Part I Introduction

Cambridge University Press
0521661471 - Evolution of Phase Transitions: A Continuum Theory
Rohan Abeyaratne and James K. Knowles
Excerpt
More information

1 Introduction

1.1 What this monograph is about

Certain crystalline materials can exist in more than one solid phase, where a phase is identified by a distinct crystal structure. Typically, one phase is preferred under certain conditions of stress and temperature, while another is favored under different conditions. As the stress or temperature varies, the material may therefore transform abruptly, from one phase to another, leading to a discontinuous change in the properties of the body. Examples of such materials include the shape-memory alloy NiTi, the ferroelectric alloy $BaTiO_3$, the ferromagnetic alloy FeNi and the high-temperature superconducting ceramic alloy $ErRh_4B_4$. In each of these examples the transition occurs without diffusion and one speaks of the transformation as being *martensitic* (or displacive).

Alloys such as Au–47.5%Cd and Cu–15.3%Sn are known to have a cubic lattice at high temperatures and an orthorhombic lattice at low temperatures. Therefore, if such a material is subjected to thermal cycling, it will transform between these two phases. Similarly, alloys such as Ni–36%Al and Fe–7%Al–2%C transform between a high-temperature cubic phase and a low-temperature tetragonal phase, whereas near-equiatomic NiTi has a high-temperature cubic phase and low-temperature monoclinic phase.

If a stress-free single crystal of such a two-phase material is slowly cooled from a sufficiently high temperature, it starts out in the high-temperature phase and at first, merely undergoes a thermal contraction. However at some critical temperature, denoted M_s in the materials science literature[1], a portion of the lattice suddenly transforms from the high-temperature structure to the low-temperature structure. Most of the specimen is still in the high-temperature phase, but a small amount is now in the low-temperature phase. The interface between the regions occupied by the two phases is a "phase boundary." As the temperature is further decreased, the phase boundary propagates into the high-temperature phase, thus transforming

---

[1] The subscripts s and f denote "start" and "finish" respectively, whereas M and A denote martensite, the low-temperature phase, and austenite, the high-temperature phase. So, for example, M_s is the temperature at which the formation of $\underline{m}$artensite $\underline{s}$tarts, etc.

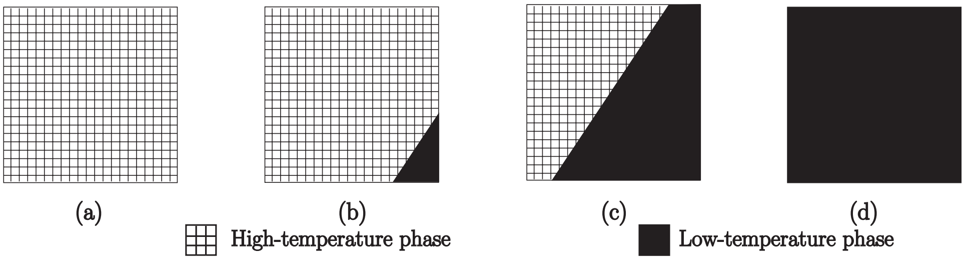

Figure 1.1. Schematic depiction of a stress-free single crystal of a two-phase material at a sequence of progressively decreasing temperatures.

it into the low-temperature phase as it crosses the moving front. Eventually, at some temperature M_f, the entire specimen is in the low-temperature phase. Further cooling simply causes thermal contraction of the low-temperature phase. This is depicted in Figure 1.1 by the sequence of diagrams (a)→(b)→(c)→(d).

If the specimen is now gradually reheated, the reverse process takes place. The high-temperature phase is nucleated at some temperature A_s, it grows by the propagation of phase boundaries, and finally the entire specimen returns to the high-temperature phase at a temperature A_f. Further heating simply causes thermal expansion.

