The origin of nucleation peak in transformational plasticity

Lev Truskinovsky[a], Anna Vainchtein[b],*

[a] *Laboratoire de Mechanique des Solides, CNRS-UMR 7649, Ecole Polytechnique, 91128, Palaiseau, France*
[b] *Department of Mathematics, University of Pittsburgh, Pittsburgh, PA 15260, USA*

Abstract

A typical stress-strain relation for martensitic materials exhibits a mismatch between the nucleation and propagation thresholds leading to the formation of the nucleation peak. We develop an analytical model of this phenomenon and obtain specific relations between the macroscopic parameters of the peak and the microscopic characteristics of the material. Although the nucleation peak appears in the model as an interplay between discreteness and nonlocality, it does not disappear in the continuum limit. We verify the quantitative predictions of the model by comparison with experimental data for cubic to monoclinic phase transformation in NiTi.
© 2003 Elsevier Ltd. All rights reserved.

1. Introduction

In displacement-controlled experiments shape memory alloys and other martensitic materials display a nucleation peak: prior to nucleation of a new phase the load reaches a maximum but then drops to a distinctly lower value. The subsequent plateau is associated with phase boundaries propagating along the specimen at essentially constant stress (Lexcellent and Tobushi, 1995; Shaw and Kyriakides, 1995, 1997b; Sun and Zhong, 2000). The nucleation peak is not unique to transformational plasticity and is also observed during initiation of a conventional plastic deformation in mild steels (Butler, 1962; Hall, 1970; Kyriakides and Miller, 2000; Froli and Royer-Carfagni, 2000), where

---

* Corresponding author.
  E-mail addresses: trusk@lms.polytechnique.fr (Lev Truskinovsky), aav4@pitt.edu (Anna Vainchtein).

it has been attributed to the fact that the stress required to release the trapped disloca-tions is higher than the stress needed to sustain their motion (Cottrell and Bilby, 1949; Johnston and Gilman, 1959). For martensitic materials the nucleation-induced load drop has been observed in 3D numerical simulations based on various plasticity-like phe-nomenological models (Shaw and Kyriakides, 1997a; Kyriakides and Miller, 2000; Sun and Zhong, 2000) but due to the complexity of these models, the physical parameters responsible for the size of the peak have not been identified. At a qualitative level, the nucleation peak in these materials has been associated with the presence of sufficiently fine grains and heuristically linked to the strong locking of phase boundaries and the relative ease of their glide upon release (Shaw and Kyriakides, 1995, 1997b).

In this paper we develop an analytical model of the nucleation peak phenomenon in martensites: our model supports the intuition developed in plasticity theory and adapts it to the case when the principal carriers of inelastic deformation are phase boundaries. Specifically, we consider a prototypical mass-spring system consisting of rigid ele-ments (crystal planes) connected by bi-stable elastic springs representing transforming shear layers. To mimic the three-dimensional nature of the actual problem, we comple-ment the up-down-up interactions between the nearest neighbors (NN) by a harmonic interaction of the next-to-nearest neighbors (NNN). The bi-stable discrete models with-out NNN interactions (e.g. Müller and Villaggio, 1977; Fedelich and Zanzotto, 1992; Puglisi and Truskinovsky, 2000, 2002a, b) capture many important features of transfor-mational plasticity but fail to predict the peak phenomenon. Recent numerical studies of the models incorporating NNN interactions showed that the nucleation peak can be recovered (e.g. Ye et al., 1991; Triantafyllidis and Bardenhagen, 1993; Rogers and Truskinovsky, 1997; Froli and Royer-Carfagni, 2000; Pagano and Paroni, 2003); none of these models, however, have been developed analytically to the extent that they could explain the necessity of the peak phenomenon and identify the microparameters controlling the size of the stress drop.

We begin by finding the limits of instability of a homogeneous state and deter-mine an analytical expression for the nucleation threshold in the most general case. In order to obtain an analytical characterization of the propagation threshold we use a piecewise linear approximation for the NN interactions. This simplification allows us to reconstruct the non-equilibrium Peierls-Nabarro landscape for the propagating phase boundaries and compute the martensitic analog of the Peierls stress. We show that the presence of nonlocal interactions in the discrete model makes nucleation and propagation thresholds different and then prove that the nucleation peak does not dis-appear in the continuum limit. An important question is whether the quasi-continuum strain-gradient approximation of the discrete model (Mindlin, 1965; Triantafyllidis and Bardenhagen, 1993) is capable of capturing the peak phenomenon. We show that the nucleation threshold is approximated well only in the case of long-wave instability (see also Triantafyllidis and Bardenhagen, 1996) and that although the peak is captured, the stress drop is grossly exaggerated. Finally, we verify the quantitative prediction of the model by using the experimental data for cubic to monoclinic phase transformation in NiTi wires. The comparison with experiment leads to the bounds for the measure of nonlocality which are in a good agreement with the independent estimates based on a realistic interatomic potential.

The proposed model stays in the same prototypical relation to transformational plasticity as the well known Frenkel-Kontorova model to classical plasticity. A nontrivial formal correspondence between the two models in the case of infinite systems has been established by Truskinovsky and Vainchtein (2003).

The structure of the paper is as follows. In Section 2 we formulate the discrete problem for a finite system and specify the boundary conditions. In Section 3 we analyze stability of the single-phase equilibrium in the general setting discovering the possibility of both macro and microinstabilities. In Section 4 we introduce a piecewise linear model for the local interactions and obtain an explicit representation of a generic metastable equilibrium state. In Section 5 we study the energy barriers between the neighboring metastable states and show that propagation always reduces to a succession of transformations inside individual elements, while nucleation can involve transformation of several elements at once. Section 6 contains the derivation of the explicit formulae for the size of the nucleation peak in both discrete and continuum problems. In Section 7 we establish a correspondence between the microparameters of the lattice and the experimental measurements of both the stress drop and the size of the nucleation band. Finally, Section 8 contains comparison of the discrete and strain-gradient models showing that the agreement is at most qualitative. The conclusions are summarized in Section 9.

2. The model

Consider a system of $N + 1$ particles linked to their nearest and next-to-nearest neighbors by elastic springs (see Fig. 1). Let u_k, $0 \leqslant k \leqslant N$, be the displacements of the particles with respect to a load-free homogeneous reference configuration with spacing ε. Denote the strain in the kth NN spring by $w_k = (u_k - u_{k-1})/\varepsilon$. Then the total energy of the system can be written as

$$\Psi = \varepsilon \sum_{k=1}^{N} \phi_1(w_k) + 2\varepsilon \sum_{k=1}^{N-1} \phi_2 \left(\frac{w_{k+1} + w_k}{2} \right) + \Psi_B(w_1, w_N), \tag{1}$$

where $\phi_1(w)$ and $\phi_2(w)$ are the energy densities of the NN and NNN interactions, respectively. The term Ψ_B corresponds to the energy of the boundary elements.
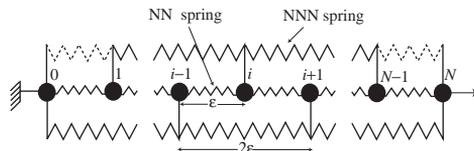

Fig. 1. A finite chain of particles with nearest-neighbor (NN) and next-to-nearest-neighbor (NNN) interactions.

We assume that the chain is placed in a hard device with the total displacement d:

$$u_N - u_0 = \varepsilon \sum_{k=1}^{N} w_k = d. \tag{2}$$

Due to the nonlocality of the model, the boundary condition (2) must be complemented by other constraints. For instance, to mimic an "extra hard device" one can additionally impose conditions $w_1 = 0$ and $w_N = 0$, making the term Ψ_B irrelevant. Alternatively, one may consider a "zero-moment device" by assuming that $\Psi_B = 0$.

Let $f_i(w) = \phi_i'(w)$, $i = 1, 2$, denote the forces in NN and NNN springs, respectively. The equations governing the equilibrium of the interior particles with $2 \leqslant k \leqslant N - 1$ have the form

$$f_1(w_k) + f_2\left(\frac{w_{k+1} + w_k}{2}\right) + f_2\left(\frac{w_k + w_{k-1}}{2}\right) = F, \tag{3}$$

where F is the total force in the system. The natural boundary conditions read

$$f_1(w_1) + f_2\left(\frac{w_2 + w_1}{2}\right) + \frac{1}{\varepsilon}\frac{\partial \Psi_B}{\partial w_1} = F,$$

$$f_1(w_N) + f_2\left(\frac{w_N + w_{N-1}}{2}\right) + \frac{1}{\varepsilon}\frac{\partial \Psi_B}{\partial w_N} = F, \tag{4}$$

An additional assumption adopted in what follows,

$$\Psi_B = \varepsilon\phi_2(w_1) + \varepsilon\phi_2(w_N), \tag{5}$$

means that the boundary NNN springs are cut in half and reconnected parallel to the NN springs (see the dashed springs in Fig. 1). The formal advantage of this choice of Ψ_B is that the boundary equations (4) can be included into the bulk equations (3) if we additionally assume the existence of fictitious 0th and $(N+1)$th springs satisfying

$$w_0 = w_1, \quad w_{N+1} = w_N. \tag{6}$$

The real advantage of (4) and (5), however, is that the corresponding boundary conditions ensure the existence of a trivial solution with the uniform strain distribution (Triantafyllidis and Bardenhagen, 1993; Charlotte and Truskinovsky, 2002). While in the rest of the paper we will be using mostly (5), the effect of switching to $\Psi_B = 0$ is briefly discussed in Section 6.

