QUASICONVEX FUNCTIONS CAN BE APPROXIMATED BY QUASICONVEX POLYNOMIALS

Sebastian Heinz[1]

Abstract. Let W be a function from the real $m \times n$-matrices to the real numbers. If W is quasiconvex in the sense of the calculus of variations, then we show that W can be approximated locally uniformly by quasiconvex polynomials.

1. Introduction

We study quasiconvexity in the calculus of variations. This concept was first introduced by Morrey [9]. He showed that quasiconvexity is an essential condition in the context of the so-called direct methods in the calculus of variations. We refer the reader to Dacorogna [3] for the full statements and a detailed introduction to direct methods as well as applications.

Despite decades of research on quasiconvexity (recent developments include Müller [11], Iwaniec [6], Faraco and Székelyhidi [4]), many problems are still open. See also the survey by Ball [2].

The investigation of quasiconvexity can be reduced to the class of polynomials according to the following new result (equivalent to Th. 6.1).

Theorem 1.1. *Let $m, n \geq 1$ be fixed positive integers.*

Every quasiconvex function $W \colon \mathbb{R}^{m \times n} \longrightarrow \mathbb{R}$ can be approximated locally uniformly by quasiconvex polynomials.

We admit that quasiconvexity remains a challenging property even for polynomials. In fact, up to now, there is no efficient way to decide whether a given polynomial is quasiconvex or not. An illustrating example of a polynomial of degree four can be found in Alibert and Dacorogna [1] (in this context see also Iwaniec and Kristensen [7], Gutiérrez [5]).

Nevertheless, the theorem indicates that the subset of quasiconvex polynomials already provides a comprehensive view on quasiconvexity. We wish to explain, why this was rather not expected to be the case.

[1] Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät II, Institut für Mathematik, DFG-Graduiertenkolleg 1128, Germany; `sheinz@mathematik.hu-berlin.de`

On one hand, Kristensen [8] shows, using Šverák's example [13], that quasiconvexity in dimensions $m \geq 3$, $n \geq 2$ cannot be a "local" property. He constructs a smooth function that is not quasiconvex. However, its restriction to a neighborhood of any fixed point can be extended smoothly to a quasiconvex function on $\mathbb{R}^{m \times n}$. On the other hand, every property is "local" if restricted to the class of polynomials. As a consequence, our density result does not seem obvious.

We remark that the subset of quasiconvex functions has an empty interior with respect to locally uniform convergence on the set of all continuous functions. In particular, every neighborhood of a quasiconvex function contains non-quasiconvex functions. In order to guarantee that an approximating polynomial is quasiconvex, we shall use only such polynomials whose difference to the function W is convex.

This idea works as long as the function W is of polynomial growth. Nevertheless, we are going to prove the result without assuming growth conditions. In order to do so, we will apply an approximation result by Müller [10] about quasiconvex functions that take the value $+\infty$ outside a convex body. This is going to imply that every quasiconvex function can be approximated locally uniformly by quasiconvex functions of polynomial growth.

The paper is organized as follows:

In Section 2 we will fix the notation and symbols that are used later. Definitions and preliminaries can be found in Section 3. In particular, we will give an application of the Stone-Weierstraß theorem. Section 4 is dedicated to a well known density result. It states that quasiconvex functions can be approximated by quasiconvex smooth functions. A result by Müller [10] combined with an idea described here forms the key to the proof of the theorem announced above. This is done in Sections 5 and 6 and completes the paper.

