PERGAMON

INTERNATIONAL JOURNAL OF
Plasticity

International Journal of Plasticity 16 (2000) 805–849

www.elsevier.com/locate/ijplas

Structural changes without stable intermediate state in inelastic material. Part I. General thermomechanical and kinetic approaches

Valery I. Levitas *

Texas Tech University, Department of Mechanical Engineering, Lubbock, Texas 79409-1021, USA

Received in final revised form 4 November 1999

Abstract

We define a structural change in inelastic material as a thermomechanical process of change in some region of material properties (elastic moduli, heat capacity, thermal expansion coefficient, entropy, yield strength), as well as a transformation strain from initial to final values. This process cannot be stopped at a material point in the intermediate state. Such a definition includes various phase transitions (martensitic, diffusional-displacive and second-order phase transitions), twinning, ductile fracture (nucleation and growth of void and cracks), solid-state chemical reactions, generation of point defects, dislocations, disclinations, deformation of amorphous materials, and so on. A theory suggested represents a generalization of the theory of martensitic phase transformations developed earlier by Levitas [Levitas, V.I., 1998a. Thermomechanical theory of martensitic phase transformations in inelastic materials. Int. J. Solids and Structures 35(9–10), 889–940] and its extension to the above phenomena. The theory includes a nonlocal thermodynamic criterion of structural changes in some finite region, a kinetic equation for the transformation rate (= inverse transformation time) and principle of minimum of transformation time derived for determination of all variable parameters (e.g. shape and volume of transforming region) and is valid for arbitrary inelastic material. Instead of surface-independent and path-independent integrals, which are an ideal tool for the description of various phenomena in elastic media (e.g. fracture, motion and interaction of various defects and singularities), the theory suggests a region-independent integral which represents a dissipation increment due to the structural change only during the complete structural change in transforming region. In some cases the principle of the minimum of transformation time reduces to the principle of the minimum of transforming mass and the characteristic size of the transforming region (nucleus) is determined either from the

---

* Tel.: + 1-806-742-3777; fax: + 1-806-742-3450.
E-mail address: valery.levitas@coe.ttu.edu.

PII: S0749-6419(99)00084-4

thermodynamic criterion of structural changes or is equal to the interatomic distance. A new kinetic concept of a thermodynamically admissible nucleus (void, crack) follows from the theory. A number of examples, related to displacive and diffusional-displacive phase transitions, strain-induced chemical reactions and ductile fracture, are considered in Part II of the paper.

Keywords: A. Phase transformation; A. Fracture; A. Chemo-mechanical processes; B. Constitutive behaviour; B. Kinetics

1. Introduction

It is known that a well-developed universal formalism can be applied to the description of the motion and interaction of various singularities in elastic materials (Eshelby, 1951, 1956, 1970; Rice, 1968a,b; Cherepanov, 1967, 1979). As examples, we can mention point defect, dislocation, crack tip, cavity, interface between different phases, which are responsible for such structural changes as plastic flow, creep, fracture, damage, phase transformations. The main idea is that with the motion of an arbitrary singularity the rate of change of the total energy can be represented as $\dot{E} = \mathbf{J} \cdot \dot{\mathbf{x}}$, where $\dot{\mathbf{x}}$ is the rate of motion of the defect as a whole with respect to the material and $\mathbf{J}$ is the generalized force acting on the defect (Eshelby, 1951, 1956, 1970). The expression for $\mathbf{J} = \int_{\Sigma} \mathbf{P} \cdot \mathbf{n} d\Sigma$ can be obtained in terms of the Eshelby energy-momentum tensor $\mathbf{P}$ integrated over any surface Σ with the unit normal $\mathbf{n}$ surrounding the defect and isolating it from others or over an actual surface defect like an interface. The results of Eshelby were extended for or rediscovered in many concrete fields. In fracture mechanics a generalized force $\mathbf{J}$ was introduced as a path-independent J-integral by Rice (1968a,b) or Γ-integral by Cherepanov (1967, 1979), for a martensitic phase transformation the same expression for J as in Eshelby's paper (1970) were obtained by many authors (see Kaganova and Roitburd, 1988 and references; Grinfeld, 1991). Further development and systematization of Eshelbian mechanics can be found in the works by Maugin (1993, 1995). From the thermodynamic point of view the independence of $\mathbf{J}$ of the surface Σ or path along which the integration is fulfilled can be explained as follows. The expression $\dot{E} = \mathbf{J} \cdot \dot{\mathbf{x}}$ is the total rate of dissipation in the volume v surrounded by the surface Σ. As for elastic materials, dissipation occurs due to the motion of singularities (defects) only, then adding to $\mathbf{J}$ the integral over the volume without moving singularities does not change the dissipation and $\mathbf{J}$. As is customary in irreversible thermodynamics, the force $\mathbf{J}$ is used to formulate the condition of defect equilibrium $\mathbf{J} = 0$ (or $\mathbf{J} \cdot \dot{\mathbf{x}} < \Gamma |\dot{\mathbf{x}}|$, where Γ is the surface energy) or the kinetic equation of defect motion $= \dot{\mathbf{x}} - \mathbf{f}(\mathbf{J})$. This is the usual way to derive the equation of motion of a point defect and dislocation, as well as an interface or crack propagation.

For inelastic materials (elastoplastic, viscoelastic, viscoplastic) dissipation due to plastic flow contributes to the total dissipation rate $\mathbf{J} \cdot \dot{\mathbf{x}}$ in the volume v surrounded by surface Σ and $\mathbf{J}$ ceases to be independent of the choice of surface Σ. Even when the volume v tends to zero, i.e. it includes the defect only, plastic dissipation in this singular point occurs as well (for example, in the crack tip) and it is difficult to separate dissipation due to the defect motion itself and inelastic deformation. That is why in a majority of papers on fracture mechanics (Rice, 1968a; Atluri, 1997; Kolednik et al., 1997) it is recognized that the strict thermodynamic criterion of ductile fracture is lacking and other approaches like critical plastic strain (Rice, 1968a; McClintock, 1971), critical crack tip opening displacements (Harrison, 1980), an energy flow in an infinitesimal or finite-sized process zone (Rice, 1968a; Cherepanov, 1979; Hutchinson, 1983; Atluri, 1997), plastic work (Kuang and Chen, 1997), total dissipation rate (Kolednik et al., 1997; Turner and Kolednik, 1997) and so on are assumed. Other authors (Cherepanov, 1979, 1997) continue to use the same fracture criterion and kinetic relations in terms of **J** (or Γ-integral) for arbitrary inelastic materials neglecting the necessity of separation of plastic dissipation. A similar situation took place in phase transformation (PT) theory. Physicists (Kaganova and Roitburd, 1989) had a feeling that plastic dissipation at a moving interface had to be excluded, but did not express this completely correctly from the point of view of thermodynamics. In papers by Fischer et al. (1994), Cherkaoui et al. (1998), Fischer and Reisner (1998) the interface propagation condition is formulated with the help of the energy-momentum tensor without excluding the plastic dissipation.

As in elastic materials the conditions both of nucleation and of equilibrium of defect [for example, the appearance of a nucleus of a new phase and its phase (interface) equilibrium] can be derived from the principle of the minimum of Gibbs free energy, these conditions coincide. Consequently, from a thermodynamic (but not a kinetic) point of view it is not necessary to consider a nucleation process; it is sufficient to introduce a nucleus and consider its interface equilibrium using the local condition $\mathbf{J} = 0$. It was mentioned in the paper by Roitburd and Temkin (1986) that nucleation and interface propagation conditions in elastoplastic materials do not coincide. Consequently, *a nucleation problem* in a finite volume (due to surface energy) of inelastic materials arises. We will call a nucleation of some defect the structural change (SC) and the region in which SC occurs the transforming region. Nucleation in inelastic materials cannot be described either by the principle of minimum of Gibbs energy, nor using an energy-momentum tensor. The known extremum principle of irreversible thermodynamics (Ziegler, 1977) cannot be applied directly either, because generalized force and rate were not defined. In fracture mechanics when a stress singularity does not exist (for instance, in perfectly plastic materials) or should be avoided (models with a cohesive zone), the dimension of the transforming (fracture) region increases from a one (crack tip line) to a two (cohesive zone surface) or to a three-dimensional process zone as well.

In order to estimate the thermodynamic possibility of defect nucleation or motion in elastic materials, it is sufficient to compare the energetic state before and after this process without knowledge of the entire process. As the behaviour of inelastic materials is history dependent, consideration of the *entire deformation process in the transforming region* is necessary. This results in several fundamental consequences. For elastic materials it is possible to avoid a consideration of a singular point and surface and to describe their energetics using an integral over some surface surrounding the singularity (for dislocation, point defect or crack tip) or values of parameters in points approaching the interface from both sides. For plastic materials it is necessary to consider *the transforming region* during the entire transformation process and consequently, it is necessary to know all the thermomechanical properties for each intermediate state of the transformation process.

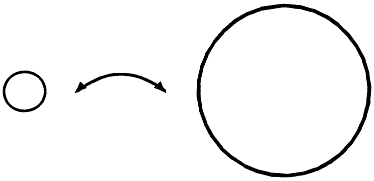
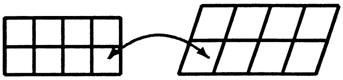
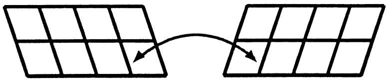
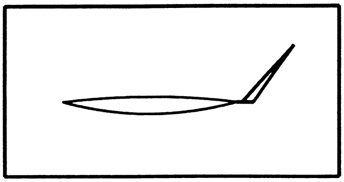
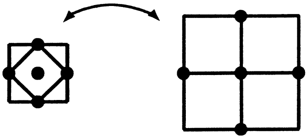
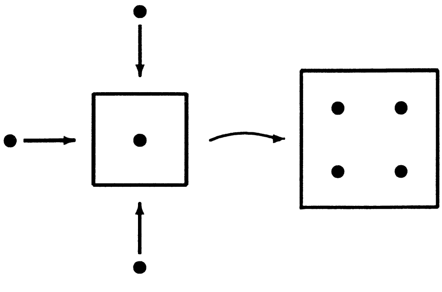
All the above remarks give the impression that it is very difficult or impossible to develop a correct thermodynamic description for each concrete SC in inelastic materials and to find a unified approach for a wide class of SC, comparable with the one known for elastic materials. The aim of this paper is an attempt to develop unified continuum thermodynamic and kinetic approaches for a wide class of SC which includes in particular displacive and diffusional-displacive PT, twinning, solid state chemical reactions and ductile fracture.

This approach will be based on the extension of a thermodynamic theory of martensitic PT in inelastic materials developed during the last six years (Levitas, 1992a, 1995a,b, 1996, 1997) and presented in the most complete form in Levitas (1998a). PT is treated as a thermomechanical deformation process of growth of transformation strain from the initial to the final value in some transforming region which is accompanied by a change in all the properties of the material. Continuum thermomechanics is extended to the description of such a process and then allows us to separate the dissipation rate due to PT only, which is the most important result for the formulation of PT criterion and kinetics. A new plausible assumption called a postulate of realizability (Levitas, 1992a, 1995a, 1998a) is formulated and applied to the derivation of an extremum principle to determine all the unknown parameters, like position, shape and orientation of the nucleus and the variation of all the fields in it. Based on this theory, a number of boundary-value problems were formulated and solved analytically (Levitas, 1997, 1998a; Levitas et al., 1998b,c) or numerically (Idesman, Levitas and Stein, 1997, 1998; Levitas et al., 1998a, 1999). A number of experimental results are explained (for instance, PT under compression and shear of materials in Bridgman anvils, strain-induced PT in steel, pressure-induced PT graphite-diamond; incoherent PT), and some of the interpretations are completely unexpected. From the other side, this theory is developed for time-independent (athermal) kinetics only. As will be shown in Section 4, this can be contradictory in some cases.

For the description of strain-induced chemical reactions in a shear band (Levitas et al., 1998b,c) the above theory was supplemented by a kinetic equation, however, not in a systematic way. In a short communication by Levitas (1998b) an example of a possible application of this theory to ductile fracture is presented and the kinetic concept is further extended.

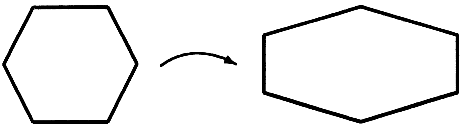
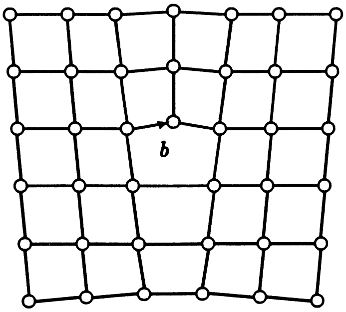
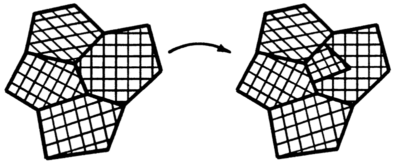
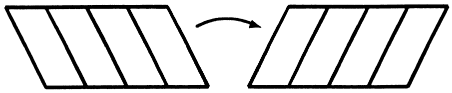
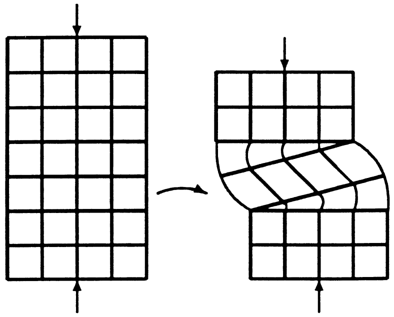
In the present paper, a systematic formulation of the thermodynamic and kinetic theory of SC in inelastic materials will be developed and new results of the application of the theory to various concrete SC will be presented. In Section 2 a model for SC without a stable local intermediate state is defined which includes such SC as displacive and diffusional-displacive PT, twinning, generalized second order PT, recrystallization, chemical reactions, ductile fracture and some other (see Table 1). The key assumption is that the transformation process in each point of the transforming region *cannot be stopped in an intermediate stage*. The thermodynamics of this kind of SC is developed in Section 3. In contrast to previous studies, the global form of

Table 1
Types of structural changes without stable intermediate state and their peculiarities

PT liquid $\leftrightarrows$ gas, crystallization, melting, cavitation 	Transformation strain — pure volumetric. Change in all physical properties
Displacive PT 	Transformation strain — arbitrary. Change in all physical properties
Twinning 	Transformation strain — simple shear. Change in tensorial properties due to lattice rotation
Generalized second-order PT	Transformation strain and entropy are zero. Change in all physical properties
Fracture: cracks, voids 	Change in tensile and shear elastic moduli, thermal expansion coefficient, heat capacity from intial value to zero
Chemical reactions 	Transformation strain — arbitrary or pure volumetric. Change in all physical properties
Diffusional–displacive PT 	Transformation strain arbitrary. Change in the volume fraction of saturating element and in all the physical properties

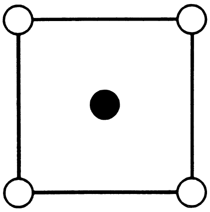
(Table continued overpage)

Table 1 (*continued*)

Plastic deformation of amorphous materials by rearrangement 	Transformation strain arbitrary
Dislocation nucleation  	Growth of Burgers vector from zero to final value **b** (or transformation shear from zero to $\frac{\lvert\mathbf{b}\rvert}{a_a}$)
Recrystallization (static, dynamic) 	Transformation deformation gradient — pure rotation. Decrease in defect density. Change in tensorial properties and surface energy due to rotation of crystal lattice
Reorientation of martensitic variants 	Transformation strain arbitrary volume preserving. Change in tensorial properties due to rotation of crystal lattice
Crystal lattice reorientation 	Transformation deformation gradient — pure rotation. Change in tensorial properties due to rotation of crystal lattice

(Table continued overpage)

Table 1 (*continued*)

Inclusions: interstitial and solute atoms 	Transformation strain — arbitrary or pure volumetric. Change in all physical properties

thermodynamic laws is used as a starting point. Even when all the constitutive equations describing deformation are local, for SC only a nonlocal equation can be derived due to surface energy. It is shown that for a model of SC without a stable local intermediate state application of a standard dissipative inequality local in time (e.g. the Clausius–Duhem inequality) is *contradictory*. It is necessary to average the thermodynamic parameters related to SC over the duration t_s of SC in order to filter off fluctuations. Consequently, we arrive at *a nonlocal formulation both in space and time*. The dissipation rate due to SC only is separated from other dissipative contributions. An explicit expression for the generalized (driving) force and rate is derived: the force is the total dissipation increment due to SC only during the complete SC in the transforming region; the rate is the inverse SC duration. As the driving force is a volume integral related to dissipation due to SC only, adding a similar integral over an arbitrary volume without SC (which is identically zero) does not change the force. Consequently for arbitrary inelastic materials the driving force for SC represents *a region-independent integral*. In Section 4 the thermodynamic criterion of SC and the general kinetic equation between generalized force and rate are formulated. It is shown that athermal (time-independent) kinetics can lead to contradictions. In Section 5 an equation obtained is transformed to the form of an interface propagation condition. In Section 6 extremum principles for the determination of all unknown parameters are derived using the postulate of realizability. The time-dependent kinetics is treated in this context for the first time and the principle of minimum of transformation time is derived.

In Section 7 an example of thermally activated kinetics is considered and the concept of effective temperature is introduced. Detailed analysis of the new interrelation between the criterion for SC, the kinetic equation and extremum principle are given in Sections 8 and 9 for the case where the local driving force is independent of the geometry of the nucleus. This is related to such processes as crystallization, melting or cavitation; however, similar considerations will be used in Part II of this paper for displacive PT, crack propagation and void nucleation. Then the principle of the minimum of the transformation time can be reduced in this case to the principle of the minimum of transforming mass. If the new nucleus does not touch the previous ones, its minimum characteristic size (radius, side of cube) is determined from the thermodynamic criterion of SC. The concept of a thermodynamically admissible nucleus introduced has nothing in common with the well-known critical nucleus. The appearance of the critical nucleus contradicts the second law of thermodynamics. If the new nucleus touches the previous one (in this way we simulate its growth), then the change in surface energy and the minimum characteristic size of the nucleus are much smaller and its size is often limited by interatomic distance. Consequently, the theory predicts accelerated growth kinetics after nucleation due to decrease in transforming volume during each event and in particular predicts the transition to atom by atom growth. This is illustrated in Section 9 when we consider the appearance and growth of a cubic nucleus.

The theory suggested will be applied in Part II of this paper to the description of displacive and diffusional-displacive PT, strain-induced chemical reactions, void nucleation and ductile crack propagation.

