Rank-one convex functions on 2×2 symmetric matrices and laminates on rank-three lines

S. Conti, D. Faraco, F. Maggi, and S. Müller

Max-Planck-Institute for Mathematics in the Sciences,
Inselstr. 22-26, 04103 Leipzig, Germany

August 13, 2004

We construct a function on the space of symmetric 2×2 matrices in such a way that it is convex on rank-one directions and its distributional Hessian is not a locally bounded measure. This paper is also an illustration of a recently proposed technique to disprove L^1 estimates by the construction of suitable probability measures (laminates) in matrix space. From this point of view the novelty is that the support of the laminate, besides satisfying a convex constraint, needs to be contained on a rank-three line, up to arbitrarily small errors.

1 Motivation and main result

Rank-one convex functions arise naturally in the study of variational problems on vector-valued maps, especially in connection with the fundamental notion of quasiconvexity. Nevertheless, very little is known on their general properties, either of global or local character. Even simple questions are still open, such as whether the Hessian of a rank-one convex function, that is always defined as a distribution, is a locally bounded measure. This statement was recently disproved for the model case of rank-one convex functions on diagonal matrices, i.e. for separately convex functions [3]. We show here that the same happens in the case of rank-one convex functions on symmetric matrices. The general case remains open.

As it is known since the work of Tartar on oscillations in nonlinear PDEs [19] and from the theory of convex integration by Gromov [8], the existence of solutions to differential inclusions of the form

$$\nabla u \in E \subset \mathbb{R}^{n \times m}, \quad u : \mathbb{R}^n \to \mathbb{R}^m, \tag{1}$$

can be shown by analyzing certain geometric properties of the set E, related to the presence in E of *rank-one connections*, i.e., of pairs of matrices $A, B \in E$ such that the difference $A - B$ has rank one. A general theory for the existence of Lipschitz solutions to differential inclusions of the above form has been developed by Müller and Šverák [15, 16], Dacorogna and Marcellini [5], Kirchheim [10, 11], and Sychev [18]. Prototypical applications of this theory are the construction of nowhere differentiable solutions to smooth elliptic systems [16, 9] and of functions whose gradients take only a finite number of non compatible (non rank-one connected) values [10, 11, 2]. These examples focus on Lipschitz solutions, and thus on the "local" rank-one geometry of E. We study rank-one convex functions with unbounded Hessian, which are also solutions to a certain differential inclusion. At variance with the previous example, we shall be interested in the study of rank-one geometry "at infinity".

A method to construct solutions u to (1) exploiting the rank-one geometry of a set of matrices E is to consider certain probability measures supported on E, the so-called *laminates*, which can be approximated by sequences of gradients. Our approach is based on laminates which include very large matrices (and which are unbounded in the limit). Analogous *unbounded laminates* have been used in [6, 7] for constructing critical solutions to elliptic equations, and in [3] to give a rank-one convex function on diagonal matrices with locally unbounded Hessian, as well as to obtain a clearer derivation of a classical counterexample to Korn's inequality in L^1 by Ornstein [17]. Here we develop the technique further, see Section 3.

A function $f : \mathbb{R}^{2\times 2} \to \mathbb{R}$ is rank-one convex if it is convex along rank-one lines, i.e.

$$f(F + ta \otimes b) \quad \text{is convex as a function of } t \in \mathbb{R} \tag{2}$$

whenever $F \in \mathbb{R}^{2\times 2}$ and $a, b \in \mathbb{R}^2$. Rank-one convex functions are automatically locally Lipschitz (see Marcellini [12, page 6] and Dacorogna [4, Th. 4.1.1]), hence the Hessian of f is defined as a vector-valued distribution $\nabla^2 f$ on $\mathbb{R}^{2\times 2}$. For a convex f, by monotonicity the distributional Hessian actually defines a locally bounded measure. We are interested in the possible validity of this property for rank-one convex functions:

Q: Is the Hessian of a rank-one convex function a locally bounded measure?

We do not know the answer to this question in general, and address instead simplified models of rank-one convexity, in which the dimension of the underlying space is reduced. More specifically, let V be a linear subspace of $\mathbb{R}^{2\times 2}$, and consider functions $f : V \to \mathbb{R}$ which are rank-one convex on V, in the sense that (2) holds for every $F \in V$ and $a, b \in \mathbb{R}^2$ with $a \otimes b \in V$. If V is spanned by rank-one matrices (as it is the case whenever the dimension of $V \subset \mathbb{R}^{2\times 2}$ is at least three) all rank-one convex functions on V are locally Lipschitz and thus have a well-defined distributional Hessian. In particular the above question makes sense on every such V. We remark that being able to answer negatively to question Q on a proper subspace V *does not* allow to do the same on $\mathbb{R}^{2\times 2}$. Indeed, we are unaware of any general method to extend a rank-one convex function on V having locally unbounded Hessian to a rank-one convex function on $\mathbb{R}^{2\times 2}$, and even if this would be possible the implication would not be immediate. Strictly rank-one convex or quasiconvex functions with bounded Hessian can instead be extended, see [21, 20, 16, 13].

When V is two-dimensional there are three distinct cases: (i) all the directions in V are of rank one, (ii) only two of them are, (iii) one or less are. The first case corresponds to convex functions on the plane, whose Hessian is locally a bounded measure. In the second case, up to a linear isomorphism which preserves rank-one convexity, we can assume that V is the space of diagonal matrices. Rank-one convexity is then equivalent to separate convexity on the plane, and a counterexample was obtained in [3]. The third case is degenerate and trivial.