2. Notation

Let $m, n \geq 1$ be fixed positive integers. We shall write a typical element A of the matrix space $\mathbb{R}^{m \times n}$ like

$$A = \left(\begin{array}{ccc} A_{11} & \cdots & A_{1n} \\ \cdots & & \cdots \\ A_{m1} & \cdots & A_{mn} \end{array} \right), \ A_{11}, \dots, A_{mn} \in \mathbb{R}.$$

We equip $\mathbb{R}^{m \times n}$ with a scalar product defined by

$$A : B = \sum_{i=1}^{m} \sum_{j=1}^{n} A_{ij} B_{ij}, \ A, B \in \mathbb{R}^{m \times n}.$$

The corresponding norm of a matrix A is the Frobenius norm $|A|$, so that

$$|A|^2 = A : A, \ A \in \mathbb{R}^{m \times n}.$$

We denote the set of all continuous functions $W : \mathbb{R}^{m \times n} \longrightarrow \mathbb{R}$ by $\mathcal{C}(\mathbb{R}^{m \times n})$ and the subset of all smooth functions by $\mathcal{C}^{\infty}(\mathbb{R}^{m \times n})$.

There is a natural embedding of the ring $\mathbb{R}[A_{11}, \dots, A_{mn}]$ of polynomials into the ring $\mathcal{C}^{\infty}(\mathbb{R}^{m \times n})$ of smooth functions. We denote the range of this embedding by $\mathcal{P}(\mathbb{R}^{m \times n})$, and we identify a polynomial $P \in \mathbb{R}[A_{11}, \dots, A_{mn}]$ with its image in $\mathcal{P}(\mathbb{R}^{m \times n})$.

3. Preliminaries

Recall that a function $W : \mathbb{R}^{m \times n} \longrightarrow \mathbb{R}$ is called *convex* if for all matrices $A, B \in \mathbb{R}^{m \times n}$ and for all $\lambda \in [0, 1]$ it holds that

$$W(\lambda A + (1 - \lambda)B) \leq \lambda W(A) + (1 - \lambda)W(B).$$

A function $W \colon \mathbb{R}^{m \times n} \longrightarrow \mathbb{R}$ is called *quasiconvex* if for all matrices $A \in \mathbb{R}^{m \times n}$ and for all smooth functions $\phi \colon \mathbb{R}^n \longrightarrow \mathbb{R}^m$ with compact support it holds that

$$0 \leq \int_{\mathbb{R}^n} \left[W(A + \mathrm{D}\phi(x)) - W(A) \right] \mathrm{d}x. \tag{3.1}$$

The calculus of variations knows at least two more "convexity concepts", namely, polyconvexity and rank-one convexity. Yet they are not studied in this paper.

We aim to discuss the approximation of quasiconvex functions, so we have to fix a suitable space as well as a topology. It is sufficient to concentrate on the set $\mathcal{C}(\mathbb{R}^{m \times n})$ of all continuous functions, since every convex function is quasiconvex and every quasiconvex function is continuous (even locally Lipschitz continuous). See for example [3], pp. 29–32 and pp. 102–106 for a proof.

Next we associate a topology to $\mathcal{C}(\mathbb{R}^{m \times n})$ that is given *via* converging sequences. We say that a sequence of functions $W_1, W_2, \ldots \in \mathcal{C}(\mathbb{R}^{m \times n})$ *converges locally uniformly* to a function $W \in \mathcal{C}(\mathbb{R}^{m \times n})$ whenever we have uniform convergence for every non-empty compact subset $K \subseteq \mathbb{R}^{m \times n}$, that means

$$\|W_s - W\|_K \longrightarrow 0 \ \text{ as } \ s \longrightarrow \infty.$$

Here $\|W_s - W\|_K$ stands for the supremum norm of the function $W_s - W$ taken over the set K

$$\|W_s - W\|_K = \sup(|W_s(A) - W(A)| \mid A \in K).$$

Consider the family of closed balls $\mathcal{B}_1, \mathcal{B}_2, \ldots \subseteq \mathbb{R}^{m \times n}$ defined by

$$\mathcal{B}_r = \left\{ A \in \mathbb{R}^{m \times n} \mid |A| \leq r \right\}, \ r = 1, 2, \ldots$$

In order to show locally uniform convergence in $\mathbb{R}^{m \times n}$, it suffices to show uniform convergence for every ball $\mathcal{B}_r$.

