

Validity and Failure of the Cauchy-Born Hypothesis in a Two-Dimensional Mass-Spring Lattice

G. Friesecke and F. Theil

Mathematics Institute, University of Warwick, Coventry CV4 7AL, UK
e-mail: gf@maths.warwick.ac.uk; theil@maths.warwick.ac.uk

Received September 27, 2001; accepted March 26, 2002
Online publication August 9, 2002
Communicated by R. Kohn

Summary. The Cauchy-Born rule postulates that when a monatomic crystal is subjected to a small linear displacement of its boundary, then all atoms will follow this displacement. In the absence of previous mathematical results, we study the validity of this rule in the model case of a 2D cubic lattice interacting via harmonic springs between nearest and diagonal neighbours. Our main result is that for favourable values of the spring constants and spring equilibrium lengths, the CB rule is actually a theorem.

Simple counterexamples show that for unfavourable spring parameters or large displacements the CB rule fails. Moreover the resulting overestimation of the lattice energy per unit volume by the CB rule cannot be cured by convexification (let alone quasiconvexification) of the CB energy.

The main tool in our proof is a novel notion of lattice polyconvexity which allows us to overcome the difficulty that the elastic energy as a function of atomic positions can never be convex, due to frame-indifference.

1. Introduction

This article is devoted to two basic questions concerning the elastostatics of a 2D lattice of particles interacting via interatomic potentials. We are not aware of any previous mathematically rigorous results on these questions, and hence study them in the innocent-looking model case where the particles have equal mass and are linked by harmonic springs between nearest and next nearest neighbours (see Figure 1).¹

The nearest neighbour springs and diagonal springs are characterized by their respective equilibrium lengths $a_1, a_2 > 0$ and spring constants $K_1, K_2 > 0$. Despite the

¹ Nearest neighbour models have no shear resistance and are hence incapable of capturing elasticity.

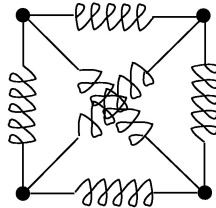


Fig. 1. Unit cell of the mass-spring lattice.

interaction force of each spring being linear with respect to interatomic distance, the model is still nonlinear. The remaining nonlinearity is of geometric origin (caused by the frame-indifference of atomic forces), and not due to any modelling assumption. It is a universal feature of all atomic models which respect the law of frame-indifference (such as Lennard-Jones models, non-pair-potential models, density functional theory, or full quantum mechanics).

Question 1. What is the ground state of a large finite lattice subject to (say linear) deformations prescribed at the boundary? In particular, does crystalline order ensue (or persist under change of the linear boundary conditions) if all nonboundary atoms are allowed to relax into their energy-minimizing positions?

Question 2. What is the energy per unit volume of the ground state when the system size gets large?

Question 1 is a variant of the well-known crystallization problem. (Another variant, obtained by replacing the linear displacement boundary conditions by zero applied forces, is studied in our companion paper [FT02].) That the question is far from trivial even for very simple-looking interactions is illustrated by the sphere packing problem, which remains unsolved in three dimensions despite a long history of attempts.

Question 2 means calculating, from atomic parameters, the continuum-mechanical stored-energy function of the crystal, i.e., the macroscopic energy per unit volume as a function of macroscopic deformation gradient. Note that by prescribing linear boundary conditions on the atomic system, one prescribes the average or macroscopic deformation gradient.²

In applications, Question 2 (extraction of a continuum-mechanical stored-energy function from an atomistic model) is usually “solved” by assuming a trivial solution to Question 1. One ignores the possibility of relaxation of atomic positions from the outset and postulates instead (“Cauchy-Born hypothesis”) that the minimum energy is attained when each unit cell individually follows the linear deformation prescribed at the boundary, see Figure 2.

² In the simpler case of elastic continua, knowledge of such a stored-energy function obtained by relaxation under linear bc’s suffices to determine the minimum energy under nonlinear bc’s as well (e.g. [Da89, Ch. 5, Thm. 2.1]).

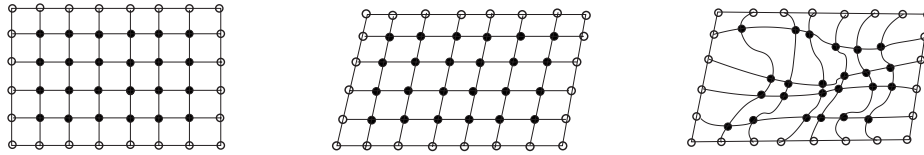


Fig. 2. Reference configuration of a finite lattice; Cauchy-Born deformation; and general deformation subject to the same boundary condition.

Our result for the 2D mass-spring lattice is that for an open set of atomic parameters (spring constants $K_1, K_2 > 0$ and spring equilibrium lengths $a_1, a_2 > 0$) and all boundary data close to the identity, the CB hypothesis is actually a theorem, while for another open set of atomic parameters the CB hypothesis fails, even at the identity. (Moreover, for any choice of parameters it fails for some boundary data.) In the failure region, energy-minimizing configurations are shown to exhibit fine-scale spatial oscillations. See Section 2 for precise statements.

Mathematically, the instability results require only finite-dimensional analysis—indeed, instability occurs already in the two-dimensional subspace of “shift-relaxation” deformations corresponding to the formation of two interpenetrating superlattices. (The corresponding energy surface at different parameter values is pictured in Figure 6.) But the stability result, i.e., the fact that in a certain parameter region the Cauchy-Born state is the unique minimizer, is an infinite-dimensional result and therefore less trivial.

The main difficulty is that the energy as a function of atomic positions is never convex (not even locally convex in any neighbourhood of the Cauchy-Born state), even when the interatomic potentials are convex with respect to interatomic distance. This nonconvexity is of universal geometric origin, and is the energetic manifestation of the geometric nonlinearity of the forces due to frame-indifference, discussed earlier. In one-dimensional lattice models such as those studied in [Tr96], [PP00], nonconvexity always leads to inhomogeneous or oscillatory atomic relaxation patterns. The two main mathematical ideas which allow us to understand how homogeneity can survive in 2D in spite of nonconvexity are the following:

- (a) A novel notion of *lattice polyconvexity*, inspired by the analogous notion for continuous systems [Ba77], [Da89]. Lattice polyconvexity is weaker than convexity (and is consistent with frame-indifference), but turns out to be sufficient to establish that Cauchy-Born states are minimizers subject to their own boundary conditions.
- (b) A careful analysis of how lattice-polyconvexity of a model is implied by a *single-well property* and *nonnegativity of the vibrational spectrum* of a local “cell energy function” on the lattice. The cell energy function accounts for all degrees of freedom of each unit cell, including those which are suppressed by the Cauchy-Born hypothesis. Its vibrational spectrum can be thought of as a local counterpart of the phonon spectrum of the lattice: Instead of global vibrational modes (phonons), one studies the local vibrational modes of a single unit cell, unconstrained by compatibility or periodicity constraints imposed by neighbouring cells.

These methods reduce Question 1, which concerns a sequence of problems of larger and larger dimension, to an auxiliary problem of fixed finite dimension. (See Theorem 5.1

for a precise statement and see Lemmas 5.2 and 5.3 for steps (a) and (b). The local cell-energy function is constructed in Section 3.) The methods should, at least in principle, be applicable to much more general interatomic force models, including anharmonic, nonpair, or further-neighbour interactions.

For our model interactions, the finite-dimensional auxiliary problem can be analyzed completely. In particular, we determine the (eight-dimensional) normal modes of the unit cell depicted in Figure 1. (See Section 6.)

Similarities and differences of our results with those of relaxation theory for continuous elastic bodies (and its central notion of quasiconvex hull) are indicated. In particular, it turns out that in the parameter region where the CB rule fails, the energy per unit volume of the lattice as a function of macroscopic deformation gradient is quasiconvex, but it is not the quasiconvex hull of the CB energy density. Physically this means that atomic relaxation effects are irrecoverable from first eliminating atomic degrees of freedom by the CB rule, and then allowing the continuum degrees of freedom to relax. (See Section 2.)

An important question left open by our analysis is the precise structure of the fine-scale spatial oscillations of the ground state (i.e., their period or degree of disorder) in the parameter region where the CB rule fails. This issue is easy to resolve only in the case of stress free boundary conditions, which will be discussed elsewhere.

Another issue beyond the scope of this paper is the elastostatic response of lattices to more general, i.e., nonaffine, boundary conditions (or applied forces). In this more complicated situation, a partial answer to Questions 1 and 2 which would bypass the finer issue of explicit atomic relaxation patterns is suggested by recent work on discrete models arising in fracture mechanics, image segmentations, and phase transitions [BDG99], [Ge99], [Go98], [PP00] (see also [Tr96], [FJ00]). In these studies, the asymptotic regime of large system size is investigated. The discrete minimizers are shown to converge (in a suitable weak topology which corresponds physically to averaging over a larger and larger number of atoms) to minimizers of a limiting continuum theory. In our 2D elastic lattice, an obvious candidate for the limit functional is $\int_{\Omega} W(Dy(x)) dx$, with $W(F)$ as arising in our analysis of the Cauchy-Born rule (see 2.5).

Finally we remark that it would be interesting to extend the present study to incorporate thermal effects, and to determine the asymptotic behaviour of lower-order contributions to the lattice energy such as surface or edge energies, which could explain e.g. some of the interesting effects observed numerically by Novak and Salje [NS98]. For their more general interactions, the construction of a local cell-energy function is explained in Section 8.

2. Lattice Energy Functional and Main Result

We begin by specifying the finite lattice occupied by the particles in a reference configuration. It is taken to be $\mathcal{L} = r^* \mathbb{Z}^2 \cap \Omega_R$, where $r^* > 0$ is a lattice parameter, $\Omega \subset \mathbb{R}^2$ is an open bounded domain with sufficiently regular (say Lipschitz continuous) boundary, and Ω_R is the dilate $\{Rx \mid x \in \Omega\}$, where $R > 0$. A deformation of the lattice is described by a vector field $y: \mathcal{L} \rightarrow \mathbb{R}^2$, where $y(x)$ denotes the position of the particle with reference position $x \in \mathcal{L}$.

The total elastic energy of the deformed lattice consists of the contributions of the nearest neighbour springs and of the diagonal springs,

$$\begin{aligned} E[\{y(x)\}_{x \in \mathcal{L}}] &= \frac{1}{2} \sum_{\substack{x, x' \in \mathcal{L} \\ |x-x'|=r^*}} \frac{K_1}{2} (|y(x) - y(x')| - a_1)^2 \\ &\quad + \frac{1}{2} \sum_{\substack{x, x' \in \mathcal{L} \\ |x-x'|=\sqrt{2}r^*}} \frac{K_2}{2} (|y(x) - y(x')| - a_2)^2. \end{aligned} \quad (2.1)$$

This system constitutes what we perceive as a minimal atomistic model for nonlinear elasticity: (1) The model is frame indifferent, i.e., $E(\{Ry(x)\}_{x \in \mathcal{L}}) = E(\{y(x)\}_{x \in \mathcal{L}})$ for all $R \in SO(2)$; (2) The functional form of the interaction potentials is the simplest possible; (3) The number of neighbours is minimal: Nearest neighbour models ($K_2 = 0$) have no shear resistance and are hence incapable of capturing elastostatics.

We study the following problem, which concerns the elastostatic response of the lattice to a prescribed linear deformation at the boundary:

$$\begin{aligned} &\text{Minimize } E[y] \text{ among deformations } y \in \mathcal{A} \\ &\text{satisfying the boundary condition (2.3).} \end{aligned} \quad (2.2)$$

The class \mathcal{A} of admissible deformations will be specified shortly and the boundary condition is

$$y(x)|_{x \in \partial \mathcal{L}} = Fx, \quad (2.3)$$

where F is a given 2×2 matrix with $\det F > 0$. Here the set of boundary atoms is given by $\partial \mathcal{L} = \{x \in \mathcal{L} \mid x \text{ possesses a nearest or next nearest neighbour } x' \in r^* \mathbb{Z}^2 \text{ with } x' \notin \mathcal{L}\}$.

The above simple energy, like all lattice models for crystals, only makes physical sense for deformations y whose magnitude does not exceed the plastic limit beyond which dislocations are created and the list of nearest neighbours of the atoms begins to change. In particular, the above model does not distinguish energetically between the identity and ‘‘folded’’ configurations such as $y(r^*me_1, r^*ne_2) = r^*ne_2$ when $m = 0 \pmod{2}$ and $r^*ne_2 + r^*e_1$ when $m = 1 \pmod{2}$. Its domain must be restricted to orientation-preserving maps

$$\mathcal{A} := \{y: \mathcal{L} \rightarrow \mathbb{R}^2 \mid \det'(D'y(x)) \geq 0 \forall x \in \mathcal{L}'\}, \quad (2.4)$$

where $\det'(D'y(x))$ is the oriented area of the deformed unit cell with corners $y(x)$, $y(x + r^*e_1)$, $y(x + r^*(e_1 + e_2))$, $y(x + r^*e_2)$ (see Section 4 below for an algebraic definition of $\det'(D'y(x))$), and \mathcal{L}' is the set of bottom left corners of unit cells, $\mathcal{L}' = \{x \in \mathcal{L} \mid x, x + r^*e_1, x + r^*(e_1 + e_2), x + r^*e_2 \in \mathcal{L}\}$.

As explained in the Introduction, we are interested in (1) periodicity properties of the minimizers $y \in \mathcal{A}$, and (2) their elastic energy per unit volume (as a function of the macroscopic deformation $y(x) = Fx$ prescribed on the boundary) as the system size gets large:

$$W(F) := \lim_{R \rightarrow \infty} W_{\Omega, R}(F), \quad \text{where } W_{\Omega, R}(F) := \min_{\substack{y \in \mathcal{A} \\ y(x)|_{x \in \partial \mathcal{L}} = Fx}} \frac{E[\{y(x)\}_{x \in \mathcal{L}}]}{\text{vol}(\Omega_R)}. \quad (2.5)$$

In particular, we seek to prove or disprove the Cauchy-Born hypothesis which postulates that the minimum in (2.5) is attained when each unit cell individually follows the prescribed linear deformation, in which case $W(F)$ coincides with the much simpler function

$$W_{CB}(F) := \lim_{R \rightarrow \infty} \frac{E[\{Fx\}_{x \in \mathcal{L}}]}{\text{vol}(\Omega_R)}. \quad (2.6)$$

The above definitions make sense because of the following lemma.

