Arch. Rational Mech. Anal. (2004)
Digital Object Identifier (DOI) 10.1007/s00205-003-0286-1

Stored Energy Functions for Phase Transitions in Crystals

JOHANNES ZIMMER

Communicated by K. BHATTACHARYA

Abstract

A method is presented to construct nonconvex free energies that are invariant under a symmetry group. Algebraic and geometric methods are used to determine invariant functions with the right location of minimizers. The methods are illustrated for symmetry-breaking martensitic phase transformations. Computer algebra is used to compute a basis of the corresponding class of invariant functions. Several phase transitions, such as cubic-to-orthorhombic, are discussed. An explicit example of an energy for the cubic-to-tetragonal phase transition is given.

1. Introduction

This article is concerned with the construction of nonconvex energy functions that are invariant under a symmetry group. The applications we have in mind are symmetry-breaking martensitic transformations occurring in active crystalline materials, such as shape-memory alloys. Yet, the methods described in this paper are applicable to a large class of other situations where symmetry breaking occurs. The approach works both for first- and second- order transitions.

In the framework of nonlinear thermoelasticity, these phase transitions are frequently described as changes in the Helmholtz free-energy density. The most common approach to finding energy functions describing transformations in crystals goes back to LANDAU [16]. Here, the temperature- and strain-dependent energy function is assumed to be analytic and therefore expanded into a power series. In practice, it is customary to use an ansatz where the energy function is polynomial in the components of the strain. For the different crystal classes, SMITH & RIVLIN [25] determine a basis for the invariant polynomials under the respective symmetry group. In Section 4, as an example of some results of group theory collected in Section 3, we show that such a polynomial basis can nowadays be easily computed by means of computer algebra. The energy functions are taken to be polynomials of an invariant polynomial basis.

The latter approach has been successfully applied to martensitic transformations in crystalline solids, see, e.g., FALK [9] and the references therein, or the monograph [18]. Yet, when using this approach, some difficulties have to be faced. The restriction to polynomials considerably reduces the degrees of freedom. By Weierstraß' theorem, polynomials approximate C^∞ functions arbitrarily closely; however, polynomials become much more rigid objects if we require them to have minima and maxima only at given points. The attempt to use a polynomial of the lowest feasible order often leads to unsatisfactory results, in particular in several space dimensions: since the minima have to concentrate on the few positions corresponding to energetically stable crystalline configurations, they tend to be degenerate. This makes it hard to fit elastic moduli, and the energy wells often become too shallow to be distinguishable in a finite-element simulation.

In this paper, we will describe an alternative approach to deriving energy functions with a given symmetry. For some transitions, like the cubic-to-tetragonal transformation, this yields a comparatively easy way to find a large number of energy functions describing this transformation. In particular, the method is to some extent intuitive since it relies on geometric arguments, and removes all constraints coming from symmetry. Sometimes, as for the cubic-to-orthorhombic transformation, an additional algebraic argument is needed to find energy functions with exactly the right number of minimizers. In Section 4, it is shown how these obstacles can be be overcome for the cubic-to-orthorhombic transition. The method presented there will also work for other symmetry breakings where geometric arguments are not enough. In Section 4, we also briefly discuss all possible symmetry breakings from a cubic parent phase. An application of this method, applied to zirconia as a material with a triple point, can be found in [7].

This paper is organized as follows: in Section 2, we briefly review some elementary facts from continuum mechanics used in the subsequent sections. Section 3 contains the algebraic framework. There, the orbit space is introduced and characterized in several ways. The orbit space will be the fundamental object used to describe symmetric functions. In Section 4, martensitic phase transitions are considered as an application. Advantages and limitations of this approach are discussed in Section 5.

2. Martensitic phase transitions

The mathematical framework of Section 3 will be applied to martensitic phase transitions in Section 4. For the convenience of the reader, we will briefly collect some essential features of these transformations. Certain crystalline solids, such as CuAlNi, are able to undergo a diffusionless temperature- or stress-induced phase transformation. A typical example is the cubic-to-tetragonal transformation. At high temperature, the cubic phase is stable in the stress-free configuration. Below a critical temperature, three tetragonal variants become stable. Typically, the macroscopic behavior of these materials is modeled within the framework of nonlinear elasticity. Let $\Omega \subseteq \mathbb{R}^n$ be a reference configuration. The energy is a function of the temperature θ and the gradient of the deformation $u\colon \Omega \to \mathbb{R}^n$. The deformation gradient is

injective (to prevent self-penetration of matter) and orientation preserving. Hence, if u is differentiable in $x \in \Omega$, $F(x) := \mathrm{grad}(u(x))$ has a positive determinant. Consequently, we seek an explicit expression for the free energy (density) Φ:

$$\mathrm{GL}(n, \mathbb{R})^+ \times \mathbb{R}^+ \to \mathbb{R},$$
$$(F, \theta) \mapsto \Phi(F, \theta)$$

($\mathrm{GL}(n, \mathbb{R})^+$ is the set of all real invertible matrices of dimension $n \times n$ with positive determinant).

Let $\mathcal{P}$ denote the point group that describes the symmetry of the crystal. For example, for a cubic crystal, $\mathcal{P}$ is the group of orientation-preserving self-mappings of the cube. Material symmetry requires that

$$\Phi(FP^{-1}, \theta) = \Phi(F, \theta) \text{ for every } P \in \mathcal{P}, F \in \mathrm{GL}(n, \mathbb{R})^+, \theta \in \mathbb{R}^+. \quad (1)$$

In what follows, we will assume that $\mathcal{P}$ is the point group of the high-temperature phase and a subgroup of $\mathrm{SO}(n)$.

