf $\lambda_{\max} > 0$ is the greatest eigenvalue of $(A-B)^T(A-B)$. Since for any $r > 0$, with $r < |A|$, the closed ball $B_r = \{X \in L, |X| \leq r\}$ in L, does not intersect K, that is $B_r K = 0$, we see that there exists $\gamma > 0$ such that the close cylinder,

$$B_{r,\gamma} = \{X - tD,\ X \in L,\ |X| \leq r;\ 0 \leq t \leq \gamma\}$$

does not intersect K. Hence, for sufficiently large s, if we set $l_r = \sqrt{s^2 + r^2}$, the intersection of the closed ball

$$B_{l_r} = \{P, X \in E_k,\ |X - P| \leq l_r\}$$

with the half-space

$$E_k^- = \{X - tD,\ X \in L,\ t \geq 0\}$$

is contained in $B_{r,\gamma}$, which is disjoint with $K, B_{lr}(P) \cap E_k^- \subset B_{r,\gamma}$. Therefore, for any $\epsilon > 0$ there exists an $s_0 > 0$, when $s \geq s_0$ and $P = sD$,

$$\mathrm{dist}^2(P, K) \geq R^2 - \epsilon,$$

where $l = \sqrt{s^2 + |A|^2}$. If we choose $\epsilon < \frac{1}{4}\lambda_{\max}$, there is an $s_0 > 0$, when $s \geq s_0$,

$$\mathrm{dist}^2(P, K) \geq l^2 - \epsilon > l^2 - \frac{1}{4}\lambda_{\max} = Rf(P) \geq R\mathrm{dist}^2(P, K).$$

The proof is complete. $\square$

3.4 Distance Functions to Subspaces $L \subset \mathbb{R}^{N \times n}$

We know that distance functions to subspaces are always convex because of being quadratic and non-negative. So in particular they are quasiconvex. However, if the subspace has no rank-one direction one can show a stronger assertion. (See the discussion at the end of Chapter 2.)

3.5 Distance Functions to Subspaces $L \subset \mathbb{R}^{2 \times 2}$ with No Rank-one Directions

Here we proceed by presenting various concrete examples and working out the details of the corresponding quadratic forms.

Example 1

We consider the 2-dimensional subspace L of $\mathbb{M}^{2\times 2}$

$$L = \left\{ \begin{pmatrix} a & b \\ b & -a \end{pmatrix} : a, b \in \mathbb{R} \right\}. \tag{3.15}$$

Let us find the basis which span the subspace L.

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \alpha \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \beta \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \gamma \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \delta \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

If we say $A = A_1 + A_2$ while $A_1 \in L$ and $A_2 \in L^{\perp}$, then

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \beta & -\alpha \end{pmatrix} + \begin{pmatrix} \gamma & \delta \\ -\delta & \gamma \end{pmatrix}.$$

Here

$$\begin{cases} \alpha = \frac{a_{11}-a_{22}}{2}, \\ \beta = \frac{a_{12}+a_{21}}{2}, \\ \gamma = \frac{a_{11}+a_{22}}{2}, \\ \delta = \frac{a_{12}-a_{21}}{2}. \end{cases}$$

We have then

$$\mathrm{dist}_L^2 = f(A) = \|A_2\|^2 = A_2.A_2 = \frac{(a_{11}+a_{22})^2}{2} + \frac{(a_{12}-a_{21})^2}{2}, \tag{3.16}$$

and

$$\mathrm{dist}_{L^{\perp}}^2 = h(A) = \|A_1\|^2 = A_1.A_1 = \frac{(a_{11}-a_{22})^2}{2} + \frac{(a_{12}+a_{21})^2}{2}. \tag{3.17}$$

Now if we take a matrix

$$E = \begin{pmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{pmatrix},$$

we have

$$\frac{d^2}{dt} f(A+tE)|_{t=0} = e_{11}^2 + e_{22}^2 + 2e_{11}e_{22} + e_{12}^2 + e_{21}^2 - 2e_{12}e_{21}.$$

As we see it is $2f(E)$ where, $f(E) = \mathrm{dist}_L^2(E)$, Since f is positive and quadratic, it is convex. Hence f is quasiconvex, hence f is rank-one convex.

Now since f is a quadratic function on $\mathbb{R}^{2\times 2}$, we can write $f(E) = \langle QE, E\rangle$. So

$$f(E) = \left\langle \begin{pmatrix} q_{11} & q_{12} & q_{13} & q_{14} \\ q_{21} & q_{22} & q_{23} & q_{24} \\ q_{31} & q_{32} & q_{33} & q_{34} \\ q_{41} & q_{42} & q_{43} & q_{44} \end{pmatrix} \begin{pmatrix} e_{11} \\ e_{12} \\ e_{21} \\ e_{22} \end{pmatrix}, \begin{pmatrix} e_{11} \\ e_{12} \\ e_{21} \\ e_{22} \end{pmatrix} \right\rangle$$

$$= q_{11}e_{11}^2 + q_{12}e_{12}e_{11} + q_{13}e_{21}e_{11} + q_{14}e_{22}e_{11} + q_{21}e_{11}e_{12} + q_{22}e_{12}^2 + q_{23}e_{21}e_{12} + q_{24}e_{22}e_{12}$$
$$+q_{31}e_{11}e_{21} + q_{32}e_{12}e_{21} + q_{33}e_{21}^2 + q_{34}e_{22}e_{21} + q_{41}e_{11}e_{22} + q_{42}e_{12}e_{22} + q_{43}e_{21}e_{22} + q_{44}e_{22}^2.$$

Then

$$\begin{cases} \frac{1}{2}q_{11}e_{11} = e_{11}^2 \to q_{11} = \frac{1}{2}, \\ \frac{1}{2}e_{22}^2 = q_{44}e_{22}^2 \to q_{44} = \frac{1}{2}, \\ -e_{12}e_{21} = q_{23}e_{21}e_{12} + q_{32}e_{12}e_{21} \to q_{23} + q_{32} = -1 \to q_{23} = q_{32} = -\frac{1}{2}, \\ \frac{1}{2}e_{12}^2 = q_{22}e_{12}^2 \to q_{22} = \frac{1}{2}, \\ \frac{1}{2}e_{21}^2 = q_{33}e_{21}^2 \to q_{33} = \frac{1}{2}, \\ e_{11}e_{22} = q_{14}e_{22}e_{11} + q_{41}e_{11}e_{22} \to q_{14} + q_{41} = 1 \to q_{14} = q_{41} = \frac{1}{2}, \end{cases}$$

and so

$$Q = \begin{pmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix}.$$

Then eigenvalues of Q are

$$\begin{cases} \lambda_1 = \lambda_2 = 0, \\ \lambda_3 = \lambda_4 = 1, \end{cases} \tag{3.18}$$

in line with Q being a projection. Then, the eigenvectors of Q are as follows

$$\lambda_1 = \lambda_2 = 0, \qquad \text{we have} \qquad \vec{V_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \text{ and } \vec{V_2} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix},$$

and

$$\lambda_3 = \lambda_4 = 1, \qquad \text{we have} \qquad \vec{V_3} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \text{ and } \vec{V_4} = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}.$$

As we see all the eigenvalues are ≥ 0. Hence Q is non-negative definite. Hence f is convex.

Restriction to rank-one matrices

Let us consider a rank-one matrix

$$D = (u \otimes v) = \begin{pmatrix} u_1v_1 & u_1v_2 \\ u_2v_1 & u_2v_2 \end{pmatrix}. \tag{3.19}$$

We have

$$\mathrm{dist}_L^2(A) = f(A) = \frac{(u_1v_1 + u_2v_2)^2}{2} + \frac{(u_1v_2 - u_2v_1)^2}{2}. \tag{3.20}$$

As we see

$$\frac{d^2}{dt}f(A+tD)|_{t=0} = u_1^2v_1^2 + u_1^2v_2^2 + u_2^2v_1^2 + u_2^2v_2^2 = 2\mathrm{dist}_L^2(D) > 0. \tag{3.21}$$

Hence f is convex.

Claim .

$$\mathrm{dist}_L^2(u \otimes v) \geq \epsilon \|u\|^2\|v\|^2.$$

Finding ϵ from (3.21), we have

$\frac{1}{2}[u_1^2v_1^2 + u_1^2v_2^2 + u_2^2v_1^2 + u_2^2v_2^2] = \frac{1}{2}[(u_1^2 + u_2^2) + (v_1^2 + v_2^2)] = \frac{1}{2}[\|u\|^2 + \|v\|^2]$. Hence $\epsilon = \frac{1}{2}$.

We found that $f(A)$ is convex, Now If we define $G(A) = f(A) - \sigma\|A\|^2$, it means $G(A)$ can't be convex. But it could be rank one convex only for $0 < \sigma \leq \frac{1}{2}$.

Since we can write $\mathrm{dist}_L(B) = \langle BD, D\rangle$, we have

$$\mathrm{dist}_L(B) = \langle \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \\ b_{41} & b_{42} & b_{43} & b_{44} \end{pmatrix} \begin{pmatrix} u_1v_1 \\ u_1v_2 \\ u_2v_1 \\ u_2v_2 \end{pmatrix}, \begin{pmatrix} u_1v_1 \\ u_1v_2 \\ u_2v_1 \\ u_2v_2 \end{pmatrix} \rangle$$

$= b_{11}u_1^2v_1^2 + b_{12}u_1^2v_2v_1 + b_{13}u_2v_1^2u_1 + b_{14}u_2v_2u_1v_1 + b_{21}u_1^2v_1v_2 + b_{22}u_1^2v_2^2 + b_{23}u_2v_1u_1v_2 + b_{24}u_1u_2v_2^2 + b_{31}u_1u_2v_1^2 + b_{31}u_1u_2v_1^2 + b_{32}u_1v_2u_2v_1 + b_{33}u_2^2v_1^2 + b_{34}u_2^2v_1v_2 + b_{41}u_1v_1u_2v_2 + b_{42}u_1u_2v_2^2 + b_{43}u_2^2v_1v_2 + b_{44}u_2^2v_2^2$.

Then

$$\begin{cases} b_{11}u_1^2v_1^2 = \frac{1}{2}u_1^2v_1^2 \rightarrow b_{11} = \frac{1}{2}, \\ b_{44}u_2^2v_2^2 = \frac{1}{2}u_2^2v_2^2 \rightarrow b_{44} = 1\frac{1}{2}, \\ b_{22}u_1^2v_2^2 = \frac{1}{2}u_1^2v_2^2 \rightarrow b_{22} = \frac{1}{2}, \\ b_{33}u_2^2v_1^2\frac{1}{2}u_2^2v_1^2 \rightarrow b_{33} = \frac{1}{2}, \end{cases}$$

and so

$$B = \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{pmatrix}.$$

The eigenvalues of B are

$$\lambda_1 = \ldots = \lambda_4 = \frac{1}{2}. \tag{3.22}$$

So the minimum eigenvalue is $\frac{1}{2}$, which is equal to the ϵ that we found earlier.

Determinat

Let us now consider the matrix

$$E = \begin{pmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{pmatrix}. \tag{3.23}$$

Clearly, $\det E = e_{11}e_{22} - e_{12}e_{21}$. and $\mathrm{dist}_L(E) = \langle CE, E\rangle$

$$= \langle \begin{pmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{21} & c_{22} & c_{23} & c_{24} \\ c_{31} & c_{32} & c_{33} & c_{34} \\ c_{41} & c_{42} & c_{43} & c_{44} \end{pmatrix} \begin{pmatrix} e_{11} \\ e_{12} \\ e_{21} \\ e_{22} \end{pmatrix}, \begin{pmatrix} e_{11} \\ e_{12} \\ e_{21} \\ e_{22} \end{pmatrix} \rangle$$

$= c_{11}e_{11}^2 + c_{12}e_{12}e_{11} + c_{13}e_{21}e_{11} + c_{14}e_{22}e_{11} + c_{21}e_{11}e_{12} + c_{22}e_{12}^2 + c_{23}e_{21}e_{12} + c_{24}e_{22}e_{12} + c_{31}e_{11}e_{21} + c_{32}e_{12}e_{21} + c_{33}e_{21}^2 + c_{34}e_{22}e_{21} + c_{41}e_{11}e_{22} + c_{42}e_{12}e_{22} + c_{43}e_{21}e_{22} + c_{44}e_{22}^2.$

If we compare it with $\det E$ we have

$$\begin{cases} c_{14}e_{22}e_{11} + c_{41}e_{11}e_{22} = e_{11}e_{22}, \\ c_{23}e_{21}e_{12} + c_{32}e_{12}e_{21} = -e_{12}e_{21}. \end{cases} \Rightarrow \begin{cases} c_{14} = c_{41} = \frac{1}{2}, \\ c_{23} = c_{32} = -\frac{1}{2}. \end{cases}$$

Hence

$$C = \begin{pmatrix} 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \end{pmatrix}.$$

So we can say

$$A = B - C. \tag{3.24}$$

If we consider

$$B = A - \mu C, \tag{3.25}$$

then

$$B = \begin{pmatrix} \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix} - \epsilon \begin{pmatrix} 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 & 0 & \frac{1-\mu}{2} \\ 0 & \frac{1}{2} & \frac{-1+\mu}{2} & 0 \\ 0 & \frac{-1+\mu}{2} & \frac{1}{2} & 0 \\ \frac{1-\mu}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix}.$$

The eigenvalues of B are

$$\Rightarrow \begin{cases} \lambda_1 = \frac{\mu}{2}, \\ \lambda_2 = 1 - \frac{\mu}{2}, \\ \lambda_3 = 1 - \frac{\mu}{2}, \\ \lambda_4 = \frac{\mu}{2}. \end{cases} \tag{3.26}$$

As we see if $0 < \mu < 2$, then $\lambda > 0$. and if $\mu = 1$ we have $\lambda > 0$. Also, If $\epsilon = 1$ then, $\lambda = \frac{1}{2}$ min e.v which is positive.

Example 2

Consider the one-dimensional subspace L of $\mathbb{M}^{2\times 2}$

$$L = \left\{ \begin{pmatrix} a & 0 \\ 0 & a \end{pmatrix} : a \in \mathbb{R} \right\}. \tag{3.27}$$

Let us find the basis which span the subspace L. We have

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \alpha \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \beta \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \gamma \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \delta \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

If we say $A = A_1 + A_2$ while $A_1 \in L$ and $A_2 \in L^{\perp}$, we have

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} \alpha & 0 \\ 0 & \alpha \end{pmatrix} + \begin{pmatrix} \beta & \varphi + \sigma \\ \varphi - \sigma & -\beta \end{pmatrix}.$$

$$\begin{cases} \alpha = \frac{a_{11}+a_{22}}{2}, \\ \beta = \frac{a_{11}-a_{22}}{2}, \\ \varphi = \frac{a_{12}+a_{21}}{2}, \\ \sigma = \frac{a_{12}-a_{21}}{2}. \end{cases}$$

We have

$$\operatorname{dist}_L^2(A) = f(A) = \|A_2\|^2 = A_2.A_2 = (\frac{a_{11} - a_{22}}{2}^2) + a_{12}^2 + a_{21}^2 + (\frac{a_{11} - a_{22}}{2})^2,$$

and

$$\operatorname{dist}_{L^{\perp}}^2(A) = h(A) = \|A_1\|^2 = A_1.A_1 = (\frac{a_{11} + a_{22}}{2})^2 + (\frac{a_{11} + a_{22}}{2})^2. \tag{3.28}$$

Now if we take a matrix

$$E = \begin{pmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{pmatrix}.$$

Then

$$\frac{d^2}{dt} f(A + tE)|_{t=0} = e_{11}^2 + e_{22}^2 - 2e_{11}e_{22} + 2e_{12}^2 + 2e_{21}^2, \tag{3.29}$$

which is 2 $\operatorname{dist}_L^2(E) = 2f(E)$ and ≥ 0. So f is convex. Hence f is quasiconvex, hence f is rank-one convex. Again, since f is a quadratic function on $\mathbb{R}^{2\times 2}$, we have

$$f(E) = \langle QE, E \rangle$$
$$= q_{11}e_{11}^2 + q_{12}e_{12}e_{11} + q_{13}e_{21}e_{11} + q_{14}e_{22}e_{11} + q_{21}e_{11}e_{12} + q_{22}e_{12}^2 + q_{23}e_{21}e_{12} + q_{24}e_{22}e_{12}$$
$$+ q_{31}e_{11}e_{21} + q_{32}e_{12}e_{21} + q_{33}e_{21}^2 + q_{34}e_{22}e_{21} + q_{41}e_{11}e_{22} + q_{42}e_{12}e_{22} + q_{43}e_{21}e_{22} + q_{44}e_{22}^2.$$

Comparing with $f(E)$ formula, we get

$$Q = \begin{pmatrix} \frac{1}{2} & 0 & 0 & -\frac{1}{2} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix}.$$

The eigenvalues of Q are $\lambda_1 = 0$ and $\lambda_2 = \lambda_3 = \lambda_4 = 1$ in line with Q being a projection matrix.

The eigenvectors of Q are as follows

$$\lambda_1 = 0 \qquad \text{we have} \quad \vec{V_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix},$$

and

$$\lambda_2 = \lambda_3 = \lambda_4 = 1, \qquad \text{we have} \quad \vec{V_2} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix} \vec{V_3} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \vec{V_4} = \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \end{pmatrix}.$$

As we see all the eigenvalues are ≥ 0. Hence Q is non-negative definite, and f is convex.

Restriction to rank-one matrices

We consider a rank-one matrix

$$D = (u \otimes v) = \begin{pmatrix} u_1v_1 & u_1v_2 \\ u_2v_1 & u_2v_2 \end{pmatrix}.$$

$$We have \operatorname{dist}^2_L(A) = f(A) = \frac{(u_1v_1 - u_2v_2)^2}{2} + u_1^2v_2^2 + u_2^2v_1^2. \tag{3.30}$$

We see

$$\frac{d^2}{dt} f(A + tD)|_{t=0} = u_1^2v_1^2 + u_2^2v_1^2 - 2u_1v_1u_2v_2 + 2u_1^2v_2^2 + 2u_2^2v_1^2 = 2\operatorname{dist}^2_L(D). \tag{3.31}$$

Hence f is convex.

Claim.

$$\operatorname{dist}^2_L(u \otimes v) \geq \epsilon \|u\|^2 \|v\|^2.$$

As we know, $f(A)$ is convex, Now If we define $G(A) = f(A) - \sigma\|A\|^2$, it means $G(A)$ can't be convex. But it could be rank one convex only for $0 < \sigma \leq \frac{1}{2}$. Findind the ϵ, we have, again f being a quadratic function, we have $\text{dist}_L(D)=\langle BD, D\rangle$

$$= \langle \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \\ b_{41} & b_{42} & b_{43} & b_{44} \end{pmatrix} \begin{pmatrix} u_1v_1 \\ u_1v_2 \\ u_2v_1 \\ u_2v_2 \end{pmatrix}, \begin{pmatrix} u_1v_1 \\ u_1v_2 \\ u_2v_1 \\ u_2v_2 \end{pmatrix} \rangle$$

$=b_{11}u_1^2v_1^2 + b_{12}u_1^2v_2v_1 + b_{13}u_2v_1^2u_1 + b_{14}u_2v_2u_1v_1 + b_{21}u_1^2v_1v_2 + b_{22}u_1^2v_2^2 + b_{23}u_2v_1u_1v_2 + b_{24}u_1u_2v_2^2 + b_{31}u_1u_2v_1^2 + b_{31}u_1u_2v_1^2 + b_{32}u_1v_2u_2v_1 + b_{33}u_2^2v_1^2 + b_{34}u_2^2v_1v_2 + b_{41}u_1v_1u_2v_2 + b_{42}u_1u_2v_2^2 + b_{43}u_2^2v_1v_2 + b_{44}u_2^2v_2^2$.

Then

$$B = \begin{pmatrix} \frac{1}{2} & 0 & 0 & \alpha \\ 0 & 1 & -\frac{1}{2}-\alpha & 0 \\ 0 & -\frac{1}{2}-\alpha & 1 & 0 \\ \alpha & 0 & 0 & \frac{1}{2} \end{pmatrix}.$$

Now using, $\det(B-\lambda I)=0$, we have,

$$[(\frac{1}{2}-\lambda)^2-\alpha^2][(1-\lambda)^2-(-\frac{1}{2}-\alpha)^2]=0$$

So the eigenvalues of B are

$$\begin{cases} \lambda_1 = \frac{1}{2}+\alpha \\ \lambda_2 = \frac{1}{2}-\alpha \\ \lambda_3 = \frac{3}{2}+\alpha \\ \lambda_4 = \frac{1}{2}+\alpha \end{cases}$$

As we want all the eigenvalues to be positive, so if the minimum eigenvalue is positive, we are done. If $\alpha = 0$ we have

$$\begin{cases} \lambda_1 = \frac{1}{2}, \\ \lambda_2 = \frac{1}{2}, \\ \lambda_3 = \frac{3}{2}, \\ \lambda_4 = \frac{1}{2}. \end{cases} \tag{3.32}$$

So substituting α by 0, minimum eigenvalue is $\frac{1}{2}$ and also

$$B = \begin{pmatrix} \frac{1}{2} & 0 & 0 & 0 \\ 0 & 1 & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 1 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{pmatrix}.$$

Determinant

Let us now consider a matrix $E = \begin{pmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{pmatrix}$. Clearly

$$\det\, E = e_{11}e_{22} - e_{12}e_{21}.$$

We can say

$$\mathrm{dist}_L(E) = \langle CE, E\rangle = \langle \begin{pmatrix} c_{11} & c_{12} & c_{13} & c_{14} \\ c_{21} & c_{22} & c_{23} & c_{24} \\ c_{31} & c_{32} & c_{33} & c_{34} \\ c_{41} & c_{42} & c_{43} & c_{44} \end{pmatrix} \begin{pmatrix} e_{11} \\ e_{12} \\ e_{21} \\ e_{22} \end{pmatrix}, \begin{pmatrix} e_{11} \\ e_{12} \\ e_{21} \\ 22 \end{pmatrix} \rangle =$$

$c_{11}e_{11}^2 + c_{12}e_{12}e_{11} + c_{13}e_{21}e_{11} + c_{14}e_{22}e_{11} + c_{21}e_{11}e_{12} + c_{22}e_{12}^2 + c_{23}e_{21}e_{12} + c_{24}e_{22}e_{12} + c_{31}e_{11}e_{21} + c_{32}e_{12}e_{21} + c_{33}e_{21}^2 + c_{34}e_{22}e_{21} + c_{41}e_{11}e_{22} + c_{42}e_{12}e_{22} + c_{43}e_{21}e_{22} + c_{44}e_{22}^2$.
Comparing the result with det E we will get

$$\begin{cases} c_{14}e_{22}e_{11} + c_{41}e_{11}e_{22} = e_{11}e_{22} \\ c_{23}e_{21}e_{12} + c_{32}e_{12}e_{21} = -e_{12}e_{21} \end{cases} \Rightarrow \begin{cases} c_{14} = c_{41} = \frac{1}{2}, \\ c_{23} = c_{32} = -\frac{1}{2}. \end{cases}$$

Hence,

$$C = \begin{pmatrix} 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \end{pmatrix}.$$

Then, we can show that

$$Q = B - C. \tag{3.33}$$

If we say

$$B = Q - \mu C, \tag{3.34}$$

then

$$B = \begin{pmatrix} \frac{1}{2} & 0 & 0 & -\frac{1}{2} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix} - \mu \begin{pmatrix} 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & 0 & 0 & \frac{-1-\mu}{2} \\ 0 & 1 & \frac{\mu}{2} & 0 \\ 0 & \frac{\mu}{2} & 1 & 0 \\ \frac{-1-\mu}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix}.$$

Finding eigenvalues of B, we have

$$\det(B - \lambda I) = 0,$$

so

$$\det\begin{pmatrix} \frac{1}{2}-\lambda & 0 & 0 & \frac{-1-\mu}{2} \\ 0 & 1-\lambda & \frac{\mu}{2} & 0 \\ 0 & \frac{\mu}{2} & 1-\lambda & 0 \\ \frac{-1-\mu}{2} & 0 & 0 & \frac{1}{2}-\lambda \end{pmatrix} = 0.$$

Then

$$\begin{cases} \lambda_1 = 1+\frac{\mu}{2}, \\ \lambda_2 = 1-\frac{\mu}{2}, \\ \lambda_3 = -\frac{\mu}{2}?, \\ \lambda_4 = 1+\frac{\mu}{2}. \end{cases} \tag{3.35}$$

If $-2 < \mu < 0$, we have $\lambda_i > 0$ where $i = 1, ..., 4$.