Lemma 2.1. *The limits in (2.5), (2.6) exist, and are independent of the region Ω .*

Here existence of the clamped-atom limit (2.6) is trivial;³ by contrast, the proof of existence of the relaxed-atom limit (2.5) requires some work and is postponed to Appendix 1.

To compare W and W_{CB} , it is convenient to choose the lattice parameter r^* in such a way that the deformation gradient $F = I$ is an equilibrium, i.e., satisfies $\partial W_{CB}(F)/\partial F = 0$. As seen a posteriori from Propositions 3.2 and 3.1 below, this is (a) always possible for any given spring constants $K_1 > 0$, $K_2 > 0$, and spring equilibrium lengths $a_1 > 0$, $a_2 > 0$, and (b) uniquely determines r^* as

$$r^* = \frac{K_1 a_1 + \sqrt{2} K_2 a_2}{K_1 + 2K_2}. \quad (2.7)$$

Remark. While the homogeneous cubic equilibrium state is geometrically trivial, it is not physically trivial, i.e., the interatomic forces are not individually zero. Instead, equilibrium occurs through detailed balance of nonvanishing forces (nonzero tensions and compressions $K_1(r^* - a_1)$, $K_2(\sqrt{2}r^* - a_2)$ in the horizontal respectively diagonal springs, balanced by the equilibrium equation (3.8) for r^*).

As a final preparation before stating our main result, we extract from the spring constants and spring equilibrium lengths the following dimensionless parameters:

$$\alpha := \frac{a_2/\sqrt{2}}{a_1}, \quad \kappa = \frac{K_2}{K_1}.$$

Physically the case $\alpha = 1$ is special; it corresponds to the cubic equilibrium state possessing simultaneously relaxed nearest-neighbour and diagonal springs. When $\alpha > 1$, the diagonal bonds are repulsive, while if $\alpha < 1$ the diagonal bonds are attractive.

Theorem 2.2. *There exist open parameter regions $\mathcal{U} \supset \{(\alpha, \kappa) \in (0, \infty) \times (0, \infty) \mid \alpha = 1\}$, $\mathcal{U}' \supset \{(\alpha, \kappa) \in (0, \infty) \times (0, \infty) \mid \alpha < \frac{1}{2}, \kappa \geq \frac{1}{2} \frac{1}{1-2\alpha}\}$ with the following property.*

(a) *(Validity of Cauchy-Born rule) For $(\alpha, \kappa) \in \mathcal{U}$,*

$$W(F) = W_{CB}(F) \text{ for all } F \text{ in some open neighbourhood of } SO(2), \quad (2.8)$$

³ In fact it is even known to hold in the much more subtle case of long-range, quantum mechanical forces [LS77], [Fe85], [CLL98].

- and the minimization problem “Minimize $E[\{y(x)\}_{x \in \mathcal{L}}]$ among $y \in \mathcal{A}$ satisfying $y(x)|_{x \in \partial \mathcal{L}} = Fx$ ” has the unique solution $y(x) = Fx$ for each finite lattice \mathcal{L} .
- (b) (Failure of Cauchy-Born rule by changes in atomic parameters) For $(\alpha, \kappa) \in \mathcal{U}'$,

$$W(F) < W_{CB}(F) \text{ for all } F \text{ in some open neighbourhood of } SO(2),$$

and the superlattice

$$y_d(x) = \begin{cases} Fx + dx \in \mathcal{L} \setminus \partial \mathcal{L}, \frac{x_1}{r^*} + \frac{x_2}{r^*} \text{ even,} \\ Fx, \text{ otherwise} \end{cases} \quad (2.9)$$

with $d \in \mathbb{R}^2$ chosen suitably, has lower energy than the Cauchy-Born state $y(x) = Fx$ for all sufficiently large $R > 0$.

- (c) (Failure of Cauchy-Born rule by changes in applied loads) For all $\alpha > 0, \kappa > 0$, if $F = \frac{r}{r^*}I$ and $r \leq a_1/2$ then $W(F) < W_{CB}(F)$, and the shift-relaxed configuration given in (b) has lower energy than the Cauchy-Born state $y(x) = Fx$ for all sufficiently large $R > 0$.

See Figure 3.

The proof of part (a) is given in Section 6, using the reduction methods developed in Sections 4 and 5. Parts (b) and (c) are proved in Section 7.

Our viewpoint is that (b) and (c) should be considered unsurprising, and that the less expected result is (a). The reason is that the energy functional E is not convex in the atomic positions $y(x)$ ($x \in \mathcal{L}$), even when the interatomic potentials are convex functions of atomic distance. As discussed in the Introduction, this nonconvexity is of geometric origin and is caused by the frame-indifference $E[\{y(x)\}_{x \in \mathcal{L}}] = E[\{Ry(x)\}_{x \in \mathcal{L}}]$ for all $R \in SO(2)$ and all $y: \mathcal{L} \rightarrow \mathbb{R}^2$.

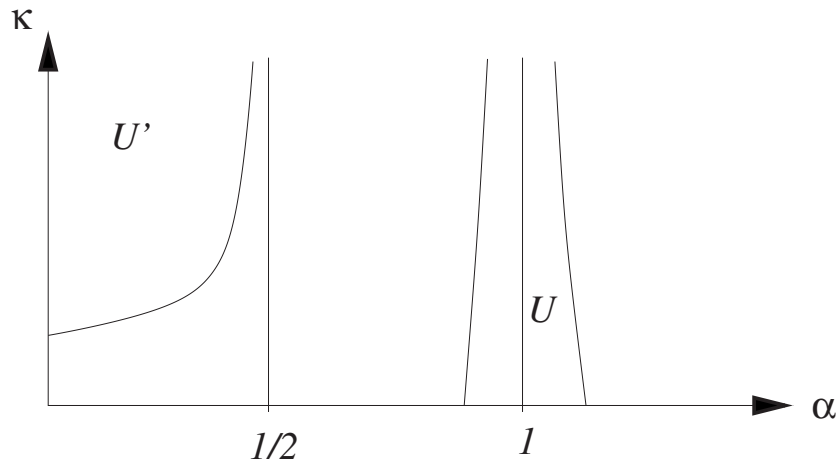


Fig. 3. Validity and failure of the CB hypothesis: In the parameter region \mathcal{U} , it is satisfied for F in some neighbourhood of $SO(2)$. In \mathcal{U}' it fails.

We do not know whether the energy can be lowered further by further symmetry breaking. An inkling of how difficult it is likely to be to determine the structure of the exact ground states can be had from the situation $\mathcal{L} = r^*\mathbb{Z}^2 \cap \{|x_1| + |x_2| \leq r^*\}$, corresponding to four boundary atoms surrounding a single “interior” atom. This problem is analyzed completely in Section 7. Depending on the choice of parameters, it contains up to seventeen different equilibrium positions of the interior atom, which correspond to the Cauchy-Born state plus four quartets of non-CB equilibria.⁴

What we can say however is that in the non-CB parameter region, the ground state is not just inhomogeneous, but necessarily contains short-scale spatial oscillations on the length scale of the lattice. To formulate this precisely, we introduce the long-wavelength relaxation (or quasiconvex hull) of $W_{CB}(F)$,

$$(W_{CB})^{qc}(F) := \inf_{\substack{y \in W^{1,2}(\Omega; \mathbb{R}^2) \\ y(x)|_{x \in \partial\Omega} = Fx, \det Dy(x) \geq 0}} \frac{1}{|\Omega|} \int_{\Omega} W_{CB}(Dy(x)) \, dx. \quad (2.10)$$

(Here $W^{1,2}(\Omega; \mathbb{R}^2)$ is the usual Sobolev space of square-integrable mappings from Ω to \mathbb{R}^2 with square-integrable gradient. The definition appears to depend on Ω , but it does not [Mo52], [Da89].) The following is an immediate consequence of Theorem 2.2 (b) and Proposition 3.1 in Section 3 below.

Corollary 2.3. *In the parameter region \mathcal{U}' (see Theorem 2.2), $W(F) < (W_{CB})^{qc}(F)$ for all $F \in SO(2)$.*

Physically, this result means that atomic relaxation effects are irrecoverable from first eliminating atomic degrees of freedom by the CB hypothesis and then allowing the continuum degrees of freedom to relax. For further information on this phenomenon, see Appendix 2.

By contrast, continuum relaxation effects are already contained in the atomic relaxation (2.5):

Proposition 2.4. *For all parameters $a_1 > 0$, $a_2 > 0$, $K_1 > 0$, $K_2 > 0$, W is quasiconvex, in the following sense:*

$$\frac{1}{|\Omega|} \int_{\Omega} W(Dy(x)) \, dx \geq W(F),$$

for all $y(x)|_{x \in \partial\Omega} = Fx$, $y \in C^1(\bar{\Omega}; \mathbb{R}^2)$, $\det Dy(x) > 0$ in $\bar{\Omega}$, and all $F \in M^{2 \times 2}$, $\det F > 0$.

The proof of the proposition is a straightforward adaptation of analogous results [Da89] for continuous elasticity models and is hence omitted. We do not know whether the class of admissible trial functions in the proposition can be enlarged to the less smooth class $y \in W^{1,2}(\Omega; \mathbb{R}^2)$ with $\det Dy(x) \geq 0$.

⁴ By the cubic symmetry of the five-atom problem, equilibria must appear as singlets, quartets, or octets.

3. Cauchy-Born Energy

Before tackling the complicated function W , which involves solving an asymptotically infinite-dimensional minimization problem, we analyze the much simpler function W_{CB} .

Proposition 3.1. *Let the lattice parameter of the reference configuration be given by (2.7). For any choice of the spring constants and spring equilibrium lengths $K_1 > 0$, $K_2 > 0$, $a_1 > 0$, $a_2 > 0$, W_{CB} is a single-well energy, that is to say*

- (a) W_{CB} is minimized on $\{F \in M^{2 \times 2} \mid \det F \geq 0\} =: \mathcal{D}$ if and only if $F \in SO(2)$.
 (b) There are no other critical points (i.e., F_0 such that $\frac{\partial W_{CB}}{\partial F}(F_0) = 0$) in the interior of \mathcal{D} .

Remarks. (1) In particular, in the parameter regime where Theorem 2.2 establishes that $W = W_{CB}$, the elastic energy per unit volume W as a function of average deformation gradient F has the single-well structure given in (a) and Figure 4. Such a single-well structure, derived here from interatomic forces, is a central feature of the classical phenomenological stored-energy functions of nonlinear elasticity theory (such as the Saint-Venant-Kirchhoff function and its refinements [Ba77].)

(2) The second neighbour interactions are essential for the single-well structure of W_{CB} . If $K_2 = 0$, then W_{CB} would fail to penalize collapse of the lattice by shearing, and the set of minimizers in (a) would be given by $\{RF_\theta \mid R \in SO(2), \theta \in [-\pi/2, \pi/2]\}$, where F_θ is the orthorhombic state

$$\begin{pmatrix} 1 & \sin \theta \\ 0 & \cos \theta \end{pmatrix}.$$

If $K_2 < 0$, then W_{CB} would promote collapse of the lattice by shearing, and the set of minimizers would be contained in the singular set $\{\det F = 0\}$.

The proof relies on two simple ideas. First, one partitions the energy functional E into a bulk part, additive with respect to unit cells, and a surface part, and observes that the latter is negligible in the limit $R \rightarrow \infty$. This partitioning gives an explicit formula for W_{CB} , and also plays an important role in our subsequent study of atomic

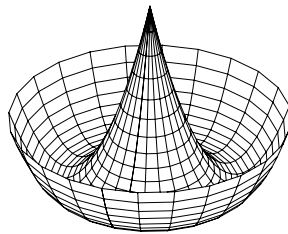


Fig. 4. The Cauchy-Born energy function restricted to the 2D subspace $\{F \in M^{2 \times 2} \mid F = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, a, b \in \mathbb{R}\}$ which contains $SO(2)$.

relaxation. Second, one avoids a direct study of the system of four nonlinear equations $\partial W_{CB}/\partial F = 0$ but instead works in carefully chosen geometric parameters suggested by the geometry of the image of the reference unit cell under the linear map $y(x) = Fx$.

Reduction to a finite-dimensional problem. Starting point is the energy functional (2.1). We rewrite the sum over “bonds” as a sum over unit cells plus a surface energy, i.e., decompose E as follows:

$$E[\{y(x)\}_{x \in \mathcal{L}}] = E_b[\{y(x)\}_{x \in \mathcal{L}}] + E_s[\{y(x)\}_{x \in \mathcal{L}}], \quad (3.1)$$

where the first term is a bulk term which can be written as a sum over unit cells,

$$E_b[\{y(x)\}_{x \in \mathcal{L}}] = \sum_{x \in \mathcal{L}'} E_{cell}(y(x), y(x+r^*e_1), y(x+r^*(e_1+e_2)), y(x+r^*e_2)) \quad (3.2)$$

(with the summation index x corresponding to the lower left corner of the cell and the summation domain \mathcal{L}' defined below), the cell energy is given by

$$E_{cell}(y_1, y_2, y_3, y_4) = \frac{1}{2} \sum_{i=1}^4 \frac{K_1}{2} (|y_{i+1} - y_i| - a_1)^2 + \sum_{i=1}^2 \frac{K_2}{2} (|y_{i+2} - y_i| - a_2)^2, \quad (3.3)$$

and the surface energy equals

$$\begin{aligned} E_s[\{y(x)\}_{x \in \mathcal{L}}] &= \frac{1}{4} \sum_{\substack{x, x' \in \partial \mathcal{L} \\ |x-x'|=r^*}} \frac{K_1}{2} (|y(x) - y(x')| - a_1)^2 \\ &\quad + \frac{1}{2} \sum_{\substack{x, x' \in \partial \mathcal{L} \\ |x-x'|=\sqrt{2}r^*}} \frac{K_2}{2} (|y(x) - y(x')| - a_2)^2. \end{aligned} \quad (3.4)$$

Here the set of lower left corners of unit cells appearing in (3.2) is given by $\mathcal{L}' = \{x \in \mathcal{L} \mid \text{the associated cell } \{x, x+r^*e_1, x+r^*(e_1+e_2), x+r^*e_2\} \text{ is contained in } \mathcal{L}\}$.

In expression (3.3) for the cell energy, note the extra factor $\frac{1}{2}$ in front of the horizontal contributions: It is necessary to avoid double-counting as each horizontal contribution features in two unit cells.

