About the origin of the nucleation peak in transformational plasticity

Lev Truskinovsky

Laboratoire de Mechanique des Solides, CNRS-UMR 7649, Ecole Polytechnique, 91128, Palaiseau, France

Anna Vainchtein *

Department of Mathematics, University of Pittsburgh, Pittsburgh, PA 15260

Abstract

We study the processes of nucleation and growth of a new phase in a prototypical discrete model of a martensitic material. The model takes nonlocality into account and exhibits all major phenomena observed during quasistatic loading of shape memory wires including rate-independent hysteresis and formation of isolated phase boundaries. We solve the associated finite difference problem exactly and show that the presence of nonlocal interactions leads to the nucleation peak phenomenon: the force at which nucleation takes place is higher than the force at which a phase boundary can propagate. We show that the difference between the nucleation and propagation thresholds persists in the continuum limit and tends to zero as the nonlocality gets weaker. The model suggests specific relations between the microscopic parameters of the lattice and the size of the nucleation peak which we verify for cubic to monoclinic phase transformation in NiTi wires.

Key words: martensitic phase transitions, lattice models, nonlocal interactions, Peierls-Nabarro landscape, nucleation

1 Introduction

Shape memory alloys and other martensitic materials undergoing diffusionless phase transformations are known to exhibit rate-independent hysteresis

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

when subjected to quasistatic cyclic loading. The corresponding displacement-controlled experiments often display the following characteristic feature: prior to nucleation of the new phase, the load reaches a peak which is followed by a plateau with a distinctly lower stress (Horikawa and Miyazaki, 1988; Lexcellent and Tobushi, 1995; Shaw and Kyriakides, 1995, 1997b; Sun and Zhong, 2000). The nucleation peaks are not unique to transformational plasticity and are also routinely observed during initiation of a conventional plastic deformation in mild (low carbon) steels; a sudden drop in stress is associated with the initiation of Lüders bands which then spread along the length of the specimen at an essentially constant stress (Butler, 1962; Hall, 1970; Kyriakides and Miller, 2000). The presence of the nucleation peak in steels has been attributed to the pinning of dislocations and the fact that the stress required to release the trapped dislocations is higher than the stress needed to sustain their motion (Cottrell and Bilby, 1949). Another theory links the sudden drop of stress with the multiplication of dislocations and the fact that the stress required to move dislocations decreases with their number (Johnston and Gilman, 1959).

Although in martensitic materials the phenomenon of nucleation-induced load discontinuity appears to be similar to the drop from "upper" to "lower" yield stress in steels, it has not been understood at the equally fundamental level. Thus the nucleation peak has been detected numerically in 3D plasticity-type models of shape memory alloys (Shaw and Kyriakides, 1997a; Kyriakides and Miller, 2000; Sun and Zhong, 2000), however the phenomenological parameters responsible for the size of the peak have not been identified. At a qualitative level, the peak 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 when released. To achieve the quantitative understanding of the factors contributing to the peak phenomenon, we develop in this paper a prototypical model of a martensitic material which underlines the important physical aspects of the problem and maximally simplifies the unimportant ones. The proposed model supports the intuition developed in plasticity theory and adapts it to the case when dislocations are replaced by phase boundaries as the principal carriers of inelastic deformation.

According to our model, the main factors responsible for the appearance of the nucleation peak are (i) discreteness, revealed through the presence of an internal length scale, (ii) nonconvexity of the energy of the elementary transforming units or bands, and (iii) the long-range or nonlocal coupling between these units. The study of one-dimensional discrete chains of bi-stable springs without nonlocal interactions has long demonstrated the ability of this class of mechanical theories to capture many important properties of shape memory alloys (e.g. Müller and Villaggio, 1977; Fedelich and Zanzotto, 1992; Făciu and Suliciu, 1994; Puglisi and Truskinovsky, 2000, 2002a,b). In particular, these purely local models generate sufficiently complex energy landscapes that are compatible with the hysteretic behavior and the fine structure of inner loops.

One common feature of these models is their permutational degeneracy: nucleation of a new phase can occur anywhere in the chain. Most importantly, the local models fail to predict the nucleation peak, at least in one dimension. Numerical studies have shown that incorporation of nonlocal interaction between the bi-stable elements leads to additional features such as selection of a particular number of interfaces and, in some cases, the nucleation peak (e.g. Ye et al., 1991; Triantafyllidis and Bardenhagen, 1993; Rogers and Truskinovsky, 1997; Froli and Royer-Carfagni, 2000; Pagano and Paroni, 2002). None of the nonlocal models, however, have been developed analytically to an extent that they could explain the necessity of the nucleation peak and identify the microparameters controlling its size.

The advantage of the prototypical model of martensitic material considered in the present paper is that it contains all three required ingredients (i)-(iii) and at the same time lends itself to a completely analytical treatment. Specifically, we consider a finite system of particles connected by bi-stable springs. To mimic the three-dimensional nature of the actual problem, we complement the anharmonic interactions between the nearest neighbors (NN) by a harmonic interaction of the next-to-nearest neighbors (NNN). We then use a piecewise linear approximation for the NN interaction which allows us to obtain a fully analytical characterization of all metastable equilibria and explicitly reconstruct the non-equilibrium Peierls-Nabarro landscape. Considering a quasistatic displacement-controlled loading of the system and selecting a maximally dissipative strategy for switching between metastable equilibria, we demonstrate the necessity of the nucleation peak in the experimentally feasible range of parameters. In addition, we show that the difference between the nucleation and propagation thresholds does not disappear in the continuum limit, although it tends to zero as the nonlocality gets weaker. The obtained exact solution of the discrete problem allows us to relate the size of the nucleation peak and the associated propagation (or Peierls) stress to the microstructural characteristics of the lattice. The model suggests specific relations between microscopic and macroscopic parameters. We verify these relations for the cubic to monoclinic phase transformation in NiTi wires. In particular, we use the experimental value for the stress drop to obtain a bound for the stiffness of the NNN springs and show that it is compatible with the independent estimate based on the lattice model with Lennard-Jones particle interactions.

In the final section of the paper we briefly compare our nonlocal discrete model with the quasi-continuum strain-gradient approximation (Mindlin, 1965; Triantafyllidis and Bardenhagen, 1993). As expected (e.g. Braides and Gelli, 2002), the strain-gradient model reproduces well the structure of the global minimum of the energy in the continuum limit but misses the rich variety of metastable equilibria. Although the nucleation peak is captured by the gradient approximation, the resulting shape of the hysteresis loop is completely

unrealistic. This suggests that the nonlocal coupling in the actual physical system is weaker than what is required for the gradient approximation to be applicable.

The structure of the paper is as follows. In Section 2 we formulate the discrete problem and discuss the origin of instability of the trivial solution leading to nucleation and growth of the new phase. The derivation of the general solution of the discrete problem for an arbitrary phase configuration is outlined in Section 3. We then discuss the energy barriers in Section 4 and construct the associated Peierls-Nabarro landscape. In Section 5 we compute the magnitude of the nucleation peak for the finite chain and then study the continuum limit. In Section 6 we link the parameters of the model to the experimental measurements in NiTi wires. Section 7 contains the comparison of the discrete and strain-gradient models. The conclusions are summarized in Section 8. Results of the technical nature are presented in the three Appendices.

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); an alternative interpretation of this mechanical system would be a set of rigid (atomic) planes connected by shear springs. Let u_k, $0 \le k \le N$, be the displacements of the particles with respect

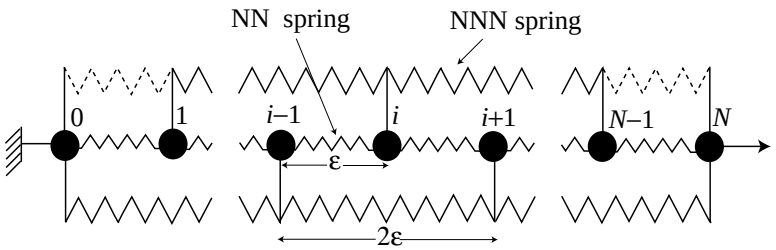

Fig. 1. A finite chain of particles with nearest-neighbor (NN) and next-to-nearest-neighbor (NNN) interactions.

to a load-free homogeneous reference configuration with spacing ε. If we denote the strain in the kth NN spring by $w_k = (u_k - u_{k-1})/\varepsilon$, the total energy of the system can be written as

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

where $\phi_1(w)$ and $\phi_2(w)$ are the energies of the NN and NNN interactions, respectively, and the term Ψ_B stands for the energy of the boundary elements (to be specified below). 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) may be complemented by other constraints. For instance, to mimic an "extra hard device" one can additionally impose $w_1 = 0$ and $w_N = 0$. Alternatively, one may consider a "zero-moment device" which means that no additional conditions are imposed and $\Psi_B = 0$.