3.6 Distance Function to Subspaces $L \subset \mathbb{R}^{3\times 3}$ with No Rank-one Directions

Now we proceed to the 3-dimensional case and show the result in this form. Here we are interested in subspace L of skew-symmetric matrices. It is easily seen that in general this subspace has no rank-one directions. Indeed consider the 3-dimensional subspace L of $M_3(\mathbb{R})$ given by

$$L = \left\{ \begin{pmatrix} 0 & a & b \\ -a & 0 & c \\ -b & -c & 0 \end{pmatrix} : a, b, c \in \mathbb{R} \right\}. \tag{3.36}$$

Distanse function

Let us find the basis for subspace L. We have

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} =$$

$$+\alpha\begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \beta\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{pmatrix} + \theta\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} + \delta\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$+\eta\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} + \tau\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} + \gamma\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \mu\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} + \zeta\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

If we say $A = A_1 + A_2$ while $A_1 \in L$ and $A_2 \in L^{\perp}$, then

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} 0 & \alpha & \beta \\ -\alpha & 0 & \theta \\ -\beta & -\theta & 0 \end{pmatrix} + \begin{pmatrix} \gamma & \delta & \eta \\ \delta & \mu & \tau \\ \eta & \tau & \zeta \end{pmatrix}$$

Then we have

$$\begin{cases} \alpha = \frac{a_{12}-a_{21}}{2} \\ \beta = \frac{a_{13}-a_{31}}{2} \\ \theta = \frac{a_{23}-a_{32}}{2} \end{cases} \quad \begin{cases} \delta = \frac{a_{12}+a_{21}}{2} \\ \eta = \frac{a_{13}+a_{31}}{2} \\ \tau = \frac{a_{23}+a_{32}}{2} \end{cases} \quad \begin{cases} \gamma = a_{11}^2 \\ \mu = a_{22}^2 \\ \zeta = a_{33}^2 \end{cases}$$

Finding distance A to L, we have

$$f(A) = \|A_2\|^2 = A_2.A_2 = a_{11}^2 + a_{22}^2 + a_{33}^2 + \frac{(a_{12}+a_{21})^2}{2} + \frac{(a_{13}+a_{31})^2}{2} + \frac{(a_{23}+a_{32})^2}{2}, \tag{3.37}$$

and also for distance A to $L^{\perp}$, we have

$$h(A) = \|A_1\|^2 = A_1.A_1 = \frac{(a_{12}-a_{21})^2}{2} + \frac{(a_{13}-a_{31})^2}{2} + \frac{(a_{23}-a_{32})^2}{2}. \tag{3.38}$$

Now if we take a matrix

$$E = \begin{pmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{pmatrix},$$

we can see

$$\begin{aligned} &\frac{d^2}{dt} f(A+tE)|_{t=0} \\ &= 2e_{11}^2 + 2e_{22}^2 + 2e_{33}^2 - 2e_{12}e_{21} + 2e_{23}^2 + e_{32}^2 - 2e_{12}e_{21} \\ &= 2\text{dist}_L^2(E) \geq 0. \end{aligned}$$

Hence

$$f \text{ is convex} \Rightarrow f \text{ is quasiconvex} \Rightarrow f \text{ is rank-one convex.}$$

Now since f is a quadratic function, we can write

$$\begin{aligned} f(E) = \langle QE, E\rangle \\ &= q_{11}e_{11}^2 + q_{12}e_{12}e_{11} + q_{13}e_{21}e_{11} + q_{14}e_{22}e_{11} \\ &+ q_{21}e_{11}e_{12} + q_{22}e_{12}^2 + q_{23}e_{21}e_{12} + q_{24}e_{22}e_{12} \\ &+ q_{31}e_{11}e_{21} + q_{32}e_{12}e_{21} + q_{33}e_{21}^2 + q_{34}e_{22}e_{21} \\ &+ q_{41}e_{11}e_{22+}q_{42}e_{12}e_{22} + q_{43}e_{21}e_{22} + q_{44}e_{22}^2. \end{aligned}$$

Finding Q and its eigenvalues, we see that all the λs are ≥ 0. So Q is non-negative definite. Hence, f is convex, in line with Q being a projection.

3.7 A Subspaces $L \subset \mathbb{R}^{2\times 2}$ with Rank-one Directions

We consider the 2-dimensional subspace L of $\mathbb{M}_2(\mathbb{R})$

$$L = \left\{ \begin{pmatrix} a & b \\ b & a \end{pmatrix} : a, b \in \mathbb{R} \right\}.$$

Distanse function

Let us find the basis for subspace L.

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \alpha \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \beta \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + \gamma \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \delta \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

If we say $A = A_1 + A_2$ while $A_1 \in L$ and $A_2 \in L^{\perp}$, then

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix} + \begin{pmatrix} \gamma & \delta \\ -\delta & -\gamma \end{pmatrix}.$$

$$\begin{cases} \alpha = \frac{a_{11}+a_{22}}{2}, \\ \beta = \frac{a_{12}+a_{21}}{2}, \\ \gamma = \frac{a_{11}-a_{22}}{2}, \\ \delta = \frac{a_{12}-a_{21}}{2}. \end{cases}$$

Finding distance A to L and $L^{\perp}$ we have

$$f(A) = \|A_2\|^2 = A_2.A_2 = \frac{(a_{11}-a_{22})^2}{2} + \frac{(a_{12}-a_{21})^2}{2},$$

and

$$h(A) = \|A_1\|^2 = A_1.A_1 = \frac{(a_{11}+a_{22})^2}{2} + \frac{(a_{12}+a_{21})^2}{2}.$$

So

$$\operatorname{dist}_L^2(A) = f(A) = \frac{(a_{11}-a_{22})^2}{2} + \frac{(a_{12}-a_{21})^2}{2}.$$

Now if we take a matrix

$$E = \begin{pmatrix} e_{11} & e_{12} \\ e_{21} & e_{22} \end{pmatrix}.$$

We have

$$\frac{d^2}{dt} f(A+tE)|_{t=0} = e_{11}^2 + e_{22}^2 - 2e_{11}e_{22} + e_{21}^2 + e_{21}^2 - 2e_{12}e_{21} = 2\ dist_L^2(E) \geq 0. \quad (3.39)$$

Hence

$$f \text{ is convex} \Rightarrow f \text{ is quasiconvex} \Rightarrow f \text{ is rank-one convex.}$$

Now since f is a quadratic function, we have

$f(E) = \langle QE, E\rangle = q_{11}e_{11}^2 + q_{12}e_{12}e_{11} + q_{13}e_{21}e_{11} + q_{14}e_{22}e_{11} + q_{21}e_{11}e_{12} + q_{22}e_{12}^2 + q_{23}e_{21}e_{12} + q_{24}e_{22}e_{12} + q_{31}e_{11}e_{21} + q_{32}e_{12}e_{21} + q_{33}e_{21}^2 + q_{34}e_{22}e_{21} + q_{41}e_{11}e_{22+}q_{42}e_{12}e_{22} + q_{43}e_{21}e_{22} + q_{44}e_{22}^2$.

So

$$Q = \begin{pmatrix} \frac{1}{2} & 0 & 0 & -\frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{2} & \frac{1}{2} & 0 \\ -\frac{1}{2} & 0 & 0 & \frac{1}{2} \end{pmatrix}.$$

The eigenvalues of Q are $\lambda_1 = \lambda_2 = 0$ and $\lambda_3 = \lambda_4 = 1$. As we see as $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ are ≥ 0, Q is non-negative definite. Hence, f is convex in line with Q being a projection.

Restriction to rank-one matrices

Let us consider a rank-one matrix

$$D = (u \otimes v) = \begin{pmatrix} u_1v_1 & u_1v_2 \\ u_2v_1 & u_2v_2 \end{pmatrix}.$$

We have

$$\begin{aligned} \operatorname{dist}_L^2(A) = f(A) &= \frac{(u_1v_1 - u_2v_2)^2}{2} + \frac{(u_1v_2 - u_2v_1)^2}{2} \\ &= [(u_1^2 + u_2^2) + (v_1^2 + v_2^2)] - 4u_1v_1u_2v_2 = [\|u\|^2 + \|v\|^2] - 2u_1v_1u_2v_2. \end{aligned}$$

As we see

$$\frac{d^2}{dt}f(A + tD)|_{t=0} = u_1^2v_1^2 + u_1^2v_2^2 + u_2^2v_1^2 + u_2^2v_2^2 - 4u_1v_1u_2v_2 = 2\operatorname{dist}_L^2(D) \geq 0,$$

hence f is convex.

Since, f is a quadratic function, we can find a matrix B from $\operatorname{dist}_L(B) = \langle BD, D\rangle$. Finding the eigenvalues of B, we notice that the following claim

$$\operatorname{dist}_L^2(u \otimes v) \geq \epsilon\|u\|^2\|v\|^2, \tag{3.40}$$

could not be true in this case.

3.8 Distance Function to the Special Orthogonal Group $K = \mathbf{SO}(n)$

Before giving the definitions, let us introduce some notations. In this section we let $O(n)$ be the set of $N \times n$ orthogonal matrices,

$$D := (0, 1)^n \subset \mathbb{R}^n$$

and $W^{1,\infty}_{per}(D;\mathbb{R}^n)$ be the space of periodic functions in $W^{1,\infty}(D;\mathbb{R}^N)$, meaning that

$$u(x) = u(x+e_i), \text{ for every } x \in \mathbb{D} \quad \text{and} \quad i = 1, ..., n,$$

where $\{e_1, ..., e_n\}$ is the standard orthogonal basis of $\mathbb{R}^n$, $\mathcal{W}_{per}$ denotes the subspace of functions in $W^{1,\infty}_{per}(D;\mathbb{R}^n)$, whose gradients take only a finite number of values. We denote by $SO(n)$ the set of all rotations in $\mathbb{M}^{N\times n}$ i.e., orthogonal matrices with unit determinants, and we let $C(K)$ be its convex hull and $\dim(C(K))$ be its dimension. A function $f : \mathbb{R}^{N\times n} \to \mathbb{R} \cup \{\infty\}$ is said to be $SO(n)$ invariant if $f(QAR) = f(A)$ for each $A \in \mathbb{M}^{N\times n}, Q, R \in SO(n)$. Of particular importance is the case $n = 2$. It is known that in the class of $SO(2)$ invariant functions the convexity, polyconvexity, and quasiconvexity are distinct.

Now we aim to find find $f_K(A)$, where $F_K(A) = \mathrm{dist}^2(A, K)$ and $K = SO(n)$. Let A be a $N \times n$ matrix. We have

$$\begin{aligned} f(A) = \mathrm{dist}^2(A, SO(n)) &= \inf_{P\in SO(n)} \|A-P\|^2 \\ &= \inf\,[(A-P):(A-P)] \\ &= \inf\,[\mathrm{trace}\ (A-P)(A-P)^t] \\ &= \inf\,[\mathrm{trace}\ (AA^t - AP^t - PA^t + I)] \\ &= \inf\,[|A|^2 - 2\,A:P + n]. \end{aligned} \tag{1}$$

Since $A : P$ is the only term which depends on P, if we find the supremum of it, we are done with finding $\inf \|A-P\|^2$.

We let $n = 2$, then we have

$$A : P = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} : \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} = (a_{11}+a_{22})\cos\theta + (a_{21}-a_{12})\sin\theta. \tag{3.41}$$

By taking derivative in respect to θ, we have

$$\frac{d}{d\theta} = -(a_{11}+a_{22})\sin\theta + (a_{21}-a_{12})\cos\theta \tag{3.42}$$

$$\frac{d}{d\theta} = 0 \Rightarrow \tan\theta = \frac{a_{21}-a_{12}}{a_{11}+a_{22}}$$

Straight-forward calculation, we have that

$$\max\ A : P = \sqrt{|A|^2 + 2\det A} \tag{3.43}$$

Hence,

$$\mathrm{dist}^2(A, SO(2)) = \inf_{P\in SO(2)} \|A-P\|^2 = |A|^2 - 2\sqrt{|A|^2 + 2\det A} + 2 \tag{3.44}$$

Proposition 3.8.1. *$dist^2(.,SO(n))$: $\mathbb{M}^{n\times n} \to \mathbb{R}$ is not quasiconvex.*

Proof. Suppose $f(P) = \min\{|P-A|^2, |P-B|^2\}$, where $P \in \mathbb{M}^{N\times n}$ and $A, B \in \mathbb{M}^{N\times n}$ are fixed matrices. The formula for the relaxation of the squared distance function to $SO(2)$ is certainly well-known to experts and can be found in various formulations in the literature. Using the formula introduced in [32] we have

$$Qf(P) = \min_{0\leq\theta\leq 1}\{|P - \theta A - (1-\theta)B|^2 + \theta(1-\theta)[|A-B|^2 - \lambda_{\max}]\}, \tag{3.45}$$

where $\lambda_{\max}$ is the greatest eigenvalue of the matrix $(A-B)^T(A-B)$. Now, let $A = I$, $B = J$, where I is the $n \times n$ identity matrix and

$$J = \begin{pmatrix} -I_2 & 0 \\ 0 & I_{n-2} \end{pmatrix}.$$

Here I_2 and I_{n-2} are 2×2 and $(n-2)\times(n-2)$ identity matrices respectively.
Consider $F(P) = \min\ \{|P-I|^2, |P-J|^2\}$, for $P \in M^{n\times n}$. Since $I, J \in SO(n)$, we can say, $F(P) \geq \text{dist}^2(P, SO(n))$. Let's say $\text{dist}^2(P, SO(n))$ is quasiconvex, then by Definition of quasiconvexification, $QF(P) \geq \text{dist}^2(P, SO(n))$, for every $P \in \mathbb{M}^{n\times n}$.
Now, we can check if it is true for the particular case when P is the $n\times n$ zero matrix. In fact, if we show $QF(0) \geq \text{dist}^2(0, SO(n))$, is not true we are done.
We have

$$\text{dist}^2(0, SO(n)) = n \tag{3.46}$$

When $A = I$, $B = J$, $\lambda_{\max}$ would be the greatest eigenvalue for the matrix $(I-J)^T(I-J)$. Since,

$$(I-J)^T(I-J) = \begin{pmatrix} 4 & 0 & 0 & \dots & 0 \\ 0 & 4 & 0 & \dots & 0 \\ 0 & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 0 \end{pmatrix},$$

then, $\lambda_{\max} = 4$.
Now we aim to find the $QF(0)$.

$$QF(0) = \min_{0\leq\theta\leq 1}\{|\theta I + (1-\theta)J|^2 + \theta(1-\theta)[|I-J|^2 - \lambda_{\max}]\}$$

$$= n - 2 + \min_{0\leq\theta\leq 1}\{|\theta I_2 - (1-\theta)I_2|^2 + \theta(1-\theta)[|2I_2|^2 - \lambda_{\max}]\}$$

$$= n - 2 + \min_{0\leq\theta\leq 1}\{2(2\theta-1)^2 + 4\theta(1-\theta)\} = n-1.$$

As we see $n - 1 < n$. So there is contradiction. The proof is complete. ☐

Claim

$$\text{dist}^2_{SO(2)}(A) = \|\pi_{L^\perp}(A)\|^2 + dist^2_{SO(2)}(\pi(A)). \tag{3.47}$$

Proof. Let $A = L + L^\perp$, where L is a subspace of $M_2(\mathbb{R})$

$$L = \left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} : a, b \in \mathbb{R} \quad , \right\}$$

and, $L^\perp$ is also a subspace of $M_2(\mathbb{R})$

$$L^\perp = \left\{ \begin{pmatrix} a & b \\ b & -a \end{pmatrix} : a, b \in \mathbb{R} \quad \right\}.$$

We can say

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} \frac{a_{11}+a_{22}}{2} & \frac{a_{12}-a_{21}}{2} \\ \frac{-a_{21}+a_{12}}{2} & \frac{a_{11}+a_{22}}{2} \end{pmatrix} + \begin{pmatrix} \frac{a_{11}-a_{22}}{2} & \frac{a_{12}+a_{21}}{2} \\ \frac{a_{21}+a_{12}}{2} & \frac{-a_{11}+a_{22}}{2} \end{pmatrix}.$$

So

$$\text{dist}^2_L(A) = \frac{(a_{11} - a_{22})^2}{2} + \frac{(a_{12} + a_{21})^2}{2},$$

and

$$\text{dist}^2_{L^\perp}(A) = \frac{(a_{11} + a_{22})^2}{2} + \frac{(a_{12} - a_{21})^2}{2}.$$

As we showed earlier

$$\begin{aligned} \text{dist}^2(A, SO(2)) &= |A|^2 - 2\sqrt{|A|^2 + 2\,\text{det}A} + 2 \\ &= a_{11}^2 + a_{22}^2 + a_{12}^2 + a_{21}^2 - 2\sqrt{a_{11}^2 + a_{22}^2 + a_{12}^2 + a_{21}^2 + 2(a_{11}a_{22} - a_{21}a_{12})} + 2. \end{aligned}$$

Calculating the right-hand side, we have

$$\begin{aligned} \|\pi_{L^\perp}(A)\|^2 &= (\frac{a_{11} - a_{22}}{2})^2 + (\frac{a_{12} + a_{21}}{2})^2 + (\frac{a_{12} + a_{21}}{2})^2 + (\frac{-a_{11} + a_{22}}{2})^2 \qquad (1) \\ &= \frac{(a_{11} - a_{22})^2}{2} + \frac{(a_{12} + a_{21})^2}{2} \\ &= \frac{|A|^2 - 2\ \text{det}|A|}{2}. \end{aligned}$$

$$\begin{aligned} \text{dist}^2_{SO(2)}(\pi(A)) &= \frac{(a_{11} + a_{22})^2}{2} + \frac{(a_{12} - a_{21})^2}{2} - \qquad (2) \\ &\quad 2\sqrt{\frac{(a_{11} + a_{22})^2}{2} + \frac{(a_{12} - a_{21})^2}{2} + 2[\frac{(a_{11} + a_{22})^2}{4} - \frac{(a_{12} - a_{21})^2}{4}] + 2} \\ &= \frac{|A|^2 + 2\ \text{det}|A|}{2} - 2\sqrt{\frac{|A|^2 + 2\,\text{det}|A|}{2} + \frac{|A|^2 + 2\text{det}|A|}{2}} + 2 \\ &= \frac{|A|^2 + 2\ \text{det}|A|}{2} - 2\sqrt{|A|^2 + 2\ \text{det}\ A} + 2. \end{aligned}$$

from, (1)+(2), we have,

$$\frac{|A|^2 - 2\ \det |A|}{2} + \frac{|A|^2 + 2\ \det |A|}{2} - 2\sqrt{|A|^2 + 2\ \det A} + 2$$
$$= |A|^2 - 2\sqrt{|A|^2 + 2\ \det A} + 2,$$

which is equal to the left-hand side and so the claim is right. $\square$

Theorem 3.8.1. *Suppose that the continuous function $f : \mathbb{M}^{N\times n} \to \mathbb{R}$ is quasiconvex and that for the same real constant α, the level set*

$$K_\alpha := \{P \in \mathbb{M}^{N\times n} : f(P) \leq \alpha\} \tag{3.48}$$

is compact. Then for every $1 \leq q < +\infty$, there is a continuous quasiconvex function $g_q \geq 0$, such that

$$-C_1 + c|P|^q \leq g_p(P) \leq C_1 + C_2|P|^q \tag{3.49}$$

(It shows we have linear growth when $q = 1$.)
and

$$K_\alpha := \{P \in \mathbb{M}^{N\times n} : g_q(P) = 0\} \tag{3.50}$$

where $C_1 \geq 0,\ c > 0,\ C_2 > 0$ are constants.

Remark 3.8.1. *We shall prove this important Theorem in the chapter 4.*

As we showed earlier distance function to $SO(n)$ is not quasiconvex. Now, if we find its quasiconvexification, it is quasiconvex and from the Theorem 3.8.1 its zero set is equal to $SO(n)$ itself.

Proposition 3.8.2. *For any $1 \leq p < \infty$*

$$QK = \{X \in \mathbb{M}^{N\times n}, Q\, dist^p(X, K) = 0\}.$$

Proof. Let K_1 be

$$K_1 = \{X \in \mathbb{M}^{N\times n}, Q\, \mathrm{dist}^p(X, K) = 0\}. \tag{3.51}$$

We know $QK \subset K_1$. Now, we define a quasiconvex function $f : M^{N\times n} \to \mathbb{R}$. Let

$$\alpha_f = \sup_{X\in K} f(X) \tag{3.52}$$

and

$$f_{\alpha f}(X) = \max\{f(X) - \alpha_f, 0\}. \tag{3.53}$$

So, f_α is quasiconvex, $QK \subset f_{\alpha f}^{-1}(0)$ and $QK = \cap_f f_{\alpha f}^{-1}(0)$. We may assume that $f_\alpha^{-1}(0)$ is compact, otherwise, take the convex function

$$g(.) = \text{dist}^2(., \text{conv}K), \tag{3.54}$$

which is the squared distance function to a convex set. Therefore $f_{\alpha f} + g$ is quasiconvex. We claim that $(f_{\alpha f} + g)^{-1}(0) \subset \text{conv}K$.

This is easy to see because $f_{\alpha f} \geq 0$, and $g^{-1}(0) = \text{conv } K$. We have, for fixed $1 \leq p < \infty$,

$$Q\,\text{dist}^p(X, f_{\alpha f}^{-1}(0)) \leq \text{dist}^p(X, f_{\alpha f}^{-1}(0)) \leq \text{dist}^p(X, K) \tag{3.55}$$

for all $X \in \mathbb{M}^{N\times n}$.

Since $Q\,\text{dist}^p(X, f_{\alpha f}^{-1}(0))$ is quasiconvex, we have

$$Q\,\text{dist}^p(X, f_{\alpha f}^{-1}(0)) \leq Q\,\text{dist}^p(X, K). \tag{3.56}$$

From Theorem 3.8.1, we see that for a compact zero set $f_{\alpha f}^{-1}(0)$ of a nonnegative quasiconvex function $f_{\alpha f}$, the quasiconvexification of distance function $Q\,\text{dist}^p(X, f_{\alpha f}^{-1}(0))$ for any $1 \leq p < \infty$, the zero set remains itself. Therefore,

$$f_{\alpha f}^{-1}(0) = \{X \in \mathbb{M}^{N\times n}, Q\,\text{dist}^p(X, f_{\alpha f}^{-1}(0)) = 0\}. \tag{3.57}$$

Hence, $K_1 \subset f_{\alpha f}^{-1}(0)$ for every quasiconvex function f, thus $K_1 \subset QK$. The proof is complete. □

Proposition 3.8.3. *Let $\Omega \subset \mathbb{R}^2$ be a unit disc $B(0,1)$, $K = SO(2)$. Then $u(x,y) = \frac{1}{2}(x, -y)\ln(x^2+y^2)$ is a minimizer for the functional*

$$I[u] = \int_\Omega F_K(Du)dy = \int_\Omega dist^2(Du, SO(2))dy \tag{3.58}$$

in $W_0^{(1,2)}(\Omega; \mathbb{R}^2)$.

Proof. We use the orthogonal decomposition with respect to the standard inner product in $\mathbb{M}^{2\times 2}$ of a 2×2 matrix X into its conformal and anticonformal part, $X = X^+ + X^-$. We can show it as below,

$$X = \begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix} = \begin{pmatrix} \frac{X_{11}+X_{22}}{2} & \frac{X_{12}-X_{21}}{2} \\ \frac{X_{21}-X_{12}}{2} & \frac{X_{11}+X_{22}}{2} \end{pmatrix} + \begin{pmatrix} \frac{X_{11}-X_{22}}{2} & \frac{X_{21}+X_{12}}{2} \\ \frac{X_{21}+X_{12}}{2} & \frac{X_{22}-X_{11}}{2} \end{pmatrix}.$$

If we assume $a = X_{11} + X_{22}$, $b = X_{21} - X_{12}$, $c = X_{11} - X_{22}$, and $d = X_{21} + X_{12}$, we have

$$X^+ = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \quad \text{and} \quad X^- = \begin{pmatrix} c & d \\ d & -c \end{pmatrix}.$$

In particular, we record

$$2\,\mathrm{det}X = |X^+|^2 - |X^-|^2,$$

Let $F_K(Du) = \mathrm{dist}^2(Du, K)$ denotes the squared Euclidean distance function to the compact set K in matrix space $\mathbb{M}^{N\times n}$ and $U : \Omega \to \mathbb{R}^N$ is a mapping in Sobolev space $W_0^{1,2}(\Omega, \mathbb{R}^2)$.