By means of the decomposition (3.1) and the fact that the surface energy (3.4) only grows linearly with R , i.e., does not contribute to the limit (2.6), we obtain the explicit formula

$$W_{CB}(F) = \frac{E_{cell}(Fx_1, Fx_2, Fx_3, Fx_4)}{r^{*2}}, \quad (3.5)$$

where

$$x_1 = a, \quad x_2 = r^*e_1 + a, \quad x_3 = r^*(e_1 + e_2) + a, \quad x_4 = r^*e_2 + a, \quad (3.6)$$

is the reference configuration of the unit cell of the lattice. (By translation invariance of the cell energy (3.3), W_{CB} is independent of the translation vector $a \in \mathbb{R}^2$.)

Affine versus nonaffine deformations of the unit cell. Analyzing W_{CB} means exploring the behaviour of the exact lattice bulk energy functional, but only for configurations

constrained in two drastic ways: (a) the reference configuration consists of a single unit cell, (b) the deformation of the unit cell is affine.

Here and below we say that a deformation $\{x_1, x_2, x_3, x_4\} \mapsto \{y_1, y_2, y_3, y_4\}$ of a unit cell is affine if there exist $F \in M^{2 \times 2}$ and $c \in \mathbb{R}^2$ such that $y_i = Fx_i + c$ for $i = 1, 2, 3, 4$. (Here the x_i are as defined in (3.6).) Geometrically this means that opposite edges of the cell must remain parallel.

The affine deformations of the unit cell constitute a 6D linear subspace

$$V_{CB} := \{y: \{x_1, x_2, x_3, x_4\} \rightarrow \mathbb{R}^2 \mid y \text{ is affine}\}$$

of the 8D space of deformations

$$V := \{y: \{x_1, x_2, x_3, x_4\} \rightarrow \mathbb{R}^2\}.$$

An interesting *nonlinear* property of this subspace follows.

Proposition 3.2. V_{CB} is invariant under the Euler-Lagrange operator

$$DE_{cell}: (\tilde{y}_1, \dots, \tilde{y}_4) \mapsto \left(\frac{\partial E_{cell}}{\partial y_1}(\underline{y}), \dots, \frac{\partial E_{cell}}{\partial y_4}(\underline{y}) \right) \Big|_{\underline{y}=\tilde{y}}$$

of the cell energy (3.3). In particular, if $F_0 \in M^{2 \times 2}$ is an equilibrium of the Cauchy-Born energy (i.e., $\frac{\partial W_{CB}(F)}{\partial F} \Big|_{F=F_0} = 0$), then the associated cell (Fx_1, \dots, Fx_4) is automatically an equilibrium of the cell energy (i.e., $(\frac{\partial E_{cell}}{\partial y_1}(\underline{y}), \dots, \frac{\partial E_{cell}}{\partial y_4}(\underline{y})) \Big|_{\underline{y}=(Fx_1, \dots, Fx_4)} = 0$).

Proof. This follows from the following dual characterization of V_{CB}

$$V_{CB} = \{y: \{x_1, x_2, x_3, x_4\} \rightarrow \mathbb{R}^2 \mid y_1 - y_2 + y_3 - y_4 = 0\},$$

the calculation

$$\nabla_{y_1} E_{cell} - \nabla_{y_2} E_{cell} + \nabla_{y_3} E_{cell} - \nabla_{y_4} E_{cell} = K_1 \sum_{i=1}^4 (-1)^i \left(1 - \frac{a_1}{|y_{i+1} - y_i|} \right) (y_{i+1} - y_i),$$

and the fact that the above right-hand side vanishes when $y \in V_{CB}$, since then the first term equals minus the third term and the second term equals minus the fourth term. \square

It will from now on be convenient to chose the reference configuration of the cell to be a stress-free state. A canonical candidate is given by the cubic cell whose lattice parameter r^* minimizes energy among cubic cells, i.e., minimizes

$$E_{cell}(y_1, y_2, y_3, y_4) = K_1(r^* - a_1)^2 + K_2(\sqrt{2}r^* - a_2)^2. \quad (3.7)$$

By solving the equation

$$0 = \frac{dE_{cell}}{dr^*}(x_1, x_2, x_3, x_4) = 2K_1(r^* - a_1) + 2\sqrt{2}K_2(\sqrt{2}r^* - a_2) \quad (3.8)$$

for r^* , we obtain the unique minimizer (2.7).

It is a fact that the so-obtained cubic cell $\{0, r^*e_1, r^*(e_1 + e_2), r^*e_2\}$ is indeed an equilibrium of E_{cell} on the full space V , i.e., it satisfies the Euler-Lagrange equations (denoting $\underline{y} = (y_1, y_2, y_3, y_4)$),

$$DE_{cell}(0, r^*e_1, r^*(e_1 + e_2), r^*e_2) = 0.$$

This follows either from a tedious direct computation, or more elegantly from Proposition 3.2 and Proposition 3.1.

Proof of Proposition 3.1 (b). It is advantageous to work not with the components of the strain matrix F as parameters, but instead in the following geometric coordinate system: For the deformed unit cell with corners Fx_1, Fx_2, Fx_3, Fx_4 , let $\theta \in (0, \pi)$ be the angle between the diagonals and let $a > 0, b > 0$ be the half lengths of the diagonals.

By elementary trigonometry, the sidelengths are

$$\sqrt{(a \pm b \cos \theta)^2 + (b \sin \theta)^2} = \sqrt{a^2 + b^2 \pm 2ab \cos \theta} =: s_{\pm}.$$

Consequently, with the shorthand notation $(y_1, y_2, y_3, y_4) =: \underline{y}$,

$$E_{cell}(\underline{y}) = 2\frac{K_1}{4}[(s_+ - a_1)^2 + (s_- - a_1)^2] + \frac{K_2}{2}[(2a - a_2)^2 + (2b - a_2)^2]. \quad (3.9)$$

A necessary condition for equilibria is

$$\begin{aligned} 0 &= \frac{d}{d\theta} E_{cell}(\underline{y}) = K_1 \left[(s_+ - a_1) \frac{-ab \sin \theta}{s_+} + (s_- - a_1) \frac{ab \sin \theta}{s_-} \right] \\ &= K_1 ab \sin \theta \left(\frac{a_1}{s_+} - \frac{a_1}{s_-} \right). \end{aligned}$$

Since $\sin \theta > 0$ in $(0, \pi)$, it follows that $s_+ = s_-$. Consequently $\cos \theta = 0$, i.e., $\theta = \pi/2$, corresponding to an orthorhombic cell. For an orthorhombic cell, the energy simplifies to

$$E_{cell}(\underline{y}) = K_1(\sqrt{a^2 + b^2} - a_1)^2 + \frac{K_2}{2}((2a - a_2)^2 + (2b - a_2)^2).$$

The necessary condition for equilibria

$$0 = \begin{pmatrix} \frac{\partial E_{cell}(\underline{y})}{\partial a} \\ \frac{\partial E_{cell}(\underline{y})}{\partial b} \end{pmatrix} = 2 \left[\left(K_1 \left(1 - \frac{a_1}{\sqrt{a^2 + b^2}} \right) + K_2 \right) \begin{pmatrix} a \\ b \end{pmatrix} - K_2 a_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right]$$

forces $a = b$, establishing that equilibria must be cubic. But the only cubic equilibria have sidelength r^* , as already established. The proof of (b) is complete. \square

Proof of Proposition 3.1 (a). To prove (a), observe first that $\lim_{|F| \rightarrow \infty} W_{CB}(F) = \infty$; consequently, it suffices to show that

$$\inf_{\det F=0} W_{CB}(F) > W(I).$$

In the polar coordinate system introduced above, the condition $\det F = 0$ implies at least one of (i) $a = 0$, (ii) $b = 0$, (iii) $\theta \in \{0, \pi\}$.

First, suppose (i) holds. Then $E_{cell}(\underline{y}) = K_1(b - a_1)^2 + K_2(a_2^2 + (2b - a_2)^2)/2$. For any $b > 0$, this is strictly bigger than the energy of the cubic cell with sidelength b , $K_1(b - a_1)^2 + K_2(\sqrt{2}b - a_2)^2$, the difference between the two energies being equal to $2K_2(\sqrt{2} - 1)ba_2$. Moreover we infer from this expression that when $b = 0$, $E_{cell}(\underline{y})$ equals the energy of a cubic cell with sidelength $b = 0$, which in turn is bigger than that of a cubic cell with sidelength r^* , as shown earlier. This establishes

$$\inf_{F|F \text{ satisfies (i)}} W_{CB}(F) > W_{CB}(I).$$

The case (ii) is analogous.

It remains to look at case (iii). Physically, one needs to compare the energy of a cubic “cluster” to that of a linear chain. Without loss of generality, we can assume $\alpha = 0$ and $a \geq b$; then $s_+ = a + b$, $s_- = a - b$, and

$$\begin{aligned} E_{cell}(\underline{y}) &= \frac{K_1}{2}[(a + b - a_1)^2 + (a - b - a_1)^2] \\ &\quad + \frac{K_2}{2}[(2a - a_2)^2 + (2b - a_2)^2]. \end{aligned} \quad (3.10)$$

The elementary calculus problem of minimizing (3.10) with respect to a and b is solved uniquely by

$$a^* = \frac{K_1 a_1 + K_2 a_2}{K_1 + 2K_2}, \quad b^* = \frac{K_2 a_2}{K_1 + 2K_2}. \quad (3.11)$$

To compare the energy of the linear chain $y_1 = -a^*v$, $y_2 = b^*v$, $y_3 = a^*v$, $y_4 = -b^*v$, $v \in \mathbb{R}^2$, $|v| = 1$, to that of the cubic equilibrium state requires a somewhat lengthy calculation, so we give some intermediate steps: Abbreviating $K' := K_1 + 2K_2$, one has $a^* + b^* - a_1 = (2K_2 a_2 - 2K_2 a_1)/K'$, $a^* - b^* - a_1 = (-2K_2 a_1)/K'$, $2a^* - a_2 = (2K_1 a_1 - K_1 a_2)/K'$, $2b^* - a_2 = (-K_1 a_2)/K'$, $r^* - a_1 = (\sqrt{2}K_2 a_2 - 2K_2 a_1)/K'$, $\sqrt{2}r^* - a_2 = (\sqrt{2}K_1 a_1 - K_1 a_2)/K'$. Consequently, by (3.7), (2.7), (3.10), (3.11),

$$\begin{aligned} &E_{cell}(-a^*v, b^*v, a^*v, -b^*v) - E_{cell}(0, r^*e_1, r^*(e_1 + e_2), r^*e_2) \\ &= \frac{K_1}{(K_1 + 2K_2)^2} \left[\frac{(2K_2 a_2 - 2K_2 a_1)^2 + (2K_2 a_1)^2}{2} - (\sqrt{2}K_2 a_2 - 2K_2 a_1)^2 \right] \\ &\quad + \frac{K_2}{(K_1 + 2K_2)^2} \left[\frac{(2K_1 a_1 - K_1 a_2)^2 + (K_1 a_2)^2}{2} - (\sqrt{2}K_1 a_1 - K_1 a_2)^2 \right]. \end{aligned}$$

It is now a straightforward matter to check that the first square bracket equals $4K_2^2(\sqrt{2} - 1)a_1 a_2 > 0$ and the second square bracket equals $2K_1^2(\sqrt{2} - 1)a_1 a_2 > 0$, completing the proof. \square

4. Discrete Null-Lagrangians

An important technical tool needed in studying the discrete boundary value problem in (2.5) is the existence of certain local nonlinear functions of the lattice deformation gradient whose sum over the lattice depends only on the boundary values of the deformation.

In modern continuum elasticity theory, such “null Lagrangians” play an important role. See in particular [Ba77], [Da89], where it is proved that in two dimensions the only such functions (i.e., the only continuous functions $f: M^{2 \times 2} \rightarrow \mathbb{R}$ such that if y, \tilde{y} are two smooth mappings from a bounded domain $\Omega \subset \mathbb{R}^2$ to \mathbb{R}^2 which agree on $\partial\Omega$, then $\int_{\Omega} f(Dy(x)) dx = \int_{\Omega} f(D\tilde{y}(x)) dx$) are of the form $a + B \cdot Dy(x) + c \det Dy(x)$ for some $a \in \mathbb{R}, B \in M^{2 \times 2}, c \in \mathbb{R}$.

We proceed to construct an analogue of the function $Dy(x) \mapsto \det Dy(x)$ on lattices. To this end, besides the standard discrete gradient of a map $y: \mathcal{L} \mapsto \mathbb{R}^2$ obtained by taking forward differences,

$$D^+y(x) := \frac{1}{r^*} (y(x+r^*e_1) - y(x) \mid y(x+r^*e_2) - y(x)) \in M^{2 \times 2}, \quad (4.1)$$

which would suffice if only Cauchy-Born deformations were under consideration, we introduce a generalized gradient which accounts for the additional atomistic degrees of freedom of the unit cell with bottom left corner x as well,

$$D'y(x) := \frac{1}{r^*} (y(x) - \bar{y} \mid y(x+r^*e_1) - \bar{y} \mid y(x+r^*(e_1+e_2)) - \bar{y} \mid y(x+r^*e_2) - \bar{y}) \in M^{2 \times 4},$$

where $\bar{y} = (y(x) + y(x+r^*e_1) + y(x+r^*(e_1+e_2)) + y(x+r^*e_2))/4$. Now for arbitrary $G = (g_1 \mid g_2 \mid g_3 \mid g_4) \in M^{2 \times 4}$ we define the following function:

$$\det'(G) := \frac{1}{2} (\det(g_2 - g_1 \mid g_4 - g_1) + \det(g_4 - g_3 \mid g_2 - g_3)).$$

Geometrically, $\det'(D'y(x))$ is the area of the deformed unit cell with bottom left corner $y(x)$ divided by the area of the reference unit cell. In particular $\det'(D'y(x)) = \det D^+y(x)$ when y is a Cauchy-Born deformation.

Lemma 4.1. *The function $D'y(x) \mapsto \det'(D'y(x))$ is a discrete null-Lagrangian, that is to say, if y and \tilde{y} are two mappings from \mathcal{L} to \mathbb{R}^2 which agree on $\partial\mathcal{L}$ then*

$$\sum_{x \in \mathcal{L}'} \det'(D'y(x)) = \sum_{x \in \mathcal{L}'} \det'(D'\tilde{y}(x)).$$

Remark. We do not know whether there exists a local function of the nongeneralized deformation gradient $D^+y(x)$ which is a null-Lagrangian.