Direct tensor notations are used throughout this paper. Vectors and tensors are denoted in boldface type; **mn** is the dyadic product of vectors **m** and **n**; $\mathbf{A}\cdot\mathbf{B}$ and $\mathbf{A}:\mathbf{B}$ are the contraction of tensors over one and two indices. A superscript -1 denotes inverse operation, *tr* the trace operation, $[\mathbf{a}]=\mathbf{a}_2-\mathbf{a}_1$ the jump of value **a** across the interface, $\Delta\mathbf{a}=\mathbf{a}_2-\mathbf{a}_1$ is the difference in **a**, **I** the unit tensor second order, *dev***A** is a deviatoric part of **A**, $|\mathbf{A}|:=(\mathbf{A}:\mathbf{A})^{1/2}$ the modulus (amplitude) of tensor $\mathbf{A}, \boldsymbol{\nabla}$ is the gradient operator, $:=$ means equals per definition, the indices 1 and 2 denote the values before and after SC.

2. Model for structural changes without stable local intermediate state

Let us consider an infinitesimal or finite volume undergoing a process of SC. The volume can belong to the new nucleus or region covered by moving interface. All the SC which can be treated in the framework developed in this paper are schematically presented in the Table 1. The following features are important for the model under consideration.

1. As a main geometrical characteristic of the SC under study we consider the transformation deformation gradient $\mathbf{F}_t$ which relates the geometry of infinitesimal material volume in the stress-free state and some reference temperature θ_o before, during the SC (current value of $\mathbf{F}_t$) and after the SC (final value of $\mathbf{F}_{t2}$ in phase 2) in the absence of other inelastic strains.

   For martensitic or so-called displacive PT, the transformation deformation gradient $\mathbf{F}_t$ transforms the crystal lattice of the parent phase into the crystal lattice of the product phase. For twinning the transformation deformation gradient is simple shear. Using the polar decomposition theorem we obtain $\mathbf{F}_t=\mathbf{R}_t\cdot\mathbf{U}_t$,, where $\mathbf{R}_t$ is the orthogonal rotation tensor and $\mathbf{U}_t$ is the symmetric right stretch transformation tensor. The determinant $det\mathbf{F}_t$ characterizes the volume change due to SC. For small strains the transformation strain is $\boldsymbol{\varepsilon}_t:=\mathbf{U}_t-\mathbf{I}$.

For martensitic PT the transformation deformation gradient after finishing the SC cannot be arbitrary (as elastic or plastic strain). For each PT the right stretch transformation tensor $\mathbf{U}_{t2}$ after SC is some fixed tensor to within symmetry operations. All intermediate values of the transformation right stretch tensor are unstable and cannot exist in an equilibrium after finishing the SC. Due to the symmetry there is a finite number (e.g. 12 for the PT from a cubic to monoclinic lattices) of crystallographically equivalent variants of martensite with the same (to within symmetry operations) $\mathbf{U}_{t2}$.

If one martensitic variant with transformation deformation gradient $\mathbf{F}^{a}_{t2}$with respect to the parent phase transforms into another one with transformation deformation gradient $\mathbf{F}^{b}_{t}$ with respect to the parent phase, then the corresponding transformation gradient is $\mathbf{F}^{b}_{t}\cdot\left(\mathbf{F}^{a}_{t2}\right)^{-1}$. As both $\mathbf{F}^{a}_{t}$ and $\mathbf{F}^{b}_{t}$ have the same volumetric part, the transformation between martensite variants is volume preserving. For reorientation of the crystal lattice without a change in the crystal cell, for example with the deformation of a single crystal (Klassen-Neklyudova, 1960) or static and dynamic recrystallization the transformation deformation gradient is an orthogonal rotation tensor. Plastic deformation of amorphous materials by rearrangement of particles can be described by the introduction of transformation strain as well (Bulatov and Argon, 1994). It is assumed that in the final state the transformation strain is a simple shear, but in some intermediate state it includes a dilatational component as well. Despite the fact that in this case (as well as for PT liquid $\leftrightarrows$ gas, melting, some chemical reactions) the transformation strain is not related to the deformation of a crystal lattice, it is not essential for a formal continuum theory.

Dislocation nucleation in a dislocation–free region of a crystal (e.g. when a Franck–Read mechanism of dislocation multiplication does not work), for example near the crack tip or near the interface between two phases, plays a very important part in the description of brittle to ductile transition at fracture (Argon, 1987; Rice, 1992) or transition from thermoelastic to nonthermoelastic (plastic) crystal growth in a martensitic PT (Olson and Cohen, 1986; Haezebrouck, 1987). For a dislocation in the simplest case (Eshelby, 1957) the transformation strain is a simple shear along the Burgers vector $\mathbf{b}$ with amplitude $\gamma = |\mathbf{b}|/a_{\mathrm{a}}$, where a_{a} is a characteristic interatomic distance. In continuum theory when $a_{\mathrm{a}} \rightarrow 0$ we have $\gamma \rightarrow \infty$ and instead of the transformation strain the displacement discontinuity $\mathbf{b}$ is a geometric characteristic of SC. Note that a dislocation with the variable Burgers vector, i.e. dislocation nucleation as the process of growth of Burgers vector, is considered by Argon (1987).

For diffusional-displacive, diffusional PT and chemical reactions it is possible to determine for some macroscopic volume the deformation gradient related to SC. In papers by Levitas et al. (1998b,c) it is assumed that for chemical reactions the transformation deformation gradient describes the change of volume only without change of shape, i.e. it is a spherical tensor $\mathbf{F}_{t} = a\mathbf{I}$.

2. Another important characteristic of SC is change in some or in all thermomechanical (which includes physical and chemical) properties, e.g. elastic properties, heat capacity, thermal expansion coefficient and yield strength,

entropy, hardening rules and so on. For some SC like first-order PT, chemical reactions, or twinning, the change in properties occurs simultaneously with the change in transformation strain; for other ones, like fracture, or second-order PT, the transformation strain is absent and change in properties is the only characteristic of SC. For an arbitrary SC we will call phase 1 the state before SC and phase 2 the state after SC; a process of SC will be called a transformation. The reason for the occurrence of SC under consideration for nondissipative materials is the loss of stability of the thermomechanical equilibrium of phase 1, for instance, of the parent crystal lattice or a state without crack. For dissipative materials it is not sufficient to consider the stability of the equilibrium state. Even if the equilibrium state of phase 1 or the parent crystal lattice is stable, loss of stability of a thermomechanical deformation process without transformation strain and change of properties can occur. For nondissipative materials the nonstable phase 1 or crystal lattice transforms into a new stable phase 2 or lattice. Dissipative materials undergo a stable deformation process with transformation strain and change of material properties.

3. We accept that the transformation process in each point of the transforming volume *cannot be stopped in an intermediate stage*. Consequently, the material in the state before SC (phase 1) and in the state after SC (phase 2) is stable only in each material point. We will call such SC *SC without a stable local intermediate state*. It is possible to consider a representative volume consisting of a mixture of phase 1 and 2 with a volume fraction c of the second phase, i.e. global (for finite volume) intermediate stable state exists, but each material point of this volume belongs either to phase 1 or to phase 2. Then during the SC in the given material point the transformation deformation gradient and material properties must reach the final values corresponding to phase 2 and cannot be fixed at any intermediate state. These features are similar to those for martensitic PT.

   Such a definition cannot be applied to the continuous transformation of the crystal lattice of one component by atoms of other substances or to spinodal decomposition. The theory for such processes is much simpler and is well developed.

Summarizing, let us give a definition of *SC without a stable local intermediate state*, which will be dealt with in the paper.

The SC will be considered as a process of variation of the transformation deformation gradient and some or all thermomechanical properties in an infinitesimal or finite transforming volume from the initial to final value. This process cannot be stopped at an intermediate state in any transforming point. The thermodynamic equilibrium for an intermediate value of the transformation deformation gradient or material properties is impossible.

Such a definition excludes from consideration the Landau–Ginzburg model (Olson and Cohen, 1986), in which due to a nonlocal term smooth transition from phase 1 to phase 2 occurs and any intermediate state can be stable inside a diffuse interface of finite thickness. We will concentrate on the case with a sharp interface and local constitutive equations describing the deformation in each material point.

3. Thermodynamics. Region–independent integral

3.1. General relations

Consider a volume V of multiphase material with a boundary S. Let on one part of surface S_p the traction vector $\mathbf{p}$ be prescribed and on the other part S_u the displacement vector $\mathbf{u}$ be given, but mixed boundary conditions are also possible. Assume that in some volume V_n with a fixed boundary Σ_n relative to the material's points, due to SC during the time t_s the new nucleus of phase 2 appeared, i.e. some material mass m_n undergoes the SC (Fig. 1).

Assume that we know the boundary Σ_n of the nucleus after the nucleation event and we will fix it during the entire nucleation event. The change in all the fields and material properties in a volume V_n are shown schematically in Fig. 2 for a one-dimensional nucleus. The line *AD* corresponds to the state before the SC when the transformation strain and all the properties are equal to those in phase 1. The line *BC* shows that after SC the transformation strain and all the properties reach their values in phase 2. Recall that the indices 1 and 2 denote the values before and after SC. The solid horizontal lines inside the rectangle *ABCD* represent the homogeneous variation of the transformation strain and all the properties in the nucleus during the SC. A model of this kind is used for example in atomistic calculations and allows us to determine the variation of material properties and stresses during the SC (Fig. 4) (e.g. Krasko and Olson, 1989). In our approach we assume that we know all the material properties not only for phases 1 and 2, but for each intermediate state of the SC. The dotted lines correspond to the nonhomogeneous variation of transformation strain and properties during the SC. In the framework of Landau–Ginzburg type theories of PT, a nonhomogeneous way of nucleus formation is called a nonclassical transformation path (Olson and Cohen, 1986). The

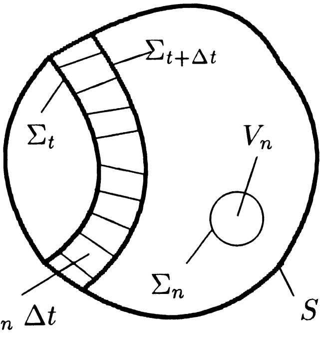


Fig. 1. Volume with SC.

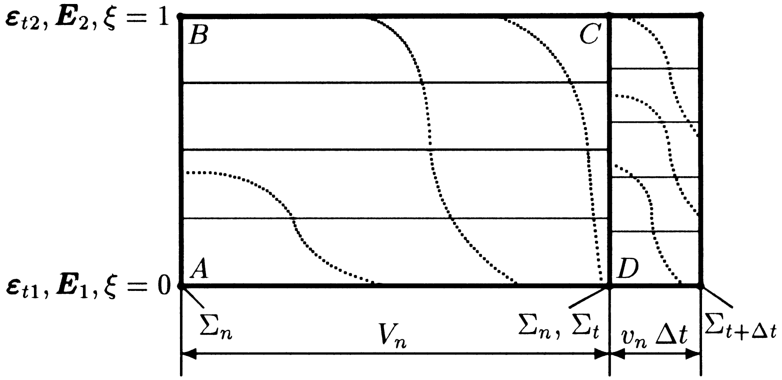


Fig. 2. Scheme of the transformation process in a one-dimensional nucleus V_n and in volume $v_n\Delta t$ covered by a moving interface.

problem of how to choose the actual variation of all the properties and fields during the SC will be considered below.

For simplicity we will consider a small strain approximation. The finite strain kinematics and the corresponding thermodynamic treatment can be taken without major changes from the paper by Levitas (1998a). We will use an energy balance equation (the first law of thermodynamics) and the entropy balance equation combined with the Clausius–Duhem inequality (the second law of thermodynamics) for the whole volume V:

$$\int_S (\mathbf{p}\cdot\mathbf{v}-\mathbf{h}\cdot\mathbf{n})\,\mathrm{d}S-\frac{\mathrm{d}}{\mathrm{d}t}\int_V \rho(U+0.5\mathbf{v}\cdot\mathbf{v})\,\mathrm{d}V-\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Sigma_\mathrm{n}} \Gamma\,\mathrm{d}\Sigma_\mathrm{n}+\int_V \rho(\mathbf{f}\cdot\mathbf{v}+r)\,\mathrm{d}V=0, \quad (1)$$

$$S_i:=\frac{\mathrm{d}}{\mathrm{d}t}\int_V \rho s\,\mathrm{d}V-\int_V \rho\frac{r}{\theta}\,\mathrm{d}V+\int_S \frac{\mathbf{h}}{\theta}\cdot\mathbf{n}\,\mathrm{d}S\geqslant 0. \quad (2)$$

Here $\mathbf{v}$ is the velocity, $\mathbf{h}$ the heat flux, $\mathbf{n}$ is the unit normal to Σ_n, t is the time, U is the specific (per unit mass) internal energy, ρ is the mass density, $\mathbf{f}$ is the body force per unit mass, s is the specific entropy, S_i is the total entropy production, r is the specific volumetric heat supply, $\theta\geqslant 0$ is the temperature and Γ is the surface energy per unit area after the SC.

As outside the nucleus V_n no SC occur and local constitutive equations are valid, both these equations can be rewritten for an arbitrary volume and for each material point without SC. In particular, we can choose a volume $V-V_n$ bounded by a S and *fixed surface* Σ_n (or by a surface lying outside the nucleus infinitesimally close to Σ_n, in order to exclude the term with the surface energy). Then extracting from Eqs. (1) and (2) similar equations for the volume $V-V_n$, we obtain the result that the integral form of the first and second laws of thermodynamics have to be used for the nucleus V_n only:

$$\int_{\Sigma_n}(\mathbf{p}\cdot\mathbf{v}-\mathbf{h}\cdot\mathbf{n})\mathrm{d}\Sigma_n-\frac{\mathrm{d}}{\mathrm{d}t}\int_{V_n}\rho(U+0.5\mathbf{v}\cdot\mathbf{v})\mathrm{d}V_n-\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Sigma_n}\Gamma\mathrm{d}\Sigma_n+\int_{Vn}\rho(\mathbf{f}\cdot\mathbf{v}+r)\mathrm{d}V_n=0, \tag{3}$$

$$S_i:=\frac{\mathrm{d}}{\mathrm{d}t}\int_{V_n}\rho s\mathrm{dV}_n-\int_{V_n}\rho\frac{\mathrm{r}}{\theta}\mathrm{d}V_n+\int_{\Sigma_n}\frac{\mathbf{h}}{\theta}\cdot\mathbf{n}\mathrm{d}\Sigma_n\geqslant 0. \tag{4}$$

Note that the surface energy varies during the transformation in some way from the initial value (in particular from 0) to the final value Γ_2. For a nonhomogeneous transformation process in the nucleus (Fig. 2), the interface between the nucleus and parent phase and consequently the surface energy are not well defined in the course of SC. The volume undergoing the SC can also be variable at the beginning of the transformation process in V_n. However, this does not affect the final result, which depends on the surface energy Γ_2 after SC related to the *final* nucleus surface Σ_n as well as on the *final* volume V_n (more precisely, on an arbitrary volume which includes V_n). That is why we consider nucleation as a process in a fixed volume V_n bounded by the fixed interface Σ_n which coincide with the volume and interface after finishing the SC. Note that the interface propagation can be considered in a similar framework as well (see Section 5). Using the Green–Gauss theorem, the relation between rate of displacements and deformations, as well as the impulse balance equation we obtain

$$\int_{V_n}\left(\boldsymbol{\sigma}:\dot{\boldsymbol{\varepsilon}}-\rho\dot{U}-\operatorname{div}\mathbf{h}+\rho r\right)\mathrm{d}V_n-\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Sigma_n}\Gamma\mathrm{d}\Sigma_n=0, \tag{5}$$

$$S_i:=\int_{V_n}\left(\rho\dot{s}-\rho\frac{r}{\theta}+\operatorname{div}\frac{\mathbf{h}}{\theta}\right)\mathrm{d}V_n\geqslant 0. \tag{6}$$

Here $\boldsymbol{\varepsilon}$ and $\boldsymbol{\sigma}$ are the strain and stress tensors. The energy balance Eq. (5) cannot be written for each material point due to the surface energy term, but the entropy inequality is valid for each point

$$\tilde{S}_i:=\rho\dot{s}-\rho\frac{r}{\theta}+\frac{1}{\theta}\operatorname{div}\mathbf{h}-\frac{\boldsymbol{\nabla}\theta}{\theta_2}\cdot\mathbf{h}\geqslant 0, \tag{7}$$

where $\tilde{S}_i$ is the local entropy production and transformation

$$\operatorname{div}\frac{\mathbf{h}}{\theta}=\frac{1}{\theta}\operatorname{div}\mathbf{h}+\mathbf{h}\cdot\boldsymbol{\nabla}\left(\frac{1}{\theta}\right)=\frac{1}{\theta}\operatorname{div}\mathbf{h}-\frac{1}{\theta^2}\mathbf{h}\cdot\boldsymbol{\nabla}\theta \tag{8}$$

was used. Excluding the expression div $\mathbf{h}-\rho r$ from Eq. (7) and substituting it in Eq. (5) we obtain after simple transformations an inequality

$$\bar{\mathcal{D}}:=\int_{V_n}\rho\theta\tilde{S}_i\mathrm{d}V_n=\int_{V_n}\left(\boldsymbol{\sigma}:\dot{\boldsymbol{\varepsilon}}-\rho\dot{U}+\rho\theta\dot{s}-\frac{\boldsymbol{\nabla}\theta}{\theta}\cdot\mathbf{h}\right)\mathrm{d}V_n-\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Sigma_n}\Gamma\mathrm{d}\Sigma_n\geqslant 0, \tag{9}$$

or in more compact form

$$\bar{\mathcal{D}} := \int_{V_n} \rho\theta\tilde{S}_i \mathrm{d}V_n = \int_{V_n} \rho\mathcal{D}\mathrm{d}V_n - \dot{\bar{\Gamma}} \geqslant 0, \tag{10}$$

where

$$\rho\mathcal{D} := \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} - \rho\dot{U} + \rho\theta\dot{s} - \frac{\boldsymbol{\nabla}\theta}{\theta}\cdot\mathbf{h} \quad \text{and} \quad \bar{\Gamma} := \int_{\Sigma_n} \Gamma\mathrm{d}\Sigma_n \tag{11}$$

is the local dissipation rate per unit volume and the total surface energy, respectively, $\bar{\mathcal{D}}$ is the total dissipation rate. In the case with neglected surface energy and a local system the global dissipation inequality (10) is equivalent to the local inequality $\mathcal{D}\geqslant 0$ along with inequality $\tilde{S}_i \geqslant 0$. In the case with $\dot{\bar{\Gamma}} \neq 0$ again $\tilde{S}_i \geqslant 0$, but the inequality $\mathcal{D}\geqslant 0$ cannot be proved.