First we find the formula for $F_K(X) = \mathrm{dist}^2(X, K)$. As we showed, $X = X^+ + X^-$ is an orthogonal decomposition, we have

$$F_K(X) = \mathrm{dist}^2(X, SO(2)) = \min_{R\in SO(2)} |X^+ - R|^2 + |X^-|^2. \tag{3.59}$$

Since,

$$\min |X^+ - R|^2 = \mathrm{dist}^2(X^+, R) \tag{3.60}$$

and,

$$\begin{aligned}
&\mathrm{dist}^2(X^+, R) = |X^+|^2 - 2\sqrt{|X^+|^2 + 2\,\mathrm{det}X^+} + 2 \\
&= a^2 + b^2 + a^2 + b^2 - 2\sqrt{a^2 + b^2 + a^2 + b^2 + 2(a^2 + b^2)} + 2 \\
&= 2(a^2 + b^2) - 2\sqrt{2(a^2 + b^2) + 2(a^2 + b^2)} + 2 \\
&= 2(a^2 + b^2) - 4\sqrt{a^2 + b^2} + 2 \\
&= |X^+|^2 - 2\sqrt{2}(\sqrt{2}\sqrt{a^2 + b^2}) + 2 \\
&= |X^+|^2 - 2\sqrt{2(a^2 + b^2)} + 2 \\
&= |X^+|^2 - 2\sqrt{2}|X^+| + 2.
\end{aligned}$$

Hence

$$F_K(X) = |X^+|^2 - 2\sqrt{2}|X^+| + 2 + |X^-|^2. \tag{3.61}$$

As we mentioned $2\,\mathrm{det}X = |X^+|^2 - |X^-|^2$, So we can say

$$F_K(X) = 2|X^+|^2 - 2\sqrt{2}|X^+| + 2 - 2\det X. \tag{3.62}$$

Lemma 3.8.1. *Let $g : \mathbb{R}^2 \to \mathbb{R}$ be a given function and define*

$$f : \mathbb{M}^{2\times 2} \mapsto \mathbb{R}, \quad f(X) = g(X^+),$$

then

$$f^c(X) = f^{pc}(X) = f^{qc}(X) = f^{rc}(X) = g^c(X^+). \tag{3.63}$$

In particular, if $g : [0, \infty) \to \mathbb{R}$ is a given function and if we define $f : \mathbb{M}^{2\times 2} \to \mathbb{R}$ by

$$f(X) = g(|X^+|)$$

then

$$f^c(X) = f^{pc}(X) = f^{qc}(X) = f^{rc}(X) = \tilde{g}(|X^+|), \tag{3.64}$$

where $\tilde{g}$ is the largest convex function below g with

$$g(0) = \underset{t \geq 0}{inf}\, \tilde{g}(t). \tag{3.65}$$

We aim to find the quasiconvex envelope of $F_K(X)$, where

$$F_K(X) = 2|X^+|^2 - 2\sqrt{2}|X^+| + 2 - 2\,\mathrm{det}X \tag{3.66}$$

As $2\det X$ is quasiaffine, it is quasiconvex, So we just need to find the quasiconvexification of $2|X^+|^2 - 2\sqrt{2}|X^+| + 2$. If we define a function $g : [0, \infty) \to \mathbb{R}$, we can say $f(X) = g(|X^+|)$. So,

$$g(X^+) = 2|X^+|^2 - 2\sqrt{2}|X^+|$$

and

$$D(g(X^+)) = 4(X^+) - 2\sqrt{2} = 0.$$

So

$$|X^+|^2 = \frac{2\sqrt{2}}{4} = \frac{1}{\sqrt{2}}.$$

Hence, from above lemma, we have

$$F_k^{pc}(X) = F_k^{qc}(X) \geq 1 - 2\ \det X \quad \text{for all } X. \tag{3.67}$$

Now if we are going to find find $Du(x, y)$, where

$$u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} x \\ -y \end{pmatrix} \ln{(x^2 + y^2)}.$$

So

$$u_1 = \frac{1}{2} x \ln{(x^2 + y^2)},$$
$$u_2 = \frac{1}{2}(-y) \ln{(x^2 + y^2)}.$$

We have

$$Du = \begin{pmatrix} \frac{\partial u_1}{\partial x_1} & \frac{\partial u_1}{\partial x_2} \\ \frac{\partial u_2}{\partial x_1} & \frac{\partial u_2}{\partial x_2} \end{pmatrix},$$

and

$$(Du)^+ = \frac{1}{2r^2}\begin{pmatrix} x^2 - y^2 & 2xy \\ -2xy & x^2 - y^2 \end{pmatrix}.$$

$$|Du^+|^2 = \frac{1}{4r^4}((x^2 - y^2) + 4x^2y^2)2 = frac12r^4((x^2 + y^2)^2) = \frac{1}{2r^4}r^4 = \frac{1}{2}. \tag{3.68}$$

Hence, $|Du^+|^2 = \frac{1}{\sqrt{2}}$. a.e. on the boundary. And consequently

$$F_K(Du) = F_K^{qc}(Du) = 1 - 2\det Du \quad \text{on}\, \Omega. \tag{3.69}$$

Now let $\varphi \in W_0^{1,2}(\Omega; \mathbb{R}^2)$. Then we estimate in a routine way

$$\begin{aligned}
&\int_\Omega F_K(Du + D\varphi)dy \geq \int_\Omega F_K^{qc}(Du + D\varphi) \\
&\geq \int_\Omega (1 - 2\det(Du + D\varphi))dy = \int_\Omega (1 - 2\det Du)dy \\
&= \int_\Omega F_K(Du)dy.
\end{aligned}$$

□

Suppose $f(P) = \min\{|P - A|^2, |P - B|^2\}$, where $P \in \mathbb{M}^{N\times n}$ and $A, B \in \mathbb{M}^{N\times n}$ are fixed matrices. As we showed the formula of the quasiconvexification of f is,

$$Qf(P) = \min_{0\leq\theta\leq 1}\{|P - \theta A - (1-\theta)B|^2 + \theta(1-\theta)[|A - B|^2 - \lambda_{\max}]\}, \tag{1}$$

where $\lambda_{\max}$ is the greatest eigenvalue of the matrix $(A - B)^T(A - B)$. Now, let $P = X$, $A = I$ and $B = -I$. We have

$$f(X) = \min\{|X - I|^2, |X + I|^2\}.$$

In order to find quasiconvex envelope for f, we have

$$Qf(X) = \min_{0\leq\theta\leq 1}\{|X - \theta I - (1-\theta)(-I)|^2 + \theta(1-\theta)[|I - (-I)|^2 - \lambda_{\max}]\}.$$

First, we find $\lambda_{\max}$.

$$I - (-I) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

So

$$(I - (-I))^T = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix},$$

and

$$(I - (-I))^T(I - (-I)) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}.$$

Hence $\lambda_{\max} = 4$.

Considering $X = X^+ + X^-$, and $2\det = |X^+|^2 + |X^-|^2$, we have

$$\begin{aligned}
Qf(X) =& \min_{0\le\theta\le 1}\{|X-\theta I + I - \theta I|^2 + \theta(1-\theta)[|2I|^2 - 4]\}\\
=& \min_{0\le\theta\le 1}\{|X+(1-2\theta)I|^2 + \theta(1-\theta)4\}\\
=& \min_{0\le\theta\le 1}\{|X^+ + (1-2\theta)I|^2 + |X^-|^2 + \theta(1-\theta)4\}\\
=& |X^-|^2 + \min_{0\le\theta\le 1}\{|X^+ + (1-2\theta)I|^2 + \theta(1-\theta)4\}\\
=& |X^-|^2 - |X^+|^2 + \min_{0\le\theta\le 1}\{|X^+|^2 + |X^+ + (1-2\theta)I|^2 + \theta(1-\theta)4\}.
\end{aligned}$$

Now we are going to find

$$\min_{0\le\theta\le 1}\{|X^+ + (1-2\theta)I|^2 + \theta(1-\theta)4\}. \tag{3.70}$$

As

$$X^+ = \begin{pmatrix} a & -b \\ b & a \end{pmatrix},$$

we have,

$$\begin{aligned}
&\min_{0\le\theta\le 1}\left\{\left|\begin{pmatrix} a & -b \\ b & a \end{pmatrix} + \begin{pmatrix} 1-2\theta & 0 \\ 0 & 1-2\theta \end{pmatrix}\right|^2 + \theta(1-\theta)4\right\}\\
=&\min_{0\le\theta\le 1}\left\{\left|\begin{pmatrix} a+1-2\theta & -b \\ b & a+1-2\theta \end{pmatrix}\right|^2 + \theta(1-\theta)4\right\}\\
=&\min_{0\le\theta\le 1}\{2[(a+1-2\theta)^2 + b^2] + \theta(1-\theta)4\}.
\end{aligned}$$

Taking derivative in respect to θ we have

$$\begin{aligned}
\frac{d}{d\theta} = 0 &\Rightarrow 4(-2)(a+1-2\theta) + 4(1-\theta) - 4\theta = 0\\
&\Rightarrow -8a + 8\theta - 4 = 0 \Rightarrow \theta = \frac{2a+1}{2}.
\end{aligned}$$

Substituting θ with $\frac{2a+1}{2}$, in the formula (1) we get

$$\begin{aligned}
Qf(X) &= \left\{\left|\begin{pmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{pmatrix} - \begin{pmatrix} \frac{2a+1}{2} & 0 \\ 0 & \frac{2a+1}{2} \end{pmatrix} - \begin{pmatrix} \frac{2a-1}{2} & 0 \\ 0 & \frac{2a-1}{2} \end{pmatrix}\right|^2 + (\frac{2a+1}{2})(\frac{-2a+1}{2})4\right\}\\
&= \left\{\left|\begin{pmatrix} X_{11}-2a & X_{12} \\ X_{21} & X_{22}-\frac{4a}{2} \end{pmatrix}\right|^2 + (2a+1)(-2a+1)\right\}
\end{aligned}$$

As

$$\begin{cases} X_{11} = \frac{a+c}{2} & X_{12} = \frac{-b+d}{2} \\ X_{21} = \frac{b+d}{2} & X_{22} = \frac{a-c}{2} \end{cases}$$

we have

$$Qf(X) = \left\{ \left| \begin{pmatrix} \frac{a+c-4a}{2} & \frac{-b+d}{2} \\ \frac{b+d}{2} & \frac{a-c-4a}{2} \end{pmatrix} \right|^2 + (2a+1)(-2a+1) \right\}$$

$$\begin{aligned}
&=(\frac{-3a+c}{2})^2 + (\frac{-b+d}{2})^2 + (\frac{b+d}{2})^2 + (\frac{-3a-c}{2})^2 + (2a+1)(-2a+1) \\
&=\frac{1}{4}[18a^2 + 2c^2 + 2b^2 + 2d^2] - 4a^2 + 1 \\
&=\frac{1}{2}[a^2 + c^2 + b^2 + d^2] + 1.
\end{aligned}$$

Finding $Qf(X)$ in another way

Let

$$2 \ \det X = |X^+|^2 - |X^-|^2, \tag{3.71}$$

where $F_K(X) = \mathrm{dist}^2(X, K)$ denotes the squared Euclidean distance function to the compact set K in matrix space $W_0^{1,2}(\Omega, \mathbb{R}^2)$.

First we find the formula for $F_K(X) = \mathrm{dist}^2(X, K)$. Now, let $K = \{I, -I\}$,

$$\begin{aligned}
F_K(X) =&\mathrm{dist}^2(X, \{I, -I\}) = \min\{|X^+ - I|^2 + |X^+ - I|^2\} + |X^-|^2 \\
=&|X^-|^2 - |X^+|^2 + |X^+|^2 + \min\{|X^+ - I|^2 + |X^+ - I|^2\} \\
=&-2 \ \det X + |X^+|^2 + \{\min\{|X^+ - I|^2 + |X^+ - I|^2\} \\
=&-2 \ \det X + g(X^+).
\end{aligned}$$

Hence by lemma (3.5.1) and a direct calculation

$$F_K^{pc} = F_K^{qc} = -2 \det X + g^c(X^+) = -2 \det + \mathrm{tr}(JX)^2 + f(\mathrm{tr}(X)), \tag{3.72}$$

where

$$f(a) = \begin{cases} a^2 + 2a + 2 \quad \text{when } a \leq -1, \\ 1 \quad \text{when } |a| \leq 1, \\ a^2 - 2a + 2 \quad \text{when } a \geq 1. \end{cases}$$

Finding $\mathrm{tr}(JX)^2$ where, J is the counterclockwise rotation by 90^o.

$$J = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

So

$$\begin{aligned}
\mathrm{tr}(JX)^2 =&\mathrm{tr} \begin{pmatrix} -X_{21} & -X_{22} \\ X_{11} & X_{12} \end{pmatrix}^2 \\
=&\mathrm{tr} \begin{pmatrix} X_{21}^2 - X_{11}X_{22} & X_{21}X_{22} - X_{12}X_{22} \\ -X_{11}X_{21} + X_{12}X_{11} & -X_{11}X_{22} + X_{12}^2 \end{pmatrix}
\end{aligned}$$

$$
\begin{aligned}
&= X_{12}^2 + X_{21}^2 - 2(X_{11}X_{22}) \\
&= \frac{1}{4}[b^2 + d^2 + 2bd + b^2 + d^2 - 2bd] - \frac{1}{2}(a+c)(a-c) \\
&= (\frac{b+d}{2}) + (\frac{-b+d}{2}) - 2(\frac{a+c}{2})(\frac{a-c}{2}) \\
&= \frac{1}{2}[2b^2 + 2d^2] - \frac{1}{2}[a^2 - c^2] \\
&= \frac{1}{2}[b^2 + d^2 - a^2 + c^2].
\end{aligned}
$$

Finding $f(\mathrm{tr}(X))$, we have

$$
\mathrm{tr}(X) = X_{11} + X_{22} = \frac{a+c}{2} + \frac{a-c}{2} = a,
$$

so

$$
f(\mathrm{tr}(X)) = f(a) = 1. \tag{3.73}
$$

If $\theta = 1$, we have

$$
\begin{aligned}
Qf(P) &= \{|X - I|^2\} \\
&= \left\{ \left| \begin{pmatrix} X_{11} - 1 & X_{12} \\ X_{12} & X_{22} - 1 \end{pmatrix} \right|^2 \right\} \\
&= \left\{ \left| \begin{pmatrix} \frac{a+c-2}{2} & \frac{-b+d}{2} \\ \frac{b+d}{2} & \frac{a-c-2}{2} \end{pmatrix} \right|^2 \right\} \\
&= (\frac{a+c-2}{2})^+ (\frac{-b+d}{2})^2 + (\frac{b+d}{2})^2 + (\frac{a-c-2}{2})^2 \\
&= \frac{1}{4}(2a^2 + 2c^2 + 2b^2 + 2d^2 - 8a + 8) \\
&= \frac{1}{2}(a^2 + c^2 + b^2 + d^2 - 4(a-1)) \\
&= \frac{1}{2}[-a^2 + c^2 + b^2 + d^2] + a^2 - 2a + 2 \\
&= \mathrm{tr}(JX)^2 + f(\mathrm{tr}(X)).
\end{aligned}
$$

If $\theta = 0$, we have

$$
\begin{aligned}
&Qf(P) = \{|X+I|^2\} \\
&= \left\{ \left| \begin{pmatrix} X_{11}+1 & X_{12} \\ X_{12} & X_{22}+1 \end{pmatrix} \right|^2 \right\} \\
&= \left\{ \left| \begin{pmatrix} \frac{a+c+2}{2} & \frac{-b+d}{2} \\ \frac{b+d}{2} & \frac{a-c+2}{2} \end{pmatrix} \right|^2 \right\} \\
&= (\frac{a+c+22^{+}}{)}(\frac{-b+d2^{2}}{)} + (\frac{b+d}{2})^2 + (\frac{a-c+2}{2})^2 \\
&= \frac{1}{4}(2a^2+2c^2+2b^2+2d^2+8a+8) \\
&= \frac{1}{2}(a^2+c^2+b^2+d^2+4(a+1)) \\
&= \frac{1}{2}[-a^2+c^2+b^2+d^2]+a^2+2a+2 \\
&= \mathrm{tr}(JX)^2 + f(\mathrm{tr}(X)).
\end{aligned}
$$

Chapter 4

Construction of Quasiconvex Fucntions: Using Maximal Function Techniques and Perturbations

4.1 Introduction

In this chapter we present some techniques for the construction of quasiconvex functions with specific qualitative properties. In a sense the chapter can be thought of as a continuation of what was described and discussed in the previous chapter. However here we go further and discuss some constructions beyond the use of distance functions and their relaxations. These not only produce quasiconvex functions with desirable properties but also provide insights into some of the open problems in the field. As such the chapter can be seen as an excursion into the works of Astala, Iwaniec, Sacksman [4],[5] Baernstein and Montgomery [6], Banuelos,[15, 16, 17, 14] Burkholder[18], Iwaniec and Martin[29], Iwaniec and Kristensen [33], Petermichel, Šverâk [48], [49], Volberg[57] and Zhang [59, 60, 63, 62].

The first topic we discuss is a classical method for constructing quasiconvex functions and brought to the fore by Kristensen and Iwaniec and is based on observation that given any suitably rank-one convex function R and any strongly quasiconvex function F the function $R + tF$ is quasiconvex for sufficiently large values of the parameter t. Naturally here the underlying methodology is that we regard the rank-one convex function R as the function that we ideally would like to show that is quasiconvex and the additional term tF as a perturbation. The method is illustrated on a family of functions that was considered by Dacorogna-Marcellini [19], namely,

$$|\xi|^4 - 2\gamma|\xi|^2\det\xi, \qquad \xi \in \mathbb{R}^{2\times 2} \tag{4.1}$$

and where $\gamma \in \mathbb{R}$ is a parameter.

Here $|\xi|$ stands for the Euclidean norm of the 2×2 matrix ξ and therefore the function is easily seen to be a homogeneous polynomial of degree 4. Let us also note that this function is polyconvex precisely when $|\gamma| \leq 1$ and rank-one convex precisely when $|\gamma| \leq 2/\sqrt{3}$ [19, 21]. Interestingly the precise range of γ for which the function is quasiconvex is still unknown. The result of Alibert and Dacorogna states that there exists a positive number $\epsilon > 0$ such that the function is quasiconvex whenever $|\gamma| \leq 1 + \epsilon$. Following Iwaniec and Krsitensen we shall see how this results can be recovered as an application of this construction. Note that upon replacing the Euclidean norm $|.|$ in (4.1) by the so-called spectral norm $\|.\|$, we obtain the function

$$\|\xi\|^4 - 2\gamma\|\xi\|^2\det\xi \tag{4.2}$$

and quite nicely apart from a constant factor this function is that considered by Burkholder in [18] in the study of martingale transforms and inequalities. See also the papers cited earlier by Baernstein and Montgomery, Banuelos, Iwaniec and Volberg.

We next present another method for constructing quasiconvex functions with interesting analytic and geometric properties and this is based on a quasiconvex modification of the squared distance function $\text{dist}^2(X, K)$ [58].

In this second part following closely the work of Zhang we present a method for designing nontrivial quasiconvex functions with a prescribed p-th growth at infinity starting from a quasiconvex function. With this results we can construct a rich class of quasiconvex functions, for example, those with linear growth at infinity and, for instance, having the two-point set $\{A, B\}$ as its zero set provided that $\text{rank(A} - \text{B)} \neq 1$. However for sets like $\mathbf{SO}(n)$ we need much deeper results to cope with the zero set. It is clear that the set $K = \mathbf{SO}(n)$ here is important for various reasons particularly in connection with the study of quasiconformal mappings. As a by-product and again following Ball and Zhang one can establish connections between these results and Tartar conjecture on sets without rank-one connections. We show that for any compact subset $K \subset R_+SO(n)$ we can construct quasiconvex functions f with $K = f^{-1}(0)$, i.e., the zero set of f, and with prescribed growth at infinity.

The main technique and basic idea for proving these results is to apply the maximal function method developed by Acerbi and Fusco in the study of weak lower semi-continuity of variational integrals [2] and an approximation result for Sobolev functions by Liu [35].

4.2 Construction by Perturbation

As mentioned earlier the discussion here we present a classical perturbation technique that has been nicely developed and implemented in the context of quasiconvex functions by Iwaniec and Kristensen. To spell out the details let $R : \mathbb{R}^{N\times n} \to \mathbb{R}$ denote a C^3-smooth function which is positively homogenous of degree $p > 3$, that is,

$$R(t\xi) = t^p R(\xi) \tag{4.3}$$

for all $\xi \in \mathbb{R}^{N\times n}$ and all $t \geq 0$.

We assume that for some $\delta > 0$ we have,

$$R''(\xi)[\eta, \eta] \geq \delta|\xi|^{p-2}|\eta|^2 \tag{4.4}$$

for all ξ and $\eta \in \mathbb{R}^{N\times n}$ with $\mathrm{rank}(\eta) \leq 1$. The left-hand side stands for the second differential of R and is defined by

$$\begin{aligned} R''(\xi)[\eta, \eta] &\equiv \frac{d^2}{dt^2} R(\xi + t\eta)|_{t=0} \\ &= \sum_{i,j=1}^{n} \sum_{a,\beta=1}^{N} \frac{\partial^2 R(\xi)}{\partial \xi_i^a \partial \xi_j^\beta} \eta_i^a \eta_j^\beta . \end{aligned}$$

The inequality (4.4) is a strict form of the well-known Legendre-Hadamard condition. Observe that this inequality extends to complex rank-one matrices of the form $\eta = A \otimes a$, where $A \in \mathbb{C}^N$ and $a \in \mathbb{R}^n$. It then reads as

$$R''(\xi)[\eta, \bar{\eta}] \geq \delta|\xi|^{p-2}|\eta|^2, \tag{4.5}$$

where $\bar{\eta}$ denotes the (component-wise) complex conjugate of η and $|\eta|^2 = \langle \eta, \bar{\eta} \rangle$.

Definition 4.2.1. *We say that a continous function $F : \mathbb{R}^{N\times n} \to \mathbb{R}$ is strongly quasiconvex of degree p if for some positive $\epsilon > 0$ we have*

$$\int_{\mathbb{R}^n} (F(\xi + \nabla\phi(x)) - F(\xi))dx \geq \epsilon \int_{\mathbb{R}^n} |\nabla\phi(x)|^p dx \tag{4.6}$$

whenever $\xi \in \mathbb{R}^{N\times n}$ and $\phi \in \mathcal{D}$.

Where $\mathcal{D}$ denotes space of maps $\phi = (\phi_1 \dots \phi_N)^T : \mathbb{R}^n \to \mathbb{R}^N$ for which each coordinate function ϕ_j vanishes outside a bounded set and has continuous partial derivatives of any order. Observe that there are functions F satisfying (4.6) for an $\epsilon > 0$, but for which $F - \rho|.|^p$ is not quasiconvex for any $\rho > 0$. In particular for $p > 3$ the condition (4.6) at a matrix $\xi \neq 0$ is a strictly weaker condition than strict, uniform quasiconvexity as defined by Evans [26].

Theorem 4.2.1. *Suppose that R is C^3 and that (4.3), (4.4) hold. Then for a function F satisfying (4.6) there exists a constant t_0 such that the function*

$$R + tF$$

is quasiconvex for each $t \geq t_0$.

4.3 Strongly Quasiconvex Functions

Let $\mathcal{A} : \mathbb{R}^{N\times n} \to \mathbb{R}^{N\times n}$ be a linear transformation. We assume that the kernel of $\mathcal{A}$ contains no rank-one matrices. Using the L^p-theory of the Riesz transforms [47] one obtains for each $p \in (1, \infty)$ and any map $\phi \in \mathcal{D}$ the bound

$$\|\mathcal{A}[\nabla\phi]\|_{L^p} \geq k_p \|\nabla\phi\|_{L^p}, \tag{4.7}$$

where $k_p = k_p(\mathcal{A})$ is a positive constant depending on p and $\mathcal{A}$ only.

Lemma 4.3.1. *The function $F : \mathbb{R}^{N\times n} \to \mathbb{R}$ defined by $F(\xi) = |\mathcal{A}\xi|^p$ is strongly quasiconvex of degree p for each $2 \leq p < \infty$.*

Proof. It is well-known that for $p \geq 2$ and for vectors X, Y in an arbitrary inner product space the inequality

$$|X|^P - |Y|^p \geq p|Y|^{p-2}\langle Y, X - Y\rangle + 2^{2-p}|X - Y|^p \tag{4.8}$$

holds. (see [37]). In particular,

$$\begin{aligned}
&\int_{\mathbb{R}^n} (F(\xi + \nabla\phi(x)) - F(\xi))dx \\
&\geq p|\mathcal{A}\xi|^{p-2}\left\langle \mathcal{A}\xi, \mathcal{A}\int_{\mathbb{R}^n} \nabla\phi(x)dx \right\rangle \\
&+ 2^{2-p}\int_{\mathbb{R}^n} |\mathcal{A}\nabla\phi(x)|^p dx \\
&= 2^{2-p}\|\mathcal{A}\nabla\phi\|^p_{L^p},
\end{aligned}$$

and involving (4.7) we obtain (4.6) with $\epsilon = 2^{2-p}k_p^p$. $\square$

The functions of Lemma 4.3.1 are convex, but not strictly convex since they are constant on translate of the kernel of $\mathcal{A}$.
Two examples of particular relevance two Lemma 4.3.1 are the functions

$$\xi \mapsto |\xi^-|^p \quad \text{and} \quad \xi \mapsto |\xi^+|^p, \qquad (1 < p < \infty) \tag{4.9}$$

defined for squared matrices

$$\xi = \begin{pmatrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{pmatrix}$$

Recall that the conformal part ξ^+ and the anticonformal part ξ^- of ξ are given by

$$\xi^\pm = 1/2 \begin{pmatrix} \xi_{11} \pm \xi_{22} & \xi_{12} \mp \xi_{21} \\ \xi_{21} \mp \xi_{12} & \xi_{22} \pm \xi_{11} \end{pmatrix}.$$

Lemma 4.3.1 applies to $\xi \mapsto |\xi^-|^p$ because the kernel of the linear transformation $\xi \mapsto \xi^-$ precisely the conformal matrices that, apart from the zero matrix, all have rank two. A similar remark applies to $\xi \mapsto |\xi^+|^p$. Using complex notation inequality (4.7) reduces to the familiar Beurling-Ahlfors inequality for the Cauchy-Riemann operators

$$\left\| \frac{\partial f}{\partial z} \right\|_{L^p} \leq A_p \left\| \frac{\partial f}{\partial \bar{z}} \right\|_{L^p}. \tag{4.10}$$

More precisely, the complex notation is facilitated via the isomorphism $i : \mathbb{R}^{2\times 2} \to \mathbb{C}^2$ defined as

$$i\left(\begin{pmatrix} \xi_{11} & \xi_{12} \\ \xi_{21} & \xi_{22} \end{pmatrix}\right) \equiv (z_1, z_2),$$

where

$$z_1 \equiv \frac{1}{2}((\xi_{11} + \xi_{22}) + i(\xi_{21} - \xi_{12})),$$
$$z_2 \equiv \frac{1}{2}((\xi_{11} - \xi_{22}) + i(\xi_{21} + \xi_{12})).$$

With usual identification $\mathbb{C} \approx \mathbb{R}^2$ we have for $f \equiv u + iv : \mathbb{C} \to \mathbb{C}$ the real Jacobi matrix

$$\nabla f = \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix}.$$

and $i(\nabla f) = (\partial f/\partial z, \bar{\partial} f/\partial \bar{z})$.