3. Nucleation threshold

It is easy to see that for arbitrary spring potentials the trivial solution of the problem (3), (4) and (5) is given by $w_k = d/L$, where $L = N\varepsilon$. The homogeneous response of the system is then characterized by the formulae $F = f_1(d/L) + 2f_2(d/L)$ and $\Psi = L(\phi_1(d/L) + 2\phi_2(d/L))$. To analyze the stability of this solution we introduce the tangential moduli of NN and NNN springs: $K(w) = \phi_1''(w)$ and $\gamma(w) = \phi_2''(w)/2$.

The homogeneous configuration is stable if and only if the quadratic form $\mathbf{Bv} \cdot \mathbf{v}$ with

$$
\mathbf{B} = \begin{bmatrix}
K+3\gamma & \gamma & 0 & \cdots & & 0 \\
\gamma & K+2\gamma & \gamma & \ddots & & 0 \\
0 & \ddots & \ddots & \ddots & & 0 \\
\vdots & 0 & \gamma & K+2\gamma & & \gamma \\
0 & \cdots & 0 & \gamma & & K+3\gamma.
\end{bmatrix}
\tag{7}
$$

is positive definite for all $\mathbf{v} \neq \mathbf{0}$ such that

$$
\sum_{k=1}^{N} v_k = 0.
\tag{8}
$$

Finding the boundaries of stability in the space of parameters is important for the subject of the paper because they coincide with the limits of a barrierless nucleation.

Before addressing the stability problem systematically, we observe that $K + 4\gamma > 0$ and $\gamma \leqslant 0$ are sufficient for stability. Indeed, in this case all terms in the quadratic form

$$
\mathbf{Bv} \cdot \mathbf{v} = (K+4\gamma) \sum_{k=1}^{N} v_k^2 - \gamma \sum_{k=1}^{N-1} (v_{k+1} - v_k)^2.
\tag{9}
$$

are nonnegative. Similarly, by writing the quadratic form as

$$
\mathbf{Bv} \cdot \mathbf{v} = K \sum_{k=1}^{N} v_k^2 + \gamma \sum_{k=1}^{N-1} (v_{k+1} + v_k)^2 + 2\gamma(v_1^2 + v_N^2).
$$

we obtain that conditions $K > 0$ and $\gamma \geqslant 0$ are also sufficient for stability. In terms of the main nondimensional parameter of the problem,

$$
\mu = \frac{K}{4\gamma},
\tag{10}
$$

these stability intervals can be written as $-\infty < \mu < -1$ and $0 < \mu < \infty$ and hence the instability limits are located in the interval

$$
-1 \leqslant \mu \leqslant 0.
\tag{11}
$$

To find the exact locations of the stability boundaries, consider the (zero) eigenvalue problem $\mathbf{Bv} = \mathbf{0}$. In the bulk of the chain ($1 \leqslant k \leqslant N-1$) this means

$$
(K+2\gamma)v_k + \gamma v_{k-1} + \gamma v_{k+1} = 0.
\tag{12}
$$

On the boundaries we obtain

$$
(K+3\gamma)v_1 + \gamma v_2 = 0,
\tag{13}
$$

$$
(K+3\gamma)v_N + \gamma v_{N-1} = 0.
\tag{14}
$$

The eigenvector $\mathbf{v}$ must be nonzero and satisfy (8). We seek solution in the form $v_k = \rho^k$ and obtain the characteristic polynomial

$$\rho^2 + (4\mu + 2)\rho + 1 = 0. \tag{15}$$

This means that $\rho_{1,2} = -1 - 2\mu \pm 2\sqrt{\mu(\mu + 1)}$, and according to (11), we have to consider the following cases:

Special case $\mu = -1$: Here $\rho_1 = \rho_2 = 1$ and $v_k = A_1 k + A_2$. The boundary equations (13) and (14) reduce to $v_1 = v_2$ and $v_N = v_{N-1}$, implying that $A_1 = 0$ and thus $v_k = A_2$. The constraint (8) implies that $A_2 = 0$. Since there are no nontrivial solutions, there is also no stability change.

Special case $\mu = 0$: Here $\rho_1 = \rho_2 = -1$ and $v_k = (A_1 k + A_2)(-1)^k$. Applying (13) and (14), we obtain $A_1 = A_2 = 0$ which again means that there is no change of stability.

Generic case $-1 < \mu < 0$. We first observe that in this interval $\rho_{1,2}$ can be written as

$$\rho_{1,2} = e^{\pm i\omega},$$

where ω is defined by

$$\mu = -\cos^2 \frac{\omega}{2}. \tag{16}$$

Applying the boundary equation (13), we obtain (up to a multiple)

$$v_k = \cos[(k - 1/2)\omega].$$

The second boundary equation (14) implies $2 \sin(\omega/2) \sin(N\omega) = 0$, and since $0 < \omega < \pi$, we obtain $\sin(N\omega) = 0$ and

$$\omega = \frac{\pi n}{N}, \tag{17}$$

for $1 \leqslant n \leqslant N - 1$. All these solutions are nontrivial and satisfy (8). Therefore the instability of the trivial solution can take place at any of the bifurcation points

$$K + 4\gamma \cos^2 \frac{\pi n}{2N} = 0.$$

The corresponding unstable modes are

$$v_k = \cos[(k - \tfrac{1}{2})\tfrac{\pi n}{N}]. \tag{18}$$

To locate the stability boundary, we begin with the case $\gamma < 0$. Then

$$K + 4\gamma \cos^2 \frac{\pi n}{2N} \geqslant K + 4\gamma \cos^2 \frac{\pi}{2N}$$

and hence

$$K + 4\gamma \cos^2 \frac{\pi}{2N} > 0 \tag{19}$$

is both necessary and sufficient for stability. The instability develops through the growth of the long-wave mode

$$v_k = \cos[(k - \tfrac{1}{2})\tfrac{\pi}{N}].$$

As N tends to infinity, the unstable wave length also becomes infinite and (19) reduces to $K + 4\gamma > 0$. At finite N we have $K + 4\gamma \cos^2 \frac{\pi}{2N} > K + 4\gamma$, meaning that the discrete homogeneous configuration may still be stable when the macroscopic modulus $E = K + 4\gamma$ is already negative. One can say that in the discrete problem the *macroscopic* instability is delayed due to the finite size of the system (see also Triantafyllidis and Bardenhagen, 1993).

The situation is different when $\gamma > 0$. In this case

$$K + 4\gamma \cos^2 \frac{\pi n}{2N} \geqslant K + 4\gamma \cos^2 \frac{\pi(N-1)}{2N} = K + 4\gamma \sin^2 \frac{\pi}{2N},$$

and hence the necessary and sufficient condition for stability is

$$K + 4\gamma \sin^2 \frac{\pi}{2N} > 0. \tag{20}$$

The instability develops through the growth of the short-wave mode

$$v_k = (-1)^k \sin\left[\left(k - \tfrac{1}{2}\right)\tfrac{\pi}{N}\right].$$

In the limit of infinite N, the wave length approaches the interatomic distance. Also (20) becomes $K > 0$, and since now $K + 4\gamma \sin^2(\pi/2N) < K + 4\gamma$, this *microscopic* instability develops before the macroscopic (or homogenized) system becomes unstable (see also Triantafyllidis and Bardenhagen, 1996). In the context of martensitic phase transitions, this effect may be linked to the observation of the pre-martensitic tweed microstructures (Kartha et al., 1995).

We now summarize the necessary and sufficient conditions for stability of the trivial solution. In terms of elastic moduli K and γ we obtain the following stability intervals

$$K + 4\gamma \cos^2 \frac{\pi}{2N} > 0, \quad K + 4\gamma \sin^2 \frac{\pi}{2N} > 0. \tag{21}$$

In the limit of infinite N the inequalities (21) reduce to the known conditions $K > 0$, $K + 4\gamma > 0$ (e.g. Mindlin, 1965). In terms of the nondimensional parameter μ the stability intervals take the form

$$\mu < -\cos^2 \frac{\pi}{2N} \ (\gamma < 0), \quad \mu > -\sin^2 \frac{\pi}{2N} \ (\gamma > 0), \tag{22}$$

which in the limit of infinite N gives $\mu < -1$ ($\gamma < 0$) and $\mu > 0$ ($\gamma > 0$). The structure of the stability domain in the plane $K - \gamma$ is illustrated in Fig. 2a for both finite and infinite N; in Fig. 2b, showing the plane $\mu - N^{-1}$, we illustrate the dependence of the stability boundaries on the size of the system. For the NNN system with the "moment free" boundary conditions ($\Psi_B = 0$) similar analysis of the stability for trivial equilibria (which now have boundary layers) can be found in Charlotte and Truskinovsky (2002).