For a local system without the surface energy the usual procedure in irreversible thermodynamics consists of the following. Dissipation rate $\mathcal{D}\geqslant 0$ or local entropy production $\tilde{S}_i = \frac{\mathcal{D}}{\theta}$ is represented as a product of generalized forces and rates. They have to be interrelated by a kinetic equation, otherwise inequality $\mathcal{D}\geqslant 0$ cannot be satisfied. Both these representations (in terms of dissipation rate $\mathcal{D}$ or local entropy production $\frac{\mathcal{D}}{\theta}$) are equivalent. In the case with the surface energy and consideration of finite volume, according to Eq. (10) the dissipation rate $\bar{\mathcal{D}}$ can only be presented as a product of generalized forces and rates (see below), for $\int_{V_n} \rho\tilde{S}_i dV_n$ it is impossible due to the surface integral. Introducing the specific Helmholtz free energy $\psi = U - \theta s$ we obtain

$$\rho\mathcal{D} = \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} - \rho\dot{\psi} - \rho s\dot{\theta} - \frac{\boldsymbol{\nabla}\theta}{\theta}\cdot\mathbf{h}. \tag{12}$$

Eq. (12) is more convenient for applications than Eq. (11), when the temperature and not the entropy is used as independent variable.

We assume an additive decomposition of a total strain tensor $\boldsymbol{\varepsilon}$ into elastic $\boldsymbol{\varepsilon}_e$, plastic $\boldsymbol{\varepsilon}_p$, thermal $\boldsymbol{\varepsilon}_\theta$ and transformational $\boldsymbol{\varepsilon}_t$ parts, i.e.

$$\boldsymbol{\varepsilon} = \boldsymbol{\varepsilon}_e + \boldsymbol{\varepsilon}_p + \boldsymbol{\varepsilon}_\theta + \boldsymbol{\varepsilon}_t. \tag{13}$$

We introduce the internal dimensionless time (order parameter) ξ $(0\leqslant\xi\leqslant 1)$ which is related to $\boldsymbol{\varepsilon}_t$ and all the variable material properties $\mathbf{A}_k, k = 1, 2, \ldots, n$ and has the following features: SC starts at $\xi = 0$ and finishes at $\xi = 1$; when ξ varies between 0 and 1 the transformation strain grows from $\boldsymbol{\varepsilon}_{t1}$ to $\boldsymbol{\varepsilon}_{t2}$, and all the thermodynamic properties of phase 1 $\mathbf{A}_{k1}, k = 1, 2, \ldots, n$, change into the properties of phase $\mathbf{A}_{k2}$. It is possible to define

$$\xi := \frac{|\boldsymbol{\varepsilon}_t|}{|\boldsymbol{\varepsilon}_{t2}|}, \quad \xi := \frac{tr\boldsymbol{\varepsilon}_t}{tr\boldsymbol{\varepsilon}_{t2}} \quad \text{or} \quad \xi := \frac{|\mathbf{A}_m - \mathbf{A}_{m1}|_m}{|\mathbf{A}_{m2} - \mathbf{A}_{m1}|_m} \tag{14}$$

for some m (e.g. for the tensile modulus or the shear modulus), where $|\ldots|_m$ is some norm, and to express all the components of tensor $\boldsymbol{\varepsilon}_t$ and materials properties $\mathbf{A}_k$ in terms of ξ in a such way that $\boldsymbol{\varepsilon}_t(0) = \boldsymbol{\varepsilon}_{t1}$, $\boldsymbol{\varepsilon}_t(1) = \boldsymbol{\varepsilon}_{t2}$, $\mathbf{A}_k(0) = \mathbf{A}_{k1}$ and $\mathbf{A}_k(1) = \mathbf{A}_{k2}$.This guarantees that ξ does not reach 1 before $\boldsymbol{\varepsilon}_t$ reaches $\boldsymbol{\varepsilon}_{t2}$ and all the properties do not reach the properties of phase 2. Definition $(14)_1$ is valid in the general case with variable transformation strain, Eq. $(14)_2$ can be applied for nonzero dilatation due to SC and, for SC with variable properties, definition $(14)_3$ is noncontradictory. Many other definitions are acceptable.

The introduction of ξ is not necessary (one can vary the transformation strain tensor and all the material properties directly); however, such a scalar measure of a transformation process simplifies some manipulations, interpretations and formulations of simplified models (see Part II of the paper). Conventionally we can consider the material at each point as a "mixture" of phase 1 with a fraction $(1 - \xi)$ and phase 2 with a fraction ξ. Then the internal time ξ formally plays a similar role in the consideration of the transformation process at a material point as a volume fraction of new phase for the averaged description of SC. Let us define the specific (per unit mass) Helmholtz free energy

$$\psi = \psi\big(\boldsymbol{\varepsilon}_e, \theta, \boldsymbol{\varepsilon}_p, \mathbf{g}, \boldsymbol{\varepsilon}_t, \xi\big), \tag{15}$$

where $\mathbf{g}$ is a set of internal variables, e.g. internal stress tensor (back stress), dislocations or point defect density. In small strain approximation $\rho = \text{const}$. The substitution of the rate of free energy

$$\dot{\psi} = \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}_e} : \dot{\boldsymbol{\varepsilon}}_e + \frac{\partial \psi}{\partial \theta}\dot{\theta} + \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}_p} : \dot{\boldsymbol{\varepsilon}}_p + \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}_t} : \dot{\boldsymbol{\varepsilon}}_t + \frac{\partial \psi}{\partial \mathbf{g}^t} : \dot{\mathbf{g}} + \frac{\partial \psi}{\partial \xi}\dot{\xi} \tag{16}$$

with taking into account the additive decomposition of $\boldsymbol{\varepsilon}$, $\boldsymbol{\varepsilon}_\theta = \boldsymbol{\varepsilon}_\theta(\theta, \xi)$ and

$$\dot{\boldsymbol{\varepsilon}}_\theta = \frac{\partial \boldsymbol{\varepsilon}_\theta}{\partial \theta}\dot{\theta} + \frac{\partial \boldsymbol{\varepsilon}_\theta}{\partial \xi}\dot{\xi} \tag{17}$$

in Eq. (12) yields

$$\begin{aligned} \rho\mathcal{D} = & \left(\boldsymbol{\sigma} - \rho\frac{\partial \psi}{\partial \boldsymbol{\varepsilon}_e}\right) : \dot{\boldsymbol{\varepsilon}}_e - \rho\left(s + \frac{\partial \psi}{\partial \theta} - \frac{\boldsymbol{\sigma}}{\rho} : \frac{\partial \boldsymbol{\varepsilon}_\theta}{\partial \theta}\right)\dot{\theta} + \left(\boldsymbol{\sigma} - \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}_p}\right) : \dot{\boldsymbol{\varepsilon}}_p \\ & - \rho\frac{\partial \psi}{\partial \mathbf{g}^t} : \dot{\mathbf{g}} - \frac{\boldsymbol{\nabla}\theta}{\theta}\cdot\mathbf{h} + \left(\boldsymbol{\sigma} : \frac{\partial(\boldsymbol{\varepsilon}_t + \boldsymbol{\varepsilon}_\theta)}{\partial \xi} - \rho\frac{\partial \psi}{\partial \boldsymbol{\varepsilon}_t} : \frac{\partial \boldsymbol{\varepsilon}_t}{\partial \xi} - \rho\frac{\partial \psi}{\partial \xi}\right)\dot{\xi}. \end{aligned} \tag{18}$$

The assumption that the global rate of dissipation in Eq. (10) is independent of $\dot{\boldsymbol{\varepsilon}}_e$ and $\dot{\theta}$ results in the hyperelasticity law and expression for entropy

$$\boldsymbol{\sigma} = \rho\frac{\partial \psi}{\partial \boldsymbol{\varepsilon}_e}; \qquad s = \frac{\boldsymbol{\sigma}}{\rho} : \frac{\partial \boldsymbol{\varepsilon}_\theta}{\partial \theta} - \frac{\partial \psi}{\partial \theta}, \tag{19}$$

as well as in the representation of the local dissipation rate as a product of generalized forces and rates:

$$\mathcal{D} = \mathbf{X}_p : \dot{\boldsymbol{\varepsilon}}_p + \mathbf{X}_g : \dot{\mathbf{g}} + \mathbf{X}_h \cdot \mathbf{h} + X_\xi \dot{\xi}, \tag{20}$$

where

$$\mathbf{X}_p = \frac{1}{\rho}\boldsymbol{\sigma} - \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}_p}; \quad \mathbf{X}_g = -\frac{\partial \psi}{\partial \mathbf{g}^t}; \quad \mathbf{X}_h = -\frac{1}{\rho}\frac{\boldsymbol{\nabla}\theta}{\theta} \tag{21}$$

and

$$\mathbf{X}_\xi = \frac{1}{\rho}\boldsymbol{\sigma} : \frac{\partial(\boldsymbol{\varepsilon}_t + \boldsymbol{\varepsilon}_\theta)}{\partial \xi} - \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}_t} : \frac{\partial \boldsymbol{\varepsilon}_t}{\partial \xi} - \frac{\partial \psi}{\partial \xi} \tag{22}$$

are the dissipative forces conjugated to dissipative rates $\dot{\boldsymbol{\varepsilon}}_p, \dot{\mathbf{g}}, \mathbf{h}$ and $\dot{\xi}$ respectively. For the global dissipation rate we have

$$\bar{\mathcal{D}} = \int_{V_n} \rho\Big(\mathbf{X}_p : \dot{\boldsymbol{\varepsilon}}_p + \mathbf{X}_g : \dot{\mathbf{g}} + \mathbf{X}_h \cdot \mathbf{h} + X_\xi \dot{\xi}\Big) \mathrm{d}V_n - \dot{\bar{\Gamma}} \geqslant 0. \tag{23}$$

The simplest assumption that each rate $\dot{\boldsymbol{\varepsilon}}_p, \dot{\mathbf{g}}$ and $\mathbf{h}$ depends locally on the conjugate force at the same material point only leads to evolution equations

$$\dot{\boldsymbol{\varepsilon}}_p = \mathbf{f}_p\big(\mathbf{X}_p, \xi\big); \quad \dot{\mathbf{g}} = \mathbf{f}_g\big(\mathbf{X}_g, \xi\big); \quad \mathbf{h} = \mathbf{f}_h(\mathbf{X}_h, \xi), \tag{24}$$

as well as to inequalities

$$\mathbf{X}_p : \dot{\boldsymbol{\varepsilon}}_p \geqslant 0, \quad \mathbf{X}_g : \dot{\mathbf{g}} \geqslant 0, \quad \mathbf{X}_h \cdot \mathbf{h} \geqslant 0. \tag{25}$$

Eq. $(24)_1$ is the flow rule, Eq. $(24)_2$ is the evolution equation for the internal variables, Eq. $(24)_3$ is the generalized Fourier law. All functions in Eq. (24) can depend on temperature, plastic strain, internal variables and so on.

3.2. Description of structural changes

The assumption that SC is thermodynamically independent of all other thermodynamic processes, i.e. $\dot{\xi}$ is independent of $\mathbf{X}_p, \mathbf{X}_g$ and $\mathbf{X}_h$, leads to the inequality

$$\mathcal{D}_\xi := \int_{V_n} \rho X_\xi \dot{\xi} \mathrm{d}V_n - \dot{\bar{\Gamma}} \geqslant 0. \tag{26}$$

The above assumptions do not exclude a mutual influence of all thermomechanical processes through the stress and temperature fields. At $\dot{\bar{\Gamma}} \neq 0$ inequality (26) cannot be localized; consequently SC have to be described by nonlocal equations.

Now let us show that even at $\bar{\Gamma} = 0$ we cannot prescribe for $\dot{\xi}$ a constitutive equation of type $\dot{\xi} = f_\xi\big(X_\xi, \xi\big)$ or any other constitutive equation. We assume that the condition $\dot{\xi} = 0$ can be satisfied with some external parameters, for example at $X_\xi = 0$. This is one scalar equation and it is always possible for each ξ and $\theta(\xi)$ to choose six components of stress tensor $\boldsymbol{\sigma}(\xi, \theta(\xi))$ to satisfy it. If the actual stress variation follows this dependence, then the phase equilibrium is possible for arbitrary ξ. According to the definition of SC without stable intermediate state accepted in Section 2, the phase equilibrium is impossible at $0 < \xi < 1$; only at $\xi = 0$ and $\xi = 1$ do we have the stable equilibrium. At $0 < \xi < 1$ a nonequilibrium process takes place, which requires energy and stress fluctuations.

Such a contradiction in the application of continuum thermodynamics was revealed for the first time for PT in (Levitas, 1996, 1998a). In this case a standard thermodynamic approach cannot be applied. It is necessary to average the thermodynamic parameters, related to SC, over the duration of SC t_s in order to filter off these fluctuations. We introduce the averaged dissipation rate due to SC

$$\begin{aligned} &\bar{\mathcal{D}}_\xi := \frac{1}{t_s}\int_0^{t_s} \mathcal{D}_\xi \mathrm{d}t = \frac{1}{t_s}\left(\int_{V_n}\int_0^{t_s} \rho X_\xi \dot{\xi} \mathrm{d}t \mathrm{d}V_n - \int_0^{t_s} \dot{\bar{\Gamma}} \mathrm{d}t\right) = \frac{X_v}{t_s} = X_v \dot{\chi}, \\ &\text{where} \quad X_v := \int_{V_n}\int_0^1 \rho X_\xi \mathrm{d}\xi \mathrm{d}V_n - \Delta\bar{\Gamma}, \quad \dot{\chi} := \frac{1}{t_s} \end{aligned} \tag{27}$$

are the averaged dissipative force and rate, $\Delta\bar{\Gamma} := \bar{\Gamma}_2 - \bar{\Gamma}_1$. The definition of dissipative rate is logical, because a variation of the parameter ξ during the time t_s is one and $1/t_s$ is in fact the mean rate of transformation. The dissipative force is defined as a conjugate variable in the expression for the rate of dissipation.

For the reverse SC the parameter ξ varies from 1 to 0 and

$$\begin{aligned} \bar{\mathcal{D}}_\xi &= \frac{1}{t_s}\left(\int_0^{t_s} \rho X_\xi \dot{\xi} \mathrm{d}t + \Delta\bar{\Gamma}\right) = \frac{1}{t_s}\left(\int_1^0 \rho X_\xi \mathrm{d}\xi + \Delta\bar{\Gamma}\right) \\ &= -\frac{1}{t_s}\left(\int_{V_n}\int_0^1 \rho X_\xi \mathrm{d}\xi \mathrm{d}V_n - \Delta\Gamma\right) = X_v \dot{\chi}, \\ \text{where } X_v &:= \int_{V_n}\int_0^1 \rho X_\xi \mathrm{d}\xi \mathrm{d}V_n - \Delta\bar{\Gamma} \leqslant 0, \quad \dot{\chi} = -\frac{1}{t_s} \leqslant 0. \end{aligned} \tag{28}$$

For direct or reverse SC without dissipation due to SC, the condition $X_v = 0$ is valid per definition. Such an SC can be called the SC which occurs at macroscopic equilibrium. Utilizing Eqs. (22) and (28), we obtain an explicit form of the SC criterion for SC without dissipation:

$$X_v = \int_{V_n}\int_{\boldsymbol{\varepsilon}_{t1}}^{\boldsymbol{\varepsilon}_{t2}} \rho\left(\frac{\boldsymbol{\sigma}}{\rho} - \frac{\partial\psi}{\partial\boldsymbol{\varepsilon}_t}\right) : \mathrm{d}\boldsymbol{\varepsilon}_t \mathrm{d}V_n + \int_{V_n}\int_0^1 \rho\left(\frac{\boldsymbol{\sigma}}{\rho} : \frac{\partial\boldsymbol{\varepsilon}_\theta}{\partial\xi} - \frac{\partial\psi}{\partial\xi}\right) \mathrm{d}\xi \mathrm{d}V_n - \Delta\bar{\Gamma} = 0. \tag{29}$$

We will call X_v the driving force of SC. Alternatively we can define

$$\bar{\mathcal{D}}_\xi := \left(\frac{1}{m_n} X_v\right)\frac{m_n}{t_s} = \bar{X}\dot{m}_n; \quad \text{where} \quad \bar{X} := \frac{1}{m_n} X_v \quad \text{and} \quad \dot{m}_n := \frac{m_n}{t_s} \tag{30}$$

is the dissipative force and rate. For the reverse SC $\dot{m}_n := -\frac{m_n}{t_s}$. Dissipative force and rates of this kind were used by Levitas (1998b, 1999). Generally, the results are dependent on the choice of the force and rate. We will show the advantage of the use of X_v and $\dot{\chi}$. Let us consider the alternative expressions for X_v. From Eqs. (12) and (20) it follows that

$$X_\xi \dot{\xi} = \mathcal{D} - \mathbf{X}_p : \dot{\boldsymbol{\varepsilon}}_p - \mathbf{X}_g : \dot{\mathbf{g}} - \mathbf{X}_h \cdot \mathbf{h} = \rho^{-1} \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} - \dot{\psi} - s\dot{\theta} - \mathbf{X}_p : \dot{\boldsymbol{\varepsilon}}_p - \mathbf{X}_g : \dot{\mathbf{g}}; \qquad (31)$$

$$X_v = \int_{V_n} \left(\int_{\boldsymbol{\varepsilon}_1}^{\boldsymbol{\varepsilon}_2} \boldsymbol{\sigma} : \mathrm{d}\boldsymbol{\varepsilon} - \rho(\psi_2 - \psi_1) - \int_{\theta_1}^{\theta_2} \rho s \mathrm{d}\theta - \int_0^{t_s} \rho\big(\mathbf{X}_p : \dot{\boldsymbol{\varepsilon}}_p + \mathbf{X}_g : \dot{\mathbf{g}}\big) \mathrm{d}t \right) \mathrm{d}V_n - \Delta\bar{\Gamma}. \qquad (32)$$

Taking into account the decomposition (13) and Eq. $(21)_1$ for $\mathbf{X}_p$ we obtain

$$\begin{aligned} X_v = \int_{V_n} \Bigg(& \int_{\boldsymbol{\varepsilon}_1}^{\boldsymbol{\varepsilon}_2} \boldsymbol{\sigma} : \mathrm{d}(\boldsymbol{\varepsilon}_e + \boldsymbol{\varepsilon}_\theta + \boldsymbol{\varepsilon}_t) - \rho(\psi_2 - \psi_1) - \int_{\theta_1}^{\theta_2} \rho s \mathrm{d}\theta \\ & + \int_{\mathbf{g}_1}^{\mathbf{g}_2} \rho \frac{\partial \psi}{\partial \mathbf{g}^t} : \mathrm{d}\mathbf{g} + \int_{\boldsymbol{\varepsilon}_{p1}}^{\boldsymbol{\varepsilon}_{p2}} \rho \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}_p} : \mathrm{d}\boldsymbol{\varepsilon}_p \Bigg) \mathrm{d}V_n - \Delta\bar{\Gamma}. \end{aligned} \qquad (33)$$

The expressions (32) and (33) are sometimes more convenient for the analysis than Eq. (29). At $\frac{\partial \psi}{\partial \boldsymbol{\varepsilon}_p} = 0$ Eq. (33) is more suitable for application because in this case the plastic strain $\boldsymbol{\varepsilon}_p$ disappears completely. Eq. (32) is more appropriate when the moving interface is considered and the Hadamard compatibility condition for $\boldsymbol{\varepsilon}_2 - \boldsymbol{\varepsilon}_1$ is taken into account (see Section 5).