If during the aforementioned cooling–heating cycle one plots some representative linear dimension ℓ of the specimen versus the temperature θ, one obtains a hysteresis loop as shown schematically in Figure 1.2. The cooling portion of the cycle corresponds to ① → ② → ③ → ④ → ⑤. Figure 1.1(a) corresponds to the initial cooling segment ① → ② → ③ during which the entire specimen is in the high-temperature phase. The low-temperature phase is nucleated at ③ and grows from ③ → ④ as the specimen evolves from (a)→(b)→(c)→(d) in Figure 1.1. The transformation is complete at ④ (Figure 1.1(d)) and thereafter undergoes pure thermal contraction from ④ → ⑤. The reverse transformation ⑤ → ④ → ⑥ → ② → ① occurs during reheating.

If, instead of being stress free, the specimen was held at some constant nonzero level of stress (in a suitable range) during the thermal cycling, one observes a similar sequence of events but the quantitative details will be different. Likewise, if instead of thermal loading, the specimen is subjected to a cyclic mechanical loading, one again observes a qualitatively similar sequence of events. One can consider a wide range of loading rates, either thermal or mechanical, with the associated response ranging from quasistatic to inertia driven, and isothermal to adiabatic, and in each case (provided one stays in a suitable range) one can induce forward and reverse phase transitions, the detailed response depending on the specifics of the loading conditions.

A conceptually helpful way in which to think about the modeling of a martensitic phase transition using the continuum theory of finite thermoelasticity is as follows: a general thermoelastic material is characterized by its free energy potential $\psi(\mathbf{F}, \theta)$, which is a function of the deformation gradient tensor $\mathbf{F}$ and the

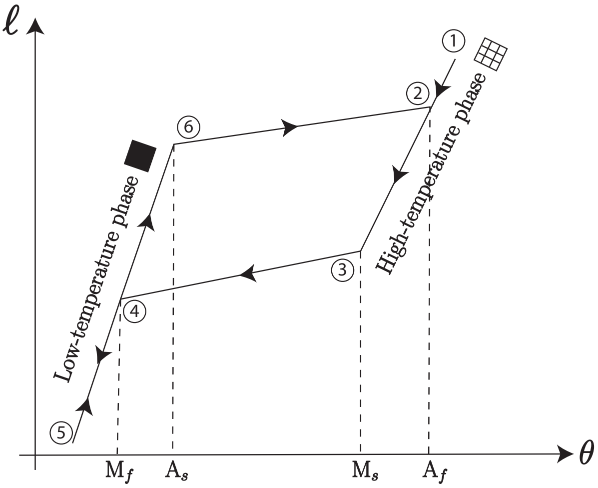

Figure 1.2. A schematic plot of a representative linear dimension ℓ of the specimen versus temperature θ during a cycle of cooling and reheating. The specimen transforms between a high-temperature phase and a low-temperature phase.

absolute temperature θ. The particular form of $\psi(\mathbf{F}, \theta)$ corresponding to a given material is determined by the microstructural details of that material. Consider a given thermoelastic body at a temperature θ. Suppose that at this θ, the potential ψ as a function of $\mathbf{F}$ has a local minimum at $\mathbf{F} = \mathbf{F}^+$. Then, since the gradient of ψ with respect to $\mathbf{F}$ is essentially the stress, the pair $(\mathbf{F}^+, \theta)$ describes a homogeneous stress-free equilibrium state of this body in the sense that if the entire body is homogeneously deformed by the deformation gradient tensor $\mathbf{F}^+$, and the temperature is held at θ, then the body will be stress free and can therefore be maintained in equilibrium without the application of any external loading.

If $\psi(\cdot, \theta)$ happens to have *two* local minima, say at $\mathbf{F}^+$ and $\mathbf{F}^-$ as depicted schematically in Figure 1.3(a), and if $\mathbf{F}^+$ and $\mathbf{F}^-$ do not differ by a rigid rotation, then there are *two* distinct stress-free homogeneous equilibrium states $(\mathbf{F}^+, \theta)$ and $(\mathbf{F}^-, \theta)$, *either of which* the body can occupy. Typically, each of these equilibrium states is associated with a distinct *phase* of the material.

Under certain circumstances the two phases can coexist, with part of the body being in the homogeneous stress-free equilibrium state $(\mathbf{F}^+, \theta)$, and the remainder being in the homogeneous stress-free equilibrium state $(\mathbf{F}^-, \theta)$; the interface separating these two parts of the body is a phase boundary. Since each part of the body is stress free, this piecewise homogeneous configuration will be in mechanical equilibrium without the application of any external loading.