Three-dimensional subspaces of $\mathbb{R}^{2\times 2}$ can be classified on the basis of whether they are orthogonal to a rank-one line or not (see the proof of Corollary 6). In the first case, up to linear isomorphisms preserving rank-one convexity, we can assume that V is the space of upper triangular matrices. The example on diagonal matrices applies also to this case (extending the function so that it does not depend on the off-diagonal entry). The second case can be reduced to the situation where V is the set of 2×2 symmetric matrices, i.e.

$$V = \mathbb{R}^{2\times 2}_{\mathrm{sym}}.$$

Here the rank-one cone is the set of matrices of the form

$$\{\pm a \otimes a : a \in \mathbb{R}^2\},$$

which is curved (whereas in the upper-triangular case one gets two planes). Our main result is the following.

Theorem 1. *There exists a rank-one convex function on symmetric 2×2 matrices whose distributional Hessian is locally unbounded on the unit ball, and that agrees with a smooth convex function elsewhere.*

The rest of the paper is organized as follows: in Section 2 we explain how Theorem 1 can be reduced to prove the existence of a sequence of laminates with certain critical properties, while in Section 3 we perform the construction.

2 Laminates and differential inclusions

Laminates are certain probability measures on matrix spaces, which are defined in purely geometric terms and can be easily approximated by sequences of gradients (in the sense that the distribution function of the gradients converges to the laminate, see (3) below). We first define a laminate by induction. Let three matrices A, B and C be given in such a way that $A = \lambda B + (1-\lambda)C$ for $\lambda \in [0,1]$. If the rank of $B - C$ is equal to one we say that we can *split* the Dirac measure δ_A into the atomic probability measure $\lambda\delta_B + (1-\lambda)\delta_C$, and write

$$\delta_A \rightsquigarrow \lambda\delta_B + (1-\lambda)\delta_C.$$

Both measures have center of mass A. We can iterate the procedure by taking a rank-one segment through B or C. The result is again a probability measure with center of mass A, as for example

$$\lambda\delta_B + (1-\lambda)\delta_C \rightsquigarrow \lambda\delta_B + (1-\lambda)\mu\,\delta_{C_1} + (1-\lambda)(1-\mu)\delta_{C_2},$$

where $C = \mu C_1 + (1-\mu)C_2$, $C_2 - C_1$ is a rank-one matrix, and $\mu \in [0,1]$. We call a *laminate of finite order on E* every probability measure that can be obtained by a finite number of splitting steps as the ones we have just described and which is supported on E. The *laminates on E* are just the weak-$*$ limits of sequences of laminates of finite order.

Given two atomic probability measures μ and ν on $\mathbb{R}^{2\times 2}$ we shall write $\nu \rightsquigarrow \mu$ if μ can be obtained from ν by replacing a Dirac mass in ν with a laminate of finite order. When $\nu \rightsquigarrow \mu$ then μ has the same center of mass of ν, and if ν is a laminate of finite order then so is μ.

Laminates on E are models for certain solutions to the differential inclusion $\nabla u \in E$. Indeed, for every laminate ν on E with center of mass at $A \in E$ and compact support, and for every open set $\Omega \subset \mathbb{R}^2$, we can find a sequence of functions $u_h \in W^{1,\infty}(\Omega; \mathbb{R}^2)$ such that

$$\lim_{h\to\infty} \int_\Omega \phi(\nabla u_h(x))dx = \int \phi(F)d\nu(F), \quad \forall \phi \in C^\infty(\mathbb{R}^{2\times 2}), \tag{3}$$

with

$$u_h(x) = Ax \quad \text{on } \partial\Omega\,, \qquad \text{and} \qquad \lim_{h\to\infty} \sup_{x\in\Omega} \operatorname{dist}(\nabla u_h(x), E^{**}) = 0 \tag{4}$$

(here E^{**} is the convex hull of E), see for example [14].

We now relate the notion of laminate to rank-one convexity on symmetric matrices. A function $f : \mathbb{R}^{2\times 2}_{\text{sym}} \to \mathbb{R}$ is rank-one convex if and only if

$$\nabla^2 f(F)(a\otimes a, a\otimes a) \geq 0 \tag{5}$$

for all $F \in \mathbb{R}^{2\times 2}_{\text{sym}}$ and $a \in \mathbb{R}^2$. This one parameter family of linear inequalities defines a six dimensional convex cone E_{sym} of quadratic forms on $\mathbb{R}^{2\times 2}_{\text{sym}}$ (a more detailed description of this cone is given in Section 3).

We now rephrase our main result in terms of laminates. In the next section we shall prove Theorem 2 and derive Theorem 1 from it.

Theorem 2. *There exists a sequence of laminates ν_k supported on the set E_{sym} with a fixed center of mass and such that*

$$\int |F| d\nu_k(F) \to \infty.$$

3 Construction of the critical laminate

We shall from now on identify $\mathbb{R}^{2\times 2}_{\text{sym}}$ with $\mathbb{R}^3$ according to

$$(\alpha, \beta, \gamma) \in \mathbb{R}^3 \mapsto \begin{pmatrix} \gamma+\beta & \alpha \\ \alpha & \gamma-\beta \end{pmatrix} \in \mathbb{R}^{2\times 2}_{\text{sym}}.$$

The rank-one cone of $\mathbb{R}^{2\times 2}_{\text{sym}}$, namely $\{\pm a\otimes a : a \in \mathbb{R}^2\}$, takes the form

$$H = \{(\alpha, \beta, \gamma) \in \mathbb{R}^3 : \alpha^2 + \beta^2 = \gamma^2\}$$

A rank-one convex function f on $\mathbb{R}^{2\times 2}_{\text{sym}}$ is therefore a function which is convex along all lines in H, i.e., it is such that

$$\nabla^2 f \in E_{\text{sym}} = \left\{ F \in \mathbb{R}^{3\times 3}_{\text{sym}} : (Fv, v) \geq 0 \text{ for all } v \in H \right\}.$$

The set E_{sym} is a convex cone. To prove Theorem 2 we have to construct a sequence of laminates ν_k in E_{sym} with a fixed center of mass and such that the sequence of first order moments is diverging. Loosely speaking, the only way to achieve these two requirements at the same time is to perform wider and wider oscillations at least along a line entirely contained in E_{sym}. Up to translations, there is only one line contained in E_{sym}:

$$e = \left\{ \begin{pmatrix} t & 0 & 0 \\ 0 & t & 0 \\ 0 & 0 & -t \end{pmatrix} : t \in \mathbb{R} \right\}.$$

This line corresponds to the Hessians of the quadratic forms which are multiples of the determinant, which is affine on rank-one lines. It is the only line contained in E_{sym}, since quadratic rank-one-convex functions are the sum of a convex function and a multiple of the determinant.

