A new approach to counterexamples to L^1 estimates: Korn's inequality, geometric rigidity, and regularity for gradients of separately convex functions

Sergio Conti,[1] Daniel Faraco[1] and Francesco Maggi[1,2]

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

[2] *Dipartimento di Matematica "U. Dini", Università di Firenze, Viale Morgagni 67/A, 50134 Firenze, Italy*

November 21, 2003

The derivation of counterexamples to L^1 estimates can be reduced to a geometric decomposition procedure along rank-one lines in matrix space. We illustrate this concept in two concrete applications. Firstly, we recover a celebrated, and rather complex, counterexample by Ornstein, proving the failure of Korn's inequality, and of the corresponding geometrically nonlinear rigidity result, in L^1. Secondly, we construct a function $f : \mathbb{R}^2 \to \mathbb{R}$ which is separately convex but whose gradient is not in BV_{loc}, in the sense that the mixed derivative $\partial^2 f/\partial x_1 \partial x_2$ is not a bounded measure.

1 Introduction

In the study of partial differential equations one often needs bounds for linear operators from L^p to L^p. The singular integral operators that typically appear can be studied in the general framework of harmonic analysis [20] if $1 < p < \infty$, whereas the limiting cases $p = 1$ and $p = \infty$ must be considered separately. In several concrete applications counterexamples can be obtained through a celebrated construction by Ornstein [18]. We present here a new method to treat this kind of examples, which leads to a simple geometric insight into the problem.

The first concrete application we present is to variational problems in elasticity, where one minimizes over vector fields $u : \Omega \subset \mathbb{R}^n \to \mathbb{R}^n$ the integral of an energy density $W(\nabla u)$ which is rotationally invariant, in the sense that $W(QF) = W(F)$ for any $Q \in SO(n)$. The energy density W vanishes on the identity matrix, and hence on the entire $SO(n)$; therefore, even strong growth assumptions on W do not deliver directly an equally strong control on ∇u. For smooth maps a classical result, due to Liouville, shows that if $\nabla u \in SO(n)$ almost everywhere then ∇u is a constant. Rigidity results of this kind play a fundamental role in the theory of elasticity, and several generalizations have been derived, among others, by Gehring [10], John [11] and Reshetnyak [19]. More recently Friesecke, James and Müller [9] have derived a quantitative rigidity estimate by showing that for any p strictly between 1 and ∞ and any bounded Lipschitz domain Ω there is a constant $c = c(p, \Omega)$ such that

$$\min_{Q \in SO(n)} \int_\Omega |\nabla u - Q|^p dx \le c \int_\Omega \mathrm{dist}^p(\nabla u, SO(n)) dx \,, \tag{1}$$

for all $u : \Omega \to \mathbb{R}^n$.

In the geometrically linear case one replaces $\mathrm{dist}(\nabla u, SO(n))$ by $|\nabla v + \nabla v^T|$, where $v(x) = u(x) - x$. This corresponds to using the distance from the tangent space to $SO(n)$ at the identity, which is the set of antisymmetric matrices. A classical result, known as Korn's inequality, states that for $1 < p < \infty$ gradients which have small symmetric part are approximately constant, in the sense that there is a constant $c = c(p, \Omega)$ such that

$$\min_{S = -S^T} \int_\Omega |\nabla u - S|^p dx \le c \int_\Omega \left|\nabla u + \nabla u^T\right|^p dx \,. \tag{2}$$

The estimate (2) does not hold for $p = 1$, as was shown by means of a rather involved construction by Ornstein in [18], see also the discussion below. We provide in Theorem 1 a simple construction showing that (2) does not hold for $p = 1$. This also shows that the nonlinear estimate (1) does not hold for $p = 1$. The combination of a similar rigidity estimate with the Sobolev embedding instead holds also in the critical case $p = 1$, as was shown by Kohn in [14].

The second part of this paper deals with a related but slightly different issue. It is well known that convex functions have a gradient which is locally of bounded variation, see e.g. [1]. Indeed, since the distributional Hessian is positive semidefinite, the total variation of the gradient on a set Ω can be controlled in terms of its oscillation on the boundary of Ω. We show that the same does not hold if convexity is replaced by separate convexity. Separate convexity was studied by Tartar in [21], in particular in connection with the issues of rank-one convexity and quasiconvexity. Tartar [21] also poses the question of which class of functions $g : \mathbb{R} \to \mathbb{R}$ can be the trace on the diagonal of a separately convex f, in the sense of $g(t) = f(t, t)$, shows that every $g \in C^2(\mathbb{R})$ has this property, and mentions additional partial results by Preiss and Šverák. The same question appears in the series of open problems proposed by Kirchheim, Müller and Šverák in [13]. Our example shows that even some functions whose second derivative is not a bounded measure can be obtained as traces of separately convex functions.