We will derive (4.10) for $p > 2$ by means of elementary properties of harmonic functions and well-known inequalities for the sharp function. (Note that it follows by partial integration that $\|\partial f/\partial z\|_{L^2} = \|\partial f/\partial \bar{z}\|_{L^2}$.) We do not suggest that this method is easier than the approach based on Riesz transforms mentioned above, however, it has the virtue of also giving a point-wise version of the inequality. Recall that for a square integrable function $f : \mathbb{C} \to \mathbb{C}$ the (centred, quadratic) sharp function is defined as

$$f^*(z) = \sup_{r>0}\Big(\int_{B_{(z,r)}} |f(x) - f_{B(z,r)}|^2 dx\Big)^{\frac{1}{2}},$$

where $f_{B(z,r)}$ denotes the average of f over $B(z,r)$. Here $\int$ is the average integral so we could show it with a bar sign. The Hardy-Littlewood-Wiener maximal inequality and the Fefferman-Stein sharp inequality imply that for each $p > 2$ there exist constants a_p, β_p such that

$$a_p\|f\|_{L^p} \leq \|f^*\|_{L^p} \leq \beta_p\|f\|_{L^p} \tag{4.11}$$

holds for all $f \in L^p(\mathbb{C}, \mathbb{C})$. (see [47].) In the statement of the next result we adopt the shorthand notation

$$\partial f \equiv \frac{\partial f}{\partial z} \quad \text{and} \quad \bar{\partial} f \equiv \frac{\partial f}{\partial \bar{z}}.$$

Lemma 4.3.2. *For smooth and compactly supported functions $f : \mathbb{C} \to \mathbb{C}$ the inequality*

$$(\bar{\partial} f)^*(z) \leq 8(\partial f)^*(z) \tag{4.12}$$

holds for all $z \in \mathbb{C}$.

Finally, the Beurling-Ahlfors inequality (4.10) is an immediate consequence of Lemma 4.3.2 and the inequalities (4.11).

Example 4.3.1. We consider the function

$$R(\xi; \gamma) \equiv |\xi|^2(|\xi|^2 - 2\gamma \det \xi) \tag{4.13}$$

defined for $\xi \in \mathbb{R}^{2\times 2}$ and where $\gamma \in \mathbb{R}$ is a parameter. In view of the identities: $|\xi|^2 = |\xi^+|^2 + |\xi^-|^2$ and $2 \det \xi = |\xi^+|^2 - |\xi^-|^2$, it takes the form

$$R(\xi; \gamma) = |\xi|^2[(1 - \gamma)|\xi^+|^2 + (1 + \gamma)|\xi^-|^2].$$

We know that $R(.; \gamma)$ is polyconvex if and only if $|\gamma| \leq 1$ and rank-one convex if and only if $|\gamma| \leq 2/\sqrt{3}$.(see [28],[30]) Observe that for $|\gamma| < 2/\sqrt{3}$ the function $R(.; \gamma)$ satisfies condition (4.4) with $p = 4$. This follows from the decomposition

$$R(\xi; \gamma) = \Big(1 - \frac{\gamma\sqrt{3}}{2}\Big)|\xi|^4 + \frac{\gamma\sqrt{3}}{2} R\big(\xi; \frac{2}{\sqrt{3}}\big). \tag{4.14}$$

However as previously mentioned, the question whether $R(.; \gamma)$ is quasiconvex for the above range of the parameter γ remains open. We fix $1 < \gamma < 2/\sqrt{3}$ and consider the following perturbation of $R(.; \gamma)$:

$$R(\xi; \gamma) + t|\xi^-|^4.$$

For t large enough, this new function becomes quasiconvex. Although since it is 4-homogeneous and it changes sign then, it is not polyconvex.

We show another example which follows by perturbing $R(.;\gamma)$ with the polyconvex function

$$F(\xi) = R(\xi;1) = 2|\xi|^2|\xi^-|^2.$$

F is also strongly quasiconvex of degree 4. Observe that

$$R(\xi;\gamma) + tF(\xi) = (1+t)|\xi|^2\left(|\xi|^2 - 2\frac{\gamma+t}{1+t}\det\xi\right).$$

If we take t sufficiently large we then recover a result of Alibert and Dacorogna, therefore, we say that the Alibert-Dacorogna-Marcellini function remains quasiconvex for some parameters larger than 1, namely

$$\gamma' \equiv \frac{\gamma+t}{1+t} > 1.$$

4.4 The Example of Alibert, Dacorogna and Marcellini

*Š*verâk proved in [48] that when $N = 2$, $n = 2$ there exist quasiconvex functions with subquadratic growth that are not polyconvex. By further developing this technique Zhang showed how to construct nontrivial quasiconvex functions with linear growth at infnity. As an application of the latter arguments Müller has constructed quasiconvex functions that are positively homogeneous of degree one that are not convex. Another particular interesting example in the list is produced by Alibert, Dacorogna and Marcellini, where again $N = 2$, $n = 2$ and here f is a homogeneous polynomial of degree four. This example allows one to illustrate the different notions of convexity by using a single real parameter γ.

Theorem 4.4.1. *Let $\gamma \in \mathbb{R}$ and let $f_\gamma : \mathbb{R}^{2\times 2} \to \mathbb{R}$ be defined as*

$$f_\gamma(\xi) = |\xi|^2(|\xi|^2 - 2\gamma\det\xi), \qquad \xi \in \mathbb{R}^{2\times 2}. \tag{4.15}$$

Then the following hold:

$$\begin{aligned}
&f_\gamma \text{ is convex} \quad \Leftrightarrow \quad |\gamma| \le \gamma_c = \frac{2\sqrt{2}}{3},\\
&f_\gamma \text{ is polyconvex} \Leftrightarrow |\gamma| \le \gamma_p = 1,\\
&f_\gamma \text{ is quasiconvex} \Leftrightarrow |\gamma| \le \gamma_q \text{ and } \gamma_q > 1,\\
&f_\gamma \text{ is rank-one convex} \Leftrightarrow |\gamma| \le \gamma_r = \frac{2}{\sqrt{3}}.
\end{aligned}$$

Note that in the above list, the conditions for f to be rank-one convex and polyconvex were established by Dacorogna and Marcellini. The other results were established by Alibert and Dacorogna. Iwaniec and Kristensen showed a method for constructing quasiconvex functions, which can also be applied to establish the third fact above. Note that as a by-product this example also provides an instance of a quasiconvex function that is not polyconvex. The question as to whether $\gamma_q = \frac{2}{\sqrt{3}}$ is still open: if it were not the case, then this would give a complete answer to Morrey's conjecture. We now proceed with the proof of the statement on the quasiconvexity of the function f_γ. Let us first start with the following theorem. This result is proved by Alibert-Dacorogna (see [3]), which is the consequence of regularity results for Laplace equation.

Theorem 4.4.2. *Let $1 < p < \infty$ and $\Omega \subset \mathbb{R}^2$ be a bounded open set. Then there exists $\epsilon = \epsilon(\Omega, p) > 0$ such that*

$$\int_\Omega \left[|\nabla\varphi(x)|^2 \pm 2 \det\,(\nabla\varphi(x))\right]^{p/2} dx \geq \epsilon \int_\Omega |\nabla\varphi(x)|^p dx \tag{4.16}$$

for every $\varphi \in W_0^{1,\infty}(\Omega;\mathbb{R}^2)$. Moreover, when $p = 4$, the inequality

$$\begin{aligned} &\int_\Omega [|\xi + \nabla\varphi(x)|^2 && (4.17)\\ &\pm 2\det(\xi + \nabla\varphi(x))]^2 dx \\ &\geq (|\xi|^2 \pm 2\,\det\xi)^2\ \mathrm{meas}\Omega + \epsilon \int_\Omega |\nabla\varphi|^4 dx && (4.18)\end{aligned}$$

holds for every $\xi \in \mathbb{R}^{2\times 2}$ and every $\varphi \in W_0^{1,\infty}(\Omega;\mathbb{R}^2)$.

The only trivial part of this (4.16) is the case $p = 2$. (In this case we can take $\epsilon = 1$ and equality, instead of inequality, holds.) Observe also that this inequality (4.16) shows that the functional on the left-hand side of (4.16) is coercive in $W_0^{1,p}(\Omega;\mathbb{R}^2)$, even though the integrand is not.

let us now, present the proof of the main theorem on the quasiconvexity of the function f_γ.

Proof. (Theorem 4.4.1: Quasiconvexity). Indeed here we have to establish the implication

$$f_\gamma \text{ is quasiconvex } \Leftrightarrow \quad \gamma \leq \gamma_q \text{ and } \gamma_q > 1.$$

In the first step we prove the existence of a γ_q with the above property and in the step 2, we show that $\gamma_q > 1$.

Step 1. *Existence of γ_q.*

We start by showing that if f_γ is quasiconvex, then f_β is quasiconvex for every $0 \le \beta \le \gamma$. Let

$$I_\gamma(\xi,\varphi) := \int_\Omega [f_\gamma(\xi + \nabla\varphi(x)) - f_\gamma(\xi)]dx$$

for every $\xi \in \mathbb{R}^{2\times 2}$ and every $\varphi \in W_0^{1,\infty}(\Omega;\mathbb{R}^2)$. We have to show that $I_\gamma(\xi,\varphi) \ge 0$ implies $I_\beta(\xi,\varphi) \ge 0$. We have to deal with two cases

Case 1. If

$$\int_\Omega [|\xi + \nabla\varphi(x)|^2 \det(\xi + \nabla\varphi(x)) \ - |\xi|^2 \det\xi]dx \le 0,$$

then the claim is trivial using the convexity of $\xi \to |\xi|^4$ and the fact that $\beta \ge 0$.

Case 2. If

$$\int_\Omega [|\xi + \nabla\varphi(x)|^2 \det(\xi + \nabla\varphi(x)) - |\xi|^2 \det\xi]dx \ge 0,$$

we observe that

$$\begin{aligned} &I_\beta(\xi,\varphi) - I_\gamma(\xi,\varphi) \\ &= 2(\gamma-\beta)\int_\Omega [|\xi + \nabla\varphi(x)|^2 \det(\xi + \nabla\varphi(x)) - |\xi|^2 \det\xi]dx \ge 0, \end{aligned}$$

as wished.

We may now define γ_q by taking the largest γ such that f_γ is quasiconvex. It exists because of the preceding observation and from the fact that

$$1 = \gamma_p \le \gamma_q \le \gamma_r = \frac{2}{\sqrt{3}}$$

and this completes step 1.

Step 2. $\gamma_q > 1$. We therefore have to show that there exists $\alpha > 0$ small enough, so that if $\gamma = 1+\alpha$, then f_γ is quasiconvex. We start with a preliminary result.

Step 2'. We prove the quasiconvexity of f_γ at 0 for $\gamma = 1+\alpha$ with $\alpha > 0$ small enough. We have to prove that

$$\int_\Omega f_\gamma(\nabla\varphi(x))dx \ge 0$$

for every $\varphi \in W_0^{1,\infty}(\Omega;\mathbb{R}^2)$ and for some $\alpha > 0$. Observe first the following algebraic inequality (we use the fact that $|\xi|^2 \ge 2\det\xi$), valid for any $\xi \in \mathbb{R}^{2\times 2}$,

$$\begin{aligned} f_\gamma(\xi) &= |\xi|^4 - 2(1+\alpha)|\xi|^2 \det\xi \\ &= \frac{1}{2}[|\xi|^4 - 4|\xi|^2\det\xi + 4(\det\xi)^2] \\ &+ \frac{1}{2}[|\xi|^4 - 4(\det\xi)^2] - 2\alpha|\xi|^2\det\xi \\ &\ge \frac{1}{2}[|\xi|^2 - 2\det\xi]^2 - \alpha|\xi|^4. \end{aligned}$$

We then integrate and use Theorem (4.4.2.) to get

$$\int_\Omega f_\gamma(\nabla\varphi(x))dx \geq (\epsilon-\alpha)\int_\Omega |\nabla\varphi(x)|^4 dx. \tag{4.19}$$

Choosing $0<\alpha<\epsilon$, we have indeed obtained the result.

Step 2". We now proceed with the general case. We already know that $\gamma_q \geq \gamma_p = 1$, so we will assume throughout this step that $\gamma \geq 1$ and we will set $\alpha = \gamma - 1$. Expanding f_γ, we find

$$\begin{aligned} f_\gamma(\xi+\eta) = f_\gamma(\xi) + \langle \nabla f_\gamma(\xi);\eta\rangle + \frac{1}{2}\langle \nabla^2 f_\gamma(\xi)\eta;\eta\rangle \\ + \langle \nabla f_\gamma(\eta);\xi\rangle + f_\gamma(\eta). \end{aligned}$$

We rewrite this as

$$f_\gamma(\xi+\eta) - f_\gamma(\xi) = A_\gamma(\xi,\eta) + B_\gamma(\xi,\eta) + C_\gamma(\xi,\eta) + D_\gamma(\eta) + E_\gamma(\eta) \tag{4.20}$$

where

$$A_\gamma(\xi,\eta) := \langle \nabla f_\gamma(\xi);\eta\rangle - 2\gamma|\xi|^2 \det\,\eta$$

$$\begin{aligned} B_\gamma(\xi,\eta) &:= \frac{1}{2}\langle \nabla^2 f_\gamma(\xi)\eta;\eta\rangle + 2\gamma|\xi|^2\det\,\eta \\ &= 4(\langle\xi;\eta\rangle)^2 + 2|\xi|^2|\eta|^2 - 4\gamma\langle\xi;\eta\rangle\langle\tilde{\xi};\eta\rangle - 2\gamma|\eta|^2\det\,\xi \\ C_\gamma(\xi,\eta) &:= \langle \nabla f_\gamma(\eta);\xi\rangle \\ &= 4\langle\xi;\eta\rangle|\eta|^2 - 4\gamma\langle\xi;\eta\rangle\det\eta - 2\gamma\langle\tilde{\xi};\eta\rangle|\eta|^2 \\ D_\gamma(\eta) &:= (1-\epsilon)f_1(\eta) + \frac{\epsilon^2}{2}|\eta|^4 \\ E_\gamma(\eta) &:= \epsilon f_1(\eta) - 2(\gamma-1)|\eta|^2\det\,\eta - \frac{\epsilon^2}{2}|\eta|^4 \\ &\geq \epsilon f_1(\eta) - (\alpha + \frac{\epsilon^2}{2})|\eta|^4. \end{aligned}$$

Observe that

$$D_\gamma(\eta) + E_\gamma(\eta) = f_\gamma(\eta).$$

From step $2'$, applying (4.18) with $\gamma = 1$ and hence $\alpha = 0$, we have that for every $\varphi \in W_0^{1,\infty}(\Omega;\mathbb{R}^2)$,

$$\int_\Omega E_\gamma(\nabla\varphi(x))dx \geq [\epsilon^2 - (\alpha+\frac{\epsilon^2}{2})]\int_\Omega |\nabla\varphi(x)|^4 dx$$

which for $\alpha > 0$ sufficiently small with respect to ϵ^2 leads to

$$\int_\Omega E_\gamma(\nabla\varphi(x))dx \geq 0. \tag{4.21}$$

We also have that for $\epsilon > 0$ and $\alpha > 0$ even smaller

$$\sigma_{\epsilon,\alpha}(\xi,\eta) = B_\gamma(\xi,\eta) + C_\gamma(\xi,\eta) + D_\gamma(\eta) \geq 0 \tag{4.22}$$

for every $\xi, \eta \in \mathbb{R}^{2\times 2}$.

By combining (4.19), (4.20) and (4.21), we see that for every $\xi \in \mathbb{R}^{2\times 2}, \varphi \in W_0^{1,\infty}(\Omega;\mathbb{R}^2)$, we have

$$\int_\Omega [f_\gamma(\xi + \nabla\varphi(x) - f_\gamma(\xi)]dx \geq \int_\Omega A_\gamma(\xi, \nabla\varphi(x))dx = 0. \tag{4.23}$$

This concludes the proof of the theorem. □

4.5 Zhang's Construction and Zhang's Lemma

Theorem 4.5.1. *Suppose that the continuous function $f : \mathbb{M}^{N\times n} \to \mathbb{R}$ is quasiconvex and that for some real constant α, the level set*

$$K_\alpha := \{P \in \mathbb{M}^{N\times n} : f(P) \leq \alpha\} \tag{4.24}$$

is compact. Then, for every $1 \leq q < +\infty$, there is a continuous quasiconvex function $g_q \geq 0$, such that

$$-C_1 + c|P|^2 \leq g_q(P) \leq C_1 + C_2|P|^q \tag{4.25}$$

and

$$K_\alpha = \{P \in \mathbb{M}^{N\times n} : g_q(P) = 0\} \tag{4.26}$$

where $c > 0, C_1 \geq 0, C_2 > 0$ are constants.

Corollary 4.5.1. *Under the assumptions of Theorem 4.5.1 without assuming that K_α is compact, for any compact subset $H \subset K_\alpha$, satisfying*

$$K_\alpha \cap (convH \backslash H) = 0, \tag{4.27}$$

and $1 \leq q < \infty$, there exists a non-negative quasiconvex function g_q satisfying (4.25) and with H as its zero set :

$$H = \{P \in \mathbb{M}^{N\times n} : g_q(P) = 0\}. \tag{4.28}$$

With these results we can construct a rich class of quasiconvex functions with linear growth.

4.6 The Maximal Function Technique

Definition 4.6.1. *(The Maximal Function).*

Let $u \in C_0^\infty(\mathbb{R}^n)$. We define

$$(M^*u)(x) = (Mu)(x) + \sum_{\alpha=1}^{n}(Mu,\alpha)(x) \tag{4.29}$$

where we set

$$(Mf)(x) = \sup_{r>0}\frac{1}{\omega_n r^n}\int\limits_{B(x,r)} |f(y)|dy \tag{4.30}$$

for every locally summable f. Here ω_n is the volume of the n dimensional unit ball.

Lemma 4.6.1. *(See [46]) If $f \in L^1(\mathbb{R}^n)$, then for every $\lambda > 0$*

$$\operatorname{meas}(\{x \in \mathbb{R}^n : (Mf)(x) > \lambda\}) \leq \frac{C(n)}{\lambda}\int\limits_{\mathbb{R}^n} |f|dx. \tag{4.31}$$

Lemma 4.6.2. *If $u \in C_0^\infty(\mathbb{R}^n)$, then $M^*u \in C^0(\mathbb{R}^n)$ and*

$$|u(x)| + \sum_{\alpha=1}^{n} +|u,\alpha| \leq (M^*u)(x) \tag{4.32}$$

for all $x \in \mathbb{R}^n$. Moreover if $p > 1$, then

$$\|M^*u\|_{L^p(\mathbb{R}^n)} \leq c(n,p)\|u\|_{W^{1,p}(\mathbb{R}^n)} \tag{4.33}$$

and if $p \geq 1$, then

$$\operatorname{meas}(\{x \in \mathbb{R}^n : (M^*u)(x) \geq \lambda\}) \leq \frac{c(n,p)}{\lambda^p}\|u\|^p_{W^{1,p}(\mathbb{R}^n)} \tag{4.34}$$

for all $\lambda > 0$.

Lemma 4.6.3. *(see [1]) Let $u \in C_0^\infty(\mathbb{R}^n)$ and $\lambda > 0$, and set*

$$H^\lambda = \{x \in \mathbb{R}^n : (M^*u)(x) < \lambda\}. \tag{4.35}$$

Then for every $x, y \in H^\lambda$ we have

$$\frac{|u(x) - u(y)|}{x - y} \leq C_n\lambda. \tag{4.36}$$

Lemma 4.6.4. *Let X be a metric space, E a subspace of X, and k a positive real number. Then any k-Lipschitz mapping from E into $\mathbb{R}$ can be extended to a k-Lipschitz mapping from X into $\mathbb{R}$.*

See [25], for the proof.

Now, we shall present the following important lemma in the prove of the Theorem 4.5.1.

Lemma 4.6.5. *Let $u_j \rightharpoonup 0$ in $W_0^{1,1}(\Omega;\mathbb{R}^N)$ and there is $K>0$ such that*

$$\int\limits_{\Omega\cap\{|Du_j|\geq K\}} |Du_j|\mathrm{d}x \to 0 \qquad \text{as } j\to\infty. \tag{4.37}$$

Then there exists a bounded sequence g_j in $W^{1,\infty}(\Omega;\mathbb{R}^N)$ such that

$$\int\limits_{\Omega} |Du_j - Dg_j|\mathrm{d}x \to 0 \qquad \text{as } j\to\infty. \tag{4.38}$$

We show the proof of the lemma.

Proof. (*Proof of the lemma 4.6.5*) For a fixed j, extend u_j by zero outside Ω so that it is defined on $\mathbb{R}^n$. Since $C_0^\infty(\mathbb{R}^n,\mathbb{R}^N)$ is dense in $W_0^{1,1}(\mathbb{R}^n,\mathbb{R}^N)$, there exists a sequence w_j in $C_0^\infty(\mathbb{R}^n,\mathbb{R}^N)$ such that

$$\|u_j - w_j\|_{W_0^{1,1}(\mathbb{R}^n;\mathbb{R}^N)} < \frac{1}{j},$$

and

$$\int\limits_{\{x\in\mathbb{R}^n:|Dw_j|\geq 2K\}} |Dw_j(x)|dx \to 0$$

as $j\to\infty$, so that we can assume that $u_j\in C_0^\infty(\mathbb{R}^n;\mathbb{R}^N)$.

For each fixed j,i, define

$$H_{i,j}^\lambda = \{x\in\mathbb{R}^n : (M^*u_j^i)(x) < \lambda\},$$

where $H_j^\lambda = \bigcap\limits_{i=1}^N H_{i,j}^\lambda$, and $\lambda\geq 4nK$.

From Lemma (4.6.3), we know that for all $x,y\in H_j^\lambda$,

$$\frac{|u_j^i(x)-u_j^i(y)|}{|x-y|} \leq C(n)\lambda. \tag{4.39}$$

Let g_j^i be a Lipschitz function extending u_j^i outside H_j^λ with Lipschitz constant not greater than $C(n)\lambda$ (Lemma 4.6.4). Since H_j^λ is an open set, we have

$$g_j^i(x) = u_j^i(x), \quad Dg_j^i(x) = Du_j^i(x)$$

for all $x\in H_j^\lambda$, and

$$\|Dg_j^i\|_{L^\infty(R^N)} \leq C(n)\lambda.$$

We may assume

$$\|g_j^i\|_{L^\infty} \leq \|u_j^i\|_{L^\infty(H_j^\lambda)} \leq C(n)\lambda \tag{4.40}$$

where set $g_j^i = (g_j^1,\dots,g_j^N)$.