The line e is composed by rank-three matrices, hence there is no nontrivial laminate of finite order supported on e (indeed, every laminate of finite order which is not a single Dirac mass contains at least a pair of Dirac masses whose supports are rank-one connected). Further, no two points of e can be joined by less than three rank-one segments (since the rank of a matrix is subadditive). It is therefore natural to seek a three-dimensional subspace W of $\mathbb{R}^{3\times 3}_{\mathrm{sym}}$ spanned by three rank-one directions which contains e, and to try to construct the laminate on $E_{\mathrm{sym}} \cap W$. We take the set of diagonal matrices as W, so that $E = E_{\mathrm{sym}} \cap W$ takes the form

$$E = \left\{ \begin{pmatrix} x & 0 & 0 \\ 0 & y & 0 \\ 0 & 0 & z \end{pmatrix} : x, y, z \in \mathbb{R},\ x + z \geq 0,\ y + z \geq 0 \right\} \subset E_{\mathrm{sym}} \subset \mathbb{R}^{3\times 3}_{\mathrm{sym}}.$$

The only rank-one directions in this cone are given by the coordinate axes x, y, z.

From now on we focus only on E, and denote its generic element by (x, y, z). We are going to construct a sequence of finite-order laminates ν_k in E with center of mass at $(0, 0, 1)$ and such that

$$\int |(x, y, z)| d\nu_k(x, y, z) \to \infty.$$

The sequence ν_k will be constructed by iterative splitting, starting from $\delta_{(0,0,1)}$. Each splitting corresponds to taking a convex combination in which only one variable between x, y and z is changed.

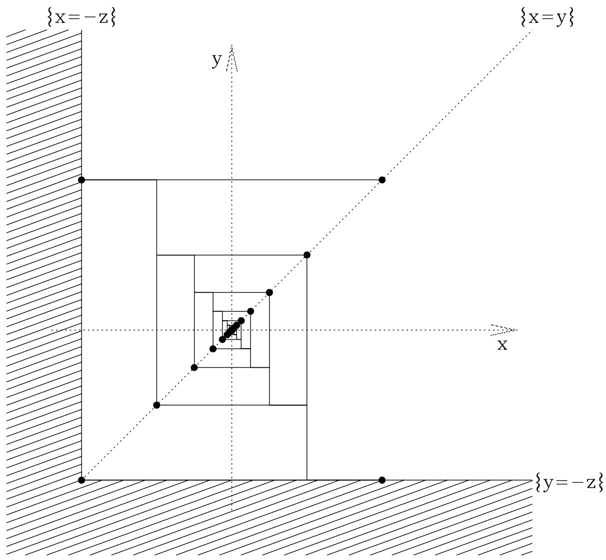

FIGURE 1: Splitting procedure at constant z in the spreading step (see Lemma 3). The dashed area represents the region outside E. Starting from the point in the origin, most of the mass is spread along the $x = y$ line, with two small errors in $(z, -z, z)$ and $(-z, z, z)$.

Before presenting the proof we sketch the main ideas of the three construction steps.

Spreading. We start by splitting $\delta_{(0,0,1)}$ inside the plane $\{z = 1\}$ by means of the rank-one directions x and y. Our goal is to obtain an oscillation along the rank-two line $\{x = y,\, z = 1\}$, without leaving the set E. The situation is here in many respects analogous to the one discussed in [3] for the case of separate convexity, and we use an adaptation of that argument (see Figure 1). We obtain a splitting of the form

$$\delta_{(0,0,1)} \rightsquigarrow \sum_i \lambda_i \delta_{(a_i,a_i,1)} + \text{error}, \quad \text{with} \quad \sum_i \lambda_i = 1 - \varepsilon \quad \text{and } |a_i| \leq 1.$$

Up to the error terms, which are supported on $(1, -1, 1)$ and $(-1, 1, 1)$ and have total mass ε, this measure is supported on the rank-two line $\{x = y,\, z = 1\}$, and has an L^1-oscillation (in the sense of the average value of $|x|$) which scales as

$\sum |a_i|\lambda_i \simeq \varepsilon \ln 1/\varepsilon$, i.e. most of the oscillation is concentrated on a rank-two line. The main difference with the construction in [3] is that instead of iterating to infinity we perform the spreading on a small scale, stopping after finitely many steps with a finite error term on the boundary of E, and then (see below) repeat it at different z-levels. This spreading step is the main new technical point in the construction of laminates that we present here, and is discussed in Lemma 3.

Vertical splitting. We now project the oscillation onto e via a splitting in the z direction, of the form

$$\delta_{(a_i,a_i,1)} \rightsquigarrow \mu_i \delta_{(a_i,a_i,-a_i)} + (1-\mu_i)\delta_{(a_i,a_i,K)},$$

where $K > 1$ is a parameter to be fixed. We have therefore

$$\delta_{(0,0,1)} \rightsquigarrow \nu_1 = \sum_i \lambda_i \mu_i \delta_{(a_i,a_i,-a_i)} + \sum_i \lambda_i (1-\mu_i)\delta_{(a_i,a_i,K)} + \text{ error}$$

The first term in ν_1 is supported on the rank-three line e, while the second term is supported on $\{x = y,\, |x| \leq 1,\, z = K\}$. This segment corresponds to the starting segment translated at an height K, and will be the starting point for the iteration. For large K the mass left in this segment scales as $1/K$, but its distance from the boundary of E, and hence the size of the oscillation that will be generated at the next step, scales as K. We choose $K = 1/\varepsilon$. This construction, which can be done with $\delta_{(b,b,1)}$ instead of $\delta_{(0,0,1)}$, is presented in Lemma 4 and illustrated in Figure 2.