In closing, we remark that Ornstein's construction provides for any $k > 0$ a smooth function $f^k : \mathbb{R}^2 \to \mathbb{R}$ such that

$$\int_{(0,1)^2} |f^k_{12}| dx \geq k \int_{(0,1)^2} |f^k_{11}| + |f^k_{22}| dx \,.$$

This implies, for example, that the space of functions of bounded deformation BD, which is used in the theory of plasticity and is defined as the subset of L^1 such that the symmetric part of the distributional gradient $\nabla u + \nabla u^T$ is a bounded measure, is larger than the space of functions of bounded variation BV, which is the subset of L^1 such that the entire gradient ∇u is a bounded measure (to see this, consider $u^k = (f^k_1, -f^k_2)$). Ornstein's example was used to prove by duality [2, 5] that there are continuous functions which are not the divergence of any Lipschitz map, and a related construction works also for the nonlinear version of this problem, namely, $\det \nabla u = f$, see [3, 15]. Our construction for the separately convex case provides a similar sequence with the additional condition that f^k_{11} and f^k_{22} are nonnegative.

2 Counterexample to Korn's inequality and to geometrically nonlinear rigidity in L^1

We start from the geometrically linear case. The following result was first proven by Ornstein [18].

Theorem 1. *For any $k > 0$ and any $n \geq 2$ there is a function $u_k \in W^{1,\infty}((0,1)^n, \mathbb{R}^n)$ such that*

$$\int_{(0,1)^n} |\nabla u_k - F| dx \geq k \int_{(0,1)^n} |\nabla u_k + \nabla u_k^T| dx$$

for all $F \in \mathbb{R}^{n\times n}$.

The main tool used in our construction is the concept of laminate, which is the simplest class of mixtures of gradients that can be realized by finely-oscillating vector fields. Consider two matrices A and B. If their difference is a rank-one matrix, i.e. $A - B = a \otimes n$, then for any $\lambda \in (0,1)$ we can construct a continuous, piecewise affine vector field u_ε such that its gradient equals

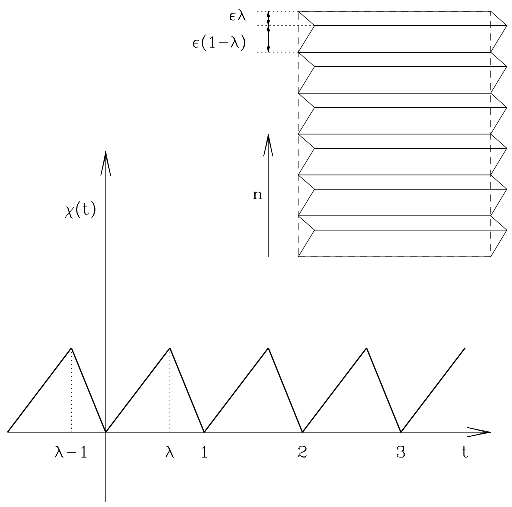

FIGURE 1: Graph of the periodic function $\chi(t)$, as defined after (3). The sketch in the upper right corner indicates the corresponding pattern generated by the deformation u_ε defined in (3). The quantities λ and $1-\lambda$ represent the volume fractions, the oscillation period is ε. The dashed line represent the average deformation Cr.

almost everywhere A or B, and therefore has average value $C = \lambda A + (1-\lambda)B$, by writing

$$u_\varepsilon(x) = Cx + \varepsilon\chi\left(\frac{x\cdot n}{\varepsilon}\right) a\,. \tag{3}$$

Here ε is a small positive parameter, representing the scale of the lamination, and $\chi : \mathbb{R} \to \mathbb{R}$ is a one-periodic function defined by $\chi(0) = 0$, $\chi'(t) = 1-\lambda$ if $t \in (0,\lambda)$ and $\chi'(t) = -\lambda$ if $t \in (\lambda, 1)$ (see Figure 1). As $\varepsilon \to 0$, the sequence $u_\varepsilon(x)$ converge weakly-$*$ in $W^{1,\infty}$ to an affine function with gradient C. Further, we can force u_ε to have C as affine boundary data, if we permit the gradient to deviate slightly on a small set, see Lemma 4 below. Therefore the construction can be iterated, by redefining u_ε, for example, in the regions where $\nabla u_\varepsilon = A$ as a function whose gradient oscillates on a scale much smaller than ε between two rank-one connected matrices whose (weighted) average is A.