If the values $\psi^+ = \psi(\mathbf{F}^+, \theta)$ and $\psi^- = \psi(\mathbf{F}^-, \theta)$ of the free energy associated with the two phases are equal, then a configuration that involves both phases is in *phase equilibrium* as well as in mechanical equilibrium: each phase has the same energy and therefore neither is preferred over the other. On the other hand if, for example, $\psi^+ > \psi^-$, as in Figure 1.3(a), then the phase associated with $\mathbf{F}^-$ is the low-energy phase and is therefore energetically favored over the other phase.

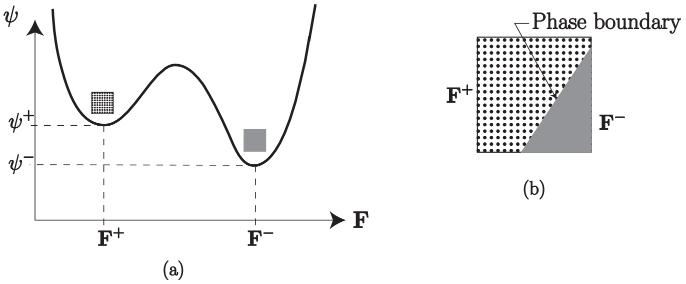

Figure 1.3. (a) A schematic plot of the free energy ψ versus the deformation gradient $\mathbf{F}$ at a fixed temperature, showing local minima at $\mathbf{F} = \mathbf{F}^+$ and $\mathbf{F} = \mathbf{F}^-$. (b) Two phases characterized by $\mathbf{F}^+$ and $\mathbf{F}^-$ coexisting in a body, separated by a phase boundary.

There is now a "thermodynamic driving force," which tends to transform the high-energy phase "plus" into the low-energy phase "minus". Accordingly, this driving force makes it favorable for the phase boundary to move into the region associated with $\mathbf{F}^+$, thereby transforming the high-energy phase into the low-energy phase as it crosses the moving interface. The driving force vanishes in phase equilibrium. How fast the phase boundary moves – the question of "kinetics" – is determined by the microstructure underlying the two phases and the details of the mechanism by which one lattice changes into the other.

In order to relate this discussion to the earlier one, we simply need to relate the plus and minus phases here to what we previously called the high-temperature and low-temperature phases. For example, if, as a function of the temperature θ, the energies $\psi(\mathbf{F}^+, \theta)$ and $\psi(\mathbf{F}^-, \theta)$ are such that $\psi^+ > \psi^-$ for high temperatures and $\psi^+ < \psi^-$ for low temperatures, then the minus phase is energetically preferred at high temperatures and can therefore be identified with what we previously called the high-temperature phase; similarly the plus phase can be identified with the low-temperature phase.

Next, consider a stress-free body composed of two coexistent phases that are in mechanical and phase equilibrium. If the body is now subjected to some stress, this alters the relative favorability of the two phases and will, in principle, lead to phase boundary motion. Similarly, this will happen if the temperature is changed.

It should be noted that in the *phenomenological* theory of thermoelasticity, if one is simply given a potential ψ with multiple local minima, one cannot say much about the different phases associated with the different minima. They might, for example, be associated with the liquid and gaseous phases of an elastic fluid, two crystallographically distinct lattices of a crystalline solid, the disordered and ordered states of a polymer, and so on. In order to physically identify the phases one must consider the underlying microstructure.

This monograph explores in detail a number of the concepts referred to above. We study questions related to the quasistatic and dynamic responses of an elastic material characterized by a free energy potential with multiple local minima. We pay particular attention to the notions of *driving force* and *kinetics*. The driving force plays a central role in characterizing the dissipation rate during time-dependent processes, and we make this notion rigorous under very general circumstances and then specialize it to various particular cases. We address both modeling and mathematical issues. For example on the modeling front, we calculate an explicit free energy function ψ for a material with cubic and tetragonal phases. We develop various models, both phenomenological and micromechanical, of the transformation kinetics. And we solve a number of boundary-initial value problems connected to experiments. On the mathematical front, we show that quasistatic and dynamic problems for two-phase materials, when formulated as in classical elasticity, suffer from a severe lack of uniqueness of solution to boundary–initial value problems. We show that the degree of nonuniqueness is *precisely* filled up by a proper description of the nucleation and kinetics of the transformation.