4. Nontrivial solutions and metastability

To model martensitic phase transitions, we assume that each NN spring has a double-well energy generating a non-monotone "up-down-up" force-strain relation depicted in Fig. 3. If the negative slope of the force-strain relation for the NN spring is sufficiently steep, it is easy to show that the total strain can reach the threshold where
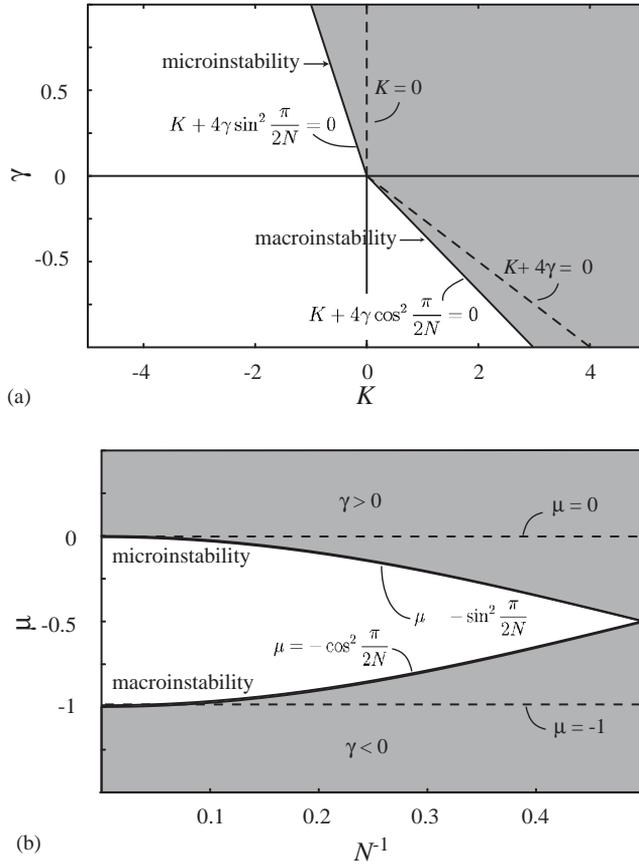
Fig. 2. (a) Domain of stability for the trivial solution in $K - \gamma$ plane: solid lines—discrete model, $N = 3$; dashed lines—continuum limit, $N \to \infty$. (b) The dependence of the stability limits on the size of the system. In both figures stability domain is in gray.
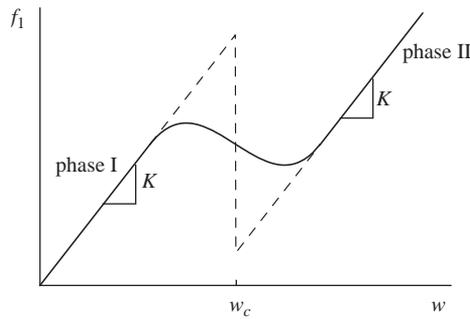

Fig. 3. Up-down-up force-strain relation in an individual NN spring (solid line) and its bilinear approximation (dashed lines).

the homogeneous phase is absolutely unstable. To study the resulting nucleation and the subsequent growth of the new phase, we need to describe the postcritical behavior of the system. More precisely, in the domain of loadings where trivial solution is unstable, we need to find the nontrivial solutions of the equilibrium equations (3) corresponding to either local or global minima of the energy.

Although the instability limits for the trivial solution can be found in the general case, we succeeded in obtaining a complete analytical description of the postcritical behavior only in the case when the functions $f_1(w)$ and $f_2(w)$ are either linear or piecewise linear. Specifically, we considered

$$f_1(w) = \begin{cases} Kw & \text{for } w < w_c \\ K(w - a) & \text{for } w \geqslant w_c \end{cases} \tag{23}$$

and

$$f_2(w) = 2\gamma w. \tag{24}$$

The parameters of the NN potential are the critical strain w_c, transformation strain a and elastic modulus K (see Fig. 3). The linear NNN spring is characterized by the elastic modulus γ which will also be used as the measure of nonlocality. We assume that $K > 0$ and $\gamma < 0$, which is suggested by the linearization of the Lennard-Jones potential (see Charlotte and Truskinovsky, 2002); to ensure that independently of N the homogeneous states are stable in their domain of definition, we also assume that $K + 4\gamma > 0$. In terms of the nondimensional parameter μ, the assumptions on the moduli can be summarized as

$$-\infty < \mu < -1. \tag{25}$$

It is not hard to see that under these assumptions the microinstability is excluded while the macroinstability always takes place at $w = w_c$.

For the piecewise linear model (23), (24) the total energy (1) reduces to

$$\Psi = \varepsilon \left\{ \frac{1}{2} \mathbf{B}\mathbf{w} \cdot \mathbf{w} - \mathbf{q} \cdot (\mathbf{w} - \mathbf{w_c}) \right\}, \tag{26}$$

where the vector $q_k = Ka\theta(w_k - w_{ck})$ prescribes distribution of phases, $\theta(x)$ is a unit step function and $w_{ck} = w_c$, $k = 1, \ldots, N$. The equilibrium equations (3) together with the boundary conditions (6) can then be rewritten as

$$\mathbf{B}\mathbf{w} = \mathbf{F} + \mathbf{q}, \tag{27}$$

where $F_k = F$, $k = 1, \ldots, N$. To eliminate the redundant parameters we rescale the variables. By selecting $L = N\varepsilon$ as the length scale and K as the scale of force, we define

$$\bar{u}_k = \frac{u_k}{N\varepsilon}, \quad \bar{d} = \frac{d}{N\varepsilon}, \quad \bar{F} = \frac{F}{K}, \quad \bar{\Psi} = \frac{\Psi}{KN\varepsilon}, \quad \bar{\mathbf{B}} = \frac{1}{K}\mathbf{B}. \tag{28}$$

Unless specially mentioned, in what follows, we will be using only rescaled variables with the bars dropped. The dimensionless problem depends on the two main parameters: μ and N.

To find the metastable states we can simply prescribe the distribution of phases $\mathbf{q}$ and solve the linear problem (27). The resulting equilibrium configurations are automatically local minimizers of the energy (26), because the stiffness matrix $\mathbf{B}$ is positive definite. To obtain the global minima of the energy, we need additionally to minimize Ψ with respect to the phase geometry $\mathbf{q}$.

The solution of the linear system (27) can be written as

$$\mathbf{w} = \mathbf{B}^{-1}\mathbf{F} + \mathbf{B}^{-1}\mathbf{q} = \mathbf{w}^0 + \mathbf{w}^1. \tag{29}$$

The first term in (29) corresponds to the uniform configuration with all springs in the first phase. Without explicitly computing $\mathbf{B}^{-1}$ one can see that

$$w_k^0 = \frac{F\mu}{1+\mu}. \tag{30}$$

It is convenient to express $\mathbf{w}^1$ not in terms of the variables $\mathbf{q}$, but in terms of their discrete derivatives $\mathbf{p}$ defined by

$$p_i = \frac{q_{i+1} - q_i}{a}. \tag{31}$$

Notice that these relations can always be explicitly inverted, yielding

$$q_k = q_1 + a \sum_{i=1}^{N} p_i \theta(k - i - 1). \tag{32}$$

The physical meaning of the variables p_i is clear from the representation

$$p_i = \begin{cases} 1 & \text{if } w_i < w_c \text{ and } w_{i+1} > w_c \text{ (I to II phase switch)} \\ 0 & \text{if } \operatorname{sign}(w_i - w_c) = \operatorname{sign}(w_{i+1} - w_c) \text{ (no phase switch)} \\ -1 & \text{if } w_i > w_c \text{ and } w_{i+1} < w_c \text{ (II to I phase switch).} \end{cases} \tag{33}$$

To simplify the subsequent formulae we also set $p_N = 0$.

In the interior nodes ($2 \leqslant k \leqslant N - 1$) w_k^1 must satisfy the difference equation

$$\left(1 + \frac{1}{2\mu}\right) w_k^1 + \frac{1}{4\mu}(w_{k+1}^1 + w_{k-1}^1) = q_1 + a \sum_i p_i \theta(k - i - 1). \tag{34}$$

We seek the general solution of (34) in the form

$$w_k^1 = w_k^h + w_k^{in}, \tag{35}$$

where w_k^h satisfies the homogeneous equation and w_k^{in} is a particular solution. The general solution of the homogeneous problem can be represented (Mickens, 1990) as a linear combination of ρ_1^k and ρ_2^k, where $\rho_{1,2}$ are the roots of the characteristic polynomial (15). By writing $\rho_{1,2} = e^{\pm\lambda}$, where

$$\lambda = 2\operatorname{arccosh}\sqrt{|\mu|}, \tag{36}$$

we obtain

$$w_k^h = C_1 e^{\lambda k} + C_2 e^{-\lambda k}. \tag{37}$$

The constants C_1 and C_2 are to be found from the boundary conditions (6).