In general, a probability measure on $n\times n$ matrices $\nu = \lambda\delta_A + (1-\lambda)\delta_B$ is a *first–order laminate with average* C if $\lambda A + (1-\lambda)B = C$, $\operatorname{rank}(A-B) = 1$, and $\lambda \in (0,1)$. Here δ_F denotes a Dirac mass supported on the matrix F. *Laminates of order* k *with average* C are then defined as the set of probability measures obtained from laminates of order $k-1$ replacing any δ_{C_j} with a first–order laminate with average C_j. For a detailed discussion of this and related issues, see e.g. [16, 17, 13]. In practice, the concept of laminate permits

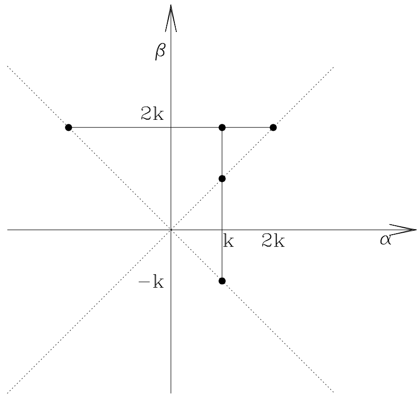

FIGURE 2: Schematic representation of the laminate used in the proof of Lemma 2. The oblique lines represent the matrices which are multiples of those in (4), on which all laminates are supported. The vertical line represents the splitting used to generate ν', the horizontal one the splitting used to generate ν''.

to construct Lipschitz maps with prescribed gradients by simply drawing the corresponding splitting of matrices along rank-one lines. The graphical representation permits to greatly simplify the construction procedure.

We construct a laminate supported on 2×2 matrices with zero diagonal entries which are either symmetric or antisymmetric, i.e., on multiples of the two matrices

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \text{and} \qquad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} . \tag{4}$$

The first is used only with a small weight δ, and with a large coefficient of order $1/\delta$. This ensures that the total L^1-norm of the symmetric part is bounded. The second matrix is used in the rest of the domain, with alternating positive and negative coefficients, which become large in significant parts of the domain, and eventually leads to a violation of the L^1 bound.

To present the precise construction, we parameterize the matrices above by

$$G_{\alpha,\beta} = \begin{pmatrix} 0 & \alpha \\ \beta & 0 \end{pmatrix}$$

and observe that in the (α, β) plane the rank-one directions are horizontal and vertical lines. Our construction is based on iteration of a basic step, which is illustrated in Figure 2 and presented in Lemma 2.

Staircase-type laminates were introduced in [6] to determine the range of exponents p for which the L^p theory of planar elliptic equations holds, depending on the ellipticity constant K, see also [7, 8]. In fact, our construction originates from the one used in [8] for the case that the ellipticity constant tends to 1. A related construction also plays a role in the determination of effective potentials in crystal plasticity [4].

Lemma 2. *For any $k > 0$ there is a laminate ν_k on 2×2 matrices, supported on multiples of the matrices in (4), such that*

$$\int F\,d\nu_k(F) = G_{k,k}\,, \quad \int \left|\frac{F+F^T}{2}\right| d\nu_k(F) = |G_{k,k}|\,,$$

$$\int |F|\,d\nu_k(F) = \frac{5}{3}|G_{k,k}|\,,$$

and which contains a Dirac mass on the matrix $G_{2k,2k}$ with weight $1/2$.

Proof. We construct the laminate explicitly, as shown in Figure 2. We first decompose $G_{k,k}$ into $G_{k,-k}$ and $G_{k,2k}$, and obtain

$$\nu' = \frac{1}{3}\delta_{k,-k} + \frac{2}{3}\delta_{k,2k}$$

which has average $G_{k,k}$ (in this section we use the shorthand notation $\delta_{\alpha,\beta} = \delta_{G_{\alpha,\beta}}$). This is possible since the lines at constant α are rank-one lines. Indeed, $G_{\alpha,\beta} - G_{\alpha,\beta'} = (\beta-\beta')e_2\otimes e_1$ is rank-one.

In a second step, we use that also lines at constant β are rank-one lines, and decompose $G_{k,2k}$ into a laminate supported on $G_{2k,2k}$ and $G_{-2k,2k}$. This gives

$$\nu'' = \frac{1}{4}\delta_{-2k,2k} + \frac{3}{4}\delta_{2k,2k}$$

which has average $G_{k,2k}$. Composing the two we obtain the final laminate,

$$\nu_k = \frac{1}{3}\delta_{k,-k} + \frac{1}{6}\delta_{-2k,2k} + \frac{1}{2}\delta_{2k,2k}\,.$$

It is now a simple check to see that all mentioned properties are satisfied. $\square$

We now iterate this construction.

