NUMERICAL OPTIMIZATION AND QUASICONVEXITY

By

Pierre-Alain Gremaud

IMA Preprint Series # 1133

April 1993

NUMERICAL OPTIMIZATION AND QUASICONVEXITY*

PIERRE-ALAIN GREMAUD†

Abstract. We present a numerical study of quasiconvexity for some functions of the type $f(\xi) = g(|\xi|^2, \det \xi)$, where ξ is a 2×2 -matrix. The corresponding global optimization problems are solved using a simulated annealing-like algorithm. The computations strongly indicate that the considered functions are quasiconvex if and only if they are rank-one convex. Relation to Morrey's conjecture, various applications and implementation problems are discussed.

Key words. global optimization, simulated annealing, quasiconvexity, rank-one convexity

AMS(MOS) subject classifications (1985 revision). 65K10, 49M10

1. Introduction

The goal of this paper is twofold, namely

- introduction and study of a global minimization algorithm,
- (numerical) determination of quasiconvexity and application of the method to examples related to Morrey's conjecture (see below).

The first point will lead to the development, test and use of a general algorithm of *global* minimization based on ideas ranging from quantum mechanics and stochastic differential equations (see [chs], [gi] or [gm]) to more traditional descent methods (gradient or conjugate gradient algorithms). The second point has been a major open problem in the Calculus of Variations for forty years. Recently, Šverák [sv] verified Morrey's conjecture in the three dimensional case; we will thus focus the application of our algorithm on the two dimensional case which is still open.

Let us be more specific. Consider the following minimization problem

$$(1.1) \quad \min\{I(u), u \in W_0^{1,\infty}(\Omega; \mathbb{R}^2)\} \quad \text{with} \quad I(u) = \int_{\Omega} f(\nabla u(x)) dx,$$

where $\Omega \subset \mathbb{R}^2$ is an open bounded set and $f : \mathbb{R}^{2 \times 2} \rightarrow \mathbb{R}$ is a continuous function; $W_0^{1,\infty}(\Omega; \mathbb{R}^2)$ denotes the space of $\mathbb{R}^2$ -valued Lipschitz functions vanishing on the boundary $\partial\Omega$. For a very large class of problems (for instance, nonlinear elasticity), the classical

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

†School of Mathematics, University of Minnesota, Minneapolis, MN 55455, USA.

method which consists in solving the Euler equation $I'(u) = 0$ fails, essentially because of the complexity of the involved calculations. The only available tool consists then in establishing the weak lower semicontinuity of I . This latter property is known to be equivalent to the quasiconvexity of f (see e.g. [da]). For the sake of completeness, let us recall the two following definitions

f is said to be quasiconvex if

$$(1.2) \quad f(\xi) \leq \frac{1}{|\Omega|} \int_{\Omega} f(\xi + \nabla \phi(x)) dx, \quad \forall \xi \in \mathbb{R}^{2 \times 2}, \forall \phi \in W_0^{1,\infty}(\Omega; \mathbb{R}^2);$$

f is said to be rank-one convex if

$$(1.3) \quad f(\lambda \xi + (1 - \lambda) \eta) \leq \lambda f(\xi) + (1 - \lambda) f(\eta), \\ \forall \lambda \in [0, 1], \forall \xi, \eta \in \mathbb{R}^{2 \times 2}, \det(\xi - \eta) = 0.$$

It turns out that it is usually extremely difficult to establish the quasiconvexity of a function. The notion of rank-one convexity is known to be slightly weaker (i.e. “quasiconvexity $\Rightarrow$ rank-one convexity”, see e.g. [da]). It was conjectured by Morrey (see [mo]), that f rank-one convex $\nRightarrow f$ quasiconvex. This problem is still open in the two dimensional case. Numerous attempts have been made to construct examples of functions which are rank-one convex but not quasiconvex (see [ad], [ddgr]). The complexity of the involved calculations have not yet permitted their complete analytical study (the example of Šverák seems to be purely three-dimensional and not “translatable” to a two dimensional setting). As will be seen from the computations, the numerical analysis is quite an efficient tool with which to study this problem.

From a numerical point of view, a quasiconvexity test will consist in the global minimization of a functional in some space $\mathbb{R}^p$. The algorithm we consider here is, in essence, not new and can be related to some kind of Monte Carlo method. The idea is based on the annealing process in which one goes from one configuration of a system to a more “stable” one by heating and then cooling down very slowly. In this frame, the unknown function in the optimization problem is viewed as the state of an imaginary physical system whose energy has to be minimized. Results of convergence towards global minima have been established, and the optimal annealing schedule (i.e. the way to change the “temperature” in order to reach equilibrium as fast as possible) is known. We refer the reader to [gi] (see also [chs] and [gm]).

Algorithms based on these ideas are referred to as simulated annealing methods (SA), and can be applied to a huge class of minimization problems. However, to our knowledge and much to our surprise, closely related methods have only been used for solving much more simple problems than the one we are interested in here (see e.g. [apz] for some low

dimensional applications; see also [lm] where a crude version of SA, namely the Metropolis algorithm, has been successfully used for a “real life” problem). On the other hand, there exists in the literature several computational and numerical studies of quite complicated variational problems (see e.g. [cck], [cl], [gr]) but the emphasis is on the use and convergence of the finite element method (i.e., in the case of problem (1.1), the construction of finite dimensional spaces approximating functions in $W_0^{1,\infty}$) rather than on the much more involved issue of the actual global minimization of the functional.