To obtain a particular solution of (34) with $q_1 = 0$ and $p_j = 0$ for $j \neq i$, we observe that in this case the right hand side of (34) is zero for $k \leqslant i$ and constant for $k \geqslant i+1$. Therefore we can write

$$
w_k^{p_i} =
\begin{cases}
w_k^{p-} = A_1 e^{\lambda k} + A_2 e^{-\lambda k} & \text{for } k \leqslant i \\[2mm]
w_k^{p+} = A_3 e^{\lambda k} + A_4 e^{-\lambda k} + \dfrac{a\mu p_i}{1+\mu} & \text{for } k \geqslant i,
\end{cases}
\tag{38}
$$

and then "glue" the two sides together by requiring that

$$
w_i^{p-} = w_i^{p+}, \qquad w_{i+1}^{p-} = w_{i+1}^{p+}.
\tag{39}
$$

Since we are looking for a particular solution, we may always let $A_1 = A_2 = 0$, and thus consider $w_k^{p-} = 0$. Solving (39) for A_3 and A_4, we obtain

$$
w_k^{p_i} = \Delta \left\{ \theta(k - i - 1/2) \left[1 - \frac{\cosh[(k - i - 1/2)\lambda]}{\cosh(\lambda/2)} \right] \right\},
$$

where

$$
\Delta = \frac{a\mu}{\mu + 1}
\tag{40}
$$

is the macroscopic transformation strain. Finally, by superposition, we obtain the particular solution in the form

$$
w_k^{in} = \frac{q_1 \mu}{1 + \mu} + \Delta \sum_{i=1}^{N} p_i \theta(k - i - 1/2) \left[1 - \frac{\cosh[(k - i - 1/2)\lambda]}{\cosh(\lambda/2)} \right].
\tag{41}
$$

The general solution of (34) is now given by (35), (37) and (41). Applying the boundary conditions (6), we obtain

$$
w_k^1 = \frac{q_1 \mu}{1 + \mu} + \Delta \sum_{i=1}^{N} p_i \left\{ \frac{\sinh[(N - i)\lambda] \cosh[(k - 1/2)\lambda]}{\cosh(\lambda/2) \sinh(N\lambda)} \right.
$$

$$
\left. + \theta(k - i - 1/2) \left(1 - \frac{\cosh[(k - i - 1/2)\lambda]}{\cosh(\lambda/2)} \right) \right\}.
\tag{42}
$$

Before combining (30) and (42), we notice that the relation between the force F and the total displacement d can be written in the form:

$$
F = \frac{1 + \mu}{\mu} \left(d - \Delta \frac{l}{N} \right),
\tag{43}
$$

where l is the number of springs in phase II related to q_1 and $\mathbf{p}$ via

$$
l = \frac{N q_1}{a} + \sum_{i=1}^{N} (N - i) p_i.
\tag{44}
$$

Finally, substituting (43), (44) into (29), (30) and (42), we obtain the representation for a generic metastable configuration

$$w_k = d + \Delta \sum_{i=1}^{N} p_i \left\{ \frac{\sinh[(N-i)\lambda]\cosh[(k-1/2)\lambda]}{\cosh(\lambda/2)\sinh(N\lambda)} + \frac{i}{N} - 1 \right.$$
$$\left. + \theta(k-i-1/2)\left(1 - \frac{\cosh[(k-i-1/2)\lambda]}{\cosh(\lambda/2)}\right) \right\}. \tag{45}$$

To compute the energy we introduce a new variable $\tilde{\mathbf{w}}$ with components $\tilde{w}_k = w_k - d$ representing the **p**-dependent part of the strain field $\tilde{\mathbf{w}}$. Then (26) can be rewritten as

$$\Psi = \frac{1+\mu}{2\mu} d^2 - a(d - w_c)\frac{l}{N} + \tilde{\Psi}, \tag{46}$$

where

$$\tilde{\Psi} = \frac{1}{N}\left(\frac{1}{2}\mathbf{B}\tilde{\mathbf{w}} - \mathbf{q}\right)\cdot\tilde{\mathbf{w}} \tag{47}$$

is the contribution due to phase changes. Observing that in equilibrium

$$\tilde{\Psi} = -\frac{1}{2N}\mathbf{q}\cdot\tilde{\mathbf{w}}$$

and using (32) to eliminate **q**, we obtain

$$\Psi = \frac{1+\mu}{\mu}\left\{\frac{1}{2}d^2 - \Delta(d - w_c)\frac{l}{N} + \frac{\Delta^2}{2N}\mathbf{Jp}\cdot\mathbf{p}\right\}, \tag{48}$$

where the kernel **J** is given by

$$J_{ki} = \frac{\sinh[(N-i)\lambda]\sinh[\lambda k]}{\sinh(N\lambda)\sinh\lambda} + k\left(\frac{i}{N} - 1\right)$$
$$+ \theta(k-i-1)\left[k-i - \frac{\sinh[(k-i)\lambda]}{\sinh\lambda}\right]. \tag{49}$$

For a given distribution of phases the loading parameter d cannot take arbitrary values since the strains must satisfy the constraint

$$p_i(w_i - w_c) < 0, \quad p_i(w_{i+1} - w_c) > 0 \quad \text{for } i: p_i \neq 0. \tag{50}$$

Conditions (50) generate bounds on $d(i_1, i_2, \ldots, i_n)$, where $i_1, i_2, \ldots, i_n$ are locations of the phase boundaries, parameterizing a particular metastable branch. For example, if $n = 1$ we obtain the limits $d_-(i) < d(i) < d_+(i)$, where

$$d_\pm(i) = w_c + \Delta\left\{\frac{\sinh[(N-i)\lambda]\cosh[(i\pm1/2)\lambda]}{\cosh\frac{\lambda}{2}\sinh(N\lambda)} + \frac{i}{N} - 1\right\}. \tag{51}$$

Formulae (43), (48) and (51) describing single-interface solutions are illustrated in Fig. 4. An equilibrium branch parametrized by i begins at $d = d_-(i)$, where $w_i = w_c$, and terminates at $d = d_+(i)$, where $w_{i+1} = w_c$. In the interval $d_-(i) < d < d_+(i)$ the force is linear, while the energy is quadratic. Due to the symmetry of the problem, the ith branch is indistinguishable from the $(N-i)$th branch.
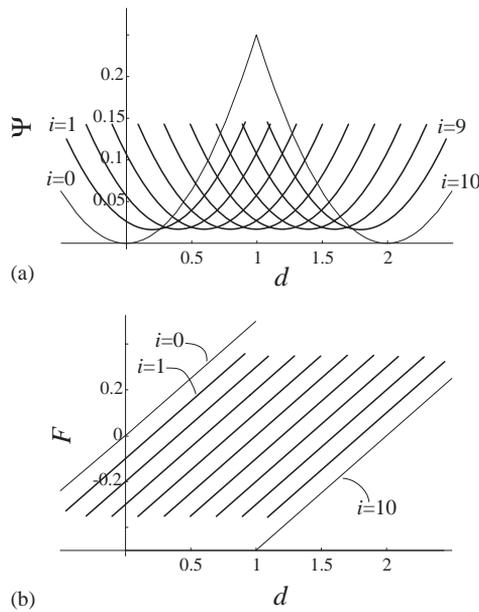
Fig. 4. The overall energy-strain and force-strain relations along the trivial ($i = 0$, N) and the single-interface ($1 < i < N$) metastable solutions. Here $N = 10$, $\mu = -2$, $w_c = a = 1$.
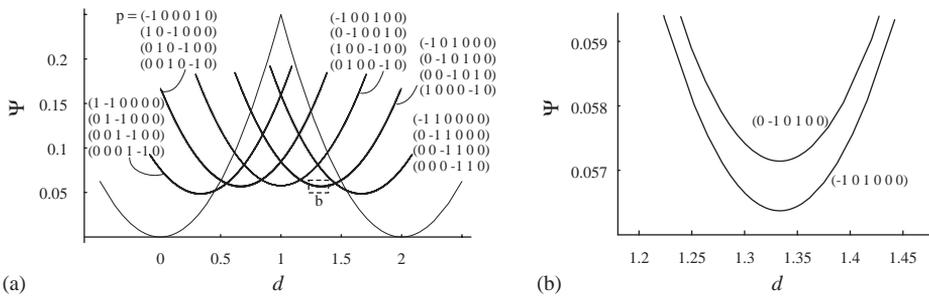

Fig. 5. (a) Energy-strain relation for the homogeneous solution (thin lines) and the two-interface solutions with various locations of the interfaces (thick lines). (b) The blow-up of (a) around configurations with $\mathbf{p} = (0 \ -1 \ 0 \ 1 \ 0 \ 0)$ and $\mathbf{p} = (-1 \ 0 \ 1 \ 0 \ 0 \ 0)$. Parameters: $\mu = -2$, $w_c = a = 1$, $N = 6$.

Equilibria with two interfaces are illustrated in Fig. 5 for a chain with $N = 6$. Notice that unlike the case without NNN interactions (e.g. Puglisi and Truskinovsky, 2002b), here not only the volume fractions but also the actual locations of the interfaces distinguish the branches. For example, the blow-up in Fig. 5b shows that the energy of the branch $\mathbf{p} = (0 \ -1 \ 0 \ 1 \ 0 \ 0)$ is higher than the energy of a branch $\mathbf{p} = (-1 \ 0 \ 1 \ 0 \ 0 \ 0)$ although both branches have the same fraction of phase II.