Iteration. The third ingredient of the construction is to iterate the process with a careful choice of the parameters at each step. The idea is to apply the spreading step to *each* point charged by ν_1 on the rank-two segment $\{x = y,\, |x| \leq 1,\, z = 1/\varepsilon\}$. These points have total mass of order $\varepsilon(1-\varepsilon)$ and are at a distance of order $1/\varepsilon$ from the boundary of E, therefore the next spreading step produces an oscillation of width $1/\varepsilon$, which compensates for the small mass. This procedure is presented in Lemma 5, and the value of ε at each step is chosen in the proof of Theorem 2.

We start from the spreading lemma that, for a better understanding, is stated on $\mathbb{R}^{n\times m}$. For brevity we write

$$\delta_{\pm F} = \delta_F + \delta_{-F} \qquad \text{and} \qquad \lambda(\pm t)\delta_{F(\pm t)} = \lambda(t)\delta_{F(t)} + \lambda(-t)\delta_{F(-t)}\,.$$

Further, in the following summations over empty sets are intended to be zero, and products to be one. For example, $\sum_{i=2}^n i$ is zero for all $n < 2$.

Lemma 3 (Spreading step). *Let $A, B \in \mathbb{R}^{n\times m}$ be two rank-one matrices, and $\varepsilon > 0$. Then for every natural number $N \geq 1$ the measure*

$$\begin{aligned}\nu_N \;=\; & \frac{1}{3}\,\delta_{\pm\varepsilon\,(A+B)} + \frac{1}{6}\sum_{n=1}^{N-1}\frac{1}{2^n}\,\delta_{\pm 2^n\,\varepsilon\,(A+B)} + \frac{1}{12}\,\frac{1}{2^N}\,\delta_{\pm 2^N\,\varepsilon\,(A+B)} \\ & + \frac{1}{2^{N+2}}\,\delta_{\pm 2^N\,\varepsilon\,(A-B)}\end{aligned}$$

is a laminate with center of mass equal to 0 and support in the parallelogram of vertices $\{\pm 2^N\varepsilon(A+B), \pm 2^N\varepsilon(A-B)\}$.

Proof. We denote by (α,β) the matrix $F = \alpha A + \beta B$. The lemma is proven by induction, starting from the inductive step.

We assume that the statement holds for some $N \geq 1$, and prove it for $N+1$. This is done by the splitting

$$\delta_{(t,-t)} \rightsquigarrow \frac{1}{3}\delta_{(t,t)} + \frac{2}{3}\delta_{(t,-2t)} \rightsquigarrow \frac{1}{3}\delta_{(t,t)} + \frac{2}{3}\left\{\frac{1}{4}\delta_{(-2t,-2t)} + \frac{3}{4}\delta_{(2t,-2t)}\right\}. \tag{6}$$

Since ν_N is a laminate with center of mass at $(0,0)$, the measure that is obtained from ν_N using on its last term the splitting above with $t = \pm 2^N\varepsilon$ is a laminate. A direct calculation shows that the result is ν_{N+1}.

Now we consider the case $N = 1$. The splitting

$$\delta_{(0,0)} \rightsquigarrow \frac{1}{2}\delta_{(\varepsilon,0)} + \frac{1}{2}\delta_{(-\varepsilon,0)} \rightsquigarrow \frac{1}{2}\left\{\frac{1}{2}\delta_{(\varepsilon,\varepsilon)} + \frac{1}{2}\delta_{(\varepsilon,-\varepsilon)}\right\} + \frac{1}{2}\left\{\frac{1}{2}\delta_{(-\varepsilon,\varepsilon)} + \frac{1}{2}\delta_{(-\varepsilon,-\varepsilon)}\right\}$$

shows that

$$\delta_{(0,0)} \rightsquigarrow \nu_0 = \frac{1}{4}\delta_{\pm(\varepsilon,\varepsilon)} + \frac{1}{4}\delta_{\pm(\varepsilon,-\varepsilon)}.$$

The splitting (6) with $t = \pm\varepsilon$ applied to the last term gives ν_1. This proves the statement for $N = 1$ and hence concludes the proof. □

We now come back to the (x, y, z) notation introduced before Lemma 3.

Lemma 4 (Vertical splitting step). *Let $0 < \varepsilon \leq 1/16$, $a \in (-1/2, 1/2)$, and $Z > 0$. Then there exists a probability measure $\mu = \mu_{\varepsilon,a,Z}$ supported on E such that*

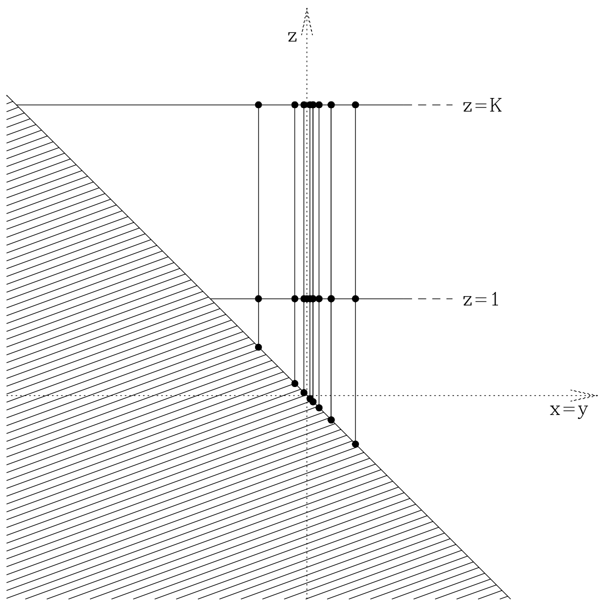

FIGURE 2: Splitting along z at constant x and y ($Z = 1$ in Lemma 4).

