Hydrodynamic description of elastic solids with open boundary conditions undergoing a phase transition

A. C. E. Reid and R. J. Gooding
Department of Physics, Queen's University, Kingston, Ontario, Canada K7L 3N6
(Received 30 November 1993)

We present a thorough examination of the deterministic dynamics of finite, nonlinear, continuum systems which undergo elastic phase transitions. Particular attention is given to the correct implementation of open boundary conditions, which are nontrivial. As examples of the application of this formalism, in this paper we model two transformations: (i) The symmetry-preserving change of volume first-order transition undergone by some mixed-valence compounds, such as $Ce_{1-x}Th_x$. This example allows us to demonstrate how a bulk fluctuation that is subcritical (viz., does not nucleate product phase) can propagate to the surface and become supercritical (viz., does nucleate the transformed state) due to the greatly reduced nucleation barrier found at the system's only heterogeneity, its surface. (ii) The late-stage dynamics of autocatalytic twinning undergone by shear martensites, such as discussed by Bales and one of us in Phys. Rev. Lett. **67**, 3412 (1991). Such complete dynamical systems of equations involving a small number of continuum fields are a prerequisite for examining the rich sequences of patterns that form at these and other first-order transitions involving strains.

I. INTRODUCTION

Structural phase transitions in solids come in a variety of forms and are mediated by various microscopic mechanisms. A particular class of these transitions are characterized principally by a discontinuous change in the shape and symmetry of the crystalline unit cell.[1] In transitions of this type, there is no net transport of atoms across unit-cell boundaries, and the lattice remains coherent throughout the transition. If one specializes to the case of first-order transitions, these begin with the creation of a localized nucleus of product phase. Then, because of the difference in shapes between the unit cells of the product and parent phases, the nucleus is accompanied by the formation of long-ranged strain fields in the parent lattice. It is this strain field that is the principal participant in the ensuing dynamics. It is not possible to characterize the nucleation rate of these domains classically because the energy of the strain field does not admit decomposition into the "surface" and "bulk" portions assumed by a classical droplet growth model.[2] Rather, it is necessary to describe the full dynamics of the elastic field of the system in order to correctly model either the nucleation and/or growth processes.

Many examples of this type of transition have been studied experimentally. Here we shall focus on idealizations of two types of such first-order strain transitions, where there is no intercell diffusion of atoms. Firstly, we consider a very unusual type of discontinuous structural transition: the mixed-valence compound $Ce_{1-x}Th_x$ undergoes a fcc-to-fcc transition where *no* change of symmetry of the unit cell occurs—only a rearrangement of the electronic filling of the d and f levels is found.[3] So, only a discontinuous length scale change is found—a spontaneously shrinking fcc unit cell. Secondly, martensitic transitions involve non-symmetry-preserving strains, usually involving shear deformations, and lead to spectacular arrangements of polydomain configurations reminiscent of pattern formation in hydrodynamic systems.[1,4] The dynamics of such transformations are described by nonlinear, nonlocal partial differential equations that lead to fascinating bifurcating fronts, and so-called autocatalytic twin formation.[5] In this paper we shall determine the *complete* dynamical evolution of two systems modeling these transitions, paying particular attention to what happens when the transformation fronts reach the boundaries.

The first model system that we study here is finite, and one dimensional, with a single-component displacement field. The undistorted system has a uniform mass density. We model the local elastic interactions by means of a local potential dependent on the *strain* of the system at a given point. The undistorted system is metastable, and the globally stable *product* system is a homogeneous length deformation of the undistorted *parent* system. This globally stable state, viz., a straight but lengthened bar, is only achievable if the end points of the system are actually free to move. This necessitates the inclusion of free boundary conditions into the dynamics. Additionally, we include a nonlocal energy density, which depends on the *strain gradients* of the system. We examine the deterministic dynamical evolution of this model system in the presence of a phenomenological elastic viscous dissipation, using various numerical techniques. We assume that the system is immersed in a heat bath, so that the dynamics are entirely isothermal. This reduces the problem to a conceptually simple but technically involved nonlinear, nonlocal continuum mechanics problem. (Further, one may argue[6] that for such transitions the forces arising from thermal fluctuations are always much smaller than the mechanical forces, and thus to a first and very good approximation it is the deterministic dynamics that

0163-1829/94/50(6)/3588(15)/$06.00

are most important.) This phenomenological approach is somewhat abstract, but it does make the problem both more tractable and more fundamental.

The second model system that we study involves shear martensites, and the phenomenology of such transitions is discussed in detail elsewhere.[1,4,5] The important aspects of this problem are associated with the degenerate low-temperature phase, and the experimental observations showing that only for shallow quenches does such a transition lead to a single domain of martensite—for deeper quenches one finds a polydomain pattern of different variants of the low-temperature phase coexisting with small regions of the high-temperature parent phase.[5] This problem, which quite naturally falls under the category of pattern formation in elastic solids, intimately involves the behavior of the transformation fronts at the boundaries. For such transitions there is no absolute reason why one must employ open-boundary conditions, although it is certainly reminiscent of an isolated sample undergoing such a transition. Thus, here we will extend the earlier work of Bales and one of the present authors[5] and display the steady-state profile of the autocatalytic twinning dynamics associated with the pattern formation found in martensite. We wish to emphasize that we will obtain the steady-state configurations via integration of the equation of motion, and thus for this system we also require the satisfaction of the open-boundary conditions at all instants of the system's evolution.

The principal technical challenge arises from the presence of boundaries in the model system. In fact, the continuum hypothesis—that nearby parts of the system have nearby displacements—is clearly violated by a termination of the system. This brings into question the validity the continuum approximation for finite systems. Consequently, it is necessary to pay close attention to the physical implications of mathematical choices made with regard to "surface terms" that arise in a variational formulation of the equation of motion. In the body of the paper, we assume that the boundaries of the system are free to move but that the system cannot exchange mechanical energy or momentum through its endpoints. This is not the most general assumption, but is physically reasonable, and for the first model that we consider it does allow the system to access the globally stable, fully transformed state of the local potential. A more general approach to arbitrary-stress boundary conditions is described in Appendix A.

Our paper is organized as follows. In Sec. II we discuss the continuum approximation for systems with open-boundary conditions, and apply this formalism to the two-model systems of interest. In Sec. III we examine the dynamics found in the discontinuous change of length transition using a mode-based approximation for the displacement field. We note here that a related approach to a fixed-boundary condition problem has been studied in a previous paper,[7] where (like Ref. 5) the importance of the hydrodynamic character of the sound waves in strain systems was again demonstrated; here we extend that mathematical (i.e., mode-based) approach[7] to the case of free boundaries. A preliminary version of this work has been reported elsewhere.[8] We then examine the late-stage dynamics of the dynamical twinning strain evolution introduced by Bales and Gooding.[5] Finally, in Sec. IV we summarize our work on this problem to date.

II. THE CONTINUUM APPROXIMATION AND THE EQUATIONS OF MOTION

A. Change of volume transition

In this section we shall present all the necessary details to derive and justify the continuum equations of motion for the displacement-field dynamics for a discontinuous change of length transition, modeling a transition such as that undergone by $Ce_{1-x}Th_x$.[3] Then, in the next section we shall simply state the relevant equations for the case of shear martensites.

To model a discontinuous change of volume transition, it will be simplest to consider a finite, one-dimensional lattice of particles; the extension to three dimensions is straightforward. In the continuum approximation for this system we introduce a displacement field $u(x,t)$ that describes the deformations of the lattice relative to some reference configuration. A point whose reference position is x will, at time t, have position $x+u(x,t)$. The strain (tensor) for this system is denoted by $e \equiv \partial_x u$. For such a one-dimensional system these strains represent deformations that take the form of local contractions ($e<0$) or dilations ($e>0$). (We note that we will be considering finite strains, and this usually entails introducing a nonlinear combination of partial derivatives of the displacement field to adequately account for the complete distortion undergone by the system—see, e.g., Ref. 9. However, for a one-dimensional system this only involves higher-order powers of $\partial_x u$, and since we are about to invoke an expansion of the elastic potential energy in powers of the strain, in ignoring the nonlinear part of the finite strain tensor all we are doing is ignoring certain higher-order terms in the potential energy.)

The symmetry of our one-dimensional system dictates that all powers of the strain tensors are invariant.[1,10] Thus, we find that an adequate local potential energy density describing deformations of the system, of extent $-l \le x \le l$, is

$$V_{\rm loc} = \int_{-l}^{l} [\tfrac{1}{2}a(\partial_x u)^2 - \tfrac{1}{3}b(\partial_x u)^3 + \tfrac{1}{4}c(\partial_x u)^4]dx \ . \quad (2.1)$$

We choose a, b, and c so that this potential energy has a metastable local minimum for $\partial_x u = 0$, and an absolute minimum for some finite homogeneous strain. Such a potential is sketched in Fig. 1. Higher-order powers of the strain may be included in the potential without changing the physics of the transition, and only complicate the mathematics describing the dynamics. This potential energy may be understood as the isothermal case of a Landau-type phenomenological free energy for a one-dimensional bar undergoing a transition involving a change of length.[8]

We also include a gradient contribution to the potential energy, analogous to the Ginzburg term in a Landau-type free-energy expansion, given by

$$V_{\rm grad} = \tfrac{1}{2}d \int_{-l}^{l} (\partial_x^2 u)^2 dx \ . \quad (2.2)$$

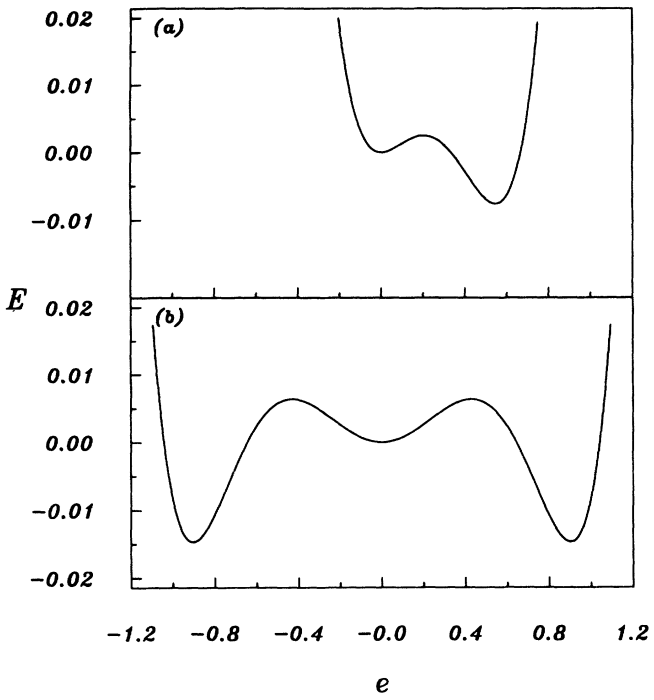

FIG. 1. The local, nonlinear elastic energy densities used for (a) the change-of-volume transition (with $a=0.2$, $b=c=1$) and (b) the shear martensite transition (with $A=0.15$, $B=C=1$), plotted as a function of strain $e\equiv\partial_x u$. In both cases $e=0$ represents the metastable parent phase. Note that the material-dependent constants b, c, B, and C may be scaled to unity, as discussed, e.g., in Refs. 5 and 7, leaving only a and A as free parameters.