The axiom of frame-indifference requires that

$$\Phi(SF, \theta) = \Phi(F, \theta) \text{ for every } S \in \mathrm{SO}(n), F \in \mathrm{GL}(n, \mathbb{R})^+, \theta \in \mathbb{R}^+.$$

It is well known that this implies that Φ is a function of the left Cauchy-Green strain tensor $C := \sqrt{F^* F}$ (F^* is the transpose of F). Since there is no danger of confusion, the energy function defined on C will also be denoted by Φ. Let $\mathrm{Sym}(n, \mathbb{R})^+$ denote the set of symmetric matrices with positive determinant. Thus, we can reformulate (1) as

$$\Phi(CP^{-1}, \theta) = \Phi(C, \theta) \text{ for every } P \in \mathcal{P}, C \in \mathrm{Sym}(n, \mathbb{R})^+, \theta \in \mathbb{R}^+.$$

Since $\mathcal{P}$ is a subgroup of $\mathrm{SO}(n)$, this is equivalent to

$$\Phi(PCP^{-1}, \theta) = \Phi(C, \theta) \text{ for every } P \in \mathcal{P}, C \in \mathrm{Sym}(n, \mathbb{R})^+, \theta \in \mathbb{R}^+. \quad (2)$$

We will use this formulation since, from a group-theoretical point of view, conjugation is the natural way for a group to act on a set.

3. Review of polynomial invariant theory

In this section, we will present the algebraic background of the method. A more detailed treatment can be found, e.g., in [26] and [6]. For the more applied reader, we start with an overview of this section where the results are motivated by, and related to, symmetries of crystals. The isotropy (sub)groups (Definition 2) are the different crystalline (sub-) symmetries; Proposition 1 states that they form a lattice. Section 3.2 deals with invariant polynomials. This is partially related to the work of SMITH & RIVLIN [25]. Here, we explain how to find a generating set of invariant polynomials for a given symmetry (Theorems 1 and 2). Definition 4 introduces the orbit space, which is a central notion in this paper. Essentially, the orbit space collapses all symmetry-related points to one point and therefore strips

off the symmetry, making it easy to define an energy. A reader interested only in the applications can skip Sections 3.3 and 3.4 and move on directly to Section 3.5, where a geometric description of the orbit space is given.

The mathematical framework can be stated as follows. Let V be an n-dimensional real vector space, and suppose $\Gamma \in \mathrm{GL}(V)$ is a finite matrix group. In Section 4, V will be the strain space (space of symmetric matrices), and Γ will be the the point group $\mathcal{P}$. The order of the group is denoted by $|\Gamma|$; Greek lower-case letters stand for elements of Γ. In particular, ε is the neutral element of Γ.

With obvious changes, all statements in this section also hold true for a compact Lie group Γ.

3.1. Background from elementary bifurcation theory

We start with some basic definitions.

Definition 1. The (Γ-)*orbit* of $v \in V$ is the set $\Gamma v := \{\gamma v \mid \gamma \in \Gamma\}$. The set of all orbits is called the *orbit space* and is denoted by $V/\Gamma := \{\Gamma v \mid v \in V\}$.

The importance of orbits and the orbit space stems from (2): we want to satisfy a symmetry requirement for $V := \mathrm{Sym}(n, \mathbb{R})$ and $\Gamma := \mathcal{P}$; but this requirement states exactly that the energy function Φ is constant on the orbits. Therefore, we will use the orbit space to define it. Since the orbit space is a quotient space, it is equipped with the quotient topology.

Definition 2. The *isotropy* (*sub*)*group* Γ_v of a point $v \in V$ is the subgroup fixing v,

$$\Gamma_v := \{\gamma \in \Gamma \mid \gamma(v) = v\}.$$

Two isotropy subgroups Γ_v, Γ_w are conjugated if there is a $\gamma \in \Gamma$ such that $\Gamma_w = \gamma \Gamma_v \gamma^{-1}$. The equivalence class of Γ_v is denoted by $[\Gamma_v]$. We omit the brackets $[\cdot]$ if the class contains only one element.

If Γ is the symmetry group of the high-symmetry phase, then the isotropy groups correspond to possible low-symmetry phases.

A short calculation shows that $\Gamma_{\gamma v} = \gamma \Gamma_v \gamma^{-1}$. Therefore, two points v, γv on the same orbit have conjugate isotropy subgroups. The set of conjugacy classes is finite (this is trivial for finite groups; for compact Lie groups, see [14, I.1.4]). Combining the last observations, we see that the set of equivalence classes of isotropy classes is a Γ-invariant finite decomposition of V. We partially order conjugacy classes by saying that $[\Gamma_v] \leqq [\Gamma_w]$ if Γ_v is conjugate to a subgroup of Γ_w. We obtain the so-called *isotropy lattice of* Γ. The proof of the following statements can, for example, be found in [21]; see also [10]. A locally closed set is a set that is the intersection of an open and a closed set.