We shall also work in the set $\mathcal{C}^\infty(\mathbb{R}^{m \times n})$ of smooth functions. In the next remark, we give a sufficient condition for convexity that can be applied to smooth functions. The second derivative of a function $W \in \mathcal{C}^\infty(\mathbb{R}^{m \times n})$ at a matrix $A \in \mathbb{R}^{m \times n}$ can be seen as a symmetric bilinear form

$$\mathrm{D}^2 W(A) \colon \mathbb{R}^{m \times n} \times \mathbb{R}^{m \times n} \longrightarrow \mathbb{R}$$

induced by the quadratic form

$$\mathrm{D}^2 W(A)[B, B] = \left. \frac{\mathrm{d}^2}{\mathrm{d}t^2} W(A + tB) \right|_{t=0}, \ B \in \mathbb{R}^{m \times n}.$$

We say that $\mathrm{D}^2 W(A)$ is positive definite if for every non-zero $B \in \mathbb{R}^{m \times n}$ it holds that

$$\mathrm{D}^2 W(A)[B, B] > 0. \tag{3.2}$$

It is not difficult to verify the validity of the following remark.

Remark 3.1. A smooth function $W \colon \mathbb{R}^{m \times n} \longrightarrow \mathbb{R}$ is convex if the second derivative $\mathrm{D}^2 W(A)$ is positive definite for every $A \in \mathbb{R}^{m \times n}$.

The result of the next lemma is an application of the Stone-Weierstraß theorem to simultaneous approximation of derivatives.

Lemma 3.2. *Let $m, n \geq 1$, $k \geq 0$ be fixed integers.*

Every smooth function $W \colon \mathbb{R}^{m \times n} \longrightarrow \mathbb{R}$ can be approximated by polynomials with respect to locally uniform convergence of all partial derivatives up to order k.

Proof. See for example Sauvigny [12], pp. 6–7. $\qquad\square$

4. Density of smooth functions

The next lemma is well known. Nevertheless, we give the argument, since we want to stress an additional feature of the approximation procedure.

Lemma 4.1. *Let $m, n \geq 1$ be fixed positive integers.*
Every quasiconvex function $W \colon \mathbb{R}^{m \times n} \longrightarrow \mathbb{R}$ can be approximated locally uniformly by quasiconvex smooth functions.

Proof. The proof is by a mollifier argument.

Fix an arbitrary quasiconvex function $W \in \mathcal{C}(\mathbb{R}^{m \times n})$ and choose a non-negative smooth function $\psi \in \mathcal{C}^{\infty}(\mathbb{R}^{m \times n})$ with compact support so that

$$\int_{\mathbb{R}^{m \times n}} \psi(A) \mathrm{d}A = 1.$$

For every $s = 1, 2, \ldots$ set $\psi_s(A) = s^{m \cdot n} \psi(sA)$, $A \in \mathbb{R}^{m \times n}$, and consider the sequence $W_1, W_2, \ldots$ of functions defined by the convolution

$$W_s(A) = \int_{\mathbb{R}^{m \times n}} W(A + T) \psi_s(T) \mathrm{d}T, \ A \in \mathbb{R}^{m \times n}. \tag{4.3}$$

Since W is uniformly continuous on the closed ball $\mathcal{B}_r$, we have uniform convergence on $\mathcal{B}_r$, $r = 1, 2, \ldots$, and we deduce that the functions $W_1, W_2, \ldots$ converge locally uniformly to W. Moreover, we get that every W_s is smooth by iteration of the formula

$$\frac{\partial W_s}{\partial A_{ij}}(A) = \int_{\mathbb{R}^{m \times n}} W(T) \frac{\partial \psi_s}{\partial A_{ij}}(T - A) \mathrm{d}T, \quad 1 \leq i \leq m, \ 1 \leq j \leq n. \tag{4.4}$$

It remains to show that every W_s is quasiconvex. In order to prove that, we apply the definition of quasiconvexity to W_s. After changing the order of integration, the right-hand side of (3.1) reads

$$\int_{\mathbb{R}^{m \times n}} \left[\int_{\mathbb{R}^n} \left[W(A + T + \mathrm{D}\phi(x)) - W(A + T) \right] \mathrm{d}x \right] \psi_s(T) \mathrm{d}T.$$