This term makes the sound waves dispersive, and may be thought of as representing bond-bending forces. Again, it is possible, although usually unnecessary, to include higher-order derivatives to accurately model the phonon dispersion relation for the system of interest. The magnitude of the phenomenological constant d may be determined from the curvature of phonon dispersion curves near $q=0$ (if the curvature of the phonon dispersion curves is negative then the higher-order gradient terms *must* be included). We note that the inclusion of the gradient term is necessary to break the scale invariance of the elastic medium, and incorporates discreteness effects of the underlying lattice into our model—we have discussed the physical implications of this term elsewhere.[7,11]

In this model, the system undergoes a transition in which it lowers its internal energy by changing its length. In order for this transition to reach a stationary, reduced-energy equilibrium state, it is necessary to dissipate the excess (internal) energy from the system. We therefore include a Rayleigh dissipation functional for the system, which models the elastic viscosity, given by

$$R=\tfrac{1}{2}\gamma\int_{-l}^{l}(\partial_t\partial_x u)^2dx\ . \tag{2.3}$$

(This term has a functional form that is *independent* of the choice of boundary conditions—this is demonstrated in Appendix A.) In terms of the sound waves in the system, the Rayleigh dissipation function damps the excitations in a form that keeps them propagating at long wavelengths, consistent with general hydrodynamic considerations.[12]

The kinetic energy is given by

$$T=\tfrac{1}{2}\rho\int_{-l}^{l}(\partial_t u)^2dx\ , \tag{2.4}$$

where ρ is the linear mass density of the undistorted bar.

We use the kinetic and potential energy terms to construct a Lagrangian, $L=T-V_{\rm loc}-V_{\rm grad}$, viz.,

$$L=\int_{-l}^{l}[\tfrac{1}{2}\rho(\partial_t u)^2-\tfrac{1}{2}a(\partial_x u)^2+\tfrac{1}{3}b(\partial_x u)^3 -\tfrac{1}{4}c(\partial_x u)^4-\tfrac{1}{2}d(\partial_x^2 u)^2]dx \tag{2.5}$$

and determine the equation of motion and boundary conditions by means of the variational principle. We have included this calculation in detail because it provides the clearest derivation of the boundary conditions, which here appear in the form of surface energy terms. These boundary conditions are one of the focuses of this paper.

The equation of motion is obtained by setting the first variational derivative of the Lagrangian, with respect to the function u, equal to the first variational derivative of the dissipation function with respect to the function $(\partial_t u)$.[9] We compute these quantities, and temporarily discard the surface terms, which arise during the integration process, in order to obtain the bulk equation of motion. The more general boundary-condition analysis in Appendix A demonstrates that retaining the surface terms at this stage is equivalent to the energy argument given in this section. The variational analysis, following the notation of Ref. 9, gives

$$\frac{\delta L}{\delta u}=-\rho(\partial_t^2 u)+(\partial_x^2 u)[a-2b(\partial_x u) +3c(\partial_x u)^2]-d(\partial_x^4 u) \tag{2.6}$$

and

$$\frac{\delta R}{\delta(\partial_t u)}=-\gamma(\partial_t\partial_x^2 u)\ , \tag{2.7}$$

which then give the bulk equation of motion for the system:

$$\rho(\partial_t^2 u)=(\partial_x^2 u)[a-2b(\partial_x u)+3c(\partial_x u)^2] -d(\partial_x^4 u)+\gamma(\partial_t\partial_x^2 u)\ . \tag{2.8}$$

By itself, this equation is incomplete—the boundary conditions, which arise from the discarded surface terms, remain to be computed. A particularly powerful and physically motivated method of obtaining these conditions begins from energy considerations. The Lagrangian dynamical formalism[9] requires that the time rate of change of the energy E of the system, as it evolves according to the equation of motion, is related to the dissipation function by $\partial_t E=-2R$. Thus, in order to derive the boundary conditions on the equation of motion, we construct the quantity

$$0\equiv\partial_t E+2R=\int_{-l}^{l}[\rho(\partial_t^2 u)(\partial_t u)+a(\partial_t\partial_x u)(\partial_x u) -b(\partial_t\partial_x u)(\partial_x u)^2+c(\partial_t\partial_x u)(\partial_x u)^3 +(d\partial_t\partial_x^2 u)(\partial_x^2 u)+\gamma(\partial_t\partial_x u)^2]dx \ . \tag{2.9}$$

Integrating by parts and substituting for $\rho\partial_t^2 u$ from the equation of motion, we find that the integrands cancel exactly, and we are left with only surface terms:

$$\{(\partial_t u)\cdot[a(\partial_x u)-b(\partial_x u)^2+c(\partial_x u)^3+\gamma(\partial_t\partial_x u) -d(\partial_x^3 u)]+d(\partial_t\partial_x u)(\partial_x^2 u)\}|_{-l}^{l}=0 \ . \tag{2.10}$$

These form the boundary conditions for the equation of motion.

While mathematically speaking, only the difference between these quantities at the two boundaries need be zero, it is easy to see that physically we must require that the conditions be met at each boundary separately. The quantity proportional to $\partial_t u$,

$$a(\partial_x u)-b(\partial_x u)^2+c(\partial_x u)^3+\gamma(\partial_t\partial_x u)-d(\partial_x^3 u) \ , \tag{2.11}$$

may be interpreted as the momentum current J of the system, conjugate to the momentum density $\rho\partial_t u$. In this interpretation, the equation of motion simply becomes the continuity equation,

$$\partial_t(\rho\partial_t u)=\partial_x J \ . \tag{2.12}$$

The boundary condition then has a natural physical interpretation (e.g., as in any hydrodynamic theory)—it is the momentum flux at the boundary. If the system has free-boundary conditions, that is, $\partial_t u$ is free to be nonzero at the boundary, then the momentum current at the boundary must be zero. Mathematically, it is allowed that the momentum current at one boundary be compensated by an opposite flux at the other boundary, but this situation is not physical.

The second part of the boundary condition, proportional to $\partial_t\partial_x u$, involves $d\partial_x^2 u$. This is the generalized stress conjugate to the strain gradient, as can be seen by taking the functional derivative of the Lagrangian with respect to $\partial_x^2 u$—it is, coincidentally, proportional to the strain gradient itself, the proportionality constant being the gradient coefficient d. If $\partial_t\partial_x u$ is also nonzero, corresponding to free-boundary conditions, then the second boundary condition is a generalized stress-free boundary condition, which again must be satisfied independently at each boundary.

The free-boundary elastic continuum system with the local potential given by Eq. (2.1) is therefore completely described by the equation of motion Eq. (2.8) subject to the *nonlinear, time-dependent* boundary conditions

$$a(\partial_x u)-b(\partial_x u)^2+c(\partial_x u)^3+\gamma(\partial_x\partial_t u)-d(\partial_x^3 u)=0 \quad \text{at } x=l \text{ and } x=-l \ , \tag{2.13a}$$

$$d(\partial_x^2 u)=0 \quad \text{at } x=l \text{ and } x=-l \ . \tag{2.13b}$$

The above equations represent the continuum approximation to the equations of motion and boundary conditions for a *finite* elastic system. However, as discussed in the introduction, the validity of this approximation is not obvious. To be specific, the continuum approximation involves representing slowly varying discrete variables, the displacement of each individual ion, by a continuum field. The termination of a system is certainly not a slowly varying perturbation, so we must justify this approximation near the boundary. With this in mind, we consider the discrete representation for this system.

The system consists of some number N of atoms linked together in a chain, and in the undistorted state is uniformly spaced with some repeat distance. Each particle has a local displacement from the undistorted configuration associated with it, and these distortions are the dynamical quantities of interest. The kinetic energy of the system is given by

$$T=\sum_{i=1}^{N}\tfrac{1}{2}m\dot{u}_i^2 \ , \tag{2.14}$$

where $m=\rho L/N$, with $L(=2l)$ the total length of the systems, and the overdot represents differentiation with respect to time, viz., $\dot{u}_i\equiv du_i/dt$, etc. The local potential of the continuum system is modeled here by a sort of "nonlinear springs" between adjacent particles. There are $N-1$ such springs, which are sensitive to the difference in the displacements between adjacent particles. We denote the undistorted distance between particles, the lattice constant, by δ. The energy of these local potential springs, as a function of the distortion difference, has the same shape as the continuum local potential. The local potential for this system is then given by

$$V_{\rm loc}=\sum_{i=1}^{N-1}\left[\tfrac{1}{2}a\left[\frac{u_{i+1}-u_i}{\delta}\right]^2-\tfrac{1}{3}b\left[\frac{u_{i+1}-u_i}{\delta}\right]^3 +\tfrac{1}{4}c\left[\frac{u_{i+1}-u_i}{\delta}\right]^4\right] . \tag{2.15}$$

The discrete analogue of the nonlocal potential is a next-nearest-neighbor energy term, which favors even spacing between adjacent springs. There are $N-2$ terms of this type, given by

$$V_{\rm grad}=\sum_{i=2}^{N-1}\tfrac{1}{2}d\left[\frac{u_{i+1}+u_{i-1}-2u_i}{\delta^2}\right]^2 . \tag{2.16}$$

Finally, there is a dissipation-function term, which damps the velocity differences between adjacent particles, given by

$$R=\sum_{i=1}^{N-1}\tfrac{1}{2}\gamma\left[\frac{\dot{u}_{i+1}-\dot{u}_i}{\delta}\right]^2 . \tag{2.17}$$

In the foregoing, the parameters a, b, c, d, and γ are *the same* as in the continuum case, which is approached in the limit of large N and small δ such that $N\delta=L$ remains constant.