The outline of the paper is as follows. In section 2, we introduce and analyze the spatial discretization using the Finite Element method. The simulated annealing method is discussed in section 3. We present the examples considered in this paper in section 4. Some theoretical results are recalled. In section 5, the detailed numerical test of quasiconvexity is introduced. The method is evaluated and the numerical results presented and discussed. Section 6 is devoted to concluding remarks. Finally, in the appendix, a PDE approach to the simulated annealing method is considered.

2. Discretization

In this section, we analyze the discretization of the problem (1.1) by the finite element method.

Let $\{\mathcal{T}_h\}$ be a family of regular triangulations of Ω , h being the size of the mesh. For the sake of simplicity, we assume that Ω is polygonal and $\bar{\Omega} = \cup_{K \in \mathcal{T}_h} K$. We set

$$\begin{aligned} V_h &= \{\phi : \Omega \rightarrow \mathbb{R}, \phi \in C^0(\bar{\Omega}), \phi|_K \in \mathcal{P}_1(K), \phi = 0 \text{ on } \partial\Omega\}, \\ W_h &= V_h^2, \\ \text{Aff}_0(\Omega; \mathbb{R}^2) &= \{\phi \in W_0^{1,\infty}(\Omega; \mathbb{R}^2), \phi \text{ piecewise affine}\}, \end{aligned}$$

$\mathcal{P}_1$ being the space of polynomials of degree 1. For any $\xi \in \mathbb{R}^{2 \times 2}$, we introduce the notations

$$\begin{aligned} Qf(\xi) &= \inf\left\{\frac{1}{|\Omega|} \int_{\Omega} f(\xi + \nabla u(x)) dx, u \in W_0^{1,\infty}(\Omega; \mathbb{R}^2)\right\}, \\ Q'f(\xi) &= \inf\left\{\frac{1}{|\Omega|} \int_{\Omega} f(\xi + \nabla u(x)) dx, u \in \text{Aff}_0(\Omega; \mathbb{R}^2)\right\}, \\ Q_h f(\xi) &= \inf\left\{\frac{1}{|\Omega|} \int_{\Omega} f(\xi + \nabla u(x)) dx, u \in W_h\right\}. \end{aligned}$$

We have the following result.

LEMMA 2.1.

$$(2.1) \quad Qf(\xi) = Q'f(\xi) \quad \forall \xi \in \mathbb{R}^{2 \times 2},$$

$$(2.2) \quad \lim_{h \rightarrow 0} Q_h f(\xi) = Qf(\xi) \quad \forall \xi \in \mathbb{R}^{2 \times 2}.$$

Proof. The relation (2.1) is proved in [da], p.207–212. It is then not difficult to establish (2.2) (see [bc]). $\square$

REMARK 2.1. Let us point out that in the scalar case (i.e. $u \in W_0^{1,\infty}(\Omega; \mathbb{R})$ and $f : \mathbb{R}^2 \rightarrow \mathbb{R}$), it is possible to prove that $Q_h f(\xi)$ tends to $Qf(\xi)$ as $h^{1/2}$, provided that $\frac{f(\xi)}{|\xi|} \xrightarrow{|\xi| \rightarrow \infty} +\infty$ (see [bc]). To our knowledge, there is no such result in the present vectorial case. Finally, we recall that Qf does not depend on Ω (see [da]). $\square$

We now turn to the elaboration of a numerical test of quasiconvexity. We will restrict ourself to homogeneous functions f , i.e., such that there exists a real number p with

$$(2.3) \quad f(\lambda \xi) = \lambda^p f(\xi) \quad \forall \lambda > 0.$$

We set

$$J(\xi, \phi) = \int_{\Omega} (f(\xi + \nabla \phi(x)) - f(\xi)) dx.$$

The quasiconvexity of f is then equivalent to

$$(2.4) \quad \inf_{\xi \in \mathbb{R}^{2 \times 2}} \inf_{\phi \in W_0^{1,\infty}(\Omega; \mathbb{R}^2)} J(\xi, \phi) = 0.$$

Let us remark that by (2.3), the left member of the previous relation is either 0 or $-\infty$.

We say that f is h -quasiconvex if

$$(2.5) \quad \inf_{\xi \in \mathbb{R}^{2 \times 2}} \inf_{\phi \in W_h} J(\xi, \phi) = 0.$$

The lemma 2.1 justifies such a denomination.

By allowing a slight stretch in the notation, a function ϕ in W_h is identified with the $2 \times N$ -matrix of its values at the nodes of $\mathcal{T}_h$, N being the number of the nodes. The $2 \times (N + 2)$ -matrix corresponding to (ξ, ϕ) is then denoted Φ and (2.5) becomes

$$(2.6) \quad \inf_{\Phi \in \mathbb{R}^{2 \times (N+2)}} J(\Phi) = 0.$$

3. Numerical optimization

In this section, the algorithm of minimization is introduced and studied.

We are interested in the numerical resolution of the following problem

$$(3.1) \quad \inf_{\Phi \in \mathbb{R}^p} J(\Phi),$$

for a family of functions J as large as possible (see e.g. (2.6)).

The traditional algorithms of descent (gradient or conjugate gradient methods) are essentially based first, on the observation that solutions to the following time dependent equation

$$(3.2) \quad \frac{d\Phi}{dt}(t) = -\nabla J(\Phi(t)), \quad t > 0,$$

(where ∇ denotes the derivative with respect to Φ) tend to local minima of J as t tends to infinity, and secondly on the discretization of this equation by a single Euler forward scheme, i.e.,

$$(3.3) \quad \Phi^{k+1} = \Phi^k - \tau^k \nabla J(\Phi^k),$$

where the time steps, τ^k , may or may not be “optimized.” A major drawback suffered by this approach is that it is obviously *local*; the missing part in (3.2), i.e., the choice of the initial condition, plays a central role.