Let $f_i(w) = \phi_i'(w)$, $i = 1, 2$, denote the forces in NN and NNN springs, respectively. By minimizing the energy (1) subject to (2), we obtain as necessary conditions the equations governing the equilibrium of the interior particles with $2 \le k \le N - 1$:

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

Here F is the total force in the system, the Lagrange multiplier associated with the constraint (2). Unless the strains in the boundary elements with $i = 1$ and $i = N$ are constrained, the minimization over these variables gives the following "natural" boundary conditions

$$\begin{aligned} f_1(w_1) + f_2\left(\frac{w_2 + w_1}{2}\right) + \frac{\partial \Psi_B}{\partial w_1} &= F, \\ f_1(w_N) + f_2\left(\frac{w_N + w_{N-1}}{2}\right) + \frac{\partial \Psi_B}{\partial w_N} &= F, \end{aligned} \tag{4}$$

If we additionally assume that

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

the conditions (4) would mean that the boundary NNN springs are cut in half and then reconnected parallel to the NN springs (see the dashed springs in Fig. 1). The advantage of the latter choice is that the boundary equations (4) can be naturally included into the bulk equations (3) which would then hold for all $1 \le k \le N$; we must assume, however, that in this case there are fictitious 0th and $(N + 1)$th springs with the strains satisfying

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

Another advantage of the boundary conditions (4) and (5) is that they suppress the boundary layers (Charlotte and Truskinovsky, 2002) and ensure the existence of a *trivial solution* with the uniform strain distribution $w_k = d/L$, where $L = N\varepsilon$, $F_{\rm tr} = f_1(d/L) + 2f_2(d/L)$ and $\Psi_{\rm tr}(w) = L(\phi_1(d/L) + 2\phi_2(d/L))$. In the following we will deal exclusively with the boundary conditions (4), (5)

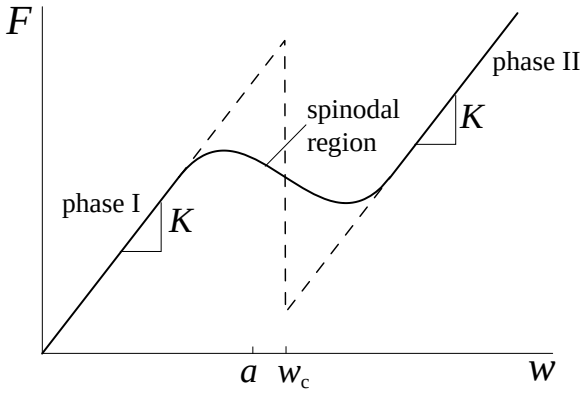

Fig. 2. Nonmonotone force-strain relation in an individual NN spring (solid line) and its bilinear approximation (dashed lines).

and only briefly consider the case $\Psi_B = 0$ (when the boundary layers reappear) in Appendix C.

To model martensitic phase transitions, we assume that each NN spring has a double-well energy generating a non-monotone force-strain relationship depicted in Fig. 2. The regions where the force increases with strain correspond to two different phases of material; the domain of decreasing force is called the *spinodal region*. We recall that in the continuum theory the spinodal region is associated with the absolute instability of the homogeneous configuration (Ericksen, 1975). To assess stability of the homogeneous (trivial) solutions in the discrete model we need to study the conditions of the positive definiteness of the matrix $\mathbf{B} = \frac{1}{\varepsilon}\frac{\partial^2 \Psi}{\partial w_i \partial w_j}$. Let $K(w) = \phi_1''(w)$ and $\gamma(w) = \phi_2''(w)/2$ denote the tangential moduli of NN and NNN springs, respectively. The necessary and sufficient conditions for stability can then be expressed in terms of these moduli as (see Appendix A for the proof)

$$K + 4\gamma > 0, \quad K + 4\gamma \sin^2 \frac{\pi}{2N} > 0. \tag{7}$$

Notice that in the limit of infinite N (7) reduce to the known conditions $K > 0$, $K + 4\gamma > 0$ (e.g Mindlin, 1965). Stability conditions similar to (7) were obtained in (Charlotte and Truskinovsky, 2002) for a finite chain with "moment free" boundary conditions ($\Psi_B = 0$).

If the chain is stretched in the spinodal region and the stability conditions (7) are violated, the trivial solution of the discrete problem becomes unstable, leading to the formation of inhomogeneities. To specify the stable response of the chain in this case, we need to find all nontrivial solutions of the equilibrium equations (3) corresponding to either local or global minima of the energy. The complete analytical treatment of this problem is possible when both the $f_1(w)$ and $f_2(w)$ are piecewise linear. Therefore, following Fedelich and Zanzotto

(1992), we choose the force-strain relation in NN springs in the form

$$f_1(w) = \begin{cases} Kw & \text{for } w < w_c \\ K(w-a) & \text{for } w \geq w_c. \end{cases} \tag{8}$$

Here the spinodal region reduces to a point: the phase state of the NN spring depends on whether the strain is below or above the critical value w_c. The other parameters in (8) are the transformation strain a and the constant elastic modulus K which is assumed to be positive and whose value is taken the same in both phases (see Fig. 2). For the NNN springs we adopt the simplest linear force-strain relation

$$f_2(w) = 2\gamma w \tag{9}$$

with $\gamma < 0$. The latter inequality is justified for nonlinear interactions of the Lennard-Jones type (see Section 6).

Under the assumptions (8), (9) the total energy (1) reduces to

$$\Psi = \varepsilon \left\{ \frac{1}{2}\mathbf{B}\mathbf{w}\cdot\mathbf{w} - \mathbf{q}\cdot(\mathbf{w}-\mathbf{w_c}) \right\}, \tag{10}$$

where $\mathbf{B}$ is a tridiagonal matrix presented explicitly in Appendix A (see (A.1)), $q_k = Ka\theta(w_k - w_{ck})$ is the vector prescribing the distribution of phases, $\theta(x)$ is a unit step function and $w_{ck} = w_c$, $k = 1, ..., N$. The equilibrium equations (3) together with the boundary conditions (6) can then be rewritten in the form

$$\mathbf{B}\mathbf{w} = \mathbf{F} + \mathbf{q}, \tag{11}$$

where $F_k = F$, $k = 1, ..., N$. Both the external force F and the phase inhomogeneity vector $\mathbf{q}$ contribute as sources in the equation (11), which remains nonlinear since $\mathbf{q}$ depends on $\mathbf{w}$. To ensure that the matrix $\mathbf{B}$ is positive definite, we assume that in addition to already stated conditions on elastic moduli, $K > 0$ and $\gamma < 0$, the inequality $K + 4\gamma > 0$ is also satisfied, which implies (7).

At this point it is convenient to rescale the variables and introduce nondimensional parameters. Selecting the length of the chain $L = N\varepsilon$ as the length scale and choosing K as the force scale, we define the new variables

$$\bar{u}_k = \frac{u_k}{N\varepsilon}, \quad \bar{d} = \frac{d}{N\varepsilon}, \quad \bar{F} = \frac{F}{K}, \quad \bar{\Psi} = \frac{\Psi}{KN\varepsilon}, \quad \bar{\mathbf{B}} = K^{-1}\mathbf{B}. \tag{12}$$

The rescaled problem contains the only nondimensional parameter

$$\beta = \frac{4\gamma}{K}. \tag{13}$$

In what follows, we omit the bars and work exclusively with the nondimensional variables, having in mind that after the rescaling $\varepsilon = 1/N$. Observe

that in terms of β our assumptions on the moduli can be summarized as

$$-1 < \beta < 0. \tag{14}$$

3 Stable and metastable equilibria

Due to the partial linearity of the force-strain relation, one can divide the problem of finding the global minimum of the energy into two steps. First for a prescribed distribution of phases $\mathbf{q}$ we need to solve the linear elastic problem (11) and compute the equilibrium function $\Psi(\mathbf{q}, d)$. The resulting equilibrium configurations are automatically local minimizers of the energy Ψ (metastable equilibria), because by virtue of (14) the stiffness matrix $\mathbf{B}$ is positive definite. Next, to obtain the global minimizers (stable equilibria), one needs to minimize $\Psi(\mathbf{q}, d)$ with respect to the phase geometry $\mathbf{q}$.

The solution of the linear system (11) can be formally written as

$$\mathbf{w} = \mathbf{B}^{-1}\mathbf{F} + \mathbf{B}^{-1}\mathbf{q} = \mathbf{w}^0 + \mathbf{w}^1. \tag{15}$$

The first term in (15) corresponds to the uniform configuration with all springs in the first phase. Without explicitly computing the inverse matrix $\mathbf{B}^{-1}$ one can show that

$$w_k^0 = \frac{F}{1+\beta}. \tag{16}$$

The second term in (15) is the non-uniform contribution to the strain field due to phase transformation. In what follows it will be convenient to express $\mathbf{w}^1$ not in terms of the variables $\mathbf{q}$, but in terms of their "derivatives" $\mathbf{p}$ defined for $1 \le i \le N-1$ by

$$p_i = q_{i+1} - q_i. \tag{17}$$

Notice that these relations can always be "integrated", yielding

$$q_k = q_1 + a\sum_{i=1}^{N} p_i \theta(k-i-1). \tag{18}$$

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

$$p_i = \begin{cases} 1 & \text{if } w_i < w_c \text{ and } w_{i+1} > w_c \text{ (I to II phase switch)} \\ 0 & \text{if } \operatorname{sign}(w_i - w_c) = \operatorname{sign}(w_{i+1} - w_c) \text{ (no phase switch)} \\ -1 & \text{if } w_i > w_c \text{ and } w_{i+1} < w_c \text{ (II to I phase switch).} \end{cases} \tag{19}$$

One can see that p_i are nonzero (± 1) only at the locations i_1, i_2, ..., i_n of the phase boundaries. To simplify the subsequent formulas we also assume that $p_N = 0$.