Note that expressions $(21)_1$ and $(22)_2$ for $\mathbf{X}_p$ and $\mathbf{X}_g$ can be obtained using the standard thermomechanical approach for materials without SC. Then a very simple method of derivation of expression for X_v is evident [see Eq.(32)]: *the dissipation increment due to SC only is the difference between the total dissipation increment and the dissipation increment due to other dissipative processes* (plastic flow, variation of internal variables and so on).

It is easy to show that an *integral in Eq. (32) can be taken over an arbitrary volume v containing one transforming region V_n with SC*, i.e.

$$X_v = \int_{v} \left(\int_{\boldsymbol{\varepsilon}_1}^{\boldsymbol{\varepsilon}_2} \boldsymbol{\sigma} : \mathrm{d}\boldsymbol{\varepsilon} - \rho(\psi_2 - \psi_1) - \int_{\theta_1}^{\theta_2} \rho s \mathrm{d}\theta - \int_0^{t_s} \rho\big(\mathbf{X}_p : \dot{\boldsymbol{\varepsilon}}_p + \mathbf{X}_g : \dot{\mathbf{g}}\big) \mathrm{d}t \right) \mathrm{d}v - \Delta\bar{\Gamma}. \qquad (34)$$

Integration over the region $v - V_n$ without SC gives zero contribution to the driving force X_v, because an integrand is the dissipation rate due to SC only [see Eqs. (31) and (32)] which is zero due to the lack of SC in the region $v - V_n$. The change in surface energy has to be taken into account over the surface where this change in fact occurs, i.e. over nucleus surface Σ_n. Using the Green–Gauss theorem and the impulse balance equation, stress work can be transformed into a surface integral over the surface bounding v, e.g. the external surface S, and change in kinetic energy and work of mass force into the volume v [see Eq.(40)]. Consequently, *instead of a surface-independent Eshelby integral in the theory of defects and path-independent J-integral in fracture mechanics for elastic materials, we introduced a region-independent integral for arbitrary inelastic materials.* In contrast to the Γ-integral (Cherepanov, 1967, 1979) for a singular defect and a region-dependent T-integral (Atluri, 1997), which are formulated for inelastic materials as well, the region-independent integral (34) separates the dissipation increment due to SC only from other dissipation contributions and is the thermodynamic driving force for SC. It consistently takes into account all types of dissipation and temperature variation in the process of SC. Let

$$\rho\psi_i = 0.5\boldsymbol{\varepsilon}_{ei} : \mathbf{E}_i : \boldsymbol{\varepsilon}_{ei} + \rho\psi_i^{\theta} = \rho\psi_i^{e} + \rho\psi_i^{\theta}, \quad i = 1, 2, \tag{35}$$

where ψ_i^e and ψ_i^θ are the elastic and thermal (independent of elastic strain) parts of free energy, $\mathbf{E}_i$ the elasticity tensors. Since

$$\begin{aligned}\int_{\boldsymbol{\varepsilon}_{e1}}^{\boldsymbol{\varepsilon}_{e2}} \boldsymbol{\sigma} : \mathrm{d}\boldsymbol{\varepsilon}_e &= \int_{\boldsymbol{\varepsilon}_{e1}}^{\boldsymbol{\varepsilon}_{e2}} \boldsymbol{\varepsilon}_e : \mathbf{E} : \mathrm{d}\boldsymbol{\varepsilon}_e = 0.5\int_{\boldsymbol{\varepsilon}_{e1}}^{\boldsymbol{\varepsilon}_{e2}} \mathrm{d}(\boldsymbol{\varepsilon}_e : \mathbf{E} : \boldsymbol{\varepsilon}_e) - 0.5\int_{\mathbf{E}_1}^{\mathbf{E}_2} \boldsymbol{\varepsilon}_e : \mathrm{d}\mathbf{E} : \boldsymbol{\varepsilon}_e \\ &= \rho\left(\psi_2^e - \psi_1^e\right) - 0.5\int_{\mathbf{E}_1}^{\mathbf{E}_2} \boldsymbol{\varepsilon}_e : \mathrm{d}\mathbf{E} : \boldsymbol{\varepsilon}_e,\end{aligned} \tag{36}$$

it follows then from Eq. (33)

$$\begin{aligned}X_v = \int_{V_n}\Bigg(&\int_{\boldsymbol{\varepsilon}_1}^{\boldsymbol{\varepsilon}_2} \boldsymbol{\sigma} : \mathrm{d}(\boldsymbol{\varepsilon}_\theta + \boldsymbol{\varepsilon}_t) - 0.5\int_{\mathbf{E}_1}^{\mathbf{E}_2} \boldsymbol{\varepsilon}_e : \mathrm{d}\mathbf{E} : \boldsymbol{\varepsilon}_e - \rho\left(\psi_2^\theta - \psi_1^\theta\right) - \int_{\theta_1}^{\theta_2} \rho s \mathrm{d}\theta \\ &+ \int_{\mathbf{g}_1}^{\mathbf{g}_2} \rho \frac{\partial \psi}{\partial \mathbf{g}^t} : \mathrm{d}\mathbf{g} + \int_{\boldsymbol{\varepsilon}_{p1}}^{\boldsymbol{\varepsilon}_{p2}} \rho \frac{\partial \psi}{\partial \boldsymbol{\varepsilon}_p} : \mathrm{d}\boldsymbol{\varepsilon}_p\Bigg)\mathrm{d}V_n - \Delta\bar{\Gamma}.\end{aligned} \tag{37}$$

When $\mathbf{E}_1 = \mathbf{E}_2 = \mathbf{const}$ (e.g. independent of θ, $\boldsymbol{\varepsilon}_p, \ldots$), then the term with $\mathrm{d}\mathbf{E}$ disappear.

4. Kinetics. Thermodynamic criterion of structural changes

As is usual in irreversible thermodynamics, the kinetic equation between force and rate $\dot{\chi} = f(X_v, \ldots)$ or $X_v = K_v(\dot{\chi}, \ldots)$ has to be given. Functions f and K_v can depend on θ, $\boldsymbol{\varepsilon}_p$, $\mathbf{g}$, $V_n, \ldots$ and should be determined experimentally. Sometimes it is convenient to present $K = \frac{K_v}{\mathrm{m}_n} = \frac{1}{m_n}\int_{V_n} \rho\tilde{K}\mathrm{d}V_n$ where $\tilde{K}$ is a locally determined function and express the kinetic equation in the form $\dot{\chi} = f(\bar{X}m_n, \ldots)$ or $\bar{X} = K(\dot{\chi}, \ldots)$. For homogeneously distributed $\tilde{K}$ we have $K = \tilde{K}$. There are only two possibilities for the behaviour of the function K as $\dot{\chi} \to 0$.

1. Let the condition $\dot{\chi} = 0$ be valid at $\bar{X} = 0$ only. Then the macroscopically equilibrium SC ($\dot{\chi} \to 0$) can be described with the equation $\bar{X} = 0$ (Fig. 3, curve 1).

2. Let

$$\begin{aligned}&\text{at} \quad \dot{\chi} = 0 \quad K^0_{2\to 1} < \bar{X} < K^0_{1\to 2},\\ &\text{at} \quad \dot{\chi} > 0 \quad \bar{X} = K_{1\to 2}(\dot{\chi}, \ldots) > 0,\\ &\text{at} \quad \dot{\chi} < 0 \quad \bar{X} = K_{2\to 1}(\dot{\chi}, \ldots) < 0,\end{aligned} \tag{38}$$

where $K^0_{1\to 2}$ and $K^0_{2\to 1}$ are the threshold values of $\bar{X}$ at the direct and the reverse SC (Fig. 3, curves 2 and 3), which can depend on temperature, internal variables, plastic strain, their history and so on. In this case there is *thermodynamic hysteresis*, i.e. direct and reverse SC begin at different values $\bar{X}$ and between these values SC is impossible. In the first case there is no hysteresis. For most martensitic transformations the values $K^0_{1\to 2}$ and $\left|K^0_{2\to 1}\right|$ are very high. The values $K^0_{1\to 2}$ and $K^0_{2\to 1}$ characterize an athermal component of the driving force for the direct and reverse SC respectively. We assume here finite K^0 as well. The case $K^0 = 0$ will be considered as a particular case of Eq. (38). Conditions

$$\bar{X} \geqslant K^0_{1\to 2} \quad \text{and} \quad \bar{X} \leqslant K^0_{2\to 1} \tag{39}$$

are the thermodynamic criteria of direct and reverse SC.

Eq. (38) means that the calculated value of the dissipation increment due to SC reaches its experimentally determined value. We assume that when $\bar{X}$ reaches K^0 the rate $\dot{\chi}$ has a jump to some finite value (Fig. 3, curve 2′). We will distinguish two types of SC kinetics:

- athermal kinetics, for which K is independent of the rate $\dot{\chi}$ (or SC duration t_s). In this case real time and rate do not play a part, SC occurs instantaneously when $\bar{X}$ reaches K (Fig. 3, curve 3). Relation $K(\dot{\chi})$ for athermal kinetics is

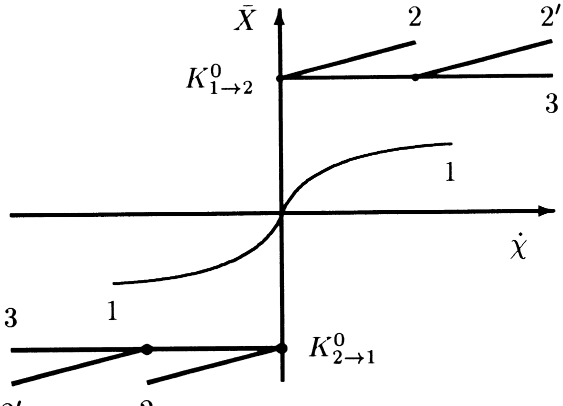


Fig. 3. Possible dependencies between $\bar{X}$ and $\dot{\chi}$.

similar to the relation between yield stress and the plastic deformation rate for rate independent plastic materials. As will be demonstrated below, such an idealization can lead to a contradiction for some SC.

- Time-dependent (true) kinetics, for which $\bar{X}$ is a function of rate $\dot{\chi}$ and this function has to be specified.

Substituting Eq. (32) for X_v in Eq. (38), we obtain for the nucleation condition during time t_s

$$\begin{aligned}&\int_{V_n}\int_{\boldsymbol{\varepsilon}_1}^{\boldsymbol{\varepsilon}_2}\boldsymbol{\sigma}:\mathrm{d}\boldsymbol{\varepsilon}\mathrm{d}V_n=\int_0^{t_s}\int_{\Sigma_n}\mathbf{p}\cdot\mathbf{v}\mathrm{d}\Sigma_n\mathrm{d}t+\int_0^{t_s}\int_{V_n}\rho\mathbf{f}\cdot\mathbf{v}\mathrm{d}V_n\mathrm{d}t\\&-\int_{V_n}0.5\rho\Delta|\mathbf{v}|^2\mathrm{d}V_n=\Delta\bar{\Gamma}+\int_{V_n}\rho\Delta\psi\mathrm{d}V_n\\&+\int_{V_n}\int_{\theta_1}^{\theta_2}\rho s\mathrm{d}\theta\mathrm{d}V_n+\int_{V_n}\int_t^{t+\Delta t}\rho\big(\mathbf{X}_p:\dot{\boldsymbol{\varepsilon}}_p+\mathbf{X}_g:\dot{\mathbf{g}}\big)\mathrm{d}t\mathrm{d}V_n+\int_{V_n}\rho\tilde{\mathrm{K}}(\dot{\chi})\mathrm{d}V_n,\end{aligned}\qquad(40)$$

note that the Green–Gauss theorem and energy balance were used.

For elastic materials without internal variables and for isothermal processes, at $K=\bar{\Gamma}=0$ it follows from Eq. (40) that

$$\int_{V_n}\left(\int_{\boldsymbol{\varepsilon}_1}^{\boldsymbol{\varepsilon}_2}\boldsymbol{\sigma}:\mathrm{d}\boldsymbol{\varepsilon}-\rho(\psi_2-\psi_1)\right)\mathrm{d}V_n=0.\qquad(41)$$

If SC proceeds at constant stress $\boldsymbol{\sigma}^{\mathrm{o}}$, which is homogeneously distributed in V_n, then $\boldsymbol{\sigma}^{\mathrm{o}}:(\boldsymbol{\varepsilon}_2-\boldsymbol{\varepsilon}_1)=\rho(\psi_2-\psi_1)$. This is the Maxwell rule. Geometrically speaking, it means that the areas (BCD) and (DEF) are the same, where OBCDEF is the shear stress τ-shear strain γ diagram (Fig. 4). Indeed, terms $\tau_{\mathrm{o}}(\gamma_2-\gamma_1)$ and $\rho(\psi_2-\psi_1)$

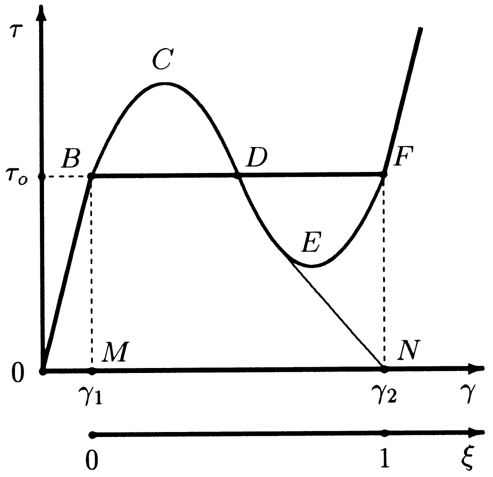


Fig. 4. Diagram of simple shearing for material with PT and fracture.

correspond to areas (BFNM) and (BCDEFNM) respectively, and their equality results in the equality of areas (BCD) and (DEF) Consequently Eq.(40) is a generalization of the Maxwell rule for dissipative materials, nonisothermal processes and for variable stresses.

Fig. 4 shows why it is impossible to apply the second law of thermodynamics for each ξ. To deform a material point in accordance with the curve BCDEF under the fixed stress τ_o fluctuations are needed to overcome the energy barrier (BCD), which will be designated as an activation energy E_a. The fluctuations can be thermo-activated or due to the local stress concentration on various defects (dislocations, grain boundaries). The material borrows energy (BCD) from the system and at the next instant returns the same energy (DEF) to it. The rate of dissipation formally calculated for each ξ in the interval $0 < \xi < \xi_D$ is negative, because standard thermodynamics does not take into account the fluctuations and necessary activation energy. The activation energy plays an important part for kinetics of SC (see Sections 8 and 9).

The stress–strain diagram which has a branch decreasing to zero stress corresponds to the case with fracture rather than PT. The Maxwell stress in this case $\tau_o \rightarrow 0$, i.e. at any $\tau > 0$ the second (fractured) phase is stable, but due to high activation energy (OBCDN) the solid state can be metastable for a very long time.

Let us show that the formulation based on time-independent kinetics can be contradictory even for an elastic material. For time-independent kinetics $\bar{X}$ cannot exceed K^0 At the same time for many problems at prescribed stresses, after the appearance of the first nucleus (i.e. when condition $\bar{X} = K^0$ is fulfilled), the driving force $\bar{X}$ grows drastically at PT in the next small volume at fixed external stresses and temperature (see for example Levitas et al., 1998a). As this growth of K^0 is physically unreal, then $\bar{X}$ exceeds K^0, which contradicts time-independent kinetics. Such a situation is typical of martensitic PT. The characteristic size of a nucleus is usually one–two orders of magnitude smaller than the plate which forms from this nucleus due to very quick spontaneous interface propagation at fixed external stresses and temperature (i.e. at $\bar{X} > K^0$). The well-known autocatalytic effect, when the appearance of one martensitic plate leads to the spontaneous appearance of several other plates, is of the same nature.

A similar phenomenon occurs during interface propagation by martensitic PT or twinning. According to dislocation representations (Christian, 1965; Boiko et al., 1991), the interface does not move as a smooth surface, but first a small jog or transformation (twinning) dislocation appears ("nucleus") which then moves spontaneously along the interface making the PT. Approximate simulation of such a mechanism by FEM (Idesman et al., 1997) demonstrates that $\bar{X}$ grows during the jog propagation along the interface.

Another reason for the spontaneous increase of $\bar{X}$ is related to the change in surface energy. The maximum change in the interface surface occurs with the appearance of the first nucleus. The growth of this nucleus can be modeled by the appearance of another nucleus which contacts the first one, then the next nucleus and so on. As the change in surface for the subsequent nuclei decreases and for some geometries reaches zero, $\bar{X}$ grows. This phenomenon will be analyzed in detail in

Sections 8 and 9 (see also the example of a strain-induced chemical reaction in (Levitas et al., 1998b,c).

5. Interface propagation and nucleus nondissapearance conditions

It seems at first glance that it is impossible to apply the results obtained above to the interface propagation, because they are obtained for the volume with a fixed interface with respect to the material. We will show that a formulation can be found which makes this possible. Assume that the volume V_n is obtained by the *interface* Σ propagation with a normal velocity v_n in time Δt, i.e. it is bounded by surfaces Σ_t and $\Sigma_{t+\Delta t}$ at time t and $t+\Delta t$ respectively, as well as by two lateral infinitesimal surfaces with the heights $v_n \Delta t$ (Fig. 1). Let us fix the surface $\Sigma_{t+\Delta t}$ and consider the SC process in a fixed volume $\mathrm{d}V_n = v_n \Delta t \mathrm{d}\Sigma_t$ (Fig. 2). It appears that due to the infinitesimal height $v_n \Delta t$ the results are independent of details of the process inhomogeneity along the height. At the same time, due to the history–dependent character of constitutive equations for inelastic materials, we must in any case consider the transformation process in the transforming points.