Lemma 3. *There is a sequence of laminates of finite order $\nu^{(n)}$ on 2×2 matrices such that*

$$\int F\,d\nu^{(n)}(F) = G_{1,1}\,, \qquad \int \left|\frac{F+F^T}{2}\right| d\nu^{(n)}(F) = |G_{1,1}|\,,$$

with

$$\lim_{n\to\infty}\int |F|\,d\nu^{(n)}(F) = \infty\,.$$

Each laminate $\nu^{(n)}$ is supported on multiples of the matrices in (4).

Proof. We define the sequence of laminates $\nu^{(n)}$ iteratively. We set $\nu^{(0)} = \delta_{1,1}$. The laminate $\nu^{(1)}$ is the one obtained in Lemma 2 for $k = 1$, and contains a term $\frac{1}{2}\delta_{2,2}$. At stage n, we replace the term $2^{-n}\delta_{2^n,2^n}$ with the corresponding laminate obtained in Lemma 2, with $k = 2^n$. The averages of F and of $|F + F^T|$ are unchanged. On the other hand,

$$\begin{aligned}\int |F| d\nu^{(n)}(F) &= \int |F| d\nu^{(n-1)}(F) + \frac{2}{3}|G_{2^n,2^n}|2^{-n} \\ &= \int |F| d\nu^{(n-1)}(F) + \frac{2}{3}|G_{1,1}|\,.\end{aligned}$$

This concludes the proof. □

Theorem 1 follows now from the general fact that for any laminate ν of finite order one can find a sequence of uniformly Lipschitz functions $u_i : (0,1)^2 \to \mathbb{R}^2$ with affine boundary data given by the average of ν, such that

$$\int_{(0,1)^2} \phi(\nabla u_i(x))dx \to \int_{\mathbb{R}^{2\times 2}} \phi(F)d\nu(F) \tag{5}$$

for any continous function $\phi : \mathbb{R}^{2\times 2} \to \mathbb{R}$ (recall that ν is supported on finitely many points). For completeness we give here a brief sketch of how this sequence can be concretely constructed.

We start from a simple laminate $\nu = \lambda A + (1-\lambda)B$, and show how to modify the function defined in (3) so that it achieves affine boundary data, without modifying significantly the gradient distribution.

Lemma 4. *Let A, $B \in \mathbb{R}^{n\times n}$ fulfill $\mathrm{rank}(A - B) = 1$, and let $C = \lambda A + (1-\lambda)B$ for some $\lambda \in (0,1)$. Then, for any open domain $\Omega \subset \mathbb{R}^n$ and $\delta > 0$ there is a function $v \in W^{1,\infty}(\Omega, \mathbb{R}^n)$ such that*

(i). $v(x) = Cx$ *if* $x \in \partial\Omega$, *and* $\|v - Cx\|_{L^\infty} \leq \delta$,

(ii). $\mathrm{dist}(\nabla v, [A,B]) \leq \delta$ *a.e.*,

(iii). there are open sets ω_A and ω_B such that $\nabla v = A$ on ω_A, $\nabla v = B$ on ω_B, and $|\omega_A| \geq (1-\delta)\lambda|\Omega|$, $|\omega_B| \geq (1-\delta)(1-\lambda)|\Omega|$.

Proof. We construct v by interpolation around the boundary between the function u_ε defined in (3) and the affine function $x \to Cx$. By a standard covering and scaling argument, it is sufficient to prove the thesis in the unit cube $Q = (0,1)^n$.

From (3) we see that $|u_\varepsilon(x) - Cx| \leq c\varepsilon$, for some constant c depending on A and B only. Let ψ_δ be an interpolation function which equals one away from a δ-neighborhood of ∂Q equals zero in a δ^2-neighborhood of ∂Q, and has gradient bounded by $1/\delta$. We set

$$v(x) = \psi_\delta(x)u_\varepsilon(x) + (1 - \psi_\delta(x))Cx\,.$$

We have only to check that the gradient of v is close to the segment $[A, B]$. But,

$$\nabla v(x) = \psi_\delta \nabla u_\varepsilon + (1 - \psi_\delta)C + (u_\varepsilon - Cx)\nabla\psi_\delta$$

where the sum of the first two terms belongs to $[A, B]$. The last term is controlled by $c\varepsilon/\delta$. Hence, by choosing ε small enough, the proof is completed. □

The rest is a simple iteration procedure, which mimics the iterative definition of a laminate and the iterative construction of our ν's. Indeed, at each step we obtain a function which is affine on two sets ω_A and ω_B, which have measures δ-close to the weights of the corresponding matrices in the laminate. In the next splitting step, we apply Lemma 4 to each of the sets ω_A and ω_B. The resulting error in the sizes of the sets is controlled, after n steps, by $(1 - \delta)^n$ (note that we work here with laminates of finite order).