Proposition 1. *For equivalence classes of isotropy groups, the following holds true:*

(i) *Every* $[\Gamma_v]$ *is a locally closed smooth* (C^∞) *manifold.*

(ii) *If* $[\Gamma_v] \neq [\Gamma_w]$ *and* $[\Gamma_v] \cap \overline{[\Gamma_w]} \neq \emptyset$, *then* $[\Gamma_v] \subseteq \overline{[\Gamma_w]}$, $\dim[\Gamma_v] < \dim[\Gamma_w]$ *and* $[\Gamma_v] < [\Gamma_w]$.

(iii) *With respect to the given partial ordering, there is a uniquely determined maximal equivalence class that is open and dense in* V.

3.2. Algebra of invariant polynomials

Definition 3. We will first consider symmetric functions built of polynomials. To formulate the mathematical framework, we will need the trivial fact that polynomials form a ring. Let $\mathbb{R}[V]$ denote the ring of all polynomials with real coefficients in n variables. The subring of Γ-invariant polynomials on V is $\mathbb{R}[V]^\Gamma$,

$$\mathbb{R}[V]^\Gamma := \{p \in \mathbb{R}[V] \mid p(\gamma(x)) = p(x) \text{ for all } \gamma \in \Gamma, x \in V.\}$$

We will first construct energy functions in $\mathbb{R}[V]^\Gamma$ and extend later our results to the space of symmetric smooth functions. Energy functions for phase transitions can be characterized as symmetric functions with (usually parameter-dependent) minimizers at well-defined positions. The first observation is that it is easy to symmetrize a given function: The *Reynolds operator* is defined as

$$\begin{aligned} *\colon \mathbb{R}[V] &\to \mathbb{R}[V]^\Gamma, \\ p \mapsto p^* &:= \tfrac{1}{|\Gamma|} \sum_{\gamma \in \Gamma} p \circ \gamma \end{aligned}$$

(for compact Lie groups, it is necessary to use the Haar integral instead of the finite sum). Therefore, the Reynolds operator applied to an arbitrary polynomial p returns a polynomial p^* with the given symmetry. But in general, there is no relation between the minimizers of p and those of p^*: the Reynolds operator may both destroy and create minimizers. Apart from that, for applications such as those in Section 4, it is difficult to find minimizers of p^* since p^* will be a high-order polynomial of several variables. Nevertheless, there are methods for finding minimizers of p^* by exploiting the symmetry. The paper of WORFOLK [28] is a clear introduction to this subject. Other references on minimizers of invariant functions are [15, 11].

A proof of the following theorem can, for example, be found in [26, Theorem 2.1.3] or the classical reference [27].

Theorem 1 (Hilbert)**.** *The invariant ring $\mathbb{R}[V]^\Gamma$ of a finite matrix group Γ is finitely generated. That is, there are finitely many polynomials $\rho_1, \dots, \rho_k \in \mathbb{R}[V]^\Gamma$, such that every $p \in \mathbb{R}[V]^\Gamma$ can be written as $p = P(\rho_1, \dots, \rho_k)$, where $P \in \mathbb{R}[x_1, \dots, x_k]$ is a polynomial in k variables.*

In Landau theory, the space of polynomials is considered as an infinite-dimensional vector space; the aim is to find energy functions in a finite-dimensional subspace. Theorem 1 is the reason why we use another approach. By considering the set of polynomials as an algebra instead of as a vector space, we automatically get a finite basis. Furthermore, as we will see, such a basis can easily be computed.

In 1916, Emmy Noether gave a version of the Finiteness Theorem 1 which allows us to give an explicit set of polynomials generating $\mathbb{R}[V]^\Gamma$.

Theorem 2 (Noether's Degree Bound)**.** *If $\dim(V) = n$, then the invariant ring $\mathbb{R}[V]^\Gamma$ is generated as an algebra by at most $\binom{n+|\Gamma|}{n}$ invariants whose degrees are bounded above by $|\Gamma|$.*

This theorem enables us to compute an algebra basis of $\mathbb{R}[V]^\Gamma$ since $\mathbb{R}[V]^\Gamma$ is a graded algebra, i.e., it is the direct sum of the finite-dimensional vector spaces $\mathbb{R}[V]^\Gamma_d$, where $\mathbb{R}[V]^\Gamma_d$ is the set of all homogeneous invariants of degree d. It is easy to obtain a basis of $\mathbb{R}[V]^\Gamma_d$ by averaging over the space of homogeneous polynomials of degree d by applying the Reynolds operator to a basis of that space. In fact, more sophisticated techniques exist. Since there are many implementations of algorithms to find an algebra basis of $\mathbb{R}[V]^\Gamma$, we will not go into detail, but consider these algorithms as a black box. The interested reader is referred to [26, 6] for more information about Gröbner bases, which provide the main technical tool for this.