In order to prove that $u_j - g_j\to 0$ strongly in $W^{1,1}(\Omega;\mathbb{R}^N)$, we have

$$\int\limits_{\Omega} |Du_j - Dg_j|dx \leq \int\limits_{\Omega\setminus H_j^\lambda} (|Du_j|+|Dg_j|)dx. \tag{4.41}$$

Considering (4.38),

$$\text{meas } (\Omega\backslash H_j^\lambda) \to 0.$$

From the definition of $H_{i,j}^\lambda$, we have

$$\Omega\backslash H_{i,j}^\lambda \subset \{x \in \Omega : (Mu_j^i)(x) \geq \lambda/2\} \cup \left\{x \in \Omega : \sum_{\alpha=1}^{n} \left(M\frac{\partial u_j^i}{\partial x_\alpha}\right)(x) \geq \lambda/2\right\},$$

and

$$\left\{x \in \mathbb{R}^n : \sum_{\alpha=1}^{n} (Mu_{j,\alpha}^i)(x) \geq \lambda/2\right\} \subset \bigcup_{\alpha=1}^{n} \left\{x \in \mathbb{R}^n : (Mu_{j,\alpha}^i)(x) \geq \frac{\lambda}{2n}\right\}.$$

Define $h : \mathbb{R}^n \to \mathbb{R}$ by

$$h(s) = \begin{cases} 0 & \text{as} \quad |s| \leq K, \\ |s| - K & \text{as} \quad |s| \geq K, \end{cases}$$

so that we can prove that

$$\left\{x \in \mathbb{R}^n : (Mu_{j,\alpha}^i)(x) \geq \frac{\lambda}{2n}\right\} \subset \left\{x \in \mathbb{R}^n : (Mh(Du_j^i))(x) \geq \frac{\lambda}{2n} - K\right\}. \tag{4.42}$$

In fact, when $Mu_{j,\alpha}^i(x) \geq \frac{\lambda}{2n}$, we have a sequence of $\epsilon_k > 0,\ \epsilon \to 0$ and a sequence of balls $B_k = B(x, \mathbb{R}_k)$ such that

$$\frac{1}{\text{meas}(B_k)} \int\limits_{B_k} |u_{j,\alpha}^i| dx \geq \frac{\lambda}{2n} - \epsilon_k$$

which implies

$$\begin{aligned} Mh(Du_j^i) &\geq \frac{1}{\text{meas}(B_k)} \int\limits_{B_k \cap \{x : |Du_j^i(x)| \geq K\}} (|Du_j^i| - K) dx \\ &\geq \frac{\lambda}{2n} - \frac{1}{\text{meas}(B_k)} \int\limits_{B_k \cap \{x : |Du_j^i| \leq K\}} |u_{j,\alpha}^i| dx \\ &- \frac{1}{\text{meas}(B_k)} \int\limits_{B_k \cap \{x : |Du_j^i(x)| \geq K\}} K dx - \epsilon_k \geq \frac{\lambda}{2n} - K - \epsilon_k. \end{aligned}$$

Now, passing to the limit $k \to \infty$, we obtain (4.42) (here we choose $\frac{\lambda}{2n} > K$). From lemma (4.6.1) we have

$$\begin{aligned} &\text{meas}\left(\left\{x \in \mathbb{R}^n : (Mh(Du_j^i))(x) \geq \frac{\lambda}{2n} - K\right\}\right) \\ &\leq \frac{1}{\frac{\lambda}{2n} - K} \int\limits_{\mathbb{R}^n} |h(Du_j^i)| dx \leq \frac{1}{\frac{\lambda}{2n} - K} \int\limits_{\{x \in \Omega : |Du_j^i| \geq K\}} |Du_j^i| dx \\ &\leq \frac{1}{\frac{\lambda}{2n} - K} \int\limits_{\{x \in \Omega : |Du_j| \geq K\}} |Du_j| dx \to 0 \end{aligned}$$

as $j \to \infty$. Also, from lemma (4.6.1) and embedding theorem, we have

$$\text{meas}(\{x \in \mathbb{R}^n : (Mu_j^i)(x) \geq \lambda/2\}) \leq \frac{1}{\lambda/2} \int_{\Omega} |u_j^i| dx \to 0,$$

as $j \to \infty$, so that we conclude that

$$\text{meas}(\Omega \backslash H_j^\lambda) \to 0 \qquad \text{as } j \to \infty.$$

□

Proof. (Proof of Theorem 4.5.1) Now we are in the position to prove the theorem (4.5.1). We can see that,

$$F(P) = \max\{0, f(P) - \alpha\} \tag{4.43}$$

is quasiconvex and satisfies assumption of theorem (4.5.1) with zero set

$$\{P \in \mathbb{M}^{n\times N} : F(P) = 0\} = K_\alpha. \tag{4.44}$$

Define

$$\tilde{f}_\alpha = \text{dist}(P; K_\alpha) \tag{4.45}$$

and

$$G(P) = Q\tilde{f}_\alpha. \tag{4.46}$$

We aim to prove that $G(P) = 0$, if and only if $P \in K_\alpha$. From the definition of quasiconvexification of $\tilde{f}_\alpha, G$ is zero on K_α. Conversely, suppose $G(P) = 0$, i.e.,

$$0 = G(P) = \inf_{\phi \in C_0^\infty(B;\mathbb{R}^N)} \frac{1}{\text{meas}(B)} \int_B \tilde{f}_\alpha(P + D\phi)\text{d}x \tag{4.47}$$

for a ball $B \subset \mathbb{R}^n$, we have a sequence $\phi_j \in C_0^\infty(\Omega; \mathbb{R}^N)$ such that for $K \geq 2\ \text{dist}(P; K_\alpha)$,

$$\begin{aligned} 0 &= \lim_{j\to\infty} \int_B \text{dist}(P + D\phi_j, K_\alpha)\text{d}x \\ &\geq \lim_{j\to\infty} \int_{B\cap\{x\in\Omega:|D\phi_j(x)|\geq K\}} [|D\phi_j| - \text{dist}(P; K_\alpha)]\text{d}x \\ &\geq \lim_{j\to\infty} K/2\ (\text{meas}\{x \in \Omega : |D\phi_j(x)| \geq K\}), \end{aligned}$$

hence,

$$\int_{\{x\in\Omega:|D\phi_j|\geq K\}} |D\phi_j|\text{d}x \to 0$$

as $j \to \infty$ and $(|D\phi_j|)$ are equi-integrable on Ω with respect to j. Then, by a vector-valued version of Dunford-Pettis theorem, [27][22], there exists a subsequence (still denoted by

ϕ_j) which converges weakly in $W_0^{1,1}(B;\mathbb{R}^N)$ to a function ϕ. Moreover, by an argument of Tartar [55], and the embedding theorems, $D\phi(x) \in \mathrm{conv}K_\alpha$ for a.e. $x \in B$, so that $\phi \in W_0^{1,\infty}(B;\mathbb{R}^N)$. Define $\sigma_j = \phi_j - \phi$. Then σ_j satisfies all assumptions of Lemma (4.6.5). Hence there exists a bounded sequence $g_j \in W_0^{1,\infty}(B;\mathbb{R}^N)$, such that

$$\int\limits_B |D\sigma_j - Dg_j|dx \to 0, \quad g_j \overset{*}{\rightharpoonup} 0 \quad \text{in } W_0^{1,\infty}(B;\mathbb{R}^N), \tag{4.48}$$

as$j \to \infty$.

Let $\{v_x\}_{x\in B}$ be the family of Young measures corresponding to the sequence Dg_j (up to a subsequence), we have

$$\limsup_{j\to\infty} \int\limits_B \tilde{f}_\alpha(P + D\phi + Dg_j)dx \tag{4.49}$$

$$\leq \lim_{j\to\infty} \int\limits_B |D\sigma_j - Dg_j|dx + \lim_{j\to\infty} \int_B \tilde{f}_\alpha(P + D\phi + D\sigma_j)dx = 0 \tag{4.50}$$

which implies

$$\int\limits_B \langle \nu_x, \tilde{f}_\alpha(P + D\phi(x) + \lambda)\rangle \; dx = 0 \tag{4.51}$$

which further implies

$$\mathrm{supp}\; \nu_x \subset K_\alpha - P - D\phi(x) \quad \text{for a.e. } \; x \in B. \tag{4.52}$$

Since $g_j \overset{*}{\rightharpoonup} 0$ in $W^{1,\infty}(B;\mathbb{R}^N)$, by Ball and Zhang [11] and (4.52), (up to a subsequence), we have,

$$0 = F(P + D\phi + Dg_j) \overset{*}{\rightharpoonup} \langle \nu_x, F(P + D\phi(x) + \lambda)\rangle \geq F(P + D\phi(x)) \tag{4.53}$$

for a.e. $x \in B$, as $j \to \infty$. By the definition of quasiconvex functions, we have

$$0 = \int\limits_B F(P + D\phi(x) \geq F(P)\mathrm{meas}(B) \tag{4.54}$$

which implies $F(P) = 0, P \in K_\alpha$.

Now, for $q > 1$, define

$$g_q(P) = \max\{[\mathrm{dist}(P, \mathrm{conv}K_\alpha)]^q, Q\mathrm{dist}(P, K_\alpha)\}. \tag{4.55}$$

As we see that g_q satisfies (4.25) and (4.26). □

4.7 Tartar's Conjecture and Some Examples

Using theorem (4.5.1), we can study the connection between out constructions and Tartar's conjecture on oscillations of gradients. [56][7]

Proposition 4.7.1. *(TARTAR'S CONJECTURE).*
Let $K \subset \mathbb{M}^{N\times n}$ be closed and has no rank-one connections, i.e. for every $A, B \in K$, rank$(A - B) \neq 1$. Let z_j be a bounded sequence in $W^{1,\infty}(\mathbb{R}^n; \mathbb{R}^N)$ and the Young measures (ν_x) associated with Dz_j satisfies $\nu_x \subset K$, and such that $f(Dz_j)$ is weak-$\star$ convergent in $L^\infty(\mathbb{R}^n)$ for every continuous $f : \mathbb{M}^{N\times n} \to \mathbb{R}$. Then (ν_x) is a Dirac mass.

The answer of this conjecture is, in general, negative.[7] However, there are a number of cases when Tartar's conjecture is known to be true for gradients under supplementary hypotheses on the set K.
(i) $K_1 = \{A, B\}$ with rank $(A - B) > 1$ [12],
(ii) $K = SO(n)$, where $n = N > 1$. [31] In fact, more generally, for $n > 1$ and

$$K_2 = \{tR : t \geq 0, R \in SO(n)\} := R_+SO(n). \tag{4.56}$$

Theorem 4.7.1. *Suppose $K \subset \mathbb{M}^{N\times n}$ has no rank-one connections and Tartar's conjecture is known to be true for K. Moreover, for any bounded $Q^{1,\infty}$ sequence with Young measures $\nu_x \subset K$ has the property that $\nu_x = \delta_T$ with T a constant matrix in K $(T = \langle \nu_x, \lambda\rangle)$. Then for any non-empty compact subset $H \subset K$, any $1 \leq p < \infty$, there exists a continuous quasiconvex function $f \geq 0$, such that*
(i) $c(p)|P|^p - C(P) \leq f(P) \leq C_1(p)(1 + |P|)^p$, with $c(p), C_1(p) > 0, C(P) \geq 0$;
(ii) $\{P \in \mathbb{M}^{N\times n} : f(P) = 0\} = H$.

Remark 4.7.1. *In the case $K = K_1$, Kohn constructs a quasiconvex function with the above properties when $p = 2$ and $n, N > 1$ arbitrary; Šverâk [48] does the same in the case $p \geq 1, n = N = 2$.*

Proof. (Proof of Theorem 4.7.1). To prove this theorem, we shall use the same argument that we used for proving theorem (4.5.1). Firstly, we construct a quasiconvex function with linear growth. Define as before

$$G(P) = \operatorname{dist}(P, H) \quad \text{and} \quad f(P) = QG(P) \tag{4.57}$$

and assume that

$$f(P) = \inf_{\phi\in C_0^\infty(B;\mathbb{R}^N)} \int_B G(P + D\phi)dx = 0 \tag{4.58}$$

to devide a sequence $\phi_j \to \phi$ in $W_0^{1,1}(B; \mathbb{R}^N)$ with $\phi \in W_0^{1,\infty}(B; \mathbb{R}^N)$. In fact, we can assume $\phi_j \in C_0^\infty(\mathbb{R}^n; \mathbb{R}^N)$ and $\phi \in W_0^{1,\infty}(\mathbb{R}^n; \mathbb{R}^N)$ supported in B. It is easy to see that $D\phi_j$ converges in measure to the set $H - P$. Let g_j be the approximate sequence in $W_0^{1,\infty}(\mathbb{R}^n; \mathbb{R}^N)$, we have the young measures $(\nu_x)_{x\in\mathbb{R}^n}$ associated with Dg_j satisfy supp

$\nu_x \subset H - P - D_\phi(x)$ for a.e. $x \in \mathbb{R}^n$. Therefore, the Young measures associated with $P + D\phi(x) + Dg_j(x)$ will be supported in $H \subset K$, so that from the assumption, they are the same Dirac measure. Since, $\langle \nu_x, \lambda \rangle = D_g(x) = 0$, $P + D\phi(x) = \text{constant} \in H$. Therefore $\phi = 0$ a.e. and $P \in H$. □

Example 4.7.1. Let $K_1 = \{A, B\}$ with $A, B \in \mathbb{M}^{N\times n}$ and assume that rank $(A-B) > 1$. It is known that there exists a non-negative quasiconvex function f with quadratic growth such that,

$$\{P \in \mathbb{M}^{N\times n} : f(P) = 0\} = \{A, B\}. \tag{4.59}$$

From Theorem (4.5.1), the zero set of quasiconvex function with linear growth $Q\text{dist}(P; K_1)$ should be K_1.

Example 4.7.2. Let $K_2 = \{P = tQ : t \geq 0,\ Q \in SO(n)\} = R_+SO(n)$ and let H be any non-empty compact subset of K_2. Then, with applying Theorem (4.7.1) and a result due to [40], [41], we can show that

$$\{P \in \mathbb{M}^{n\times n} : Q\text{dist}(P, H) = 0\} = H. \tag{4.60}$$

Here we employ the approach based on an argument of Ball [7]. Following the proof of Theorem (4.5.1), the Young measures $\{\nu_x\}_{x\in B}$ associated with Dg_j are supported in $H - P - D\phi(x)$ for a.e. $x \in B$. Let us consider the quasiconvex function (see[7]).

$$F(P) = |P|^n - n^{n/2}\det P \tag{4.61}$$

which is non-negative and has K_2 as its zero set. We have

$$0 = \liminf_{j\to\infty} \int_B F(P + D\phi + Dg_j)\mathrm{d}x \tag{4.62}$$

$$= \int_B \langle \nu_x, F(P + D\phi(x) + \lambda)\rangle \mathrm{d}x \geq \int_B F(P + D\phi(x))\mathrm{d}x \geq F(P)\ \text{meas}(B). \tag{4.63}$$

Since the function $|.|^n$ is strictly convex and

$$\int_B \langle \nu_x, \det(P + D\phi(x) + \lambda)\rangle dx = \int_B \det(P + D\phi(x))dx = \det\ P\ \text{meas}(B), \tag{4.64}$$

we have

$$\int_B \langle \nu_x, |P + D\phi(x) + \lambda|^n\rangle \mathrm{d}x = \int_B |P + D\phi(x)|^n \mathrm{d}x = |P|^n\ \text{meas}(B), \tag{4.65}$$

which implies

$$\nu_x = \delta_0,\ \textit{and}\ D\phi(x) = 0 \quad \textit{a.e.} \tag{4.66}$$

so that $P \in H$.

Remark 4.7.2. Since any non-empty compact subset of $R_+SO(n)$ can be the zero set of some negative quasiconvex function, the topology of zero sets for quasiconvex functions can be very complicated. For example let K be any compact subset of $\mathbb{R}^2$, define

$$K_1 = \left\{ \begin{pmatrix} a & b \\ -b & a \end{pmatrix} : (a,b) \in K \right\}, \tag{4.67}$$

then $K_1 \subset R_+SO(2)$ and has the same topology as K.

Remark 4.7.3. The method used in theorem (4.5.1) depends heavily on the compactness of the level set k_α.

4.8 Müler's Improvement and a Variant of Zhang's Lemma

Let $\{u_j\}$ be a sequence of weakly differentiable functions $u_j : \mathbb{R}^n \to \mathbb{R}^N$ whose gradients approach the ball $B(0,R)$ in the mean, i.e.

$$\int_{\mathbb{R}^n} \operatorname{dist}(Du_j, B(0,R))dx \to 0. \tag{4.68}$$

Motivated by work of Acerbi and Fusco [2],[1] and Liu [35], Kewei Zhang showed that the sequence can be modified on a small set in such a way that the new sequence is uniformly Lipschitz. The following theorem is a slight variant of Zhang's lemma which is used in [58].

Theorem 4.8.1. *There exists a constant $c(n,N)$ with the following property. If* (4.68) *holds, then there exists a sequence of functions $v_j : \mathbb{R}^n \to \mathbb{R}^N$ such that*

$$\|Dv_j\|_\infty \le c(n,N)R, \quad \mathcal{L}^n(\{u_j \ne v_j\}) \to 0. \tag{4.69}$$

In fact one has the seemingly stronger conclusions

$$\mathcal{L}^n(\{u_j \ne v_j \text{ or } Du_j \ne Dv_j\}) \to 0, \qquad \int_{\mathbb{R}^n} |Du_j - Dv_j| dx \to 0. \tag{4.70}$$

For the first conclusion it suffices that for weakly differentiable functions u and v the implication

$$u = v \text{ a.e. in } A \quad \Rightarrow \quad Du = Dv \text{ a.e. in } A \tag{4.71}$$

holds.(See [12].) For the second conclusion observe that

$$|Du_j - Dv_j| \le |Dv_j| + |Du_j| \le c(n,N)R + R + \operatorname{dist}(Du_j, B(0,R))$$

and integrate over the set $\{Du_j \neq Dv_j\}$. Theorem (4.8.1) has found important applications to the calculus of variations, in particular the study of quasiconvexity, lower semicontinuity, relaxation and gradient Young measures. The purpose of this section is to show that the constant $c(n, N)$ can be chosen arbitrarily close to 1 and that the ball $B(0, R)$ can be replaced by a compact, convex set.

Theorem 4.8.2. *Let K be a compact convex set in $\mathbb{R}^{N\times n}$. Suppose $u_j \in W^{1,1}_{\text{loc}}(\mathbb{R}^n, \mathbb{R}^N)$ and*

$$\int_{\mathbb{R}^n} \operatorname{dist}(Du_j, K)dx \to 0. \tag{4.72}$$

Then there exists a sequence v_j of Lipschitz functions such that

$$\|\operatorname{dist}(Dv_j, K)\|_\infty \to 0, \qquad \mathcal{L}^n\{u_j \neq\ v_j\} \to 0.$$

Remark 4.8.1. A more natural and apparently much harder question is whether the same assertion holds if K is quasiconvex rather than convex. Let us denote the convex hull of K by $C(K)$ and $C\operatorname{dist}_K$ be the convex relaxation of the distance function.

$$\begin{aligned}
&\inf\{\int_B \operatorname{dist}(Dv, K_\gamma)dy \ : \ v = u \text{ on } \partial B\} \\
&= \inf\{\int_B C\operatorname{dist}(Dw, K_\gamma) \ : \ w = u \text{ on } \partial B\} \\
&\leq \int_B \operatorname{dist}(D\tilde{u}, (CK)_\gamma) \\
&\leq (1 - 3^{-n}) \int_B \operatorname{dist}(Du, CK)dx.
\end{aligned}$$

A similar argument can be applied for $N > 1$ provided that a condition holds which is slightly stronger than the requirenment that CK agrees with the quasiconvex hull QK of K.

Application to Quasiconvex Functions

As we mentioned earlier, quasiconvexity plays a crucial role in the vector-valued calculus of variations.[8][50] It states that affine functions minimize the functional $u :\to \int_\Omega f(Du)$ subject to their own boundary conditions. We also noticed that quasiconvexity is difficult to handle since no local characterization is known for $n, N > 1$ (and can not exist for $N \geq 3, n \geq 2$; (see [34]). Even the approximation of general quasiconvex functions by a more manageable subclass is a largely open question. We now comment further on this in the following corollary, as a result of Theorem 4.8.2.. We remark that every real-valued quasiconvex function is continuous and even locally Lipschitz, since it is rank-one convex. (see [20].)

Corollary 4.8.1. *Let $K \subset \mathbb{R}^{N\times n}$ be a convex, compact set with non-empty interior. Let $f : \mathbb{R}^{N\times n} \to \mathbb{R} \cup \{-\infty, \infty\}$ be a quasiconvex that satisfies*

$$f \in C(K;\mathbb{R}), \qquad f = +\infty \quad \text{on} \quad \mathbb{R}^{N\times n}\backslash K. \tag{4.73}$$

Then, for all $F \in K$,

$$f(F) = \sup\{g(F) = \mid g : \mathbb{R}^{N\times n} \to \mathbb{R}, \quad g \leq f \text{ on } K,\ g \text{ quasiconvex}\}. \tag{4.74}$$

Proof. **1.** We may assume $0 \in$ int K, since quasiconvexity is invariant under translation in $\mathbb{R}^{N\times n}$. We have

$$K \subset \lambda \text{int } K, \quad \forall \lambda > 1. \tag{4.75}$$

Indeed, if $A \in \partial K$, then $tA + (1-t)B \in K$ for all $t \in (0,1)$, and all B in a small neighbourhood of zero. Hence $tA \in$ int K, $\forall t \in (0,1)$. Thus (4.75) holds.

2. Let G_∞ denote the right hand side of (4.74) and let P denote the nearest neighbour projection onto K. For $k \in \mathcal{N} \cup \{0\}$ define

$$h_k(F) = f(PF) + k \ \text{dist}(F, K) \leq f(F).$$

Let $g_k = h_k^{qc}$ denote the quasiconvex envelope of h_k, i.e. the largest quasiconvex function below h_k. Thus $g_k(F) \leq G_\infty$. On the other hand, by standard relaxation results

$$g_k(F) = \inf\{\int\limits_Q h_k(Du)dx : u - Fx \in W_0^{1,\infty}(Q, \mathbb{R}^N)\},$$

where $Q = (0,1)^n$. Hence there exist Lipschitz functions u_k such that

$$\limsup_{k\to\infty} \int\limits_Q h_k(Du_k)dx \leq G_\infty, \quad u_k = Fx \text{ on } \partial Q. \tag{4.76}$$

In particular,

$$\int\limits_Q \text{dist}(Du_k, K)dx \to 0. \tag{4.77}$$

Hence Du_k is bounded in L^1, and after possible passage to the subsequence we may assume that $u_k \to u_0$ in L^1.

3. Let us first present the following Theorem.

Theorem 4.8.3. *Let K be a compact, convex set in $\mathbb{R}^{N\times n}$, let $\Omega \subset \mathbb{R}^n$ be open and let $\{u_j\}$ be a sequence in $W^{1,1}_{loc}(\Omega;\mathbb{R}^N)$ that satisfies*

$$u_j \to u_0 \text{ in } L^1_{loc}(\Omega;\mathbb{R}^N),$$

$$dist(Du_j, K) \to 0 \text{ in } L^1_{loc}(\Omega).$$

Then there exists an increasing sequence of open sets U_j, compactly contained in Ω, and functions $v_j \in W^{1,1}_{loc}(\Omega;\mathbb{R}^N)$ such that

$$v_j = u_0 \text{ on } \Omega \setminus U_j,$$

$$\mathcal{L}^n(\{u_j \neq v_j\} \cap U_j) \to 0,$$

$$\|dist(Dv_j, K)\|_{\infty,\Omega} \to 0.$$

By Theorem 4.8.3 there exist $v_k \in W^{1,\infty}(Q, \mathbb{R}^N)$ which satisfy

$$\mathcal{L}(\{u_k \neq v_k\}) \to 0, \qquad v_k = Fx \text{ on } \partial Q, \tag{4.78}$$

$$\|\text{dist}(Dv_k, K)\|_\infty \to 0. \tag{4.79}$$

Taking into account (4.71), the uniform continuity of h_0 and the inequality $h_0 \leq h_k$, we see that

$$\limsup_{k\to\infty} \int_Q h_0(Dv_k)dx = \limsup_{k\to\infty} \int_Q h_0(Du_k)dx \leq G_\infty. \tag{4.80}$$

In view of (4.75) and (4.79), there exist $\lambda k \searrow 1$ such that $\lambda_k^{-1} Dv_k \in K$, $\lambda_k^{-1} F \in K$. Using the uniform continuity of h_0 as well as quasiconvexity and continuity of f, we obtain

$$f(F) = \lim_{k\to\infty} f(\lambda_k^{-1}F) \leq \limsup_{k\to\infty} \int_Q f(\lambda_k^{-1}Dv_k)dx \tag{4.81}$$

$$= \limsup_{k\to\infty} \int_Q h_0(\lambda_k^{-1}Dv_k)dx \leq G_\infty. \tag{4.82}$$

The proof is finished. $\square$

Remark 4.8.2. The Iwaniec conjecture has close links with rank-one convexity and quasiconvexity in the calculus of variations, specifically, to those of Morrey and Šverâk's counterexample on the one hand, and to the geometric function theory, especially planar theory of quasiconformal mappings and the precise L^p-norm $(1 < p \leq 2)$ of the Beurling-Ahlfors operator. In this direction there are a number of related conjectures, namely those of Šverâk and Banuelos-Wang on the Burkholder functional (see Baernstein & Montgomery [6]). Note that if the Banuelos-Wang conjecture is true, then the Iwaniec conjecture will be true and if the Banuelos-Wang conjecture is not true, then Morrey's conjecture would be settled for the case $N = n = 2$. The truth of the Iwaniec conjecture would also have consequences for quasiconformal mappings in $\mathbb{R}^n$. If the Iwaniec conjecture does hold, then it would be a stronger variation of Astala's area distortion theorem on quasiconvex mappings.

Chapter 5

SO(n) and a Class of Geometric Maps into Spheres

5.1 Introduction

In this chapter we turn to the study of p-harmonic maps from a generalised annulus to an sphere and examine a geometric class of maps in connection with the Euler-Lagrange equation associated with the p-Dirichlet energy. In this regard our investigations continues the work of Taheri [52], [54], [53] as well as Shahrokhi-Taheri [43], [45], [42], [44] in a slightly different context. The compact Lie group $\mathbf{SO}(n)$ and the structure of its closed geodesics will play an interesting role in this study.

5.2 Spherical Twists and $W^{1,p}(\mathbb{X}, \mathbb{S}^{n-1})$

Let $\mathbb{X} := \mathbb{X}[a, b] = \{\mathrm{x} \in \mathbb{R}^n : a < |x| < b\}$ with $0 < a < b < \infty$ be a generalised annulus in $\mathbb{R}^n$ (with $n \geq 2$) and for $1 < p < \infty$ fixed consider the p-energy

$$\mathbb{E}_p[u; \mathbb{X}] := p^{-1} \int_{\mathbb{X}} |\nabla u|^p \, dx \tag{5.1}$$

over the space of admissible maps

$$\mathcal{A}_p(\mathbb{X}) = \left\{ u \in W^{1,p}(\mathbb{X}, \mathbb{S}^{n-1}) : u|_{\partial \mathbb{X}} = x|x|^{-1} \right\}. \tag{5.2}$$

Note that here $\partial \mathbb{X} = \mathbb{S}_a^{n-1} \cup \mathbb{S}_b^{n-1}$ is the union of two disjoint spheres centered at the origin having radii a and b respectively. (It is not difficult to see that the space of admissible maps $\mathcal{A}_p(\mathbb{X})$ is non-empty.) Now a standard and straightforward calculation shows that the Euler-Lagrange equation associated with this p-energy over the space of admissible maps $\mathcal{A}_p(\mathbb{X})$ takes the form

$$\Delta_p u + |\nabla u|^p u = 0, \tag{5.3}$$

where for the sake of brevity and convenience we have set $\Delta_p u = \mathrm{div}(|\nabla u|^{p-2}\nabla u)$.

Definition 5.2.1. (Spherical twist)

Let $\mathbb{X} = \mathbb{X}[a,b] = \{x \in \mathbb{R}^n : a < |x| < b\}$ with $0 < a < b < \infty$ and $n \geq 2$. A map $y \in \mathrm{C}(\overline{\mathbb{X}}, \mathbb{S}^{n-1})$ is called a *spherical twist* if and only if it can be expressed in the form

$$y(x) = \mathrm{Q}(r)\frac{x}{|x|} = \mathrm{Q}(r)\theta, x \in \mathbb{X}, \tag{5.4}$$

where here and in sequel $r = |x|$, $\theta = x|x|^{-1}$ and $\mathrm{Q} \in \mathrm{C}([a,b], \mathbf{SO}(n))$.