REMARK 3.1. Such methods may, nevertheless, be used for general problems, but with the two possible *major* restrictions :

- to repeat the computations for a large number of initial conditions (the hope being that one will eventually choose a starting point in the “basin of attraction” of a global minimizer);
- to have an *a priori* knowledge of the solution (!) or other closely related informations (e.g. Young measures associated to minimizing sequences, see e.g. [cl2]). $\square$

If one aims at finding global minimizers, starting with “arbitrary” initial conditions, one obviously has to be able to overcome potential barriers (i.e., “to climb hills”). In order to do this, we introduce a perturbation in the relation (3.2) in the form of a stochastic fluctuation

$$(3.4) \quad d\Phi(t) = -\nabla J(\Phi(t))dt + \sqrt{2\vartheta(t)}dw(t),$$

where the “temperature” $\vartheta(t)$ is a given annealing rate and $w(t)$ is a standard Brownian motion in $\mathbb{R}^p$. We refer the reader to [sc], chap.8, for various comments on the physical significance of (3.4).

Using stochastic analysis techniques, it is possible to prove, for suitable functions J and ϑ , that Φ , solution of (3.4), tends to a global minimizer “in some probabilistic sense” (see e.g. [chs], [gi] or [gm]). These results are beyond the scope of this paper, and we refer the reader to the abovementioned papers for more details (see also the naive approach

presented in the appendix). Let us however remark that, central in these papers, is the fact that the annealing rate $\vartheta(t)$ has to be very slowly decreasing in order to ensure convergence. We accordingly choose

$$(3.5) \quad \vartheta(t) = \frac{c}{\log(2+t)}, \quad c \geq 0.$$

The minimizing algorithm is then :

find $\{\Phi^k\}_{k \in \mathbb{N}} \subset \mathbb{R}^p$ such that

$$(3.6) \quad \begin{cases} \Phi^{k+1} = \Phi^k - \tau^k g^k + \sqrt{2\vartheta(t^k)} \tilde{w}^k, \\ \Phi^0 = \Phi_0, \end{cases}$$

where ϑ is given by (3.5) and g^k is given by

$$g^k = \nabla J(\Phi^k) \quad \text{or} \quad g^k = \nabla J(\Phi^k) + \beta^k g^{k-1},$$

where $\beta^k = \frac{(\nabla J(\Phi^k) - \nabla J(\Phi^{k-1}), \nabla J(\Phi^k))}{\|\nabla J(\Phi^{k-1})\|^2}$, i.e., using respectively the gradient or conjugate gradient algorithm. Moreover, we have set in (3.6) $t^k = \sum_{i=0}^{k-1} \tau^i$, where the time steps τ^k are experimentally determined; $\tilde{w}^k = w(t^{k+1}) - w(t^k)$ is a randomly generated vector in $\mathbb{R}^p$ such that $\tilde{w}^k \in [-1, 1]^p$ and $\Phi_0 \in \mathbb{R}^p$ is an initial condition.

4. The examples of Dacorogna & Marcellini

In this section, we consider functionals of the type (1.1) where $f : \mathbb{R}^{2 \times 2} \rightarrow \mathbb{R}$ is as follows

$$(4.1) \quad f(\xi) = g(|\xi|^2, \det \xi),$$

form some function $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ and where $|\cdot|$ denotes the euclidean norm. Such functions are important in elasticity or in optimal design (see [da], appendix, or [ks]).

Our first and principal example will be functions of the type

$$(4.2) \quad f(\xi) = |\xi|^{2\alpha} (|\xi|^2 - 2\gamma \det \xi) \quad \alpha \geq 1/2, \gamma \in \mathbb{R}.$$

Let us recall the following result.

THEOREM 4.1. *Let f be given as in (4.2), with $\alpha = 1$. Then*

$$\begin{aligned} f \text{ convex} &\Leftrightarrow |\gamma| \leq \frac{2}{3}\sqrt{2}, \\ \exists \varepsilon > 0 \text{ such that } f \text{ quasiconvex} &\Leftrightarrow |\gamma| \leq 1 + \varepsilon, \\ f \text{ rank-one convex} &\Leftrightarrow |\gamma| \leq \frac{2}{\sqrt{3}}. \end{aligned}$$

Proof. See [ad], theorem 1. $\square$

The most difficult point in the above result is obviously the quasiconvexity. We give in the next section an estimation of the parameter ε .

The general case (4.2) is treated in [ddgr], theorem 1.8, where necessary and sufficient conditions for rank-one convexity are established. Optimal quasiconvexity conditions are unknown and will be numerically determined in the next section.

In a second example, the following kind of functions is studied

$$(4.3) \quad f(\xi) = |\xi|^{2\alpha} - 2^{\alpha-1}\gamma|\det \xi|^\alpha.$$

Rank-one convexity conditions for the above functions f are established in [ddgr], theorem 1.6. Here also, no quasiconvexity condition is known.

5. Numerical results

In this section, we first briefly evaluate the simulated annealing method (SA) with respect to other more traditional descent methods. The examples of the previous section are used as a bench-mark. In a second part, we discuss computational results showing that those functions are “ h -quasiconvex” if and only if they are rank-one convex.

Throughout this section, we set $\Omega = (0, 1)^2$ and $\mathcal{T}_h = \{K_{l,m,p}, 1 \leq l, m \leq N, p = 1, 2\}$ where $K_{l,m,1}$ (resp. $K_{l,m,2}$) is the triangle with vertices $(lh, (m-1)h)$, $((l-1)h, mh)$ and $((l-1)h, (m-1)h)$ (resp. (lh, mh)) and where $h = 1/N$.