The dynamics for this discrete approximation to the

continuum system can then be obtained by standard, finite-dimensional Lagrangian analysis on the dynamical variables u_i. Since the individual terms in the Lagrangian must involve variables at two or more sites, there are different kinds of resulting equations of motion. For an index k corresponding to the bulk regime, that is, k not within two spaces of the end points, we obtain the equation of motion

$$m\ddot{u}_k=\frac{a}{\delta^2}(u_{k+1}+u_{k-1}-2u_k)-\frac{b}{\delta^3}[(u_{k+1}+u_{k-1}-2u_k)(u_{k+1}-u_{k-1})]$$
$$+\frac{c}{\delta^4}[(u_{k+1}+u_{k-1}-2u_k)(u_{k+1}^2+u_k^2+u_{k-1}^2-u_{k+1}u_k-u_ku_{k-1}-u_{k+1}u_{k-1})]$$
$$-\frac{d}{\delta^4}(u_{k-2}-4u_{k-1}+6u_k-4u_{k+1}+u_{k+2})+\frac{\gamma}{\delta^2}(\dot{u}_{k+1}+\dot{u}_{k-1}-2\dot{u}_k)\ . \tag{2.18}$$

For the left boundary, the local potential part of the Lagrangian, given by Eq. (2.15), has only one term containing u_1, so that the equation of motion for u_1 does not have the same structure as the above. The nonlocal term, Eq. (2.16), has only one term containing u_1, and two terms containing u_2, whereas it contributes three terms for a typical bulk site. The equation of motion for the end particle is therefore

$$m\ddot{u}_1=\frac{a}{\delta^2}(u_1-u_2)-\frac{b}{\delta^3}(u_1-u_2)^2+\frac{c}{\delta^4}(u_1-u_2)^3+\frac{d}{\delta^4}(u_3+u_1-2u_2)+\frac{\gamma}{\delta^2}(\dot{u}_1-\dot{u}_2)\ . \tag{2.19}$$

The equation of motion for the second particle in from the boundary has the bulk form for the terms derived from the local potential and a modified nonlocal contribution,

$$m\ddot{u}_2=\frac{a}{\delta^2}(u_3+u_1-2u_2)-\frac{b}{\delta^3}[(u_3+u_1-2u_2)(u_3-u_1)]$$
$$+\frac{c}{\delta^4}[(u_3+u_1-2u_2)(u_3^2+u_2^2+u_1^2-u_3u_2-u_2u_1-u_3u_1)]-\frac{d}{\delta^4}(u_4-4u_3+5u_2-2u_1)-\frac{\gamma}{\delta^2}(\dot{u}_3+\dot{u}_1-2\dot{u}_2)\ . \tag{2.20}$$

The right end involves analogous equations.

The principal advantage of this formulation is that the boundary conditions are intuitively straightforward. In particular, free boundaries are implemented simply by not constraining the end particles—that's it. This system, being a well-posed Lagrangian dynamical problem, will then show the correct behavior, approaching the continuum behavior as the number of particles is increased for a given (undistorted) physical length. This makes this system an excellent "benchmark" with which one can scrutinize the boundary conditions, which were derived above to append the bulk equation of motion for the continuum. To be specific, if one integrates the equations for the set of u_i, as well as for the continuum field $u(x,t)$, and finds identical descriptions of the displacements, then clearly our continuum approximation *including the boundary conditions* is vindicated.

Figure 2 shows a series of displacement fields of the system, each taken at a different time for the dynamical evolution of the system. This is an integration of the discrete equation of motion. One can use these discrete variable to evaluate the boundary conditions stated in Eq. (2.13)—to do this one must use finite difference approximations for the spatial derivatives (see Ref. 13 for these approximations to arbitrarily high order in δ). We find that the boundary conditions of Eq. (2.13) are indeed satisfied by our integration of the discrete displacement variables, to within the accuracy our temporal integration scheme, thus justifying our use of the continuum approximation to represent a finite system.

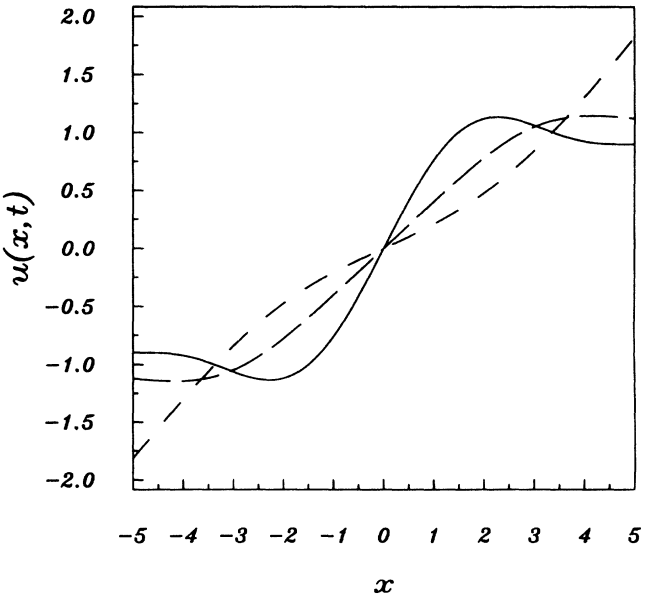

FIG. 2. Nonlinear dynamical evolution of the change-of-length system, derived from the integration of the explicitly discretized equation of motion. The solid curve is the zero-velocity initial displacement field. The long-dashed curve is the displacement field after two time units have elapsed, and the short-dashed curve is the displacement field after four time units. The initial condition was chosen such that the early dynamics exhibited dramatic changes at the boundary, in order to test the validity of the continuum dynamical equation of motion and boundary conditions. This system ultimately evolves to the fully transformed state.

(The primary disadvantage of using the discrete displacement variables, and the reason it is not the only approach discussed here, is that a large number of particles are required to represent any reasonably sized system, and it is such systems that form the variety of complicated patterns observed in, say, transmission electron micrographs of interesting first-order strain transformations.[4] For example, the number of arithmetic operations required to implement this method grows superlinearly in the number of points N, which is already a large number. Also, the numerical problem of cancellation errors in the computation of the high powers of the first differences in the local potential leads to a round-off problem at long times[14]—no such problem arises when integrating $(1+1)$-dimensional partial differential equations, as discussed, e.g., in Ref. 14.)

B. Shear transformations

The first-order shear transformations have a very simple one-dimensional Landau-type potential,[15] which Bales and one of us have studied previously.[5] The local potential is given by

$$V_{\mathrm{LOC}}=\int_{-l}^{l}[\tfrac{1}{2}A(\partial_x u)^2-\tfrac{1}{4}B(\partial_x u)^4+\tfrac{1}{6}C(\partial_x u)^6]dx \ . \tag{2.21}$$

This potential, for appropriately chosen A, B, and C, has a metastable minimum at zero strain, and two degenerate stable minimum at the martensitic strains $e=\pm e_M$. This potential is sketched in Fig. 1. The transition we will follow involves beginning a system in the metastable unstrained state, and watching it transform into either or both of the stable strained states. Using the same strain-gradient energy density (with coefficient G), kinetic energy density (with mass density ρ'), and Rayleigh dissipation function (with sound-wave viscosity Γ) as given above, one finds the equation of motion studied in Ref. 5, viz.,

$$\rho'\partial_t^2 u=(\partial_x^2 u)[A-3B(\partial_x u)^2+5C(\partial_x u)^4] -G(\partial_x^4 u)+\Gamma(\partial_t\partial_x^2 u) \ . \tag{2.22}$$

We append this equation with the boundary conditions, viz.,

$$A(\partial_x u)-B(\partial_x u)^3+C(\partial_x u)^5+\Gamma(\partial_x\partial_t u)-G(\partial_x^3 u)=0 \quad \text{at } x=l,-l \ , \tag{2.23a}$$

$$G(\partial_x^2 u)=0 \quad \text{at } x=l,-l \tag{2.23b}$$

and thus have a complete set of dynamical equations plus boundary conditions to describe such transitions.

III. DYNAMICAL EVOLUTION

A. Surface nucleation in a change of volume transition

At this point we define our method of studying the growth dynamics that takes a system with a single supercritical nucleus to its final state structure—for the change of volume transition this steady-state arrangement is fully transformed product phase.

As mentioned in the Introduction, nonlinear elastic systems do not readily admit the decomposition of the energy into bulk and surface terms conventionally used in classical nucleation theory—this is due to the presence of nonlocal terms in the energy, and also due to the distortion of the parent system arising from the inclusion of a localized region of product phase.[11] It is nevertheless possible to classify perturbations of the metastable parent system as supercritical or subcritical, depending on the behavior of the system under the *dynamics* with the given perturbation as an initial condition. If the perturbation relaxes back into the metastable phase, it is subcritical. If it grows towards the globally stable product phase (or some inhomogeneous structure involving coexisting parent and product phase—see the next section), it is supercritical. In principle, any perturbation may be used as an initial condition, but since we are primarily interested in the nucleation of regions of product state in the system, we restrict ourselves to strongly localized displacement-field perturbations. We have shown elsewhere[7] in a similar system that the distinction between supercritical and subcritical perturbations can depend on purely dynamical features of the system, such as the magnitude of the sound-wave viscosity, thus emphasizing our posture that it is essential to study the dynamics in order to be able to properly characterize nucleation.

In order to propagate a given initial condition forward in time, we first express the initial displacement field in terms of orthogonal functions on a finite interval, the modes of the system, and then numerically integrate the mode amplitudes in time using an appropriate amplitude-based equation of motion. This method is not trivial to implement, and the derivation of this equation of motion is technically involved. However, a substantial advantage results from a mode-decomposition of the displacement field: for a complete basis not having any restrictions on the value of the basis functions at the boundaries, the boundary conditions [e.g., Eq. (2.13)] are automatically satisfied—thus, only the bulk equation of motion need be integrated. A detailed description of this method is given in Appendix B. Within this functional decomposition, the full dynamics of the system are well described by relatively few degrees of freedom—the mode amplitudes—without sacrificing the ability to exactly represent the metastable, homogeneous parent phase and the globally stable, homogeneous transformed phase, which are important "landmarks" in the phase space of the system.

