Progress in Nonlinear Differential Equations
and Their Applications, Vol. 25

Remarks on a Numerical Study of Convexity, Quasiconvexity, and Rank One Convexity*

BERNARD DACOROGNA* AND JEAN-PIERRE HAEBERLY*

∗ Département de Mathématiques, Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne, Switzerland

⋆ Department of Mathematics, Fordham University, Bronx, NY 10458-5165, U.S.A.

1. Introduction

In the calculus of variations, the notions of convexity, quasiconvexity, and rank one convexity play an important role. In this paper we will discuss some numerical results concerning these notions.

We begin by recalling some definitions. Let $\mathbf{R}^{2\times 2}$ denote the set of 2×2 real matrices. Let $f : \mathbf{R}^{2\times 2} \longrightarrow \mathbf{R}$ be continous.

DEFINITION *The function f is* rank one convex *if*

$$f(tA + (1-t)B) \leq tf(A) + (1-t)f(B) \tag{1.1}$$

for every $t \in [0,1]$, and every $A, B \in \mathbf{R}^{2\times 2}$ with $\det(A-B) = 0$, where $\det A$ denotes the determinant of the matrix A.

DEFINITION *The function f is* quasiconvex *if*

$$f(A) \leq \frac{1}{\text{meas}\,\Omega} \int_\Omega f(A + \nabla\varphi(x))\, dx \tag{1.2}$$

for every bounded open set $\Omega \subset \mathbf{R}^2$, every $A \in \mathbf{R}^{2\times 2}$, and every $\varphi \in C_0^\infty(\Omega; \mathbf{R}^2)$.

It is well known that 1.2 implies 1.1, while, obviously, the convexity of f implies 1.2. In concrete examples it might be very hard to decide whether a given function is quasiconvex or not. In particular, it is a long standing question to know if 1.1 and 1.2 are equivalent. (In higher dimensions they are not equivalent, see Sverak [13].)

We propose some numerical algorithms to determine the convexity, quasiconvexity, and rank one convexity of a function. In particular, we study two families

*This research was supported in part by a grant from the *Fonds National Suisse pour la Recherche Scientifique*.

of examples, namely

$$f(A) = |A|^{2\alpha}\left(|A|^2 - \gamma \det A\right) \tag{1.3}$$

$$f(A) = |A|^{4\alpha} - 2^{2\alpha-1}\gamma\left((\det A)^2\right)^{\alpha} \tag{1.4}$$

where $\alpha \geq 0$, $\gamma \in \mathbf{R}$, and $|A|$ denotes the Euclidean norm of a matrix (i.e. $|A|^2 = \sum_{i,j=1}^{2} A_{ij}^2$).

These families of examples have been studied analytically for the rank one convexity by Dacorogna, Douchet, Gangbo, and Rappaz [9]. For the particular case $\alpha = 1$ in the first example, the convexity, quasiconvexity, and rank one convexity were analyzed by Dacorogna and Marcellini [7], and by Alibert and Dacorogna [1]. The case $\alpha < 1$ in the second example has been studied by Ball and Murat [2].

Concerning numerical computations on these examples and on related questions see Dacorogna, Douchet, Gangbo, and Rappaz [9], Gremaud [10], Brighi and Chipot [3], Collins, Kinderlehrer, and Luskin [5], Collins and Luskin [6], Chipot and Lécuyer [4].

Our numerical results are in accordance with the analytical ones whenever these are known. They also tend to show that for the families of examples studied, quasiconvexity and rank one convexity are equivalent. In a forthcoming paper, we will present the numerical analysis of other families of functions, as well as examples in higher dimensions. We will also discuss another notion of convexity, namely polyconvexity, that is relevant in the calculus of variations (see [8] for a precise definition).

2. Analytical Results

We briefly recall the known results on the convexity properties of the two families of examples mentioned above (see [9] for details).