Proof. We begin by noting that for discrete vector fields $v: \mathcal{L} \rightarrow \mathbb{R}^2$ one can easily derive a discrete version of Gauss’s theorem:

$$\sum_{x \in \mathcal{L}'} \operatorname{div}^+ v(x) = \sum_{x \in \mathcal{L} \cap \partial\Omega'} v(x) \cdot n(x). \quad (4.2)$$

Here $\operatorname{div}^+ v(x)$ is the standard discrete divergence obtained by taking forward differences,

$$\operatorname{div}^+ v(x) = \frac{1}{r^*}(v(x + r^*e_1) - v(x)) \cdot e_1 + \frac{1}{r^*}(v(x + r^*e_2) - v(x)) \cdot e_2,$$

the region $\Omega' \subset \mathbb{R}^2$ whose boundary appears in the discrete surface integral is the union of the squares $[x, x + r^*e_1] \times [x, x + r^*e_2]$ corresponding to unit cells with corner $x \in \mathcal{L}'$, and the normal vector field $n(x)$ on $\mathcal{L} \cap \partial\Omega'$ is defined as follows. On “edges,” i.e., when x has both a left and a right neighbour $\in \mathcal{L} \cap \partial\Omega'$, or both a top and a bottom neighbour $\in \mathcal{L} \cap \partial\Omega'$, $n(x)$ is simply the outward unit normal to $\partial\Omega'$. On “corners,” i.e., when x has a horizontal and a vertical neighbour $\in \mathcal{L} \cap \partial\Omega'$, the definition of n depends on the type of corner. One sets $n(x) = -e_1 - e_2$ for bottom left outward corners (i.e., when the neighbours of x on $\mathcal{L} \cap \partial\Omega'$ are $x + r^*e_1$ and $x + r^*e_2$ and the cell $[x, x + r^*e_1] \times [x, x + r^*e_2]$ belongs to Ω'), $n(x) = e_1 + e_2$ for top right inward corners (i.e., when the neighbours of x are as above but the cell $[x, x + r^*e_1] \times [x, x + r^*e_2]$ does not belong to Ω'), $n(x) = 0$ for bottom left inward corners and top right outward corners, $n(x) = e_1$ for bottom right outward corners, $n(x) = -e_2$ for bottom right inward corners, $n(x) = e_2$ for top left outward corners, and $n(x) = -e_1$ for top left inward corners.

The assertion of the lemma now follows upon noting that the function $x \mapsto \det'(D'y(x))$ can be written as a divergence,

$$\det'(D'y(x)) = \operatorname{div}^+ \left(\begin{pmatrix} \left(\frac{y(x) + y(x + r^*e_2)}{2} \cdot e_1 \right) (\operatorname{cof} D^+ y(x))_{11} \\ \left(\frac{y(x) + y(x + r^*e_1)}{2} \cdot e_1 \right) (\operatorname{cof} D^+ y(x))_{12} \end{pmatrix} \right), \quad (4.3)$$

where

$$\begin{aligned} (\operatorname{cof} D^+ y(x))_{11} &= \frac{y(x + r^*e_2) - y(x)}{r^*} \cdot e_2, \\ (\operatorname{cof} D^+ y(x))_{12} &= -\frac{y(x + r^*e_1) - y(x)}{r^*} \cdot e_2. \end{aligned}$$

To verify (4.3), one expands both left-hand side and right-hand side into products of components of y . This yields sixteen terms on each side. On each side, four terms appear twice, with opposite sign, leaving eight terms on each side. Each of these eight terms appears on both sides, with the same sign. The proof of the lemma is complete. \square

Remark. In a continuum limit the above identity (4.3) reduces to the well-known identity $\det Dy(x) = \operatorname{div}(y_1(\text{1st row of } \operatorname{cof}(Dy)))$. Note however the somewhat subtle way in which the symmetry in the continuous case that the same scalar factor, y_1 , appears in both components of the vector field is broken in the discrete case.

5. Finite-Dimensional Sufficient Condition for Validity of the CB Hypothesis

Rather than just proving Theorem 2.2(a), we derive a more general sufficient condition, not limited to harmonic potentials or central forces, which guarantees validity of the Cauchy-Born hypothesis (2.8) for deformations F near $SO(2)$. (For our particular mass-spring model, we will verify in Section 6 below that the sufficient condition holds in an open parameter region.) The sufficient condition only concerns the behaviour of a finite-dimensional auxiliary problem of fixed dimension, while the CB hypothesis concerns a sequence of larger and larger systems whose dimension tends to infinity.

Theorem 5.1. *Recall the cell energy $E_{cell}(y_1, y_2, y_3, y_4)$ defined in (3.3). Assume that*

(i) $E_{cell}(y_1, y_2, y_3, y_4)$ *is minimized among orientation-preserving mappings $\{x_1, x_2, x_3, x_4\} \mapsto \{y_1, y_2, y_3, y_4\}$ if and only if there exist $R \in SO(2)$ and $c \in \mathbb{R}^2$ such that $y_i = Rx_i + c$ ($i = 1, \dots, 4$).*

(ii) (Nondegeneracy) *The Hessian $D^2 E_{cell}(y_1, y_2, y_3, y_4)|_{y=\underline{x}}$ is positive-definite on the orthogonal complement of the 3D space spanned by the translations (e_1, e_1, e_1, e_1) and (e_2, e_2, e_2, e_2) and the linearized rotation $(e_1 - e_2, e_1 + e_2, -e_1 + e_2, -e_1 - e_2)$.*

Then for F in an open neighbourhood of $SO(2)$, (2.8) holds, and the unique minimizer of the variational problem (2.2) is the linear map $y(x) = Fx$.

Assumptions (i) and (ii) ensure that the cell energy $E_{cell}(y_1, y_2, y_3, y_4)$ looks qualitatively like the Cauchy-Born energy although it is defined on a higher-dimensional domain.

The proof of Theorem 5.1 is based on a definition and two lemmas, which allow us to overcome the fact that the cell energy $E_{cell}(y_1, y_2, y_3, y_4)$ is not convex (not even locally convex in any neighbourhood of the cubic equilibrium positions x_1, \dots, x_4), due to its frame-indifference $E_{cell}(y_1, y_2, y_3, y_4) = E_{cell}(Ry_1, Ry_2, Ry_3, Ry_4)$ for all $R \in SO(2)$.

Definition 1. (Recall that both the cell-energy function and the generalized determinant \det' associated with our 2D lattice are functions on $M^{2 \times 4}$, not $M^{2 \times 2}$.) Let $\tilde{\mathcal{D}}$ be a subset of $M^{2 \times 4}$ which is invariant under translations $(g_1, g_2, g_3, g_4) \mapsto (g_1 + d, g_2 + d, g_3 + d, g_4 + d)$, $d \in \mathbb{R}^2$. A translation-invariant function $\tilde{E}: \tilde{\mathcal{D}} \rightarrow \mathbb{R}$ is called *lattice-polyconvex* if there exists a convex function

$$f: \tilde{\mathcal{D}} \times \{a \in \mathbb{R} \mid a = \det'(F') \text{ for some } F' \in \tilde{\mathcal{D}}\} \rightarrow \mathbb{R},$$

such that $\tilde{E}(G) = f(G, \det'(G))$. (Here a function $f: \mathcal{D}_0 \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ with not necessarily convex domain \mathcal{D}_0 is called convex if whenever X, Y, Z are three distinct points in the domain \mathcal{D}_0 of f and $Y = \lambda X + (1 - \lambda)Z$ for some $\lambda \in (0, 1)$, then $f(Y) \leq \lambda f(X) + (1 - \lambda)f(Z)$.)

Definition 2. If there exists an f as in Definition 1 which satisfies in addition

$$f(G, \det G) > t(G, \det' G)$$

for all affine functions $t \leq f$ which agree with f at some point $G_0 \neq G$, and all $G, G_0 \in \tilde{\mathcal{D}} \cap \{G' = (g_1, g_2, g_3, g_4) \mid g_1 + g_2 + g_3 + g_4 = 0\}$, then \tilde{E} is said to be strictly lattice-polyconvex.

On generalized deformation gradients of affine (i.e., Cauchy-Born) lattice deformations, the notion of lattice polyconvexity introduced here reduces to classical [Ba77] continuum polyconvexity. Precisely, given $F \in M^{2 \times 2}$, denote by F' the generalized gradient $D'y(x)$ of the lattice map $y(x) = Fx + c$ ($x \in \mathcal{L}$), explicitly

$$F' = \frac{1}{2}(F(-e_1 - e_2) \mid F(e_1 - e_2) \mid F(e_1 + e_2) \mid F(-e_1 + e_2)). \quad (5.1)$$

Then a function $F \mapsto \tilde{E}(F)$ on $M^{2 \times 2}$ is polyconvex if and only if the function $F' \mapsto \tilde{E}(F)$ is lattice-polyconvex on $\{F' \mid F \in M^{2 \times 2}\}$.

Our first lemma is an obvious consequence of the definition of lattice polyconvexity and of the null-Lagrangian property of the lattice determinant; analogous—indeed much more general—statements are known in a continuum context [KS84]. Our second lemma does not seem obvious; the analogous statement in a continuum context (see Appendix 2) appears to be new as well.

Below, \mathcal{D}' denotes the set of values of generalized deformation gradients $D'y(x)$ of orientation-preserving mappings $y: \mathcal{L} \rightarrow \mathbb{R}^2$, i.e.,

$$\mathcal{D}' = \left\{ G = (g_1 \mid g_2 \mid g_3 \mid g_4) \in M^{2 \times 4} \mid \sum_{i=1}^4 g_i = 0, \det' G > 0 \right\}.$$

Lemma 5.2. *Suppose $E_{cell}: \mathcal{D}' \rightarrow \mathbb{R}$ is a continuous function which agrees in an open neighbourhood of $SO(2)'$ ($:= \{F' \mid F \in SO(2)\}$) with a strictly lattice-polyconvex function $\tilde{E}: \mathcal{D}' \rightarrow \mathbb{R}$ with $\tilde{E} \leq E_{cell}$, and let the lattice energy functional E be given by (3.1), (3.2), with the surface energy $E_s[\{y(x)\}_{x \in \mathcal{L}}]$ depending only on $\{y(x)\}_{x \in \partial \mathcal{L}}$ and satisfying $E_s = O(\frac{1}{R})$ as $R \rightarrow \infty$. Then for all F in some open neighbourhood of $SO(2)$, the CB hypothesis (2.8) holds and the unique minimizer of the variational problem (2.2) is the linear map $y(x) = Fx$.*

Remark. Local agreement of a function E_{cell} with a lattice-polyconvex function $\tilde{E} \leq E_{cell}$ is not a local condition on E_{cell} , but requires information about E_{cell} on all of \mathcal{D}' .

Lemma 5.3. *Suppose a function $\tilde{E}: \{G \in M^{2 \times 4} \mid \det'(G) > 0\} \rightarrow \mathbb{R}$ is continuous, rotation- and translation-invariant, C^3 in a neighbourhood of $SO(2)'$, satisfies conditions (i) and (ii) in Theorem 5.1, and obeys the following growth condition at infinity:*

$$\liminf_{|G| \rightarrow \infty, G \in V_0} \frac{\tilde{E}(G)}{|G|^2} > 0, \quad \limsup_{|G| \rightarrow \infty, G \in V_0} \frac{\tilde{E}(G)}{|G|^2} < \infty, \quad (5.2)$$

where $V_0 = \{G = (g_1, g_2, g_3, g_4) \mid g_1 + g_2 + g_3 + g_4 = 0\}$ is the orthogonal subspace to the translations. Then \tilde{E} agrees in an open neighbourhood of $SO(2)'$ with a strictly lattice-polyconvex function $\tilde{E}_0: \{G \in M^{2 \times 4} \mid \det'(G) > 0\} \rightarrow \mathbb{R}$, with $\tilde{E}_0 \leq \tilde{E}$.

Proof of Lemma 5.2. We need to show that for all F in a neighbourhood of $SO(2)$,

$$E[y] \geq E[y_0] \forall y \in \mathcal{A} \cap \{y(x) \mid_{x \in \partial \mathcal{L}} = y_0(x)\}, \quad (5.3)$$

where $y_0(x) = Fx$. Since all lattice deformations satisfying the above boundary condition have identical surface energy E_s , it suffices to prove (5.3) with the total energy E replaced by the bulk energy E_b (see (3.2)). By a simple approximation argument we can without loss of generality assume that f is differentiable. We denote the partial derivatives of $f: \mathcal{D}' \times \mathbb{R}^+ \rightarrow \mathbb{R}$ with respect to its first and second argument with $\frac{\partial f}{\partial G}$ respectively $\frac{\partial f}{\partial d}$. We calculate

$$\begin{aligned} E_b[y] &= \sum_{x \in \mathcal{L}'} E_{cell}(D'y(x)) = \sum_{x \in \mathcal{L}'} f(D'y(x), \det'(D'y(x))) \\ &\geq \sum_{x \in \mathcal{L}'} \left(f(D'y_0(x), \det' D'y_0(x)) + \frac{\partial f}{\partial G}(D'y_0(x), \det' D'y_0(x)) \right. \\ &\quad \cdot (D'y(x) - D'y_0(x)) + \frac{\partial f}{\partial d}(D'y_0(x), \det' D'y_0(x))(\det' D'y(x) \\ &\quad \left. - \det' D'y_0(x)) \right), \end{aligned} \quad (5.4)$$

with strict inequality unless $D'y(x) = D'y_0(x)$ for all $x \in \mathcal{L}'$. The first term on the right-hand side equals $E_b[y_0]$. As for the second term, it is an easy exercise to check that it vanishes: For instance, introducing analogously to (4.1) the backward discrete gradient $D^-y(x) := r^{*-1}(y(x) - y(x - r^*e_1) \mid y(x) - y(x - r^*e_2))$, the first column of $D'y(x)$ equals $-\frac{3}{8}D^+y(x)e_1 - \frac{3}{8}D^+y(x)e_2 - \frac{1}{8}D^-y(x + r^*(e_1 + e_2)) - \frac{1}{8}D^-y(x + r^*(e_1 + e_2))$ and each component of each term can be written as a divergence, e.g. $(D^+y(x)e_1) \cdot e_i = \operatorname{div}_+(y(x) \cdot e_1 e_i)$. Finally, the last term on the right-hand side of (5.4) vanishes, by Lemma 4.1. The proof of Lemma 5.2 is complete. \square

Proof of Lemma 5.3. The proof is based on a careful analysis of the geometry of the mapping \det' (which, being an indefinite quadratic form, canonically splits $M^{2 \times 4}$ into a negative and a positive subspace) relative to the geometry of the embedded submanifold $SO(2)' \subset M^{2 \times 4}$.