Proof of Theorem 1. The proof amounts to choosing a suitable diagonal subsequence. Precisely, let $\nu = \nu^{(n)}$ be a laminate as in Lemma 3 such that $\int |F| d\nu > k$. By the argument above there is a sequence $u_i : (0, 1)^2 \to \mathbb{R}^2$ with affine boundary data given by $G_{1,1}$ which satisfies (5), so that

$$\int_{(0,1)^2} \nabla u_i dx = G_{1,1}\,, \qquad \lim_{i\to\infty} \int_{(0,1)^2} |\nabla u_i| dx > k\,,$$

and

$$\lim_{i\to\infty} \int_{(0,1)^2} \left| \frac{\nabla u_i + \nabla u_i^T}{2} \right| dx = |G_{1,1}|\,.$$

Choosing a large enough i the theorem is proven. □

We now come to the geometrically nonlinear case.

Theorem 5. *For any $k > 0$ and any $n \geq 2$ there is a function $v_k \in W^{1,\infty}((0, 1)^n, \mathbb{R}^n)$ such that*

$$\int_{(0,1)^n} |\nabla v_k - Q| dx \geq k \int_{(0,1)^n} \operatorname{dist}(\nabla v_k, SO(n)) dx$$

for all rotations $Q \in SO(n)$.

Proof. Also in this case we can restrict ourselves to $n = 2$. The function v_k is constructed by a simple modification of the function u_k constructed in Theorem 1. We set, for some $\varepsilon_k > 0$,

$$v_k(x) = x + \varepsilon_k u_k(x)$$

and compute

$$\nabla v_k = \mathrm{Id} + \varepsilon_k \nabla u_k \,.$$

Now we observe that for any matrix F we have, by Taylor expansion,

$$\mathrm{dist}(\mathrm{Id} + F, SO(2)) \le \frac{1}{2}|F + F^T| + c|F|^2 \,.$$

Substituting $F = \varepsilon_k \nabla u_k$ and integrating we get

$$\int_{(0,1)^2} \mathrm{dist}(\nabla v_k, SO(2))dx \le \frac{\varepsilon_k}{2} \int_{(0,1)^2} |\nabla u_k + \nabla u_k^T|dx + c\varepsilon_k^2 \int_{(0,1)^2} |\nabla u_k|^2 dx \,,$$

and since $u_k \in W^{1,\infty}$ we can choose ε_k such that the first term is larger than the second one. On the other hand,

$$\begin{aligned} \min_{F\in\mathbb{R}^{2\times 2}} \int_{(0,1)^2} |\nabla v_k - F|dx &= \varepsilon_k \min_{F\in\mathbb{R}^{2\times 2}} \int_{(0,1)^2} |\nabla u_k - F|dx \\ &\ge \varepsilon_k k \int_{(0,1)^2} |\nabla u_k + \nabla u_k^T|dx \,. \end{aligned}$$

This concludes the proof. □

In closing, we remark that both Korn's inequality and the Friesecke-James-Müller rigidity hold in the weak-L^1 sense, i.e., the norm of $\nabla u - F$ in weak-L^1 is controlled by the L^1 norm of $|\nabla u + \nabla u^T|$ or of $\mathrm{dist}(\nabla u, SO(n))$, respectively, for a suitable F depending on ∇u.

3 A separately convex function whose gradient is not in BV_{loc}

Theorem 6. *There is a function $f : \mathbb{R}^2 \to \mathbb{R}$ which is separately convex but whose gradient is not in $BV_{\mathrm{loc}}(\mathbb{R}^2)$. Precisely, f_{11}, $f_{22} \in BV_{\mathrm{loc}}(\mathbb{R}^2)$, f_{11}, $f_{22} \ge 0$ (distributionally), but*

$$\int_{(0,1)^2} |f_{12}| = \infty \,.$$

The function f satisfies $f(x) = x^2$ for large $|x|$.

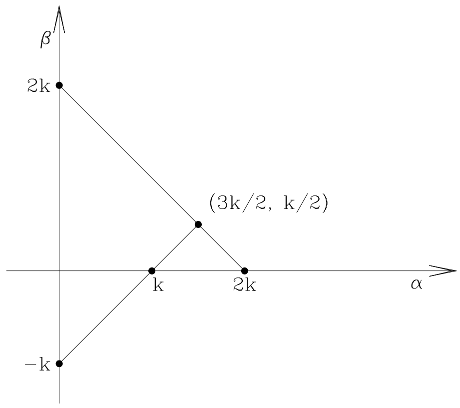

FIGURE 3: Sketch of the laminate used in the proof of Lemma 7.

Here and below, we denote by $\int_\Omega |g|$ the total variation of the distribution g on the domain Ω.