Example 1: Let $\alpha \geq 1$ and define

$$\gamma_1 = \left(1 + \frac{1}{\alpha}\right) \min_{t>0} \left\{ \frac{t^4 + 2(\alpha+1)t^2 + 2\alpha + 1}{3t^3 + (2\alpha+1)t} \right\}$$

and

$$\gamma_2 = 1 + \sqrt{1 - \frac{1}{2\alpha} - \frac{1}{2\alpha^2}}$$

The function

$$f_\gamma(A) = |A|^{2\alpha}\left(|A|^2 - \gamma \det A\right)$$

is rank one convex if and only if $|\gamma| \leq \gamma_r$, where

$$\gamma_r = \begin{cases} \gamma_1 & \text{if } \alpha \in \left[1, \frac{9+5\sqrt{5}}{4}\right), \\ \gamma_2 & \text{if } \alpha \geq \frac{9+5\sqrt{5}}{4} \end{cases}$$

Also, there exist constants γ_c and γ_q such that f_γ is convex, respectively quasiconvex, if and only if $|\gamma|$ is less than or equal to γ_c, respectively γ_q.

When $\alpha = 1$, the value of γ_c is known to be equal to $\frac{4}{3}\sqrt{2}$, while $\gamma_r = \frac{4}{\sqrt{3}}$ in this case. Furthermore, there is an $\epsilon > 0$ such that $\gamma_q = 2 + \epsilon$, but the value of ϵ is not known. When $\alpha > 1$ the values of γ_c and γ_q have not been determined.

Example 2: Let $\alpha \geq \frac{1}{4}$ and define

$$\gamma_1 = \frac{\frac{2(2\alpha-1)}{4\alpha-1}}{\left(1 - \frac{\alpha}{(2\alpha-1)^2}\right)^{2(\alpha-1)}}$$

and

$$\gamma_2 = \frac{2(2\alpha + (2\alpha - 1)z)}{(2\alpha - 1)(1 - z)(1 - z^2)^{\alpha-1}}$$

where

$$z = \frac{-(4\alpha^2 - 1) + \left[(4\alpha^2 - 1)^2 - 8(2\alpha - 1)(\alpha - 1)(4\alpha - 1)\right]^{1/2}}{4(2\alpha - 1)(\alpha - 1)}$$

The function

$$f_\gamma(A) = |A|^{4\alpha} - 2^{2\alpha-1}\gamma\left((\det A)^2\right)^\alpha$$

is rank one convex if and only if $\gamma \leq \gamma_r$, where

$$\gamma_r = \begin{cases} 0 & \text{if } \frac{1}{4} \leq \alpha < 1 \\ 1 & \text{if } \alpha = 1 \\ \gamma_2 & \text{if } 1 < \alpha < 1 + \frac{1}{\sqrt{2}} \\ \gamma_1 & \text{if } \alpha \geq 1 + \frac{1}{\sqrt{2}} \end{cases}$$

When $\alpha \in [\frac{1}{4}, 1]$, f_γ is convex if and only if it is rank one convex.

In general, there exist γ_q, $\gamma_c^+ \geq 0$, and $\gamma_c^- \leq 0$, such that f_γ is quasiconvex if and only if $\gamma \leq \gamma_q$, and is convex if and only if $\gamma \in [\gamma_c^-, \gamma_c^+]$. But for $\alpha > 1$, the values of γ_q, γ_c^+, and γ_c^-, are not known.

Observe that f_γ is homogeneous of degree $2(\alpha + 1)$ in example 1 and homogeneous of degree 4α in example 2. Also, f_γ is always affine in γ, i.e. it can be written as

$$f_\gamma(A) = f_1(A) + \gamma f_2(A)$$

for some functions f_1 and f_2 independent of γ.

3. Numerical Methods

We first discuss numerical methods for computing γ_c (γ_c^+) and γ_r.