Our goal is to show that the function

$$f(G, d) := (\tilde{E} + \epsilon \det')(G) - \epsilon d$$

(where $\epsilon > 0$ is a parameter to be specified below), which obviously satisfies $\tilde{E}(G) = f(G, \det' G)$, is bounded from below by a strictly lattice-polyconvex function with which it agrees in an open neighbourhood of $SO(2)'$. That is to say, our goal is to show that for some $\epsilon > 0$, $\delta > 0$,

$$\begin{aligned} (\tilde{E} + \epsilon \det')(H) - (\tilde{E} + \epsilon \det')(G) - D(\tilde{E} + \epsilon \det')(G) \cdot (H - G) &\geq 0, \\ \forall |G - SO(2)'| < \delta, G \in \mathcal{D}', \forall H \in \mathcal{D}', \text{ with equality iff } G = H. \end{aligned} \quad (5.5)$$

Different arguments are needed depending on whether H is close to G or H is far away from G .

Case 1: H is close to G . We show that there exist $\epsilon, \delta > 0$ such that for all $R' \in SO(2)'$ (5.5) holds whenever $|G - R'|, |H - R'| < \delta$.

By frame-indifference we can without loss of generality assume $R' = I'$. Recall the subspace V_0 in the statement of the lemma, which is the minimal subspace containing \mathcal{D}' . By translation invariance of \tilde{E} and \det' , we can view these functions as functions on the subspace V_0 . Let $\phi_G \in V_0$ be a lowest eigenvector of the Hessian $D^2\tilde{E}(G)$, and let $\phi_G^\perp \in V_0$ be a vector in the orthogonal complement of the lowest eigenspace. When $G \in SO(2)'$, the lowest eigenvalue of the Hessian equals 0 and all other eigenvalues are $\geq \gamma$ for some constant $\gamma > 0$. It follows from the local boundedness of $D^3\tilde{E}$ near $SO(2)'$ that there exist constants $\delta > 0$, $C > 0$ such that for all $|G - SO(2)'| < \delta$

$$D^2\tilde{E}(G)(\phi_G, \phi_G) \geq -C|G - SO(2)'|,$$

and

$$D^2\tilde{E}(G)(\phi_G^\perp, \phi_G^\perp) \geq \frac{\gamma}{2}. \quad (5.6)$$

holds. As for \det' , it is elementary to check that $\det'(\phi_{I'}) = \frac{\lambda}{2}|\phi_{I'}|^2$ for some $\lambda > 0$, whence

$$D^2\det'(G)(\phi_G, \phi_G) = 2\det'(\phi_G) \geq \frac{\lambda}{2}|\phi_G|^2 \quad \forall |G - SO(2)'| < \delta,$$

while due to the quadraticity of \det' there exists $\Lambda > 0$ such that

$$|D^2\det'(G)(\phi, \tilde{\phi})| \leq \Lambda|\phi||\tilde{\phi}| \quad \forall G, \phi, \tilde{\phi} \in V_0.$$

Armed with these four inequalities and the fact that $D^2\tilde{E}(G)(\phi_G, \phi_G^\perp) = 0$ (by symmetry of the Hessian and the fact that ϕ_G is an eigenvector), the Hessian of $\tilde{E} + \epsilon \det'$ can be estimated from below as follows:

$$\begin{aligned} & D^2(\tilde{E} + \epsilon \det')(G)(\phi_G + \phi_G^\perp, \phi_G + \phi_G^\perp) \\ &= D^2(\tilde{E} + \epsilon \det')(G)(\phi_G, \phi_G) + 2\epsilon D^2\det'(G)(\phi_G, \phi_G^\perp) + D^2(\tilde{E} + \epsilon \det')(\phi_G^\perp, \phi_G^\perp) \\ &\geq \left(\frac{\epsilon\lambda}{2} - C|G - SO(2)'|\right)|\phi_G|^2 - 2\epsilon\Lambda|\phi_G||\phi_G^\perp| + \left(\frac{\gamma}{2} - \epsilon\Lambda\right)|\phi_G^\perp|^2. \end{aligned}$$

By the Cauchy-Schwarz inequality, the middle term is bounded from below by $-\epsilon((\lambda/4)|\phi_G|^2 + (4\Lambda^2/\lambda)|\phi_G^\perp|^2)$. Consequently the right-hand side is bounded from below by

$$\left(\frac{\epsilon\lambda}{4} - C|G - SO(2)'|\right)|\phi_G|^2 + \left(\frac{\gamma}{2} - \epsilon\Lambda\left(1 + \frac{4\Lambda}{\lambda}\right)\right)|\phi_G^\perp|^2.$$

It follows that when

$$\epsilon < \frac{\gamma}{2\Lambda(1 + 4\frac{\Lambda}{\lambda})},$$

the Hessian of $\tilde{E} + \epsilon \det'$ on V_0 is positive-definite for all base points in the ball $B_{\lambda\epsilon/(4c)}(R')$ with center R' and radius $\lambda\epsilon/(4c)$. Consequently (5.5) holds for G, H in this ball.

Case 2: H is not close to G . This is the harder part, since local analysis will not suffice. We begin by noting the following global fact, which follows from assumptions (i), (ii) and the growth condition (5.2): There exist constants $C, c > 0$ such that

$$C|G - SO(2)'|^2 \geq \tilde{E}(G) - \tilde{E}(I') \geq c|G - SO(2)'|^2 \quad \forall G \in V_0. \quad (5.7)$$

Next, it follows from the fact that rigid rotations are the unique minimizers of \tilde{E} (see hypothesis (ii)) that $D\tilde{E}(G)|_{G \in SO(2)'} = 0$. Consequently by the local boundedness of $D^2\tilde{E}$ near $SO(2)'$,

$$D\tilde{E}(G) \cdot (H - G) \geq -C|G - SO(2)'||H - G|. \quad (5.8)$$

Next, we will need the following nontrivial property concerning the geometric relationship of \det' and the submanifold $SO(2)'$: For some constants $C, c > 0$, without loss of generality identical to those in (5.7),

$$\det'(H - G) \geq c|H - G|^2 - C(|H - SO(2)'|^2 + |G - SO(2)'|^2). \quad (5.9)$$

Finally, we will use a trivial property of \det' , which is true for any quadratic form $Q(G)$, by Taylor's theorem:

$$\det'(H) - \det'(G) - D\det'(G) \cdot (H - G) = \det'(H - G). \quad (5.10)$$

Proof of (5.9). For two rotations

$$R = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \quad \tilde{R} = \begin{pmatrix} \tilde{a} & -\tilde{b} \\ \tilde{b} & \tilde{a} \end{pmatrix} \in SO(2),$$

$\det'(R' - \tilde{R}') = \det(R - \tilde{R}) = 2((a - \tilde{a})^2 + (b - \tilde{b})^2) = |R - \tilde{R}|^2 = |R' - \tilde{R}'|^2$. Hence, choosing closest points R'_G, R'_H on $SO(2)'$ to G, H ,

$$\begin{aligned} \det'(H - G) &= \det'((R'_H - R'_G) + (H - R'_H) - (G - R'_G)) \\ &\geq |R'_H - R'_G| - 2\Lambda|R'_H - R'_G| \cdot (|H - R'_H| + |G - R'_G|) \\ &\quad - \Lambda(|H - R'_H| + |G - R'_G|)^2, \end{aligned}$$

from which the desired inequality (5.9) follows by straightforward Cauchy-Schwarz arguments.

It follows from (5.7), (5.8), (5.9), (5.10) that

$$\begin{aligned} &\tilde{E}(H) - \tilde{E}(G) - D\tilde{E}(G) \cdot (H - G) \\ &\quad + \epsilon (\det'(H) - \det'(G) - D\det'(G) \cdot (H - G)) \\ &\geq c|H - SO(2)'|^2 - C|G - SO(2)'|^2 - C|G - SO(2)'||H - G| \\ &\quad + \epsilon (c|H - G|^2 - C|H - SO(2)'|^2 - C|G - SO(2)'|^2) \\ &= (c - \epsilon C)|H - SO(2)'|^2 - (C + \epsilon c)|G - SO(2)'|^2 \\ &\quad + (\epsilon c|H - G| - C|G - SO(2)'|)|H - G|. \end{aligned}$$

Now let R' be an arbitrary element on $SO(2)'$. We claim that when $\epsilon \leq c/C$, if $H \notin B_{\lambda\epsilon/(4c)}(R')$ and $|G - R'| < \delta$ with δ sufficiently small, the right-hand side is positive. Indeed, in the above situation the right-hand side is bounded from below by $-(C +$

$\epsilon c \delta + (\epsilon c (\frac{\lambda \epsilon}{4C} - \delta) - C \delta) \cdot (\frac{\lambda \epsilon}{4C} - \delta)$, which tends to $\epsilon c (\frac{\lambda \epsilon}{4C})^2 > 0$ as $\delta \rightarrow 0$. The proof of Lemma 5.3 is complete. \square

Proof of Theorem 5.1. This is an immediate consequence of Lemma 5.2 and Lemma 5.3. \square

6. Verification of the Sufficient Condition for Validity of CB in an Open Parameter Region

Proof of Theorem 2.2 (a). We need to verify that for an open neighbourhood U of the parameter line $\{(\kappa, \alpha) \in (0, \infty)^2 \mid \alpha = 1\}$, the cell energy (3.3) of the mass-spring model satisfies the assumptions of Theorem 5.1. These assumptions consist of a global assumption, (i), and a local assumption, (ii).

The fact, proved below, that both (i) and (ii) fail in an open parameter region means that their verification must invoke subtle features of the interatomic forces and cannot, for instance, rely on simple properties such as cubic symmetry or convexity of bond potentials with respect to interatomic distance, which do not change as the transition occurs.

We begin by establishing the local assumption (ii). The Euler-Lagrange operator of the cell energy reads







$$\begin{aligned} \frac{\partial E_{cell}}{\partial y_i}(y_1, \dots, y_4) &= \frac{K_1}{2} \left(- \left(1 - \frac{a_1}{|y_{i+1} - y_i|} \right) (y_{i+1} - y_i) \right. \\ &\quad \left. + \left(1 - \frac{a_1}{|y_i - y_{i-1}|} \right) (y_i - y_{i-1}) \right) \\ &\quad - K_2 \left(1 - \frac{a_2}{|y_{i+2} - y_i|} \right) (y_{i+2} - y_i). \end{aligned}$$

The Hessian at the cubic equilibrium state $(x_1, \dots, x_4) = (0, r^*e_1, r^*(e_1 + e_2), r^*e_2)$ is

$$\begin{aligned} \frac{\partial^2 E_{cell}}{\partial y_i \partial y_i}(x_1, \dots, x_4) &= \left(K_1 - \frac{K_1 a_1}{2r^*} + K_2 \right) I - \frac{K_2 a_2}{\sqrt{2}r^*} \frac{x_{i+1} - x_{i-1}}{|x_{i+1} - x_{i-1}|} \otimes \frac{x_{i+1} - x_{i-1}}{|x_{i+1} - x_{i-1}|}, \\ \frac{\partial^2 E_{cell}}{\partial y_{i+1} \partial y_i}(x_1, \dots, x_4) &= \frac{K_1}{2} \left(-I + \frac{a_1}{r^*} \frac{x_i - x_{i-1}}{|x_i - x_{i-1}|} \otimes \frac{x_i - x_{i-1}}{|x_i - x_{i-1}|} \right), \\ \frac{\partial^2 E_{cell}}{\partial y_{i-1} \partial y_i}(x_1, \dots, x_4) &= \frac{K_1}{2} \left(-I + \frac{a_1}{r^*} \frac{x_{i+1} - x_i}{|x_{i+1} - x_i|} \otimes \frac{x_{i+1} - x_i}{|x_{i+1} - x_i|} \right), \\ \frac{\partial^2 E_{cell}}{\partial y_{i+2} \partial y_i}(x_1, \dots, x_4) &= K_2 \left(-I + \frac{a_2}{\sqrt{2}r^*} \frac{x_{i+1} - x_{i-1}}{|x_{i+1} - x_{i-1}|} \otimes \frac{x_{i+1} - x_{i-1}}{|x_{i+1} - x_{i-1}|} \right). \end{aligned}$$

Despite the high dimension of the Hessian (namely, 8×8), its eigenvalues and eigenvectors can be determined explicitly, for all values of the parameters K_1, K_2, a_1, a_2 .

Lemma 6.1. (Spectral analysis of Hessian of cell energy) *The eigenvalues and eigenmodes of the Hessian $D^2 E_{cell}(x_1, \dots, x_4)$ of the cell energy (3.3) at the cubic equilibrium state (3.6), (2.7) are*

Eigenvalue	Eigenvector	Description	
0	$(v, v, v, v), v \in \mathbb{R}^2$	Translations	
0	$((\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}), (\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}), (\begin{smallmatrix} -1 \\ 1 \end{smallmatrix}), (\begin{smallmatrix} -1 \\ -1 \end{smallmatrix}))$	Infinitesimal rotation	
$K_1 + 2K_2$	$((\begin{smallmatrix} -1 \\ -1 \end{smallmatrix}), (\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}), (\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}), (\begin{smallmatrix} -1 \\ 1 \end{smallmatrix}))$	Homogeneous dilation	
$\frac{K_1 a_1}{r^*}$	$((\begin{smallmatrix} -1 \\ 1 \end{smallmatrix}), (\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}), (\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}), (\begin{smallmatrix} -1 \\ -1 \end{smallmatrix}))$	$\langle 1, 0 \rangle$ -Poisson mode	
$\frac{\sqrt{2} K_2 a_2}{r^*}$	$((\begin{smallmatrix} -1 \\ -1 \end{smallmatrix}), (\begin{smallmatrix} -1 \\ 1 \end{smallmatrix}), (\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}), (\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}))$	$\langle 1, 1 \rangle$ -Poisson mode	
$K_1 \left(2 - \frac{a_1}{r^*} \right)$	$(v, -v, v, -v), v \in \mathbb{R}^2$	Non-Cauchy-Born (shift-relaxation) modes	

In particular, the Hessian restricted to the orthogonal complement of translations and infinitesimal rotations is positive-definite when $2 - \frac{a_1}{r^*} > 0$ but possesses a negative eigenvalue when $2 - \frac{a_1}{r^*} < 0$.