Quasiconvexity of W implies that the inner integral is non-negative. In addition, we know that ψ_s is a non-negative function. Hence, the function W_s is quasiconvex for every $s = 1, 2, \ldots$ $\qquad\square$

We say that a function $W \in \mathcal{C}(\mathbb{R}^{m \times n})$ has *polynomial growth* if there exists a polynomial $P \in \mathcal{P}(\mathbb{R}^{m \times n})$ so that for all $A \in \mathbb{R}^{m \times n}$ it holds that

$$|W(A)| \leq P(A).$$

Remark 4.2. Assume that the function $W \colon \mathbb{R}^{m \times n} \longrightarrow \mathbb{R}$ in the preceding lemma has polynomial growth. Then (4.3) and (4.4) imply that W_s and every partial derivative of W_s have polynomial growth for every $s = 1, 2, \ldots$

5. Polynomial growth

The proof of the next lemma makes heavily use of techniques applied by Müller [10]. In order to simplify a comparison, we take the same notation as in [10].

Lemma 5.1. *Let $m, n \geq 1$ be fixed positive integers.*
Every quasiconvex function $W \colon \mathbb{R}^{m \times n} \longrightarrow \mathbb{R}$ can be approximated locally uniformly by quasiconvex functions of polynomial growth.

Proof. Fix a positive integer $r \geq 1$ and set $K = \mathcal{B}_r$.

We define $f : \mathbb{R}^{m \times n} \longrightarrow \mathbb{R} \cup \{+\infty\}$ by

$$f(F) = \begin{cases} W(F) & \text{for } F \in K \\ +\infty & \text{for } F \in \mathbb{R}^{m \times n} \setminus K. \end{cases}$$

The function f is quasiconvex, *i.e.* f fulfills the generalized definition of quasiconvexity where the function value $+\infty$ is allowed.

Consider the monotonously increasing sequence $h_0, h_1, \ldots \in \mathcal{C}(\mathbb{R}^{m \times n})$ defined by

$$h_k(F) = f(PF) + k \operatorname{dist}(F, K), \ F \in \mathbb{R}^{m \times n},$$

where P denotes the nearest neighbor projection onto the convex set K and $\operatorname{dist}(F, K) = \min(|F - A| \mid A \in K)$ the distance function.

Let $g_k = h_k^{qc}$, $k = 0, 1, \ldots$, be the quasiconvex envelopes. Every function g_k is continuous, quasiconvex and has polynomial growth, since h_k has polynomial growth. In addition, we conclude that the functions $g_0, g_1, \ldots$ yield a monotonously increasing sequence which is bounded from above by f. See [3] for properties of the quasiconvex envelope.

In order to prove the lemma, we are going to show that the functions $g_0, g_1, \ldots$ converge uniformly to f on the set K. Applying Dini's theorem, it is sufficient to show that $g_0, g_1, \ldots$ converge pointwise to f on K.

We denote by $g_* : K \longrightarrow \mathbb{R}$ the pointwise limit so that

$$g_k(F) \longrightarrow g_*(F), \ F \in K.$$

Now [10], Corollary 9, implies that for every $F \in K$ it holds

$$f(F) = \sup \left\{ g(F) \mid g : \mathbb{R}^{m \times n} \longrightarrow \mathbb{R}, \ g \leq f \text{ on } K, \ g \text{ quasiconvex} \right\}. \tag{5.5}$$

Fix a matrix $F \in \mathbb{R}^{m \times n}$ and denote the right-hand side of (5.5) by G_∞. Müller [10] proves that $f(F) = G_\infty$. An analysis of the proof shows that we can replace G_∞ by $g_*(F)$. Hence, we conclude that $f(F) = g_*(F)$. $\square$

6. Density of polynomials

As the main result, we prove the following theorem.