Given a function $f : \mathbf{R}^{2\times 2} \longrightarrow \mathbf{R}$, it is easy to show that it is rank one convex if and only if the function

$$\sigma(t) = f(A + tu \otimes v)$$

is convex in t for every $A \in \mathbf{R}^{2\times 2}$ and for every $u, v \in \mathbf{R}^2$. Here $u \otimes v$ is the rank one matrix obtained from the vectors $u = (u_1, u_2)$, $v = (v_1, v_2)$ as follows:

$$u \otimes v = \begin{pmatrix} u_1 v_1 & u_1 v_2 \\ u_2 v_1 & u_2 v_2 \end{pmatrix}$$

Let us assume that f is C^1, and let us define a function

$$h : \mathbf{R}^{2\times 2} \times \mathbf{R}^{2\times 2} \longrightarrow \mathbf{R}$$

as follows:

$$h(A,B) = f(A+B) - f(A) - \langle \nabla f(A), B \rangle.$$

Here $\langle \; ; \; \rangle$ denotes the scalar product on $\mathbf{R}^{2\times 2}$. It is well-known that f fails to be convex if and only if

$$\inf_{A,B} h(A,B) < 0 \tag{3.5}$$

while it fails to be rank one convex if and only if

$$\inf_{A,u,v} h(A, u \otimes v) < 0. \tag{3.6}$$

Now let $f = f_\gamma = f_1 + \gamma f_2$ be a function from example 1 or 2, and let us write h_γ for h. We have

$$h_\gamma(A,B) = h_1(A,B) + \gamma h_2(A,B)$$

where, in example 1,

$$h_1(A,B) = |A+B|^{2(\alpha+1)} - |A|^{2(\alpha+1)} - 2(\alpha+1)\langle A, B\rangle |A|^{2\alpha}$$

$$h_2(A,B) = |A|^{2(\alpha-1)} \left[|A|^2 \langle \tilde{A}, B \rangle + |A|^2 (\det A) + 2\alpha (\det A) \langle A, B \rangle \right] - |A+B|^{2\alpha} \det(A+B)$$

while in example 2,

$$h_1(A,B) = |A+B|^{4\alpha} - |A|^{4\alpha} - 4\alpha \langle A, B\rangle |A|^{4\alpha-2}$$

$$h_2(A,B) = 2^{2\alpha-1} \left[2\alpha (\det A)^{2\alpha-1} \langle \tilde{A}, B \rangle + (\det A)^{2\alpha} - (\det(A+B))^{2\alpha} \right].$$

Here $\tilde{A}$ is the gradient of the determinant function evaluated at A, namely

$$\tilde{A} = \begin{pmatrix} A_{22} & -A_{21} \\ -A_{12} & A_{11} \end{pmatrix}$$

To compute γ_c (or γ_c^+) we may proceed as follows.

ALGORITHM *1. Choose an initial value of γ large enough that f_γ is not convex.*

2. Use a minimization algorithm to compute a pair of matrices (A, B) for which $h_\gamma(A,B) < 0$.

3. Set γ *equal to* $-\frac{h_1(A,B)}{h_2(A,B)}$.

4. Go to step 2.

The algorithm stops when the minimization procedure fails to produce (A, B) with $h_\gamma(A, B) < 0$. The current value of γ is then the computed approximation to γ_c.

An identical appoach can be used to compute γ_r but with $h_1(A, B), h_2(A, B)$ replaced with $h_1(A, u \otimes v), h_2(A, u \otimes v)$.

This method works well and is perfectly suitable to compute γ_c or γ_r, where the optimization takes place in a space of very small dimension. The method is inherently inefficient, however. Indeed, the minimization procedure needs to be applied to the function h_γ for a large number of values of γ. This amounts to a considerable computational effort, even if one uses the computed pair (A, B) of the previous step as the starting point of the minimization process of the current step. In the case of quasiconvexity, where the optimization takes place in a space of very large dimension, it is thus much preferable to compute the threshold values γ_q directly as the minima of certain functions. In order to describe how to construct these functions, we observe the following. Suppose that we are given a family of function $\phi_\gamma : \mathbf{R}^n \longrightarrow \mathbf{R}$ of the form

$$\phi_\gamma(x) = \phi_1(x) + \gamma\phi_2(x)$$

where ϕ_1, ϕ_2 are independent of γ.