Let us transform the stress work integral in Eq. (40). We neglect for simplicity surface and kinetic energy. At the moving coherent interface, the position vector and the traction vector $\mathbf{p}$ are continuous, $\mathbf{p}_2 = \mathbf{p}_1$ and due to Hadamard compatibility condition

$$[\mathbf{F}] = -[\mathbf{v}]\mathbf{n}/v_n, \quad \text{whence } [\mathbf{v}] = -[\mathbf{F}]\cdot\mathbf{n}v_n \quad \text{and} \quad [\mathbf{F}] = [\mathbf{F}]\cdot\mathbf{nn}, \tag{42}$$

where $\mathbf{F} = \mathbf{I} + \boldsymbol{\varepsilon} + \boldsymbol{\omega}$ is the deformation gradient, $\boldsymbol{\omega}$ is the rotation tensor. For the infinitesimal layer the same relations apply while neglecting the terms of order Δt, where $[\mathbf{a}]$ means the difference in values $\mathbf{a}$ at the surfaces Σ_t and $\Sigma_t + \Delta t$. Then neglecting all the terms of order Δt^2 we obtain

$$\begin{gathered}\int_t^{t+\Delta t}\int_{\Sigma_n} \mathbf{p}\cdot\mathbf{v}\mathrm{d}\Sigma\mathrm{d}t = \left(\int_{\Sigma_{t+\Delta t}} \mathbf{p}\cdot\mathbf{v}_1\mathrm{d}\Sigma_{t+\Delta t} - \int_{\Sigma_t} \mathbf{p}_2\cdot\mathbf{v}_2\mathrm{d}\Sigma_t\right)\Delta t = -\int_{\Sigma_t}[\mathbf{p}\cdot\mathbf{v}]\Delta t\mathrm{d}\Sigma_t, \\ -[\mathbf{p}\cdot\mathbf{v}] = -\mathbf{p}\cdot[\mathbf{v}] = \mathbf{n}\cdot\boldsymbol{\sigma}\cdot[\mathbf{F}]\cdot\mathbf{n}v_n = \boldsymbol{\sigma} : [\mathbf{F}]\cdot\mathbf{nn}v_n = \boldsymbol{\sigma} : [\boldsymbol{\varepsilon}]v_n,\end{gathered} \tag{43}$$

i.e. work-producing components of the stress tensor are fixed in the course of SC. Substitution of Eq. (43) into Eq. (40) with taking into account of $\mathrm{d}V_n = v_n\Delta t\mathrm{d}\Sigma_t$ yields

$$\begin{aligned}\int_{V_n} \mathbf{n}\cdot\boldsymbol{\sigma}\cdot[\mathbf{F}]\cdot\mathbf{n}\mathrm{d}V_n = \int_{V_n} \boldsymbol{\sigma} : [\boldsymbol{\varepsilon}]\mathrm{d}V_n = \int_{V_n} \rho[\psi]\mathrm{d}V_n + \int_{V_n}\int_{\theta_1}^{\theta_2} \rho s\mathrm{d}\theta\mathrm{d}V_n \\ + \int_{V_n}\int_t^{t+\Delta t} \rho\Big(\mathbf{X}_p : \dot{\mathbf{F}}_p + \mathbf{X}_g : \dot{\mathbf{g}}\Big)\mathrm{d}t\mathrm{d}V_n + \int_{V_n} \rho\tilde{K}(\dot{\chi})\mathrm{d}V_n\end{aligned} \tag{44}$$

and

$$\mathbf{n}\cdot\boldsymbol{\sigma}\cdot[\mathbf{F}]\cdot\mathbf{n} = \boldsymbol{\sigma} : [\boldsymbol{\varepsilon}] = \rho[\psi] + \int_{\theta_1}^{\theta_2} \rho s \mathrm{d}\theta + \int_{t}^{t+\Delta t} \rho\big(\mathbf{X}_p : \dot{\boldsymbol{\varepsilon}}_p + \mathbf{X}_g : \dot{\mathbf{g}}\big)\mathrm{d}t + \tilde{K}(v_n). \qquad (45)$$

The terms $\boldsymbol{\sigma} : [\boldsymbol{\varepsilon}] - \rho[\psi]$ in Eq. (45) coincide with those in the standard consideration of the moving interface in elastic materials (Eshelby, 1970; Kaganova and Roitburd, 1988 and references; Grinfeld, 1991), which confirms the validity of our approach based on the consideration of a process in the fixed volume. All the other terms can be obtained by considering the transformation process only. More detailed analysis of an interface propagation in inelastic materials at finite strains can be found in the paper by Levitas (1998a). The transition from an integral form (44) to a local one Eq. (45) is based on the assumption that each point of an interface can move independently of another points and the volume V_n can be chosen arbitrarily. If the interface can move without variation of its shape only (e.g. as a plane), i.e. the motion of interface points are not mutually independent, then only the integral form (44) is valid. As will be seen from the simple example in Section 9 related to crystallization, the interface does not move as a smooth surface, but first a small jog appears which then moves along the interface making the PT. The same is true of the martensitic PT or twinning, for which the jog represents a transformation or twinning dislocation. It is impossible to apply Eq. (45) to this situation, because the PT progress occurs along the interface and not in the normal to the interface direction. In the case when the time of appearance of the jog is much higher than the time of its spreading along the interface, Eq. (45) is a good approximation.

For the interface propagation in the case of time-independent kinetics it is necessary for Eq. (45) to be met at time $t + \Delta t$ as well

$$\rho_{\Delta}^{-1}\boldsymbol{\sigma}_{\Delta} : [\boldsymbol{\varepsilon}_{\Delta}] = [\psi_{\Delta}] + \int_{\theta_{1\Delta}}^{\theta_{2\Delta}} s\mathrm{d}\theta + \int_{\boldsymbol{\varepsilon}_{p1\Delta}}^{\boldsymbol{\varepsilon}_{p2\Delta}} \mathbf{X}_p : d\boldsymbol{\varepsilon}_p + \int_{\mathbf{g}_{1\Delta}}^{\mathbf{g}_{2\Delta}} \mathbf{X}_g : \mathrm{d}\mathbf{g} + \tilde{K}_{\Delta}. \qquad (46)$$

Eq. (46) is a counterpart of the consistency condition in plasticity theory, where subscript Δ denotes that a parameter is determined at time $t + \Delta t$. For an infinitesimal Δt Eqs. (45) and (46) can be transformed into

$$\begin{aligned}
&\rho^{-1}\big(\dot{\boldsymbol{\sigma}} + \mathbf{n}\cdot\nabla\boldsymbol{\sigma}v_n\big) : [\boldsymbol{\varepsilon}] + \rho^{-1}\boldsymbol{\sigma} : \big([\dot{\boldsymbol{\varepsilon}}] + \mathbf{n}\cdot[\nabla\boldsymbol{\varepsilon}]v_n\big) - \big([\dot{\psi}] + \mathbf{n}\cdot[\nabla\psi]v_n\big)\\
&- \big([s\dot{\theta}] + \mathbf{n}\cdot[s\nabla\theta]v_n\big) - \big([\mathbf{X}_p : \dot{\boldsymbol{\varepsilon}}_p] + [\mathbf{X}_p : (\mathbf{n}\cdot\nabla\boldsymbol{\varepsilon}_p)]v_n\big)\\
&- \big([\mathbf{X}_g : \dot{\mathbf{g}}] + [\mathbf{X}_g : (\mathbf{n}\cdot\nabla\mathbf{g})]v_n\big) = \frac{\partial\tilde{K}}{\partial u_n}v_n + \frac{\partial\tilde{K}}{\partial\mathbf{y}}\cdot\big([\dot{\mathbf{y}}] + \mathbf{n}\cdot[\nabla\mathbf{y}]v_n\big),
\end{aligned} \qquad (47)$$

where u_n is the interface displacement $\dot{u}_n = v_n$, $\mathbf{y} = \{\theta, \mathbf{F}_p, \mathbf{g}, \ldots\}$ is the set of arguments of $\tilde{K}$, $\tilde{K} = \tilde{K}(\mathbf{y})$. We have taken into account

$$\mathbf{a}_{\Delta} = \mathbf{a} + (\dot{\mathbf{a}} + \mathbf{n}\cdot\nabla\mathbf{a}v_n)\Delta t \qquad (48)$$

for all parameters, equation

$$\begin{aligned}\int_{\varepsilon_{p1\Delta}}^{\varepsilon_{p2\Delta}} \mathbf{X}_p : \mathrm{d}\boldsymbol{\varepsilon}_p &= \int_{\varepsilon_{p1}}^{\varepsilon_{p}2} \mathbf{X}_p : \mathrm{d}\boldsymbol{\varepsilon}_p + \int_{\varepsilon_{p2}}^{\varepsilon_{p2\Delta}} \mathbf{X}_p : \mathrm{d}\boldsymbol{\varepsilon}_p - \int_{\varepsilon_{p1}}^{\varepsilon_{p1\Delta}} \mathbf{X}_p : \mathrm{d}\boldsymbol{\varepsilon}_p \\ &= \int_{\varepsilon_{p1}}^{\varepsilon_{p2}} \mathbf{X}_p : \mathrm{d}\boldsymbol{\varepsilon}_p + \mathbf{X}_{p2} : \left(\dot{\boldsymbol{\varepsilon}}_{p2} + \mathbf{n}\cdot\nabla\boldsymbol{\varepsilon}_{p2} v_n\right)\Delta t - \mathbf{X}_{p1} : \left(\dot{\boldsymbol{\varepsilon}}_{p1} + \mathbf{n}\cdot\nabla\boldsymbol{\varepsilon}_{p1} v_n\right)\Delta t\end{aligned} \tag{49}$$

and similar equations for all other parameters. Here the term $\nabla\mathbf{a}$ appears due to the fact that the tensor $\mathbf{a}_\Delta$ is determined on the Σ_Δ surface, i.e. at point $\mathbf{r} + v_n\mathbf{n}\Delta t(\mathbf{r} \in \Sigma(t))$.

It is not necessary that after nucleation Eqs. (45) and (46) be valid for the points of the interface Σ_n because after nucleation the interface Σ_n can be fixed. But we should be sure that under the same boundary conditions the interface Σ_n does not move back and the nucleus does not disappear. The condition of nondisappearance of a nucleus is a violation of the propagation condition, when $v_n < 0$ and SC $2 \to 1$ occurs, i.e.

$$X \geqslant \tilde{K}^0_{2\to 1} < 0; \quad X_\Delta > \tilde{K}^0_{2\to 1\Delta}. \tag{50}$$

As the second non-disappearance condition we should exclude the possibility of appearance and growth under the same boundary conditions of the parent phase inside the nucleus of the product phase, i.e. Eq. $(38)_3$ for the reverse SC should be violated.

For time-dependent kinetics the non-disappearance condition is not as important as for the time-independent case. Even if this condition is violated and reverse SC can occur, the nucleus can remain in phase 2 for a long time. Consequently the kinetics of both nucleation and reverse SC have to be considered.

6. Extremum principle for determination of unknown parameters

The SC criterion and kinetic Eq. (38) is only one scalar equation which is not sufficient for the determination of all unknown parameters. To determine all unknown parameters **b**, e.g. position, shape, volume (mass) and orientation of nucleus, jump of $\boldsymbol{\varepsilon}$ and plastic deformation $\boldsymbol{\varepsilon}_p$, related to $\boldsymbol{\varepsilon}_p$ temperature evolution, functions $\boldsymbol{\varepsilon}(\xi)$, $\boldsymbol{\varepsilon}_p(\xi), \theta(\xi)$ and so on, as well as rate $\dot{\chi}$, let us use the postulate of realizability (Levitas, 1992a, 1995a, 1998a). Generally both $\bar{X}$ [calculated with the help of Eqs. (30) and (32)] and K depend on **b** and $\dot{\chi}$. The dependence of $\bar{X}$ on $\dot{\chi}$ is for the following reason. For time-dependent materials due to the dependence of constitutive equations on rates, $\bar{X}$ is some function or functional of time of SC t_s. Time dependence can arise even for elastoplastic materials due to the temperature evolution equation (if the adiabatic approximation is not valid) or due to the diffusion equation for diffusional-displacive PT.

The postulate of realizability is formulated as follows: *Let us start with such boundary conditions and some fixed* $\dot{\chi}$ *that*

$$\bar{X}(\dot{\chi}, \mathbf{b}^*) - K(\dot{\chi}, \mathbf{b}^*) < 0 \tag{51}$$

for all possible SC parameters $\mathbf{b}^*$, *i.e. SC cannot occur with this* $\dot{\chi}$. *If in the course of continuous variation of boundary conditions and for each boundary conditions continuous variation of all admissible parameters* $\mathbf{b}^*$ *the condition*

$$\bar{X}(\dot{\chi}, \mathbf{b}^*) - K(\dot{\chi}, \mathbf{b}^*) = 0 \tag{52}$$

is fulfilled for the **first** *time for some of parameters* $\mathbf{b}$, *then SC will occur with these* $\mathbf{b}$ *and* $\dot{\chi}$. If, in the course of continuous variation of boundary conditions and for each boundary conditions continuous variation of all admissible parameters $\mathbf{b}^*$ the equality (52) is met the first time for one or several $\mathbf{b}$ then for arbitrary other $\mathbf{b}^*$ inequality (51) should be held, as in the opposite case for this $\mathbf{b}^*$ condition (52) had to be met before it was satisfied for $\mathbf{b}$. Consequently, we obtain the extremum principle for determination of all unknown parameters $\mathbf{b}$:

$$\bar{X}(\dot{\chi}, \mathbf{b}) - K(\dot{\chi}, \mathbf{b}) = 0 > \bar{X}(\dot{\chi}, \mathbf{b}^*) - K(\dot{\chi}, \mathbf{b}^*) \quad \forall \mathbf{b}^* \neq \mathbf{b}. \tag{53}$$

The same procedure can be repeated for arbitrary $\dot{\chi}$. Inequality (53) (i.e. extremum principle) allows us to find actual parameters $\mathbf{b}$ for arbitrary $\dot{\chi}$, and equality (53) determines the transformation rate $\dot{\chi}$. Principle (53) after multiplication of the left part by $m_n > 0$ and the right part by $m_n^* > 0$ can be represented in the following explicit form

$$\begin{aligned}
&\int_{V_n}\int_{\boldsymbol{\varepsilon}_1}^{\boldsymbol{\varepsilon}_2} \boldsymbol{\sigma} : \mathrm{d}\boldsymbol{\varepsilon}\,\mathrm{d}V_n - \int_{V_n} \rho(\psi_2 - \psi_1)\mathrm{d}V_n - \int_{V_n}\int_{\theta_1}^{\theta_2} \rho s\,\mathrm{d}\theta\,\mathrm{d}V_n \\
&-\int_{V_n}\int_0^{t_s} \rho\big(\mathbf{X}_p : \dot{\boldsymbol{\varepsilon}}_p + \mathbf{X}_g : \dot{\mathbf{g}}\big)\mathrm{d}t\,\mathrm{d}V_n - \Delta\bar{\Gamma} - \int_{V_n} \rho\tilde{K}_{1\rightarrow 2}(\dot{\chi})\mathrm{d}V_n = 0 > \\
&\int_{V_n^*}\int_{\boldsymbol{\varepsilon}_1^*}^{\boldsymbol{\varepsilon}_2^*} \boldsymbol{\sigma}^* : \mathrm{d}\boldsymbol{\varepsilon}^*\mathrm{d}V_n - \int_{V_n} \rho\big(\psi_2^* - \psi_1^*\big)\mathrm{d}V_n - \int_{V_n^*}\int_{\theta_1^*}^{\theta_2^*} \rho s^*\mathrm{d}\theta^*\mathrm{d}V_n \\
&-\int_{V_n^*}\int_t^{t+\Delta t} \rho\Big(\mathbf{X}_p^* : \dot{\boldsymbol{\varepsilon}}_p^* + \mathbf{X}_g^* : \dot{\mathbf{g}}^*\Big)\mathrm{d}t\,\mathrm{d}V_n - \Delta\bar{\Gamma}^* - \int_{V_n^*} \rho\tilde{K}_{1\rightarrow 2}^*(\dot{\chi})\mathrm{d}V_n
\end{aligned} \tag{54}$$

for nucleation and the form of

$$\begin{aligned}
&\rho^{-1}\boldsymbol{\sigma} : [\boldsymbol{\varepsilon}] - [\psi] - \int_{\theta_1}^{\theta_2} s\,\mathrm{d}\theta - \int_t^{t+\Delta t} \big(\mathbf{X}_p : \dot{\boldsymbol{\varepsilon}}_p + \mathbf{X}_g : \dot{\mathbf{g}}\big)\mathrm{d}t - \tilde{K}_{1\rightarrow 2}(\dot{\chi}) = 0 > \\
&\rho^{-1}\boldsymbol{\sigma}^* : [\boldsymbol{\varepsilon}^*] - [\psi^*] - \int_{\theta_1^*}^{\theta_2^*} s^*\mathrm{d}\theta^* - \int_t^{t+\Delta t} \Big(\mathbf{X}_p^* : \boldsymbol{\varepsilon}_p^* + \mathbf{X}_g^* : \dot{\mathbf{g}}^*\Big)\mathrm{d}t - \tilde{K}_{1\rightarrow 2}^*(\dot{\chi})
\end{aligned} \tag{55}$$

for interface propagation. As shown by Levitas (1998a), for elastic materials and isothermal processes as well as at $K = 0$ the extremum principle (53) is equivalent to the principle of the minimum of Gibbs energy. For time-independent kinetics the extremum principles obtained play a similar role in the description of SC in dissipative materials, as does the principle of the minimum of Gibbs energy for non-dissipative materials. Application of principle (53) for time-dependent kinetics allows us to overcome some contradictions that are typical of the energetic approach and obtain new results (see Sections 8 and 9).

For the reverse SC similar reasoning leads to the following extremum principle

$$\bar{X}(\dot{\chi}, \mathbf{b}) - K_{2\rightarrow 1}(\dot{\chi}, b) = 0 < \bar{X}(\dot{\chi}, \mathbf{b}^*) - K_{2\rightarrow 1}(\dot{\chi}, \mathbf{b}^*) \quad \forall \mathbf{b}^* \neq \mathbf{b}. \tag{56}$$

and its explicit forms.