As pointed out in [gi], the optimal annealing schedule is determined by the competition between the two following effects

- if the temperature is lowered too abruptly, the solution may get stuck in some local minimum (“adiabatic” effect);
- if the temperature is lowered too slowly, the convergence (if any) towards a global minimum might be too slow to be of any practical use (“supercooling” effect).

In our case, the speed of cooling is measured by the coefficient c in (3.5). The value $c = 0$ corresponds to the traditional descent algorithms.

We now illustrate the previous remark by considering the algorithm (3.5), (3.6) applied to the example (4.2) with $\alpha = 1$, $\gamma = 1.4$. The functional is defined according to (2.6) and we use $\tau = 10^{-4}$ and $N = 20$. According to theorem 4.1, the continuous corresponding functional is not quasiconvex and thus we should find

$$(5.1) \quad \inf_{\Phi \in \mathbb{R}^{2 \times (N+2)}} J(\Phi) = -\infty.$$

In view of figure 1, the traditional method does not “converge” after the 1000 displayed iterations (in this example, J becomes negative after ≈ 1400 iterations). On the other hand, SA gives very satisfactory answers after only a few hundred iterations.

Figure 1. Simulated annealing with $c = 0, 10^{-6}, 10^{-4}, 10^{-3}$ (example 4.2 $\alpha = 1, \gamma = 1.4$).

If the annealing coefficient c is further increased, then oscillations appear and rapidly lead to an overflow.

REMARK 5.1. The algorithm is actually very sensitive to the way random numbers are generated. This could somehow render the algorithm quite “machine dependent” (at least if one uses a built-in random number generator). For instance, we use in this paper a generator of the type GFSR (generalized feedback shift register) with period $2^{521} - 1$, see [ri], appendix; this leads to a substantial improvement with respect to the algorithm obtained in conjunction with the built-in random number generator RANF provided by CRAY (which has a period of $2^{48} - 1$). The computations presented below have been obtained on a CRAY-2/4-512. $\square$

Let us conclude this brief evaluation by the following remarks. The advantages and disadvantages of SA versus the classical descent methods have to be analyzed and weighed from case to case. For instance, for convex optimization problems, there is no point in adding some fluctuation terms in the algorithm. On the other hand, the more “nonconvex” the functional is, the more needed those terms will be. We doubt that any much more precise, definite and serious statement can be offered in the *general* case.

The applications presented hereafter are indeed a perfect example of a “mixed” approach, where traditional and stochastic algorithms are both used to solve a same problem.

We now consider the application of SA to the examples of Dacorogna & Marcellini (see section 4). As recalled in the previous section, necessary and sufficient conditions for rank-one convexity are known if f is given by (4.2) (or (4.3)). Our goal is to find if there exists a rank-one convex function which is not quasiconvex.

Roughly speaking, the strategy is as follows : if γ is “large” (see e.g. theorem 4.1), then the function is not rank-one convex and thus not quasiconvex; consequently, for “large” γ , the relation (5.1) has to be found. Therefore, for various fixed values of α , we are looking for the smallest γ , denoted γ_h^* such that (5.1) is satisfied. The theoretical critical value for rank-one convexity is denoted γ_r . Three cases are then possible.

- γ_h^* is much larger than γ_r : the numerical test is not precise enough (since, neglecting numerical errors, γ_h^* has to satisfy $\gamma_h^* \leq \gamma_r$, see also remark 5.2);
- γ_h^* is reasonably close to γ_r : this would tend to imply that the function is quasiconvex if and only if it is rank-one convex;
- γ_h^* is smaller than γ_r .

REMARK 5.2. Even if the continuous function was such that $\gamma^* = \gamma_r$ (where γ^* is the largest number such that f is quasiconvex), we should have $\gamma_h^* \gtrsim \gamma^*$ because of the “convexifying” effect of the discretization itself. $\square$

For a given fixed value of α in (4.2) or (4.3), the numerical test of quasiconvexity is as follows

```

0.  CHOOSE  $h$  AND  $\gamma$  ( $\gamma \geq \gamma_r$ )
1.  INITIALIZE  $\Phi^0, k = 0$ 
2.  COMPUTE  $\nabla J(\Phi^k)$ 
3.  COMPUTE  $\Phi^{k+1}$  ACCORDING TO (3.5), (3.6)
4.  IF  $J(\Phi^{k+1}) \ll 0$  (NON QUASICONVEXITY) THEN
      SET  $\gamma \leftarrow \gamma - \delta\gamma, \Phi^0 \leftarrow \Phi^{k+1}, k = 0$ 
      GO TO 2
    ELSE
      SET  $k \leftarrow k + 1$ 
      GO TO 2 (TILL # ITERATIONS = MAX)
    ENDIF .

```

Some comments are in order. First, SA is most useful for going from the initial condition to a well where a global minimum lies. This means that after a first successful

computation with a value γ , one can, if the decrement $\delta\gamma$ is small, switch off the noise term afterwards (the hope being that the structure of J does not change too much if γ is slightly changed). Secondly, the use of SA is here more than a mere numerical gadget. Indeed, for small h and large α , for instance, the functional is so steep that an arbitrary choice of an initial condition leads to unsolvable numerical problems (the energy being simply too large to be efficiently numerically handled). A natural choice for the initial condition is then $\Phi^0 = (\xi^0, 0)$, where ξ^0 is any 2×2 matrix. It is easy to check that Φ^0 is a local minimum of J and satisfies $J(\Phi^0) = 0$. Contrary to traditional methods, SA allows us to start from such a local minimum without getting stuck there (see figure 2).