If $\phi_1(x) \geq 0$ for all $x \in \mathbf{R}^n$, then

$$\begin{aligned}
\phi_\gamma(x) \geq 0 \text{ for all } x &\iff \phi_1(x) \geq -\gamma\phi_2(x) \text{ for all } x \\
&\iff \begin{cases} -\dfrac{\phi_1(x)}{\phi_2(x)} \geq \gamma \text{ for all } x \text{ with } \phi_2(x) < 0 \\[2ex] -\dfrac{\phi_1(x)}{\phi_2(x)} \leq \gamma \text{ for all } x \text{ with } \phi_2(x) > 0 \end{cases} \\
&\iff \gamma^- \equiv \sup_{\phi_2(x)>0} -\frac{\phi_1(x)}{\phi_2(x)} \leq \gamma \leq \gamma^+ \equiv \inf_{\phi_2(x)<0} -\frac{\phi_1(x)}{\phi_2(x)}.
\end{aligned}$$

By assumption, $0 \in [\gamma^-, \gamma^+]$. Hence

$$\min\{\gamma^+, -\gamma^-\} = \inf_x \frac{\phi_1(x)}{|\phi_2(x)|}.$$

In particular, if $\gamma^+ = -\gamma^- = \gamma$, then

$$\gamma = \inf_x \frac{\phi_1(x)}{|\phi_2(x)|}. \tag{3.7}$$

We now show how 3.7 can be used to estimate γ_q in examples 1 and 2. Recall that in the definition of quasiconvexity it is sufficient to take $\Omega = [0, 1] \times [0, 1]$, the unit square in $\mathbf{R}^2$. Thus the function $f_\gamma : \mathbf{R}^{2\times2} \longrightarrow \mathbf{R}$ is quasiconvex if and only if

$$\int_\Omega f_\gamma(A + \nabla\varphi(x_1, x_2))\, dx_1 dx_2 \; - \; f_\gamma(A) \geq 0 \tag{3.8}$$

for all $A \in \mathbf{R}^{2\times 2}$ and all $\varphi \in C_0^\infty(\Omega; \mathbf{R}^2)$, the space of smooth functions from Ω to $\mathbf{R}^2$ with compact support in Ω. Recall further that it is possible to use Lipshitz functions that are periodic on the boundary instead. In order to approximate the left hand side of 3.8, we use finite elements. More precisely, we fix an integer $m \geq 2$, and we construct a regular mesh of triangles of type 1 on Ω as follows. We partition the interval $[0,1]$ into m subintervals of equal length, and then divide each square of the resulting grid into two triangles, for a total of $2m^2$ triangles. We then replace the function φ with its values at the $(m-1)^2$ interior nodes in the case of zero boundary conditions, or at the m^2 independent nodes in the case of periodic boundary conditions. Our approximation is then a function

$$k_{\gamma,m} : \mathbf{R}^n \longrightarrow \mathbf{R}$$

where n is either $(m-1)^2 + 4$ or $m^2 + 4$ depending on the choice of boundary conditions. Note that $k_{\gamma,m}$ is also affine in γ. As before, we write $k_{\gamma,m}(x) = k_{1,m}(x) + \gamma k_{2,m}(x)$. Note further that $k_{1,m}(x) \geq 0$ for all $x \in \mathbf{R}^n$ since f_1 is a convex function in examples 1 and 2. Now let

$$\gamma_{q,m} = \inf_x \frac{k_{1,m}(x)}{|k_{2,m}(x)|}.$$

By approximation, we have the following:

$$\gamma_q = \lim_{m \longrightarrow \infty} \gamma_{q,m}.$$

For the minimization, we have used an extension of the conjugate gradient algorithm of Polak and Ribière due to Shanno [11], and implemented in the subroutine CONMIN developed by Shanno and Phua [12]. The algorithm uses a cubic interpolation for the line search.

4. Numerical results: example 1

The algorithms were programmed in FORTRAN, and all computations were done on Silicon Graphics workstations equipped with R4000 and R8000 processors.