Let us derive an extremum principle for $\dot{\chi}$. If we define formally the rate $\dot{\chi}^0$ from condition $\bar{X}(\dot{\chi}^0, \mathbf{b}^*) - K(\dot{\chi}^0, \mathbf{b}^*) = 0$, then using Eq. (53) we obtain the following extremum principle

$$\bar{X}(\dot{\chi}^0, \mathbf{b}^*) - K(\dot{\chi}^0, \mathbf{b}^*) = 0 > \bar{X}(\dot{\chi}, \mathbf{b}^*) - K(\dot{\chi}, \mathbf{b}^*) \quad \forall \mathbf{b}^* \neq \mathbf{b}, \quad \dot{\chi}^0 \neq \dot{\chi}. \tag{57}$$

We would like to find how $\dot{\chi}^0$ will change at $\mathbf{b}^* \rightarrow \mathbf{b}$. We will limit ourselves to the case when the functional $\bar{X}(\dot{\chi}, \mathbf{b}^*) - K(\dot{\chi}, \mathbf{b}^*)$ decreases with growing $\dot{\chi}$ for each fixed $\mathbf{b}^*$. Then it follows from principle (57)

$$\dot{\chi}^0 < \dot{\chi} \quad \forall \mathbf{b}^* \neq \mathbf{b}. \tag{58}$$

As at $\mathbf{b}^* \rightarrow \mathbf{b}$ we have $\dot{\chi}^0 \rightarrow \dot{\chi}$ (i.e. $\dot{\chi}^0$ tends to its possible maximum value), this means we have derived the *principle of maximum of transformation rate*

$$\dot{\chi}^0 \rightarrow \max_{\mathbf{b}^*}. \tag{59}$$

It is clear that this principle is equivalent to the *principle of minimum of transformation time*

$$t_s \rightarrow \min_{\mathbf{b}^*}. \tag{60}$$

Function $K(\dot{\chi}^0, \mathbf{b})$ is the given material property and it is easy to analyze whether it is a growing function of $\dot{\chi}$ or not. The functional $\bar{X}$ can be estimated after the solution of the boundary-value problems. It seems intuitive that the higher $\dot{\chi}$ and shorter transformation time, the higher resistance to plastic flow and the smaller decrease in free energy due to diffusion (for diffusional-displacive PT) and consequently at the same boundary conditions $\bar{X}$ will be smaller. However, this may also be not the case. If $\bar{X}$ is independent of $\dot{\chi}$ (for time-independent materials and processes), then the principle (57) can be simplified

$$K(\dot{\chi}^0, \mathbf{b}^*) < K(\dot{\chi}, \mathbf{b}^*) \quad \forall \mathbf{b}^* \neq \mathbf{b}. \tag{61}$$

Principles (58) and (59) follow from Eq. (61) for the functions K which grow with growing $\dot{\chi}$. The case when the functional $\bar{X}(\dot{\chi}, \mathbf{b}^*) - K(\dot{\chi}, \mathbf{b}^*)$ increases with growing

$\dot{\chi}$ for each fixed $\mathbf{b}^*$ or is a nonmonotonous function of $\dot{\chi}$ will be treated in detail elsewhere.

The essence of the postulate of realizability is very simple: *as soon as some dissipative process (plastic flow, PT, fracture) can occur from the viewpoint of thermodynamics and kinetics, it will occur (or be realized)*, i.e. the first fulfillment of the necessary energetic condition is sufficient for the beginning of the dissipative process. A number of applications of the postulate (Levitas, 1992a, 1995a, 1997, 1998a,b,c) give the impression that such an essence is a property of dissipative systems.

Let us consider an alternative, more natural way to derive the principle (60), which, however, cannot be applied to time-independent systems. Assume that kinetic equation $\bar{X}(\dot{\chi}, \mathbf{b}^*) - K(\dot{\chi}, \mathbf{b}^*) = 0$ (or equation $\bar{X}(\dot{m}_{\mathbf{n}}, \mathbf{b}^*) - K(\dot{m}_{\mathrm{n}}, \mathbf{b}^*) = 0$ with an alternative choice of force $\bar{X}$ and rate $\dot{m}_n$) can be resolved with respect to the transformation time, i.e. $t_s = q(EC, \mathbf{b}^*)$, where EC means "the external conditions". In particular, when $\bar{X}$ is independent of $\dot{\chi}$, $t_s = f^{-1}\left(\bar{X}, \mathbf{b}^*\right)$. We formulate the postulate of realizability in the following form: *as soon as SC can occur (i.e. when the SC criterion is satisfied), it will occur during the time determined by equation* $t_s = q(EC, \mathbf{b}^*)$ for each $\mathbf{b}^*$.

Consequently, for each thermodynamically admissible parameter $\mathbf{b}^*$ SC can occur during the corresponding time (Fig. 5). We are allowed to do nothing and simply wait. Then we will observe the SC which appears in the shortest time, i.e. we arrive at the *principle of minimum of transformation time* (60). The occurrence of this SC changes the initial condition (and consequently *EC*) for the next SC in another volume, which again we will observe in the shortest time.

If we choose alternatively $\bar{X}$ and $\dot{m}_n$ as force and rate and apply the postulate of realizability in the form given before Eq. (53) at fixed $\dot{m}_n$ instead of $\dot{\chi}$, we obtain the principle of maximum of $\dot{m}_n$ instead of maximum of $\dot{\chi}$, i.e. the results depend on the choice of generalized rate. The advantage of the alternative derivation is that we did not fixed any rate, i.e. result is independent of the choice of generalized rate. Nevertheless, the principle of minimum of transformation time derived coincides with the principle obtained with the choice of $\dot{\chi}$ as a generalized rate. That is why we made such a choice at the very beginning [after Eq. (30)]. Another advantage is related to the independence of the result of increasing or decreasing the functional $\bar{X}(\dot{\chi}, \mathbf{b}^*) - K(\dot{\chi}, \mathbf{b}^*)$ with growing $\dot{\chi}$.

A general scheme for the application of PT or SC criterion and extremum principle (53) is as follows. All material properties and constitutive Eqs. (15), (19) and (24) must be given as a function of ξ. Then assume that at some initial conditions and prescribed boundary conditions some SC occurs in some region V_n. We introduce incrementally in some way transformation strain and change the elastic moduli, heat capacity, thermal expansion coefficient and yield strength from initial to final values in a possible SC region. After solving the elastoplastic and thermal boundary-value problems for each step we determine the variation of all fields, calculate the driving force $\bar{X}$ and resistance K and determine the SC rate $\dot{\chi}$ and consequently the SC time t_s from equality $\bar{X}(\dot{\chi}) = K(\dot{\chi})$. Then we vary the possible SC region and way of variation of transformation strain, elastic moduli, heat capacity, thermal expansion

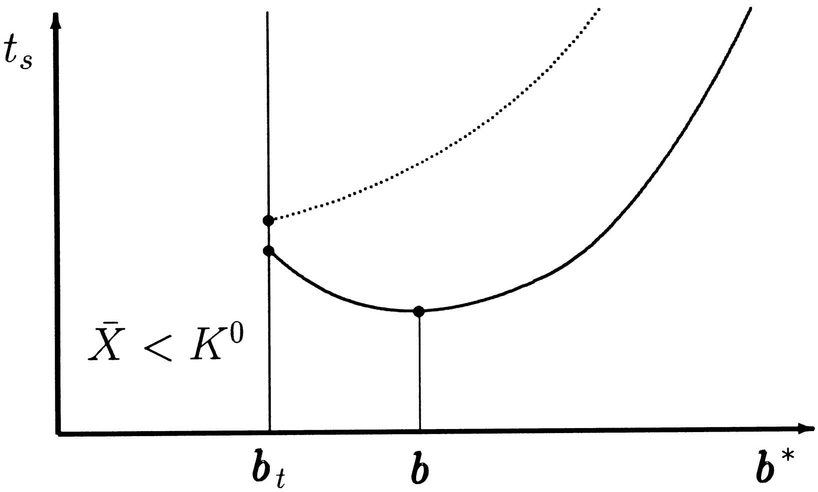


Fig. 5. Scheme for derivation of the principle of minimum of transformation time.

coefficient and yield strength from initial to final values in it (homogeneously or heterogeneously) and find such a SC region and way of varying of transformation strain and properties which maximize the difference $\bar{X} - K$ or minimize the transformation time t_s. The final SC rate and time in this region are determined again by equality $\bar{X}(\dot{\chi}) = K(\dot{\chi})$.

7. Thermally activated kinetics: concept of effective temperature

As an example of a kinetic equation we consider

$$\bar{X} = K^0 + E_a + \frac{n}{m_n N} R\theta_{\mathrm{ef}} \ln \frac{\dot{\chi}}{\dot{\chi}_o}, \tag{62}$$

which leads to size-dependent Arrhenius-type kinetics and includes both thermal activation and athermal threshold K^0:

$$\frac{1}{\dot{\chi}} = t_s = t_o \exp\left(- \frac{\left(\bar{X} - K^0 - E_a\right) m_n}{\mathrm{R}\theta_{\mathrm{ef}}} \frac{N}{\mathrm{n}} \right) \quad \text{at} \quad 0 \leqslant \bar{X} - K^0 \leqslant E_a, \tag{63}$$

$$\dot{\chi} = 0 \quad \text{at} \quad \bar{X} - K^0 < 0. \tag{64}$$

Here E_a is the activation energy per unit mass at $\bar{X} - K^0 = 0$, $R = 8.314$ J/(k mol) is the gas constant, $t_o = \frac{1}{\dot{\chi}_o}$ is some characteristic time, n the number of atoms in volume V_n which undergo thermal fluctuations, $N = 6.02 \cdot 10^{23}$ is Avogadro's number

(number of atoms in 1 mol), θ_{ef} is the effective temperature. In Eq. (63) the actual activation energy includes difference $K^0 - \bar{X}$ to take into account the effect of the driving force, dissipative threshold and surface energy on kinetics. The lower bound for $\bar{X} - K^0$ in Eq. $(63)_2$ is related to the fulfillment of the SC criterion; the upper bound is due to the requirement of positiveness of an actual activation energy, since otherwise the process does not need thermal activation and Eq. (63) loses its sense.

Sometimes the transforming volume or mass is absent in Arrhenius-type kinetic equations and activation energy is defined per unit atom (or per mol when N is omitted). Such an equation can be applied to atom by atom SC only (see Sections 8 and 9); its application to the appearance of a multiatomic nucleus is contradictory. In most cases an activation energy of the whole transforming volume is taken into account as in Eq. (63), e.g. when the activation energy is an energy of a critical nucleus at PT (Christian, 1965; Olson and Cohen, 1986), for dislocation motion (Kocks et al., 1975) or for plastic deformation of polymers (Bulatov and Argon, 1994). As will be demonstrated in Sections 8 and 9 and in Part II of this paper, accounting for the transforming mass in Eq. (63) in combination with the extremum principle derived plays crucial role in the qualitative and quantitative description of SC kinetics.

By introducing the *effective temperature* we also take into account the fact that temperature can vary significantly during the SC. In the standard approach describing kinetics in terms of rate of product phase volume fraction change, we can use ordinary temperature, averaged over some representative volume, because small change of volume fraction causes small temperature variation and this can be taken into account by solving coupled kinetic and temperature evolution equation. In a local description we must take into account the local temperature variation in the transforming region during the entire SC, which for example for reactions in Ti–Si and Nb–Si systems exceeds 1500 K (Levitas et al., 1998b,c).

We assume that during the SC (growth of ξ from 0 to 1) multiple thermal activations are necessary for each small increment $\Delta\xi$. Then as the simplest variant we define

$$\theta_{\mathrm{ef}} := \int_0^1 \theta_m \mathrm{d}\xi; \quad \theta_m^{-1} := \frac{1}{3V_n}\int_{\Sigma_n} \theta^{-1}\mathbf{r}\cdot\mathbf{n}d\Sigma_n = \frac{1}{V_n}\int_{V_n}\left(\theta^{-1} + \frac{1}{3}\nabla\theta^{-1}\cdot\mathbf{r}\right)\mathrm{d}V_n. \tag{65}$$

Here θ_m is the macroscopic temperature for a finite volume V_n with heterogeneously distributed temperature. Eq. $(65)_2$ for macroscopic temperature is derived by Levitas (1992b) from the condition that the first and the second laws of thermodynamics for a finite volume can be written in the same way as for the material point. This definition has a similar form as for macroscopic stress and heat flux. For a cubic shape of the volume V_n Eq. $(65)_2$ is equivalent to ordinary averaging over the surface Σ_n for the inverse temperature. It is clear that alternative definitions and descriptions to Eq. $(65)_1$ are possible, e.g.

$$\theta_{\mathrm{ef}}^{-1} := \int_0^1 \theta_m^{-1}\mathrm{d}\xi. \tag{66}$$

The introduction in Eq. (63) of the number n of atoms in volume V_n which undergo thermal fluctuations allows us to consider transformation in a macroscopic volume in which multiple thermal fluctuations occur. It is necessary to mention that we have no intention to derive Eq. (63) using the Boltzmann statistic. The later is used as a hint for postulating some plausible macroscopic kinetic equation which can be reduced to classical results. In contrast to known approaches describing kinetics in terms of rate of change of product phase volume fraction or rate of interface propagation, in our approach we determine the time of structural change in some transforming volume. This is consistent with the standard method of nonequilibrium thermodynamics which relates conjugate force and rate in the expression for dissipation rate.

Of course, other types of kinetic relations between $\bar{X}$ and $\dot{\chi}$ can be considered, see for example Kocks et al. (1975) for dislocation motion or Ghosh and Olson (1994) for martensitic PT. For fracture, for example, the actual activation energy can depend on stresses rather than on or additional to the driving force $\bar{X}$.

Generally, for thermally activated kinetics we can decompose SC time $t_s = t_w + t_p$, where t_w is the waiting time (i.e. time of waiting of a fluctuation with the required activation energy) and t_p is the process time (i.e. the actual time of the transformation process). We assume that for each t_w and t_p the same form of kinetic Eq. (63) is valid, but with different predexponential multipliers t_{ow} and t_{op}, and $t_o = t_{\mathrm{ow}} + t_{0\mathrm{p}}$. When calculating $\bar{X}$ according to definition Eq. (29) and Eq. (30) we should use $1/t_p$, because all the fields depend on the actual process time rather than on the waiting time or t_s, i.e. $\bar{X} = \bar{X}\left(1/t_{\mathrm{p},}\mathbf{b}\right)$ When calculating of t_w the temperature θ_1 probably has to be used, because the waiting time depends on the first fluctuation at the beginning of the SC.

The principle of minimum of transformation time (59) for $\bar{X}$ independent of t_s results in

$$t_s = t_o \exp - \frac{\left(\bar{X}(\mathbf{b}^*) - K^0(\mathbf{b}^*) - E_a(\mathbf{b}^*)\right) m_n^*}{\mathrm{R}\theta^*_{\mathrm{ef}}} \frac{N}{\mathrm{n}} \quad \rightarrow \quad \min . \tag{67}$$

An extremum principle (67) has to be used under the constraint that the thermodynamic criterion of SC is fulfilled and, as will be explained below, that each size of transforming volume cannot be smaller than the size of atom (n atoms) in the same direction.

For the time-independent case kinetic Eq. (63) and (64) and extremum principle (56) reduce to

$$\bar{X}(\mathbf{b}) - K^0(\mathbf{b}) = 0 > \bar{X}(\mathbf{b}^*) - K^0(\mathbf{b}^*). \tag{68}$$

If SC does not require thermal activation, then Eq.(63) has to be substituted with another one.

8. New kinetic concept of thermodynamically admissible nucleus

Let us consider a problem on SC in a sphere with a radius r inside the infinite space under homogeneous pressure at the boundary. We will not concentrate on the concrete rheology of the material, type of SC and concrete stress–strain fields. As SC we can consider melting, solidification, vaporization, condensation, displacive PT or void nucleation. It is convenient to introduce the local dissipation increment due to SC only

$$X := \int_0^1 X_\xi \mathrm{d}\xi. \tag{69}$$

Assume that we have found X for SC and the material under consideration and X and $\tilde{K}$ are homogeneously distributed in a volume V_n and are independent (along with θ_{ef}) of the radius r or any other geometric parameter of nucleus for a non-spherical transforming region; X is independent of $\dot{\chi}$ as well. The kinetic equation and transformation condition (63) read as

$$t_s = t_o \ \exp\left(-\frac{\left(X - K^0 - E_a\right)\frac{4}{3}\pi r^3 \rho - 4\Gamma\pi r^2}{R\theta_{\mathrm{ef}}}\frac{N}{\mathrm{n}}\right) \tag{70}$$

at

$$\left(X - K^0\right)\frac{4}{3}\pi r^3 \rho - 4\Gamma\pi r^2 \geqslant 0, \tag{71}$$

$$\left(X - K^0 - E_a\right)\frac{4}{3}\pi r^3 \rho - 4\Gamma\pi r^2 < 0. \tag{72}$$

Minimizing t_s with respect to r [according to principle (67)] we obtain the result that at $X - K^0 - E_a \leqslant 0$ the radius $r \to \min$, because the function under consideration decreases monotonically with the decrease in r. The minimum value of r can be found from SC condition $\bar{X} = K^0$, i.e.

$$\left(X - K^0\right)\frac{4}{3}\pi r_t^3 \rho - \Gamma 4\pi r_t^2 = 0 \quad \text{or} \quad r_t = \frac{3\Gamma}{\rho\left(X - K^0\right)}. \tag{73}$$

Substitution of (73) in Eq. (70) results in

$$t_s = t_o \exp\left(\frac{\rho \mathrm{E}_a}{\mathrm{R}\theta_{\mathrm{ef}}}\frac{N}{n}\frac{4}{3}\pi r_t^3\right). \tag{74}$$

In the case $X - K^0 - E_a > 0$ the minimum of t_s in Eq. (70) will be at $r \to \infty$, then condition (72) is violated, the nucleation does not require the thermal activation and Eq. (70) has to be replaced by another one. Consequently, in the framework of consideration based on Eq. (70) we have to require

$$X - K^0 - E_a \leqslant 0. \tag{75}$$

Let us analyze the results. If at a given boundary condition the SC condition (73) is satisfied, then a new nucleus with radius r_t appears during the shortest time determined by Eq. (74). We will call a nucleus (void) with a radius r_t the *thermodynamically admissible nucleus*. It has nothing in common with the usual critical nucleus (Fig. 6), which is determined by maximization of the Gibbs potential

$$G = -X\frac{4}{3}\pi r^3 \rho + \Gamma 4\pi r^2, \tag{76}$$

with respect to r, i.e.

$$r_c = \frac{2\Gamma}{\rho X}. \tag{77}$$

It is evident that for a critical nucleus $\bar{X}_c m_n = -\frac{16}{3}\frac{\pi\Gamma^3}{(\rho X)^2} < 0$, i.e. its appearance is impossible from the point of view of the nucleation criterion and contradicts the second law of thermodynamics. If a nucleus with size exceeding r_c is introduced in the material under consideration from outside (e.g. for liquids and gases), it will grow (at $K^0 = 0$), but it cannot appear from a thermodynamic point of view. In contrast to the usual approach, we determine a thermodynamically admissible nucleus in terms of kinetics: this is a nucleus with the minimum possible size which can appear from the point of view of thermodynamics and appears during the shortest possible time. According to Eq. (73), X or $\bar{X}$ satisfy "static" SC condition $\bar{X} = K^0$ (as for rate-independent SC). Increase in X reduces the radius of the thermodynamically admissible nucleus, which leads to growth in the rate of transformation in Eq. (74). In contrast to known approaches (Zeldovich, 1943; Christian, 1965; Volmer, 1983; Olson and Cohen, 1986) the value $\bar{X}_c m_n$ does not appear in the

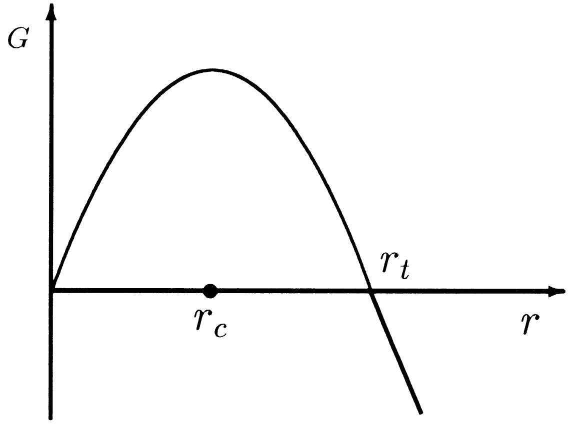


Fig. 6. Typical dependence of Gibbs free energy on radius of nucleus and definition of radius r_c of critical nucleus and radius r_t of thermodynamically admissible nucleus.

activation energy. Activation energy E_a is related to overcoming the energy barrier in Fig. 4 in each point of volume of the thermodynamically admissible nucleus and such an activation energy is absent in the above papers. However, it is taken into account in numerous papers by Abeyaratne and Knowles (1993 and reference) on interface propagation, see also Levitas (1992a, 1995a) and Levitas et al. (1996). It is clear from Eqs. (73) and (77) that at $K^0 = 0$ the radius of the equilibrium nucleus is 1.5 of the radius of critical nucleus; at $K^0 \neq 0$ the difference can be much larger. The activation energy E_a also depends on the heterogeneous defect distribution in the transforming volume and the interaction between SC and these defects (Ghosh and Olson, 1994).