We have used this integration scheme to model the dynamics of the change-of-volume system for a variety of initial conditions and have observed an extremely important nucleation event. This "surface nucleation" process is illustrated in Fig. 3. This system begins from a zero-velocity, "bump" initial condition. The system first converts the gradient potential energy of the bump to kinetic energy, which then propagates through the system until it reaches the boundary. Here a remarkable "cracking the whip" effect is observed, viz., the boundary of the system has its slope (i.e., strain) changed from one sign to another very quickly. This corresponds to the end points

reaching high velocities and subsequently being carried by their inertia into a stable well of the local potential. The system continues pulling more of the system into this well until, finally, the system is homogeneously transformed. Clearly, this is an example of heterogeneous nucleation, and it is the system's only heterogeneity, viz., its surface, which serves as the nucleation center.

This dynamical evolution demonstrates an important concept. We began with an initial state that subsequently transformed for a system *of a given length.* However, if we begin with the same localized initial condition for a *longer system* such nucleating dynamics do not always occur—only if the propagating bump reaches the surface with a sufficient strain and/or velocity does the transition to the elongated state occur. Thus, here we have an example of a supercritical nucleus for a system of finite spatial extent, where the nucleation proceeds at the surface, but homogeneous (bulk) nucleation would not occur for a longer system (of course, any homogeneous nucleation event must be independent of the size of the system or it's not really homogeneous nucleation). Thus, the bulk fluctuation only becomes supercritical due to its proximity to the surface.

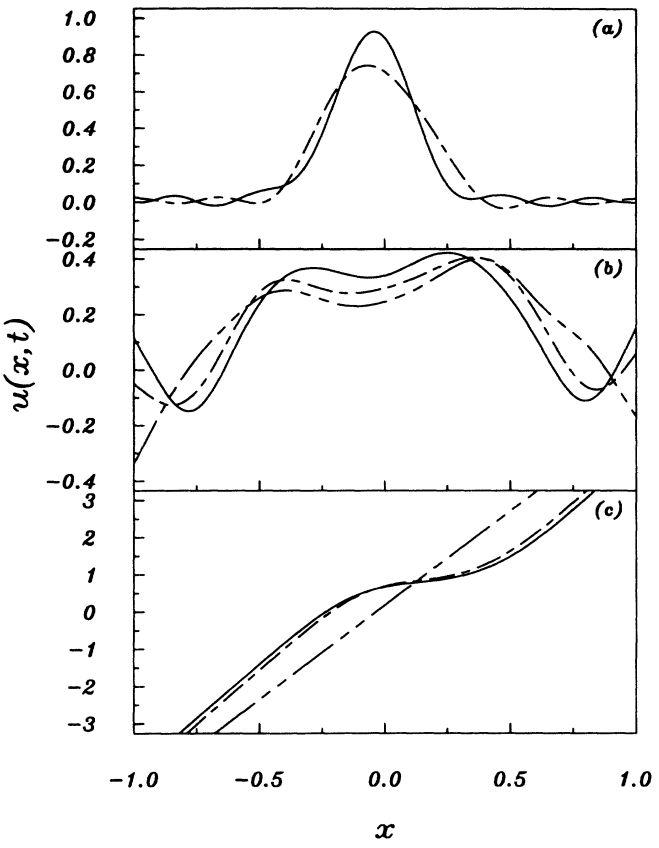

FIG. 3. The phenomenon of edge-nucleation, exhibited by the change-of-length system with a globally stable side-well state. The material constants in the equation of motion, viz., Eq. (2.8), are $\rho=b=c=g=1$, $a=0.075$, and $\gamma=0.16$. Panel (a) depicts the early dynamics of the system. The solid curve in this panel is the zero-velocity initial displacement field, and the dashed curve is the displacement field at $t=2$. The initial response of the system is to flatten out the bump. Panel (b) contains the surface nucleation event at the left-hand edge. The solid curve is the $t=5$ state, the long- and short-dashed curve is the $t=6$ state, and the double-dashed curve is the $t=7$ state. The left-hand edge of the system is propelled into the stable well of the potential by the spreading central bump. Panel (c) shows the long-time limit of the dynamics. The solid curve is at $t=35$, and the long- and short-dashed curve is at $t=36$, indicating that the system is evolving towards the globally stable lengthened state, the latter indicated by the double-dashed curve.

We note that having established the dynamical system of equations describing a system with open boundary conditions and being able to integrate these equations accurately, we have quantitatively studied heterogeneous vs homogeneous nucleation. To be specific, by using the linear response theory of Ref. 11 to define potential nuclei at both the surface and in the bulk, for the first time we have been able to quantitatively estimate the reduced heterogeneous nucleation barrier. This work will be reported elsewhere.

B. Late stage dynamics of martensitic pattern formation

Martensitic phase transitions lead to marvelous patterns of juxtaposed polydomain structures, such as those seen in electron microscopy experiments.[16] The main question motivating most of the work in this section, as well as our earlier publications,[5,17,7,8] is, how does the system *grow* into such complicated arrangements? Previous work by others has focused only on the energies of the domain walls separating different variants of the low-temperature parent (martensitic) phase, as well as interfaces separating the high-temperature parent and the low-temperature phases.[18] Our work attempts to describe how such structures evolve from an arbitrarily chosen supercritical nucleus into the, e.g., twinned bands of martensite.[19]

Work by Bales and one of us[5] was the first step in answering our questions concerning the growth leading to polydomain martensite. Firstly, it was demonstrated that previous work[20] on the motion of a parent-phase product-phase interface based on time-dependent Ginzburg-Landau theory was incorrect because it did not properly incorporate the hydrodynamic character[12] of sound waves (see our discussion of this point in Ref. 7). Then, by beginning the system in a state with a static configuration corresponding to a supercritical nucleus, we followed the deterministic dynamics of the growth of martensite. Figure 1 of Ref. 5 shows the main result of this work—for sufficiently deep quenches (viz., temperatures well below the first-order phase transition), a dynamical twinning occurs whereby alternating domains of martensite of strain $e=\pm e_M$ are produced *dynamically.* Figure 1(a) of Ref. 5 displays the evolution of the system that arises when the inertia of the displacement field is ignored, viz., a single domain of martensite is formed (this is the growth that would be predicted by Chan's theory[20]). That this path can never be accessed in nature may be understood by noting that this growth corresponds to a state in which the crystal has a kinetic energy that scales with the size of the system and is thus infinite

for any bulk sample.

We now present results for this growth at very long times for a finite sample—by including the boundary conditions given in Eq. (2.25), we can allow the growth front studied in Ref. 5 to reach the boundary, and thus allow the system to approach its steady-state configuration. In fact, despite the large body of work that has studied the stabilities of various transformed structures, the existence of certain twinned states is still in question—for example, some work[21] has predicted that very special surface stresses must be present for a twinned crystal to be stable. Then, the very natural question that arises is, does the structure that is created in the dynamical twinning process found in Ref. 5 survive when the transformation interface reaches an open boundary and the system is allowed to proceed to a steady state? To address this question, and thus determine the late-stage dynamics of dynamical twinning, we have integrated Eq. (2.22) subject to the boundary conditions given in Eq. (2.23). In this case we have integrated the equation of motion using a variant of the method of lines that has recently been proposed[14]—we note that identical results are obtained if the mode-expansion technique is employed.

We have chosen the initial state to be a static localized strain in a system bounded by $-l \le x \le l$. An antisymmetric initial state, viz., one where the localized strain is antisymmetric with respect to inversion about $x=0$, is the initial condition most reminiscent of a small local fluctuation of the displacement field, and thus here we present the results obtained from integrating such a state. However, we note that for a symmetric initial condition, similar dynamics are found. For the symmetric initial state, the boundaries move once the transformation front has propagated to the boundary, and our results are shown in Fig. 4. The $t \to \infty$ graph is the *static* configuration in which the system finishes. Thus, the dynamical twinning *concludes* with a twinned product configuration. This is clearly not the absolute minimum of the elastic potential but rather is a metastable state, which, in the absence of thermal fluctuations, has an infinite lifetime. The greater the length of the system studied, the more regular the spacing between the martensitic variants—thus, the irregularity of the spacings seen in this figure is a finite-size effect.

As mentioned above, for all initial conditions with a localized fluctuation of the strain, when the system is quenched to be at a sufficiently low temperature, we find a final steady-state configuration with a twinned morphology—obviously, such structures possess a robust stability, in contradiction to Ref. 21. It is unclear of the reason for the differences between the analytic conclusions of Cao *et al.*, and our own numerical results, although admittedly we have only presented results for a one-dimensional system, and higher dimensionalities may change our conclusions—our dynamics for two-dimensional shear transformations will be discussed in a future publication, and there a complete comparison to the work of Ref. 21 will be more suitable.

Lastly, in Ref. 5 a dynamical excitation named a two-kink solution was found (see Fig. 2 of Ref. 5). Such dynamics result if the temperature quench through the first-order transition is shallow. We have determined the late stage dynamics for such systems for both symmetric and antisymmetric initial conditions. For symmetric initial conditions the final state is a single domain of martensite—*no* domain walls are present. For an antisymmetric strain configuration, the dynamics are shown in Fig. 5. Here the final state is a single-domain wall separating the two martensitic strain states. Note that this final state is the same as that found in Ref. 7, where fixed-boundary conditions were used, and thus we have clear evidence that the inhomogeneous product configurations that result from our dynamics can be independent of boundary conditions.