We have computed γ_c, γ_r, γ_q, for α between 1 and 8 in steps of 0.01. For the case of convexity and rank one convexity, the minimization routine was initialized with a random starting point for every value of α. For the case of quasiconvexity, the minimization algorithm was initialized with a random starting point for $\alpha = 1$. Henceafter, the starting point was taken to be the computed solution at the previous value of α.

When $\alpha = 1$, so that the correct values of γ_c and γ_r are known, the results were as follows:

	computed value	correct value	error
γ_c	1.88561822236594	1.88561808316413	$\sim 1.4 \times 10^{-7}$
γ_r	2.30941239674591	2.3094010767585	$\sim 1.2 \times 10^{-5}$
γ_q	2.30940590626079		

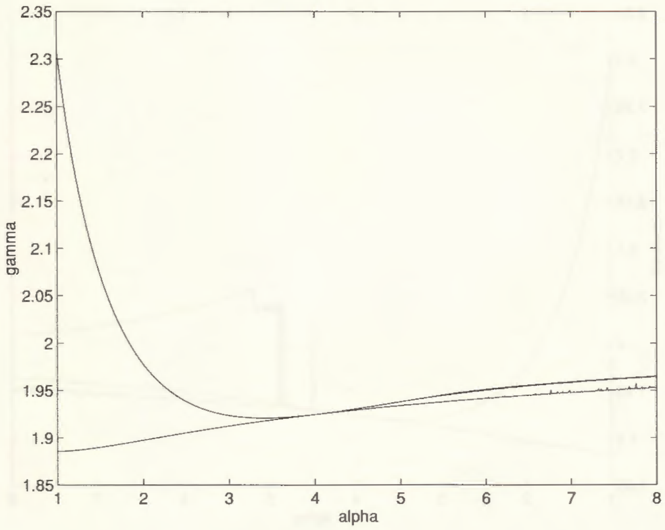

Figure 1: Graphs of γ vs α for convexity, quasiconvexity, and rank one convexity for example 1 (note that the curves for rank one convexity and quasiconvexity are, here, indistinguishable).

More generally, the worst error between the computed values of γ_r and the exact values, when known, was about 1.4×10^{-3}, while the largest gap between the computed values of γ_r and γ_q was about 9.2×10^{-4}.

The computed values of γ_c, γ_r, γ_q are plotted versus α in figure 1. We used periodic boundary conditions. The meshsize for the computation of γ_q was $m = 20$, resulting in 404 variables. Observe that the curve for γ_q and γ_r are nearly identical, and are thus indistinguishable on the plot.

Figure 2 exhibits an interesting phenomenon. The data is the same as that in figure 1 except that the minimization routine for the case of quasiconvexity is initialized with a random starting point for every value of α. As before we observe very good agreement between the γ_r and γ_q curves for $1 \leq \alpha \leq 4.5$. As α passes through the value at which the rank one convexity and the convexity curves become tangent, the computation of γ_q encounters some serious difficulties. This is due to the appearance of local minima whose basin of attraction grows progressively larger as α increases. Eventually, the algorithm, when started at a random point, can no longer compute a meaningful γ_q. An attempt to remedy this deficiency was proposed in [10], replacing the deterministic minimization algorithm with a stochastic one based on the method of simulated annealing. The computational cost is enormous, however, and the results obtained are no better than those displayed in figure 1.