It is necessary to note that the concrete size of the thermodynamically admissible nucleus can be obtained in the framework of a time-dependent thermoactivated kinetic description only. For time-independent kinetics, when SC starts with the fulfillment of thermodynamic criterion of SC (71) and the transformation time is undetermined, extremum principle (68) (or principle of the minimum of Gibbs free energy) results in $r \rightarrow \infty$ and transition in the whole sphere. In this case SC starts at $X = K^0$ as in the case with neglected surface energy. A similar problem in PT in elastic or elastoplastic materials gives the same result. It is evident that in the homogeneous fields during the PT in the whole uniform body (i.e. when the whole body transforms simultaneously), all the continuum mechanics equations are satisfied and it is easy to show using the simplest example that the homogeneous solution maximizes the functional (68). This is related to the fact that for a nonhomogeneous solution the positive energy of internal (eigen) stresses produces the negative contribution to the driving force of PT. As the corroborating example we can use the solution of a spherically symmetrical problem of the appearance of the spherical nucleus of a new phase with the radius r in sphere with the radius R_o loaded by homogeneously distributed pressure p (Roitburd and Temkin, 1986). The result is obtained that for elastic (and elastoplastic) material at $K^0 = 0$ the PT pressure monotonically decreases with the growth of radius r, and the maximum in the principle (53) will be attained at $r = R_o$. The kinetic consideration only, due to its accounting for the activation energy $E_a m_n$, results in the finite size of the nucleus.

A similar consideration is valid for the nucleus with arbitrary shape. Optimum shape is determined by extremum principle (53) and the optimum size can be determined from the principle of the minimum of transforming mass (volume). The kinetic equation and transformation condition (63) for an arbitrary shape appear as

$$t_s = t_o \exp\left(-\frac{\left(X - K^0 - E_a\right)\rho \Sigma_n \bar{r} - \Gamma \Sigma_n}{R\theta_{\mathrm{ef}}} \frac{N}{n}\right); \tag{78}$$

$$\left(X - K^0\right)\rho \Sigma_n \bar{r} - \Gamma \Sigma_n \geqslant 0, \tag{79}$$

where the characteristic size $\bar{r}$ is introduced by formula $\bar{r} = \frac{V_n}{\Sigma_n}$. According to principle (67), the minimum of the left part of Eq. (78) with respect $\bar{r}$ and Σ_n under the constraint (75) corresponds again to $\bar{r} \rightarrow \min$ and $\Sigma_n \rightarrow \min$. The minimum value of $\bar{r}$ (characteristic size of the thermodynamically admissible nucleus) can be found from condition $\bar{X} = K^0$, i.e.

$$\left(X - K^0\right)\rho\Sigma_n\bar{r} - \Gamma\Sigma_n = 0 \quad \text{or} \quad \bar{r}_t = \frac{\Gamma}{\rho(X - K^0)}. \tag{80}$$

Instead of Eq. (74) we obtain for an arbitrary form of nucleus

$$t_s = t_o \exp\left(\frac{\rho E_a}{\mathrm{R}\theta_{\mathrm{ef}}}\frac{N}{n}\Sigma_n\bar{r}_t\right). \tag{81}$$

Let us now consider the general case when $\bar{X}$, K^0 and E_a depend on some variable parameters $\mathbf{b}$ which in particular include the transformation mass m_n. After multiplication of the left part of Eq.(53) by $\frac{m_n}{\theta_{\mathrm{ef}}} > 0$ and the right part of inequality (53) by $\frac{m_n^*}{\theta_{\mathrm{ef}}^*} > 0$ we obtain an equivalent extremum principle

$$\left(\bar{X}(\dot{\chi}, \mathbf{b}^*) - K(\dot{\chi}, \mathbf{b}^*)\right)\frac{m_n^*}{\theta_{\mathrm{ef}}^*} \rightarrow \max. \tag{82}$$

Such a form is more convenient, because when substituting in it kinetic Eq. (62), the term with $\ln\frac{\dot{\chi}}{\dot{\chi}_o}$ is fixed (n is not varied) and extremum principle (82) reduces to

$$\left(\bar{X}(\dot{\chi}, \mathbf{b}^*) - K^0(\mathbf{b}^*) - E_a(\mathbf{b}^*)\right)\frac{m_n^*}{\theta_{\mathrm{ef}}^*} \rightarrow \max. \tag{83}$$

Let a characteristic size and consequently volume V_n and mass m_n of nucleus not be solely determinable from principle (83), because the thermodynamic criterion of SC is violated for the solution of the extremum problem (83). Then the thermodynamic restriction

$$\bar{X}(\dot{\chi}, \mathbf{b}^*) - K^0(\mathbf{b}^*) = 0 \tag{84}$$

has to be taken into account. Combining of Eqs. (83) and (84) results in principle

$$\frac{E_a(\mathbf{b}^*)m_n^*}{\theta_{\mathrm{ef}}^*} \rightarrow \min \tag{85}$$

or at mutually independent E_a, m_n and θ_{ef} in

$$E_a(\mathbf{b}^*) \rightarrow \min; \quad m_n^* \rightarrow \min; \quad \theta_{\mathrm{ef}}^* \rightarrow \max. \tag{86}$$

Due to Eq. (84) kinetic equation (63) can be simplied

$$t_s = t_o \exp\left(\frac{E_a m_n}{R\theta_{\mathrm{ef}}}\frac{N}{n}\right) \tag{87}$$

When an increment of surface energy is zero, it follows from Eq. (80) that $\bar{r}_t = m_n = 0$, total activation energy is zero, i.e. the process ceases to be thermo-activated and Eq. (64) cannot be applied. From a physical point of view it is clear that the transforming mass cannot be less than a mass of one atom (or molecule) when only one atom undergoes thermal fluctuation or in the case of the validity of Eq. (64) cannot be less than a mass of n atoms which undergo thermal fluctuations. In both cases the lower bound for the ratio $\frac{m_n}{n}$ when applying the extremum principle (67) is the mass of a single atom m_a. Consequently, when the increment of surface energy is so small that $\bar{r}_t = \frac{\Gamma}{\rho(\bar{X}-K^0)}$ is smaller than characteristic size of atom (or interatomic distance), then according to extremum principle (67) and constraint $\frac{m_n}{n} \geqslant m_a$ it follows

$$\frac{m_n}{n} = m_a; \quad t_s = t_o \exp\left(-\frac{\left(\bar{X} - K^0 - E_a\right)m_a}{\mathrm{k}\theta_{\mathrm{ef}}}\right) \quad \text{at} \quad 0 \leqslant \bar{X} - K^0 \leqslant E_a, \tag{88}$$

when $k = \frac{R}{N} = 1.380 \cdot 10^{-23}$ is the Boltzmann constant and $(\bar{X} - K^0 - E_a)m_a$ is the actual activation energy per atom. In this case growth of $\bar{X}$ reduces an actual activation energy per unit mass at fixed transforming mass m_a in contrast to the situation when $\bar{r}_t$ exceeds the characteristic size of the atom and constraint $\frac{m_n}{n} > m_a$ does not play any part.

A stronger constraint is that each size of nucleus a_i cannot be smaller than size of an atom a_{ai} (n atoms) in the same direction. Consequently, the principle of minimum transformation time (60) results in particular in the principle of the minimum of transformation mass

$$m_n^* \rightarrow \min \quad \text{at} \quad a_i > a_{ai}. \tag{89}$$

9. Appearance and growth of a cubic nucleus

It is clear that maximum change in surface energy will occur when a new nucleus of phase 2 appears without contact to a previously formed phase 2, because the change in surface is maximal. *Growth* of this nucleus can be modeled by the *appearance of another nucleus* which contacts the first one, then the next nucleus and so on. Assuming for simplicity that parameter X remains constant and the surface energy is isotropic, consider a model problem of nucleation in the form of a rectangular parallelepiped with two equal sides with length b and another side with length a. Assuming that the surface energy is large enough and constraint $\frac{m_n}{n} > m_a$ is satisfied, the principle of minimum of mass (volume) and the SC criterion read as follows

$$ab^2 \rightarrow \min; \quad \left(X - K^0\right)ab^2 - \Gamma\left(b^2 l + abz\right) = 0, \tag{90}$$

where l and z are the number of new sides with the area b^2 and ab respectively (Fig. 7). For the new nucleus $l = 2$ and $z = 4$ for the nucleus at the plane surface

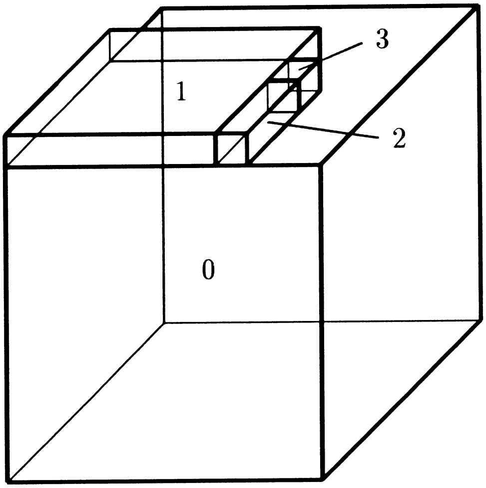


Fig. 7. Subsequent transformed volumes 0→1→2→3.

$l = 0$ and $z = 4$. Expressing a in terms of b using Eq. $(90)_2$ and substituting it in condition $\frac{\partial ab^2}{\partial b} = 0$ we obtain

$$b = \frac{3z\Gamma}{2(X - K^0)}; \quad a = b\frac{2l}{z}; \quad V_n = b^3\frac{2l}{z} = \frac{27}{4} \quad \frac{lz^2\Gamma^3}{(X - K^0)^3}. \tag{91}$$

It is clear that $X - K^0 > 0$. For the new nucleus

$$l = 2, \;\; z = 4, \;\; a = a_o = b = b_0 = \frac{6\Gamma}{(X - K^0),} \quad V_n = 216\frac{\Gamma^3}{(X - K^0)^3} = V_{n0}. \tag{92}$$

For the nucleus 1 which appears at the existing plane $l = 0$ (there are no new surfaces in the direction orthogonal to existing plane) $z = 4, b = b_0, a = 0$, we consequently have to limit a by the interatomic distance $a = a_a$ and determine corresponding b from the SC criterion, i.e.

$$b_1 = \frac{4\Gamma}{X - K^0} = \frac{2}{3}b_0; \quad V_n = a_a b_1^2 = \frac{4a_a}{9\mathrm{b}_0}V_{n0}. \tag{93}$$

Consequently the height of the nucleus which appears at the existing plane is equal to the interatomic distance and transforming volume and total activation energy can be much smaller than for the new nucleus which appears without contact to the previously formed phase 2. For nucleus 2 which characterizes the growth of nucleus 1 along the plane, there are no new surfaces in the growth direction as well and two sizes of nucleus are equal to interatomic spacing, i.e.

$$l=2,\quad z=0,\quad b_2=a_a,\quad a_2=\frac{2\Gamma}{(X-K^0)}=\frac{b_0}{3}=\frac{b_1}{2},\quad V_n=\frac{a_a^2 b_0}{3}=\frac{a_a^2 V_{n0}}{3b_o^2}. \tag{94}$$

For new nucleus, as well as for nuclei 1 and 2, kinetic Eq. (87) has to be used. For volume 3 (Fig. 7) $l=z=0, a=b=a_a, V_n=a_a^3=(\frac{a_a}{b_0})^3 V_{n0}$,, i.e. the number of new sides and consequently the increment in surface energy are zero and atom by atom transition will occur, which is according to Eqs. (87) and (88) much faster.

Consequently after the appearance of the initial nucleus it will spontaneously grow with increasing velocity because of the decrease in nucleus dimension and transforming mass at each transformation event. If X decreases or K^0 increases during the growth process (due to change of boundary conditions, heterogeneities and so on), equilibrium can occur.

Let us consider the model problem of the appearance and growth of nucleus on a crystal plane (Fig. 8) assuming $n=1$. As follows from the above considerations, the height of nucleus is equal to a_a. At the beginning nucleus 1 appears (with sides $b_1 =$

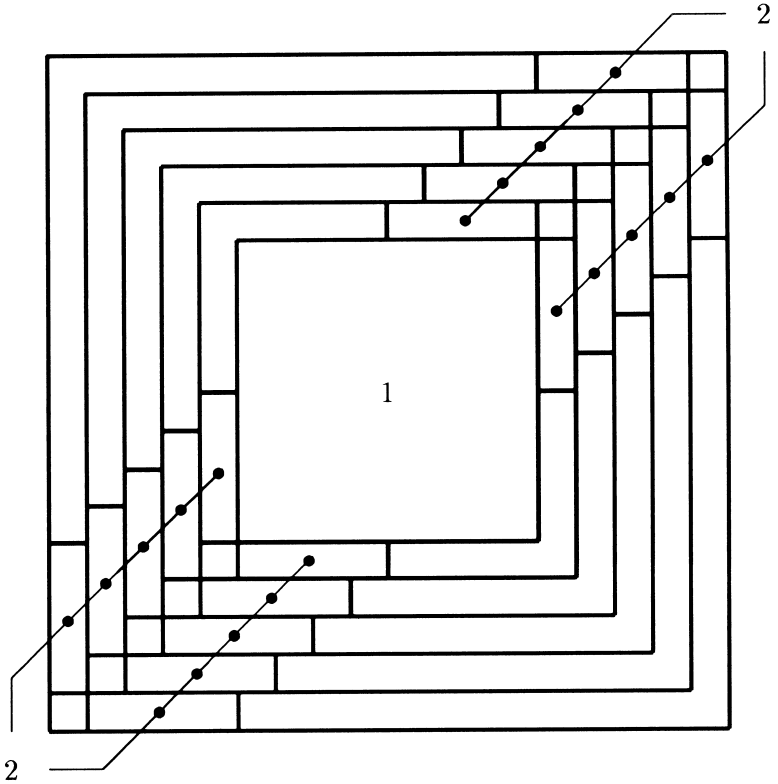


Fig. 8. Appearance of square nucleus 1 on a crystal plane and its subsequent growth by appearance of four nucleus 2 and atom by atom transformation in the remaining parts of each square.

$2/3b_0$ determined by Eq. (93) and volume $V_n = b_1^2 a_a)$ during the time t_1 determined by

$$t_1 = t_0 \exp\left(\frac{\rho E_a}{k\theta_{\text{ef}}} a_a \left(\frac{4\Gamma}{X - K^0}\right)^2\right). \tag{95}$$

Then assume that four nuclei 2 appear at each side of nucleus 1. For these nuclei $b_2 = a_a$, $a_2 = 1/3b_0 = 1/2b_1$, $V_n = 1/3b_0 a_a^2$ [see Eq. (94)] and they appear during time t_2 determined by

$$t_2 = t_0 \exp\left(\frac{\rho E_a}{\text{k}\theta_{\text{ef}}} a_a^2 \left(\frac{2\Gamma}{X - K^0}\right)\right). \tag{96}$$

Then the remaining part of new square will be transformed during atom by atom growth. The time t_3 of a single event can be determined by Eq. (88), the maximum number of events with two-sided growth of each nucleus 2 is equal to $(\frac{b_1}{2a_a} + 1)$[in the case when two pair of nuclei 2 appear near the two corners lying on diagonal of the square (Fig. 8)], the minimum number of events is equal to $(\frac{b_1}{4a_a} + 1)$ (when all nuclei 2 appear at the middle of each side of the square). Assume for definiteness that the maximum number of events is realized, and the time of atom by atom growth is $t_3(\frac{b_1}{2a_a} + 1)$. After the formation of each complete square, the appearance of 4 nuclei 2 at each side of the square with $b_2 = a_a$, $a_2 = 1/3b_0 = 1/2b_1$ during time t_2 will occur. Then the atom by atom transformation will transform the remaining part of each square. Each one-atomic rod is two atoms longer than the previous one. The possible maximum time to complete a rod number l by atom by atom growth is $t_3(\frac{b_1}{2a_a} + 2l - 1)$. After summing the time of all the events for the growth of l one-atom rods we obtain the total time

$$\begin{aligned} T &= t_1 + t_2 l + t_3 \sum_{i=1}^{l} \left(\frac{b_1}{2a_a} + 2i - 1\right) \\ &= t_1 + t_2 l + t_3 l \left(\frac{b_1}{2a_a} + l\right). \end{aligned} \tag{97}$$

Dividing the displacement a_a by T we obtain the mean value of interface propagation rate v_n. It is clear that in the general case the interface propagation rate depends on the size $\frac{lb_1}{2}$ of the transforming region, i.e. the prescription of the local kinetic equation for each interface point of type $v_n = f(X)$ is not correct. Such an assumption can be used only when the first term in Eq. (97) significantly exceeds two other terms, i.e. at $T \simeq t_1$ In this case

$$v_n := \frac{a_a}{t_1} = \frac{a_a}{t_0} \exp\left(-\frac{\rho E_a}{\text{k}\theta_{\text{ef}}} a_a \left(\frac{4\Gamma}{X - K^0}\right)^2\right). \tag{98}$$

For martensitic PT and twinning, additional change in elastic energy and plastic relaxation have to be taken into account during the calculation of local dissipation

X. Elastic energy reduces X with the appearance of an initial nucleus and nucleus 1 at the existing plane of the new phase; however, it promotes growth of the nucleus. Atom by atom growth is produced by the motion of transformation or twinning dislocations.