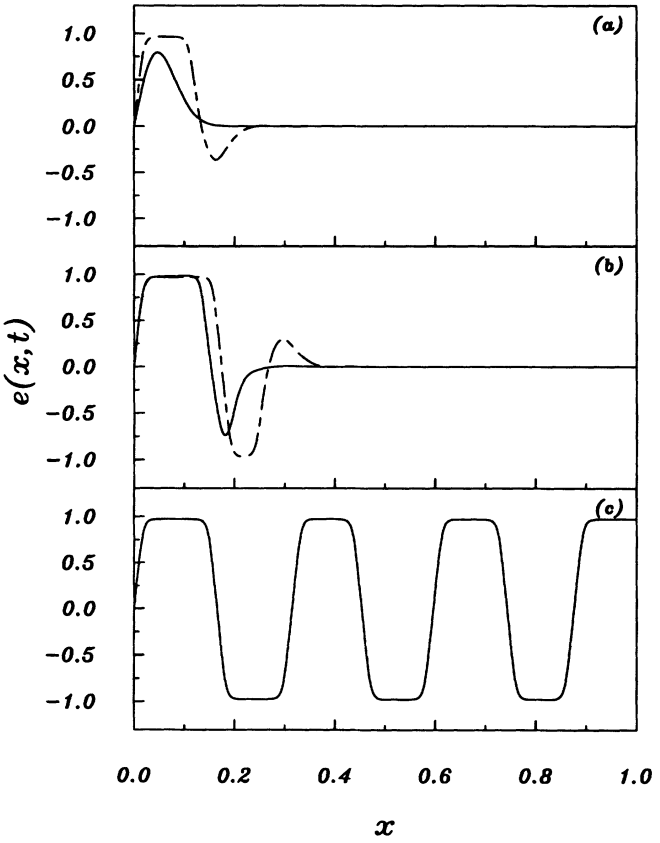

FIG. 4. The evolution of a shear martensite, described by Eqs. (2.22) and (2.23), undergoing the dynamical twinning sequence discussed in Ref. 5, eventually reaching its steady-state configuration. We only show the $x \ge 0$ section of the system. In Eq. (2.22) we have used $\rho'=1$, $A=0.05$, $B=C=G=\Gamma=1$. For these parameters the strains of the degenerate product states are $e=\pm e_M \sim \pm 1$. Panel (a) shows the initial antisymmetric localized strain state as a solid curve, and the $t=4$ curve as the dashed curve. This demonstrates the growth front beginning to twin as the portion of the system just ahead of the growth front is pushed into the strain state opposite to that of the growing domain. In panel (b), $t=6$ is the solid curve and $t=10$ is the dashed curve. Panel (c) represents the final, static, twinned state of the system, and is found at any time of the dynamical evolution for $t \gtrsim 130$.

IV. SUMMARY

The one-dimensional systems studied in this paper exhibit the important qualitative features of mesoscopic, finite, continuum elastic systems. We have modeled the full nonlinear dynamics for this system in the presence of free boundary conditions, corresponding to the situation of an isolated single-grain crystal, and have seen how the intrinsic inhomogeneity associated with the boundary can give rise to qualitatively new effects. We have employed this formalism and have observed the phenomenon of "surface nucleation" in a system with only one heterogeneity, viz., its surface. We have extended this work to yield a complete characterization of the reduced heterogeneous nucleation barrier for a free surface, and these results will be presented elsewhere. Further, we have been able to demonstrate the stability of interesting inhomogeneous structures for shear martensites in the absence of surface stresses. This work is being repeated for two- and three-dimensional systems, where the problem of so-called accommodation leads to even more complicated patterns being formed—this work will also be published elsewhere.

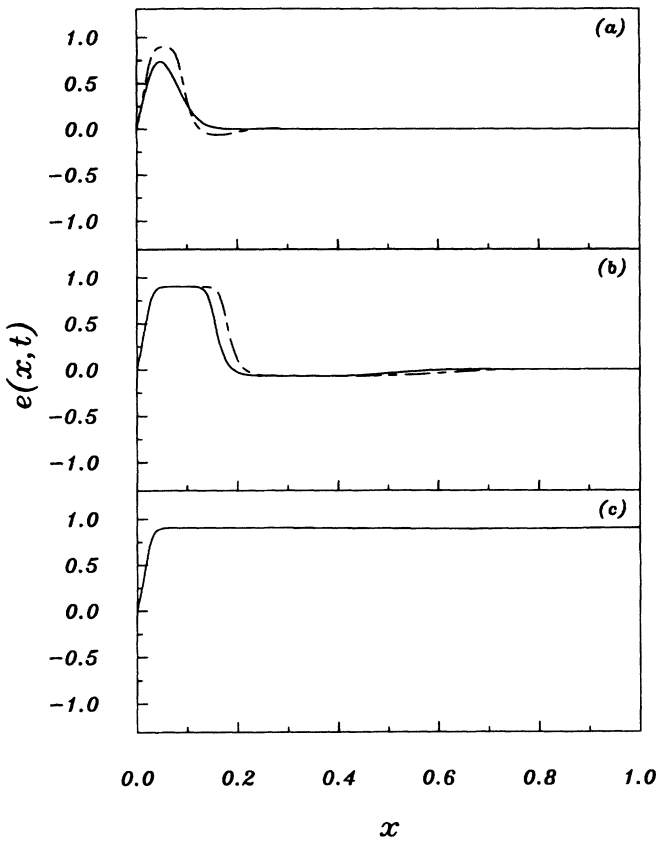

FIG. 5. The evolution of the double-kink system. The parameters for this system are the same as those of Fig. 4, except that now $A=0.15$, and again the martensitic strain is $e_M \sim \pm 1$. Panel (a) shows the early time dynamics—the solid curve in this panel in the initial, antisymmetric, zero-velocity strain profile of the system, identical to that of Fig. 4. The figure shows only the right-hand half of the system. The dashed curve of panel (a) is the $t=4$ state of the system, showing the formation of the two boundaries, one between the positively strained region of the system and the negatively strained region, and a second between the negatively strained region and the unstrained region. Panel (b) shows the propagation of these two boundaries. The solid curve is at $t=16$, the dashed curve at $t=20$. The boundary between the unstrained region and the negatively strained region propagates rapidly towards the edge of the system, while the transformation front closer to $x=0$ propagates more slowly. Panel (c) shows the final state for this system, in which only the martensitic twins and one twin boundary are present.

The methods described here give numerical access to the dynamics of finite continuum systems, and these methods are reviewed in Appendix B. The finite-difference method of Sec. II provides an intuitively straightforward but time-consuming solution to the dynamical problem, and is restricted to systems of severely limited spatial extent. However, results from such work serve as a benchmark through which other numerical methods can be tested. (In fact, here we used results obtained from this technique to justify the nonlinear, time-dependent boundary conditions that we have found in Sec. III.) Our proposed solution to the integration of such continuum systems involves expressing the continuum fields as a modelike expansion in terms of a complete set of basis functions, where there are no restrictions on the functions, or any of the function's derivatives, at the boundaries. Then, one must express the Lagrangian in terms of these basis functions, and then integrate the Euler-Lagrange equations for the mode amplitudes. This provides a complete description of the dynamics of the system, *including* the boundary conditions. The utility of this approach to higher-dimensional systems is straightforward.

ACKNOWLEDGMENTS

We wish to thank Steve Shapiro for helpful discussions on the change of volume transition, and Nick Schryvers and Lee Tanner for valuable information on TEM experiments studying the pattern formation problem in martensite. This work was supported by the NSERC of Canada, and the Advisory Research Committee of Queen's University.

APPENDIX A: GENERALIZED BOUNDARY CONDITIONS

Our primary motivation for modeling elastic systems at the mesoscopic level is to understand the nature of structural phase transitions. The choice of free-boundary conditions for the principal analysis in this paper is a somewhat idealized choice, assuming a completely coherent initial state consisting of a single, unconstrained, initial domain. In practice, experimental samples consist of numerous crystalline domains in mechanical contact with each other. A first step towards modeling this situation can be taken by generalizing the boundary conditions.

Consider the one-dimensional, nonlinear, dissipative "bar" system in contact at either end with linear, nondissipative bars, which model adjacent domains. The effect of these domains will be to act as springs, introducing

two new surface terms, which will constrain the motion of the end points of the primary system. The springs will have some spring constant k, which characterizes the new, more general boundary condition. The limit $k \to 0$ is the free-boundary case already studied, and the opposite limit $k \to \infty$ is that of fixed boundary conditions.

Including these new surface terms, assuming differing spring constants k_1 on the left and k_2 on the right, the more general Lagrangian is

$$L = \int_{-l}^{l} [\tfrac{1}{2}\rho(\partial_t u)^2 - \tfrac{1}{2}a(\partial_x u)^2 + \tfrac{1}{3}b(\partial_x u)^3 - \tfrac{1}{4}c(\partial_x u)^4 - \tfrac{1}{2}d(\partial_x^2 u)^2]dx - \tfrac{1}{2}[k_1 u^2(-l) + k_2 u^2(l)] \ . \quad \text{(A1)}$$

The dissipation function is not altered, and is still given by

$$R = \tfrac{1}{2}\gamma \int_{-l}^{l} (\partial_t \partial_x u)^2 dx \ . \quad \text{(A2)}$$

In this appendix, we shall derive the boundary conditions by retaining the surface terms in the Lagrangian analysis, rather than using the energy arguments of a previous section. The first variation of the Lagrangian is given by the terms linear in h under the substitution $u \to u + h$, where $h(x)$ is some arbitrary function of x.

$$\delta L = \int_{-l}^{l} [\rho(\partial_t u)(\partial_t h) - a(\partial_x u)(\partial_x h) + b(\partial_x u)^2(\partial_x h) - c(\partial_x u)^3(\partial_x h) - d(\partial_x^2 u)(\partial_x^2 h)]dx - k_1 u(-l)h(-l) - k_2 u(l)h(l) \ . \quad \text{(A3)}$$

For the dissipation function, we make the substitution $\partial_t u \to \partial_t u + h$,

$$\delta R = \gamma \int_{-l}^{l} (\partial_t \partial_x u)(\partial_x h)dx \ , \quad \text{(A4)}$$

using the same arbitrary function h as in the case of Eq. (A3). The Lagrangian dynamical method requires that the variational derivative of the Lagrangian, $\delta L / \delta u$, be equal to the variational derivative of the dissipation function, $\delta R / \delta(\partial_t u)$. It follows from this that the equation of motion can equally well be obtained by equating the first variations δL and δR as above. Integrating by parts to eliminate derivatives of the function h, one obtains a bulk term, the equation of motion, and two surface terms, one proportional to $\partial_x h$ and one to h. The first surface term is identical to the boundary condition Eq. (2.13b). The second surface term is a generalization of Eq. (2.13a) and is given by

$$a(\partial_x u) - b(\partial_x u)^2 + c(\partial_x u)^3 - d(\partial_x^3 u) + \gamma(\partial_t \partial_x u) - k_2 u = 0 \quad \text{at } x = l \ , \quad \text{(A5a)}$$

$$a(\partial_x u) - b(\partial_x u)^2 + c(\partial_x u)^3 - d(\partial_x^3 u) + \gamma(\partial_t \partial_x u) + k_1 u = 0 \quad \text{at } x = -l \ . \quad \text{(A5b)}$$

The gradient boundary condition is as in the previous case. These generalized boundary conditions clearly reduce to those of Sec. II in the case $k_1 = k_2 = 0$. The case $k_1, k_2 \to \infty$ is most clearly understood by dividing Eq. (A5a) through by k_2 and Eq. (A5b) through by k_1 and then taking the large k_1, k_2 limit. In both cases, the equations reduce to $u = 0$, corresponding to fixed boundary conditions, as expected.