Finally, in figures 4 and 5 we illustrate the differences between the two choices of boundary conditions. Figure 4 displays one component of the solution for the computed γ_q in example 1 with $\alpha = 1$ when one imposes zero boundary conditions,

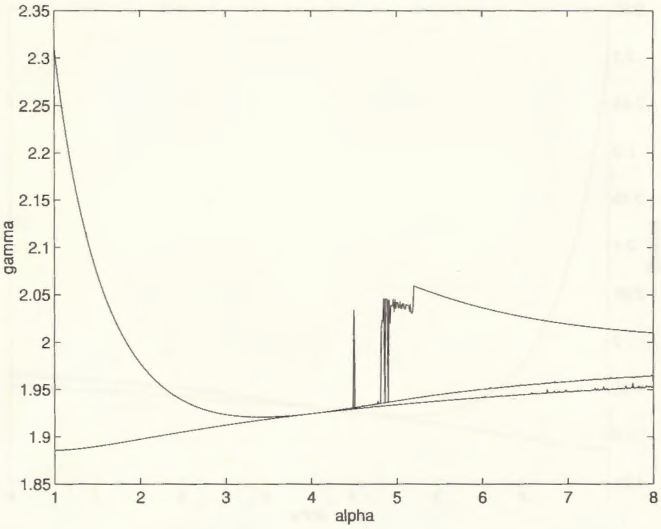

Figure 2: Graphs of γ vs α for convexity, quasiconvexity, and rank one convexity for example 1 with random starting point. The bottom curve is γ_c, the next highest γ_r, and the top one γ_q.

while figure 5 shows the solution with periodic boundary conditions. Each graph is displayed as viewed from two different view angles to emphasize the oscillatory structure. Here $m = 100$ resulting in 9805 nad 10004 variables respectively.

5. Numerical results: example 2

The results of the computations in the case of example 2 are very similar to those obtained for example 1. We computed γ_c, γ_r, and γ_q for α between 1 and 3 in increment of 0.01, using the same procedure as that used to produce figure 1. These values are plotted against α in figure 3. We took $m = 50$ and periodic boundary conditions, resulting in 2504 variables. The worst error between the computed value of γ_r and the correct values was about 1.8×10^{-3}, while the largest gap between the computed values of γ_r and γ_q was about 1.4×10^{-3}. When $\alpha = 1$, the computed value of γ_q was 1.00000808500197, yielding an error of order 10^{-6}. The gap between the computed γ_r and γ_q increases as α approaches 3.

Overall, example 2 appears slightly less amenable to numerical investigation than example 1. Local minima cause difficulties for all values of α in the chosen interval [1,3]. Furthermore, with $m = 20$, the gap between the γ_r and γ_q curves becomes apparent on the plot as α approaches 2. This gap can be very effectively controlled by the choice of m, however, yielding the plot of figure 3 for $m = 50$.

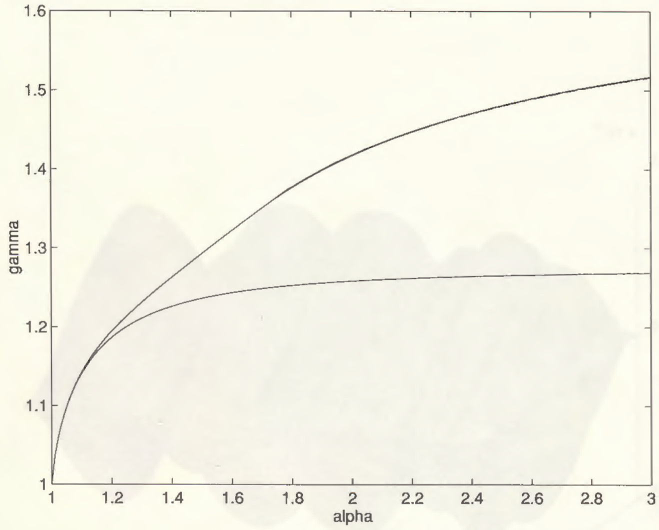

Figure 3: Graphs of γ vs α for convexity, quasiconvexity, and rank one convexity for example 2 (again the curves for rank one convexity and quasiconvexity are, here, indistinguishable).

6. Conclusions

We conclude with a few remarks about the behavior of the numerical methods and the quality of the numerical results.

- Agreement between the computed values and the correct values, when known, are very good.
- When the gap between the computed values of γ_q and γ_r is small for a certain value of m, increasing m further yields no significant gain. When it is large, it can be successfully reduced by increasing m.
- Using periodic boundary conditions gives slightly better values of γ_q than using zero boundary conditions.