Proof of lemma. With respect to the canonical basis $\{(e_1, 0, 0, 0), (e_2, 0, 0, 0), (0, e_1, 0, 0), (0, e_2, 0, 0), \dots\}$ of $\mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2 \times \mathbb{R}^2$, the Hessian is an 8×8 matrix. We obtain a compact form by introducing the parameters

$$x := \frac{K_1 a_1}{2r^*}, \quad y := \frac{K_2 a_2 / \sqrt{2}}{2r^*}, \quad z := K_1 + K_2 - x - y.$$

The Hessian then reads

$$D^2 E_{cell}(x_1, \dots, x_4) = \begin{pmatrix} z & y & -\frac{K_1}{2} & 0 & -K_2 + y & -y & -\frac{K_1}{2} + x & 0 \\ y & z & 0 & -\frac{K_1}{2} + x & -y & -K_2 + y & 0 & -\frac{K_1}{2} \\ -\frac{K_1}{2} & 0 & z & -y & -\frac{K_1}{2} + x & 0 & -K_2 + y & y \\ 0 & -\frac{K_1}{2} + x & -y & z & 0 & -\frac{K_1}{2} & y & -K_2 + y \\ -K_2 + y & -y & -\frac{K_1}{2} + x & 0 & z & y & -\frac{K_1}{2} & 0 \\ -y & -K_2 + y & 0 & -\frac{K_1}{2} & y & z & 0 & -\frac{K_1}{2} + x \\ -\frac{K_1}{2} + x & 0 & -K_2 + y & y & -\frac{K_1}{2} & 0 & z & -y \\ 0 & -\frac{K_1}{2} & y & -K_2 + y & 0 & -\frac{K_1}{2} + x & -y & z \end{pmatrix}.$$

It is obvious from the translation- and rotation-invariance of E_{cell} that the dimension of the kernel of $D^2 E_{cell}(x_1, \dots, x_4)$ has dimension greater than or equal to three, with

corresponding zero-modes given by translations and by the infinitesimal rotation listed above.

Further, it follows from Proposition 3.2 that the 6D space V_{CB} (which contains the above three-dimensional space of zero-modes) is an invariant subspace of $D^2 E_{cell}(x_1, \dots, x_4)$. As would in fact be true for any elasticity tensor of a cubically symmetric material in two dimensions, further eigenmodes in V_{CB} turn out to be given by a uniform dilation, a volume-preserving combination of compression and expansion along the edges ($\langle 1, 0 \rangle$ -Poisson mode), and a volume-preserving combination of compression and expansion along the diagonals ($\langle 1, 1 \rangle$ -Poisson mode); this corresponds to eigenmodes 3, 4, and 5 in the list above. From the explicit formula for the Hessian, it is straightforward to check that these are indeed eigenmodes and that the eigenvalues are, respectively, $K_1 + 2K_2$, $K_1 + 2K_2 - 4y$, and $K_1 + 2K_2 - 2x$. These expressions appear at first sight to be different from those given in the lemma, but they are in fact the same, by virtue of the following linear relationship between the parameters K_1 , K_2 , x , and y :

$$K_1 + 2K_2 = 2x + 4y,$$

which follows from the equilibrium equation (3.8) for r^* .

It remains to find the non-Cauchy-Born eigenmodes. To this end, note that $D^2 E_{cell}(x_1, \dots, x_4)$ equals its own transpose, i.e., it is a self-adjoint operator on \mathbb{R}^8 ; consequently the orthogonal complement of the invariant subspace V_{CB} must also be an invariant subspace. But as noted earlier (see Section 3), the orthogonal complement is nothing but the space of shift-relaxation deformations $V_{SR} = \{(v, -v, v, -v) \mid v \in \mathbb{R}^2\}$. From the explicit formula for the Hessian, one sees that on V_{SR} the Hessian equals $2(K_1 - x)I$, completing the proof of the lemma. \square

Proof of Theorem 2.2 (a), cont. It remains to verify hypothesis (i) of Theorem 5.1. It is convenient to introduce the normalization

$$\tilde{E}_{cell}(y_1, \dots, y_4) := E_{cell}\left(\frac{y_1}{r^*}, \dots, \frac{y_4}{r^*}\right).$$

By translation invariance, it suffices to verify that in an open neighbourhood of the parameter region $\{(\alpha, \kappa) \in (0, \infty)^2 \mid \alpha = 1\}$, \tilde{E} restricted to the subspace

$$\tilde{V} := \left\{ \underline{y} = (y_1, \dots, y_4) \in \mathbb{R}^2 \times \dots \times \mathbb{R}^2 \mid \sum_{i=1}^4 y_i = 0 \right\}$$

is minimized iff $(y_1, \dots, y_4) \in SO(2)'$ (see Lemma 5.2 for the definition of $SO(2)'$).

By scaling we can assume $a_1 = K_1 = 1$, $a_2/\sqrt{2} = \alpha$, $K_2 = \kappa$. We fix $\kappa_2 > 0$ and study $\tilde{E} = \tilde{E}^{(\alpha, \kappa)}$ for parameters (α, κ) near $(1, \kappa_0)$.

By dropping the (nonnegative) contributions of the diagonal springs, we have $E_{cell}(\underline{y}) \geq \frac{1}{4}(\frac{1}{2} \sum_{i=1}^4 |y_{i+1} - y_i|^2 - 4)$. Since $\sum_{i=1}^4 y_i = 0$, a short calculation shows that the right-hand side is bounded from below by $\frac{1}{4}(\sum_{i=1}^4 |y_i|^2 - 4)$. It follows that

$$\tilde{E}(\underline{y}) \geq 1 \quad \text{for } |\underline{y}|^2 \geq 8r^*.$$

We claim next that there exist $\delta > 0, \eta > 0$ such that when

$$|(\alpha, \kappa) - (1, \kappa_0)| \leq \eta, \tag{6.1}$$

\tilde{E} is minimized on $\{\underline{y} \in (\mathbb{R}^2)^4 \mid \text{dist}(\underline{y}, SO(2)') < \delta\}$ if and only if $\underline{y} \in SO(2)'$. Indeed, as proved in Section 3, each $\underline{y} \in SO(2)'$ is a solution to the Euler-Lagrange system $D\tilde{E}(\underline{y}) = 0$, while the spectral result of Lemma 6.1 implies that $D^2\tilde{E}(\underline{y})$ is 1-1 on the orthogonal subspace to the tangent space to $SO(2)'$ at $\underline{y} \in SO(2)'$, whence—by the implicit function theorem—there are no further solutions near $SO(2)'$.

Now note that $\inf_{\text{dist}(\underline{y}, SO(2)') \geq \delta, \det'(\underline{y}) \geq 0} \tilde{E}^{(1, \kappa_0)}(\underline{y}) =: \gamma$ is positive, since for parameters on the line $\alpha = 1$ it is obvious that $SO(2)'$ is the set of global minimizers. Assume without loss of generality $\gamma < 1$. Now choose η so small that for all (α, κ) satisfying (6.1), $\tilde{E}^{(\alpha, \kappa)}|_{SO(2)'} \leq \gamma/2$ and $\inf_{\text{dist}(\underline{y}, SO(2)') \geq \delta, |\underline{y}|^2 \leq 8r^*, \det'(\underline{y}) \geq 0} \tilde{E}^{(\alpha, \kappa)}(\underline{y}) \geq \gamma/2$. But in $|\underline{y}|^2 \geq 8r^*, \tilde{E} \geq 1$. This shows that for all (α, κ) satisfying (6.1), the set of global minimizers of $\tilde{E}^{(\alpha, \kappa)}$ equals $SO(2)'$. The proof of Theorem 2.2 (a) is complete. \square

7. Failure of the Cauchy-Born Rule

As announced in the Introduction, the Cauchy-Born rule already fails in simple low-dimensional subspaces which allow for non-CB behaviour.

Recall from Lemma 6.1 that the cubic unit cell loses local stability for $r^* < a_1/2$, due to the appearance of an unstable non-CB eigenmode. Evidently, unit cells undergoing this failure mode can be combined (after rotating every other cell by 180°) into a coherent lattice, and the failure mode simply becomes a shift relaxation where 50% of the atoms are translated coherently against the remaining ones; see Figure 5.

We show here that this loss of local stability is *preceded* by a loss of global stability. In other words, the transition to non-CB minimizers is first order, not second order.

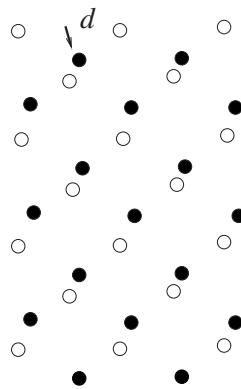


Fig. 5. Shift relaxation: The black atoms are shifted coherently with respect to the white ones.

To see this (and to complete the proof of Theorem 2.2), it will suffice to study the lattice energy (2.1) on the above 2D subspace of shift-relaxation deformations,

$$y_d(x) = \begin{cases} \frac{r}{r^*}x + d, & x \in \mathcal{L} \setminus \partial\mathcal{L}, \frac{x_1}{r^*} + \frac{x_2}{r^*} \text{ even,} \\ \frac{r}{r^*}x, & \text{otherwise.} \end{cases} \quad (7.1)$$

Here $r > 0$ is a dilation parameter which allows us to treat the more general boundary condition $y(x)|_{x \in \partial\mathcal{L}} = \frac{r}{r^*}x$ in place of $y(x)|_{x \in \partial\mathcal{L}} = x$.

We have, by the decomposition (3.1), (3.2), (3.3), (3.4) and the fact that under (7.1) all unit cells which do not contain a boundary atom have the same energy,

$$\lim_{R \rightarrow \infty} \frac{E[\{y_d(x)\}_{x \in \mathcal{L}}]}{\text{vol}(\Omega_R)} = \frac{E_{\text{cell}}(d, re_1, r(e_1 + e_2) + d, re_2)}{r^{*2}} =: \frac{E_{\text{shift}}(d)}{r^{*2}}. \quad (7.2)$$

The function $E_{\text{shift}}(d)$ relates to the stored-energy function (2.5) and the Cauchy-Born energy (2.6) as follows:

$$W\left(\frac{r}{r^*}I\right) \leq \inf_{d \in \mathbb{R}^2} r^{*2} E_{\text{shift}}(d) \leq r^{*2} E_{\text{shift}}(0) = W_{\text{CB}}\left(\frac{r}{r^*}I\right). \quad (7.3)$$

Thus in order to disprove the Cauchy-Born hypothesis at $\frac{r}{r^*}I$, it suffices to show that $d = 0$ is not the global minimizer of E_{shift} . (A global minimizer always exists, since $E_{\text{shift}}(d)$ depends continuously on d and tends to infinity as $|d| \rightarrow \infty$.)

Proposition 7.1. (See Figure 6.) Let

$$C_* := \max_{\rho > 0} \frac{\sqrt{1 + \sqrt{2}\rho + \rho^2} - 2 + \sqrt{1 - \sqrt{2}\rho + \rho^2}}{\rho^2}.$$

(Numerically, $C_* = 0.62\dots$)

- (a) If $\frac{r}{a_1} > C_*$, then $d = 0$ is the unique global minimizer of E_{shift} .
- (b) If $\frac{r}{a_1} = C_*$, then $d = 0$ is a global minimizer, but not unique.
- (c) If $\frac{1}{2} \leq \frac{r}{a_1} < C_*$, then $d = 0$ is a local minimizer but not a global minimizer.
- (d) If $\frac{r}{a_1} < \frac{1}{2}$, then $d = 0$ is not a local minimizer.

It is not difficult to understand physically the origin of the instability. Note first that $2E_{\text{shift}}(d)$ equals the sum of the interaction energies of an atom at position d with four surrounding atoms clamped at the positions $\pm re_1, \pm re_2$. Hence the problem “Minimize $E_{\text{shift}}(d)$ over $d \in \mathbb{R}^2$ ” is a special case of the boundary-value problem (2.2), by taking $\Omega_R = \{|x_1| + |x_2| \leq r^*\}$ (so that $\mathcal{L} = r^*\mathbb{Z}^2 \cap \Omega_R$ equals the five-atom configuration $\{0, \pm r^*e_1, \pm r^*e_2\}$ and $\mathcal{L} \setminus \partial\mathcal{L}$ reduces to a single atom) and choosing the boundary data $y(x)|_{x \in \partial\mathcal{L}} = \frac{r}{r^*}x$.

Due to the invariance of each interaction potential under rigid body rotations, the set of positions of the centre atom where it is minimal forms a circle—see Figure 7. If $r \ll 1$, the circles converge toward each other and the positions close to the circles become favourable compared with the symmetric centre position.

We postpone momentarily the proof of the proposition to give the following proof.