This monograph is not meant to provide an encyclopedic treatment of phase transitions. Rather, it simply collects together, and retells as a single story, the results from a series of studies mostly by the present authors on the motion of phase boundaries in solids.

1.2 Some experiments

The article by Schetky [45] provides a general introduction to the subject of martensitic transformations while the book by Duerig et al. [18] describes many engineering applications. Comprehensive overviews of the materials science literature on this subject can be found in Christian [14], Funakubo [25], and Otsuka and Wayman [42] among others

Experiments that investigate the thermomechanical response of martensitic materials can, roughly, be grouped into three sets: uniaxial tensile loading experiments, multiaxial loading experiments, and high-rate experiments.

There is an extensive experimental literature devoted to tensile loading and unloading of bars made of martensitic materials; often, the materials studied are technologically important shape-memory alloys such as NiTi. For a small sample of this literature, see, for example, Gao and Brinson [26], Krishnan and Brown [32], Nakanishi [39], Shaw and Kyriakides [46], and Leo *et al.* [36]. The loading in such experiments is slow, in the sense that inertia is insignificant. Some of these experiments are so slow that the process is essentially isothermal; but others, such as those of Shaw and Kyriakides [46] and Leo *et al.* [36], are sufficiently fast that the effects of local heating/cooling at the phase boundary become significant. The objective of these experiments is typically to determine the relation between the applied stress and the overall elongation of the bar, though in some studies such as that of Shaw and Kyriakides [46], local strain and temperature measurements along the length of the specimen are also made. The observed stress–elongation relation exhibits hysteresis, and the phase transition represents the primary mechanism responsible for this dissipative behavior. For a given material, the size and other features of the hysteresis loop depend on the loading rate and the temperature at which the test takes place.

Investigations on the multiaxial loading of martensitic materials are not as numerous. Among the comparatively few that do exist are the tension-torsion experiments of Lim and McDowell [37] and the biaxial tensile loading experiments of Chu and James [1, 15]. During a typical test, two stress components are prescribed as functions of time and the response of the conjugate kinematic quantities is measured, or vice versa, resulting, again, in a hysteretic response.

There is a large body of experimental literature describing the response of solids to shock- or impact-loading. Much of it is motivated by questions concerning the behavior of materials at extremely high pressures, as occurs, for example, deep in the earth; see for example, Meyers [38]. In a typical experiment, a specimen in the form of a cylindrical disk is subjected to a high-speed normal impact on one of its plane faces. The collision generates a compressive disturbance that propagates into the disk, ultimately reaching the rear face of the target. Propagation speeds of the disturbances involved are measured, as are particle velocities at the rear face. Depending on the material, a phase transition may occur, as, for example, in the case of the graphite-to-diamond transition studied by Erskine and Nellis [20], or the austenite-to-martensite transition in CuAlNi studied by Escobar and Clifton [21]. One goal of such experiments is to determine the relation between the impact speed and the resulting speed of the phase transformation.

Some of the aforementioned experiments have been carried out on polycrystalline specimens of the material, while others use single crystals. Experiments on single crystals usually aim to understand the local processes associated with the propagation of phase boundaries and to relate this to the macroscopic response, for example, Chu and James [1, 15], Gao and Brinson [26].

In most martensitic materials, one does *not* observe a simple piecewise homogeneous deformation during phase transformation involving two homogeneously deformed phases separated by a planar phase boundary. Usually, the crystallographic structure on one side of a phase boundary is more complicated than this. In some cases, for example, In–22.73%Tl, one observes a homogeneously deformed high-temperature phase separated by a planar phase boundary from a low-temperature phase that has a “twinned” fine structure. Another common microstructure involves a wedge of the twinned low-temperature material surrounded by the homogeneously deformed high-temperature phase, seen, for example, in Ni–36%Al. However there do exist some exceptional materials, such as some of the titanium–tantalum alloys studied by Bywater and Christian [13], which do exhibit simple piecewise homogeneous deformations during phase transformation. As will be illustrated in Chapter 12, the ability to form an interface between the high-temperature and low-temperature phases without twinning requires that the lattice parameters of the two phases be related in special ways. It is conceivable that one could adjust the composition of a shape memory alloy so as to make the lattices satisfy these special relationships, and therefore to make the material exhibit phase boundaries that do not involve finely twinned structures.