We would like to point out that, to the best of our knowledge, this work is the first systematic numerical investigation of convexity, rank one convexity, and quasiconvexity. In prior work, only γ_q has been computed in a few cases. We have computed γ_c, γ_r, and γ_q, for about a thousand different functions. Moreover, the results we obtain are sharper than older results we are aware of.

It is currently not known whether a function

$$f : \mathbf{R}^{2\times 2} \longrightarrow \mathbf{R}$$

is rank one convex if and only if it is quasiconvex. The computations presented above provide some evidence that, for the families of functions in examples 1 and 2, this is indeed the case.

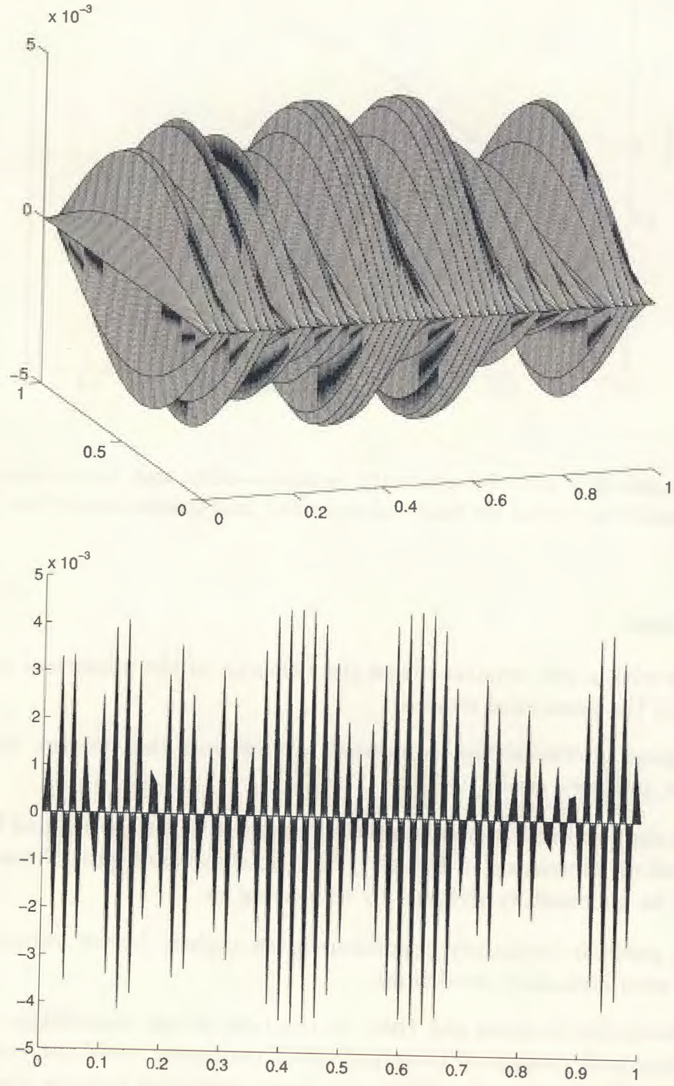

Figure 4: Graph of one component of the solution with 0 boundary conditions for quasiconvexity in example 1 with $\alpha = 1$, viewed from two different angles. The bottom figure illustrates clearly the oscillatory structure of the function. Here $m = 100$ and the computed value of γ_q is 2.3095955016.