APPENDIX B: INTEGRATION OF THE CONTINUUM EQUATION OF MOTION

1. Naive-mode expansion

Here we shall discuss a method for solving the continuum equation of motion that satisfies the boundary conditions. The method involves projecting the equation of motion, Eq. (2.8), onto a set of basis functions, and then finding the equation of motion for the basis function amplitudes. This method has the advantage of a greatly reduced number of degrees of freedom, as well as possibly generalizing to higher-dimensional systems. A similar approach was successfully used by us for the problem of fixed boundary conditions.[7] [The primary difficulty encountered in any attempt to construct this type of implementation is that the first boundary condition in Eq. (2.13a) is a function of both $\partial_x u$ and $\partial_t \partial_x u$, and is nonlinear.]

The displacement field $u(x,t)$ can be expanded into component functions

$$u(x,t) = \sum_i n_i(t) f_i(x) \ , \quad \text{(B1)}$$

where the functions $f_i(x)$ form a complete basis on the appropriate interval. These functions need not be orthogonal on the interval. Within the framework of these functions, the physical state of the system is completely specified by the amplitudes n_i, and the velocities $\dot{n}_i$, corresponding to each of the basis functions.

This expansion substituted into the continuum equation of motion, Eq. (2.8), then gives the mode expansion,

$$\rho \ddot{n}_i f_i = n_i f_i''(a - 2bn_j f_j' + 3cn_j f_j' n_k f_k') - gn_j f_j'''' + \gamma \dot{n}_j f_j'' \ , \quad \text{(B2)}$$

where $f_i' \equiv df_i(x)/dx$, etc. We implicitly sum over repeated indices. Taking the inner product of these expression with a particular basis function $f_n(x)$ on the interval of interest yields

$$\rho \ddot{n}_i \int f_i f_n dx = an_i \int f_i'' f_n dx - bn_i n_j \int f_i'' f_j' f_n dx + cn_i n_j n_k \int f_i'' f_j' f_k' f_n dx + gn_i \int f_i'''' f_n dx + \gamma \dot{n}_i \int f_i'' f_n dx \ . \quad \text{(B3)}$$

Each term in Eq. (B3) contains an integration, over the relevant interval, of a product of derivatives of the basis functions. These integrals are independent of the amplitudes n_i, but give rise to tensors (see, e.g., Ref. 7), which couple the amplitudes to each other in the equation of motion. These tensors depend only on the choice of basis functions $\{f\}$ and the form of the potential energy.

This mode-based equation of motion, Eq. (B3), can

then be integrated in time to find the dynamics of the system. The right-hand side depends only on the system state. The left-hand side is straightforward if the functions $\{f\}$ are chosen to be orthogonal. In this case, left-hand-side integration gives a quantity proportional to $\delta_{i,n}$. In general, however, the full left-hand-side term will be a linear combination of the accelerations of the various mode amplitudes. In this case, a matrix equation involving the symmetric matrix

$$M_{\text{in}}=\int f_i f_n dx \tag{B4}$$

must be solved before the acceleration values can be assigned. (This requirement may seem to strongly favor the choice of a set of functions, which are orthogonal on the interval of integration, but, in practice, the matrix M_{in} need be inverted only once and stored, and the left-hand-side evaluation process represents a small cost compared to the effort required to evaluate the right-hand side. Thus, nonorthogonal basis sets do not cause any difficulty.)

The second ingredient in this method are the boundary conditions, which can be expressed in terms of the mode amplitudes by again substituting the mode expansion, Eq. (B1), into the appropriate continuum equation from Eq. (2.13). This gives, for example at the right-hand boundary $x=l$,

$$an_i f_i'(l)-bn_i n_j f_i'(l)f_j'(l)+cn_i n_j n_k f_i'(l)f_j'(l)f_k'(l) +\gamma \dot{n}_i f_i'(l)-dn_i f_i'''(l)=0 \tag{B5}$$

and

$$dn_i f_i''(l)=0 \tag{B6}$$

with a similar constraint at the left-hand boundary.

The first boundary condition presents the most complex problem, since it is, in principle, a relation between products of sums of coefficients. Probably the most effective method of coping with this type of boundary condition is by means of a judicious choice of basis functions. We choose a primary set of functions to be orthonormal among themselves and to have no first or third derivative at the boundaries. These functions will then automatically satisfy the first boundary condition, and can be linearly combined so as to satisfy the second boundary condition by means of a projection operator. However, such a set of functions is not by itself an appropriate choice to represent the dynamics of the system—it is known that the final, fully transformed state of the system is the one for which $u(x)$ is linear, with a slope corresponding to the bottom of the stable well in the local potential. This state has a slope at the boundary, so that it is reasonable to suppose that the dynamical system will, in the coarse of the motion, acquire such a slope. The dynamics must allow for this.

Our method is to augment the basis function set with additional functions, not constrained to be orthogonal to the original basis set, which have slopes at the boundary. To account for the velocity degree of freedom, we use two such functions, chosen to be orthogonal to each other. We represent the displacement field as a function of time using the functions

$$u(x,t)=\alpha(t)g(x)+\beta(t)h(x)+n_i f_i(x)\ , \tag{B7}$$

where

$$f_i(x)=\sqrt{2/L}\sin\left[\frac{(2i+1)\pi x}{L}\right]\ , \tag{B8}$$

$$g(x)=x\ , \tag{B9}$$

and

$$h(x)=\frac{-L^3}{24\pi^3}\left[\sin\left[\frac{2\pi x}{L}\right]+\tfrac{1}{2}\sin\left[\frac{4\pi x}{L}\right]\right]\ . \tag{B10}$$

These functions are not completely general but describe only odd-parity configurations. This is physically correct if we restrict ourselves to odd-parity initial conditions, since the dynamics preserves parity. The function g has been chosen such that it has a unit first derivative at the boundary and zero second and third derivatives there, and the function h has been chosen to have a unit third derivative and zero first and second derivatives at the boundary. The principal basis functions Eq. (B8), already discussed, have been contrived to satisfy both boundary conditions. The functions g and h are not orthogonal to any of the f's, and trivially satisfy the second boundary condition. This augmentation of the basis gives two additional degrees of freedom, which allows the two boundary conditions to be incorporated into the equation of motion. We take the linear case as an example.

If we assume that the initial conditions satisfy the boundary conditions, it is sufficient to require that the second time derivative of the boundary condition remain zero throughout the dynamics. The boundary condition thus cannot acquire any velocity, and will be satisfied at all times.

Using the orthonormality of the main basis functions, and the property that trigonometric functions are proportional to their even-numbered derivatives, specifically, $f_i''=-k_i^2 f_i$, with $k_i=(2i+1)\pi/L$, we have that the mode-based equation of motion becomes

$$\begin{aligned}\ddot{\alpha}g_i^{(0)}+\ddot{\beta}h_i^{(0)}+\ddot{n}_i=&a[\alpha g_i^{(2)}+\beta h_i^{(2)}-k_i^2 n_i]\\&-d[\alpha g_i^{(4)}+\beta h_i^{(4)}-k_i^4 n_i]\\&+\gamma[\dot{\alpha}g_i^{(2)}+\dot{\beta}h_i^{(2)}-k_i^2\dot{n}_i]\ ,\end{aligned} \tag{B11}$$

where

$$g_i^{(n)}=\int_{-l}^{l}\left[\frac{d^n}{dx^n}g(x)\right]f_i(x)dx$$

and

$$h_i^{(n)}=\int_{-l}^{l}\left[\frac{d^n}{dx^n}h(x)\right]f_i(x)dx\ . \tag{B12}$$

The appropriate time derivatives of the boundary conditions can be expressed in terms of modes as well. The first time derivative of the first boundary condition gives a relation involving second time derivatives of the parameters α and β,

$$\ddot{\alpha}g'(l)+\ddot{\beta}h'(l)=-\frac{a}{\gamma}[\dot{\alpha}g'(l)+\dot{\beta}h'(l)] + \frac{d}{\gamma}[\dot{\alpha}g'''(l)-\dot{\beta}h'''(l)] , \tag{B13}$$

and the second time derivative of the second boundary condition gives

$$\sum_{i=0} \ddot{n}_i f_i''(l)=0 . \tag{B14}$$

The full dynamics for the system can be conveniently specified in matrix form. Labeling the right-hand side of Eq. (B11) by r_i and the right-hand side of the boundary condition Eq. (B13) by d_1, we may write

$$\begin{bmatrix} 0 & 0 & f_0''(l) & f_1''(l) & \cdots & f_n''(l) \\ g'(l) & h'(l) & 0 & 0 & \cdots & 0 \\ g_0 & h_0 & 1 & 0 & \cdots & 0 \\ g_1 & h_1 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ g_n & h_n & 0 & 0 & \cdots & 1 \end{bmatrix} \begin{bmatrix} \ddot{\alpha} \\ \ddot{\beta} \\ \ddot{n}_0 \\ \ddot{n}_1 \\ \vdots \\ \ddot{n}_n \end{bmatrix} = \begin{bmatrix} 0 \\ d_1 \\ r_0 \\ r_1 \\ \vdots \\ r_n \end{bmatrix} . \tag{B15}$$

This is the full equation of motion, including boundary conditions. The system can be reduced to a 2×2 matrix by means of row reduction. The appropriate multiple of each the third and subsequent rows can be subtracted from the first so that the top two rows of the matrix multiply only $\ddot{\alpha}$ and $\ddot{\beta}$. The resulting system is

$$\begin{bmatrix} \sum_i g_i f_i''(l) & \sum_i h_i f_i''(l) \\ g'(l) & h'(l) \end{bmatrix} \begin{bmatrix} \ddot{\alpha} \\ \ddot{\beta} \end{bmatrix} = \begin{bmatrix} \sum_i r_i f_i''(l) \\ d_1 \end{bmatrix} . \tag{B16}$$

This system is readily solved for $\ddot{\alpha}$ and $\ddot{\beta}$, and the remaining coefficients are then determined by back substitution,

$$\ddot{n}_i=r_i-\ddot{\alpha}g_i-\ddot{\beta}h_i . \tag{B17}$$

The dynamics of this system can then be computed by using the given accelerations to compute new velocities, and the velocities to compute new positions, iterating forward from one time step to another. The boundary conditions will be satisfied exactly at all times.