1.3 Continuum mechanics

In 1975, Ericksen [19] examined the equilibrium of an elastic bar composed of a material with a nonmonotonic stress–strain relation. The strain-energy function associated with this stress–strain relation had two local minima and could therefore be used to describe a two-phase material. Following on this seminal paper, there has been a vast amount of activity in the continuum mechanics community on the modeling of solid–solid phase transitions.

Work on this subject can broadly be divided into two parts. One set of studies has been focused on examining equilibrium states of a two-phase solid using energy minimization as the underlying principle. The richness of this subject stems from the fact that, as mentioned previously, observed microstructures can be quite complicated. The low-temperature phase often involves a fine structure ("twinning") and in some cases even involves a fine structure within a fine structure. Understanding and modeling these complex microstructures has been a challenging and rewarding subject. The recent book by Bhattacharya [10] and the earlier review article by James and Hane [29] provide comprehensive discussions of the theory of equilibrium microstructures.

The second set of studies concerns the time-dependent response of these materials corresponding to various mechanical and/or thermal loading processes. At the microscale, the goal is to understand the evolution of the microstructures, while at the macroscale one wishes to describe various aspects of the hysteretic macroscopic response. Internal-variable-based models have been developed and explored by many investigators, for example, Brinson, Lagoudas, Müller, and co-workers (e.g. [12, 11, 8]). Phase field models have also been successfully used by others including Fried, Khachaturyan, Voorhees, and co-workers, for example, [24, 9, 49]. Models that use viscosity to capture dissipation and a strain-gradient term to enable propagation have been particularly popular and feature prominently in the work of, for example, Shearer [47], Slemrod [48], and Truskinovsky [51].

Irrespective of the details of a specific material system, a hysteretic response involves energy dissipation, and one of the fundamental questions pertains to how one characterizes dissipation in a system where the primary source of dissipation is phase transformation. Knowles and Sternberg [31] and Knowles [30] showed that, much like a propagating shock wave in an inviscid compressible fluid, or a propagating crack in an elastic solid, the strain discontinuity that occurs across a propagating phase boundary can be a source of dissipation even in slow processes. This, coupled with the basic notions of irreversible thermodynamics, leads one to the notion of the "driving force" on a phase boundary – the primary agent responsible for entropy production due to phase transformation; see Abeyaratne and Knowles [2, 6], Heidug and Lehner [28], Truskinovsky [50]. It is related to the general concept of configurational forces as developed by Gurtin [27], as well as to notions introduced much earlier by Eshelby [22]. Thus, during the process of phase transformation, there is a driving force on a phase boundary, and as material crosses this moving interface, it transforms from the high-energy phase to the low-energy phase. The flux that is conjugate to this driving force is the rate at which the material transforms, or equivalently, the (Lagrangian) speed of the interface. The process of phase transformation, like any nonequilibrium process, is governed by a kinetic law that relates the flux to the driving force and other appropriate variables. Therefore in order to fully model the time evolution of a phase transition, one must specify an elastic potential that characterizes each phase and a kinetic law that characterizes the evolution of the phases. The detailed form of the kinetic law depends on the microscale dynamic process by which one lattice transforms to the other. Some preliminary efforts to derive kinetic models from lattice considerations can be found in Abeyaratne and Vedantam [7], and Purohit [44].

1.4 Quasilinear systems

The second-order quasilinear system of partial differential equations

$$\frac{\partial h(u_1)}{\partial x} = \frac{\partial u_2}{\partial t}, \quad \frac{\partial u_2}{\partial x} = \frac{\partial u_1}{\partial t}$$

for two unknowns $u_1(x, t)$, $u_2(x, t)$ has been extensively studied in the applied mathematics literature, for example [23]. This system arises in many physical contexts, for example, in gas dynamics, it describes the flow of a compressible inviscid fluid in which case one identifies u_1 with mass density, u_2 with particle velocity, and $h(u_1)$ with the function characterizing the pressure–density relation of the gas. Likewise, by identifying u_1 with strain, u_2 with particle velocity and $h(u_1)$ with the function that describes the stress–strain relation of a solid, it can be used to describe the dynamics of a nonlinearly elastic bar.