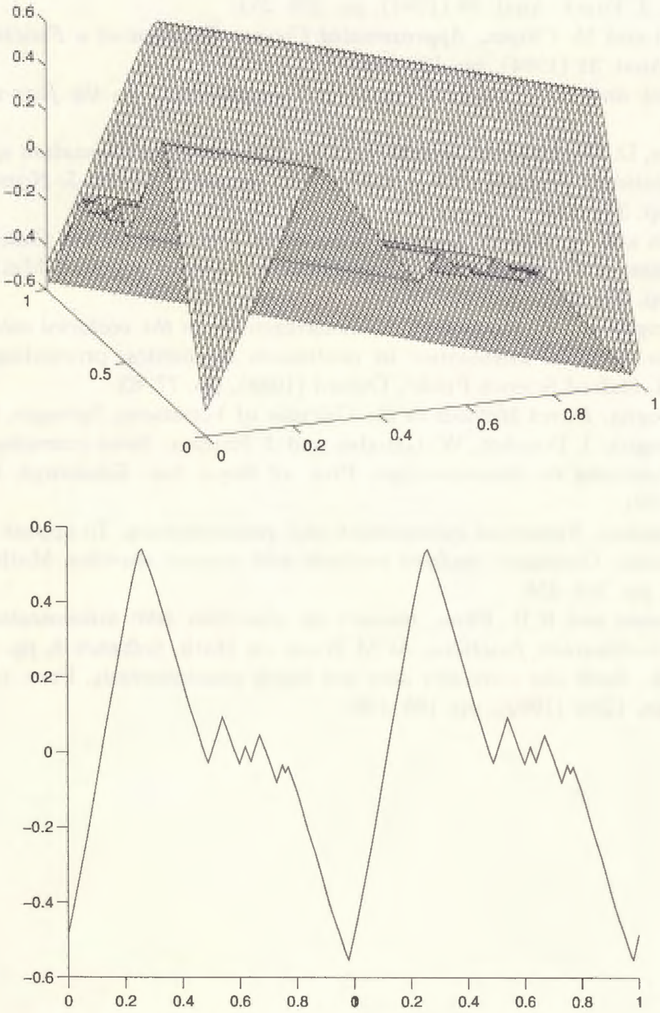

Figure 5: Graph of one component of the solution with periodic boundary conditions for quasiconvexity in example 1 with $\alpha = 1$, viewed from two different angles. The bottom figure illustrates clearly the oscillatory structure of the function. Here $m = 100$ and the computed value of γ_q is 2.3094670086.

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), pp. 155–166.

[2] J.M. Ball and F. Murat, *$W^{1,p}$ quasiconvexity and variational problems for multiple integrals*, J. Funct. Anal. 58 (1984), pp. 225–253.

[3] B. Brighi and M. Chipot, *Approximated Convex Envelope of a Function*, SIAM J. Numer. Anal. 31 (1994), pp. 128–148.

[4] M. Chipot and V. Lécuyer, *Analysis and computations in the four-well problem*, Preprint.

[5] C. Collins, D. Kinderlehrer, and M. Luskin, *Numerical approximation of the solution of a variational problem with a double well potential*, SIAM J. Numer. Anal. 28 (1991), pp. 321–332.

[6] C. Collins and M. Luskin, *Optimal order error estimates for the finite element approximation of the solution of a nonconvex variational problem*, Math. Comp. 57 (1991), pp. 621–637.

[7] B. Dacorogna and P. Marcellini, *A counterexample in the vectorial calculus of variations*, in *Material instabilities in continuum mechanics*, proceedings edited by J.M. Ball, Oxford Science Publc., Oxford (1988), pp. 77–83.

[8] B. Dacorogna, *Direct Methods in the Calculus of Variations*, Springer, Berlin, 1989.

[9] B. Dacorogna, J. Douchet, W. Gangbo, and J. Rappaz, *Some examples of rank one convex functions in dimension two*, Proc. of Royal Soc. Edinburgh 114A (1990), pp. 135–150.

[10] P.A. Gremaud, *Numerical optimization and quasiconvexity.* To appear.

[11] D.F. Shanno, *Conjugate gradient methods with inexact searches*, Math. Oper. Res. 3 (1978), pp. 244–256.

[12] D.F. Shanno and K.H. Phua, *Remark on algorithm 500: minimization of unconstrained multivariate functions*, ACM Trans. on Math. Software 6, pp. 618–622.

[13] V. Sverak, *Rank one convexity does not imply quasiconvexity*, Proc. of Royal Soc. Edinburgh, 120A (1992), pp. 185–189.