The explicit representation of $\mathbf{w}^1$ in terms of q_1 and p_i is derived in the Appendix B. Here we present the final result

$$\begin{aligned} w_k^1 =& \frac{q_1}{1+\beta} + \Delta \sum_{i=1}^{N} p_i \Bigg\{ \frac{\sinh[(N-i)\lambda]\cosh[(k-1/2)\lambda]}{\cosh(\lambda/2)\sinh(N\lambda)} \\ &+\theta(k-i-1/2)\left(1 - \frac{\cosh[(k-i-1/2)\lambda]}{\cosh(\lambda/2)}\right)\Bigg\}, \end{aligned} \tag{20}$$

where

$$\Delta = \frac{a}{\beta+1} \tag{21}$$

is the macroscopic transformation strain and

$$\lambda = 2\text{arccosh}\frac{1}{\sqrt{|\beta|}}. \tag{22}$$

Solution of (11) can now be obtained as the sum of (16) and (20). By applying the constraint (2) we can also obtain the relation between the force F and the total displacement d:

$$F = (\beta+1)(d - \Delta\frac{l}{N}), \tag{23}$$

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

$$l = \frac{Nq_1}{a} + \sum_{i=1}^{N}(N-i)p_i. \tag{24}$$

Substituting (23), (24) into (15), (16) and (20) we obtain the general solution of our discrete problem

$$\begin{aligned} w_k =& d + \Delta \sum_{i=1}^{N} p_i \Bigg\{ \frac{\sinh[(N-i)\lambda]\cosh[(k-1/2)\lambda]}{\cosh(\lambda/2)\sinh(N\lambda)} + \frac{i}{N} - 1 \\ &+\theta(k-i-1/2)\left(1 - \frac{\cosh[(k-i-1/2)\lambda]}{\cosh(\lambda/2)}\right)\Bigg\}. \end{aligned} \tag{25}$$

To compute the total energy, let $\tilde{\mathbf{w}}$ with components $\tilde{w}_k = w_k - d$ denote the $\mathbf{p}$-dependent part of the strain field $\tilde{\mathbf{w}}$. Then (10) can be rewritten as

$$\Psi = \frac{\beta+1}{2}d^2 - a(d-w_c)\frac{l}{N} + \tilde{\Psi}, \tag{26}$$

where

$$\tilde{\Psi} = \frac{1}{N}\left(\frac{1}{2}\mathbf{B}\tilde{\mathbf{w}} - \mathbf{q}\right)\cdot\tilde{\mathbf{w}} \tag{27}$$

is the portion of energy due to phase changes (note that $\tilde{\Psi} = 0$ at a trivial solution). Observing that in equilibrium (Clapeyron's theorem)

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

and using (18) to eliminate $\mathbf{q}$, we finally obtain

$$\Psi = (1+\beta)\left\{\frac{1}{2}d^2 - \Delta(d-w_c)\frac{l}{N} + \frac{\Delta^2}{2N}\mathbf{J}\mathbf{p}\cdot\mathbf{p}\right\}, \tag{28}$$

where the kernel matrix $\mathbf{J}$ is given by

$$J_{ki} = \frac{\sinh[(N-i)\lambda]\sinh[\lambda k]}{\sinh(N\lambda)\sinh\lambda} + k\left(\frac{i}{N}-1\right) + \theta(k-i-1)\left[k-i-\frac{\sinh[(k-i)\lambda]}{\sinh\lambda}\right]. \tag{29}$$

Notice that for a given distribution of phases the loading parameter d cannot take arbitrary values due to the requirement that strains on different sides of a phase boundary must correspond to different energy wells ($w_k < w_c$ for phase I and $w_k > w_c$ for phase II). This leads to the constraint

$$p_i(w_i - w_c) < 0, \quad p_i(w_{i+1} - w_c) > 0 \quad \text{for } i\text{: } p_i \neq 0, \tag{30}$$

which in turn generates specific bounds on $d(i_1, i_2, ..., i_n)$. For example, for a single-interface solution ($|\mathbf{p}| = 1$) we must have $d_-(i) < d(i) < d_+(i)$, where

$$d_\pm(i) = w_c + \Delta\left\{\frac{\sinh[(N-i)\lambda]\cosh[(i\pm 1/2)\lambda]}{\cosh\dfrac{\lambda}{2}\sinh(N\lambda)} + \frac{i}{N} - 1\right\}. \tag{31}$$

Similar explicit bounds can be found for configurations with two or more interfaces.

Formulae (23), (28) and (31) are illustrated in Fig. 3 which shows the homogeneous solution and solutions with a single phase boundary at various locations. Each of the single-interface branches, parametrized by i - the num-

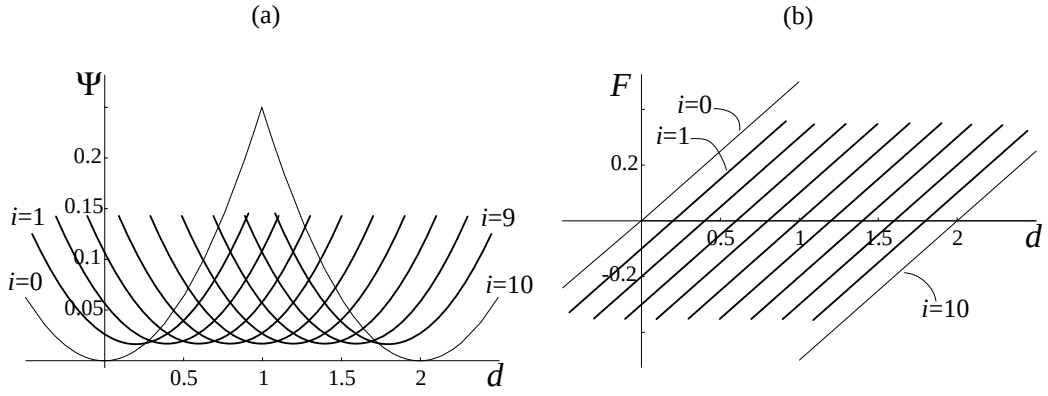

Fig. 3. The overall energy-strain and force-strain relations along the trivial ($i = 0$, N) and the single-interface ($1 < i < N$) metastable solutions. Here $N = 10$, $\beta = -1/2$, $w_c = a = 1$.

ber of springs in phase II, begins at $d = d_-(i)$, where $w_i = w_c$, and ends at

$d = d_+(i)$, where $w_{i+1} = w_c$. In the interval $d_-(i) < d < d_+(i)$ the force generated by these nontrivial solutions increases linearly with d, while the energy is parabolic. Notice that due to the symmetry of the problem, the ith branch presented in Fig. 3 with phase II located to the left of the interface is indistinguishable from the $(N-i)$th branch with phase II located to the right of the interface.

Equilibria with two interfaces and various arrangements of phases are shown in Fig. 4 for a chain with six springs. Although the structure of the corresponding

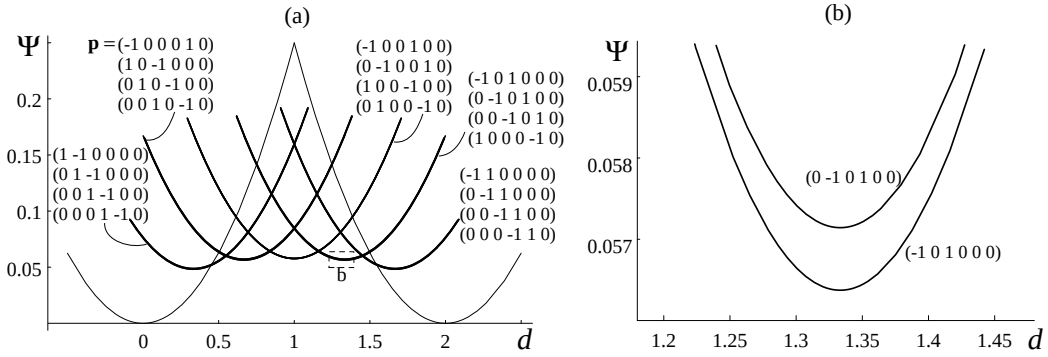

Fig. 4. (a) Energy-strain relation for the homogeneous solution (thin line) and two-interface solutions with various locations of the interfaces (various $\mathbf{p}$, thick lines). (b) The blow-up of (a) around configurations with $\mathbf{p} = (0\ -1\ 0\ 1\ 0\ 0)$ and $\mathbf{p} = (-1\ 0\ 1\ 0\ 0\ 0)$. Parameters: $\beta = -1/2$, $w_c = a = 1$, $N = 6$.

energy-strain diagram is similar to the one shown in Fig. 3a, the two-interface branches are located higher than the branches with a single interface. Observe that unlike the case with no NNN interactions (e.g. Puglisi and Truskinovsky, 2002b), in which the energy of an equilibrium configuration is determined solely by the volume fractions of the phases, in the nonlocal model the actual locations of the interfaces play an important role. For example, the blow-up in Fig. 4b shows that the energy of the branch with $\mathbf{p} = (0\ -1\ 0\ 1\ 0\ 0)$ (phase II for $k < 2$ and $k > 4$ and phase I for $2 \le k \le 4$) is higher than the energy of a branch with $\mathbf{p} = (-1\ 0\ 1\ 0\ 0\ 0)$ which has the same volume fraction of phase II, but different phase geometries.