$\delta_{(aZ,aZ,Z)} \rightsquigarrow \mu = \mu^1 + \mu^2$, with $\mu^1 \geq 0$, $\mu^2 \geq 0$,

$$\begin{aligned} \int |(x,y,z)| d\mu^1(x,y,z) &\geq \frac{Z}{3}\varepsilon \ln \frac{1}{\varepsilon} && (7) \\ \mu^2(\mathbb{R}^3) &\geq \varepsilon\,(1+a)(1-3\varepsilon) && (8) \\ \operatorname{spt}\mu^2 &\subset \{x = y,\, |x| \leq Z,\, z = Z/\varepsilon\} && (9) \\ \int (x - aZ) d\mu^2 &\geq 0\,, && (10) \end{aligned}$$

and for fixed (ε, Z) the function $a \to \mu^2(\mathbb{R}^3)$ is increasing.

Proof. By scaling we can assume without loss of generality $Z = 1$. We adopt the notation

$$F(b,z) = (b,b,z) \qquad \text{and} \qquad G(b) = (b,b,-b)\,.$$

We apply Lemma 3 with respect to the directions $A = (1,0,0)$ and $B = (0,1,0)$, with ε and a as in the hypothesis and

$$N = \left\lfloor \frac{\log(1/\varepsilon)}{\log 2} \right\rfloor - 1 \,,$$

where $\lfloor x \rfloor$ denotes the greatest integer no larger than x. This implies

$$\frac{1}{4} < 2^N \varepsilon \leq \frac{1}{2}.$$

We translate the resulting laminate by the vector $F(a,1) = (a,a,1)$, and denote the result by μ_N. We obtain

$$\delta_{F(a,1)} \rightsquigarrow \mu_N = \frac{1}{3}\,\delta_{F(a\pm\varepsilon,1)} + \frac{1}{6}\sum_{n=1}^{N-1} \frac{1}{2^n}\,\delta_{F(a\pm 2^n\varepsilon,1)} + \frac{1}{12}\,\frac{1}{2^N}\,\delta_{F(a\pm 2^N\varepsilon,1)} + \mu^3 \,,$$

where

$$\mu^3 = \frac{1}{2^{N+2}}\,\delta_{F(a,1)\,\pm\,2^N\varepsilon\,(A-B)} \,.$$

Note that $\operatorname{spt}(\mu_N - \mu^3) \subset \{x = y,\ |x| \leq 1,\ z = 1\} \subset E$ and $\operatorname{spt}\mu^3 \subset E$.

Fix $K > 1$. By splitting in the z direction each matrix $(b,b,1)$ (with $b > -1$) into $(b,b,-b)$ and (b,b,K) we see that

$$\delta_{F(b,1)} \rightsquigarrow \frac{K-1}{K+b}\,\delta_{G(b)} + \frac{1+b}{K+b}\,\delta_{F(b,K)}. \tag{11}$$

The laminate μ is obtained by applying this splitting to each matrix in the support of $\mu_N - \mu^3$ (see Figure 2). We denote by μ^2 the sum of all terms of the form $F(b,K)$, and by μ^1 the sum of μ^3 with all terms of the form $G(b)$. Up to this point we have shown that

$$\delta_{(a,a,1)} \rightsquigarrow \mu = \mu^1 + \mu^2 \,,$$

and choosing $K = 1/\varepsilon$ the condition (9) on the support of μ^2 is also immediate.

It only remains to check the two quantitative estimates. We start with (8). To see this it is convenient to write μ_N shortly as

$$\mu_N = \sum_i w_i \delta_{F(b_i,1)} + \mu^3 \,, \qquad \{b_i\} = \{a \pm 2^n\varepsilon\} \,.$$

Since $\int(x-a)d\mu^3 = \int(x-a)d\delta_{(a,a,1)} = 0$, it follows that

$$0 = \int(x-a)d\mu_N = \sum_i w_i(b_i - a)\,, \quad \text{i.e.} \quad \sum_i b_i w_i = a\sum_i w_i\,.$$

Further, since the total mass of μ^3 is $2^{-(N+1)}$, we get $\sum_i w_i = 1 - 2^{-(N+1)} \geq 1 - 2\varepsilon$. The measures μ^1 and μ^2 are obtained applying the splitting (11) to each term of μ_N. We have

$$\mu^2 = \sum_i \frac{1+b_i}{K+b_i} w_i \delta_{F(b_i,1)}\,,$$

and the total mass is given by

$$\mu^2(\mathbb{R}^3) = \sum_i \frac{1+b_i}{K+b_i} w_i \geq \sum_i \frac{1+b_i}{K+1} w_i = \frac{1+a}{K+1}\sum_i w_i \geq \frac{1+a}{K+1}(1-2\varepsilon)\,.$$

Since $K = 1/\varepsilon$ we get (8). Further, to get the monotonicity in a simply observe that changing a with fixed ε and Z amounts to a translation of all the b_i's, with K and w_i fixed. Therefore the result follows from the fact that $(1+b)/(K+b)$ is increasing in b.

Analogously we derive (10). Indeed,

$$\int(x-a)d\mu^2 = \sum_i (b_i - a)\frac{1+b_i}{K+b_i} w_i = \frac{1}{2}\sum_i w_i\left(\eta_i \frac{1+a+\eta_i}{K+a+\eta_i} - \eta_i \frac{1+a-\eta_i}{K+a-\eta_i}\right)$$

where $\eta_i = b_i - a$, and in the second step we used the symmetry in μ_N to average over the two terms with $\pm\eta_i$. A simple expansion shows that the last parenthesis is nonnegative for each i, and concludes the proof of (10).