Theorem 6.1. *Let $m, n \geq 1$ be fixed positive integers.*

Every quasiconvex function $W : \mathbb{R}^{m \times n} \longrightarrow \mathbb{R}$ can be approximated locally uniformly by quasiconvex polynomials.

Proof. As a consequence of Lemmas 4.1, 5.1 and Remark 4.2, we can assume that W is smooth and all its partial derivatives have polynomial growth.

Fix a positive integer $r \geq 1$ and consider the ball $\mathcal{B}_r$. In order to prove the theorem, we are going to show that there exists a sequence of quasiconvex polynomials that converges uniformly to W on $\mathcal{B}_r$.

Consider the functions $W_s \in \mathcal{C}(\mathbb{R}^{m \times n})$, $s = 1, 2, \ldots$, defined by a convex perturbation of W

$$W_s(A) = W(A) + \frac{1}{s} |A|^2, \ A \in \mathbb{R}^{m \times n}.$$

These functions converge locally uniformly to W as $s \longrightarrow \infty$.

Fix a positive integer $s \geq 1$ and a real number $\epsilon > 0$. We aim to construct a polynomial $P \in \mathcal{P}(\mathbb{R}^{m \times n})$ that fulfills the inequality

$$\epsilon > \|P - W_s\|_{\mathcal{B}_r}.$$

In addition, we want to ensure that the difference $P - W$ is a convex function. We wish to apply Remark 3.1. In order to do so, we will show that the second derivative $D^2(P - W)(A)$ is positive definite for every $A \in \mathbb{R}^{m \times n}$.

By definition of W_s, the form $D^2(W_s - W)(A)$ is positive definite for every A in $\mathbb{R}^{m \times n}$. We apply Lemma 3.2 for $k = 2$ and conclude that we can find a polynomial $Q \in \mathcal{P}(\mathbb{R}^{m \times n})$ so that

$$\frac{\epsilon}{2} > \|Q - W_s\|_{\mathcal{B}_{2r}}$$

and that for all $1 \leq i_1, i_2 \leq m$ and all $1 \leq j_1, j_2 \leq n$ it holds that

$$\delta > \left\| \frac{\partial^2}{\partial A_{i_1 j_1} \partial A_{i_2 j_2}} (Q - W_s) \right\|_{\mathcal{B}_{2r}},$$

where we choose $\delta = \delta(r, s) > 0$ sufficiently small so that $D^2(Q - W)(A)$ is positive definite for every $A \in \mathcal{B}_{2r}$. Note that there exists such a number δ, since $\mathcal{B}_{2r}$ is a compact subset of $\mathbb{R}^{m \times n}$ and the left-hand side of (3.2) is a continuous function in the partial derivatives of order two.

The polynomial Q is an appropriate approximation of the function W_s on $\mathcal{B}_{2r}$. However, the function $Q - W$ need not to be convex, since we do not control the second derivative $D^2(Q - W)$ outside $\mathcal{B}_{2r}$. In the last step, we show how to overcome this problem.

Therefore, we consider the convex polynomials $R_t \in \mathcal{P}(\mathbb{R}^{m \times n})$, $t = 2, 3, \ldots$, defined by

$$R_t(A) = \left(\frac{2}{3r} \right)^{2t} |A|^{2t}, \ A \in \mathbb{R}^{m \times n}.$$

They are designed to fulfill two requirements. On one hand, they converge uniformly to zero on $\mathcal{B}_r$ as $t \longrightarrow \infty$. On the other hand, the calculation of the second derivative shows that for every matrix $B \in \mathbb{R}^{m \times n}$ it holds that

$$D^2 R_t(A)[B, B] = \left(\frac{2}{3r} \right)^{2t} \left(4t(t - 1)|A|^{2t-4}(A : B)^2 + 2t|A|^{2t-2}|B|^2 \right).$$

If we skip one non-negative term of the right hand side, we end up with the inequality

$$D^2 R_t(A)[B, B] \geq 2t \left(\frac{2}{3r} \right)^{2t} |A|^{2t-2}|B|^2.$$

We conclude that, for large t, the second derivative of R_t "dominates" the second derivative of $Q - W$ outside the ball $\mathcal{B}_{2r}$, since the partial derivatives of the function $Q - W$ have polynomial growth. In other words, the second derivative $D^2(Q + R_t - W)$ becomes positive definite on the set $\mathbb{R}^{m \times n} \setminus \mathcal{B}_{2r}$ as long as the integer $t \geq 2$ is sufficiently large.