Proposition 5.2.1. *A spherical twist y lies in $\mathcal{A}_p = \mathcal{A}_p(\mathbb{X})$ with $1 < p < \infty$ provided that the following two conditions hold.*

[**1**] $\mathrm{Q} \in W^{1,p}([a,b], \mathrm{SO}(n))$,

[**2**] $\mathrm{Q}(a) = \mathrm{Q}(b) = \mathrm{I}_n$.

Note that in view of [**2**] above the "*curve*" $r \mapsto \mathbf{Q}(r)$ with $a < r < b$ forms a closed *loop* in the pointed space $(\mathbf{SO}(n), \mathrm{I}_n)$.

Proof. Let $y = y(x)$ be a spherical twist as defined above. Then referring to definitions it is plain that

$$y \in \mathcal{A}_p(\mathbb{X}) \Longleftrightarrow \begin{cases} y = x|x|^{-1} \text{ on } \partial\mathbb{X}, \\ |y| = \sqrt{\langle y, y\rangle} = 1 \text{ in } \mathbb{X}, \\ ||y||_{W^{1,p}(\mathbb{X},\mathbb{S}^{n-1})} < \infty. \end{cases} \tag{5.5}$$

The first two conditions are evidently true as a result of $\mathbf{Q}$ being an orthogonal matrix valued map. Regarding the third condition a straight-forward calculation gives

$$\nabla y = \frac{1}{r}\left[\mathrm{Q} + (r\dot{\mathrm{Q}} - \mathrm{Q})\theta \otimes \theta\right]. \tag{2.3}$$

where $r = |x|$, $\theta = x|x|^{-1}$ and $\dot{\mathrm{Q}} := \frac{d}{dr}\mathrm{Q}$. Thus we note that

$$\begin{aligned} |\nabla y|^2 &= tr\left\{|\nabla y||\nabla y|^t\right\} \\ &= \frac{1}{r^2} tr\left\{\mathrm{I}_n - \mathrm{Q}\theta \otimes \mathrm{Q}\theta + r^2\dot{\mathrm{Q}}\theta \otimes \dot{\mathrm{Q}}\theta\right\} \\ &= \frac{1}{r^2}\left\{n - \langle \mathrm{Q}\theta, \mathrm{Q}\theta\rangle + r^2\langle\dot{\mathrm{Q}}\theta, \dot{\mathrm{Q}}\theta\rangle\right\}. \end{aligned}$$

Again recalling that Q is an orthogonal matrix valued map we have $\langle Q\theta, Q\theta\rangle = 1$ and $\langle\dot{Q}\theta, Q\theta\rangle = 0$ for all $\theta \in \mathbb{S}^{n-1}$. Hence for $1 < p < \infty$ fixed we have that

$$|\nabla y|^p = \left[\frac{1}{r^2}(n-1) + |\dot{\mathrm{Q}}\theta|^2\right]^{\frac{p}{2}}. \tag{2.4}$$

Now in view of the pointwise condition $|y|^2 = 1$ an application of Proposition 5.1 in [] gives

$$
\begin{aligned}
\|y\|^p_{W^{1,p}(\mathbb{X})} &= \int_{\mathbb{X}} (|y|^p + |\nabla y|^p)\, dx \\
&= \int_a^b \int_{\mathbb{S}^{n-1}} \left\{ 1 + \left[\frac{1}{r^2}(n-1) + |\dot{\mathrm{Q}}\theta|^2 \right]^{\frac{p}{2}} \right\} r^{n-1} d\mathcal{H}^{n-1}(\theta)\, dr,
\end{aligned}
$$

and so referring to [**1**] in the statement of the proposition the conclusion follows. $\square$

Proposition 5.2.2. *Suppose that y is a spherical twist on $\mathbb{S}^{n-1}$ with $\mathrm{Q} \in \mathrm{C}^2(]a, b[, \mathrm{SO}(n))$. Then we have that*

$$
\begin{aligned}
\Delta_p y :=& div\, (|\nabla y|^{p-2} \nabla y) \\
=& \mathrm{Q} \left[\frac{1}{r} \nabla \mathrm{s} \otimes \theta - \frac{1}{r^n} \frac{d}{dr}(r^{n-1}\mathrm{s}) \mathrm{I}_n + \frac{1}{r^{n-1}} \frac{d}{dr}(r^{n-1}\mathrm{sA}) + \mathrm{sA}^2 \right] \theta,
\end{aligned}
\tag{5.6}
$$

where $\mathrm{A} = \mathrm{Q}^{\mathrm{t}}\dot{\mathrm{Q}}$ *and*

$$
\mathrm{s} = s(r, \theta) := \left[\frac{n-1}{r^2} + |\dot{\mathrm{Q}}\theta|^2 \right]^{\frac{p-2}{2}}. \tag{5.7}
$$

Proof. We begin by first computing Δy and then $\Delta_p y$.

Using the notation $y = (y_1, y_2, ..., y_n)$ we can write

$$
\begin{aligned}
\Delta y_i =& \sum_{j=1}^n \partial \partial x_j \left\{ \frac{1}{r} \mathrm{Q}_{ij} - \frac{1}{r} \sum_{k=1}^n \mathrm{Q}_{ik} \theta_k \theta_j + \sum_{k=1}^n \dot{\mathrm{Q}}_{ik} \theta_k \theta_j \right\} \\
=& \sum_{j=1}^n \left\{ \frac{-1}{r^2} \mathrm{Q}_{ij} \theta_j + \frac{1}{r} \dot{\mathrm{Q}}_{ij} \theta_j + \frac{1}{r^2} \sum_{k=1}^n \mathrm{Q}_{ik} \theta_k \theta_j^2 - \right. \\
& \frac{1}{r} \left[\sum_{k=1}^n \dot{\mathrm{Q}}_{ik} \theta_k \theta_j^2 + \frac{1}{r} \sum_{k=1}^n \mathrm{Q}_{ik} (\delta_{kj} - \theta_k \theta_j) \theta_j + \right. \\
& \left. \frac{1}{r} \sum_{k=1}^n \mathrm{Q}_{ik} \theta_k (1 - \theta_j^2) \right] + \sum_{k=1}^n \ddot{\mathrm{Q}}_{ik} \theta_k \theta_j^2 + \\
& \left. \frac{1}{r} \sum_{k=1}^n \dot{\mathrm{Q}}_{ik} (\delta_{kj} - \theta_k \theta_j) \theta_j + \frac{1}{r} \sum_{k=1}^n \dot{\mathrm{Q}}_{ik} \theta_k (1 - \theta_j^2) \right\}.
\end{aligned}
$$

Hence after some basic manipulations and simplifications we have that

$$
\Delta y_i = \sum_{k=1}^n \left[\frac{(1-n)}{r^2} \mathrm{Q}_{ik} + \frac{(n-1)}{r} \dot{\mathrm{Q}}_{ik} + \ddot{\mathrm{Q}}_{ik} \right] \theta_k.
$$

As this is true for $1 \le i \le n$ using the more convenient vector notation we can write

$$
\Delta y = \mathrm{Q} \left[\frac{(1-n)}{r^2} \mathrm{I}_n + \frac{(n-1)}{r} \mathrm{Q}^t \dot{\mathrm{Q}} + \mathrm{Q}^t \ddot{\mathrm{Q}} \right] \theta.
$$

Now using the substitutions $\dot{\mathrm{Q}} = \mathrm{QA}$ and $\ddot{\mathrm{Q}} = \mathrm{Q}[\dot{\mathrm{A}} + \mathrm{A}^2]\theta$, this reads as

$$
\begin{aligned}
\Delta y &= \mathrm{Q} \left[\frac{(1-n)}{r^2} \mathrm{I}_n + \frac{(n-1)}{r} \mathrm{A} + \dot{\mathrm{A}} + \mathrm{A}^2 \right] \theta, \\
&= \mathrm{Q} \left[-\frac{1}{r^n} \frac{d}{dr}(r^{n-1}) \mathrm{I}_n + \frac{1}{r^{n-1}} \frac{d}{dr}(r^{n-1}\mathrm{A}) + \mathrm{A}^2 \right] \theta.
\end{aligned}
$$

Note that in this case we have $\mathrm{s}(r,\theta)=1$. Now in the general case $1<p<\infty$ we have that

$$\Delta_p y = div\,(|\nabla y|^{p-2}\nabla y) = div\,(\mathrm{s}\nabla y) = \nabla y \nabla s + \mathrm{s}\Delta y.$$

Hence we can write

$$\begin{aligned}
\Delta_p y =& \nabla y \nabla s + \mathrm{s}\nabla y \\
=& \mathrm{Q}\left[\frac{1}{r}\mathrm{I}_n - \frac{1}{r}\theta\otimes\theta + \mathrm{A}\theta\otimes\theta\right]\nabla\mathrm{s}+ \\
& \mathrm{sQ}\left[-\frac{1}{r^n}\frac{d}{dr}(r^{n-1})\mathrm{I}_n + 1r^{n-1}\frac{d}{dr}(r^{n-1}\mathrm{A}) + \mathrm{A}^2\right]\theta \\
=& \mathrm{Q}\left[\frac{1}{r}\nabla s\otimes\theta - \frac{1}{r}\langle\nabla\mathrm{s},\theta\rangle + \langle\nabla\mathrm{s},\theta\rangle\mathrm{A}\right]\theta+ \\
& \mathrm{sQ}\left[-\frac{1}{r^n}\frac{d}{dr}(r^{n-1})\mathrm{I}_n + \frac{1}{r^{n-1}}\frac{d}{dr}(r^{n-1}\mathrm{A}) + \mathrm{A}^2\right]\theta.
\end{aligned}$$

Now an easy and short calculation gives

$$\nabla\mathrm{s} = \frac{1}{r}\left[r\mathrm{s}_r\mathrm{I}_n - \beta(\mathrm{A}^2 + |\mathrm{A}\theta|^2\mathrm{I}_n)\right]\theta,$$

where we have set $\mathrm{s}_r = \frac{\partial\mathrm{s}}{\partial r}$ and

$$\beta = \beta(r,\theta,p) := (p-2)\left[\frac{n-1}{r^2} + |\dot{\mathrm{Q}}\theta|^2\right]^{\frac{p-4}{2}}.$$

On the other hand in view of A being a skew-symmetric matrix it can be easily seen that $\langle\nabla\mathrm{s},\theta\rangle = \mathrm{s}_r$. Therefore substituting back gives

$$\begin{aligned}
\Delta_p y =& \mathrm{Q}\left[\frac{1}{r}\nabla\mathrm{s}\otimes\theta - \frac{1}{r}\mathrm{s}_r\mathrm{I}_n + \mathrm{s}_r\mathrm{A} - \mathrm{s}\frac{1}{r^n}\frac{d}{dr}(r^{n-1})\mathrm{I}_n + \right. \\
& \left.\mathrm{s}\frac{1}{r^{n-1}}\frac{d}{dr}(r^{n-1}\mathrm{A}) + \mathrm{sA}^2\right]\theta \\
=& \mathrm{Q}\left[\frac{1}{r}\nabla\mathrm{s}\otimes\theta - \frac{1}{r^n}\frac{d}{dr}(r^{n-1}\mathrm{s})\mathrm{I}_n + \frac{1}{r^{n-1}}\frac{d}{dr}(r^{n-1}\mathrm{s}(A) + \mathrm{sA}^2\right]\theta,
\end{aligned}$$

which is the required identity. □

5.3 The p-Energy Restricted to the Space of Spherical Twists

As before let us fix $1<p<\infty$ and consider the p-energy $\mathbb{E}_p$ (as defined earlier in the chapter) and let y be a spherical twist in $\mathcal{A}_p(\mathbb{X})$. Then we can write

$$\begin{aligned}
\mathbb{E}_p[y;\mathbb{X}] &= \frac{1}{p}\int_{\mathbb{X}} |\nabla y|^p\,dx \\
&= \frac{1}{p}\int_a^b \int_{\mathbb{S}^{n-1}} \left[\frac{n-1}{r^2} + |\dot{\mathrm{Q}}\theta|^2\right]^{\frac{p}{2}} r^{n-1}d\mathcal{H}^{n-1}(\theta)\,dr \\
&=: \int_a^b \mathrm{E}(r,\dot{\mathrm{Q}})r^{n-1}\,dr =: \mathbb{E}_p[\dot{\mathrm{Q}}],
\end{aligned} \tag{5.8}$$

where in the last line we have set

$$\mathrm{E}(r,\xi) = \int_{\mathbb{S}^{n-1}} \left[\frac{n-1}{2} + |\xi\theta|^2\right]^{\frac{p}{2}} d\mathcal{H}^{n-1}(\theta). \tag{5.9}$$

Now we consider the energy $\mathbb{E}_p$ restricted to the space of admissible *loops* which is defined as

$$\mathcal{E}_p := \left\{ \begin{array}{l} \mathrm{Q} = \mathrm{Q}(r) \in W^{1,p}([a,b], \mathrm{SO}(n)), \\ \mathrm{Q}(a) = \mathrm{Q}(b) = \mathrm{I}_n. \end{array} \right\} \tag{5.10}$$

Our aim here is to derive the Euler-Lagrange equation associated with this restricted energy and analyse its solutions and their qualitative properties.

Proposition 5.3.1. *Let* $\mathrm{Q} \in \mathcal{E}_p$ *with* $\mathrm{Q} \in \mathrm{C}^2(]a,b[, \mathrm{SO}(n))$*. Then the Euler-lagrange equation associated with* $\mathbb{E}_p$ *over* $\mathcal{E}_p$ *at* Q *takes the form*

$$\mathbb{EL}[\mathrm{Q}] = 0, \tag{5.11}$$

that is

$$\frac{d}{dr}\left\{ r^{n-1}\left[\mathrm{E}_\xi(r,\dot{\mathrm{Q}})\mathrm{Q}^t - \mathrm{Q}\mathrm{E}_\xi^t(r,\dot{\mathrm{Q}})\right]\right\} = 0, \tag{5.12}$$

on $]a,b[$.

Proof. First fix Q as described and for $\varepsilon \in \mathbb{R}$ put $\mathrm{Q}_\varepsilon = \mathrm{Q} + \varepsilon \mathrm{F}\mathrm{Q}$ where $\mathrm{F} \in \mathrm{C}_0^\infty(]a,b[, \mathbb{M}_{n\times n})$ is a skew-symmetric matrix. Then in view of F being a skew-symmetric matrix it can be easily seen that

$$\mathrm{Q}_\varepsilon \mathrm{Q}_\varepsilon^t = \mathrm{I}_n + O(\varepsilon^2).$$

Thus we can write

$$\begin{aligned} 0 &= \frac{d}{d\varepsilon}\mathbb{E}_p[\mathrm{Q}_\varepsilon]\Big|_{\varepsilon=0} \\ &= \frac{d}{d\varepsilon}\int_a^b \mathrm{E}(r,\dot{\mathrm{Q}}_\varepsilon) r^{n-1} dr\Big|_{\varepsilon=0} \\ &= \int_a^b \langle \mathrm{E}_\xi(r,\dot{\mathrm{Q}}_\varepsilon), \frac{d}{d\varepsilon}\dot{\mathrm{Q}}_\varepsilon\rangle r^{n-1} dr\Big|_{\varepsilon=0} \\ &= \int_a^b \langle \mathrm{E}_\varepsilon(r,\dot{\mathrm{Q}}), \dot{\mathrm{F}}\mathrm{Q} + \mathrm{F}\dot{\mathrm{Q}}\rangle r^{n-1} dr \\ &=: \mathbf{I} + \mathbf{II}. \end{aligned}$$

The next aim is to simplify each of the two terms in last above equation. For the first term we have

$$\begin{aligned} \mathbf{I} &= \int_a^b \langle \mathrm{E}_\xi(r,\dot{\mathrm{Q}})\mathrm{Q}^t, \dot{\mathrm{F}}\rangle r^{n-1} dr \\ &= \int_a^b \left\langle -\frac{d}{dr}\left[r^{n-1}\mathrm{E}_\xi(r,\dot{\mathrm{Q}})\mathrm{Q}^t\right], \mathrm{F}\right\rangle dr. \end{aligned}$$

Note that in deducing the second identity we used integration by parts together with the boundary conditions $\mathrm{F}(a) = \mathrm{F}(b) = 0$. On the other hand for the second term we have

$$\begin{aligned}
\mathbf{II} &= \int_a^b \left\langle \mathrm{E}_\xi(r, \dot{\mathrm{Q}}), \mathrm{F}\dot{\mathrm{Q}} \right\rangle r^{n-1}\, dr \\
&= \int_a^b \int_{\mathbb{S}^{n-1}} \left\langle p\mathrm{s}\dot{\mathrm{Q}}\theta \otimes \theta, \mathrm{F}\dot{\mathrm{Q}} \right\rangle r^{n-1} d\mathcal{H}^{n-1}(\theta)\, dr \\
&= \int_a^b \int_{\mathbb{S}^{n-1}} p\mathrm{s}\langle \dot{Q}\theta, \mathrm{F}\dot{\mathrm{Q}}\theta\rangle r^{n-1} d\mathcal{H}^{n-1}(\theta)\, dr = 0,
\end{aligned}$$

where the last identity is a result of the F being skew-symmetric. Therefore putting the two terms together we have that

$$\begin{aligned}
0 &= \frac{d}{d\varepsilon}\mathbb{E}_p[\mathrm{Q}_\varepsilon]\Big|_{\varepsilon=0} \\
&= \int_a^b \left\langle -\frac{d}{dr}\left[r^{n-1}\mathrm{E}_\xi(r, \dot{\mathrm{Q}})\mathrm{Q}^t \right], \mathrm{F} \right\rangle\, dr.
\end{aligned}$$

As this true for every skew-symmetric matrix $\mathrm{F} \in \mathrm{C}_0^\infty(]a, b[, \mathbb{M}_{n\times n})$ it follows that the skew-symmetric part of the tensor field in the brackets is zero. Which is the equation as required. □

Proposition 5.3.2. *The Euler-Lagrange equations associated with $\mathbb{E}_p$ over $\mathcal{E}_p$ can be alternatively expressed as*

$$\int_a^b \int_{\mathbb{S}^{n-1}} \langle \left[\frac{d}{dr}(r^{n-1}\mathrm{sA}) \right] \theta, \mathrm{F}\theta\rangle\, d\mathcal{H}^{n-1}(\theta) dr = 0, \tag{5.13}$$

for all skew-symmetric matrix $\mathrm{F} \in \mathrm{C}_0^\infty(]a, b[, \mathbb{M}_{n\times n})$ and $r \in]a, b[$.

Proof. Referring to the proof of the last proposition and using the same notations for A and s we can write

$$\begin{aligned}
0 &= \frac{d}{d\varepsilon}\mathbb{E}_p[\mathrm{Q}_\varepsilon]\Big|_{\varepsilon=0} \\
&= \int_a^b \left\langle \mathrm{E}_\xi(r, \dot{\mathrm{Q}}), \dot{\mathrm{F}}\mathrm{Q} \right\rangle r^{n-1}\, dr \\
&= \int_a^b \int_{\mathbb{S}^{n-1}} p\langle r^{n-1}\mathrm{sA}\theta, \dot{\mathrm{F}}\theta\rangle d\mathcal{H}^{n-1}(\theta)\, dr \\
&= \int_a^b \int_{\mathbb{S}^{n-1}} -p\langle \left\{ \frac{d}{dr}(r^{n-1}f^2\mathrm{sA}) \right\} \theta, \mathrm{F}\theta\rangle d\mathcal{H}^{n-1}(\theta)\, dr,
\end{aligned}$$

which is the required equation. □

Any twist loop Q forming a solution to the Euler-Lagrange equation associated with the energy $\mathbb{E}_p$ over $\mathcal{E}_p$ will be referred to as a p-stationary loop. Now, in view of the previous proposition it is evident that a sufficient condition on an admissible loop $\mathrm{Q} \in \mathcal{E}_p$ to be a p-stationary loop is that it satisfies

$$\frac{d}{dr}(r^{n-1}\mathrm{sA}) = 0. \tag{5.14}$$

5.4 Minimizing p-Stationary Loops in Homotopy Classes

Consider the energy functional

$$\mathbb{E}_p[\mathrm{Q}] := \int_a^b \mathrm{E}(r, \dot{\mathrm{Q}}) r^{n-1}\, dr, \tag{5.15}$$

where the integrand is given by

$$\mathbb{E}(r, \xi) = \int_{\mathbb{S}^{n-1}} \left[(n-1)(\frac{1}{r})^2 + |\xi\theta|^2 \right]^{\frac{p}{2}} d\mathcal{H}^{n-1}(\theta).$$

Furthermore recall that we are considering this p-energy over the space of admissible loops

$$\mathcal{E}_p := \left\{ \begin{array}{l} \mathrm{Q} = \mathrm{Q}(r) \in W^{1,p}([a,b], \mathrm{SO}(n)), \\ \\ \mathrm{Q}(a) = \mathrm{Q}(b) = \mathrm{I}_n. \end{array} \right\} \tag{5.16}$$

Proposition 5.4.1. *Let $1 \leq p < \infty$ and consider the p-energy $\mathbb{E}_p$ as defined above. Then there exists $d = d(p, a, b) > 0$ such that*

$$\mathbb{E}_p[\mathrm{Q}] \geq d \|\mathrm{Q}\|^p_{W^{1,p}} \tag{5.17}$$

for all $\mathrm{Q} \in \mathcal{E}_p$. Thus in particular the p-energy is $W^{1,p}$-coercive.

Proof. First note that for any$\mathrm{Q} \in \mathcal{E}_p$ we can write

$$\begin{aligned} \mathbb{E}_p[\mathrm{Q}] &= \int_a^b \int_{\mathbb{S}^{n-1}} \left[(n-1)(\frac{1}{r})^2 + |\dot{\mathrm{Q}}\theta|^2 \right]^{\frac{p}{2}} r^{n-1} d\mathcal{H}^{n-1}(\theta)\, dr \\ &\geq \int_a^b \int_{\mathbb{S}^{n-1}} r^{n-1} |\dot{\mathrm{Q}}\theta|^p d\mathcal{H}^{n-1}(\theta) dr. \end{aligned}$$

It follows that for some suitable constant $c > 0$ we can write

$$\mathbb{E}_p[\mathrm{Q}] \geq \int_a^b \int_{\mathbb{S}^{n-1}} r^{n-1} |\dot{\mathrm{Q}}\theta|^p d\mathcal{H}^{n-1}(\theta)\, dr \geq c \int_a^b r^{n-1} |\dot{\mathrm{Q}}|^p\, dr$$

Thus an application of the Poincaré inequality completes the proof. ☐

Recall that $\pi_1[\mathbf{SO}(2)] \cong \mathbb{Z}$ and $\pi_1[\mathbf{SO}(n)] \cong \mathbb{Z}_2$ for $n \geq 3$. Let us denote the homotopy classes of closed curves in the pointed space $(\mathbf{SO}(n), \mathrm{I}_n)$ by $\mathfrak{c}_k[\mathcal{E}_p]$ with $k \in \mathbb{Z}$ when $n = 2$ and $\mathfrak{c}_\alpha[\mathcal{E}_p]$ with $\alpha \in \mathbb{Z}_2$ when $n \geq 3$. Then an application of the direct method of the calculus of variations gives the following result.

Theorem 5.4.1. *Let $1 < p < \infty$. Then the following hold.*

[**1**] $(n = 2)$ *for each $k \in \mathbb{Z}$ there exists $\mathrm{Q}_k \in \mathfrak{c}_k[\mathcal{E}_p]$ such that*

$$\mathbb{E}_p[\mathrm{Q}_k] = \inf_{\mathfrak{c}_k[\mathcal{E}_p]} \mathbb{E}_p, \tag{5.18}$$

[**2**] $(n \geq 3)$ *for each $\alpha \in \mathbb{Z}_2$ there exists $\mathrm{Q}_\alpha \in \mathfrak{c}_\alpha[\mathcal{E}_p]$ such that*

$$\mathbb{E}_p[\mathrm{Q}_\alpha] = \inf_{\mathfrak{c}_\alpha[\mathcal{E}_p]} \mathbb{E}_p. \tag{5.19}$$

5.5 Spherical Twists on as p-Harmonic Maps

The aim of this section is to give a complete characterization of all those p-stationary loops $\mathrm{Q} \in \mathcal{E}_p$ whose resulting spherical twist

$$y(r\theta) = \mathrm{Q}(r)\theta \tag{5.20}$$

furnishes a solution to the Euler-Lagrange equation associated with the p-energy $\mathbb{E}_p$ over the space $\mathcal{A}_p(\mathbb{X})$. We begin with the proposition below.