Inside the family of metastable solutions with a given number of interfaces n one can now select the global minimizers of $\Psi(i_1, ..., i_n, d)$. The resulting curves $\hat{\Psi}(n, d)$, $n \le 4$, are shown in Fig. 5 for a chain with six springs. We observe that among the nontrivial optimal equilibria, single-interface solutions have the lowest energy, which is an expected result for large N (Carr et al., 1984; Triantafyllidis and Bardenhagen, 1993). In the limit of small γ branches with different number of interfaces tend to overlap, and single-interface solutions eventually lose their special status.

Finally, we remark that the structure of of local minimizers obtained an-

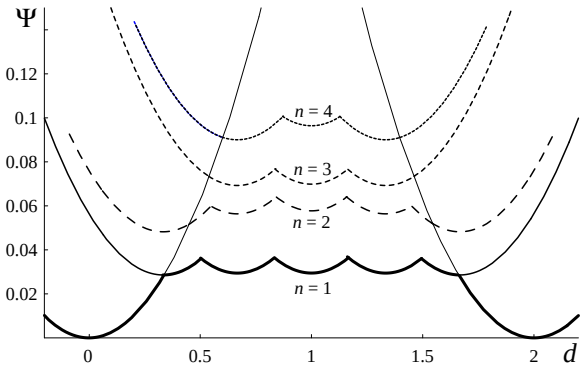

Fig. 5. Energy-strain relation for the global minimizers of the families of solutions with up to four interfaces. Here $N = 6$, $\beta = -1/2$, $w_c = a = 1$.

alytically in this section is very similar to the numerical results of Rogers and Truskinovsky (1997) for a discrete model with nonlocal interactions that spread beyond next-to-nearest neighbors.

4 Energy barriers

In this section we go beyond the analysis of local minima and reconstruct the relevant sections of the energy landscape for our discrete system. In particular, we fix the loading parameter d and find the energy barriers separating the neighboring metastable states. This information will be used in the next section where we discuss different strategies of switching between metastable configurations in the quasistatically loaded system.

For simplicity, we restrict our attention to the configurations with at most one phase boundary, which is consistent with experimental observations (e.g. Krishnan, 1985) at a very slow loading (see Appendix C for the treatment of the case where two-interface solutions may become relevant). More precisely, we assume that at a given d, the chain can be equilibrated in either homogeneous state ($i = 0$ or N) or in a variety of two-phase states with a single phase boundary at $k = i$. The generic picture is illustrated in Figure 6a where we see that at $d = 0.5$ a chain with $N = 10$ can assume metastable configurations with $i = 2$ (or 3), $i = 1$ (or 4), $i = 0$, $i = 5$ and $i = 6$, where the equilibria are listed in the order of increasing energy.

To evaluate the barriers separating neighboring local minima, consider a generic one-interface equilibrium configuration $w_k(i)$, with all springs to the left of the ith particle stretched in phase II and try to connect it with another one-interface equilibrium, $w_k(i+1)$, having one extra spring in phase II. Notice that in order to switch between $w_k(i)$ and $w_k(i+1)$, the system must follow some non-equilibrium path along which the $(i+1)$th spring changes phase.

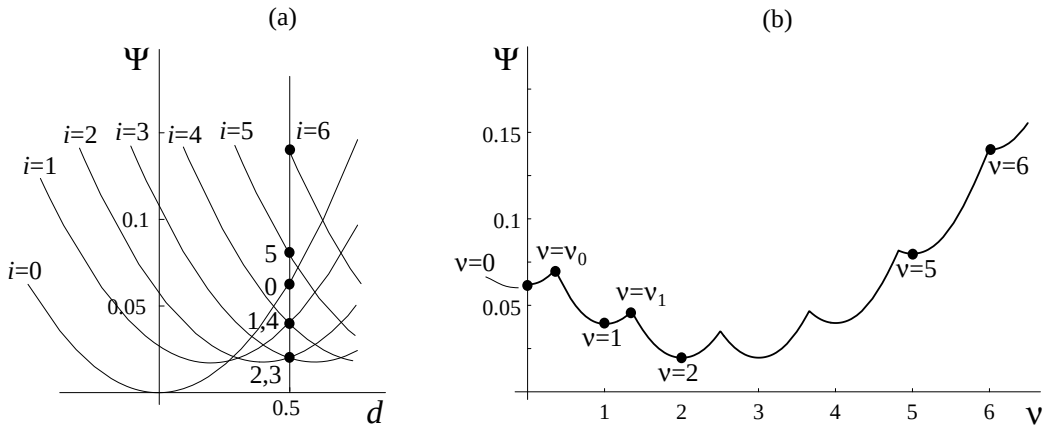

Fig. 6. (a) Energies of the single- and zero-interface branches of equilibria available at $d = 0.5$. (b) Peierls-Nabarro energy landscape along the path $w_k(\nu)$ connecting various metastable equilibria at $d = 0.5$. Parameters: $N = 10$, $\beta = -1/2$, $w_c = a = 1$.

Among infinitely many possible paths connecting the configurations $w_k(i)$ and $w_k(i+1)$, we are interested in the one passing through the minimal energy barrier. To generate such a path, it is natural to choose the strain w_{i+1} as the order parameter and minimize the total energy with respect to all "nontransforming" variables w_k with $k \le i$ and $k \ge i+2$ for a fixed w_{i+1}. The necessary conditions of this constrained equilibrium take the form

$$(1+\frac{\beta}{2})w_k + \frac{\beta}{4}(w_{k+1}+w_{k-1}) = \begin{cases} F+a \text{ for } k \le i \\ F \qquad \text{for } k \ge i+2, \end{cases} \tag{32}$$

Note that equations (32) must be supplemented by the boundary conditions (2) and (6).

To show that the configurations satisfying (32), (2) and (6) have the lowest energy among all constrained states with a given w_{i+1} we need to check the positive definiteness of the matrix $\mathbf{B}^{(i+1)}$ obtained from (A.1) by deleting the $(i+1)$th row and the $(i+1)$th column. In the generic case when $0 < i < N-1$ the matrix $\mathbf{B}^{(i+1)}$ is indeed positive definite as can be easily seen by writing the quadratic form

$$\mathbf{B}^{(i+1)}\mathbf{w}\cdot\mathbf{w} = (1+\beta)\sum_{k=1}^{N-1} w_k^2 - \frac{\beta}{4}\left(\sum_{k=1,k\neq i}^{N-2}(w_{k+1}-w_k)^2 + w_{i+1}^2 + w_i^2\right),$$

and applying the condition that β is in the range (14). Then, if $i = N-1$ the first $N-1$ principal minors of $\mathbf{B}$ coincide with the principal minors of $\mathbf{B}^{(i+1)} = \mathbf{B}^{(N)}$ and the positive definiteness of $\mathbf{B}^{(i+1)}$ follows from the positive definiteness of $\mathbf{B}$. By symmetry, the same argument applies to the matrix $\mathbf{B}^{(1)}$ $(i = 0)$.

The general solution of (32), (2) and (6) can be represented in the parametric form

$$w_k(\nu) = \begin{cases} d - \Delta\Big\{\dfrac{\nu}{N} - 1 + C_1(\sinh(\lambda k) - \sinh(\lambda(k-1)))\Big\}, & k \le [\nu]+1 \\ d - \Delta\Big\{\dfrac{\nu}{N} + C_2(e^{\lambda k} - e^{\lambda(k-1)} - e^{\lambda(2N-k)} + e^{\lambda(2N-k+1)})\Big\}, & k \ge [\nu]+1. \end{cases} \tag{33}$$

where $[\nu] = i$ and $[\nu]$ denotes the integer part of ν. It can be directly checked that (33) solves (32) for $k \le [\nu]$ and $k \ge [\nu]+2$, with the Lagrange multiplier F given by

$$F = (1+\beta)\Big(d - \frac{\nu}{N}\Delta\Big). \tag{34}$$

Now, by matching the strains at $k = i+1 = [\nu]+1$ and imposing (2), we obtain

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

and

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

In particular, we have

$$w_{[\nu]+1}(\nu) = d - \Delta\Big(\frac{\nu}{N} - 1\Big) + C_1(\sinh[([\nu]+1)\lambda] - \sinh[[\nu]\lambda]). \tag{35}$$

which furnishes a relation between the order parameters w_{i+1} and ν. Observe that parameter w_{i+1} oscillates as the function of i with period 1, while ν increases monotonically. We also notice that the integer values of ν correspond to the metastable equilibrium configurations. The energy $\Psi(\nu)$ obtained by substituting (33) in (26) can be interpreted as a Peierls-Nabarro (PN) landscape for our discrete system with the valleys at $\nu = i$.

One can see that in order to move from one local minimum of the PN landscape at, say, $\nu = i$, to the next one at $\nu = i+1$ the system must climb a Peierls barrier located at $\nu = \nu_i$ which is defined by

$$w_{i+1}(\nu_i) = w_c. \tag{36}$$

The height of the Peierls barrier is then given by $\delta\Psi^{i\to i+1} = \Psi(\nu_i) - \Psi(i)$ and can be explicitly computed using (26) and (27).