To find the global minimum of the energy we can first minimize among metastable solutions with a given number of interfaces n. The resulting lower envelopes $\hat{\Psi}(n, d)$
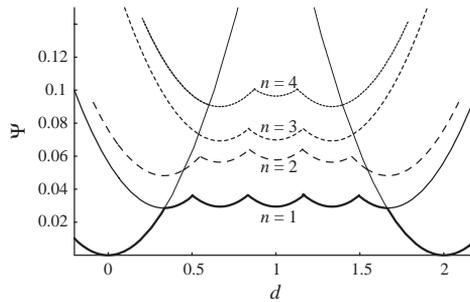
Fig. 6. Energy-strain relation for the energy minimizers in the families of solutions with up to four interfaces. Here $N = 6$, $\mu = -2$, $w_c = a = 1$.

are shown in Fig. 6 for $N = 6$ and $n \leqslant 4$. One can see that the single-interface solutions have the lowest energy, which is expected since our choice of nonlocal interactions leads to the penalization of interfaces. Similar observation has been made by Rogers and Truskinovsky (1997) for a discrete model with long-range forces.

5. Peierls-Nabarro landscape

To understand the degree of relative stability of various metastable states we need to evaluate the energy barriers. Since configurations with at most one interface are expected to have the lowest energy, it is instructive to start the analysis with the corresponding section of the energy landscape. By focusing on the configurations with a single phase boundary we obtain an analog of the Peierls-Nabarro landscape known in the theory of dislocations; it is also consistent with experimental observations for martensitic materials (e.g. Krishnan, 1985).

Consider a generic metastable configuration with a single phase boundary at $k = i$. To evaluate the barrier between this and the neighboring local minimum we need to choose a path connecting the configuration $w_k(i)$ to the configuration $w_k(i + 1)$ with one extra spring in phase II. Since the $(i + 1)$th spring must change phase, it is natural to choose the strain w_{i+1} as a parameter and minimize the total energy with respect to all w_k with $k \leqslant i$ and $k \geqslant i + 2$. We obtain the system of equations

$$\left(1 + \frac{1}{2\mu}\right) w_k + \frac{1}{4\mu}(w_{k+1} + w_{k-1}) = \begin{cases} F + a & \text{for } k \leqslant i \\ F & \text{for } k \geqslant i + 2, \end{cases} \tag{52}$$

which must be supplemented by the boundary conditions (2) and (6).

The configurations satisfying (52), (2) and (6) have the lowest energy among all states with a given w_{i+1}. This follows from the positive definiteness of the matrix $\mathbf{B}^{(i+1)}$ obtained from (7) by deleting the $(i + 1)$th row and the $(i + 1)$th column. It is

convenient to write the explicit representation for these configurations by using instead of w_{i+1} another parameter v:

$$
w_k(v) = \begin{cases} d - \Delta \left\{ \dfrac{v}{N} - 1 + C_1(\sinh(\lambda k) - \sinh(\lambda(k-1))) \right\}, & k \leqslant [v]+1 \\[3mm] d - \Delta \left\{ \dfrac{v}{N} + C_2(e^{\lambda k} - e^{\lambda(k-1)} - e^{\lambda(2N-k)} + e^{\lambda(2N-k+1)}) \right\}, & k \geqslant [v]+1, \end{cases}
$$

(53)

where $[v]$ denotes the integer part of v. It can be directly checked that (53) solves (52) for $k \leqslant [v]$ and $k \geqslant [v]+2$, with $[v]=i$ and the total force F given by

$$
F = \frac{1+\mu}{\mu} \left(d - \frac{v}{N} \Delta \right).
$$

(54)

By matching the strains at $k = i+1 = [v]+1$ and imposing (2), we compute the values of the constants

$$
C_1 = \frac{2\{e^{(2N-[v])\lambda}[v-[v]-e^{\lambda}(v-[v]-1)] - e^{\lambda([v]+1)}[v-[v]-1-e^{\lambda}(v-[v])]\}}{(e^{2\lambda}-1)(e^{2N\lambda}-1)}
$$

and

$$
C_2 = \frac{2\{(v-[v])\sinh[([v]+1)\lambda] - (v-[v]-1)\sinh[[v]\lambda]\}}{(e^{2\lambda}-1)(e^{2N\lambda}-1)}.
$$

In particular, we obtain an explicit relation between the parameters w_{i+1} and v:

$$
w_{[v]+1}(v) = d - \Delta \left(\frac{v}{N} - 1 \right) + C_1(\sinh[([v]+1)\lambda] - \sinh[[v]\lambda]).
$$

(55)

One can see that parameter w_{i+1} oscillates as the function of i with period 1, while v increases monotonically. Since the integer values of v correspond exactly to metastable configurations, the function $\Psi(v)$ obtained by substituting (53) in (46) is exactly the Peierls-Nabarro (PN) potential of our system. In order to move from one valley of this potential at $v=i$ to the neighboring one at $v=i+1$ the system must overcome the Peierls barrier. It is located at $v=v_i$ defined by

$$
w_{i+1}(v_i) = w_c.
$$

(56)

The height of the Peierls barrier $\delta\Psi^{i\to i+1} = \Psi(v_i) - \Psi(i)$ can be explicitly computed from (46) and (47).

The typical structure of the PN landscape is illustrated in Fig. 7 for a chain with $N=10$. One can see that at $d=0.5$ the metastable configurations correspond to the integer values $v=2$ (or 3), $v=1$ (or 4), $v=0$, $v=5$ and $v=6$. The corresponding strain profiles are shown in Fig. 8 for the metastable states $v=0$, $v=1$ and $v=2$ and for the saddle point configurations with $v=v_0$ and $v=v_1$.
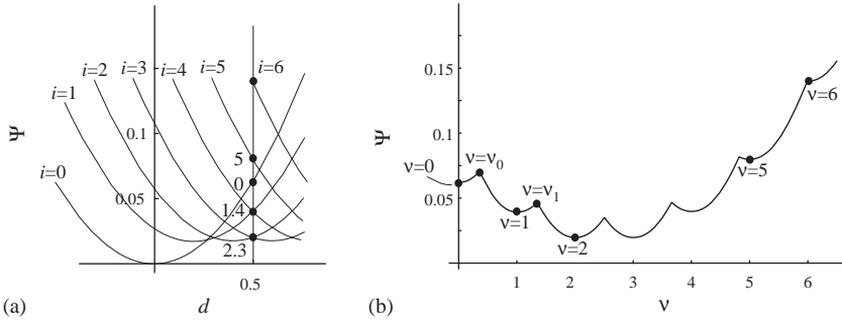
Fig. 7. (a) Energies of the single- and zero-interface branches of equilibria available at $d = 0.5$. (b) Peierls-Nabarro energy landscape along the path $w_k(v)$ connecting metastable equilibria at $d = 0.5$. Parameters: $N = 10$, $\mu = -2$, $w_c = a = 1$.
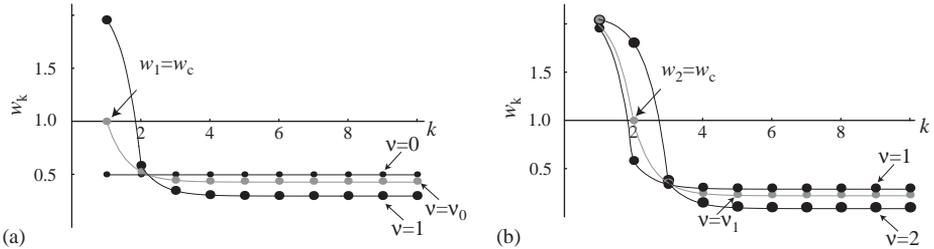

Fig. 8. Strain profiles associated with nucleation and incremental growth of a new phase. Configurations $v = 0$, $v = 1$ and $v = 2$ are local minima; $v = v_0$ and $v = v_1$ are saddle points. Parameters are the same as in Fig. 7.

6. Propagation threshold and nucleation peak

It is realistic to assume that the system driven by external loading d remains in a local minimum configuration until the minimal energy barrier around this state reaches a critical threshold H determined by the level of fluctuations (imperfections). In particular, the maximal delay strategy, associated with the gradient-flow dynamics, requires that the system stays on a given metastable branch until it becomes absolutely unstable ($H = 0$) and then evolves towards the nearest local minimum along the path of steepest descent. Various resulting force-strain paths for the NN system are discussed in Puglisi and Truskinovsky (2002a, b). For the NNN system the two characteristic paths with and without a threshold are shown in Fig. 9. One can see that as the elongation increases the system initially stays in the trivial configuration but eventually reaches the state where the smallest energy barrier becomes equal to H. Then nucleation takes place and the system escapes from the local minimum through the first saddle point with a subcritical height. After the nucleation event the phase boundary propagates along the chain in a stick-slip fashion, with the system getting temporarily trapped in each of the single interface metastable equilibria (parametrized by i). The resulting
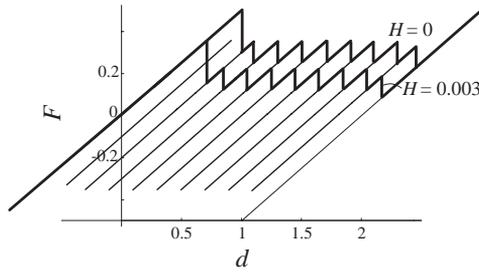
Fig. 9. Branch-switching sequence for the path of maximal delay, $H = 0$, and for the path with a finite critical energy barrier $H = 0.003$. Parameters: $\mu = -2$, $N = 10$, $w_c = a = 1$.
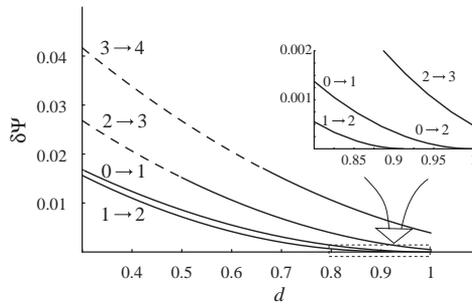

Fig. 10. Peierls barriers for several transitions $i \rightarrow i + 1$ parametrized by d. The dashed lines indicate the intervals where the transition is energetically unfavorable. The insert is the blow-up of the selected region. Parameters: $N = 10$, $\mu = -2$, $w_c = a = 1$.

graphs of the force $F(d)$ exhibit characteristic serrations. A visible nucleation peak originates from the fact that the nucleation event involves two springs transforming at once while the propagation involves only one spring changing phase at a time.