Proposition 5.5.1. *Let* $\mathbb{X} := \mathbb{X}[a,b] = \{x \in \mathbb{R}^n : a < |x| < b\}$ *and consider the vector field* $\mathrm{v} \in \mathrm{C}^1(\Omega, \mathbb{R}^n)$ *defined in spherical coordinates through*

$$\mathrm{v} = \mathrm{tA}\theta, \tag{5.21}$$

where $r \in]a,b[$, $\theta \in \mathbb{S}^{n-1}$, $\mathrm{A} = \mathrm{A}(r) \in \mathrm{C}^1(]a,b[, \mathbb{M}_{n\times n})$ *is skew-symmetric and*

$$\mathrm{t} := \mathrm{t}(r, |\mathrm{A}\theta|^2) > 0. \tag{5.22}$$

Then we have the following:
[**1**] $\mathrm{A} = 0$ *if and only if for any close path* $\gamma \subset \mathbb{S}^{n-1}$

$$\int_{r\gamma} \langle \mathrm{v}(r\gamma), r\gamma' \rangle = 0, \tag{5.23}$$

[**2**] $\frac{d}{dr}(\mathrm{tA}) = 0$ *if and only if for any close path* $\gamma \subset \mathbb{S}^{n-1}$

$$\int_{r\gamma} \langle \frac{d}{dr}[\mathrm{v}(r\gamma)], r\gamma' \rangle = 0. \tag{5.24}$$

Proof. Indeed we first note that in view of A being skew-symmetric it can be orthogonally diagonalised, i.e., we can write $\mathrm{A} = \mathrm{PDP}^t$ where $P = P(r) \in \mathrm{SO}(n)$ and $\mathrm{D} = \mathrm{D}(r) \in \mathbb{M}_{n\times n}$ is in special block digonal form, i.e.,
[**1**] $(n = 2k)$

$$\mathrm{D} = diag(d_1\mathrm{J}, d_2\mathrm{J}, ..., d_k\mathrm{J}),$$

[**2**] $(n = 2k+1)$

$$\mathrm{D} = diag(d_1\mathrm{J}, d_2\mathrm{J}, ..., d_k\mathrm{J}, 0),$$

with $\{\pm d_1 i, \pm d_2 i, \dots, \pm d_k i\}$ or $\{\pm d_1 i, \pm d_2 i, \dots, \pm d_k i, 0\}$ denoting the eigenvalues of the skew-symmetric matrix A [as well as D] respectively. Now consider a parameterised family of closed paths $\rho \in \mathrm{C}^\infty([0,2\pi], \mathbb{S}^{n-1})$ give by

$$\rho : [0, 2\pi] \ni t \to \rho(t) \in \mathbb{S}^{n-1} \subset \mathbb{R}^n \tag{5.25}$$

such that

$$\begin{cases} \rho_1 = \sin t \sin\varphi_2 sin\varphi_3 \dots \sin\varphi_{n-1}, \\ \rho_2 = \cos t \sin\varphi_2 \sin\varphi_3 \dots \sin\varphi_{n-1}, \\ \rho_3 = \cos\varphi_2 \sin\varphi_3 \dots \sin\varphi_{n-1}, \\ \vdots \\ \rho_{n-1} = \cos\varphi_{n-2} \sin\varphi_{n-1}, \\ \rho_n = \cos\varphi_{n-1}, \end{cases}$$

where $\varphi_i \in [0,\pi]$ for all $2 \leq \ j \leq n-1$. For fix $1 \leq p <, q \leq n$ we introduce the matrix Γ^{pq} as that obtained by simultaneously interchanging the first and p-th and the second and q-th rows of I_n, i.e.,

$$\Gamma^{pq} e_j = \begin{cases} e_p & \text{if } j = 1, \\ e_1 & \text{if } j = p, \\ e_q & \text{if } j = 2, \\ e_2 & \text{if } j = q, \\ e_j & \text{otherwise}, \end{cases}$$

where $\{e_j\}_{i=1}^n$ denotes the standard basis of $\mathbb{R}^n$. Now in view of $\Gamma^{p,q} \in \mathrm{O}(n)$ setting $\omega = \Gamma^{pq}\rho$ it is clear that ω is closed path in $\mathrm{C}^\infty([0,2\pi], \mathbb{S}^{n-1})$.

[**1**] For prove this part via above notation it is sufficient show that $d_i = 0$ for all i, therefore with sufficient selection for p and q it means $p = 2j-1, q = 2j$ for some $1 \leq j \leq k = [n/2]$ and $\gamma = \mathrm{P}\omega$, then we have

$$\begin{aligned} 0 &= \int_0^{2\pi} \mathrm{t}(r, |\mathrm{PDP}^t\gamma|^2)\langle \mathrm{PDP}^t\gamma, \gamma' \rangle dt \\ &= \int_0^{2\pi} \mathrm{t}(r, |\mathrm{D}\omega|^2)\langle \mathrm{D}\omega, \omega' \rangle dt, \end{aligned}$$

but it easy show that $\langle \mathrm{D}\omega, \omega' \rangle = d_i^2(\rho_1^2 + \rho_2^2)$ another hand t does not depend on variable t because of $|D\omega|^2$ does not depend on t therefore in view of $\mathrm{t} > 0$ and $\rho_1^2 + \rho_2^2 \neq 0$ we can write

$$\begin{aligned} 0 &= \int_0^{2\pi} \mathrm{t}(r, |\mathrm{D}\omega|^2)\langle \mathrm{D}\omega, \omega' \rangle dt \\ &= \mathrm{t}(r, |\mathrm{D}\omega|^2) \int_0^{2\pi} d_i^2(\rho_1^2 + \rho_2^2) dt \\ &= d_i^2(\rho_1^2 + \rho_2^2) 2\pi = d_i^2. \end{aligned}$$

Thus $\mathrm{D} = 0$ and this shows that $\mathrm{A} = 0$.

[**2**] First we introduce the skew-symmetric matrix

$$\mathrm{F} = \mathrm{F}(r,\theta) := \mathrm{P}^t \frac{d}{dr}(\mathrm{tA})\mathrm{P}.$$

Then straight-forward differentiation shows that

$$\mathrm{F} = \mathrm{t}_r \mathrm{D} + \mathrm{t}\mathrm{P}^t \dot{\mathrm{A}}\mathrm{P},$$

it is clear that $\frac{d}{dr}(\mathrm{tA}) = 0$ is equivalent to showing that $\mathrm{F}(r,\theta) = 0$ for all $r \in]a,b[$ and $\theta \in \mathbb{S}^{n-1}$. On other hand we setting $\gamma = \mathrm{P}\omega$ with same ω is [**1**] we have

$$\begin{aligned}
0 &= \int_0^{2\pi} \langle \frac{d}{dr}(\mathrm{tA})\gamma, \gamma' \rangle dt \\
&= \int_0^{2\pi} \langle \frac{d}{dr}(\mathrm{tA})\mathrm{P}\omega, \mathrm{P}\omega' \rangle dt \\
&= \int_0^{2\pi} \langle \mathrm{P}^t \frac{d}{dr}(\mathrm{tA})\mathrm{P}\omega, \omega' \rangle dt = \int_0^{2\pi} \langle \mathrm{F}\omega, \omega' \rangle dt
\end{aligned}$$

we remind that in above $\mathrm{t} = \mathrm{t}(r, \mathrm{P}\omega)$ and $\mathrm{F} = \mathrm{F}(r, \mathrm{P}\omega)$. we now want to show $\mathrm{F} = 0$ but in view of F being skew-symmetric matrix it suffices to justify the latter in the from $\mathrm{F}_{pq}(r,\theta) = 0$ only when $1 \le p < q \le n$. We consider the following two distinct case.

[**2a**] ($p = 2j-1, q = 2j$ for some $1 \le j \le k = [n/2]$) In this case again t does not depend on variable t also it is true for $\mathrm{F}(r, \mathrm{P}\omega)$ therefore we can write

$$\begin{aligned}
0 &= \int_0^{2\pi} \langle \mathrm{F}(r, \mathrm{P}\omega)\omega, \omega' \rangle dt \\
&= \int_0^{2\pi} \langle \mathrm{F}(r, \mathrm{P}\Gamma^{pq}\rho(t))\Gamma^{pq}\rho(t), \Gamma^{pq}\rho'(t) \rangle dt \\
&= 2\pi(\rho_1^2 + \rho_2^2)\mathrm{F}(r, \mathrm{P}\omega)
\end{aligned}$$

which in turn for $\rho_1^2 + \rho_2^2 \neq 0$ gives

$$\mathrm{F}_{pq}(r, \mathrm{P}\omega) = 0.$$

[**2b**] (p, q not as in [**2a**]) In this case t is depend on variable t but $\mathrm{D}_{pq} = 0$ as can verified by inspecting its block diagonal representation. Now for tis case we can write

$$\begin{aligned}
0 &= \int_0^{2\pi} \langle \mathrm{F}\omega, \omega' \rangle dt \qquad (5.5) \\
&= \int_0^{2\pi} \left\{ \sum_{j=0}^{n} \mathrm{F}_{pj}\omega_j\omega_p' + \sum_{j=0}^{n} \mathrm{F}_{qj}\omega_j\omega_q' \right\} dt \\
&= \int_0^{2\pi} \left\{ (\mathrm{F}_{pq}\rho_2^2 - \mathrm{F}_{qp}\rho_1^2) + \rho_2 \sum_{\substack{j=0 \\ j\neq q}}^{n} \mathrm{F}_{pj}\omega_j - \rho_1 \sum_{\substack{j=0 \\ j\neq p}}^{n} \mathrm{F}_{qj}\omega_j \right\} dt \\
&= \mathbf{I} + \mathbf{II} + \mathbf{III}
\end{aligned}$$

In order to evaluate the above terms we first observe that here t takes the form

$$\mathrm{t} = \mathrm{t}(r, |\mathrm{AP}\omega(t)|^2) := \mathrm{t}(\sin^2 t, \cos^2 t),$$

because of

$$\begin{aligned}
|\mathrm{AP}\omega(t)|^2 &= |\mathrm{PDP}^t\mathrm{P}\omega(t)|^2 \\
&= -\langle \mathrm{D}^2\omega(t), \omega(t)\rangle \\
&= -\langle \mathrm{D}^2\Gamma^{pq}\rho(t), \Gamma^{pq}\rho(t)\rangle \\
&= d_1^2\rho_p^2 + d_2^2\rho_q^2 + \ldots + d_\xi^2\rho_1^2 + \ldots + d_\zeta^2\rho_2^2 + \ldots
\end{aligned}$$

Now returning to (5.5) we have

$$\begin{aligned}
\mathbf{II} &= \int_0^{2\pi} \rho_2 \sum_{\substack{j=0 \\ j\neq q}}^{n} \mathrm{F}_{pj}\omega_j \, dt \\
&= \int_0^{2\pi} \rho_2 \sum_{\substack{j=0 \\ j\neq q}}^{n} [\mathrm{P}^t \frac{d}{dr}(\mathrm{tA})\mathrm{P}]_{pj}\omega_j \, dt \\
&= \sum_{\substack{j=0 \\ j\neq q}}^{n} [\mathrm{P}^t \frac{d}{dr}(\Big\{\int_0^{2\pi} \rho_2 \mathrm{t} \, dt\Big\} \mathrm{A})\mathrm{P}]_{pj}\omega_j,
\end{aligned}$$

and in a similar way

$$\begin{aligned}
\mathbf{III} &= \int_0^{2\pi} \rho_1 \sum_{\substack{j=0 \\ j\neq p}}^{n} \mathrm{F}_{qj}\omega_j \, dt \\
&= \int_0^{2\pi} \rho_1 \sum_{\substack{j=0 \\ j\neq p}}^{n} [\mathrm{P}^t \frac{d}{dr}(\mathrm{tA})\mathrm{P}]_{qj}\omega_j \, dt \\
&= \sum_{\substack{j=0 \\ j\neq p}}^{n} [\mathrm{P}^t \frac{d}{dr}(\Big\{\int_0^{2\pi} \rho_1 \mathrm{t} \, dt\Big\} \mathrm{A})\mathrm{P}]_{qj}\omega_j,
\end{aligned}$$

however in view of the specific manner in which t depends on t it follows that both integrals vanish and so as a result $\mathbf{II} = \mathbf{III} = 0$. It can be easily shown that as a result of periodicity the following identities hold:

$$\begin{aligned}
&\int_0^{2\pi} t(\sin^2 t, \cos^2 t) \sin t \, dt = 0, \\
&\int_0^{2\pi} t(\sin^2 t, \cos^2 t) \cos t \, dt = 0.
\end{aligned}$$

Hence returning to (5.5) and in view of F being a skew-symmetric we can write

$$\begin{aligned}
\mathbf{I} &= \int_0^{2\pi} (\mathrm{F}_{pq}\rho_2^2 - \mathrm{F}_{qp}\rho_1^2)\, dt \\
&= \int_0^{2\pi} (\rho_1^2 + \rho_2^2)\mathrm{F}_{pq}\, dt \\
&= \int_0^{2\pi} (\rho_1^2 + \rho_2^2)\mathrm{t}[\mathrm{P}^t\dot{\mathrm{A}}\mathrm{P}]_{pq}]\, dt \\
&= (\rho_1^2 + \rho_2^2)\left\{\int_0^{2\pi} \mathrm{t}\ dt\right\}[P^t\dot{\mathrm{A}}\mathrm{P}]_{pq} = 0.
\end{aligned}$$

Thus as $\mathrm{t} > 0$ for $\rho_1^2 + \rho_2^2 \neq 0$ it shows that $[\mathrm{P}^t\dot{\mathrm{A}}\mathrm{P}]_{pq} = 0$. But for range of p, q we have that $\mathrm{D}_{pq} = 0$ and it immediately implying that $F_{pq} = 0$.

Hence summarising we have shown in both case [**2a**] and [**2b**] foe fix $r \in]a, b[$ we have $\mathrm{F}_{pq}(r, .) = 0$ outside a copy of $\mathbb{S}^{n-3}$. By continuity of $F_{pq}(r, .)$ on $\mathbb{S}^{n-1}$ this gives $\mathrm{F}(r, \theta) = 0$ for all $r \in]a, b[$ and $\theta \in \mathbb{S}^{n-1}$ and as a result we have that

$$\frac{d}{dr}(\mathrm{tA}) = 0.$$

The proof is therefore complete. $\square$

Theorem 5.5.1. *Let y be a spherical twist and suppose that the twist loop Q lies in $\mathcal{E}_p$ and that $\mathrm{Q} \in \mathrm{C}^2(]a, b[, \mathrm{SO}(n))$. Then we have*

$$\mathbb{EL}[y] = 0 \Longleftrightarrow \left\{\begin{array}{l} (i)\ \frac{d}{dr}(r^{n-1}\mathrm{sA}) = 0, \\ (ii)\ \dot{\mathrm{Q}}(r) \in \mathbb{R}\mathrm{SO}(n)\ for\ all\ r \in]a, b[, \end{array}\right\} \tag{5.26}$$

where $\mathrm{A} = \mathrm{Q}^t\dot{\mathrm{Q}}$ *and*

$$\mathrm{s} = \mathrm{s}(r, \theta) := \left[\frac{n-1}{r^2} + |\dot{\mathrm{Q}}\theta|^2\right]^{p-2} 2. \tag{5.27}$$

Proof. Let $y = \mathrm{Q}(r)\theta$ be a generalised twist. Then an application of Proposition 2.2 we can write

$$\begin{aligned}
\mathbb{EL}[y] = 0 \Longleftrightarrow & |\nabla y|^p y + \Delta_p y = 0 \\
\Longleftrightarrow & \left[\frac{1}{r^2}(n-1) + |\dot{\mathrm{Q}}\theta|^2\right]^{\frac{p}{2}} \mathrm{Q}\theta + \mathrm{Q}\{\frac{1}{r}\nabla\mathrm{s} \otimes \theta - \\
& \frac{1}{r^n}\frac{d}{dr}(r^{n-1}\mathrm{s})\mathrm{I}_n + \frac{1}{r^{n-1}}\frac{d}{dr}(r^{n-1}\mathrm{sA}) + \mathrm{sA}^2\}\theta \\
\Longleftrightarrow & \frac{1}{r}\nabla\mathrm{s} + \{\left[\frac{n-1}{r^2} + |\mathrm{A}\theta|^2\right]^{\frac{p}{2}} \mathrm{I}_n - \frac{1}{r^n}\frac{d}{dr}(r^{n-1}\mathrm{s})\mathrm{I}_n + \\
& \frac{1}{r^{n-1}}\frac{d}{dr}(r^{n-1}\mathrm{sA}) + \mathrm{sA}^2\}\theta = 0.
\end{aligned}$$

$(\Longleftarrow)$

By condition (ii) and this fact that $\mathrm{Q}^t\dot{Q}$ is skew-symmetric matrix, there exists some

$0 < \sigma \in \mathbb{R}$ such that $\mathrm{A}^2 = -\sigma \mathrm{I}_n$. Now returning to (5.6) and using this note and condition (i) together we have

$$\begin{aligned}\mathbb{EL}[\mathrm{y}] &= \frac{1}{r}\mathrm{s}_r\theta + \left\{ \left[\frac{n-1}{r^2} + |\mathrm{A}\theta|^2 \right]^{\frac{p}{2}} \mathrm{I}_n - \frac{1}{r}\mathrm{s}_r\mathrm{I}_n - \frac{n-1}{r^2}\mathrm{s}\mathrm{I}_n + \mathrm{s}\mathrm{A}^2 \right\}\theta \\ &= \mathrm{s}\left\{ \frac{n-1}{r^2}\mathrm{I}_n + \sigma^2\mathrm{I}_n - \frac{n-1}{r^2}\mathrm{I}_n - \sigma^2\mathrm{I}_n \right\}\theta = 0.\end{aligned}$$

$(\Longrightarrow)$

Assume that $\mathbb{EL}[\mathrm{y}] = 0$. For the sake of clarity and convenience we break this part into two steps.

Step1. [Justification of (i)]

We begin by extracting a gradient out of right side in 6.8 and hence rewrite it in the form

$$\begin{aligned}\mathbb{EL}[\mathrm{y}] =& r\nabla \mathrm{t} + \frac{1}{r}\nabla \mathrm{s} + \left\{ \left[\frac{n-1}{r^2} + |\mathrm{A}\theta|^2 \right]^{\frac{p}{2}} \mathrm{I}_n - \frac{1}{r^n}\frac{d}{dr}(r^{n-1}\mathrm{s})\mathrm{I}_n + \right. \\ & \left. \frac{1}{r^{n-1}}\frac{d}{dr}(r^{n-1}\mathrm{s}\mathrm{A}) + p^{-1}r\mathrm{t}_r\mathrm{I}_n - \mathrm{s}|\mathrm{A}\theta|^2\mathrm{I}_n \right\}\theta \\ =:& r\nabla \mathrm{t} + \frac{1}{r}\nabla \mathrm{s} + \left[\frac{1}{r^{n-1}}\frac{d}{dr}(r^{n-1}\mathrm{s}\mathrm{A}) + H(r,\theta)\mathrm{I}_n \right]\theta = 0,\end{aligned}$$

where $\mathrm{t} = -p^{-1}\left[(n-1)(\frac{1}{r})^2 + |\mathrm{A}\theta|^2\right]^{\frac{p}{2}}$. We consider the vector field

$$\mathrm{v} := r\nabla \mathrm{t} + \frac{1}{r}\nabla \mathrm{s} + \left[\frac{1}{r^{n-1}}\frac{d}{dr}(r^{n-1}\mathrm{s}\mathrm{A}) + H(r,\theta)\mathrm{I}_n \right]\theta.$$

Now for this vector field we assign the differential 1-form $\omega = v_1 dx_1 + \ldots + v_n dx_n$. Then in view of v being zero, for any closed path $\gamma \in \mathrm{C}^1([0,2\pi], \mathbb{S}^{n-1})$ it must be that

$$\begin{aligned}0 &= \int_{r\gamma} \omega = \int_0^{2\pi} \langle \mathrm{v}(r\gamma(t)), r\gamma'(t)\rangle \, dt \\ &= \int_0^{2\pi} \langle \frac{d}{dr}\left[r^{n-1}\mathrm{s}(r,\gamma(t))\mathrm{A}\right]\gamma(t), r\gamma'(t)\rangle \, dt + \\ & \quad r\int_0^{2\pi} H(r,\gamma(t))\langle \gamma(t), \gamma'(t)\rangle \, dt \\ &= \int_0^{2\pi} \langle \frac{d}{dr}\left[r^{n-1}\mathrm{s}(r,\gamma(t))\mathrm{A}\right]\gamma(t), r\gamma'(t)\rangle \, dt\end{aligned}$$

where in concluding the last line we have used the identity $\langle \gamma, \gamma' \rangle = 0$ that is true as a result of $\gamma \subset \mathbb{S}^{n-1}$. Thus we prove for any closed path $\gamma \in \mathrm{C}^1([0,2\pi], \mathbb{S}^{n-1})$

$$\int_0^{2\pi} \langle \frac{d}{dr}\left[r^{n-1}\mathrm{s}(r,\gamma(t))\mathrm{A}\right]\gamma(t), r\gamma'(t)\rangle dt = 0.$$

An application of Proposition 5.1 part [**2**] gives

$$\frac{d}{dr}(r^{n-1}\mathrm{s}\mathrm{A}) = 0.$$

Step2. [Justification of (ii)]

Again referring to (5.6) and using the result in above we can write for every $\theta \in \mathbb{S}^{n-1}$

$$
\begin{aligned}
0 &= \frac{1}{r}\nabla \mathrm{s} + \left\{ \left[\frac{n-1}{r^2} + |\mathrm{A}\theta|^2\right]^{\frac{p}{2}} \mathrm{I}_n - \frac{1}{r^n}\frac{d}{dr}(r^{n-1}\mathrm{s})\mathrm{I}_n + \mathrm{s}\mathrm{A}^2 \right\}\theta \\
&= \frac{1}{r}\nabla \mathrm{s} + \left\{ \mathrm{s}\left[\frac{n-1}{r^2} + |\mathrm{A}\theta|^2\right] \mathrm{I}_n - \frac{n-1}{r^2}\mathrm{s}\mathrm{I}_n - \frac{1}{r}\mathrm{s}_r\mathrm{I}_n + \mathrm{s}\mathrm{A}^2 \right\}\theta \\
&= \frac{1}{r}\left[\mathrm{s}_r\mathrm{I}_n - \frac{\beta}{r}(\mathrm{A}^2 + |\mathrm{A}\theta|^2\mathrm{I}_n)\right]\theta + \left\{ \mathrm{s}(\mathrm{A}^2 + |\mathrm{A}\theta|^2\mathrm{I}_n) - \frac{1}{r}\mathrm{s}_r\mathrm{I}_n \right\}\theta \\
&= (\mathrm{s} - \frac{\beta}{r^2})\left[\mathrm{A}^2 + |\mathrm{A}\theta|^2\mathrm{I}_n\right]\theta,
\end{aligned}
$$

it can be easy check that

$$
\begin{aligned}
\mathrm{s} - \frac{1}{r^2}\beta &= \mathrm{s}^{\frac{(p-4)}{p-2}}\left[(n-p+1)\frac{1}{r^2} + |\mathrm{A}\theta|^2\right] \\
&=: \mathrm{s}^{\frac{(p-4)}{p-2}}\left[g(r) + |\mathrm{A}\theta|^2\right],
\end{aligned}
$$

where $g = (n-p+1)r^{-2}$. In view of $\mathrm{s} > 0$ and latest identity we arrive to

$$
0 = \left[g + |\mathrm{A}\theta|^2\right]\left[\mathrm{A}^2 + |\mathrm{A}\theta|^2\mathrm{I}_n\right]\theta, \tag{5.7}
$$

this identity is true for all $\theta \in \mathbb{S}^{n-1}$ and $r \in]a,b[$. Fix $r \in]a,b[$ and since A is skew-symmetric it follows that there exist $\mathrm{P} \in \mathrm{SO}(n)$ and block diagonal matrix D such that

$$
\mathrm{A} = \mathrm{P}\mathrm{D}\mathrm{P}^t
$$

where

$$
\mathrm{D} = \begin{cases} diag(d_1J, d_2J, \ldots, d_kJ) & n = 2k \\ diag(d_1J, d_2J, \ldots, d_kJ, 0) & n = 2k+1 \end{cases}
$$

with $\{\pm d_1 i, \pm d_2 i, \ldots, \pm d_k i\}$ or $\{\pm d_1 i, \pm d_2 i, \ldots, \pm d_k i, 0\}$ denoting the eigenvalues of the skew-symmetric matrix A. We setting $\omega = \mathrm{P}^t\theta$ and substituting in the (5.7) we will get

$$
0 = (g + |\mathrm{D}\omega|^2)\left[\mathrm{D}^2 + |\mathrm{D}\omega|^2\mathrm{I}_n\right]\omega.
$$

Since, above identity is true for all $\omega \in \mathbb{S}^{n-1}$ if we for $1 \le i \le [n/2]$ we set

$$
\omega_k = \begin{cases} \sin t & k = 2i, \\ \cos t & k = 2j := 2(i+1), \\ 0 & otherwise, \end{cases}
$$

where $t \in \mathbb{R}$ and then put this ω in (5.7), an easy calculation shows that

$$
(d_i^2 - d_j^2)(g + d_i^2\cos^2 t + d_j^2\sin^2 t)\cos t \sin t = 0,
$$

thus must be

$$(d_i^2 - d_j^2)(g + d_i^2 \cos^2 t + d_j^2 \sin^2 t) = 0. \tag{5.8}$$

Similar for

$$\omega_k = \begin{cases} \cos t & k = 2i, \\ \sin t & k = 2j, \\ 0 & otherwise, \end{cases}$$

implying that

$$(d_i^2 - d_j^2)(g + d_i^2 \sin^2 t + d_j^2 \cos^2 t) = 0. \tag{5.9}$$

But 5.8 and 5.9 together gives

$$(d_i^2 - d_j^2) \cos 2t = 0$$

therefore $d_i^2 - d_j^2 = 0$ for all $1 \le i < j \le n$ thus it implying that exists $\sigma(r) \in \mathbb{R}$ such that

$$\mathrm{D}^2 = -\sigma^2 \mathrm{I}_n$$

another hand $\mathrm{A}^2 = \mathrm{PD}^2\mathrm{P}^{\mathrm{t}}$ so A^2 is in $\mathbb{R}\mathrm{SO}(n)$. □

Theorem 5.5.2. *Let $\mathbb{X} = \mathbb{X}[a,b] = \{x \in \mathbb{R}^n \ : \ a < |x| < b\}$ and suppose $y \in \mathcal{A}_p(\mathbb{X})$ with $1 < p < \infty$ is a spherical twist whose corresponding twist loop Q satisfies the following assumptions:*

$$\left\{ \begin{aligned} &(i)\ \mathrm{Q} \in \mathrm{C}^2(]a,b[, \mathrm{O}(n)), \\ &(ii)\ \mathrm{Q} \in \mathcal{E}_p \end{aligned} \right\} \tag{5.28}$$

Then the following are equivalent.