It only remains to check (7). Since $|a+t| + |a-t| \geq 2|t|$, we have

$$\int |x| d\mu_N \geq \frac{1}{6}\sum_{n=0}^{N-1} \frac{|a+2^n\varepsilon| + |a-2^n\varepsilon|}{2^n} \geq \frac{1}{3}\varepsilon N \geq \frac{1}{5}\varepsilon \ln\frac{1}{\varepsilon}\,.$$

The vertical splitting (11) is done at constant x, and for each term a fraction $(K-1)/(K+b_i) \geq (K-1)/(K+1) \geq 5/6$ of the weight goes into μ^1. Therefore

$$\int |x| d\mu^1 \geq \frac{5}{6}\int |x| d\mu_N \geq \frac{1}{6}\varepsilon \ln\frac{1}{\varepsilon}\,,$$

and the same for $\int |y| d\mu^1$. This concludes the proof. □

We come to the third and last preparing step for the proof of Theorem 2.

Lemma 5 (Iteration). *Let* $\{\varepsilon_k\}_{k=1}^{\infty}$ *be a sequence decreasing to zero, with* $\varepsilon_1 \leq 1/16$*. Then, for every* $k \geq 1$ *there exists a finite-order laminate* ν_k *with support in* E *and center of mass at* $(0,0,1)$ *such that* $\nu_k = \nu_k^1 + C_k \nu_k^2$*, where* $\nu_k^1 \geq 0$*,* $\nu_k^2 \geq 0$*,* $\nu_k^2(\mathbb{R}^3) = 1$*,*

$$\begin{aligned}
\int |(x,y,z)| d\nu_k^1(x,y,z) &\geq \frac{1}{3} \sum_{i=2}^{k} \left\{ \prod_{j=1}^{i-1} (1-3\varepsilon_j) \right\} \varepsilon_i \log \frac{1}{\varepsilon_i}, \\
\operatorname{spt} \nu_k^2 &\subset \left\{ x = y,\, |x| \leq \prod_{j=1}^{k-1} \varepsilon_j^{-1},\, z = \prod_{j=1}^{k} \varepsilon_j^{-1} \right\}, \\
C_k &\geq \prod_{j=1}^{k} \varepsilon_j (1-3\varepsilon_j)\,, \\
\int x \, d\nu_k^2 &\geq 0\,.
\end{aligned}$$

Proof. The proof is by induction. We first show that if the thesis is true for a certain $k \geq 1$ then it is true also for $k+1$. Let $\nu_k = \nu_k^1 + C_k \nu_k^2$ be the measure given by the induction hypothesis. We write

$$\nu_k^2 = \sum_i \lambda_i \delta_{F(b_i, Z_k)}, \quad \text{where} \quad Z_k = \prod_{j=1}^{k} \varepsilon_j^{-1}\,, \quad |b_i| \leq Z_{k-1} = \varepsilon_k Z_k, \quad \sum_i \lambda_i = 1.$$

We split each $\delta_{F(b_i,Z_k)}$ according to Lemma 4, with $\varepsilon = \varepsilon_{k+1}$ and $a_i = b_i/Z_k$. We get

$$\delta_{F(b_i,Z_k)} \rightsquigarrow \mu_i^1 + \mu_i^2$$

where μ_i^1 and μ_i^2 are as in the statement of Lemma 4. We then define

$$\nu_{k+1}^1 = \nu_k^1 + C_k \sum_i \lambda_i \mu_i^1\,, \qquad \nu_{k+1}^2 = \frac{C_k}{C_{k+1}} \sum_i \lambda_i \mu_i^2\,, \qquad C_{k+1} = C_k \sum_i \lambda_i \mu_i^2(\mathbb{R}^3)\,.$$

It remains to check that $\nu_{k+1} = \nu_{k+1}^1 + C_{k+1}\nu_{k+1}^2$ has the claimed properties. The first one follows directly from (7) and the inductive hypothesis; the second follows from (9) and the definition of Z_k.

In order to prove the bound on C_{k+1}, we start from (8) and compute

$$C_{k+1} = C_k \sum_i \lambda_i \mu_i^2(\mathbb{R}^3) \geq C_k \sum_i \lambda_i \varepsilon_{k+1} \left(1 + \frac{b_i}{Z_k}\right)(1 - 3\varepsilon_{k+1}) .$$

Since the average of x over ν_k^2 is positive we get $\sum_i \lambda_i b_i \geq 0$, and therefore

$$C_{k+1} \geq C_k \sum_i \lambda_i \varepsilon_{k+1}(1 - 3\varepsilon_{k+1}) \geq C_k \varepsilon_{k+1}(1 - 3\varepsilon_{k+1}) .$$

The thesis follows by the inductive hypothesis on C_k.

Finally, we check that

$$\int x \, d\nu_{k+1}^2 = \frac{C_k}{C_{k+1}} \sum_i \lambda_i \int (x - b_i + b_i) \, d\mu_i^2 \geq \frac{C_k}{C_{k+1}} \sum_i \lambda_i b_i \mu_i^2(\mathbb{R}^3) \geq 0 .$$

In the last step we used that the μ_i^2 are obtained from Lemma 4 with the same ε and $a_i = b_i/Z_k$. Therefore all $\mu_i^2(\mathbb{R}^3)$ which correspond to positive b_i are larger than all those corresponding to negative b_i. The conclusion follows from $\sum \lambda_i b_i \geq 0$.