A typical structure of the PN landscape is illustrated in Fig. 6b. One can see that due to piecewise linear structure of the model the PN landscape consists of a set of parabolae with local minima at $\nu = i$ (metastable equilibria) and sharp local maxima at $\nu = \nu_i$ (saddle points of the original energy). The corresponding strain profiles are shown in Figure 7.

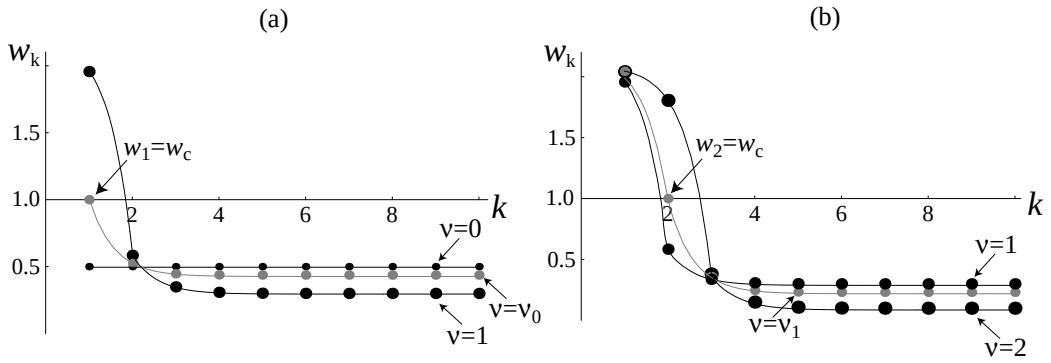

Fig. 7. Strain profiles associated with nucleation and incremental growth of the new phase. Configurations with $\nu = 0$, $\nu = 1$ and $\nu = 2$ are metastable equilibria. The saddle point configurations have $\nu = \nu_0$ and $\nu = \nu_1$. Parameters are the same as in Fig. 6.

The dependence of the height of the Peierls barriers on d is illustrated in Fig. 8. One can make a general observation that for a given i, the Peierls barrier between the ith and $(i + 1)$th metastable branches is maximal at the point of their intersection and decreases to zero at the end of the branch, where ν_i coincides with i. We also observe that at sufficiently large $|\beta|$ the Peierls barrier for the transition $0 \to 1$ is higher than for several subsequent transitions. For example, at $N = 10$ and $\beta = -1/2$, the barrier for the transition $1 \to 2$ is lower than the barrier for the transition $0 \to 1$ (see Fig. 8). Notice that the barrier $1 \to 2$ vanishes at $d_+(1)$ and beyond this point, the branch $i = 1$ does not exist any more. Therefore the transition $0 \to 1$ deteriorates into the transition

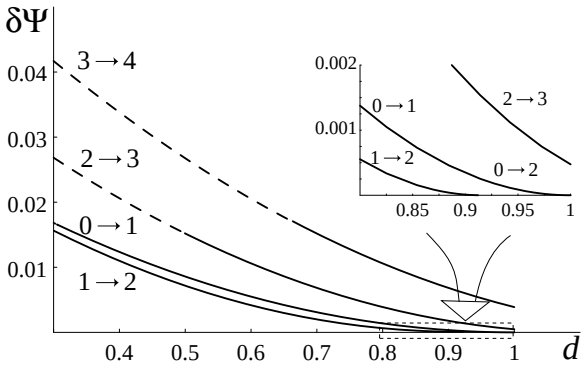

Fig. 8. Peierls barriers for several transitions $i \to i + 1$ parametrized by d. The dashed lines indicate the intervals where the transition is energetically unfavorable. The insert is the blow-up of the selected region. Parameters: $N = 10$, $\beta = -1/2$, $w_c = a = 1$.

$0 \to 2$ with the barrier vanishing exactly at $d = w_c$. A similar calculation for $N = 20$ and $\beta = -1/2$ shows that the Peierls barrier for the transition $0 \to 1$ is higher than two subsequent barriers for the transitions $1 \to 2$ and $2 \to 3$; as the last two barriers vanish at sufficiently large d, the transition $0 \to 1$ first transforms into $0 \to 2$ and then into $0 \to 3$. In general, as we show in the next

section, the number of springs involved in the first nucleation event increases with N and tends to infinity in the continuum limit.

5 Nucleation peak and hysteresis

We have shown that if the chain, driven by external loading d, is constrained to stay in the global minimum of the energy (Maxwell strategy), it has to overcome maximally high energy barriers during each switching event. Alternatively, the system may be constrained to stay in a given local minimum until the minimal energy barrier around this local minimum reaches a critical threshold. In particular, the *maximal delay strategy* would require that the system follows each metastable branch until the minimal barrier is equal to zero (so that the corresponding state becomes absolutely unstable) and then switches to the nearest local minimum by following a path of the steepest descent. For chains with only local (NN) interactions, various aspects of these strategies and their relation to the dynamical extensions of the model are discussed in Puglisi and Truskinovsky (2002a,b). Here we generalize these results for the case when the system is nonlocal.

Suppose that the branch switching takes place when the energy barrier equals a certain critical value H which is determined by the level of fluctuations or imperfections in the system. Two of the resulting paths are shown in Fig. 9 which illustrates the overall behavior of the NNN chain subjected to the quasistatic loading in a hard device. One can see that as d increases from zero,

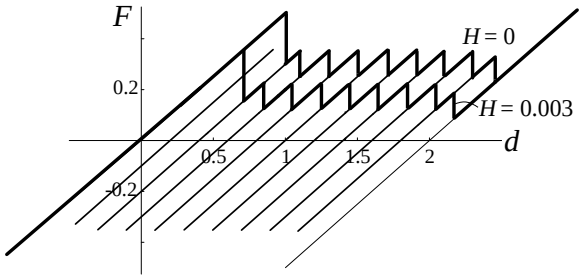

Fig. 9. Branch-switching sequence for the path of maximal delay, $H = 0$, and for the path with a finite critical energy barrier $H = 0.003$. Parameters: $\beta = -1/2$, $N = 10$, $w_c = a = 1$.

the system initially stays in the trivial (homogeneous) configuration with all springs in phase I. As the energy barriers decrease with d, the chain eventually reaches the state when the smallest energy barrier in the vicinity of the trivial metastable equilibrium becomes equal to H. At this moment nucleation takes place and the system escapes from the local minimum through the first saddle point with a subcritical height. As the loading continues after the nucleation event, the phase boundary propagates along the chain in a stick-slip fash-

ion, with the system getting temporarily trapped in each of the metastable equilibria (parametrized by i). The resulting graphs of the force $F(d)$ exhibit serrations which are due to the discreteness of the system and smoothen in the continuum limit.

Although the overall behavior of the nonlocal (NNN) system is in many respects similar to the case of the local (NN) system, there are also some noticeable differences. The most important one is the presence of the nucleation peak: the force required for phase nucleation is distinctly higher than the average force at which subsequent propagation of a phase boundary takes place. Observe also that along both paths shown in Fig. 9 the first two springs change phase simultaneously, which never happens in the NN system. Such massive nucleation occurs when the energy barrier for the transition $0 \to 1$ is higher than the barriers for several subsequent transitions which takes place only when NNN interactions are sufficiently strong (high enough $|\beta|$) and the total number of springs is sufficiently large. Notice, however, that after the first nucleation event the subsequent propagation of the interface in the NNN system takes place at the smaller average force but involves, exactly as in the NN system, only one spring changing phase at a time.

The physical reason for the appearance of a nucleation peak is the nonlocal character of interparticle interactions leading to the formation of the internal boundary layers around the phase boundary (see a different nonlocal model exhibiting similar behavior in Rogers and Truskinovsky (1997)). Thus, before the first nucleation event all springs are stretched uniformly, whereas after the nucleation, the springs in phase I that are closer to the phase boundary have higher strain and hence are closer to the critical threshold than the springs far away. This helps to trigger the subsequent switching events and results in the smaller force required for the propagation of a phase boundary than the associated nucleation force. In the limiting case of no NNN interactions ($\beta = 0$), 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 non-transformed springs simultaneously. Thus in this case phase propagation effectively reduces to successive nucleation events in the shorter and shorter chains which requires the same critical force.

To obtain an analytical upper bound for the size of the nucleation peak, we restrict our attention to the barrierless path (maximum delay strategy, $H = 0$). Recall that in the present piecewise linear model the barrierless nucleation event occurs when the force reaches the spinodal limit $F_{\max} = (\beta + 1)w_c$. On the other hand, the jump of the interface from $k = i$ to $k = i + 1$ takes place at $F(i, N) = F_{\rm M} + F_{\rm P}(i, N)$, where

$$F_{\rm M} = (\beta + 1)w_c - \frac{a}{2} \tag{37}$$

is the *Maxwell force* and $F_{\rm P}(i,N)$ is the *Peierls force* given by

$$F_{\rm P}(i,N) = F(d_+(i),i) - F_{\rm M} = a\left(\frac{\cosh\left[\left(\frac{1}{2}+i\right)\lambda\right]\sinh[\lambda(N-i)]}{\cosh\frac{\lambda}{2}\sinh(N\lambda)} - \frac{1}{2}\right). \quad (38)$$

Numerical computations show that the function $F_{\rm P}(i,N)$ has almost no dependence on i except in the narrow boundary layers near $i=1$ and $i=N-1$. In the limit of infinite N the Peierls force (38) approaches the limiting value

$$\lim_{N\to\infty} F_{\rm P}(i,N) = F_{\rm P} = \frac{a}{2}\sqrt{1+\beta}. \quad (39)$$

One can see that as NNN interactions get weaker (β approaches zero), the Peierls force tends to the spinodal limit $F_{\max} - F_M$ and the nucleation peak disappears.