To understand why the first two springs change phase simultaneously, which never happens in the NN system (Puglisi and Truskinovsky, 2002a, b), we need to compare the height of the energy barrier for the transition $0 \rightarrow 1$ with the height of the barriers for several subsequent transitions. The height of the barriers is shown in Fig. 10 for the chain with $N = 10$ and $\mu = -2$. Notice that the barrier for the transition $1 \rightarrow 2$ is lower than the barrier for the transition $0 \rightarrow 1$, moreover, the barrier $1 \rightarrow 2$ vanishes at $d_+(1)$ and beyond this point the branch $i = 1$ does not exist any more (transition $0 \rightarrow 1$ deteriorates into $0 \rightarrow 2$). A similar calculation for $N = 20$ and $\mu = -2$ shows that the barrier for the transition $0 \rightarrow 1$ is higher than two subsequent barriers for the transitions $1 \rightarrow 2$ and $2 \rightarrow 3$; as the last two barriers vanish at sufficiently large d, the transition $0 \rightarrow 1$ first deteriorates into $0 \rightarrow 2$ and then into $0 \rightarrow 3$. One can see that in this case nucleation event involves simultaneous transformation of three springs.

Our computations show that "massive" nucleation and the associated nucleation peak phenomenon occur only when NNN interactions are sufficiently strong (large $|\mu^{-1}|$). Physically, it is the consequence of the nonlocal character of the model. The nonlocality

does not have to be of the NNN type: a different nonlocal model with long-range interactions also exhibits the nucleation peak (Rogers and Truskinovsky, 1997).

To understand what is going on, we observe that before the nucleation all springs are stretched uniformly, whereas after the nucleation, due to the presence of the internal boundary layers, the springs that are closer to the phase boundary have higher strain and hence are closer to the critical threshold than the springs far away. This facilitates the subsequent switching events and results in the smaller force required for the propagation of a phase boundary comparing to nucleation. In the system without nonlocality ($\mu = -\infty$), all springs outside the interface are stretched uniformly, and therefore the propagation of the interface does not take place until the critical strain is reached in all springs simultaneously. In this case the nucleation peak is absent: phase propagation effectively reduces to successive nucleation events in the shorter and shorter chains which requires the same critical force.

To obtain the upper bound for the size of the nucleation peak it is enough to consider the maximum delay strategy. In our model barrierless nucleation takes place when the force reaches the spinodal limit $F_{\max} = (1 + \mu^{-1})w_c$. On the other hand, the advance of the interface from $k = i$ to $k = i + 1$ takes place at $F(i, N) = F_M + F_P(i, N)$, where

$$F_M = \frac{1 + \mu}{\mu} w_c - \frac{a}{2} \tag{57}$$

is the *Maxwell force* and $F_P(i, N)$ is the *Peierls force* given by

$$F_P(i, N) = F(d_+(i), i) - F_M = a \left(\frac{\cosh[(\frac{1}{2} + i)\lambda] \sinh[\lambda(N - i)]}{\cosh \frac{\lambda}{2} \sinh(N\lambda)} - \frac{1}{2} \right). \tag{58}$$

One can show that the function $F_P(i, N)$ depends on i weakly away from the narrow boundary layers near $i = 1$ and $i = N - 1$. In the limit of infinite N the Peierls force (58) approaches the constant value (Truskinovsky and Vainchtein, 2003)

$$\lim_{N \to \infty} F_P(i, N) = F_P = \frac{a}{2} \sqrt{\frac{1 + \mu}{\mu}}. \tag{59}$$

As the NNN interactions get weaker ($\mu \to -\infty$), the Peierls force tends to the spinodal limit $F_{\max} - F_M$ and the nucleation peak disappears.

The configuration of the resulting hysteresis loop in the continuum limit is shown in Fig. 11. Although the serrations disappeared, the nucleation peak remains finite with the force dropping by the amount

$$\tau = \frac{1 + \mu}{\mu} w_c - F_P = \frac{1}{2} a \left(1 - \sqrt{\frac{1 + \mu}{\mu}} \right). \tag{60}$$

This quantity is positive as long as $-\infty < \mu < -1$ and is always less than $a/2$—the difference between the spinodal and Maxwell forces. The half-height of the narrow part of the hysteresis loop is given in the continuum limit by (59).

To estimate the number of springs involved in the nucleation event, we recall that the nucleation always takes place at $d = w_c$. Therefore we must find the metastable branch with the smallest nonzero i which is defined at this value of d. Setting $d_\pm$ from
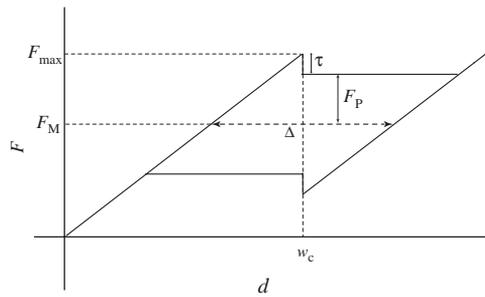
Fig. 11. The maximum hysteresis loop in the continuum limit for the nonlocal (NNN) model with $-\infty < \mu < -1$.

(51) equal to w_c, we obtain the equation for the number of participating springs i

$$\frac{\sinh[(N - i)\lambda]\cosh[(i \pm 1/2)\lambda]}{\cosh\frac{\lambda}{2}\sinh(N\lambda)} = 1 - \frac{i}{N}. \tag{61}$$

For sufficiently large N and generic μ the solution of this equation can be approximated by

$$i_{\text{nuc}} = \frac{N}{1 + e^\lambda}. \tag{62}$$

In particular, (62) implies that in the continuum limit the number of springs involved in the nucleation event is infinite. The dimensionless size of the transformed portion of the chain (martensite band) remains finite:

$$l_0 = \frac{1}{1 + e^\lambda}. \tag{63}$$

When nonlocal interactions are absent ($\mu = -\infty$), the nucleus contains only one spring and $l_0 = 0$.

To illustrate the effect of the outside boundary layers on the nucleation phenomenon we replace the special boundary conditions (6), which suppress boundary layers, by the "zero-moment" conditions (4) with $\Psi_B = 0$. In this case the zero-interface solution is no longer trivial because the strain decreases exponentially near the boundaries (see Fig. 12b).

Suppose that the chain is originally in phase I. When the critical value $d = d_+(0)$ is reached, the strains in the middle of the chain pass the threshold $w = w_c$ and the one-phase solution becomes unstable (point A in Fig. 12). This leads to the formation of either two symmetric interfaces in the center (point C in Fig. 12) or of a single interface near one of the boundaries (point B in Fig. 12). Our computations show that the single-interface configuration B has a lower energy than the two-interface configuration C. However, during transition from A to C only two springs (7th and 9th) transform into phase II, while transition from A to B requires transformation of at least three springs which are also initially farther below the threshold. As a result, transition from A to B may encounter a higher barrier than transition from A to C, and then the nucleation will take place in the interior of the chain. While this possibility
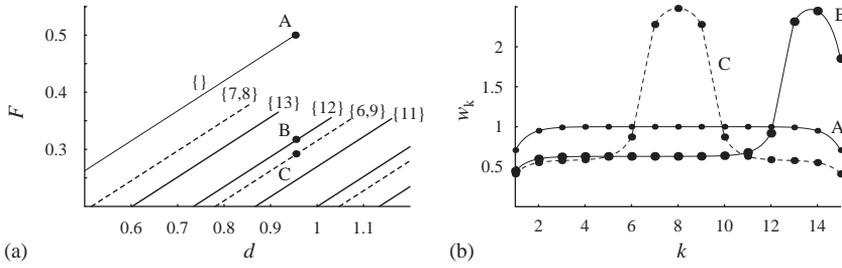
Fig. 12. Nucleation behavior of the system with external boundary layers: (a) single-interface (solid lines) and two-interface (dashed lines) metastable states; (b) the corresponding strain profiles. The numbers in brackets indicate locations of phase the boundaries for configurations with phase I on the left end. Parameters: $N = 15$, $\mu = -2$, $w_c = a = 1$.

can be investigated rigorously, for the subject of this paper it is enough to mention that the nucleation peak survives in both cases (see Fig. 12a).