This type of approach can be generalized to include the nonlinear part of the equation of motion. Formally, this is a simple matter of altering the definition of the right-hand side quantities r_i. In practice, however, the auxiliary function method is not stable in the nonlinear regime. While the nature of this instability is not entirely understood, it almost certainly involves the appearance of a divergent (with increasing numbers of modes) term in the first row of the 2×2 equation of motion, Eq. (B16).

The mode-expanded equation of motion is subject to a particular form of pathology involving the auxiliary functions. The sum

$$\sum_i g_i f_i''(l) , \tag{B18}$$

in fact, increases linearly with the number of modes included in the system. This effect is present in all implementations of this method, including the linear method, which we have already seen, but only leads to numerical instability in the nonlinear case. In the linear case, the force field of the system at any given time must morphologically resemble the displacement field and, in particular, will not contain any new Fourier coefficients. This is conjectured to give rise to equivalent divergences in the r_i and g_i sums in Eq. (B16).

The divergence of the sum, Eq. (B18), arises whenever the corresponding function $g(x)$ has a slope at the boundary. To see this, we rewrite the sum given by Eq. (B18) as

$$\sum_i \left[\int_{-l}^{l} g(x)f_i(x)dx\right] f_i''(l) , \tag{B19}$$

and, since $f_i(x)$ is the sine function of Eq. (B8), we replace it by

$$f_i(x)=\frac{-f_i''(x)}{k_i^2} \tag{B20}$$

giving

$$\sum_i \int_{-l}^{l} [g(x)f_i''(x)dx]\frac{-f_i''(l)}{k_i^2} . \tag{B21}$$

The integral can be rearranged by parts twice, using

$$g(x)f_i''(x)=\frac{d}{dx}[g(x)f_i'(x)]-\frac{d}{dx}[g'(x)f_i(x)] + g''(x)f_i(x) , \tag{B22}$$

and, absorbing the k_i into the quantity outside the integral, we have finally that

$$\sum_i \int_{-l}^{l} [g(x)f_i(x)dx]f_i'' = \sum_i \left[g(x)f_i'(x)\Big|_{-l}^{l} - g'(x)f_i(x)\Big|_{-l}^{l} + \int_{-l}^{l} g''(x)f_i(x)dx\right] f_i(l) . \tag{B23}$$

For the particular basis functions chosen here, the derivative $f'(x)$ is zero at the boundaries, and thus the first surface term in Eq. (B23) is zero. The second surface term, however, is nonzero. Because the functions $f_i(x)$ are odd parity, and the derivative $g'(x)$ is even parity, the second surface term gives the same sign contribution for every index i. For the particular basis functions we have chosen, which always have the same magnitude for the displacement at the boundary, this quantity is proportional to the number of terms in the sum, which means that it is simply proportional to the number of modes included.

This instability can be quite pronounced. A robust mode-based numerical integration method must satisfy

the requirement that increasing the number of modes in the system increases the fidelity of the dynamics. Figure 6 shows a failure of the auxiliary-function mode-based dynamics to satisfy this requirement. The curves shown are difference curves between a real-space integration and two mode-based auxiliary-function integrations. This figure should be viewed in light of Fig. 7, which is the corresponding figure for the mode-based-Lagrangian dynamical method. It is clear that the residuals for the auxiliary-function method are quite large and are not reduced but rather *increased* by the addition of more modes to the dynamics. This renders this method unsuitable for a description of the transformation problem.

A variant on this method, involving a full basis set with no auxiliary functions, was considered as a solution to the divergence problem, but expanding the continuum equation of motion, Eq. (2.8), in any basis set for which the functions have slopes at the boundaries gives rise to the difficulty that *all* of the basis functions must simultaneously satisfy the boundary conditions, Eq. (2.13), one of which is nonlinear. Thus, this type of approach requires that a nonlinear projection problem—that of projecting the basis-function coefficients onto the subspace of functions that satisfy the boundary conditions—must be performed at every time step. This is numerically unstable and conceptually hideous. It has the advantage, however, that it is one small step away from the next, more successful method.

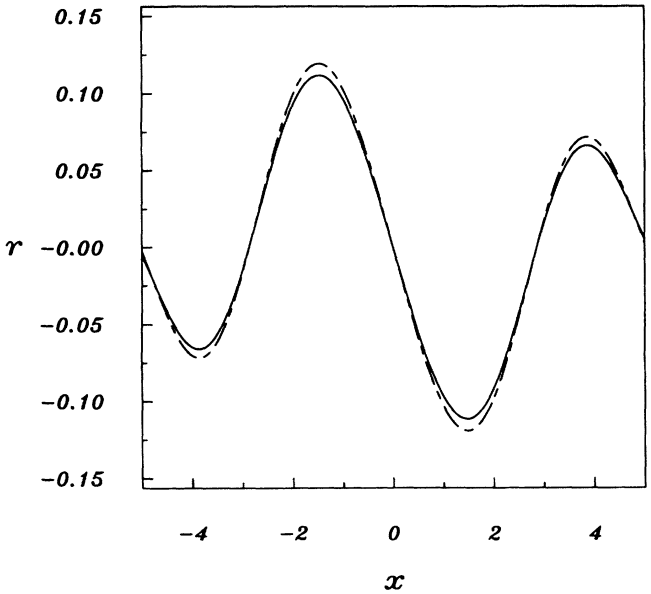

FIG. 6. Difference curves between a spatially discrete integration of a nonlinear system over a fixed-time interval, using an initial condition similar to that of Fig. 2, and the naive mode-based dynamics. The solid curve is the difference between the discrete dynamics and an eight-mode integration, and the dashed curve is the difference between the discrete dynamics and a twelve-mode integration. This figure shows that the addition of a greater number of modes to the naive mode-based dynamics *reduces* the accuracy of the approximation, thus indicating that this scheme is not robust.

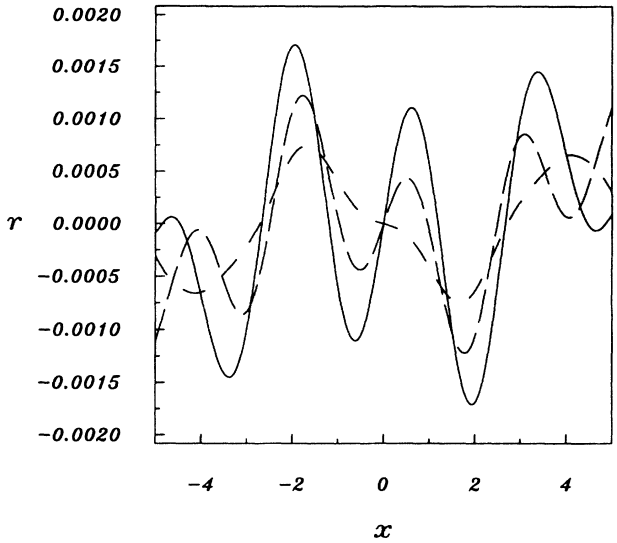

FIG. 7. Difference curves between the spatially discrete integration of a nonlinear system over a fixed-time interval, using an initial condition similar to that of Fig. 2, and mode-based Lagrangian dynamics, for various numbers of modes. The solid curve shows the residuals for a 12-mode integration, the long-dashed curve for a 14-mode integration, and the short-dashed curve for a 16-mode integration. For the mode-based Lagrangian dynamical method, the addition of a greater number of modes to the dynamics *increases* the accuracy of the approximation, indicating that the Lagrangian method is robust.

2. Mode-based Lagrangian

The final method used to describe the dynamical evolution of the bent-bar system is an outgrowth of the preceding method. It combines the advantage of the mode-based description—a small number of degrees of freedom—with the conceptual simplicity of a straightforward Lagrangian dynamical approach.

We begin by expanding the function u in some basis set, as before, but this time substitute the resulting expansion directly into the Lagrangian rather than into the equation of motion. The Lagrangian then has a finite number of degrees of freedom, and the conventional analysis can proceed. The full Lagrangian is

$$L=\int_{-l}^{l}[\tfrac{1}{2}\rho\dot{n}_i\dot{n}_j f_i f_j-\tfrac{1}{2}a n_i n_j f_i' f_j'+\tfrac{1}{3}b n_i n_j n_k f_i' f_j' f_k' -\tfrac{1}{4}c n_i n_j n_k n_l f_i' f_j' f_k' f_l'-\tfrac{1}{2}g n_i n_j f_i'' f_j'']dx\ , \tag{B24}$$

and the corresponding dissipation function is given by

$$R=\int_{-l}^{l}[\tfrac{1}{2}\gamma\dot{n}_i\dot{n}_j f_i' f_j']dx\ . \tag{B25}$$

In each of these quantities, the spatial integration can be carried out, and there are no resulting surface terms. The equation of motion is given by the standard Lagrangian method for systems with a finite number of degrees of freedom:

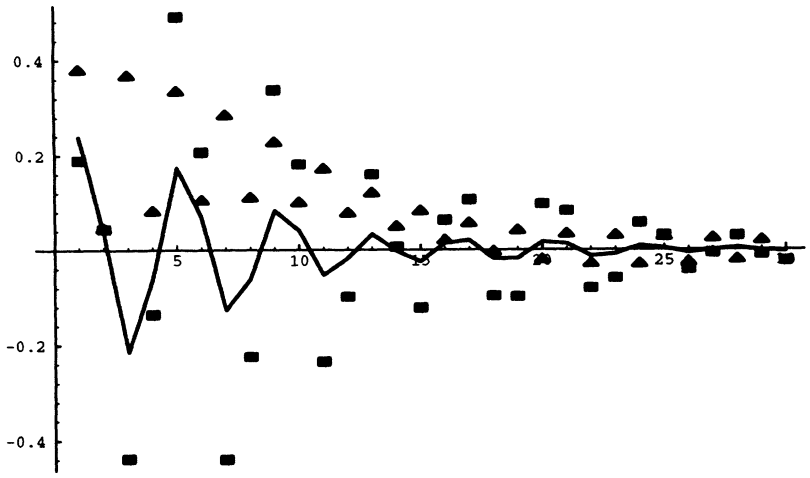

FIG. 8. Mode coefficient amplitudes as a function of mode index, fitted to a bump initial condition similar to that of Fig. 3, plotted as a function of mode index. The solid triangles are the amplitude values for trigonometric functions, and the squares are the coefficient values for a fit to Legendre polynomials. The solid line connects coefficient values for a fit to Chebyshev type-II polynomials. The Chebyshev polynomial coefficients approach zero most quickly with increasing index, indicating that these polynomials are the best choice with which to implement any mode-based dynamics.