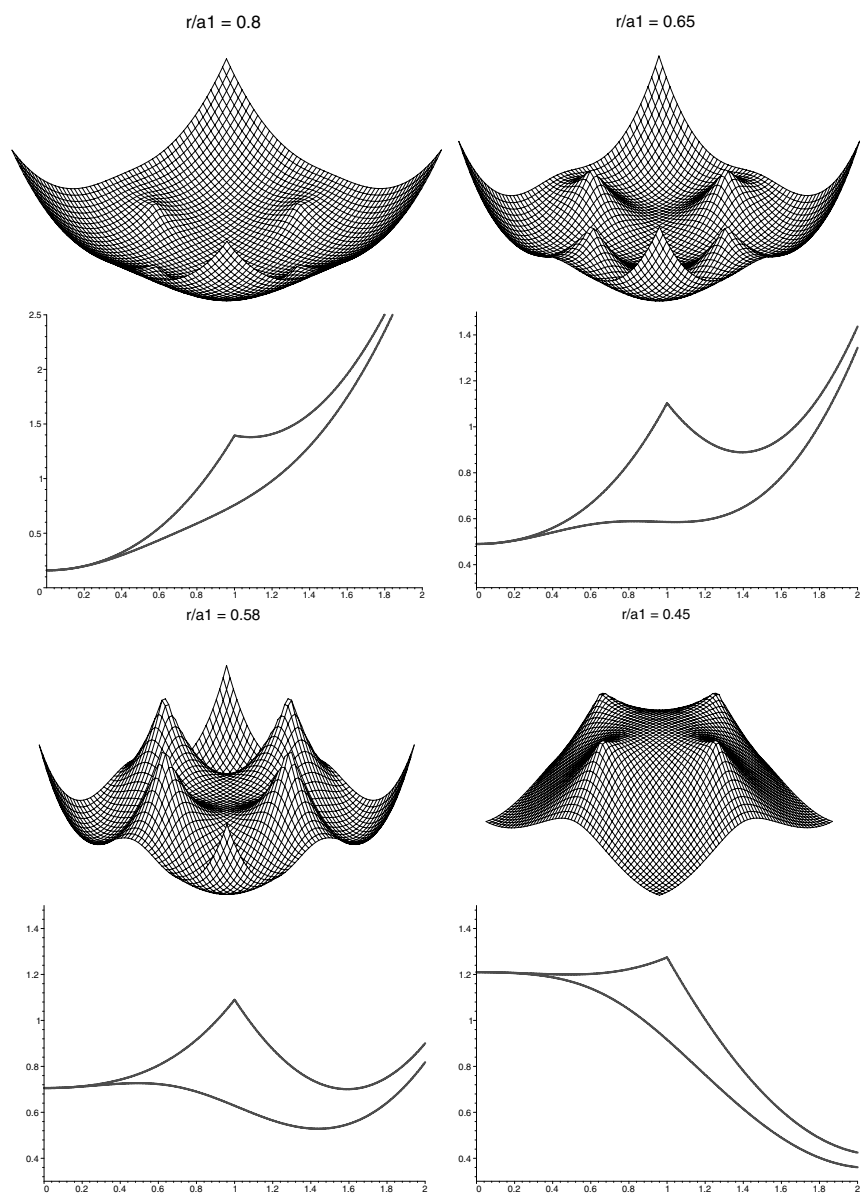


Fig. 6. The energy landscape $E_{\text{shift}}(d)$ as a function of the shift vector d (see Figure 5) illustrates the large complexity created by geometric nonlinearity alone—recall that the interaction forces are linear in interatomic distance. For $r/a_1 = 0.8$ and $r/a_1 = 0.65$, the CB state $d = 0$ is globally stable; for $r/a_1 = 0.58$ it is only locally stable; and for $r/a_1 = 0.45$ it has become locally unstable. The upper and lower curves give, respectively, the energy on the critical lines through zero in $(1, 0)$ direction (“trapezoids”) and $(1, 1)$ direction (“kites”), as a function of $|d|/a_1$.

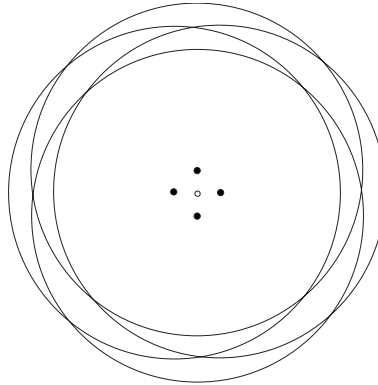


Fig. 7. Inadequacy of the CB rule for a 5-atom configuration under strong compression: If the black (boundary) atoms are moved inwards, the circles indicating the optimal position of the white (centre) atom with respect to each boundary atom converge to each other and become favourable compared with the symmetric centre position postulated by the CB rule.

Proof of Theorem 2.2 (b) and (c). To prove (b), take $r = r^*$ to infer from Proposition 7.1 that $W(F) < W_{CB}(F)$ if $F \in SO(2)$ and $\frac{r^*}{a_1} < C_*$. The latter condition is equivalent to $0 < \alpha < C_*$, $\kappa < \frac{1-C_*}{2(C_*-\alpha)}$, and this parameter region contains the region $0 < \alpha < 1/2, \kappa \geq \frac{1-1/2}{2(1/2-\alpha)}$ (or equivalently: $\frac{r^*}{a_1} \leq \frac{1}{2}$) in the theorem. This establishes (b). Finally, (c) is immediate from the proposition in the larger parameter region $r < C_*a_1$ in place of $r \leq a_1/2$. \square

Proof of Proposition 7.1. This is, “in principle,” only an elementary calculus exercise, but this very low-dimensional problem with simple interactions already requires a surprising amount of work. \square

Lemma 7.2. *For all parameter values $K_1, K_2, a_1, a_2 > 0$, and all $d \neq 0$, the first variation of $E_{shift}(d)$ with respect to polar angle is zero if and only if d lies on the four lines $d_1 = 0, d_2 = 0, d_1 + d_2 = 0, d_1 - d_2 = 0$. Moreover, if d lies on the two lines $d_1 = 0, d_2 = 0$, the second variation with respect to polar angle is negative (and hence d is angularly unstable).*

Remark. Angular criticality of points on these lines follows from the symmetry of E_{shift} under reflection of the unit cell at them; the point is that there are no further critical points. Geometrically, the unit cell of a critical point on the lines $d_1 \pm d_2 = 0$ is a kite, while the lines $d_1 = 0, d_2 = 0$ correspond to trapezoids.

Proof. As suggested by the above statement, we work in polar coordinates, $d = (r\rho \cos \theta, r\rho \sin \theta)$. The four sidelengths of the deformed unit cell $\{d, re_1, d + r(e_1 + e_2), re_2\}$ then become

$$r\sqrt{1 \pm 2\rho \cos \theta + \rho^2}, \quad r\sqrt{1 \pm 2\rho \sin \theta + \rho^2}.$$

Consequently,

$$\begin{aligned} E_{shift}(d) &= \frac{K_1}{4} \left(\sum_{\sigma \in \{\pm 1\}} (r\sqrt{1+2\sigma\rho \cos \theta + \rho^2} - a_1)^2 + \sum_{\sigma \in \{\pm 1\}} (r\sqrt{1+2\sigma\rho \sin \theta + \rho^2} - a_1)^2 \right) \\ &\quad + K_2(\sqrt{2}r - a_2)^2 \\ &= \frac{K_1 r a_1}{2} (f(\cos \theta) + f(\sin \theta)) + K_1(r^2(1 + \rho^2) + a_1^2) + K_2(\sqrt{2}r - a_2)^2, \end{aligned}$$

where

$$f(z) = -\sqrt{1 + 2\rho z + \rho^2} - \sqrt{1 - 2\rho z + \rho^2}.$$

The angular derivative is

$$\frac{\partial}{\partial \theta} E_{shift}(d) = \frac{K_1 r a_1}{2} (-\sin \theta f'(\cos \theta) + \cos \theta f'(\sin \theta)).$$

Now $f'(z) = \rho(1 - 2\rho z + \rho^2)^{-1/2} - \rho(1 + 2\rho z + \rho^2)^{-1/2}$, whence $f'(0) = 0$, $f'(z) > 0$ for $z > 0$, $f'(z) < 0$ for $z < 0$. Consequently $\frac{\partial}{\partial \theta} E_{shift}(d) = 0$ if and only if one of the following holds: (a) $\cos \theta = 0$, (b) $\sin \theta = 0$, (c) $\cos \theta, \sin \theta$ both $\neq 0$ and $\frac{f'(\cos \theta)}{\cos \theta} = \frac{f'(\sin \theta)}{\sin \theta}$. The proof of the lemma is complete if we can show that (c) implies $\cos \theta = \pm \sin \theta$. But this follows from the fact that $\frac{f'(z)}{z}$ is an even function which is strictly increasing with $z > 0$. The evenness is obvious and the strict monotonicity follows from the formula

$$\frac{d}{dz} \frac{f'(z)}{z} = \frac{\rho}{z^2} \left(\frac{1 + \rho^2 + 3\rho z}{\sqrt{1 + \rho^2 + 2\rho z}} - \frac{1 + \rho^2 - 3\rho z}{\sqrt{1 + \rho^2 - 2\rho z}} \right),$$

whose right-hand side is easily seen to be positive for $z > 0$.

Finally, the angular instability assertion follows from the calculations

$$\begin{aligned} \frac{\partial^2}{\partial \theta^2} E_{shift}(d) &= \frac{K_1 r a_1}{2} (-\cos \theta f''(\cos \theta) - \sin \theta f''(\sin \theta) \\ &\quad + \sin^2 \theta f''(\cos \theta) + \cos^2 \theta f''(\sin \theta)), \\ f''(z) &= \frac{\rho^2}{\sqrt{1 + 2\rho z + \rho^2}^3} + \frac{\rho^2}{\sqrt{1 - 2\rho z + \rho^2}^3}, \end{aligned}$$

whence

$$\begin{aligned} \frac{\partial^2}{\partial \theta^2} E_{\text{shift}}(d) \Big|_{\theta=k\pi/2, k \in \mathbb{Z}} &= \frac{K_1 r a_1}{2} (-f'(1) + f''(0)) \\ &= \frac{K_1 r a_1}{2} \left(\frac{\rho}{|1+\rho|} - \frac{\rho}{|1-\rho|} + \frac{2\rho^2}{\sqrt{1+\rho^2}} \right), \end{aligned}$$

which is easily seen to be negative for all $\rho > 0$. (Indeed, for $\rho \leq 1$ negativity is equivalent to the evident inequality $\sqrt{1+\rho^2} > 1-\rho^2$, and for $\rho > 1$ it is equivalent to the evident inequality $\sqrt{1+\rho^2} > \rho(\rho^2-1)$.) The proof of the lemma is complete. \square

Proof of Proposition 7.1. By the lemma, the set of global minimizers must be contained in the two lines $d_1 \pm d_2 = 0$. On these lines,

$$\begin{aligned} E_{\text{shift}}(d) - E_{\text{shift}}(0) &= K_1 r a_1 (f(\sqrt{2}) + 2) + K_1 r^2 \rho^2 \\ &= K_1 r a_1 \left(\frac{r}{a_1} \rho^2 - \left(\sqrt{1+\sqrt{2}\rho+\rho^2} - 2 + \sqrt{1-\sqrt{2}\rho+\rho^2} \right) \right). \end{aligned}$$

Hence for $\frac{r}{a_1} < C_*$, C_* as in Proposition 7.1, $d = 0$ is the unique global minimizer, while for $\frac{r}{a_1} = C_*$ (respectively $\frac{r}{a_1} < C_*$) any ρ achieving the maximum in the definition of C_* has equal (respectively lower) energy. (That such a maximizer exists is easy to show.) This establishes (a), (b), (c). Finally, as calculated earlier in Lemma 6.1 for $r = r^*$, $D^2 E_{\text{shift}}(0) = r^2 K_1 (2 - \frac{a_1}{r}) I$, establishing (d). \square

8. Conclusions

- (1) For our simple model, if the nearest neighbour and next-nearest neighbour distances are roughly compatible, the CB rule can be justified rigorously and the ground state subject to all linear boundary data close to the identity can be proven to be simply periodic.
- (2) For unfavourable parameters or large deformations, the CB rule fails. We found that the energy can be lowered by increasing the period. Moreover the overestimation of the true ground state energy by the CB rule is not cured by replacing the CB energy by its quasiconvex hull (or even its convex hull).
- (3) The transition to non-CB ground states is first order, not second order. In particular the parameter values and loads at which the instability occurs cannot be detected by local considerations (such as soft-phonon analysis, i.e., linearization at the CB state).
- (4) The next-nearest neighbour structure of our model was convenient but not essential for the construction of a local cell-energy function. For instance, if third-nearest neighbour interactions $\frac{1}{2} \sum_{x, x' \in \mathcal{L}, |x-x'|=2r^*} V^{(3)}(|y(x) - y(x')|)$ are included into the energy, and if the first and second neighbour potentials $\frac{K_i}{2} (|y(x) - y(x')| - a_i)^2$ ($i = 1, 2$) are replaced by more general potentials $V^{(i)}(|y(x) - y(x')|)$, analoga

of the partitioning formulae (3.1), (3.2) remain valid; see below for the technical details. These more general interactions in particular cover the model investigated numerically in [NS98].

Technical details of (4). We define $\mathcal{C} := [-r^*, r^*]^2 \cap r^*\mathbb{Z}^2$ and set

$$E_{cell}(\{y(x)\}_{x \in \mathcal{C}}) := \frac{1}{2} \sum_{x, x' \in \mathcal{C}} \sigma_{x, x'} V_{|x-x'|}(|y(x) - y(x')|),$$

where the interaction potential $V_\alpha(\rho)$ equals $V^{(1)}(\rho)$ when $\alpha = r^*$ (nearest neighbours), $V^{(2)}(\rho)$ when $\alpha = \sqrt{2}r^*$ (second neighbours), $V^{(3)}(\rho)$ when $\alpha = 2r^*$ (third neighbours), and zero otherwise, and the weight coefficients are

$$\sigma_{x, x'} = \begin{cases} 1 & , |x - x'| = 2, x \notin \text{vertices or } x' \notin \text{vertices,} \\ \frac{1}{2} & , |x - x'| = 2, x \in \text{vertices and } x' \in \text{vertices,} \\ \frac{1}{2} & , |x - x'| = \sqrt{2}, \\ \frac{1}{2} & , |x - x'| = 1, x \notin \text{edges or } x' \notin \text{edges,} \\ \frac{1}{4} & , |x - x'| = 1, x \in \text{edges and } x' \in \text{edges,} \\ 0 & \text{else.} \end{cases}$$

Then $E = E_b + E_s$ with bulk energy $E_b = \sum_{x \in \mathcal{L}'} E_{cell}(\{y(x')\}_{x' \in x + \mathcal{C}})$, where the sum is taken over the interior atoms $\mathcal{L}' = \{x \in \mathcal{L} \mid (x + [-2r^*, 2r^*]^2) \cap r^*\mathbb{Z}^2 \text{ is contained in } \mathcal{L}\}$ and the surface energy E_s depends only on the positions $y(x)$ of boundary atoms, i.e., x with $(x + [-3r^*, 3r^*]^2) \cap r^*\mathbb{Z}^2$ not contained in \mathcal{L} .

Appendix 1

We give here a proof of the nontrivial part of Lemma 2.1 (see Section 2) which asserts that as the system size gets large, the limit of lattice ground state energy divided by system volume exists and is independent of the shape of domain.