The idea is the following. Instead of constructing $f : \mathbb{R}^2 \to \mathbb{R}$, we construct a function $u : (0,1)^2 \to \mathbb{R}^2$ such that $\nabla u = \nabla u^T$, and $u(x) = 2x$ around the boundary. Then it is clear that there is an f such that $\nabla f = u$, which can be extended by $f(x) = x^2$ to the whole of $\mathbb{R}^2$. The construction of u is based on a laminate, that corresponds to the one used in the previous section, after a change of coordinates. We consider symmetric matrices which have equal diagonal entries, namely,

$$M_{\alpha,\beta} = \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix}, \qquad \alpha \geq 0\,,$$

which are second gradients of separately convex functions.

Lemma 7. *For any $k > 0$, there is a laminate ν_k supported on the matrices $M_{\alpha,\beta}$ such that*

$$\int F\, d\nu_k(F) = M_{k,0}\,, \quad \int |F_{11}| + |F_{22}| d\nu_k(F) = 2k\,,$$

with

$$\int |F| d\nu_k(F) = \frac{5}{3}|M_{k,0}|\,,$$

and which contains a Dirac mass on the matrix $M_{2k,0}$ with weight $1/2$.

Proof. The laminate is shown in Figure 3. Rank-one directions are, in the (k,h) plane, parallel to the $k = \pm h$ lines. Therefore we can decompose $M_{k,0}$

into $M_{0,-k}$ and $M_{3k/2,k/2}$, and obtain the laminate

$$\nu' = \frac{1}{3}\delta_{0,-k} + \frac{2}{3}\delta_{\frac{3}{2}k,\frac{1}{2}k}$$

which has average $M_{k,0}$ (we use here and below the shorthand notation $\delta_{k,h} = \delta_{M_{h,k}}$). In a second step, we decompose $M_{3k/2,k/2}$ into a laminate supported on $M_{2k,0}$ and $M_{0,2k}$,

$$\nu'' = \frac{1}{4}\delta_{0,2k} + \frac{3}{4}\delta_{2k,0}\,,$$

which has average $M_{3k/2,k/2}$. Composing the two we obtain the final laminate,

$$\nu_k = \frac{1}{3}\delta_{0,-k} + \frac{1}{6}\delta_{0,2k} + \frac{1}{2}\delta_{2k,0}\,.$$

□

The construction continues as in Lemma 3 and the following argument. Note, however, that the cutoff procedure if Lemma 4 violates, albeit slightly, the symmetry constraint on ∇u, hence produces functions which are not exact gradients. This problem can be solved by a simple modification of the construction used in the proof of Lemma 4. Precisely, the function u_ε as given in (3) now has symmetric gradient, and therefore we can find $f_\varepsilon \in W^{2,\infty}((0,1)^2,\mathbb{R})$ such that $\nabla f_\varepsilon = u_\varepsilon$, $f_\varepsilon(0) = 0$. For the same interpolation function ψ_δ, we define

$$g_\varepsilon(x) = \psi_\delta(x) f_\varepsilon(x) + (1-\psi_\delta(x))\frac{1}{2}x \cdot Cx \tag{6}$$

and $u_\varepsilon = \nabla g_\varepsilon$. The remaining details are identical to Lemma 4. Taking a diagonal sequence as above one then obtains for any $k > 0$ a separately convex function f^k such that

$$\int_Q |f^k_{12}| \geq k \int_Q |f^k_{11}| + |f^k_{22}|\,.$$

We now give a more detailed account how one can construct explicitly a separately convex function whose gradient is not in BV_{loc}.

Lemma 8. *Let $\Omega \subset \mathbb{R}^2$ and $k > 0$. For any $\varepsilon > 0$ there is a Lipschitz function $u : \Omega \to \mathbb{R}^2$ such that, for some open subset $\omega \subset \Omega$, we have*

(i). $u(x) = M_{k,0}x$ and $\nabla u(x) = M_{k,0}$ if $x \in \partial\Omega$,

(ii). $\|u(x) - M_{k,0}x\|_{L^\infty} \leq \varepsilon$,

(iii). $|\nabla u| \le 3|M_{k,0}|$ *a.e., and* $\nabla u = \nabla u^T$ *a.e.,*

(iv). $\nabla u = M_{2k,0}$ *a.e. on* ω*, and* $|\omega| \ge (1-\varepsilon)|\Omega|/2$*,*

(v). $\int_{\Omega\setminus\omega} |\nabla u| dx \ge \frac{1}{2}|M_{k,0}||\Omega \setminus \omega|$ *,*