The case $k = 1$ follows by application of Lemma 4 to $a = 0$, $Z = 1$, $\varepsilon = \varepsilon_1$. □

Proof of Theorem 2. By Lemma 5, it suffices to show that there exists a strictly decreasing sequence $\omega_k \to 0$ such that

$$\omega = \sum_{k=2}^{\infty} \left\{ \prod_{j=1}^{k-1} (1 - \omega_j) \right\} \omega_k \log \frac{1}{\omega_k} = \infty.$$

We define ω_k as

$$\omega_k = \min \left\{ \frac{1}{16}, \frac{1}{k \log(k+1)} \right\} .$$

We notice that

$$\begin{aligned} \prod_{j=1}^{k-1} (1 - \omega_j) &= \exp\left(\sum_{j=1}^{k-1} \log(1 - \omega_j) \right) \geq \exp\left(-\sum_{j=1}^{k-1} (\omega_j + c\omega_j^2) \right) \\ &\geq \exp\left(-\sum_{j=1}^{k-1} \omega_j \right) \exp\left(-c \sum_{j=1}^{\infty} \omega_j^2 \right) \geq C \exp\left(-\sum_{j=1}^{k-1} \omega_j \right) \end{aligned}$$

as

$$\sum_{j=1}^{\infty} \omega_j^2 \approx \int_2^{\infty} \frac{dx}{x^2 \log^2 x} < \infty.$$

On the other hand

$$\sum_{j=1}^{k-1} \omega_j \approx \int_2^{k} \frac{dx}{x \log x} = \log \log k - c.$$

Therefore for large k

$$\prod_{j=1}^{k-1} (1-\omega_j) \geq \frac{C}{\log k}\,,$$

and

$$\omega \;\; \geq \;\; C \sum_{k=3}^{\infty} \frac{1}{\log k} \frac{1}{k \log k} \log\left(k \log k\right) \geq C \sum_{k=2}^{\infty} \frac{1}{k \log k} = \infty,$$

as $\log(k \log k) \geq \log k$ for large k. □

Proof of Theorem 1. Theorem 1 follows from Theorem 2 by means of a standard construction, which is explained e.g. in the proof of Theorem 2 of [3]. For the convenience of the reader we sketch here the main steps. Let B_j be a sequence of disjoint balls contained in the unit ball, and A denote the common center of mass of the measures ν_k. For a fixed j and k, we find $u_{j,k,h}$ satisfying (3) and (4) with $\Omega = B_j$ and $\nu = \nu_k$. Since ν_k is supported on symmetric matrices, the sequence $u_{j,k,h}$ can be chosen so that $u_{j,k,h} = \nabla f_{j,k,h}$, and $f_{j,k,h}(x) = \frac{1}{2} Ax \cdot x$ on ∂B_j. We first choose h so that

$$\int_{B_j} |\nabla^2 f_{j,k,h}|\, dx \geq \frac{1}{2} |B_j| \int |F| d\nu_k(F)\,, \qquad \mathrm{dist}(\nabla u_{j,k,h}, E) \leq 1\,.$$

For each j we choose k so that

$$|B_j| \int |F| d\nu_k(F) \geq 1\,.$$

and define $f(x) = f_{j,k,h}(x)$ on each B_j, and $f(x) = \frac{1}{2} Ax \cdot x$ elsewhere. The boundary conditions on $f_{j,k,h}$ and $u_{j,k,h}$ ensure continuous matching at the boundaries. The function $\tilde{f}(x) = f(x) + |x|^2$ satisfies $\nabla^2 \tilde{f} \in E$ and $\int_{B_1} |\nabla^2 \tilde{f}| = \infty$. Details are given in the proof of Theorem 2 of [3].

Finally, by convex integration or with a Baire Category argument along the lines of [11], one can show that there is a rank-one convex $g : \mathbb{R}^{2\times 2}_{\text{sym}} \to \mathbb{R}$ such that

$$\int_\Omega |\nabla^2 g| = \infty \qquad \text{for all open } \Omega \subset B_1$$

with $g(x) = |x|^2$ outside B_1. For the proof see [3, Remark 9] or [1, Section 3]. □

In closing, we show how the present result can be extended to any three-dimensional linear subspace of $\mathbb{R}^{2\times 2}$.

Corollary 6. *Let V be a three-dimensional linear subspace of $\mathbb{R}^{2\times 2}$. Then there is a rank-one convex function $f : V \to \mathbb{R}$ which is smooth outside of the unit ball B_1 and such that $\int_{B_1} |\nabla^2 f| = \infty$.*

Proof. The corollary follows from Theorem 1 and Theorem 6 of [3] via the classification of three-dimensional subspaces mentioned in the Introduction. For the convenience of the reader we give some details on this classification.

Any three-dimensional subspace of $\mathbb{R}^{2\times 2}$ can be identified by a nonvanishing normal $G \in \mathbb{R}^{2\times 2}$. We distinguish two cases, depending on whether the rank of G is one or two.

Case 1: $\operatorname{rank} G = 1$**.** We write $G = a \otimes b$ for a, $b \in S^1$, and observe that the three linearly independent matrices

$$F_1 = a \otimes b^\perp , \qquad F_2 = a^\perp \otimes b^\perp , \qquad F_3 = a^\perp \otimes b$$

are in V, hence are a basis for V (we recall that $(a \otimes b) \cdot (c \otimes d) = (a \cdot c)(b \cdot d)$). Using the relation

$$\det(a \otimes b + c \otimes d) = (a \wedge c)(b \wedge d)$$

we see that a generic $F = \alpha F_1 + \beta F_2 + \gamma F_3 \in V$ has rank one if and only if

$$\det F = \det([\alpha a + \beta a^\perp] \otimes b^\perp + \gamma a^\perp \otimes b) = -\alpha\gamma$$

vanishes. Let $\mathbb{R}^{2\times 2}_{\text{tri}}$ denote the lower triangular matrices and e_1, e_2 the canonical basis of $\mathbb{R}^2$. Then the linear map $L : V \to \mathbb{R}^{2\times 2}_{\text{tri}}$ that maps the three basis vectors into $e_1 \otimes e_1$, $e_2 \otimes e_1$, and $e_2 \otimes e_2$ has the property that $\operatorname{rank}[L(F)] = 1$ if and only if $\operatorname{rank} F = 1$. Therefore $f : V \to \mathbb{R}$ is rank-one convex if and only if $f \circ L^{-1} : \mathbb{R}^{2\times 2}_{\mathrm{tri}} \to \mathbb{R}$ is.