$$\rho\ddot{n}_j\int_{-l}^{l} f_i f_j dx = -an_j\int_{-l}^{l} f_i' f_j' dx + bn_j n_k\int_{-l}^{l} f_i' f_j' f_k' dx - cn_j n_k n_l\int_{-l}^{l} f_i' f_j' f_k' f_l' dx - dn_j\int_{-l}^{l} f_i'' f_j'' dx - \gamma\dot{n}_j\int_{-l}^{l} f_i' f_j' dx \ . \tag{B26}$$

This equation of motion can be seen to have a similar structure to the equation of motion obtained by direct substitution of the mode decomposition into the continuum equation of motion, given by Eq. (B3). These two expressions of the equation of motion are, in fact, related through the boundary conditions. If the integrals on the right-hand side of the equation of motion derived from the mode-based Lagrangian, Eq. (B26), are integrated by parts, differentiating one of the factors in each integral and differentiating the rest of the integrand, the resulting bulk terms give Eq. (B3), and the surface terms that arise are $f_i(l)$ times the first, nonlinear boundary condition, Eq. (B5), and $f_i'(l)$ times the second boundary condition, Eq. (B6). This new equation of motion, Eq. (B26), is therefore equivalent to the mode-substituted continuum equation of motion, Eq. (B3), *including* the boundary conditions. It is therefore, by itself, a complete description of the dynamics of the system, except insofar as only a finite number of modes are retained. The boundary conditions will be automatically satisfied for any given basis set $\{f\}$, although, in practice, explicit enforcement of the second-derivative boundary condition has been found to be useful.

The integrals in the equation of motion Eq. (B26) have fixed values, which do not depend on the dynamics, and so can be computed in advance. They are the "mode-coupling tensors" of the nonlinear dynamics (see, e.g., Ref. 7), and provide the numerical mechanism that couples different basis functions to each other. The nature of the finite-mode approximation is most apparent here.

Since every mode can, in principle, be coupled to every other mode, a system whose initial condition can be exactly described by a finite number of modes need not have a dynamical evolution that retains this feature. For this reason, it is of some importance to select basis function sets for which a few functions can plausibly be thought to account for most or all of the dynamical evolution. The trigonometric functions used for illustration in the mode-based equation of motion, Eq. (B11), have a direct physical interpretation in terms of the phonon modes of the linear system. With other basis sets this identification is somewhat more problematic, but it is nevertheless clear that basis functions with large curvatures or many closely spaced zeros are likely to be high-energy states of the system and therefore not relevant to the dynamics. Further, and in some sense more importantly, it is *unlikely* that

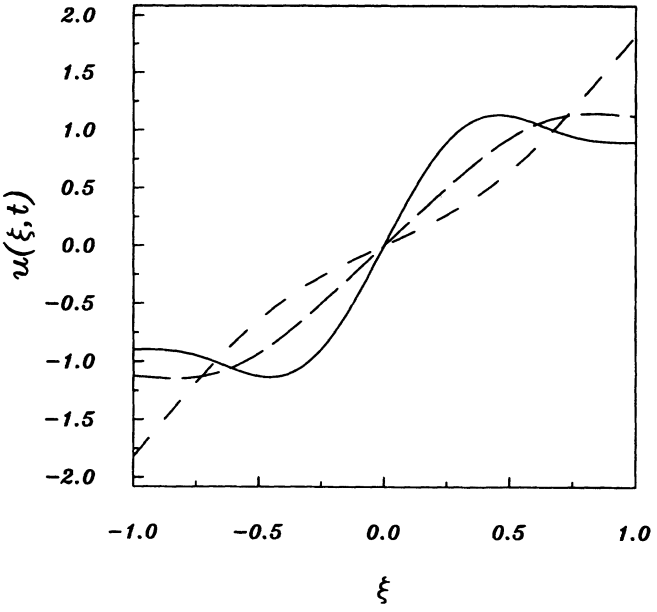

FIG. 9. Dynamical evolution of the system shown in Fig. 2, but obtained using the mode-amplitude-based integration scheme. This system correctly reproduces the dynamics of the corresponding spatially discrete dynamical evolution for a system, thus indicating that the choice of boundary conditions for this method is correct. This plot is of the orthogonal polynomials used in the dynamics, and thus is scaled in x relative to Fig. 2. However, the identities of the curves are the same as that figure.

products of these high-index states and lower states will yield low-energy states. The curvature of the central potential well, and the coefficient of the gradient term, provide a physical length-scale for the system. As long as sufficiently many modes have been retained so that structures on this length scale or longer are well represented, the mode approximation will be a good one.

Three possible choices of basis were investigated for implementation using this method, trigonometric functions, Legendre polynomials and Chebyshev type-II polynomials. To select a basis set, a bump initial condition was fit using each of the basis sets, and the resulting coefficient sets were examined for rapidity of convergence. This process is illustrated in Fig. 8, where the coefficients are plotted as a function of index. It can be seen in that figure that the coefficients of the Chebyshev type-II fit converge most rapidly, indicating that loss of some accuracy in the highest coefficients would not significantly alter the qualitative features of the fitted polynomial. The Chebyshev type-II basis set is therefore the most robust of the three fits, and these polynomials were used on the subsequent dynamics.

The need to choose between polynomial basis sets is not obvious, since the set of the first n Legendre polynomials and the set of the first n Chebyshev polynomials span the same set of possible functions—the best fit to a given initial condition using a polynomial of order n is, after all, unique. However, smallness of the higher-order coefficients gives a more robust fit and contributes to the numerical stability of the integration process.

Figure 9 illustrates the dynamics of the system as derived from the integration of the equation of motion obtained from the mode-based Lagrangian, using Legendre polynomials. These dynamics are very close to the linear discrete-system dynamics of Fig. 2. Furthermore, this system also exhibits agreement between the dissipation function and the actual numerically measured rate of energy loss. The boundary conditions are satisfied by these dynamics as well.

One final and important point: if this mode-based expansion is to lead to an accurate representation of the true dynamics, we must be sure that as we increase the number of modes, the resulting configurations of the displacement field do indeed converge to some asymptotic state, viz., the true state of the system. Figure 7 demonstrates the convergence of the dynamics with increasing mode numbers for the Chebyshev type-II basis function. The curves are difference curves between final-state configurations for a real-space nonlinear integration, and several mode-based-Lagrangian integrations. The residuals are small, and become smaller *as more modes are added to the expansion,* indicating that the quality of the approximation is increasing with added modes. This is one result that demonstrates that the configurations of the system, including its intrinsic nonlinearities, are adequately described using a finite number of basis functions.

[1]See, e.g., G. R. Barsch and J. A. Krumhansl, Phys. Rev. Lett. **53**, 1069 (1984).

[2]M. B. Olson, G. B. Olson, and P. C. Clapp, *Proceedings of the ICOMAT-79, Cambridge, 1979* (MIT Press, Cambridge, MA, 1979), p. 1.

[3]See, e.g., S. M. Shapiro, J. D. Axe, R. J. Birgeneau, J. M. Lawrence, and R. D. Parks, Phys. Rev. B **16**, 2225 (1977).

[4]See, e.g., L. E. Tanner, D. Schryvers, and S. M. Shapiro, Mater. Sci. Eng. A **127**, 205 (1990), and references therein.

[5]G. S. Bales and R. J. Gooding, Phys. Rev. Lett. **67**, 3412 (1991).

[6]We have formulated a Langevin description of elastic systems and have compared typical thermal forces with the mechanical forces that arise in our deterministic dynamics. Unless one is dealing with transitions that occur at very high temperatures, the mechanical forces are always much larger. Even for some high-temperature transitions involving only strain, e.g., the bcc-to-fcc transition undergone by La, the random forces are still sufficiently small that, to a very good aproximation, they can be ignored.

[7]A. C. E. Reid and R. J. Gooding, Physica D **66**, 180 (1993).

[8]R. J. Gooding and A. C. E. Reid, in *Nonlinear Coherent Structures in Physics and Biology,* edited by F. G. Mertens and K. H. Spatschek (Plenum, London, 1994).

[9]L. D. Landau and E. M. Lifschitz, *Theory of Elasticity,* 3rd ed. (Pergamon, Oxford, 1986).

[10]R. A. Cowley, Phys. Rev. B **13**, 4877 (1976).

[11]A. C. E. Reid and R. J. Gooding, Phys. Rev. B **46**, 6045 (1992).

[12]P. C. Martin, O. Parodi, and P. S. Pershan, Phys. Rev. A **6**, 2401 (1972).

[13]See, e.g., *Handbook of Mathematical Functions* (National Bureau of Standards, Washington, 1964).

[14]J. H. Yao and R. J. Gooding, J. Comp. Phys. **112**, 382 (1994).

[15]F. Falk, Acta Metall. **28**, 1773 (1980).

[16]C. Manolikas and S. Amelinckx, Phys. Status Solidi A **60**, 607 (1980).

[17]R. J. Gooding and G. S. Bales, Physica D **55**, 251 (1992).

[18]For a modern treatment of this problem, see, e.g., B. Horowitz, G. R. Barsch, and J. A. Krumhansl, Phys. Rev. B **43**, 1021 (1991), and references therein.

[19]See, e.g., Fig. 4 of L. Delaey, G. Guénin, Y. Murakami, P. F. Gobin, G. Van Tendeloo, and J. Van Landuyt, Phys. Status Solidi A **89**, 457 (1985).

[20]S. K. Chan, J. Chem. Phys. **67**, 5755 (1977).

[21]W. Cao and G. R. Barsch, Phys. Rev. B **41**, 4334 (1990).