(vi). $u_{1,1} \ge -\varepsilon$*,* $u_{2,2} \ge -\varepsilon$ *.*

Proof. This follows from the application of the construction in (6) to both lamination stages as discussed above in Lemma 7. Precisely, first consider $A_1 = M_{0,-k}$, $B_1 = M_{3k/2,k/2}$, $\lambda = 1/3$ and a small ε_1 (to be chosen later). This gives a function u_1, which on a set ω_1 has gradient B_1. Now we use the construction above on ω_1, with matrices $A_2 = M_{2k,0}$ and $B_2 = M_{0,2k}$, $\lambda_2 = 1/4$, and another small ε_2. This concludes the construction: (i), (ii) and (iii) are immediate. To obtain (iv) we need that $(1-\varepsilon_1)(1-\varepsilon_2) \ge (1-\varepsilon)$. Analogously, (v) holds provided that ε_1 and ε_2 are small enough (compared to 1), and if additionally $\varepsilon_1, \varepsilon_2 \le \varepsilon$ we also get (vi). $\square$

Proof of Theorem 6. We first construct $g : (0,1)^2 \to \mathbb{R}$ such that it agrees with x^2 around the boundary, satisfies $g_{11} \ge -1$ and $g_{22} \ge -1$ everywhere, and the integral of $|g_{12}|$ is infinite. Then the function

$$f(x) = \begin{cases} \frac{1}{2}g(x) + \frac{1}{2}x^2 & \text{if } x \in (0,1)^2 \\ x^2 & \text{else} \end{cases}$$

will satisfy the statement of the theorem. In turn, the function g is constructed as the limit of functions $g^{(n)}$, each of which is obtained as the potential of a suitable vector field $u^{(n)}$ with symmetric gradient.

We start with $u^{(0)} : \Omega^{(0)} = (0,1)^2 \to \mathbb{R}^2$ defined by $u^{(0)}(x) = x$. Let $\Omega^{(1)}$ be an open set compactly contained in $(0,1)^2$, and such that $|\Omega^{(1)}| \ge 1-\varepsilon^{(1)}$, for some small $\varepsilon^{(1)}$ to be fixed later. We define $u^{(1)}$ in $\Omega^{(1)}$ as the function obtained from Lemma 8 with $k = 1$ and $\varepsilon = \varepsilon^{(1)}$, and in $\Omega^{(0)} \setminus \Omega^{(1)}$ as $u^{(0)}$, and then iterate using the set ω for $\Omega^{(2)}$.

At stage n, we apply Lemma 8 with $k = 2^{n-1}$ and $\varepsilon = \varepsilon^{(n)}$ to construct the function $u^{(n)}$ on the set $\Omega^{(n)}$. The latter is then extended as $u^{(n-1)}$ to the rest of the unit square. The next set $\Omega^{(n+1)}$ is the set ω in the statement of Lemma 8.

We claim that we can chose iteratively $\varepsilon^{(n)}$ so that at stage $n \ge 1$ the following holds: $u^{(n)}_{1,1} \ge -1$, $u^{(n)}_{2,2} \ge -1$ a.e.,

$$\frac{1}{2^n} < |\Omega^{(n)}| < \frac{1}{2^{n-1}}\,, \qquad \nabla u^{(n)} = M_{2^n,0} \text{ on } \Omega^{(n+1)}\,,$$

and

$$\int_{\Omega^{(n)}\setminus\Omega^{(n+1)}} |\nabla u^{(n)}| dx \geq \frac{1}{\sqrt{2}} 2^{n-1} |\Omega^{(n)} \setminus \Omega^{(n+1)}| \leq c > 0\,.$$

The only delicate point is the control from below of the size of $\Omega^{(n)}$. To guarantee the inequality above, it is sufficient at step n to choose $\varepsilon^{(n)}$ so that $(1-\varepsilon^{(n)})|\Omega^{(n)}| > 2^{-n}$. This is always possible since by induction $|\Omega^{(n)}| > 2^{-n}$, and guarantees that $|\Omega^{(n+1)}| > 2^{-n-1}$.

Finally, if the sequence $\varepsilon^{(n)}$ is chosen summable the sequence $u^{(n)}$ converges to a limit u, and at the same time the sequence of primitives $g^{(n)}$ converges to a limit g. To see this, it is sufficient to observe that $\|u^{(n)} - u^{(n-1)}\|_{L^\infty} \leq \varepsilon^{(n)}$. Therefore $u^{(n)} \to u$ in L^∞. This implies that the $g^{(n)}$ converge strongly in $W^{1,\infty}$, and hence the limit is also separately convex. Indeed, the pointwise limit of convex functions is convex, and $W^{1,\infty}$ convergence implies pointwise convergence. □