7. Verification of the model

The explicit formulae from the previous section can be used to obtain bounds on the value of the nonlocality measure γ. The expression for γ from (10) and (60) has the form

$$\gamma = \frac{E}{4}\left[1 - \left(1 - \frac{2\tau}{E\Delta}\right)^{-2}\right], \tag{64}$$

where $E = K + 4\gamma$ is the macroscopic elastic modulus of the homogeneous chain and Δ is the macroscopic transformation strain. The magnitude of the stress drop at the peak τ and the transformation strain Δ are available from the experimental data of Shaw and Kyriakides (1995) on NiTi wires. For instance, in the experiment conducted at $70°$ and the loading rate 4×10^{-5} s^{-1}, the measurements gave $\tau = 0.039$ GPa, and $\Delta = 3.97\%$. The Young's moduli of the two participating phases are different, $E = 56.7$ and 27.5 GPa. By using two separate values we estimate γ to be between -1.1 and -1 GPa, which implies that μ is in the range $-14.7 < \mu < -7.3$. Similar comparison of experiment and theory for CuAlNi yields $\mu = -33.4$ (Truskinovsky and Vainchtein, 2003).

To obtain an independent estimate of γ we can assume that the interactions between particles are governed by the Lennard-Jones potential. In this case we have $\varepsilon\phi_1(r/\varepsilon - 1) = 2\varepsilon\phi_2(r/2\varepsilon - 1) = U(r)$, where $\phi_1(w)$ and $\phi_2(w)$ are the energies of the NN and NNN springs, respectively, and $U(r)$ has the form

$$U(r) = \frac{K\varepsilon}{72}\left[\left(\frac{\varepsilon}{r}\right)^{12} - 2\left(\frac{\varepsilon}{r}\right)^{6}\right]. \tag{65}$$

The coefficients in (65) are chosen to ensure that the elastic modulus at equilibrium $r = \varepsilon$ equals K. Linearizing around the unstretched homogeneous state with the spacings

$r = \varepsilon$ and $r = 2\varepsilon$, we obtain (Charlotte and Truskinovsky, 2002)

$$\mu = \frac{U''(\varepsilon)}{4U''(2\varepsilon)}.$$

This yields $\mu = -56.5$. Taking $E = 56.7$ GPa as the value of the Young's modulus we obtain $\gamma = -0.255$ GPa which despite the rather rigid form of the potential (65) is within a reasonable range from the values obtained above.

Finally, we can use the above estimates of μ to predict the initial size of the martensite band L_0. By taking $-14.7 < \mu < -7.3$ we obtain $0.001 < L_0/L < 0.005$, where L is the size of the specimen. While we could not find direct experimental measurements for L_0 in the literature, this parameter is usually estimated to be of the order of the specimen's diameter D (e.g. Sun and Zhong, 2000). This would be in agreement with our estimate since for thin NiTi wires used in experiment $D/L \sim 0.0025$ (Leo et al., 1993).

8. Strain-gradient approximation

Now we can compare the exact results obtained in the discrete model with the predictions of the quasi-continuum strain gradient approximation which neglects discreteness but retains the original internal length scale. To obtain the approximate model we temporarily reintroduce dimensional variables and perform the Taylor expansion in the small parameter $\varepsilon = L/N$. By preserving the first nonlocal term in the energy functional we obtain

$$\Psi = \int_0^L \left[\phi(w) + \frac{1}{2} A\varepsilon^2 (w')^2 \right] dx, \tag{66}$$

where we omitted nonessential null Lagrangian contributions and introduced

$$A = -\frac{1}{12}(K + 16\gamma). \tag{67}$$

The functional (66) must be minimized subject to the constraint

$$\int_0^L w(x)\,dx = d \tag{68}$$

and the clamping boundary conditions

$$w'(0) = w'(L) = 0 \tag{69}$$

which represent the continuum analog of (6). The Euler-Lagrange equations can be written in the form $\phi'(w) - A\varepsilon^2 w'' = F$ (analog of (3)), where $\phi'(w) = Ew - Ka\theta(w - w_c)$. Recall that $E = K + 4\gamma$ is the macroscopic modulus.

The trivial solution of the Euler-Lagrange equations $w(x) = d/L$ is stable whenever the second variation of energy

$$\delta^2 \Psi = \int_0^L (Ev^2 + \varepsilon^2 A v'^2)\,dx \tag{70}$$

is positive for all $v(x) \neq 0$ satisfying the constraint

$$\int_0^L v \, dx = 0. \tag{71}$$

The zero eigenvalue problem determining the stability boundaries reduces to solving linear equation $Ev' - \varepsilon^2 A v''' = 0$. The general nontrivial solution of this equation compatible with (71) takes the form

$$v(x) = \cos \frac{\pi n x}{L},$$

where $n \geqslant 1$ is an integer. The solution exists for

$$E + \left(\frac{\pi n \varepsilon}{L} \right)^2 A = 0. \tag{72}$$

If $A = 0$, the stability condition obviously reduces to $E > 0$. When $A > 0$, we obtain $E + (\pi n \varepsilon / L)^2 A \geqslant E + (\pi \varepsilon / L)^2 A$ and the trivial state is stable if and only if

$$E + \left(\frac{\pi \varepsilon}{L} \right)^2 A > 0. \tag{73}$$

The instability takes place through the growth of a long-wave mode $v = \cos(\pi x / L)$. When $A < 0$, the trivial solution is always unstable because one can always find large enough n at which $E + (\pi n \varepsilon / L)^2 A < 0$; the corresponding instability is of the short-wave type.

   To compare the stability limits in the strain gradient model with the ones obtained in the discrete model, we shall first rewrite them in terms of dimensionless μ, and replace L/ε by N. In the case when $A > 0$, we obtain

$$-4 < \mu < - \frac{1 - \pi^2 / 3N^2}{1 - \pi^2 / 12N^2} \quad (\gamma < 0). \tag{74}$$

Recall that in the discrete case the corresponding interval is $-\infty < \mu < - \cos^2 \pi / 2N$ $(\gamma < 0)$. It is easy to see that for large N both models predict the same upper boundary $\mu = -1$ ($E = 0$), which means that the strain-gradient model captures the onset of macroscopic instability in the discrete model rather well.

   As we showed above, when $A < 0$ the stability range does not exist due to the short-wave instability. However, in view of the underlying discrete structure, a short-wave instability is unphysical if its wave length is less than the length scale of the lattice ε. If we assume that the mode number cannot exceed $n = N$, we obtain the following stability condition:

$$E + \pi^2 A > 0, \tag{75}$$

which in terms of μ can be written as a combination of two intervals,

$$\mu < -4 \quad (\gamma < 0) \tag{76}$$

and

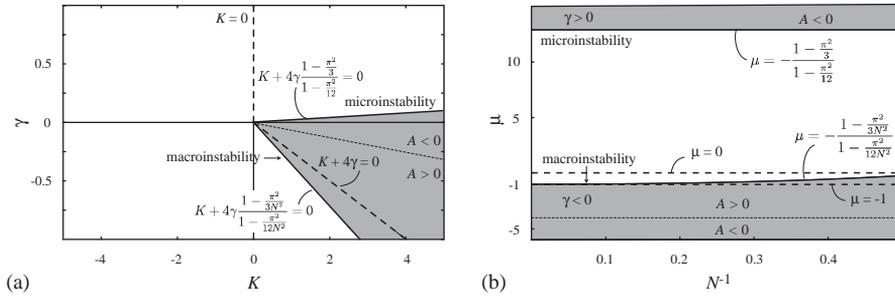$$-\frac{1 - \pi^2 / 3}{1 - \pi^2 / 12} < \mu < \infty \quad (\gamma > 0). \tag{77}$$

Fig. 13. (a) Domain of stability for the trivial solution in $K - \gamma$ plane in the strain-gradient model with $L/\varepsilon = N = 3$ and in the continuum limit ($N \to \infty$, dashed lines). (b) The dependence of the stability boundaries on the size of the system. In both figures stability domain is in gray.

While the former stability range obviously complements (74), the latter has its own analog in the discrete model, $-\sin^2 \pi/2N < \mu < \infty$ ($\gamma > 0$). Observe that the lower limit in the strain-gradient model is significantly higher than in the discrete model suggesting that even with the restriction on the minimum wave length, the strain-gradient approximation grossly exaggerates the instability domain on the side of the short-wave instability. The general trend, however, is predicted correctly: as in the discrete model, the short-wave instability in the gradient model develops before the strain enters the macroscopic spinodal region (see also Bardenhagen and Triantafyllidis, 1994).

Fig. 13 summarizes the stability analysis in both the $K - \gamma$ and $\mu - N^{-1}$ planes. It should be compared with the analogous stability diagrams for the discrete model (Fig. 2).