Figure 2. SA started at a local minimum (example 4.2).

The results of the quasiconvexity test are shown in the following table.

	$h=1/20$	$h=1/30$	$h=1/40$	γ_r
$\alpha = 1$	1.1571	1.1553	1.1550	1.1547
$\alpha = 2$	0.9906	0.9896	0.9894	0.9889
$\alpha = 6$	0.9759	0.9754	0.9753	0.9751

Example 4.2 : values of γ_h^* .

The above results have been obtained by starting with values of γ such that the corresponding function is “seriously non quasiconvex”. Typically, 10 to 15 decreasing values of γ are used and for each of these values several thousand iterations are needed.

The values of γ_h^* in the above table are such that : for the value itself, non quasiconvexity has been observed, moreover if the last digit is diminished by one unit, $J(\Phi^k)$ is still positive for $k = 10000$. The agreement between the computed values and the theoretical values for rank-one convexity (see section 4) is very good. The considered functions seem thus to be quasiconvex if and only if they are rank-one convex, i.e., $\lim_{h \rightarrow 0} \gamma_h^* = \gamma_r$.

In figure 3, level curves of one of the two components of the unknown ϕ are represented for various values of h . The picture shows that, as the mesh is refined, new modes appear that were not allowed with coarser meshes.

Figure 3. Level curves of the first component of ϕ for $h = 1/20, 1/24, 1/30, 1/40$ (example 4.2 $\alpha = 1, \gamma = 1.158$).

In all the above cases, the solution is obviously affected by various “mesh effects.” The following question naturally arises : have we really found a “minimizer” when we claim that the function is quasiconvex (or more precisely “ h -quasiconvex”)? Would we still observe quasiconvexity if the number of iterations is further increased or the mesh refined? We strongly believe it is the case, and the agreement between the theoretical values for rank-one convexity and the experimental values for quasiconvexity observed in

Figure 4. Evolution of J for various values of γ ($\alpha = 1$, $h = 1/20$)

the above table is more than a mere coincidence.

The figure 4 backs this belief. The time at which J becomes negative goes to infinity as γ decreases. We should however point out that the curves in figure 4 have been obtained by taking the same initial condition for each value of γ (which is not consistent with our algorithm).

In the case of the example 4.3, the numerical results are of the same nature, although more delicate to obtain (the “tricky” determinant term is here more dominant). The results for a mesh size $h = 1/20$ are given in the following table.

	$h=1/20$	γ_r
$\alpha = 3$	1.2946	1.2926
$\alpha = 4$	1.4237	1.4169

Example 4.3 : values of γ_h^* .

Here also, the numerical results suggest that the considered functions are quasiconvex if and only if they are rank-one convex.

6. Conclusions

We have successfully used the simulated annealing method (SA) in the framework of a numerical test of quasiconvexity. The algorithm has been tested for functions which were considered in relation to Morrey's conjecture (i.e., possible candidates for non quasiconvex functions which are rank-one convex).

The computations strongly suggest that those functions are quasiconvex if and only if they are rank-one convex, leaving thus Morrey's conjecture unanswered in dimension 2, but ruling out a family of possible candidates. The generality of this "result" is difficult to evaluate and the general case $f(\xi) = g(|\xi|^2, \det \xi)$ undoubtedly deserves further studies, despite the theoretical and numerical difficulties.

As far as the numerical method is concerned, the following remarks are worth considering. First, SA allows a lot of freedom in the choice of a starting point without getting stuck in some local minimum (even local minima are possible as initial conditions). In the case of "nasty" functionals, we have shown that, properly tuned, it works far better than the classical descent methods : SA allows the iteration to follow "the short way". The second point is that SA may turn out to be quite tricky to tune. This, in our view, constitutes its main disadvantage, but it should not prevent its broader use in the framework of global optimization.

Finally, we have focused here on an application of numerical analysis to mathematics. However, and not surprisingly, the problems encountered (global optimization, mesh effects,...) are by no means limited to this theoretical frame.

ACKNOWLEDGMENTS

I would like to thank Bernard Dacorogna for having brought this problem to my attention, as well as for useful comments. I am also grateful to Mitchell Luskin for several remarks which led to a better presentation of this paper.

References

- [ad] J.-J. ALIBERT & B. DACOROGNA, *An Example of a Quasiconvex Function that is not Polyconvex in Two Dimensions*, Arch.Rat.Mech.Anal., 117 (1992), pp. 155-166.
- [apz] F. ALUFFI-PENTINI, V. PARISI & F. ZIRILLI, *Global Optimization and Stochastic Differential Equations*, J. Optim. Theory Applications, 47 (1985), pp. 1-16.
- [bc] B. BRIGHI & M. CHIPOT, *Approximated Convex Envelope of a Function*, preprint.
- [cck] M. CHIPOT, C. COLLINS & D. KINDERLEHRER, *Numerical Analysis of Oscillations in Multiple Well Problems*, in preparation.
- [chs] TZUU-SHUH CHIANG, CHIH-RUEY HWANG & SHUENN-JYI SHEU, *Diffusion for Global Optimization in $\mathbb{R}^n$* , SIAM J. Control and Optimization, 25 (1987), pp. 737-753.
- [cl] C. COLLINS & 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.