Definition 4. Any finite set of generators $\{\rho_1, \dots, \rho_k\}$ of the $\mathbb{R}$-algebra $\mathbb{R}[V]^\Gamma$ is a *Hilbert basis*. The *Hilbert map* ρ is defined as

$$\begin{aligned} V &\to \mathbb{R}^k, \\ v &\mapsto (\rho_1(v), \dots, \rho_k(v)). \end{aligned}$$

In some sense, the image of V under the Hilbert map is similar to the more familiar notion of a fundamental domain. A *fundamental domain* contains exactly one element of every orbit. The Hilbert map ρ, however, is constant on the orbits of Γ. The next lemma shows that the Hilbert map provides an embedding of the orbit space in a vector space:

Lemma 1. *A Hilbert map $\rho\colon V \to \mathbb{R}^k$ is Γ-invariant and separates the orbits. That is, for $v, w \in V$, $\rho(v) = \rho(w)$ holds if and only if v and w are on the same orbit. The map $\overline{\rho}\colon V/\Gamma \to \mathbb{R}^k$ is hence a bijection between the orbit space and $\rho(V)$.*

Proof. This is taken from [21, 2.5.1]; see also [24]. For the reader's convenience, we repeat the argument. For two different orbits Γv, Γw, there exists a continuous function $f\colon V \to \mathbb{R}$ such that $f\big|_{\Gamma v} = 0$ and $f\big|_{\Gamma w} = 1$. Obviously, this holds true for the symmetrized function f^* as well. Hence, we can assume without loss of generality that f is symmetric. Since polynomials are dense in the space of continuous functions, and since the Reynolds operator is continuous, it is possible to find an invariant polynomial p such that $p\big|_{\Gamma v} < \frac{1}{2}$ and $p\big|_{\Gamma w} > \frac{1}{2}$. Since p can be expressed in terms of the Hilbert basis, we have $\rho(v) \neq \rho(w)$. □

We can try to determine a fundamental domain for a given action of a group on a space and define the energy function on the fundamental domain (see [5] for such an approach for $\Gamma = \mathrm{GL}(2, \mathbb{Z})$). The difficulty is that the fundamental domain is usually neither open nor closed: some parts of the boundary belong to the fundamental domain, others do not. Therefore, we have to proceed with care when defining a function on the fundamental domain. The embedding in $\mathbb{R}^k$ provided by the Hilbert map avoids this problem.

Since the orbit space V/Γ is a locally compact Hausdorff space, the Hilbert map ρ induces a homeomorphism of V/Γ and the image $\rho(V)$ [24].

3.3. The Cohen-Macaulay property

Later on, we will use the fact that rings of invariants are Cohen-Macaulay. Therefore, the minimal background is gathered here; further information can be found in [3, 8].

A ring $R \subset \mathbb{R}[x_1, \ldots, x_n]$ of polynomials has the *Cohen-Macaulay property* if there are l polynomials $\eta_1, \ldots, \eta_l \in R$ such that

$$R = \bigoplus_{j=1}^{l} \eta_j \mathbb{R}[\rho_1, \ldots, \rho_n]. \tag{3}$$

[1]

In this case, in the notation used above, $\rho_1, \ldots, \rho_n$ are *primary invariants* and $\eta_1, \ldots, \eta_l$ are *secondary invariants*. The following theorem is due to HOCHSTER & EAGON [13].

Theorem 3. *The invariant ring $\mathbb{R}[V]^\Gamma$ of a finite matrix group is Cohen-Macaulay.*

3.4. Relations between polynomials

As mentioned before, we will use the orbit space to define energy functions. To do so, we need to know its structure. In this subsection, we discuss one possibility of determining the geometry of the orbit space. Eventually, this method results in a set of equations defining the orbit space. Using version 2.0 of Singular [12], it was possible to determine these equations for the phase transitions considered in Section 4. The resulting equations, however, needed several MB of storage space. Therefore, the methods discussed in this subsection will not be used in the below.

A Hilbert basis is usually not algebraically independent: in general, there are nonzero polynomials $r \in \mathbb{R}[x_1, \ldots, x_k]$ such that $r(\rho_1, \ldots, \rho_k)$ is identically zero. Such a polynomial is a (*nontrivial*) *relation* among $\rho_1, \ldots, \rho_k$. The set of all relations among $\rho_1, \ldots, \rho_k$ is an ideal; it is called the *ideal of relations* or *syzygy ideal* and denoted by I_ρ:

$$I_\rho := \{r \in \mathbb{R}[x_1, \ldots, x_k] \mid r(\rho_1, \ldots, \rho_k) = 0 \text{ in } \mathbb{R}[V]\}.$$

Let $V(I_\rho)$ denote the set of common zeros of all elements of I_ρ. By definition, $V(I_\rho)$ is an algebraic subset of V (a variety in the language of algebraic geometry), and $\rho(V) \subseteq V(I_\rho)$. This can be improved as in the next proposition, which is taken from [6, p. 337].

Proposition 2. *Let I_ρ and $Z := V_\rho$ be as above. Let $\mathbb{R}[Z]$ be the restriction of $\mathbb{R}[V]$ to Z. Then*

[1] The formal definition is as follows: A set $\rho_1, \ldots, \rho_n$ of n homogeneous invariant polynomials is said to be a *homogeneous system of parameters* (*h.s.o.p.*) if $\mathbb{R}[V]^\Gamma$ is a finitely generated module over its subring $\mathbb{R}[\rho_1, \ldots, \rho_n]$. A ring R is *Cohen-Macaulay* if it is generated as a *free* module over the ring of any h.s.o.p..

(i) *V_ρ is the smallest variety in $\mathbb{R}^k$ containing the orbit space $\rho(V)$;*
(ii) $\mathbb{R}[Z] \simeq \mathbb{R}[V]^\Gamma$.

Part (ii) of this Proposition shows that V_ρ is essentially determined by $\mathbb{R}[V]^\Gamma$, which means we can use V_ρ to describe invariant polynomials if V_ρ can be understood in detail.

3.5. Inequalities defining the orbit space

We will use the fact that the orbit space can be described as a set defined by inequalities (in terms of algebraic geometry, this means it is a semi-algebraic variety).