The limiting configuration of a hysteresis loop in the continuum limit is shown in Fig. 10. Notice that although the serrations disappear, the nucleation peak remains finite and the force drops by the amount

$$\tau = (1+\beta)w_c - F_P = \frac{1}{2}a\left(1-\sqrt{1+\beta}\right). \quad (40)$$

This quantity is positive as long as $-1<\beta<0$ and is always less than $a/2$ - the difference between the spinodal and Maxwell forces. The half-height of

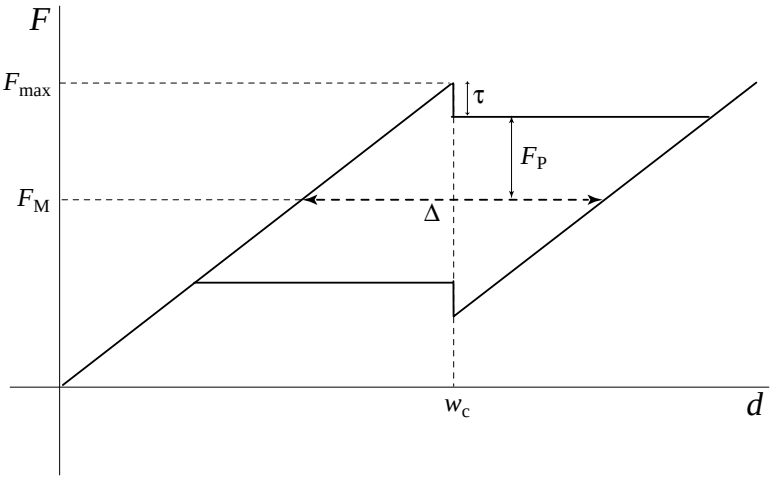

Fig. 10. The maximum hysteresis loop in the continuum limit for the nonlocal (NNN) model with $-1<\beta<0$.

the narrow part of the limiting loop is then given by $F_{\rm P}$ from (39).

To estimate the number of springs participating in the nucleation event in the continuum limit we begin with a finite N and recall that barrierless the

nucleation takes place at $d = w_c$. Therefore we must find the first equilibrium branch with nonzero i which is defined at this value of d. Setting $d_\pm$ from (31) equal to w_c, we obtain the equation for i

$$\frac{\sinh[(N-i)\lambda]\cosh[(i\pm 1/2)\lambda]}{\cosh\dfrac{\lambda}{2}\sinh(N\lambda)} = 1 - \frac{i}{N}. \tag{41}$$

Now, for $N >> i$ and sufficiently large i the left-hand side of (41) can be approximated by $(1+e^{-\lambda})^{-1}$. Hence the number of springs changing phase during nucleation in a sufficiently long chain is given by

$$i_{\text{nuc}} = \left[\frac{N}{1+e^{\lambda}}\right] + 1. \tag{42}$$

This implies that in the continuum limit the number of springs involved in nucleation is infinite. The size of the transformed portion of the chain, however, is finite and equals to

$$l_0 = \frac{1}{1+e^{\lambda}}. \tag{43}$$

Formula (43) gives the upper bound for the size of the martensite band formed during the nucleation event. When nonlocal interactions are absent (NN model, $\beta = 0$, $\lambda = \infty$) the nucleus contains only one spring and hence $l_0 = 0$.

6 Verification of the model

We now use the explicit formulae from the previous section for the interpretation of the experiments reporting the height of the nucleation peak. The goal is to obtain a bound on the value for the nonlocality measure γ and compare it with results of the independent estimates. The expression for γ from (13) and (40) takes the form

$$\gamma = \frac{E}{4}\left[1 - \left(1 - \frac{2\tau}{E\Delta}\right)^{-2}\right], \tag{44}$$

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 data of Shaw and Kyriakides (1995) on NiTi wires. Thus, in the experiment conducted at 70° and the loading rate $4\times10^{-5}\ s^{-1}$, the measurements gave $\tau =$ 0.039 GPa, and $\Delta = 3.97\%$. The Young's moduli of austenite and martensite material phases are different, $E_A = 56.7$ GPa and $E_M = 27.5$ GPa at 70°, respectively. Since our model assumes equal moduli in the two phases, we can

obtain only upper and lower bounds for γ. By using separately the moduli for austenite and martensite, we estimate γ to be between -1.1 and -1 GPa, which implies that β is in the range $-0.137 < \beta < -0.068$. Similar estimate in Truskinovsky and Vainchtein (2003) for CuAlNi yields $\beta = -0.0299$.

One can independently estimate γ by assuming that the interactions between particles are governed by the Lennard-Jones potential. In this case we have $\varepsilon\phi_1(\frac{r}{\varepsilon} - 1) = 2\varepsilon\phi_2(\frac{r}{2\varepsilon} - 1) = U(r)$, where $\phi_1(w)$ and $\phi_2(w)$ are the energy densities for NN and NNN springs, respectively, and $U(r)$ has the form

$$U(r) = \frac{K\varepsilon}{72}\left[\left(\frac{\varepsilon}{r}\right)^{12} - 2\left(\frac{\varepsilon}{r}\right)^{6}\right]. \tag{45}$$

The coefficients in $U(r)$ are selected in such a way that elastic modulus in the potential well located at $r = \varepsilon$ equals K. Linearizing around the unstretched homogeneous state with the spacings $r = \varepsilon$ and $r = 2\varepsilon$, we obtain (Charlotte and Truskinovsky, 2002)

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

This yields $\beta = -0.0177$. With the Young's modulus $E_A = 56.7$ GPa, this is equivalent to $\gamma = -0.255$ GPa which despite the rather rigid form of the potential (45) is within a reasonable range from the values obtained above from the experimental data.

Finally, one can use the above estimates of β and our formula (43) to predict the initial size of the martensite band L_0. The range $-0.137 < \beta < -0.068$ obtained above yields $0.001 < L_0/L < 0.005$, where L is the size of the specimen. While we could not find direct experimental measurements of L_0 in the literature, this parameter was estimated (Sun and Zhong, 2000) to be on the order of the specimen's diameter D. This appears to be roughly in agreement with our own estimate in the case of thin NiTi wires where in a typical sample $D/L \sim 0.0025$ (Leo et al., 1993)).

7 Comparison with the strain-gradient model

In this section we compare the exact results for the discrete model with the predictions of a strain-gradient quasi-continuum approximation which retains the length scale of the discrete model while simplifying its dispersive properties. The formal derivation of the strain-gradient model from the discrete model can be found in Mindlin (1965). By neglecting the terms of third and higher order in ε, we obtain the energy functional (up to a null Lagrangian)

$$\Psi = \int_0^1 [\phi(w) + \frac{1}{2}\alpha\varepsilon^2(w')^2]dx, \tag{46}$$

where

$$\alpha = -\frac{1}{12}(1 + 4\beta). \tag{47}$$

This functional must be minimized subject to the constraint

$$\int_0^1 w(x)dx = d. \tag{48}$$

The natural boundary conditions

$$w'(0) = w'(1) = 0 \tag{49}$$

represent the continuum analog of the boundary equations (6) in the discrete problem. The Euler-Lagrange equations resulting from this quasi-continuum formulation can be written in the form $\phi'(w) - \alpha\varepsilon^2 w'' = F$ (analog of (3)), where $\phi'(w) = (\beta + 1)w - a\theta(w - w_c)$.

The variational problem (46), (48) is well posed when α is positive which implies $\beta < -1/4$. Observe that the estimate for β obtained in the previous section from experimental data is outside this range which implies that at least some features of the relevant physical situation may be misrepresented by the quasi-continuum approximation.

It is instructive to compare the general structure of the metastable solutions in the discrete model with the corresponding solutions in the strain-gradient approximation (46). 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. 11, with the parameters chosen to match those for the discrete model with $N = 20$. We observe that while the quasi-continuum 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.

Observe that 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.