It remains to prove the result on triangular matrices. Let $g : \mathbb{R}^2 \to \mathbb{R}$ be the function constructed in Theorem 6 of [3], which is separately convex and has unbounded Hessian. We define $f : \mathbb{R}^{2\times 2}_{\mathrm{tri}} \to \mathbb{R}$ by

$$f\left(\begin{pmatrix} x & 0 \\ z & y \end{pmatrix}\right) = g(x, y)\,.$$

The thesis follows.

Case 2: $\operatorname{rank} G = 2$**.** We can write

$$G = a \otimes b + t a^\perp \otimes b^\perp$$

for some $a, b \in S^1$ and $t \in \mathbb{R}$, $t \neq 0$. The three matrices $F_1 = t^{1/2} a \otimes b - t^{-1/2} a^\perp \otimes b^\perp$, $F_2 = a^\perp \otimes b$ and $F_3 = a \otimes b^\perp$ form a basis for V, and computing as above we get

$$\det(\alpha F_1 + \beta F_2 + \gamma F_3) = \det\left(\left[t^{1/2}\alpha a + \beta a^\perp\right] \otimes b + \left[\gamma a - \frac{\alpha}{t^{1/2}} a^\perp\right] \otimes b^\perp\right) = -\beta\gamma - \alpha^2\,.$$

The linear map $L : V \to \mathbb{R}^{2\times 2}_{\mathrm{sym}}$ defined by

$$L f_1 = e_1 \otimes e_2 + e_2 \otimes e_1\,, \quad L f_2 = e_1 \otimes e_1\,, \quad L f_3 = -e_2 \otimes e_2$$

preserves the rank-one cone and hence rank-one convexity. The result follows from Theorem 1. □

In closing we remark that a direct extension of this construction to the case of rank-one convex functions on $\mathbb{R}^{2\times 2}$ does not seem possible. Indeed, one key ingredient of our method is the splitting in the z direction, which is a rank-one direction connecting the critical line e with the interior of the cone E_{sym}, but no such rank-one direction exists in the corresponding cone for $\mathbb{R}^{2\times 2}$.

Acknowledgements: The work of S.C. and F.M. was partially supported by the Deutsche Forschungsgemeinschaft through the Schwerpunktprogramm 1095 *Analysis, Modeling and Simulation of Multiscale Problems.* D.F was partially supported by the EU Research Training Networks *HYperbolic and Kinetic Equations*, contract HPRN-CT-2002-00282 and *Phase Transitions in crystalline solids*, contract FMRX-CT 98-0229.

References

[1] K. Astala, D. Faraco, and L. Székelyhidi Jr., *Convex integration and L^p theory of elliptic equations*, in preparation.

[2] M. Chlebík and B. Kirchheim, *Rigidity for the four gradient problem*, J. reine angew. Math. **551** (2002).

[3] S. Conti, D. Faraco, and F. Maggi, *A new approach to counterexamples to L^1 estimates: Korn's inequality, geometric rigidity, and regularity for gradients of separately convex functions*, MPI-MIS Preprint 93/2003.

[4] B. Dacorogna, *Direct methods in the calculus of variations*, Springer-Verlag, New York, 1989.

[5] B. Dacorogna and P. Marcellini, *Implicit partial differential equations*, Progress in Nonlinear Differential Equations and their Applications, 37, Birkhäuser, 1999.

[6] D. Faraco, *Milton's conjecture on the regularity of solutions to isotropic equations*, Ann. I. H. Poincaré **20** (2003), 889–909.

[7] ______, *Tartar conjecture and Beltrami operators*, Michigan Math. Journal **52** (2004), 83–104.

[8] M. Gromov, *Partial differential relations*, Springer-Verlag, 1986.

[9] L. Székelyhidi Jr., *The regularity of critical points of polyconvex functionals*, Arch. Rat. Mech. Anal. **172** (2004), 133–152.

[10] B. Kirchheim, *Deformations with finitely many gradients and stability of quasiconvex hulls*, C. R. Acad. Sci. Paris Sér. I Math. **332** (2001), 289–294.

[11] ______, *Rigidity and geometry of microstructures*, MPI-MIS Lecture notes n. 16, 2002.

[12] P. Marcellini, *Approximation of quasiconvex functions, and lower semicontinuity of multiple integrals*, Manuscripta Math. **51** (1985), 1–28.

[13] G. Dal Maso, I. Fonseca, G. Leoni, and M. Morini, *Higher-order quasiconvexity reduces to quasiconvexity*, Arch. Rat. Mech. Anal. **171** (2004), 55–81.

[14] S. Müller, *Variational models for microstructure and phase transitions*, in: Calculus of variations and geometric evolution problems (F. Bethuel et al., eds.), Springer Lecture Notes in Math. 1713, Springer-Verlag, 1999, pp. 85–210.

[15] S. Müller and V. Šverák, *Convex integration with constraints and applications to phase transitions and partial differential equations*, J. Eur. Math. Soc. (JEMS) **1** (1999), 393–442.

[16] ———, *Convex integration for Lipschitz mappings and counterexamples to regularity*, Ann. Math. **157** (2003), 715–742.

[17] D. A. Ornstein, *A non-inequality for differential operators in the L^1-norm*, Arch. Rat. Mech. Anal. **11** (1962), 40–49.

[18] M. A. Sychev, *Comparing two methods of resolving homogeneous differential inclusions*, Calc. Var. Part. Diff. Eq. **13** (2001), 213–229.

[19] L. Tartar, *Compensated compactness and applications to partial differential equations*, Nonlinear analysis and mechanics (R. J. Knops, ed.), Pitman, London, 1978, pp. 136–212.

[20] V. Šverák, *New examples of quasiconvex functions*, Arch. Rat. Mech. Anal. **119** (1992), 293–300.

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