By introducing coordinates if necessary, we can assume $V = \mathbb{R}^n$. Let J be the Jacobian matrix of ρ,

$$J := \left(\frac{\partial \rho_j}{\partial v_k} \right)_{j,k=1,\ldots,n}.$$

Let us take all pointwise inner products of $\operatorname{grad}(\rho_j)$ and $\operatorname{grad}(\rho_k)$ for $j, k = 1, \ldots, n$. Then for $G := JJ^*$, we have

$$(G)_{j,k} = \left\langle \operatorname{grad}\left(\rho_j\right), \operatorname{grad}\left(\rho_k\right) \right\rangle,$$

where $\langle \cdot, \cdot \rangle$ is the inner product in $\mathbb{R}^n$.

Theorem 4. *The orbit space is given by $\rho(V) = \{z \in Z \mid G(z)$ is positive semidefinite$\}$.*

This has been conjectured by ABUD & SARTORI [1, 2]. See [19] for the proof; applications of this theorem to gauge symmetry breaking are given in [20].

Having several descriptions of the orbit space at our disposal, it is in principle easy to define energy functions: by defining them on the orbit space, we do not have to take symmetry constraints into account (similar to the construction using a fundamental domain). The main point is now to determine the position of the minimizers. In the next section, this is carried out for an example.

4. Application: Energy functions for martensitic transformations

In this section, the methods presented in Section 3 are applied to energy functions describing martensitic phase transitions. More precisely, we study symmetry breaking with a cubic high-symmetry phase. That is, we assume $\mathcal{P} = O$, where O is the group of orientation-preserving mappings of the cube to itself. In applications, this is the most important case. By (2), O acts on $C \in \operatorname{Sym}(3, \mathbb{R})^+$ by conjugation:

$$\begin{aligned} \mathcal{P} \times \operatorname{Sym}(3, \mathbb{R})^+ &\to \operatorname{Sym}(3, \mathbb{R})^+, \\ (P, C) &\mapsto PCP^{-1} \end{aligned}$$

(PCP^{-1} is symmetric since $P^* = P^{-1}$. Therefore, this really defines an action of $\mathcal{P}$ on $\mathrm{Sym}(3,\mathbb{R})^+$). Since $\mathrm{Sym}(3,\mathbb{R})^+$ is invariant in $\mathrm{Sym}(3,\mathbb{R})$ under conjugation by $\mathcal{P}$, we can extend this action to an action on $\mathrm{Sym}(3,\mathbb{R})$.

In $\mathrm{Sym}(3,\mathbb{R})$, we introduce coordinates as follows: for $(m_{jk})_{1\leqq j,k\leqq 3} \in \mathrm{Sym}(3,\mathbb{R})$, let us define

$$\begin{aligned} e_1 &:= m_{11},\ e_2 := m_{22},\ e_3 := m_{33},\\ e_4 &:= m_{12},\ e_5 := m_{23},\ e_6 := m_{13}. \end{aligned} \tag{4}$$

(From an engineering point of view, this is just Voigt's notation of the strain. From a mathematical point of view, we are defining a representation of $\mathrm{Sym}(3,\mathbb{R})$.) Since the (symmetrized) strain space is six-dimensional, we have $V := \mathbb{R}^6$.

It is easy to see that O is generated by two elements, namely

$$\alpha := \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ and } \beta := \begin{pmatrix} 0 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

(Of course, this choice is far from being unique. See [17] for a discussion of the octahedral group O.)

Theorem 5. *The ring of invariant polynomials is generated by*

$$\begin{aligned}
\rho_1 &:= e_1 + e_2 + e_3,\\
\rho_2 &:= e_1^2 + e_2^2 + e_3^2,\\
\rho_3 &:= e_4^2 + e_5^2 + e_6^2,\\
\rho_4 &:= e_1^3 + e_2^3 + e_3^3,\\
\rho_5 &:= e_4 e_5 e_6,\\
\rho_6 &:= e_4^4 + e_5^4 + e_6^4,\\
\rho_7 &:= 1,\\
\rho_8 &:= e_1 e_4^2 + e_2 e_4^2 + e_2 e_5^2 + e_3 e_5^2 + e_1 e_6^2 + e_3 e_6^2,\\
\rho_9 &:= e_1^2 e_4^2 + e_2^2 e_4^2 + e_2^2 e_5^2 + e_3^2 e_5^2 + e_1^2 e_6^2 + e_3^2 e_6^2,\\
\rho_{10} &:= e_1 e_4^4 + e_2 e_4^4 + e_2 e_5^4 + e_3 e_5^4 + e_1 e_6^4 + e_3 e_6^4,\\
\rho_{11} &:= e_1^2 e_4^4 + e_2^2 e_4^4 + e_2^2 e_5^4 + e_3^2 e_5^4 + e_1^2 e_6^4 + e_3^2 e_6^4,\\
\rho_{12} &:= e_1^2 e_2 e_4^4 e_5^2 + e_2 e_3^2 e_4^2 e_5^4 + e_1 e_2^2 e_4^4 e_6^2 + e_2^2 e_3 e_5^4 e_6^2 + e_1 e_3^2 e_4^2 e_6^4 + e_1^2 e_3 e_5^2 e_6^4.
\end{aligned}$$

Proof. This basis can be computed by the methods presented in Section 3. To avoid tedious computations, we used the software package Singular [12], version 1.0.2. The computation is based on an algorithm by STURMFELS [26]. □

The invariants given in Theorem 5 are a Cohen-Macaulay basis: the first 6 invariants $\rho_1, \ldots, \rho_6$ are primary invariants, $\rho_7, \ldots, \rho_{12}$ are secondary invariants. By (3), there exist uniquely determined polynomials $P_7, \ldots, P_{12}$ such that every invariant polynomial p can be written as

$$p = \sum_{j=7}^{12} \rho_j P_j(\rho_1, \ldots, \rho_6).$$

It should be remarked that the basis given in Theorem 5 is not a minimal basis. In fact, it is easy to give a minimal basis, consisting of 9 elements. However, we will exploit the Cohen-Macaulayness of the basis given above.