While stability of configurations with multiple interfaces is common to both the NN and the NNN models, strong coupling among the bi-stable elements, reduces the number of metastable states and narrows the hysteresis. Moreover, as the coupling gets stronger, the quasi-continuum approximation becomes

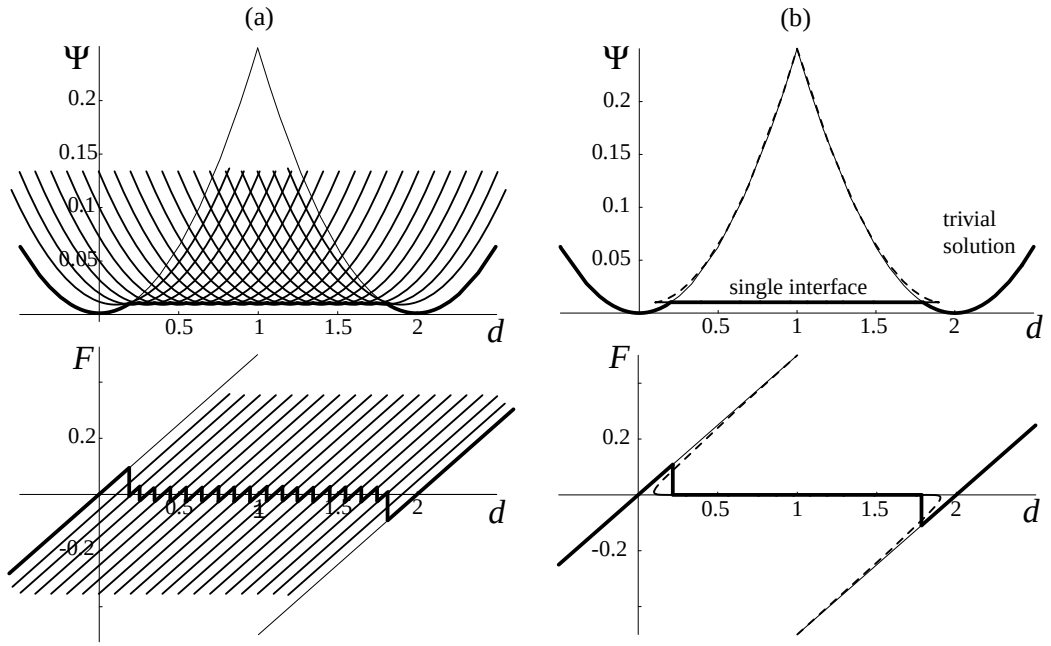

Fig. 11. The energy-strain and the force-strain relations for the absolute energy minimizers (thick line) and the metastable single-interface solutions (thin lines) for the discrete chain with $N = 20$ (a) and for the strain-gradient approximation, with $\varepsilon = 1/20$(b). Other parameters: $\beta = -1/2$, $w_c = a = 1$. Unstable single-interface solutions are shown by dashed lines.

more and more appropriate and the hysteresis asymptotically approaches the shape predicted by the strain gradient model. This is clearly seen in the numerical experiments reported by Ye et al. (1991).

8 Conclusions

In this paper we studied the effect of the harmonic coupling between bi-stable elements on the macroscopic force-strain response of a discrete chain placed in a hard loading device. The choice of the piecewise linear constitutive relations allowed us to obtain explicit solutions describing inhomogeneous metastable configurations with an arbitrary number of phase boundaries. While the general response was found to be similar in both local and nonlocal systems, we have also indicated important differences. The main feature of the system with nonlocal interactions is that 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 what one obtains in the local model. Another important difference is that in the nonlocal model more than one element may have to be deformed during the nucleation event. This leads to the appearance of a nucleation peak which has been experimentally detected not only in shape memory alloys but also

in mild steels. The fact that the propagation of the phase boundary following the nucleation takes place at a lower load than the nucleation itself is one of the main signatures of the nonlocal models. 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. The availability of the exact formula relating the size of the nucleation peak to the microscopic parameters of the lattice allowed us to estimate the value of the stiffness of the NNN springs which produced results compatible with the independent lattice computations. We have shown that the strain-gradient continuum approximation captures the global minimum of the discrete problem but fails to predict a realistic hysteresis. This systematic discrepancy points towards the importance of alternative quasi-continuum techniques that would result in the models capable of capturing local minimizers.

Acknowledgments. This work was supported by the NSF grants DMS-0102841 (L.T.) and DMS-0137634 (A.V.).

Appendix

A Necessary and sufficient conditions of stability for the trivial solution of the discrete problem

The trivial (homogeneous) solution of the problem (3), (2) and (6) is stable if and only if the $N \times N$ tridiagonal matrix

$$\mathbf{B} = \frac{1}{\varepsilon}\left[\frac{\partial^2 \Psi}{\partial w_i \partial w_j}\right]_{w_k = w} = \begin{bmatrix} K+3\gamma & \gamma & 0 & \dots & 0 \\ \gamma & K+2\gamma & \gamma & \ddots & 0 \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & 0 & \gamma & K+2\gamma & \gamma \\ 0 & \dots & 0 & \gamma & K+3\gamma \end{bmatrix} \tag{A.1}$$

is positive definite. We now show that the conditions (7) are both necessary and sufficient for stability.

First consider the case $\gamma \le 0$. Then $K + 4\gamma \sin^2(\pi/(2N)) \ge K + 4\gamma$, and it suffices to show that $(7)_1$ is necessary and sufficient. Indeed, in this case if $K + 4\gamma > 0$ both terms in the quadratic form

$$\mathbf{Bw} \cdot \mathbf{w} = (K + 4\gamma) \sum_{k=1}^{N} w_k^2 - \gamma \sum_{k=1}^{N-1} (w_{k+1} - w_k)^2. \tag{A.2}$$

are positive (unless $\mathbf{w} = \mathbf{0}$), implying positive-definiteness of $\mathbf{B}$. Conversely, if the trivial solution is stable, we have $\mathbf{Bw} \cdot \mathbf{w} > 0$ for all nonzero vectors $\mathbf{w}$. Choosing $w_k = \delta/\sqrt{N} \neq 0$, we obtain $\mathbf{Bw} \cdot \mathbf{w} = (K + 4\gamma)\delta^2 > 0$, which implies $(7)_1$.

Now, if $\gamma > 0$, we have $K + 4\gamma \sin^2(\pi/(2N)) \leq K + 4\gamma$ and we need to show that it is $(7)_2$ which is both necessary and sufficient. First observe that in this case $K > 0$ is sufficient for stability. This follows from writing the quadratic form (A.2) as

$$\mathbf{Bw} \cdot \mathbf{w} = K \sum_{k=1}^{N} w_k^2 + \gamma \sum_{k=1}^{N-1} (w_{k+1} + w_k)^2 + 2\gamma(w_1^2 + w_N^2).$$

By the same argument as above $K + 4\gamma > 0$ is again necessary. Therefore, we consider the range

$$\{K < 0, \quad K + 4\gamma > 0\} \tag{A.3}$$

and compute the principal minors D_n of (A.1). Recalling the recursion relation between principal minors of a tridiagonal matrix (e.g., Gelfand and Fomin (1963)), we obtain

$$D_1 = K + 3\gamma, \tag{A.4}$$

$$D_k = (K + 2\gamma)D_{k-1} - \gamma^2 D_{k-2}, \quad 2 \leq k \leq N - 1, \tag{A.5}$$

$$D_N = (K + 3\gamma)D_{N-1} - \gamma^2 D_{N-2}. \tag{A.6}$$

Seeking solutions of the difference equation (A.5) in the form $D_k = \rho^k$, we obtain the characteristic equation $\rho^2 - (K + 2\gamma)\rho + \gamma^2 = 0$, which implies

$$\rho_{1,2} = \frac{1}{2}(K + 2\gamma \pm \sqrt{K(K + 4\gamma)}). \tag{A.7}$$

In the parameter range (A.3) both $\rho_{1,2}$ are complex and we can write

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

where

$$\omega = 2\arccos\sqrt{-\frac{K}{4\gamma}}. \tag{A.8}$$

The principal minors D_k, $1 \leq k \leq N - 1$, can then be written in the form

$$D_k = (-\gamma)^k (c_1 \cos(\omega k) + c_2 \sin(\omega k)).$$

Using $D_0 = 1$ and (A.4) as the boundary conditions to determine c_1 and c_2, we obtain

$$D_k = (-\gamma)^k \frac{\cos(\omega(k + \frac{1}{2}))}{\cos\frac{\omega}{2}}, \tag{A.9}$$

where $1 \leq k \leq N - 1$. Finally, (A.6) and (A.9) together yield

$$D_N = -2(-\gamma)^N \tan\frac{\omega}{2} \sin(\omega N). \tag{A.10}$$

The trivial solution is stable if and only if all principal minors of $\mathbf{B}$ are positive. Observe that $(-\gamma)^k$ is positive at even k and negative otherwise. It is easy to show that $0 < \omega < \pi$ in the range (A.3), so that $\cos\frac{\omega}{2}$ and $\tan\frac{\omega}{2}$ are both positive. Hence stability conditions reduce to the requirement that $\cos(\omega(k+\frac{1}{2}))$, $1 \le k \le N-1$, is positive for even k, negative otherwise and that $\sin(\omega N)$ is negative for even N and positive otherwise. In other words, we obtain the following intervals of stability:

$$\begin{aligned}
&\frac{-\frac{\pi}{2}+2\pi m_k}{k+\frac{1}{2}} < \omega < \frac{\frac{\pi}{2}+2\pi m_k}{k+\frac{1}{2}}, \quad \text{for } 1 \le k \le N-1, \text{ even} \\
&\frac{\frac{\pi}{2}+2\pi m_k}{k+\frac{1}{2}} < \omega < \frac{\frac{3\pi}{2}+2\pi m_k}{k+\frac{1}{2}}, \quad \text{for } 1 \le k \le N-1, \text{ odd} \\
&\frac{\pi(1+2M_N)}{N} < \omega < \frac{2\pi(M_N+1)}{N}, \quad \text{for } N \text{ even} \\
&\frac{2\pi M_N}{N} < \omega < \frac{\pi(1+2M_N)}{N}, \quad \text{for } N \text{ odd},
\end{aligned}$$

where the choice of the integers m_k, M_N makes the above inequalities compatible with $0 < \omega < \pi$. It is not hard to see that all the above inequalities are satisfied if and only if

$$\omega > \frac{\pi(N-1)}{N}.$$

By combining this inequality with (A.8), we obtain $(7)_2$.