In addition, we know that all polynomials R_t, $t \geq 2$, are convex and that the form $D^2(Q - W)(A)$ is positive definite for every $A \in \mathcal{B}_{2r}$. Together with Remark 3.1, this implies that the function $Q + R_t - W$ is convex for large t.

Now we can choose an integer $t_0 = t_0(r, \epsilon, Q) \geq 2$ in such a way that

$$Q + R_{t_0} - W \text{ is convex and } \epsilon > \|Q + R_{t_0} - W_s\|_{\mathcal{B}_r}.$$

In order to obtain the desired polynomial, we set

$$P = Q + R_{t_0}.$$

Let s tend to infinity and let ϵ tend to zero. By a diagonal argument, we get an approximating sequence of polynomials that converges uniformly to W on the ball $\mathcal{B}_r$.

It remains to prove that the polynomial $P \in \mathcal{P}(\mathbb{R}^{m \times n})$ is quasiconvex for every $r, s \geq 1$ and for every $\epsilon > 0$. However, this is clear, since we have shown that P can be written as the sum of a quasiconvex function and a convex function: $P = W + (Q + R_{t_0} - W)$. $\qquad \square$

References

[1] J.-J. Alibert and B. Dacorogna, An example of a quasiconvex function that is not polyconvex in two dimensions. *Arch. Rational Mech. Anal.* **117** (1992) 155–166.

[2] J.M. Ball, Some open problems in elasticity, in *Geometry, Mechanics, and Dynamics – Volume in Honor of the 60th Birthday of J.E. Marsden*, P. Newton, P. Holmes and A. Weinstein Eds., Springer-Verlag (2002) 3–59.

[3] B. Dacorogna, *Direct Methods in the Calculus of Variations.* Springer-Verlag (1989).

[4] D. Faraco and L. Székelyhidi, *Tartar's conjecture and localization of the quasiconvex hull in* $\mathbb{R}^{2 \times 2}$. Max Planck Institute for Mathematics in the Sciences, Preprint N° 60 (2006).

[5] S. Gutiérrez, A necessary condition for the quasiconvexity of polynomials of degree four. *J. Convex Anal.* **13** (2006) 51–60.

[6] T. Iwaniec, Nonlinear Cauchy-Riemann operators in $\mathbb{R}^n$. *Trans. Amer. Math. Soc.* **354** (2002) 1961–1995.

[7] T. Iwaniec and J. Kristensen, A construction of quasiconvex functions. *Rivista di Matematica Università di Parma* **4** (2005) 75–89.

[8] J. Kristensen, On the non-locality of quasiconvexity. *Ann. Inst. H. Poincaré Anal. Non Linéaire* **16** (1999) 1–13.

[9] C.B. Morrey, Quasi-convexity and the lower semicontinuity of multiple integrals. *Pacific J. Math.* **2** (1952) 25–53.

[10] S. Müller, A sharp version of Zhang's theorem on truncating sequences of gradients. *Trans. Amer. Math. Soc.* **351** (1999) 4585–4597.

[11] S. Müller, Rank-one convexity implies quasiconvexity on diagonal matrices. *Internat. Math. Res. Not.* **20** (1999) 1087–1095.

[12] F. Sauvigny, *Partial differential equations, Foundations and Integral Representations* **1**. Springer-Verlag (2006).

[13] V. Šverák, Rank-one convexity does not imply quasiconvexity. *Proc. Roy. Soc. Edinburgh* **120A** (1992) 185–189.