By Lemma 1, we can consider $\rho(V)$ as orbit space. Since energy functions are defined only on $\mathrm{Sym}(3,\mathbb{R})^+$ instead of $\mathrm{Sym}(3,\mathbb{R})$, we have to determine the image of $\mathrm{Sym}(3,\mathbb{R})^+$ in the orbit space $\rho(V)$.

Lemma 2. *For the given Hilbert map, the image of* $\mathrm{Sym}(3,\mathbb{R})^+$ *in the orbit space is characterized by the inequalities*

$$\begin{aligned} \rho_1 &> 0, \\ \tfrac{1}{2}\left(\rho_1^2 - \rho_2\right) + \rho_3 &> 0, \\ \tfrac{1}{6}\left(\rho_1^3 - 3\rho_1\rho_2 + 2\rho_4\right) + 2\rho_5 + \rho_8 - \rho_1\rho_3 &> 0. \end{aligned}$$

Proof. Let $M \in \mathrm{Sym}(3,\mathbb{R})^+$ be a symmetric matrix. We use Voigt's notation. Since M is positive definite, the principal minors are positive by Jacobi's criterion. These minors are

$$\begin{aligned} e_1, e_2, e_3 > 0 \quad & (1\times 1 \text{ minors}), \\ e_1e_2 - e_4^2 > 0, \quad e_1e_3 - e_6^2 > 0, \quad e_2e_3 - e_5^2 > 0 \quad & (2\times 2 \text{ minors}), \\ e_1e_2e_3 + 2e_4e_5e_6 - e_2e_6^2 - e_1e_5^2 - e_3e_4^2 > 0 \quad & (\text{determinant}). \end{aligned}$$

Summing up the 1×1 minors, we obtain

$$\rho_1 = e_1 + e_2 + e_3 > 0 \tag{5}$$

and analogously in the 2×2 case

$$e_1e_2 - e_4^2 + e_1e_3 - e_6^2 + e_2e_3 - e_5^2 = \tfrac{1}{2}\left(\rho_1^2 - \rho_2\right) - \rho_3 > 0. \tag{6}$$

A simple computation yields

$$\begin{aligned} e_1e_2e_3 + 2e_4e_5e_6 - e_2e_6^2 - e_1e_5^2 - e_3e_4^2 & \\ = \tfrac{1}{6}\left(\rho_1^3 - 3\rho_1\rho_2 + 2\rho_4\right) + 2\rho_5 + \rho_8 - \rho_1\rho_3 &> 0. \end{aligned} \tag{7}$$

This proves that $\rho(\mathrm{Sym}(3,\mathbb{R})^+)$ is contained in the set characterized by the inequalities. We have to show that it is not a strict inclusion. To do so, we express inequalities (5)–(7) in terms of eigenvalues. This gives

$$\begin{aligned} \lambda_1 + \lambda_2 + \lambda_3 &> 0, \\ \lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3 &> 0, \\ \lambda_1\lambda_2\lambda_3 &> 0. \end{aligned}$$

The last inequality shows that either all eigenvalues are positive or negative eigenvalues appear pairwise. In the latter case, possibly after relabeling the eigenvalues, we can assume $\lambda_1, \lambda_2 < 0$ and $\lambda_1 \leqq \lambda_2$. The first inequality shows $\lambda_3 \geqq -2\lambda_2$ and hence $\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3 \leqq 0$, contradicting the second inequality. □

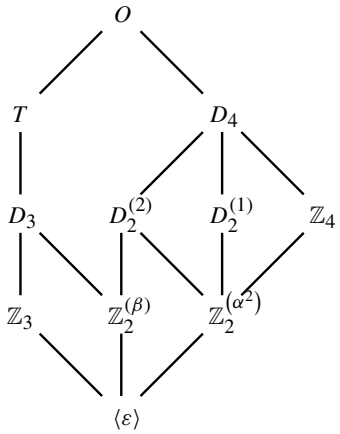

Fig. 1. Group diagram of O. The group-subgroup relations are shown by lines; $\mathbb{Z}_k$ denotes the cyclic group of order k, while D_k stands for a dihedral group. The symmetry group of a tetragon is denoted T. Upper indexes distinguish groups that are not conjugated. For $\mathbb{Z}_2$ subgroups, upper indexes also indicate the generators.