- [cl2] C. COLLINS & M. LUSKIN, *The Computation of Austenitic-Martensitic Phase Transition*, in Partial Differential Equations and Continuum Models of Phase Transitions, M. Rascle, D. Serre & M. Slemrod eds., Lecture Notes in Physics, Springer Verlag, 1989, pp. 34–50.
- [da] B. DACOROGNA, *Direct Methods in the Calculus of Variations*, Springer Verlag, 1989.
- [ddgr] B. DACOROGNA, J. DOUCHET, W. GANGBO & J. RAPPAPAZ, *Some Examples of Rank One Convex Functions in Dimension Two*, Proc. of Royal Soc. Edinburgh, 114A (1990), pp. 135–150.
- [fr] A. FRIEDMAN, *Stochastic Differential Equations and Applications*, vol.1, Academic Press, 1975.
- [gi] B. GIDAS, *Metropolis-type Monte Carlo Simulation Algorithms and Simulated Annealing*, Preprint, Brown University (1991).
- [gm] S.B. GELFAND & S.K. MITTER, *Recursive Stochastic Algorithms for Global Optimization in $\mathbb{R}^d$* , SIAM J. Control and Optimization, 29 (1991), pp. 999–1018.
- [gr] P.A. GREMAUD, *Numerical Analysis of a Nonconvex Variational Problem Related to Solid-Solid Phase Transitions*, IMA Preprint Series # 1001 (1992), to be published in SIAM J. Numer. Anal..
- [ks] R.V. KOHN & G. STRANG, *Optimal Design and Relaxation of Variational Problems*, Comm. Pure Appl. Math., 39 (1986).
- [lm] M. LUSKIN & LING MA, *Numerical Optimization of the Micromagnetics Energy*, Preprint.
- [mo] C.B. MORREY, *Quasiconvexity and the Semicontinuity of Multiple Integrals*, Pacific J. Math., 2 (1952), pp. 25–53.
- [ri] B.D. RIPLEY, *Thoughts on Pseudorandom Number Generators*, J. Comput. Appl. Math., 31 (1990), pp. 153–163.
- [sc] Z. SCHUSS, *Theory and Applications of Stochastic Differential Equations*, Wiley, 1980.
- [sv] V. ŠVERÁK, *to appear*.

Appendix : a naïve PDE approach to Simulated Annealing

To fix the ideas and for the sake of completeness, we briefly present a sketch proof of convergence of the simulated annealing method. The following approach has been suggested in [gi] to which we refer for further details. We give here only formal arguments.

Let us consider once again the stochastic differential equation

$$(a.1) \quad d\Phi(t) = -\nabla J(\Phi(t))dt + \sqrt{2\vartheta(t)}dw(t), \quad \Phi(0) = \Phi_0 \in \mathbb{R}^p.$$

We refer the reader to [fr] or [sc] for existence and uniqueness results of a Markov process solution to (a.1).

For any Borel set $A \subset \mathbb{R}^p$, we define

$$P(0, \Phi_0, t, A) = \mathbb{P}\{\Phi(t) \in A\},$$

where $\mathbb{P}\{\cdot\}$ is the probability of $\{\cdot\}$ and $P(0, \Phi_0, t, A)$ is the transition probability of $\Phi(t)$.

LEMMA. *If $\vartheta : \mathbb{R} \rightarrow \mathbb{R}^+$ and $\nabla J : \mathbb{R}^p \rightarrow \mathbb{R}^p$ are bounded, continuous, uniformly Lipschitz continuous on compact subsets of $\mathbb{R}$ and $\mathbb{R}^p$ respectively, then*

$$P(0, \Phi_0, t, A) = \int_A \rho(x, t) dx,$$

where the transition probability density $\rho = \rho(x, t)$ satisfies the Fokker-Planck equation

$$(a.2) \quad \begin{aligned} \frac{\partial \rho}{\partial t} - \vartheta(t)\Delta \rho - \nabla \cdot (\nabla J \rho) &= 0, \\ \lim_{t \downarrow 0} \rho(\cdot, t) &= \delta_{\Phi_0}, \end{aligned}$$

where δ_{Φ_0} is the Dirac mass at Φ_0 .

Proof. See e.g. [fr], chap. 6. $\square$

Let us now for the sake of the argument, assume $\vartheta(t) = \vartheta = \text{constant}$. We also consider the linear operator $\mathcal{L}_\vartheta = -\vartheta\Delta - \nabla \cdot (\nabla J \cdot)$. If $\int_{\mathbb{R}^p} \exp(-J(x)/\vartheta) dx = m_\vartheta < \infty$, then $\rho_0(x) = \frac{1}{m_\vartheta} \exp(-J(x)/\vartheta)$ is a density that satisfies $\mathcal{L}_\vartheta \rho_0 = 0$.

We next (formally) write ρ as an eigenfunction expansion

$$(a.3) \quad \rho(x, t) = \rho_0(x) + \sum_{n=1}^{\infty} \rho_n(x) \rho_n(\Phi_0) \exp(-J(x)/\vartheta) e^{-\lambda_n t},$$

where λ_n and ρ_n denote the n th eigenvalue and eigenvector of $\mathcal{L}_\vartheta$, i.e.

$$(a.4) \quad -\vartheta\Delta\rho_n - \nabla \cdot (\nabla J \rho_n) = \lambda_n \rho_n, \quad \int_{\mathbb{R}^p} \rho_n(x) dx = 1.$$

In (a.3), the function ρ_0 is the eigenfunction corresponding to the first eigenvalue $\lambda_0 = 0$. Under reasonable assumptions upon J (see e.g. [gi]), we have $0 < \lambda_1 \leq \lambda_2 \leq \dots$. Consequently, we get

$$\rho(x, t) \xrightarrow[t \rightarrow \infty]{} \rho_0(x).$$

It remains to study the behavior of ρ_0 as $\vartheta \rightarrow 0$. Using the Taylor formula, it is possible to prove, in the case of a suitable multiple well function J for instance, that ρ_0 tends in the distribution sense to a sum of Dirac masses, each of which is concentrated on a global minimizer of J , with a weight proportional to the “width” of the corresponding well (see e.g. [apz]).