Proof of Lemma 2.1. The assertion is reminiscent of the well-known result in the theory of nonlinear partial differential equations that the quasiconvex envelope of a function $W: M^{m \times n} \rightarrow [0, \infty)$,

$$W_\Omega^{qc}(F) = \inf_{y(x)|_{\partial\Omega} = Fx} \frac{1}{|\Omega|} \int_\Omega W(Dy(x)) dx,$$

is independent of the (open, bounded, Lipschitz) domain Ω . The standard proof (e.g. [Da89, Section 5.1]) begins by noting that it suffices to show

$$W_{\Omega'}(F) \leq W_\Omega(F) \tag{8.1}$$

for any pair of domains Ω, Ω' . To prove this, one uses the translation- and dilation-invariance $W_\Omega = W_{x+\epsilon\Omega}$ and Vitali's covering theorem (which allows one to cover Ω' with scaled copies $x_j + \epsilon_j\Omega$ of Ω) to obtain a trial function in the variational principle for

$W_{\Omega'}$ consisting of scaled copies of an (exact or approximate) minimizer of the variational principle for W_{Ω} . A small technical difficulty arises when trying to combine this argument with the thermodynamic limit procedure arising for lattice systems. Instead of (8.1), one now needs to establish

$$\limsup_{R \rightarrow \infty} W_{\Omega', R}(F) \leq \liminf_{R \rightarrow \infty} W_{\Omega, R}(F),$$

where $W_{\Omega, R}$ is the ground state energy per unit volume of the finite lattice $\mathcal{L} = r^*\mathbb{Z}^2 \cap \Omega_R$ (see Section 2 for notation). But a priori the \liminf on the right is only known to be realized by a subset of scales $\{R_\alpha\}$. Therefore one is only allowed to use a subset of scales in the Vitali covering of Ω' by dilates of Ω . There are two obvious strategies to overcome this: (a) Prove a version of Vitali's theorem delivering coverings with controlled scales, (b) Compare with a special reference domain where the scales can be trivially controlled. Taking the viewpoint (b), we will prove Lemma 2.1 by establishing the following three statements:

- Upper bound by energy of cube $Q := (0, 1)^2$:

$$\limsup_{R \rightarrow \infty} W_{\Omega, R}(F) \leq \liminf_{R \rightarrow \infty} W_{Q, R}(F) \text{ for all } \Omega; \quad (8.2)$$

- Existence of thermodynamic limit for cube:

$$\lim_{R \rightarrow \infty} W_{Q, R}(F) \text{ exists}; \quad (8.3)$$

- Lower bound by energy of cube:

$$\liminf_{R \rightarrow \infty} W_{\Omega, R} \geq \lim_{R \rightarrow \infty} W_{Q, R}(F). \quad (8.4)$$

To prove (8.2), we begin by choosing a sequence $\ell_j \rightarrow \infty$ of integers realizing the \liminf for Q , i.e., satisfying $\lim_{\ell_j \rightarrow \infty} W_{Q, \ell_j r^*}(F) = \liminf_{R \rightarrow \infty} W_{Q, R}(F)$. Given any positive integer ℓ and any R , introduce an associated coarse lattice $\tilde{\mathcal{L}} = \{x_0 \in \ell r^*\mathbb{Z}^2 \cap \Omega_R \mid x_0 + \ell e_1, x_0 + \ell(e_1 + e_2), x_0 + \ell e_2 \in \Omega_R\}$. For $x_0 \in \tilde{\mathcal{L}}$, let $Q(x_0)$ be the cube with corners $x_0, x_0 + \ell e_1, x_0 + \ell(e_1 + e_2), x_0 + \ell e_2$. It follows that $\Omega_R = \bigcup_{x_0 \in \tilde{\mathcal{L}}} Q(x_0) \cup N$, where the remainder N satisfies

$$\#\{x \in r^*\mathbb{Z}^2 \cap \Omega_R \mid x \in N\} = O(R\ell)$$

(the factor R accounting for the surface area of $\partial\Omega_R$ and the factor ℓ accounting for the maximal distance of points in N from $\partial\Omega_R$). Consider now the trial function

$$y(x) = \begin{cases} Fx, & x \in N, \\ Fx_0 + y_\ell(x - x_0), & x \in Q(x_0), \end{cases}$$

where y_ℓ is the minimizer of E on $Q_{\ell r^*}$ subject to the boundary condition $y_\ell(x) = Fx$ on $\partial(r^*\mathbb{Z}^2 \cap Q_{\ell r^*})$. Then

$$\begin{aligned} W_{\Omega, R}(F) &\leq \frac{E[y_\ell]}{|\Omega_R|} = \frac{W_{Q, \ell r^*}(F)|\Omega_R \setminus N| + O(R\ell)}{|\Omega_R|} \\ &= W_{Q, \ell r^*}(F) \left(1 + O\left(\frac{\ell}{R}\right)\right) + O\left(\frac{\ell}{R}\right). \end{aligned}$$

(For longer-range forces without exact additivity with respect to unit cells, there would be an $O(1/\ell)$ term as well.) Now first let $R \rightarrow \infty$ at fixed ℓ , yielding $\limsup_{R \rightarrow \infty} W_{\Omega,R}(F) \leq W_{Q,\ell r^*}(F)$, for all $\ell \in \mathbb{N}$. Now choosing $\ell = \ell_j$, where ℓ_j is our sequence of integers realizing the lim inf for Q , and letting $\ell_j \rightarrow \infty$, we obtain (2.5).

As an immediate consequence we also obtain (8.3), by applying (8.2) to $\Omega = Q$.

It remains to establish the lower bound (8.4). Let Q_a be a cube with center 0 and sidelength a which contains Ω . Then Q_{aR} contains Ω_R , for every R . We claim that

$$W_{Q,aR}(F)|Q_{aR}| \leq W_{\Omega,R}(F)|\Omega_R| + W_{Q_a \setminus \Omega,R}(F)|Q_{aR} \setminus \Omega_R| + O(R).$$

This is because the functions y_1 and y_2 which minimize E on $r^*\mathbb{Z}^2 \cap \Omega_R$ respectively $r^*\mathbb{Z}^2 \cap (Q_{aR} \setminus \Omega_R)$ can be combined into a trial function

$$y(x) = \begin{cases} y_1(x), & x \in \Omega_R, \\ y_2(x), & x \in Q_{aR} \setminus \Omega_R, \end{cases}$$

for the variational principle for $W_{Q,aR}$. Dividing by $|\Omega_R|$, one has

$$W_{\Omega,R}(F) \geq W_{Q_a,R}(F) \frac{|Q_a|}{|\Omega|} - W_{Q_a \setminus \Omega,R}(F) \frac{|Q_a \setminus \Omega|}{|\Omega|} + O\left(\frac{1}{R}\right).$$

Consequently,

$$\liminf_{R \rightarrow \infty} W_{\Omega,R}(F) \geq \left(\liminf_{R \rightarrow \infty} W_{Q_a,R}(F) \right) \frac{|Q_a|}{|\Omega|} - \limsup_{R \rightarrow \infty} W_{Q_a \setminus \Omega,R}(F) \frac{|Q_a \setminus \Omega|}{|\Omega|}.$$

But by (8.2) and (8.3), the first term on the right-hand side equals $\lim_{R \rightarrow \infty} W_{Q,R}(F) \frac{|Q_a|}{|\Omega|}$, and the second term is bounded from below by $-\lim_{R \rightarrow \infty} W_{Q,R}(F) \frac{|Q_a \setminus \Omega|}{|\Omega|}$. This establishes (8.4) and completes the proof of Lemma 2.1. \square

Appendix 2

We make explicit a purely continuum-mechanical corollary of our sufficient condition for lattice polyconvexity in Lemma 5.3.

Corollary A1. *Let $\mathcal{D} = M^{2 \times 2}$ or $\mathcal{D} = \{F \in M^{2 \times 2} \mid \det F > 0\}$. Let $\varphi: \mathcal{D} \rightarrow \mathbb{R}$ be continuous, frame-indifferent (i.e., $\varphi(RF) = \varphi(F)$ for all $F \in \mathcal{D}$ and all $R \in SO(2)$), C^3 in a neighbourhood of $SO(2)$. In case $\mathcal{D} \neq M^{2 \times 2}$, define $\varphi(F) = +\infty$ for $F \notin \mathcal{D}$. Assume φ satisfies*

- (i) (single-well property) $\varphi(F) \geq 0$, $\varphi(F) = 0$ if and only if $F \in SO(2)$;
- (ii) (positivity of elasticity tensor) $\sum_{i,j,k,l=1}^2 \frac{\partial^2 \varphi}{\partial F_{ij} \partial F_{kl}}(I) A_{ij} A_{kl} > 0$ for all $A = A^T > 0$;
- (iii) (quadratic growth) $\liminf_{|F| \rightarrow \infty, F \in \mathcal{D}} \frac{\varphi(F)}{|F|^2} > 0$, $\limsup_{|F| \rightarrow \infty, F \in \mathcal{D}} \frac{\varphi(F)}{|F|^2} < \infty$.

Then φ agrees in an open neighbourhood U of $SO(2)$ with a polyconvex function $\varphi_0(F) = f(F, \det F)$, $\varphi_0: M^{2 \times 2} \rightarrow \mathbb{R}$ (respectively $\varphi_0: M^{2 \times 2} \rightarrow \mathbb{R} \cup \{+\infty\}$ in case $\mathcal{D} \neq M^{2 \times 2}$), which satisfies $\varphi_0 \leq \varphi$ and the following strictness property:

$$f(F, \det F) > t(F, \det F)$$

for all affine functions $t \leq f$ which agree with f at some point $(F_0, \det F_0) \neq (F, \det F)$, and all $F, F_0 \in \mathcal{D}$. In particular,

$$\varphi(F) = \varphi^{qc}(F) \text{ for all } F \in U, \quad (8.5)$$

and the variational problem occurring in the definition (2.10) of the quasiconvex hull of φ has the unique solution $y(x) = Fx$ for all $F \in U$ and all bounded Lipschitz domains $\Omega \subset \mathbb{R}^2$.

Proof. The assertions regarding polyconvexity are immediate from Lemma 5.3, by considering (in the notation of Section 5) the following function on $M^{2 \times 4}$: $\tilde{\varphi}(G') := \varphi(F)$, where F' is the orthogonal projection of G' onto the subspace given by (5.1), and F is related to F' as in (5.1). (Alternatively, the interested reader might wish to simplify the proof of Lemma 5.3 into a direct argument in $M^{2 \times 2}$.) The remaining assertions follow from the obvious continuum analogue of the calculation (5.4), which exhibits in two dimensions the general fact that polyconvexity implies quasiconvexity [Ba77], [Da87]. \square

Remark. We believe that the growth condition at infinity can be weakened.

Corollary A2. Recall the Cauchy-Born energy W_{CB} of our mass-spring model, defined in (2.6) and given explicitly by

$$\begin{aligned} W_{CB}(F) = & \frac{K_1}{2} \left((|Fe_1| - \frac{a_1}{r^*})^2 + (|Fe_2| - \frac{a_1}{r^*})^2 \right) \\ & + \frac{K_2}{2} \left((|F(e_1 + e_2)| - \frac{a_2}{r^*})^2 + (|F(e_1 - e_2)| - \frac{a_2}{r^*})^2 \right), \end{aligned}$$

when $\det F \geq 0$, and $W_{CB}(F) = +\infty$ otherwise. (Here r^* is as in (2.7). Also, recall from Proposition 3.1 that W_{CB} is minimized if and only if $F \in SO(2)$.) For any choice of the parameters $K_1, K_2, a_1, a_2 > 0$, W_{CB} agrees with its own quasiconvex hull in some open neighbourhood of $SO(2)$.

Proof. The function W_{CB} fulfils the hypotheses of Corollary A1, due to Proposition 3.1 and Lemma 6.1. \square

Physically, this result sharpens Corollary 2.3 in the parameter region where the CB rule fails. It says that allowing the continuum degrees of freedom to relax (after having eliminated the discrete degrees of freedom by the CB rule) does not just fail to lower the CB energy to the level of the correct lattice energy; in fact it fails to lower the CB energy at all.

References

- [Ba77] J. M. Ball, Convexity conditions and existence theorems in nonlinear elasticity, *Arch. Ratl. Mech. Analysis* 63, 337–403, 1977.
- [BDG99] A. Braides, G. Dal Maso, A. Garroni, Variational formulation of softening phenomena in fracture mechanics: The one-dimensional case, *Arch. Ratl. Mech. Analysis* 146, 23–58, 1999.
- [CLL98] I. Catto, C. Le Bris, P.-L. Lions, *The Mathematical Theory of Thermodynamic Limits: Thomas-Fermi Type Models*, Oxford University Press, New York, 1998.
- [Da89] B. Dacorogna, *Direct Methods in the Calculus of Variations*, Springer-Verlag, Berlin, 1989.
- [Fe85] C. Fefferman, The thermodynamic limit for a crystal, *Commun. Math. Physics* 98, 289–311, 1985.
- [FJ00] G. Friesecke, R. D. James, A scheme for the passage from atomic to continuum theory for thin films, nanotubes and nanorods, *J. Mech. Phys. Solids* 48, 1519–1540, 2000.
- [FT02] G. Friesecke, F. Theil, in preparation, 2002.
- [Ge99] M. S. Gelli, Variational limits of discrete systems, Ph.D. thesis, Scuola Internazionale Superiore die Studi Avanzati, Trieste, 1999.
- [Go98] M. Gobbino, Finite difference approximation of the Mumford-Shah functional, *Commun. Pure Appl. Math.* 51, 197–228, 1998.
- [KS84] R. Knops, C. A. Stuart, Quasiconvexity and uniqueness of equilibrium solutions in nonlinear elasticity, *Arch. Ratl. Mech. Analysis* 86, no. 3, 233–249, 1984.
- [LS77] E. H. Lieb, B. Simon, The Thomas-Fermi theory of atoms, molecules and solids, *Adv. Math.* 23, 22–116.
- [Mo52] C. B. Morrey, Quasi-convexity and the lower semicontinuity of multiple integrals, *Pacific J. Math.* 2, 25–53, 1952.
- [NS98] J. Novak, E. K. H. Salje, Simulated mesoscopic structures of a domain wall in a ferroelastic lattice, *Eur. Phys. J. B* 4, 279–284, 1998.
- [PP02] S. Pagano, R. Paroni, A simple model for phase transitions: From the discrete to the continuum problem, to appear in *Quart. Appl. Math.*, 2002.
- [Tr96] L. Truskinovsky, Fracture as a phase transition, in R. C. Batra and M. F. Beatty (eds.), *Contemporary Research in the Mechanics and Mathematics of Materials*, CIMNE, Barcelona, 322–332, 1996.