Now consider the nontrivial solutions of the variational problem (66), (68). As shown in Carr et al. (1984), the strain-gradient model allows for local minima with at most one interface. For the piecewise linear model all these metastable configurations can be obtained analytically (Truskinovsky and Zanzotto, 1996). A representative picture is shown in Fig. 14, with the parameters chosen to match those for the discrete model with $N = 20$. We observe that while the strain-gradient approximation captures the structure of the absolute minimizers of the discrete problem, it fails to reproduce the rich structure of the metastable equilibria which in the quasi-continuum model all collapse into a single branch. This is the reason why the strain-gradient approximation does not generate a realistic hysteresis. Also, while all multiple-interface equilibria of the discrete model are metastable, configurations with more than one interface are absolutely unstable in the quasi-continuum approximation. This difference is due to the fact that in the discrete model one cannot vary the volume fraction of phases continuously, and thus certain perturbations that make multiple-interface solutions unstable in the continuum case are impossible in the discrete model. Finally, we remark that the estimate $-14.7 < \mu < -7.3$ obtained in the previous section from the experimental data implies that $A < 0$, contrary to the typical assumption in the strain-gradient models that the coefficient A is positive.
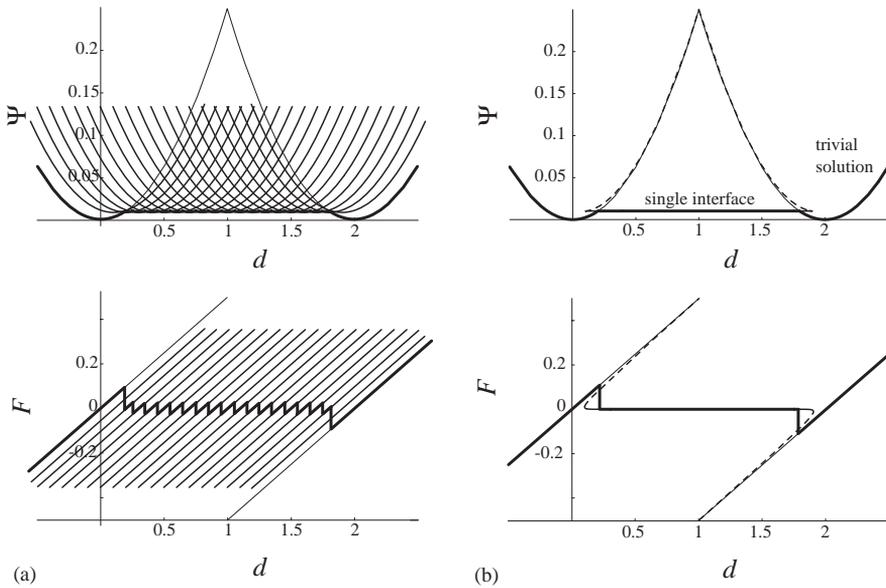
Fig. 14. The rescaled energy-strain and force-strain relations for the absolute energy minimizers (thick line) and the metastable single-interface solutions (thin lines) for (a) the discrete chain with $N = 20$ and (b) the strain-gradient approximation with $\varepsilon = \frac{1}{20}$. Other parameters: $\mu = -2$, $w_c = a = 1$. Unstable single-interface solutions are shown by dashed lines.

9. Conclusions

Our study clarifies why the behavior of the local and nonlocal systems with bi-stable elements is different. In the case of nonlocal interactions the elements around the phase-boundary are "pre-conditioned" due to the presence of the boundary layers. This results in smaller energy barriers for the propagation of the phase boundary compared to the local model and gives rise to the finite mismatch between the spinodal stress and the Peierls stress. Another difference is that in the nonlocal models the nucleation event typically involves transformation of more than one element. This produces nucleation peak which has been long known from experiment. We showed that the nucleation peak persists in the continuum limit where it manifests itself through an instantaneous formation of a finite band scaled with the length of the sample. Although the nucleation peak is captured by the strain-gradient approximation, the resulting shape of the hysteresis loop is unrealistic. This suggests that the nonlocal coupling in the actual physical system is weaker than what is required for the gradient model to be applicable.

Acknowledgements

References

Bardenhagen, S., Triantafyllidis, N., 1994. Derivation of higher order gradient continuum theories in 2,3-D non-linear elasticity from periodic lattice models. J. Mech. Phys. Solids 42 (1), 111–139.

Butler, J.F., 1962. Lüders front propagation in low carbon steels. J. Mech. Phys. Solids 10, 313–334.

Carr, J., Gurtin, M., Slemrod, M., 1984. Structured phase transitions on a finite interval. Arch. Rational Appl. Mech. Anal. 86, 317–351.

Charlotte, M., Truskinovsky, L., 2002. Linear elastic chain with a hyper-pre-stress. J. Mech. Phys. Solids 50, 217–251.

Cottrell, A.H., Bilby, B.A., 1949. Dislocation theory of yielding and strain ageing of iron. Proc. Phys. Soc. London 62/I-A, 49–62.

Fedelich, B., Zanzotto, G., 1992. Hysteresis in discrete systems of possibly interacting elements with a two well energy. J. Nonlinear Sci. 2, 319–342.

Froli, M., Royer-Carfagni, G., 2000. A mechanical model for the elastic-plastic behavior of metallic bars. Internat. J. Solids Struct. 37, 3901–3918.

Hall, E.O., 1970. Yield Point Phenomena in Metals and Alloys. Plenum Press, New York.

Johnston, W.G., Gilman, J.J., 1959. Dislocation velocities, dislocation densities and plastic flow in lithium fluoride crystals. J. Appl. Phys. 30, 129–144.

Kartha, S., Krumhansl, J., Sethna, J., Wickham, L., 1995. Disorder driven pretransitional tweed pattern in martensitic transformations. Phys. Rev. B: Condens. Matter 52 (2), 803.

Krishnan, R., 1985. Stress induced martensitic transformations. Mater. Sci. Forum 3, 387–398.

Kyriakides, S., Miller, J.E., 2000. On the propagation of luders bands in steel strips. J. Appl. Mech. 67 (4), 645–654.

Leo, P.H., Shield, T.W., Bruno, O.P., 1993. Transient heat transfer effects on the pseudoelastic behavior of shape-memory wires. Acta Metall. Mater. 41, 2477–2485.

Lexcellent, C., Tobushi, H., 1995. Internal loops in pseudoelastic behaviour of ti-ni shape memory alloys: experiment and modelling. Meccanica 30, 459–466.

Mickens, R.E., 1990. Difference Equations: Theory and Applications, 2nd Edition. Van Nostrand Reinhold Co., New York.

Mindlin, R.D., 1965. Second gradient of strain and surface-tension in linear elasticity. Internat. J. Solids Struct. 1, 417–438.

Müller, I., Villaggio, P., 1977. A model for an elastic-plastic body. Arch. Rational Appl. Mech. Anal. 65, 25–46.

Pagano, S., Paroni, S., 2003. A simple model for phase transformations: from discrete to the continuum problem. Quart. Appl. Math. 54, 328–348.

Puglisi, G., Truskinovsky, L., 2000. Mechanics of a discrete chain with bi-stable elements. J. Mech. Phys. Solids 48, 1–27.

Puglisi, G., Truskinovsky, L., 2002a. A mechanism of transformational plasticity. Continuum Mech. Thermodynamics 14, 437–457.

Puglisi, G., Truskinovsky, L., 2002b. Rate independent hysteresis in a bi-stable chain. J. Mech. Phys. Solids 50, 165–187.

Rogers, R., Truskinovsky, L., 1997. Discretization and hysteresis. Physica B 233, 370–375.

Shaw, J.A., Kyriakides, S., 1995. Thermomechanical aspects of NiTi. J. Mech. Phys. Solids 43, 1243–1281.

Shaw, J.A., Kyriakides, S., 1997a. Initiation and propagation of localized deformation in elasto-plastic strips under uniaxial tension. Internat. J. Plasticity 13 (10), 837–871.

Shaw, J.A., Kyriakides, S., 1997b. On the nucleation and propagation of phase transformation fronts in a NiTi alloy. Acta Mater. 45, 683–700.

Sun, Q.-P., Zhong, Z., 2000. An inclusion theory for the propagation of martensite band in NiTi shape memory alloy wires under tension. Internat. J. Plasticity 16, 1169–1187.

Triantafyllidis, N., Bardenhagen, S., 1993. On higher order gradient continuum theories in 1-D nonlinear elasticity. Derivation from and comparison to the corresponding discrete models. J. Elasticity 33 (3), 259–293.

Triantafyllidis, N., Bardenhagen, S., 1996. The influence of scale size on the stability of periodic solids and the role of associated higher order gradient continuum models. J. Mech. Phys. Solids 44, 1891–1928.

Truskinovsky, L., Vainchtein, A., 2003. Peierls-Nabarro landscape for martensitic phase transitions. Phys. Rev. B 67, 172103.

Truskinovsky, L., Zanzotto, G., 1996. Ericksen's bar revisited: Energy wiggles. J. Mech. Phys. Solids 44 (8), 1371–1408.

Ye, Y.Y., Chan, C.T., Ho, K.M., 1991. Effect of phonon anomalies on the shear response of martensitic crystals. Phys. Rev. Lett. 66, 2018–2021.