B Explicit general solution for the discrete problem

To find the function $\mathbf{w}^1$, we observe that in the interior nodes ($2 \le k \le N-1$) w_k^1 must satisfy the difference equation

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

We seek the general solution of (B.1) in the form

$$w_k^1 = w_k^h + w_k^{in}, \tag{B.2}$$

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

$$\beta\rho^2 + 2(2+\beta)\rho + \beta = 0 \tag{B.3}$$

By writing $\rho_{1,2} = e^{\pm\lambda}$, where λ is given by (22) we obtain

$$w_k^h = C_1 e^{\lambda k} + C_2 e^{-\lambda k}, \tag{B.4}$$

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

Next, we construct a particular solution of (B.1) with $q_1 = 0$ and $p_j = 0$ for all j except $j = i$. Observe that in this case the right hand side of (B.1) is zero for $k \leq i$ and constant for $k \geq i+1$. Therefore we can seek the solution in the form

$$w_k^{p_i} = \begin{cases} w_k^{p-} = A_1 e^{\lambda k} + A_2 e^{-\lambda k} & \text{for } k \leq i \\ w_k^{p+} = A_3 e^{\lambda k} + A_4 e^{-\lambda k} + \dfrac{a p_i}{1+\beta} & \text{for } k \geq i, \end{cases}, \tag{B.5}$$

and then "glue" both sides by requiring that

$$w_i^{p-} = w_i^{p+}, \quad w_{i+1}^{p-} = w_{i+1}^{p+}. \tag{B.6}$$

Since we are looking for a particular solution, we may let $A_1 = A_2 = 0$, and thus consider $w_k^{p-} = 0$. Solving (B.6) for A_3 and A_4, we obtain

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

with Δ defined by (21). Finally, by superposition, we obtain the particular solution of (B.1)

$$w_k^{in} = \frac{q_1}{1+\beta} + \Delta\sum_{i=1}^{N} p_i \theta(k-i-1/2)\left[1 - \frac{\cosh[(k-i-1/2)\lambda]}{\cosh(\lambda/2)}\right]. \tag{B.7}$$

The general solution of (B.1) is now given by (B.2), (B.4) and (B.7). Applying the boundary conditions (6), we obtain (20).

C Effects of the boundary layers

To study the effect of the boundary layers on the nucleation phenomenon in general and the magnitude of the nucleation peak in particular, we can replace the special boundary conditions (6) aimed at suppressing boundary layers by the "zero-moment" conditions (4) with $\Psi_B = 0$. In this case, due to the missing NNN springs on the boundary the zero-interface solution is no longer trivial and possesses two symmetric boundary layers illustrated in Fig. C.1b. Next, consider the process of quasistatic loading of the chain originally in phase I.

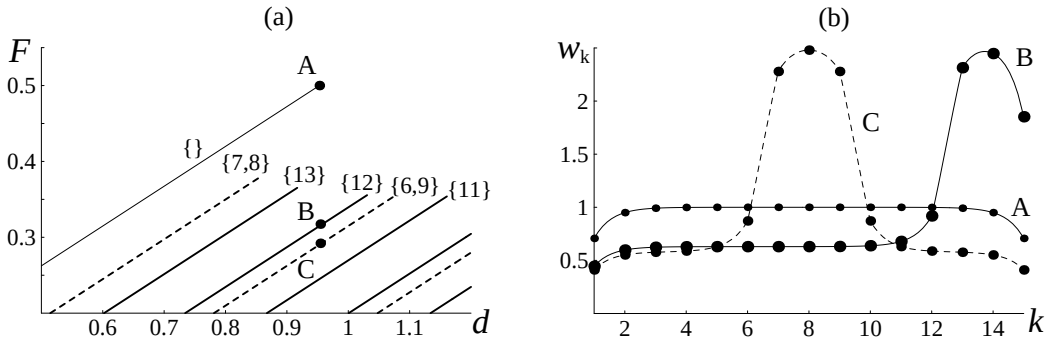

Fig. C.1. (a) The fragment of the force-elongation plane near the nucleation peak representing: single-interface (solid lines) and two-interface (dashed lines) solutions and (b) the corresponding strain profiles. The numbers in curly brackets indicate locations of phase boundaries for configurations with phase I in the left end of the chain. Parameters: $N = 15$, $\eta = -1/2$, $w_c = a = 1$.

When the critical value of the average strain $d = d_+(0)$ is reached, at least one spring in the middle of the chain passes the critical threshold $w = w_c$ and the one-phase solution becomes unstable (point A in Fig. C.1). This may lead to the formation of either two symmetric interfaces in the center (point C in Fig. C.1 with two phase boundaries located at $k = 6$ and $k = 9$) or of a single interface near one of the the boundaries (point B in Fig. C.1 with one phase boundary at $k = 12$). The computation shows that configuration B with a single interface has a slightly lower energy than the two-interface configuration C. However, the transition from A to B may require overcoming a higher barrier than in the case of the transition from A to C. Indeed, during the transition from A to C only two springs (7th and 9th, which are both slightly below w_c in A) must transform into phase II: the 8th spring is already at the critical strain. At the same time the transition from A to B requires at least three springs, that are initially much below the threshold to change phase. This suggests that in the presence of symmetric boundary layers the nucleation may takes place in the interior of the chain and this possibility can be investigated rigorously. For the subject of this paper, however, it is important to notice that the nucleation peak will survive in both cases (see Fig. C.1a).

References

Braides, A., Gelli, M. S., 2002. Limits of discrete systems with long-range interactions. Journal of Convex Analysis (to appear).

Butler, J. F., 1962. Lüders front propagation in low carbon steels. Journal of the Mechanics and Physics of Solids 10, 313–334.

Carr, J., Gurtin, M., Slemrod, M., 1984. Structured phase transitions on a finite interval. Archive for Rational and Applied Mechanics and Analysis

86, 317–351.
Charlotte, M., Truskinovsky, L., 2002. Linear elastic chain with a hyper-prestress. Journal of the Mechanics and Physics of Solids 50, 217–251.
Cottrell, A. H., Bilby, B. A., 1949. Dislocation theory of yielding and strain ageing of iron. Proceedings of Physical Society London 62/I-A, 49–62.
Ericksen, J., 1975. Equilibrium of bars. Journal of Elasticity 5, 191–202.
Fedelich, B., Zanzotto, G., 1992. Hysteresis in discrete systems of possibly interacting elements with a two well energy. Journal of Nonlinear Science 2, 319–342.
Froli, M., Royer-Carfagni, G., 2000. A mechanical model for the elastic-plastic behavior of metallic bars. International Journal of Solids and Structures 37, 3901–3918.
Făciu, C., Suliciu, I., 1994. A Maxwellian model for pseudoelastic materials. Scripta Metallurgica 31, 1399–1404.
Gelfand, I. M., Fomin, S. V., 1963. Calculus of Variations. Prentice-Hall, Englewood Cliffs.
Hall, E. O., 1970. Yield point phenomena in metals and alloys. Plenum Press, New York.
Horikawa, H., Miyazaki, S., 1988. Metallurgical Transactions A 19, 915–923.
Johnston, W. G., Gilman, J. J., 1959. Dislocation velocities, dislocation densities and plastic flow in lithium fluoride crystals. Journal of Applied Physics 30, 129–144.
Krishnan, R., 1985. Stress induced martensitic transformations. Material Science Forum 3, 387–398.
Kyriakides, S., Miller, J. E., 2000. On the propagation of luders bands in steel strips. Journal of Applied Mechanics 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 Metallurgica et Materialia 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. International Journal of of Solids and Structures 1, 417–438.
Müller, I., Villaggio, P., 1977. A model for an elastic-plastic body. Archive for Rational and Applied Mechanics and Analysis 65, 25–46.
Pagano, S., Paroni, S., 2002. A simple model for phase transformations: from discrete to the continuum problem. Tech. rep., Mathematical Institute, University of Oxford.
Puglisi, G., Truskinovsky, L., 2000. Mechanics of a discrete chain with bi-stable elements. Journal of the Mechanics and Physics of Solids 48, 1–27.
Puglisi, G., Truskinovsky, L., 2002a. A mechanism of transformational plasticity. Continuum Mechanics and Thermodynamics 14, 437–457.

Puglisi, G., Truskinovsky, L., 2002b. Rate independent hysteresis in a bi-stable chain. Journal of the Mechanics and Physics of 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. Journal of the Mechanics and Physics of Solids 43, 1243–1281.

Shaw, J. A., Kyriakides, S., 1997a. Initiation and propagation of localized deformation in elasto-plastic strips under uniaxial tension. International journal of 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 Materialia 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. International Journal of 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. Journal of Elasticity 33 (3), 259–293.

Truskinovsky, L., Vainchtein, A., 2003. Peierls-Nabarro landscape for martensitic phase transitions. Physical Review B In press.

Truskinovsky, L., Zanzotto, G., 1996. Ericksen's bar revisited: Energy wiggles. Journal of the Mechanics and Physics of 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. Let. 66, 2018–2021.