Group Γ'	Generator	$F(\Gamma') := \mathrm{Fix}(\Gamma')$	$\dim(F(\Gamma'))$	Iso.
$\mathbb{Z}_2^{(\beta)}$	β	$\{(a,b,a,c,-c,d)\}$	4	yes
$\mathbb{Z}_2^{(\alpha^2)}$	α^2	$\{(a,b,c,d,0,0)\}$	4	yes
$\mathbb{Z}_3$	$\gamma := (\beta\alpha)^{-1}$	$\{(a,a,a,b,b,-b)\}$	2	no
$\mathbb{Z}_4$	α	$\{(a,a,c,0,0,0)\}$	2	no
$D_2^{(1)}$	$\alpha^2, \beta\alpha^2\beta$	$\{(a,b,c,0,0,0)\}$	3	yes
$D_2^{(2)}$	$\beta, \alpha^2\beta\alpha^2\alpha^2$	$\{(a,b,a,0,0,c)\}$	3	yes
D_3	$\beta, \alpha\beta\alpha^{-1}$	$\{(a,a,a,b,-b,-b)\}$	2	yes
D_4	β, α^2	$\{(a,b,a,0,0,0)\}$	2	yes
T	$\gamma, \alpha^2\gamma\alpha^2$	$\{(a,a,a,0,0,0)\}$	1	no
O	α, β	$\{(a,a,a,0,0,0)\}$	1	yes

Fig. 2. Subgroups of O, their generators, their fixed-point space and the dimension of the fixed-point space. Isotropy groups are marked by "yes" in the last column.

As the next step, we determine the isotropy subgroups since they describe possible symmetry breakings. The list of isotropy subgroups can be generated automatically, e.g., by using XGAP and GAP [23]. Here, we carry out the elementary computation. Let Γ' be a subgroup of Γ. The *fixed-point space* $\mathrm{Fix}(\Gamma')$ is defined as $\mathrm{Fix}(\Gamma') := \{v \in V \mid \gamma(v) = v \text{ for all } \gamma \in \Gamma'\}$. The fixed-point space of Γ_v is the linear subspace of V containing all points with the same symmetry as v. To find all isotropy subgroups, we consider the lattice of all subgroups of O. See Fig. 1. For a given point $v \in V$, its isotropy subgroup is by definition the maximal subgroup leaving v invariant. Hence, isotropy subgroups are those groups in the group diagram whose fixed-point space has a smaller dimension than the fixed-point space of the group containing the group under consideration.

For our application, all subgroups of O, their generators, their fixed-point space and the dimension of the fixed-point space are listed in Fig. 2; isotropy groups are marked by "yes" in the last column.

Thus, we have the following symmetry types:

$$O, \quad D_4, \quad D_3, \quad \mathbb{Z}_2^{(\beta)}, \quad D_2^{(1)}, \quad D_2^{(2)}, \quad \mathbb{Z}_2^{(\alpha^2)}, \quad \langle \varepsilon \rangle.$$

Symmetry types and their images under the Hilbert map will be called *strata*.

Since $\rho(V)$ is embedded in a 12-dimensional vector space, it is difficult to visualize. To use geometric ideas as much as possible, we restrict ourselves to the projection on the primary invariants. Here, the situation is particularly easy: The representation induced by the operation of O on $\mathrm{Sym}(3, \mathbb{R})$ decomposes in three representations. Namely, the representation of the trace and one more representation on the elements e_1, e_2, e_3 representing the diagonal, and a third representation on the off-diagonals e_4, e_5, e_6. The two representations on the diagonals are orthogonal to each other, as we will see below. We will study these representations separately.

Let us start with the representation on the off-diagonals. Proceeding as before, we see that the following groups are isotropy subgroups for the restriction to the off-diagonals:

Isotropy group	O	$[D_3]$	$\left[D_2^{(2)}\right]$	$\left[\mathbb{Z}_2^{(\beta)}\right]$	$\langle \varepsilon \rangle$
dim(fixed-point space)	0	1	1	2	3

Primary invariants for the off-diagonals are

$$\begin{aligned} \rho_1^{\mathrm{o}} &:= \rho_3 = e_4^2 + e_5^2 + e_6^2, \\ \rho_2^{\mathrm{o}} &:= \rho_5 = e_4 e_5 e_6, \\ \rho_3^{\mathrm{o}} &:= \rho_6 = e_4^4 + e_5^4 + e_6^4. \end{aligned}$$

This can be seen from Theorem 5; but it is necessary to prove that this is really a basis for the off-diagonals and not just an invariant set. This can be done using the tools presented in Section 3, but there is an easier way: restricted to the off-diagonals, O is a reflection group. By a theorem of CHEVALLEY [4, Theorem (A)], for reflection groups, there is an algebraically independent Hilbert basis of n elements, where n is the dimension of the underlying vector space. This shows that for the off-diagonals, there is a basis of three algebraically independent invariants, and it is possible to convince oneself that $\rho_1^{\mathrm{o}}, \ldots, \rho_3^{\mathrm{o}}$ is such a basis.