It is not obligatory that with the appearance of a finite new nucleus the process in it should occur homogeneously. One of the ways to determine the actual history of all the fields is to apply the extremum principle (67). For this purpose it is necessary to know how the actual activation energy depends on the process history in the transforming volume. It may happen that atom by atom transformation or fracture will occur again, but the criterion of SC has to be fulfilled *not for each step but rather for the final thermodynamically admissible nucleus*. In this case energy of a critical nucleus may appear in the kinetic equation. Another way is to use atomistic theories, like Volmer (1983) theory of liquid-gas PT or the Zeldovich (1943) theory of cavitation (see also Christian, 1965). These theories consider atom by atom growth of a subcritical nucleus which is accompanied by an increase in the total energy, i.e. contradicts the second law of thermodynamics. They can be made consistent with our approach and macroscopic thermodynamics, if we apply thermodynamics not to each growth event, but to the entire process of the appearance of a thermodynamically admissible (and not critical) nucleus. Any relation to thermodynamics for a subcritical nucleus is contradictory, because the appearance and growth of a subcritical nucleus is accompanied by negative dissipation. One such error is contained in Zeldovich's (1943) paper. In order to relate the microscopic diffusion coefficient and macroscopic properties, equality between rate of energy change of gas ball due to PT liquid-gas (with the sign "-") and rate of viscous dissipation in the surrounding liquid is assumed. As the first term is the dissipation due to PT with the sign "-", this assumption means that total dissipation due to PT and viscous flow is zero, i.e. negative dissipation due to PT is compensated for by positive dissipation due to viscous flow. This assumption is contradictory, because first, it assumes nonlocal thermodynamical coupling in a local system (PT and viscous flow occurs in different regions) and second, in an ideal liquid the dissipation due to PT has to be zero, i.e. a critical nucleus which requires negative dissipation cannot appear.

In reality nucleation represents a stochastic process and the above equations estimate averaged nucleation time and rate over some ensemble. Stochasticity of the process can be taken into account by a stochastic distribution of K^0 and E_a in the material. In order to avoid evident contradictions described below in the framework of a deterministic macroscopic model we have to introduce heterogeneities in some way. In the case of uniform material and homogeneous stresses at the boundary we can divide the whole volume V into small subvolumes with size $\bar{r}_t$ and transform all of them independently and simultaneously. Particular cases of the above boundary conditions are realized in the PT under hydrostatic pressure (in liquid or gas) and temperature-induced PT (stress-free boundary). Such a process proceeds more quickly than transformation in a single volume V_n. In gas or liquid the acceleration is due to decrease for each transforming element the change in surface energy (for internal volumes it is equal to zero). For PT in elastic or elastoplastic materials an additional positive contribution to the driving force is related to the decrease to zero of the energy of internal (eigen) stresses.

To avoid such a contradiction we have to forbid the touching of simultaneously transforming volumes and distribute heterogeneously parameters K^0 and E_a. Impurities in the gas or liquid can be modeled by heterogeneities of surface energy. In solids various types of defects (dislocations, point defects, grain boundaries) can produce a nonuniform stress field.

In the second simplified variant we can assume that parameters K^0 and E_a depend on the volume fraction c (or volume) of the second phase. The place of nucleation is defined mostly by a maximization of driving force $\bar{X}$ (due to homogeneity of K^0 and E_a distributions) and sufficient increase of K^0 and E_a and c growth lead to the stability of the two-phase equilibrium for martensitic PT or to proper macroscopic kinetics for time-dependent structural changes.

10. Conclusions

1. The aim of this paper is to develop a unified continuum thermodynamic and kinetic approaches for a wide class of SC in arbitrary inelastic materials. In comparison with the corresponding method for elastic materials (Eshelby, 1956, 1970; Cherepanov, 1967, 1979; Rice, 1968a,b) the following new problem arises: consideration of a deformation process in the transforming region; consideration of the nucleation problem in a finite region which has to be determined; separation of the dissipation rate due to SC only from other dissipative processes.
2. The class of *SC without stable local intermediate state* is defined and treated. The SC are considered as a process of variation of the transformation deformation gradient and some or all thermomechanical properties in infinitesimal or finite transforming volume from the initial to final value. This process *cannot be stopped at an intermediate state in any transforming point*. The thermodynamical equilibrium for intermediate value of transformation deformation gradient or material properties is impossible.

   Such a definition includes various phase transitions (martensitic, diffusional-displacive and second-order phase transitions), twinning, ductile fracture (nucleation and growth of void and cracks), solid-state chemical reactions, generation of point defects, dislocations, disclinations, recrystallization, deformation of amorphous materials, and so on. Spinodal decomposition and continuous saturation of the crystal lattice of one component by atoms of other substances, as well as PT with a diffuse interface of finite thickness are excluded.
3. The thermodynamics of SC without a stable local intermediate state is developed. It is shown that for such a model the application of a standard dissipative inequality local in time (e.g. Clausius–Duhem inequality) is contradictory. It is necessary to average the thermodynamic parameters, related to SC, over duration t_s of SC in order to filter off fluctuations. Even when all constitutive equations describing deformation are local, for SC the nonlocal

equation only can be derived due to surface energy. Consequently, we arrive to nonlocal both in space and time formulation. The dissipation rate due to SC only is separated from another dissipative contributions. Explicit expression for generalized (driving) force and rate is derived: the force is the total dissipation increment due to SC only during the complete SC in transforming region; the rate is the inverse transformation time. For arbitrary inelastic materials the driving force for SC represents a region-independent integral.

4. A thermodynamic criterion of SC and a general thermodynamically consistent kinetic equation between generalized force and rate are formulated. The kinetic equation includes both athermal and thermally activated contributions. It is shown that pure athermal (time-independent) kinetics can lead to contradiction. As an example a size-dependent Arrhenius-type kinetics which includes athermal threshold is suggested. The actual activation energy is determined for the whole transforming region (i.e. is proportional to transforming mass or volume) and includes additionally to the local activation energy the difference between dissipative threshold and driving force. A concept of effective temperature is introduced which takes into account the fact that temperature can vary significantly during the SC.
5. Extremum principles for determination of all unknown parameters (position, shape and orientation of nucleus, variation of all fields in it and so on) are derived using the postulate of realizability. In particular, a principle of the minimum of transformation time is obtained. When the local driving force is independent of the geometry of nucleus, a principle of the minimum of transformation time is reduced to the principle of minimum transforming mass. If a new nucleus does not touch the previous ones, its minimum characteristic size (radius, side of cube) is determined from the thermodynamic criterion of SC which leads to the kinetic concept of a thermodynamically admissible nucleus. We do not operate with the well-known critical nucleus. Appearance of the critical nucleus contradicts to the second law of thermodynamics. The kinetic description only allows us to determine the concrete finite size of the nucleus, while a pure thermodynamic consideration leads to an infinite size of nucleus. If new nucleus touches the previous one (in this way we simulate its growth), then change in surface energy and minimum characteristic size of nucleus are much smaller and size is often limited by interatomic distance. Consequently, the theory predicts accelerated growth kinetics after nucleation due to decrease in transforming volume during each event and in particular transition to atom by atom growth. This is illustrated by the solution to the problem of the appearance and growth of cubic nucleus. The result is obtained that the prescription of the local kinetic equation for each point of the moving interface is not always correct.

A generalization of the theory for the case with large deformations and displacement discontinuity across the interface can be made in a similar way as for PT (Levitas, 1998a). A more detailed model of an interface has to be developed and included in the general description. A number of examples, related to displacive, diffusional-displacive and second order PT, strain-induced chemical reactions and ductile fracture, will be considered in Part II of the paper.

Acknowledgements

The support of Volkswagen Foundation and German Research Society during the work of the author at the Institute of Structural and Computational Mechanics (University of Hannover, Germany), support of Texas Tech University, as well as discussions with Prof. Erwin Stein and Dr. Alexandr Idesman are gratefully acknowledged.

References

Abeyaratne, R., Knowles, J.K., 1993. A continuum model of a thermoelastic solid capable of undergoing phase transitions. J. Mech. Phys. Solids 44, 541–571.

Argon, A.S., 1987. Brittle to ductile transition in cleavage fracture. Acta Metall. 35, 185–196.

Atluri, S.N., 1997. Structural Integrity & Durability. Tech Science Press, Forsyth.

Boiko, V.S., Garber, R.I., Kosevich, A.M., 1991. Reversible Plasticity of Crystals. Nauka, Moscow.

Bulatov, V.V., Argon, A.S., 1994. A stochastic model for continuum elasto-plastic behaviour: I. Numerical approach and strain localization. Modelling Simul. Mater. Sci. Eng. 2, 167–184.

Cherkaoui, M., Berveiller, M., Sabar, 1998. Micromechanical modeling of the martensitic transformation induced plasticity (TRIP) in austenitic single crystals. Int. J. Plasticity 14, 597–626.

Cherepanov, G.P., 1967. On crack propagation in continuous media. Appl. Math. and Mech. (PMM) 31 (3), 476–488.

Cherepanov, G.P., 1979. Mechanics of Brittle Fracture. McGraw Hill, New York.

Cherepanov, G.P., 1997. Methods of Fracture Mechanics: Solid Matter Physics. Kluwer Academic Publishers, Dordrecht.

Christian, J.W., 1965. The Theory of Transformation in Metals and Alloys. Pergamon Press, Oxford.

Eshelby, J.D., 1951. The force on an elastic singularity. Phil. Trans. Roy. Soc. A. 244, 87–111.

Eshelby, J.D., 1956. The continuum theory of lattice defects. Solid State Physics 3, 79–144.

Eshelby, J.D., 1957. The determination of the elastic field of an ellipsoidal inclusion, and related problems. Proceedings of the Royal Society A 241, 376–396.

Eshelby, J.D., 1970. Energy relations and the energy-momentum tensor in continuum mechanics. In: Kanninen, M.F. (Ed.), Inelastic Behaviour of Solids. McGraw Hill, New York, pp. 77–115.

Fischer, F.D., Berveiller, M., Tanaka, K., Oberaigner, E., 1994. Continuum mechanical aspects of phase transformations in solids. Arch. Appl. Mech. 64, 54–85.

Fischer, F.D., Reisner, G., 1998. A criterion for the martensitic transformation of a microregion in an elastic–plastic material. Acta Mater. 46, 2095–2102.

Ghosh, G., Olson, G.B., 1994. Kinetics of f.c.c.→b.c.c. heterogeneous martensitic nucleation — I. The critical driving force for athermal nucleation. Acta Metal. Mater. 42 (10), 3361–3370.

Grinfeld, M.A., 1991. Thermodynamic Methods in the Theory of Heterogeneous Systems. Longman, Sussex.

Haezebrouck, D.M., 1987. Nucleation and Growth of a Single Martensitic Particle. Northwestern University, Materials Research Center, Steel Research Group, Evanston

Harrison, J.D., 1980. State-of-the-Art in Crack Tip Opening Displacement Testing and Analysis. The Welding Institute, Abington.

Hutchinson, J.W., 1983. Fundamentals of the phenomenological theory of nonlinear fracture mechanics. Trans. ASME, ser. E., J. Appl. Mech. 50, 1042–1051.

Idesman, A.V., Levitas, V.I., Stein, E., 1997. Simulation of martensitic phase transition progress with continuous and discontinuous displacements at the interface. Computational Materials Science 9 (1–2), 64–75.

Idesman, A.V., Levitas, V.I., Stein, E., 1999. Elastoplastic materials with martensitic phase transition and twinning at finite strains: numerical solution with the finite element method. Comp. Meth. in Appl. Mech. and Eng. 173 (1–2), 71–98.

Kaganova, I.M., Roitburd, A.L., 1988. Equilibrium of elastically interacting phases. Sov. Physics JETP 67, 1174–1186.

Kaganova, I.M., Roitburd, A.L., 1989. Effect of plastic deformation on the equilibrium shape of a new phase inclusion and thermodynamic hysteresis. Sov. Phys. Solid State 31, 545–550.

Klassen-Neklyudova, M.V., 1960. Mechanical Twinning of Crystals. Academy of Science of the USSR, Moscow.

Kocks, U.F., Argon, A.S., Ashby, M. F., 1975. Thermodynamics and Kinetics of Slip. In: Chalmers, B., Christian, J.W., Massalski, T.B. (Eds.), Progress in Materials Science, vol. 19. Pergamon Press, Oxford.

Kolednik, O., Shan, G., Fischer, D.F., 1997. The energy dissipation rate — a new tool to interpret geometry and size effects. In: Piascik, J., Newman, C., Dowling, N.E. (Eds.), Fatigue and Fracture Mechanics, vol. 27 (ASTM STP 1296). pp. 126–151.

Krasko, G.L., Olson, G.B., 1989. Energetics of bcc-fcc lattice deformation in iron. Physical Review B 40 (17), 11536–11545.

Kuang, J.H., Chen, Y.C., 1997. The tip plastic strain energy applied to ductile fracture initiation under mixed-mode loading. Engineering Fracture Mechanics 58, 61–70.

Levitas, V.I., 1992a. Post-bifurcation behaviour in finite elastoplasticity. Applications to Strain Localization and Phase Transitions. Universität Hannover. Institut für Baumechanik und Numerische Mechanik, IBNM-Bericht 92/5, Hannover.

Levitas, V.I., 1992b. Laws of thermodynamics for a finite volume of microheterogeneous medium. Mechanics of Solid 27, 37–45.

Levitas, V.I., 1995a. The postulate of realizability: formulation and applications to post-bifurcation behaviour and phase transitions in elastoplastic materials. Part I and II. Int. J. Eng. Sci. 33, 921–971.

Levitas, V.I., 1995b. Thermomechanics of martensitic phase transitions in elastoplastic materials. Mech. Res. Commun. 22, 87–94.

Levitas, V.I., 1996. Phase transitions in inelastic materials at finite strains: a local description. J. de Physique IV, Colloque C1, supplément au J. de Physique III, 6, 55–64.

Levitas, V.I., 1997. Phase transitions in elastoplastic materials: continuum thermomechanical theory and examples of control. J. Mech. Phys. Solids, Part I: 45(6), 923–947; Part II: 45(7), 1203–1222.

Levitas, V.I., 1998a. Thermomechanical theory of martensitic phase transformations in inelastic materials. Int. J. Solids and Structures 35 (9–10), 889–940.

Levitas, V.I., 1998b. Thermomechanics and kinetics of generalized second-order phase transitions in inelastic materials. Application to ductile fracture. Mech. Res. Commun. 25 (4), 427–436.

Levitas, V.I., 1998c. A new look at the problem of plastic spin based on stability analysis. J. Mech. Phys. Solids 46, 557–590.

Levitas, V.I., 1999. General thermomechanical and kinetic approach to structural changes in inelastic material. In: Khan, A. (Ed.), Constitutive and Damage Modeling of Inelastic Deformation and Phase Transformation. Proceedings of Plasticity '99. Neat Press Fulton, Maryland, pp. 235–238.

Levitas, V.I., Stein, E., Lengnick, M., 1996. On a unified approach to the description of phase transitions and strain localization. Arch. Appl. Mech. 66, 242–254.

Levitas, V.I., Idesman, A.V., Stein, E., 1998a. Finite element simulation of martensitic phase transitions in elastoplastic materials. Int. J. Solids and Structures 35 (9–10), 855–887.

Levitas, V.I., Nesterenko, V.F., Meyers, M.A., 1998b. Strain-induced structural changes and chemical reactions. I. Thermomechanical and kinetic models. Acta Materialia 46 (16), 5929–5946.

Levitas, V.I., Nesterenko, V.F., Meyers, M.A., 1998c. Strain-induced structural changes and chemical reactions. II. Modeling of reactions in shear band. Acta Materialia 46 (16), 5947–5963.

Levitas, V.I., Idesman, A.V., Olson, G.B., 1999. Continuum modeling of strain-induced martensitic transformation at shear-band intersections. Acta Materialia 47 (1), 219–233.

Maugin, G.A., 1979. Material Inhomogeneities in Elasticity. Chapman and Hall, London.
Maugin, G.A., 1995. Material forces: Concepts and applications. Applied Mechanics Reviews 48 (5), 213–245.
McClintock, F.A., 1971. Plasticity aspects of fracture. Fracture. In: Liebowitz, H. (Ed.) An Advanced Treatise, vol. 3. Academic Press, NY, London.
Olson, G.B., Cohen, M., 1986. Dislocation theory of martensitic transformations. In: Nabarro, F.R.N. (Ed.), Dislocations in Solids. Elsevier Science Publishers, Amsterdam, pp. 297–407.
Rice, J.R., 1968a. Mathematical analysis in the mechanics of fracture. Fracture. In: Liebowitz, H. (Ed.), An Advanced Treatise. Academic Press, NY, London, pp. 191–311.
Rice, J.R., 1968b. A path independent integral and the approximate analysis of strain concentration by notches and cracks. Journal of Applied Mechanics 35, 379–386.
Rice, J.R., 1992. Dislocation nucleation from a crack tip: an analysis based on the Peierls concept. J. Mech. Phys. Solids 40, 239–271.
Roitburd, A.L., Temkin, D.E., 1986. Plastic deformation and thermodynamic hysteresis at phase transformations in solids. Sov. Phys. Solid State 28, 432–436.
Turner, C.E., Kolednik, O., 1997. A simple test method for energy dissipation rate, CTOA and the study of size and transferability effects for large amounts of ductile crack growth. Fatigue Fract. Engng. Mater. Struct. 20, 1507–1528.
Volmer, M., 1983. Zur Kinetik der Phasenbildung und der Elektrodengeneration: 8 Arbeiten von Max Volmer. Leipzig, Geest und Portig.
Zeldovitch, Ya.B., 1943. To the theory of new phase formation. Cavitation. Acta Physicochim. USSR 18, 1–22.
Ziegler, H., 1997. An Introduction to Thermomechanics. North Holland, Amsterdam.