Remark 9. Following the approach of Kirchheim [12, Chapter 3] based on Baire category it can be also shown that there are separately convex functions $f : \mathbb{R}^2 \to \mathbb{R}$ such that f_{12} is not a bounded measure on any open set. Indeed, let X^0 be the set of $W^{2,\infty}$ functions which coincide with x^2 up to the gradient on the boundary of $(0,1)^2$, which satisfy f_{11}, $f_{22} > 0$, and which have second gradient piecewise constant. The closure X of X^0 under the strong $W^{1,\infty}$ topology is a complete metric space. Any function in X can be extended to a separately convex function on $\mathbb{R}^2$ by

$$f(x+z) = f(x) + 2x\cdot z + z^2 \qquad \text{for } x \in [0,1)^2\,,\ \ z \in \mathbb{Z}^2\,.$$

We now define, for any open ball $B \subset \Omega$ with rational center and radius, and for any integer $M \geq 0$, the closed set

$$X_{B,M} = \left\{ f \in X : \int_B |\nabla^2 f| \leq M \right\}\,.$$

By using the construction above one can prove that the interior of $X_{B,M}$, under the strong $W^{1,\infty}$ topology, is empty. It follows that the complement of the union of the $X_{B,M}$ is residual, and hence nonempty, by the Baire category theorem. In particular, there is a function $f \in X \setminus \cup X_{B,M}$. It is separately convex on $\mathbb{R}^2$ and obeys

$$\int_B |f_{12}| = \infty \text{ for any open ball } B \subset \mathbb{R}^2.$$

Acknowledgements

We thank Bernd Kirchheim for bringing the question of regularity of separately convex functions to our attention. This work was partially supported by the EU Research Training Network *HYperbolic and Kinetic Equations*, contract HPRN-CT-2002-00282, and by the Deutsche Forschungsgemeinschaft through the Schwerpunktprogramm 1095 *Analysis, Modeling and Simulation of Multiscale Problems*.

References

[1] G. Alberti and L. Ambrosio, *A geometrical approach to monotone functions in* $\mathbb{R}^n$, Math. Z. **230** (1999), 259–316.

[2] J. Bourgain and H. Brezis, *On the equation* $\operatorname{div} y = f$ *and application to control of phases*, J. of the Am. Math. Soc. **16** (2002), 393–426.

[3] D. Burago and B. Kleiner, *Separated nets in Euclidean space and Jacobians of bilipschitz maps*, GAFA, Geom. funct. anal. **8** (1998), 273–282.

[4] S. Conti and M. Ortiz, *Dislocation microstructures and the effective behavior of single crystals*, in preparation.

[5] B. Dacorogna, N. Fusco, and L. Tartar, *On the solvability of the equation* $\operatorname{div} u = f$ *in* L^1 *and in* C^0, preprint (2002).

[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 (2003), to appear.

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

[9] G. Friesecke, R. James, and S. Müller, *A theorem on geometric rigidity and the derivation of nonlinear plate theory from three dimensional elasticity*, Comm. Pure Appl. Math **55** (2002), 1461–1506.

[10] F. W. Gehring, *Rings and quasiconformal mappings in space*, Proc. Nat. Acad. Sci. U.S.A. **47** (1961), 98–105.

[11] F. John, *Rotation and strain*, Comm. Pure Appl. Math. **14** (1961), 391–413.

[12] B. Kirchheim, *Rigidity and geometry of microstructures*, MPI-MIS Lecture notes n. 16, 2002.

[13] B. Kirchheim, S. Müller, and V. Šverák, *Studying nonlinear pde by geometry in matrix space*, Geometric analysis and nonlinear partial differential equations (S. Hildebrandt and H. Karcher, eds.), Springer, 2003, pp. 347–395.

[14] R. V. Kohn, *New integral estimates for deformations in terms of their nonlinear strains*, Arch. Rat. Mech. Anal. **78** (1982), 131–172.

[15] C. T. McMullen, *Lipschitz maps and nets in Euclidean space*, GAFA, Geom. funct. anal. **8** (1998), 304–314.

[16] 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, Berlin, 1999, pp. 85–210.

[17] S. Müller and V. Šverák, *Convex integration for Lipschitz mappings and counterexamples to regularity*, Ann. of Math. **157** (2003), 715–742.

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

[19] Y. G. Reshetnyak, *Liouville's theory on conformal mappings under minimal regularity assumptions*, Sibirskii Math. J. **8** (1967), 69–85.

[20] E. M. Stein, *Singular integrals and differentiability properties of functions*, Princeton University Press, 1970.

[21] L. Tartar, *Some remarks on separately convex functions*, Microstructure and phase transition (D. Kinderlehrer et al., eds.), IMA Vol. Mat. Appl., vol. 54, Springer, 1993, pp. 191–204.