In gas dynamics, it is well known that the Cauchy problem for this system suffers from nonuniqueness of solution, but that there is a unique weak solution that also satisfies the associated entropy inequality. More generally, for any function h that is either a strictly convex or a strictly concave function of u_1, the Cauchy problem for this system can be shown to have at most one weak solution fulfilling a so-called "admissibility criterion", Oleinik [40]. On the other hand, when h is not strictly convex or strictly concave, the Cauchy problem need not have a unique solution, *even with* the entropy inequality in force, Dafermos [17].

The issue of selecting a unique solution to a quasilinear system in the case where h is not a strictly convex or strictly concave function of u_1 has been addressed by many authors, usually by replacing the entropy inequality by a stronger "admissibility criterion," for example Dafermos [16], Lax [33], and LeFloch [35]. Among the most common such admissibility criteria are the viscosity criterion, the viscosity–capillarity criterion, the entropy-rate admissibility criterion, and the chord criterion, for example [43, 48, 16, 41].

The elastic bar considered by Ericksen [19] in his study of a two-phase material can be described by a function h that, with increasing strain u_1, first increases, then decreases, and finally increases again. Such a function h is neither strictly convex nor strictly concave, and so the quasilinear system suffers from the lack of uniqueness alluded to above, even with the entropy inequality in place. Rather than replacing the entropy inequality with a "uniqueness-generating admissibility criterion," we take the point of view that the lack of uniqueness reflects a deficiency in the physical model. The entropy inequality, being a statement of the second law of thermodynamics, must be retained. Rather, additional information from the physics of phase transformations needs to enter the mathematical model. Important insight into the information that is lacking can be obtained if one can determine *all* solutions to the quasilinear system and organize them in some illuminating way. We were able to do this explicitly in a special, algebraically simple case, [3], and thereby conclude that the nonuniqueness arises for two distinct reasons: it pertains to (i) the question of whether or not the motion involves a phase boundary and (ii) if the motion does involve a phase boundary, then the indeterminacy of the phase boundary speed. This suggests, exactly as the materials scientists would tell us, that the quasilinear system must be supplemented with a "nucleation condition" that addresses item (i) and a "kinetic relation" that addresses item (ii). In [3] we demonstrate that the solution to the Riemann problem for this quasilinear system is unique when the problem statement is supplemented in this way; see also LeFleeh [34, 35]. Each of the previously mentioned admissibility criteria can be viewed as corresponding to a particular (mathematical) kinetic relation, [4, 5], though deriving the kinetic law from the underlying physics is largely an open question.

1.5 Outline of monograph

This monograph is organized into four informal sections. The first section consists of Chapters 2–4 and introduces the main concepts of the subject within the simplest possible setting, viz. a purely mechanical theory for a one-dimensional elastic bar. The second section, consisting of Chapters 5–8, presents the general three-dimensional theory including both mechanical and thermal effects. In Chapters 9–11, which form the third section, we specialize the general theory to one dimension and then study some specific thermomechanical initial-boundary-value problems. In the final section of the monograph, Chapters 12–14, we consider some three-dimensional problems.

The first and third sections both concern one-dimensional settings; the former excludes thermal effects and the latter includes them. The second and fourth sections both concern three-dimensional settings; the former concerns the general theory and the latter concerns specific problems.

In slightly more detail, Chapter 2 sets out the purely mechanical theory for a one-dimensional elastic bar, Chapter 3 examines static configurations and quasistatic processes for it, and Chapter 4 examines a dynamic problem where inertial effects are important. In Chapter 5, we discuss, within a general three-dimensional context that includes both mechanical and thermal effects, a multiple-well Helmholtz free energy potential; two specific examples, one concerning a van-der Waals fluid and the other a martensitic material, are also presented. Chapter 6 is devoted to