[**1**] *y satisfies to full Euler-Lagrange equation associated with $\mathbb{E}_p$ over $\mathcal{A}_p$,*

[**2**] *depending on n being odd or even we have that*

[**2a**] *($n = 2k$) there exists $g = g(r) \in C^2[a,b]$ with $g(a), g(b) \in 2\pi\mathbb{Z}$ and $\mathrm{P} \in \mathrm{O}(n)$ such that*

$$\mathrm{Q} = \mathrm{P} diag(\mathfrak{R}(g), \dots, \mathfrak{R}(g))\mathrm{P}^t,$$

where g is a solution on (a,b) to

$$\frac{d}{dr}\left[r^{n-1}[(n-1)r^{-2} + g'^2]^{\frac{p-2}{2}} g'\right] = 0. \tag{5.29}$$

[**2b**] *($n = 2k+1$) in this case we have $\mathrm{Q} = \mathrm{I}_n$ or equivalently $y(x) = x|x|^{-1}$.*

5.6 Alternate Construction of Multiple p-Stationary Loops

In this section we consider special case for p-stationary loops where dimension is even. Let $p \in [1, \infty[$ and for $m \in \mathbb{Z}$ set

$$\mathcal{G}_{m,p} = \mathcal{G}(m, p, \mathbb{X}) := \begin{Bmatrix} g = g(r) \in W^{1,p}[a, b], \\ g(a) = 0, \ g(b) = 2\pi m \end{Bmatrix}. \tag{5.30}$$

Now for any $g \in \mathcal{G}_{m,p}$ and $\mathrm{P} \in \mathrm{O}(n)$ we make Q as below

$$\mathrm{Q} = \mathrm{P} diag[\mathfrak{R}(g), \ldots, \mathfrak{R}(g)]\mathrm{P}^t.$$

Evidently for each m we get that Q lies $\mathcal{E}_p$. Now we considering an energy functional $\mathbb{G}_p$ over teh space $\mathcal{G}_{m,p}$ as described above

$$\begin{aligned} \mathbb{G}_p[g] := \mathbb{E}_p[\mathrm{Q}] = &\int_a^b \int_{\mathbb{S}^{n-1}} \left[(n-1)(\frac{1}{r})^2 + |\dot{\mathrm{Q}}\theta|^2\right]^{\frac{p}{2}} r^{n-1} \, d\mathcal{H}^{n-1}(\theta) dr \\ =& n\omega_n \int_a^b \left[(n-1)(\frac{1}{r^2}) + g'^2\right]^{\frac{p}{2}} r^{n-1} \, dr. \end{aligned} \tag{5.31}$$

Theorem 5.6.1. *Suppose that $1 < p < \infty$. Consider the energy functional $\mathbb{G}_p$ over the space $\mathcal{G}_{m,n}$. Then for each $m \in \mathbb{Z}$ there exists $g \in \mathcal{G}_{m,p}$ such that*

$$\mathbb{G}_p[g] = \underset{\mathcal{G}_{m,p}}{inf} \ \mathbb{G}_p[.]. \tag{5.32}$$

In additional $g \in \mathrm{C}^\infty[a, b]$ and satisfies in corresponding Euler-Lagrange equation

$$\frac{d}{dr}\left[r^{n-1}[(n-1)r^{-2} + g'^2]^{\frac{p-2}{2}} g'\right] = 0. \tag{6.11}$$

on $[a, b]$.

Proof. This is an immediate consequence of applying the direct method of the calculus of variation and a standard regularity theory. □

Theorem 5.6.2. *Let $\mathbb{X} = \mathbb{X}[a, b]$ and consider the energy $\mathbb{E}_p$ over the space $\mathcal{A}_p(\mathbb{X})$ with $1 < p < \infty$. Then if the spherical twist y is a solution to the corresponding Euler-Lagrange equation we have following:*

[**1**] ($n = 2k$) *there is infinite many generalised twist solution for corresponding Euler-Lagrange equation and they can be described as*

$$\begin{aligned} y &= \mathrm{Q}(r; a, b, m)\theta \\ &= \mathrm{P} diag[\mathfrak{R}(g), \ldots, \mathfrak{R}(g)]\mathrm{P}^t\theta, \end{aligned} \tag{5.33}$$

where $\mathrm{P} \in \mathrm{SO}(n)$ *and* $g \in \mathrm{C}^\infty[a,b]$ *satisfies*

$$\frac{d}{dr}\left[r^{n-1}[(n-1)r^{-2} + g'^2]^{\frac{p-2}{2}} g'\right] = 0. \tag{5.34}$$

[**2**] $(n = 2k+1)$ $y = x|x|^{-1}$.

Proof. This is an immediate consequence of the previous theorem and Theorem 5.2. $\square$

Bibliography

[1] E. Acerbi, N. Fusco. *Semicontinuity Problems in the Calculus of Variations. Arch. Rational. Mech. Anal.*, 86:125–145, 1984. 6, 10, 64, 71

[2] E. Acerbi, N. Fusco. *A regularity Theorem for Minimizers of Quasiconvex Integrals. Arch. Rational. Mech. Anal.*, 99:261–281, 1987. 55, 71

[3] J. Alibert, B. Dacorogna. *An Example of a Quasiconvex Function that is not Polyconvex in Two Dimensions. Arch. Rational. Mech. Anal.*, 117:155–166, 1992. 60

[4] K. Astala, T. Iwaniec and E. Saksman. *Beltrami Operators in the Plane. Duke Math. J.107*, 1:27–56, 2001. 53

[5] K. Astala, T. Iwaniec, E. Saksman, I. Prause. *Burkholder Integrals, Morrey's Problem and Quasiconformal Mappings. J. Amer. Math. Soc.*, 25:507–531, 2012. 53

[6] A. Baernstein, S. Montgomery. *Some conjectures about integral means of* ∂f *and* $\bar{\partial} f$. *Akta Universitatis Upsaliensis.*, 64:92–109, 1999. 53, 74

[7] J. M. Ball. *Sets of Gradients with no Rank-one Connections. J. Math. Pures et Appliquees.*, 69:241–259, 1990. 68, 69, 70

[8] J.M. Ball. *Convexity Conditions and Existence Theorems in Nonlinear Elasticity. Arch. Rational. Mech. Anal.*, 63:337–403, 1977. 72

[9] J.M Ball, B. Kirchheim, J. Kristensen. *Regularity of Quasiconvex Envelopes. Calc. Var.*, 11:333–359, 2000. 18

[10] J.M. Ball, F. Murat. $W^{1,p}$*-quasiconvexity and Variational Problems for Multiple Integrals. J. Funct. Anal.*, 58:225–253, 1984. 7

[11] J.M. Ball, K.Zhang. *Lower Semicontinuity of Multiple Integrals and the Biting Lemma. . Proc. Roy. Soc. Edinburgh.*, 114:367–379, 1990. 25, 68

[12] J.M. Ball, R.D. James. *Fine Phase Mixtures as Minimizers of Energy. Arch. Rat. Mech. Anal.*, 100:13–52, 1987. 18, 69, 71

[13] J.M. Ball, R.D James. *Proposed Experimental Tests of a Theory of Fine Microstructure and the Two-well Problem. Phil. Trans. Roy. Soc.*, 338:389–450, 1992. 18

[14] R. Banuelos, G. Wang. *Sharp Inequalities for Martingales with Applications to the Beurling-Ahlfors and Riesz transforms. Duke Math. J.*, 80:575–600, 1995. 53

[15] R. Banuelos, P. Janakiraman. *L^p-bounds for the Beurling-Ahlfors Transform. Trans. Amer. Math. Soc.*, 360:3603–3612, 2008. 53

[16] R. Banuelos, P. Janakiraman. *On the Weak-type Constant of the Beurling-Ahlfors Transform. Preprint.*, pages 1–22, 2008. 53

[17] R. Banuelos, P.J Mendez-Hernandez. *Space-Time Brownian Motion and the Beurling-Ahlfors Transform. Indiana Univ. Math. J.*, 52:981–990, 2003. 53

[18] J.V. Burke. *On the Clarke Subdifferential of the Distance Function of a Closed Set. Journal Of Mathematical Analysis and Application.*, 166:199–213, 1992. 53, 54

[19] B. Dacorogna. *Quasiconvexity and Relaxation of Non-Convex Problems in the Calculus of Variations. Journal of Functional Analysis.*, 46:102–118, 1982. 54

[20] B. Dacorogna. *Direct Methods in the Calculus of Variations*, volume 78. Springer second ed., 2008. 9, 18, 73

[21] B. Dacorogna, G. Pisante, A.M. Ribeiro. *On Non Quasiconvex Problems of the Calculus of Variations. Discrete and Continuous Dynamical Systems - Series A (DCDS-A).*, 13:961–983, 2005. 54

[22] J. Diestel and J. J. Uhl, Jr. *Vector Measures. American Mathematical Society, Mathematics Surveys and Monographs.*, 15:11–16, 1977. 67

[23] G. Dolzmann, B. Kirchheim, J. Kristensen. *Conditions for Equality of Hulls in the Calculus of Variations. Arch. Ration. Mech. Anal.*, 154:93–100, 2000. 18

[24] G. Dolzmann, J. Kristensen, K. Zhang. *BMO and Uniform Estimates for Multi-well Problems. Manuscripta Math. Springer-Verlag.*, 140:83–114, 2012. 20

[25] I. Ekeland, R. Temam. *Convex Analysis and Variational Problems.* 1999. 64

[26] L.C. Evans. *Quasiconvexity and Partial Regularity in the Calculus of Variations. Arch. Ration. Mech. Anal.*, 95:227–252, 1986. 56

[27] A. Ionescu Tulcea, C. Ionescu Tulcea. *Topics in the Theory of Lifting. Springer-Verlag.*, 1969. 67

[28] T. Iwaniec, A. Lutoborski. *Integral Estimates for Null Lagrangian. Arch. Ration. Mech. Anal.*, 125:25–79, 1993. 58

[29] T. Iwaniec, J. Kristensen. *A Construction of Quasiconvex Functions . Riv. Mat. Univ. Parma.*, 4:75–89, 2005. 53

[30] T. Iwaniec, A. Lutoborski. *Polyconvex Functionals for Nearly Conformal Deformations. SIAM J. Math. Anal.*, 27:609–619, 1996. 58

[31] D. Kinderlehrer. Remarks about Equilibrium Configurations of Crystals. In: J.M. Ball (ed.). *Material Instabilities in Continuum Mechanics. Oxford Univ. Press.*, pages 217–241, 1988. 69

[32] R. V. Kohn. *The Relaxation of a Double-well Energy. Cont. Mech. Therm.*, 3:193–236, 1991. 20, 42

[33] J. Kristensen. *Lower Semicontinuity in Spaces of Weakly Differentiable Functions. J. Math. Ann.*, 313:653–710., 1999. 9, 53

[34] J. Kristensen. *On the Nonlocality of Quasiconvexity. Ann. Inst. H. Poincaré Anal. NonLinéair.*, 16:1–13, 1999. 72

[35] F.C. Liu. *A Lusin Property of Sobolev Fucntions. Indiana Math.*, 26:645–651, 1977. 55, 71

[36] S. *Müller*. *Variational Models for Microstructure and Phase Transitions. Lecture notes in Math. Springer.*, 1713:85–210, 1999. 18

[37] S. *Müller*, J Sivaloganathan, S.J. Spector. *An Isoperimetric Estimate and $W^{1,p}$-Quasiconvexity in Nonlinear Elasticity. Calc. Var. Partial Differential Equations.*, 8:159–176, 1999. 56

[38] C. B. Morrey. *Quasiconvexity and the Lower Semicontinuity of Multiple Integrals. Pacific J. Math.*, 2:25–53, 1952. 6

[39] C. B. Morrey. *Multiple integrals in the calculus of variations. Springer.*, 130, 1966. 6

[40] Yu.G. Reshetnyak. *On the Stability of Conformal Mappings in the Multidimensional Spaces. Siberian. Math. J.*, 8:69–85, 1967. 70

[41] Yu.G. Reshetnyak. *Stability Theorems for Mappings with Bounded Excursion. Siberian. Math. J.*, 9:499–512, 1968. 70

[42] M.S. Shahrokhi-Dehkordi, A. Taheri. *Generalised Twists, Stationary Loops and the Dirichlet Energy on a Space of Measure Preserving Maps. Calc. Var. and PDEs.*, 35:191–213, 2009. 4, 76

[43] M.S. Shahrokhi-Dehkordi, A. Taheri. *Generalised Twists, $SO(n)$, and the p-Energy over a Space of Measure Preserving Maps. Ann. I. H. Poincáre AN.*, 26:1897–1924, 2009. 4, 76

[44] M.S. Shahrokhi-Dehkordi, A. Taheri. *Quasiconvexity and Uniqueness of Stationary Points on a Space of Measure Preserving Maps. Journal of Convex Analysis.*, 17:69–79, 2010. 4, 76

[45] M.S. Shahrokhi-Dehkordi, A Taheri. *Generalised Twists as Elastic Energy Extremals Annuli, Quaternions and Lifting Twist Loops to the Spinor Groups. Analysis and Applications.*, 11:1350008, 2013. 4, 76

[46] E.M. Stein. *Singularity Integrals and Differentiability Properties of Functions. Princeton University Press.*, 1970. 64

[47] E.M. Stein. *Harmonic Analysis: Real-variable Methods, Orthogonality, and Oscillatory Integrals. Princeton University Press.*, 1993. 56, 58

[48] V. Sverâk. *Quasiconvex Functions with Sub-quadratic Growth. Proc. Math. Soc.*, 433:723–725, 1991. 53, 59, 69

[49] V. Sverâk. *Rank-one Convexity does not Imply Quasiconvexity. Proc. R. Soc. Edinburgh.*, 120:185–189, 1992. 53

[50] V. Sverâk. *Lower Semicontinuity of Variational Integrals and Compensated Compactness. Proceedings of the International Congress of Mathematicians.*, 2:1153–1158, 1994-1995. 72

[51] A. Taheri. *Local Minimizers and Quasiconvexity the Impact of Topology. Arch. Ration. Mech. Anal.*, 176:363–414, 2005. 4

[52] A. Taheri. *Homotopy Classes of Self-maps of Annuli, Generalised Twists and Spin Degree. Arch. Ration. Mech. Anal.*, 197:239–270, 2010. 4, 76

[53] A. Taheri. *Spherical Twists, Stationary Loops and Harmonic Maps from Generalised Annuli into Spheres. Nonlinear Diff. Eqns Appl.*, 19:79–95, 2011. 4, 76

[54] A. Taheri. *Spherical Twists, SO*(n) *and the Lifting of Twist Paths to Spin(n) in Low Dimensions. Quart. J. Math.*, 63:723–751, 2012. 4, 76

[55] L. Tartar. *Compensated Compactness and Applications to Partial Differential Equations. . R. J. Knops, Ed., Research Notes in Mathematics, Nonlinear Analysis and Mechanics.*, 4, 1979. 68

[56] L. Tartar. *The Compensated Compactness Method Applied to Systems of Conservation Laws. In Systems of nonlinear partial Differential Equations. NATO. Sci. Inst. Ser. C.*, 111:263–285, 1983. 68

[57] A. Volberg. *Ahlfors-Beurling Operator on Radial Functions. Math. CA.*, 2012. 53

[58] K. Zhang. *A Construction of Quasionvex Functions with Linear Growth at Infinity. Ann. Scuola Norm. Sup. Pisa Cl. Sci.*, 19:313–326, 1992. 54, 71

[59] K. Zhang. *On Connected Subsets of* $M^{2\times 2}$ *without Rank-one Connections. Proc. Roy. Soc. Edinburgh Section.*, 127:207–216, 1997. 53

[60] K. Zhang. *Quasiconvex Functions, SO*(n) *and Two Elastic Wells. Ann. Inst. H. Poincaré Anal. Non linéair.*, 14:759–785, 1997. 18, 53

[61] K. Zhang. *On Some Quasiconvex Functions with Linear Growth. J. Convex Anal.*, 5:133–146, 1998. 19

[62] K. Zhang. *On the Structure of Quasiconvex Hulls. Ann. Inst. H. Poincaré, Anal. Non Linéair.*, 15:663–686, 1998. 53

[63] K. Zhang. *On Various Semiconvex Relaxations of the Squared Distance Function. Proc. Roy. Soc. Edinburgh Section.*, 129:1309–1323, 1999. 18, 53

Appendix A

Convex Hull and Carathéodory's Theorem

Definition A.0.1. (Convex Hull)
Let K be a subset of $\mathbb{R}^n$. We can call the smallest (nonempty) convex set containing of K the convex hull of K and denote it by $C(K)$ or $co(K)$.

In other words, the convex hull of a set $K \in \mathbb{R}^n$, $C(K)$ is the intersection of all convex subsets in $\mathbb{R}^n$ that contain K.

Carathéodory's theorem is one of the most important characterizations of the convex hull. If $K \subset \mathbb{R}^n$, then the convex combinations of at most $n+1$ points in K are sufficient to describe $C(K)$. In fact, the following holds.

Theorem A.0.1. *Let K be a subset of n-dimensional vector space, $\mathbb{R}^n$. Then, the convex hull of K is equal to the set of convex combinations of at most $n+1$ points. Let K be a subset of $\mathbb{R}^n$. Then,*

$$C(K) = \left\{ a \in \mathbb{R} \middle| \; a := \sum_{i=1}^{n+1} \lambda_i a_i \;\; \lambda_i \geq 0 \;\; a_i \in K \;\; \forall i \quad \sum_{i=1}^{n+1} \lambda_i a_i = 1 \right\}. \tag{A.1}$$

Appendix B

Proof of the Relaxation (Quasiconvexification) Formula for Functions in the Form $f(x) = g(X^+)$

Proof. (Proof of Lemma 3.8.1) In this proof we consider $\mathbb{R}^2$ is the space of all conformal matrices. From the assumption, function $f : \mathbb{M}^{2\times 2} \to \mathbb{R}$ is defined by $f(X) = g(X+)$, where $g : \mathbb{R}^2 \to \mathbb{R}$. Function $X \to g^c(X^2)$ is convex and hence $g^c(X^+) \leq f^c(X)$. From the chain of inequalities relating the semiconvex envelope of f, we can show that

$$f^{rc}(X) \leq g^c(X^+). \tag{B.1}$$

Considering $X = X^+ + X^-$, and fixing $\epsilon > 0$. It suffices to prove that

$$f^{rc}(X) \leq g^c(X^+) + \epsilon. \tag{B.2}$$

By Carethéodory's theorem we find matrices C_i and parameters $\lambda_i \in [0,1], i = 1,2,3$ such that

$$X^+ = \sum_{i=1}^{3} \lambda_i C_i, \qquad g^c(X^+) + \epsilon \geq \sum_{i=1}^{3} \lambda_i g(C_i). \tag{B.3}$$

We may assume that $\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq 0$. There is nothing to prove for $\lambda_2 = 0$ since $g(C_1) = f(X) \geq f^{rc}(X)$. Suppose next that $\lambda_3 = 0$ and $\lambda_2 > 0$. For simplicity we write $\lambda = \lambda_1$ and $1 - \lambda = \lambda_2$. The assertion is immediate if $C_1 = C_2$ and we may therefore assume that $\alpha = |C_1 - C_2| \neq 0$. Let

$$A_1 = \frac{(1-\lambda)\alpha}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad A_2 = -\frac{\lambda\alpha}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{B.4}$$

By construction,

$$
\begin{aligned}
2\det(A_1 + C_1 - (A_2 + C_2)) &= |C_1 - C_2|^2 - |A_1 - A_2|^2 \\
&= \alpha^2 - \left| \frac{\alpha}{\sqrt{2}} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right|^2 = 0.
\end{aligned}
$$

Moreover $\lambda A_1 + (1-\lambda)A_2 = 0$ and hence

$$
\begin{aligned}
g^c(X^+) + \epsilon &\geq \lambda g(C_1) + (1-\lambda) g(C_2) \\
&= \lambda f(C_1 + A_1 + X^-) + (1-\lambda) f(C_2 + A_2 + X^-) \\
&\geq f^{rc}(X),
\end{aligned}
$$

and this establishes (B.2). It remains to consider the case $\lambda_3 > 0$. The idea is again to construct anticonformal matrices A_i such that the pairs $\{(\lambda_i, C_i + A_i + X^-)\}_{i=1,2,3}$ satisfy condition $\mathcal{H}_3$ and

$$
\sum_{i=1}^{3} \lambda_i (C_i + A_i + X^-) = X. \tag{B.5}
$$

A sufficient condition for this to hold is that the matrices A_i solve the system of equations

$$
\sum_{i=1}^{3} \lambda_i A_i = 0, \quad \operatorname{rank}(C_1 + A_1 - (C_2 + A_2)) = 1, \tag{B.6}
$$

and

$$
\operatorname{rank}\left(\frac{\lambda_1}{\lambda_1 + \lambda_2}(C_1 + A_1) + \frac{\lambda_2}{\lambda_1 + \lambda_2}(C_2 + A_2) - (C_3 + A_3) \right) = 1. \tag{B.7}
$$

Which is equivalent to

$$
\left| \frac{\lambda_1}{\lambda_1 + \lambda_2} C_1 + \frac{\lambda_2}{\lambda_1 + \lambda_2} C_2 - C_3 \right|^2 = \left| \frac{\lambda_1}{\lambda_1 + \lambda_2} A_1 + \frac{\lambda_2}{\lambda_1 + \lambda_2} A_2 - A_3 \right|^2, \tag{B.8}
$$

or

$$
|\lambda_1 A_1 + \lambda_2 A_2 - (1 - \lambda_3) A_3|^2 = |A_3|^2 = |X^+ - C_3|^2. \tag{B.9}
$$

This implies that the system for the matrices A_i is equivalent to

$$
\sum_{i=1}^{3} \lambda_i A_i = 0, \quad |A_1 - A_2|^2 = |C_1 - C_2|^2 = \alpha^2, \quad |A_3|^2 = |X^+ - C_3|^2 = \beta^2. \tag{B.10}
$$

If $\alpha = 0$, then $C_1 = C_2$ and therefore

$$
\sum_{i=1}^{3} \lambda_i g(C_i) = (\lambda_1 + \lambda_2) g(C_1) + \lambda_3 g(C_3), \tag{B.11}
$$

and in this situation we employ the argument with two matrices to establish (B.2). Suppose now that $\alpha > 0$ and that $\beta = 0$. It follows that $C_3 = X^+$ and that

$$
\frac{\lambda_1}{\lambda_1 + \lambda_2} C_1 + \frac{\lambda_2}{\lambda_1 + \lambda_2} C_2 = X^+ \tag{B.12}
$$

We define A_1 and A_2 as in the case for two matrices with $\lambda = \lambda_1/(\lambda_1 + \lambda_2)$ and obtain

$$\begin{aligned}\lambda g(C_1) + (1-\lambda)g(C_2) &= \lambda f(C_1 + A_1 + X^-) + (1-\lambda)f(C_2 + A_2 + X^-)\\ &\geq f^{rc}(X^+ + X^-)\end{aligned}$$

and hence

$$\begin{aligned}\sum_{i=1}^{3} \lambda_i g(C_i) &= (\lambda_1 + \lambda_2)\left(\frac{\lambda_1}{\lambda_1+\lambda_2} g(C_1) + \frac{\lambda_2}{\lambda_1+\lambda_2} g(C_2)\right) + \lambda_3 g(C_3)\\ &\geq (\lambda_1+\lambda_2) f^{rc}(X^+ + X^-) + \lambda_3 f(X^+ + X^-)\\ &\geq f^{rc}(X).\end{aligned}$$

It remains to consider the case $\alpha, \beta \neq 0$. Without loss of generality we may assume that

$$A = \mathrm{diag}(c_1, -c_1) \quad A = \mathrm{diag}(c_2, -c_1) \quad A = \frac{\beta}{\sqrt{2}}\mathrm{diag}(1, -1).$$

With this notation we may rewrite the system as

$$\lambda_1 C_1 + \lambda_2 C_2 + \frac{\beta \lambda_3}{\sqrt{2}} = 0, \quad 2(c_1 - c_2)^2 = \alpha^2.$$

We solve for c_1 and c_2 and obtain

$$c_1 = \frac{\lambda_2}{\lambda_1 + \lambda_2}\left(\frac{\alpha}{\sqrt{2}} - \frac{\beta\lambda_3}{\sqrt{2}\lambda_2}\right), \quad c_2 = -\frac{1}{\lambda_1}\left(\lambda_1 c_1 \frac{\beta\lambda_3}{\sqrt{2}}\right).$$

We conclude that

$$\sum_{i=1}^{3} \lambda_i g(C_i) = \sum_{i=1}^{3} \lambda_i f(C_i + A_i + X^-) \tag{B.13}$$

$$\geq f^{rc}\left(\sum_{i=1}^{3} \lambda_i (C_i + A_i + X^-)\right) = f^{rc}(X). \tag{B.14}$$

This establishes (B.1) in the general case and concludes the proof of the lemma. □