The general case $\vartheta = \vartheta(t)$ requires more technicalities and we refer the reader to [gi] and the references quoted therein, where convergence results are obtained for functions ϑ satisfying (3.5).

#	Author/s	Title
1050	J.E. Dunn & Roger Fosdick,	The Weierstrass condition for a special class of elastic materials
1051	Bei Hu & Jianhua Zhang,	A free boundary problem arising in the modeling of internal oxidation of binary alloys
1052	Eduard Feireisl & Enrique Zuazua,	Global attractors for semilinear wave equations with locally distributed nonlinear damping and critical exponent
1053	I-Heng McComb & Chjan C. Lim,	Stability of equilibria for a class of time-reversible, $D_n \times O(2)$ -symmetric homogeneous vector fields
1054	Ruben D. Spies,	A state-space approach to a one-dimensional mathematical model for the dynamics of phase transitions in pseudoelastic materials
1055	H.S. Dumas, F. Golse, and P. Lochak,	Multiphase averaging for generalized flows on manifolds
1056	Bei Hu & Hong-Ming Yin,	Global solutions and quenching to a class of quasilinear parabolic equations
1057	Zhangxin Chen,	Projection finite element methods for semiconductor device equations
1058	Peter Gutterp,	Statistical analysis of biological monitoring data
1059	Wensheng Liu & Héctor J. Sussmann,	Abnormal sub-Riemannian minimizers
1060	Chjan C. Lim,	A combinatorial perturbation method and Arnold's whiskered Tori in vortex dynamics
1061	Yong Liu,	Axially symmetric jet flows arising from high speed fiber coating
1062	Li Qiu & Tongwen Chen,	$\mathcal{H}_2$ and $\mathcal{H}_\infty$ designs of multirate sampled-data systems
1063	Eduardo Casas & Jiongmin Yong,	Maximum principle for state-constrained optimal control problems governed by quasilinear elliptic equations
1064	Suzanne M. Lenhart & Jiongmin Yong,	Optimal control for degenerate parabolic equations with logistic growth
1065	Suzanne Lenhart,	Optimal control of a convective-diffusive fluid problem
1066	Enrique Zuazua,	Weakly nonlinear large time behavior in scalar convection-diffusion equations
1067	Caroline Fabre, Jean-Pierre Puel & Enrike Zuazua,	Approximate controllability of the semilinear heat equation
1068	M. Escobedo, J.L. Vazquez & Enrike Zuazua,	Entropy solutions for diffusion-convection equations with partial diffusivity
1069	M. Escobedo, J.L. Vazquez & Enrike Zuazua,	A diffusion-convection equation in several space dimensions
1070	F. Fagnani & J.C. Willems,	Symmetries of differential systems
1071	Zhangxin Chen, Bernardo Cockburn, Joseph W. Jerome & Chi-Wang Shu,	Mixed-RKDG finite element methods for the 2-D hydrodynamic model for semiconductor device simulation
1072	M.E. Bradley & Suzanne Lenhart,	Bilinear optimal control of a Kirchhoff plate
1073	Héctor J. Sussmann,	A cornucopia of abnormal subriemannian minimizers. Part I: The four-dimensional case
1074	Marek Rakowski,	Transfer function approach to disturbance decoupling problem
1075	Yuncheng You,	Optimal control of Ginzburg-Landau equation for superconductivity
1076	Yuncheng You,	Global dynamics of dissipative modified Korteweg-de Vries equations
1077	Mario Taboada & Yuncheng You,	Nonuniformly attracting inertial manifolds and stabilization of beam equations with structural and Balakrishnan-Taylor damping
1078	Michael Böhm & Mario Taboada,	Global existence and regularity of solutions of the nonlinear string equation
1079	Zhangxin Chen,	BDM mixed methods for a nonlinear elliptic problem
1080	J.J.L. Velázquez,	On the dynamics of a closed thermosyphon
1081	Frédéric Bonnans & Eduardo Casas,	Some stability concepts and their applications in optimal control problems
1082	Hong-Ming Yin,	$\mathcal{L}^{2,\mu}(Q)$ -estimates for parabolic equations and applications
1083	David L. Russell & Bing-Yu Zhang,	Smoothing and decay properties of solutions of the Korteweg-de Vries equation on a periodic domain with point dissipation
1084	J.E. Dunn & K.R. Rajagopal,	Fluids of differential type: Critical review and thermodynamic analysis
1085	Mary Elizabeth Bradley & Mary Ann Horn,	Global stabilization of the von Kármán plate with boundary feedback acting via bending moments only
1086	Mary Ann Horn & Irena Lasiecka,	Global stabilization of a dynamic von Kármán plate with nonlinear boundary feedback
1087	Vilmos Komornik,	Decay estimates for a petrovski system with a nonlinear distributed feedback
1088	Jesse L. Barlow,	Perturbation results for nearly uncoupled Markov chains with applications to iterative methods
1089	Jong-Shenq Guo,	Large time behavior of solutions of a fast diffusion equation with source
1090	Tongwen Chen & Li Qiu,	$\mathcal{H}_\infty$ design of general multirate sampled-data control systems
1091	Satyanad Kichenassamy & Walter Littman,	Blow-up surfaces for nonlinear wave equations, I
1092	Nahum Shimkin,	Asymptotically efficient adaptive strategies in repeated games, Part I: certainty equivalence strategies
1093	Caroline Fabre, Jean-Pierre Puel & Enrique Zuazua,	On the density of the range of the semigroup for semilinear heat equations

- 1094 Robert F. Stengel, Laura R. Ray & Christopher I. Marrison, Probabilistic evaluation of control system robustness
- 1095 H.O. Fattorini & S.S. Sridharan, Optimal chattering controls for viscous flow
- 1096 Kathryn E. Lenz, Properties of certain optimal weighted sensitivity and weighted mixed sensitivity designs
- 1097 Gang Bao & David C. Dobson, Second harmonic generation in nonlinear optical films
- 1098 Avner Friedman & Chaocheng Huang, Diffusion in network
- 1099 Xinfu Chen, Avner Friedman & Tsuyoshi Kimura, Nonstationary filtration in partially saturated porous media
- 1100 Walter Littman & Baisheng Yan, Rellich type decay theorems for equations $P(D)u = f$ with f having support in a cylinder
- 1101 Satyanad Kichenassamy & Walter Littman, Blow-up surfaces for nonlinear wave equations, II
- 1102 Nahum Shimkin, Extremal large deviations in controlled I.I.D. processes with applications to hypothesis testing
- 1103 A. Narain, Interfacial shear modeling and flow predictions for internal flows of pure vapor experiencing film condensation
- 1104 Andrew Teel & Laurent Praly, Global stabilizability and observability imply semi-global stabilizability by output feedback
- 1105 Karen Rudie & Jan C. Willems, The computational complexity of decentralized discrete-event control problems
- 1106 John A. Burns & Ruben D. Spies, A numerical study of parameter sensitivities in Landau-Ginzburg models of phase transitions in shape memory alloys
- 1107 Gang Bao & William W. Symes, Time like trace regularity of the wave equation with a nonsmooth principal part
- 1108 Lawrence Markus, A brief history of control
- 1109 Richard A. Brualdi, Keith L. Chavey & Bryan L. Shader, Bipartite graphs and inverse sign patterns of strong sign-nonsingular matrices
- 1110 A. Kersch, W. Morokoff & A. Schuster, Radiative heat transfer with quasi-monte carlo methods
- 1111 Jianhua Zhang, A free boundary problem arising from swelling-controlled release processes
- 1112 Walter Littman & Stephen Taylor, Local smoothing and energy decay for a semi-infinite beam pinned at several points and applications to boundary control
- 1113 Srdjan Stojanovic & Thomas Svobodny, A free boundary problem for the Stokes equation via nonsmooth analysis
- 1114 Bronislaw Jakubczyk, Filtered differential algebras are complete invariants of static feedback
- 1115 Boris Mordukhovich, Discrete approximations and refined Euler-Lagrange conditions for nonconvex differential inclusions
- 1116 Bei Hu & Hong-Ming Yin, The profile near blowup time for solution of the heat equation with a nonlinear boundary condition
- 1117 Jin Ma & Jiongmin Yong, Solvability of forward-backward SDEs and the nodal set of Hamilton-Jacobi-Bellman Equations
- 1118 Chaocheng Huang & Jiongmin Yong, Coupled parabolic and hyperbolic equations modeling age-dependent epidemic dynamics with nonlinear diffusion
- 1119 Jiongmin Yong, Necessary conditions for minimax control problems of second order elliptic partial differential equations
- 1120 Eitan Altman & Nahum Shimkin, Worst-case and Nash routing policies in parallel queues with uncertain service allocations
- 1121 Nahum Shimkin & Adam Shwartz, Asymptotically efficient adaptive strategies in repeated games, part II: Asymptotic optimality
- 1122 M.E. Bradley, Well-posedness and regularity results for a dynamic Von Kármán plate
- 1123 Zhangxin Chen, Finite element analysis of the 1D full drift diffusion semiconductor model
- 1124 Gang Bao & David C. Dobson, Diffractive optics in nonlinear media with periodic structure
- 1125 Steven Cox & Enrique Zuazua, The rate at which energy decays in a damped string
- 1126 Anthony W. Leung, Optimal control for nonlinear systems of partial differential equations related to ecology
- 1127 H.J. Sussmann, A continuation method for nonholonomic path-finding problems
- 1128 Yung-Jen Guo & Walter Littman, The null boundary controllability for semilinear heat equations
- 1129 Q. Zhang & G. Yin, Turnpike sets in stochastic manufacturing systems with finite time horizon
- 1130 I. Györi, F. Hartung & J. Turi, Approximation of functional differential equations with time- and state-dependent delays by equations with piecewise constant arguments
- 1131 I. Györi, F. Hartung & J. Turi, Stability in delay equations with perturbed time lags
- 1132 F. Hartung & J. Turi, On the asymptotic behavior of the solutions of a state-dependent delay equation
- 1133 Pierre-Alain Gremaud, Numerical optimization and quasiconvexity
- 1134 Jie Tai Yu, Resultants and inversion formula for N polynomials in N variables
- 1135 Avner Friedman & J.L. Velázquez, The analysis of coating flows in a strip
- 1136 Eduardo D. Sontag, Control of systems without drift via generic loops
- 1137 Yuan Wang & Eduardo D. Sontag, Orders of input/output differential equations and state space dimensions
- 1138 Scott W. Hansen, Boundary control of a one-dimensional, linear, thermoelastic rod