To characterize the orbit space of the off-diagonals, we use Theorem 4. A computation of the principal minors of $\left(\left\langle \mathrm{grad}\left(\rho_i^{\mathrm{o}}\right), \mathrm{grad}\left(\rho_j^{\mathrm{o}}\right)\right\rangle\right)_{i,j}$ gives conditions for this matrix to be positive semidefinite. A short calculation shows that the following inequalities are active:

$$\begin{aligned} \rho_1^{\mathrm{o}} &\geqq 0, \\ \rho_3^{\mathrm{o}} &\leqq \rho_1^{\mathrm{o}2}, \\ \rho_1^{\mathrm{o}6} - 20\rho_1^{\mathrm{o}3}\rho_2^{\mathrm{o}2} - 4\rho_1^{\mathrm{o}4}\rho_3^{\mathrm{o}} + 36\rho_1^{\mathrm{o}}\rho_2^{\mathrm{o}3}\rho_3^{\mathrm{o}} + 5\rho_1^{\mathrm{o}2}\rho_3^{\mathrm{o}2} + 108\rho_2^{\mathrm{o}4} - 2\rho_3^{\mathrm{o}3} &\leqq 0. \end{aligned}$$

From this, we can almost immediately obtain the following geometric interpretation of this orbit space as an image of the fixed-point spaces:

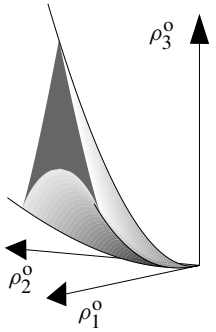

Fig. 3. Orbit space of the off-diagonals. The tip is the stratum of O, the surfaces are strata of $\left[\mathbb{Z}_2^{(\beta)}\right]$, the upper edge is the image of $\left[D_2^{(2)}\right]$, and $[D_3]$ corresponds to the two lower edges.

Stratum	Parametrization
O	$\rho_1^{\mathrm{o}} = \rho_2^{\mathrm{o}} = \rho_3^{\mathrm{o}} = 0$
$\left[D_2^{(2)}\right]$	$\rho_1^{\mathrm{o}} = t, \quad \rho_2^{\mathrm{o}} = 0, \quad \rho_3^{\mathrm{o}} = t^2$
$[D_3]$	$\rho_1^{\mathrm{o}} = t, \quad \rho_2^{\mathrm{o}} = \pm\frac{\sqrt{3}}{9}\sqrt{t^3}, \quad \rho_3^{\mathrm{o}} = \frac{1}{3}t^2$
$\left[\mathbb{Z}_2^{(\beta)}\right]$	$\rho_1^{\mathrm{o}} = t, \quad \frac{1}{3}t^2 < \rho_3^{\mathrm{o}} < t^2,$
	$\rho_2^{\mathrm{o}} =$
	$\pm\frac{1}{18}\sqrt{30t^3 - 54t\rho_3^{\mathrm{o}} \pm 6\sqrt{-2t^6 + 18t^4\rho_3^{\mathrm{o}} - 54t^2\rho_3^{\mathrm{o}2} + 54\rho_3^{\mathrm{o}3}}}$
$\langle\varepsilon\rangle$	Interior

This can be visualized as in Fig. 3.

We repeat this for the restriction of the representation to the diagonal elements e_1, e_2, e_3. This case is simpler since the representation (4) decomposes into a one-dimensional and a two-dimensional representation.

We have the following strata:

Stratum	O	$[D_4]$	$\left[D_2^{(1)}\right]$
Dimension of fixed-point space	1	2	3

A Hilbert basis of invariant polynomials is given by

$$\begin{aligned}\rho_1^{\mathrm{d}} &:= \rho_1 = e_1 + e_2 + e_3,\\ \rho_2^{\mathrm{d}} &:= \rho_2 = e_1^2 + e_2^2 + e_3^2,\\ \rho_3^{\mathrm{d}} &:= \rho_4 = e_1^3 + e_2^3 + e_3^3.\end{aligned}$$

We introduce new coordinates to use the fact that the representation decomposes. Here, s parametrizes the one-dimensional fixed-point space and t and u are chosen to be orthogonal to s. Explicitly,

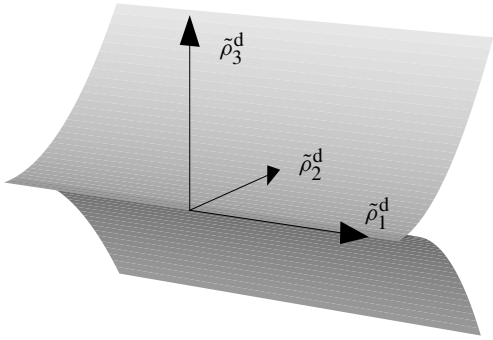

Fig. 4. Orbit space of the diagonals. The line is the stratum of O; the surfaces correspond to $[D_4]$; the interior is the image of $\left[D_2^{(1)}\right]$. The surface is cut for better inspection.

$$
\begin{aligned}
s(e_1, e_2, e_3) &:= \frac{1}{\sqrt{3}}(e_1 + e_2 + e_3),\\
t(e_1, e_2, e_3) &:= \frac{1}{\sqrt{2}}(e_1 - e_2),\\
u(e_1, e_2, e_3) &:= \sqrt{\frac{2}{3}}\left(\frac{1}{2}e_1 + \frac{1}{2}e_2 - e_3\right).
\end{aligned}
$$

A Hilbert basis with respect to the new coordinates is, for example, given by

$$\tilde{\rho}_1^{\mathrm{d}} := s = \frac{1}{\sqrt{3}}(e_1 + e_2 + e_3), \tag{8}$$

$$\tilde{\rho}_2^{\mathrm{d}} := t^2 + u^2, \tag{9}$$

$$\tilde{\rho}_3^{\mathrm{d}} := \frac{3}{\sqrt{2}}t^2u - \frac{1}{\sqrt{2}}u^3. \tag{10}$$

Proceeding as before, we arrive at

$$\left(\left\langle \operatorname{grad}\left(\tilde{\rho}_i^{\mathrm{d}}\right), \operatorname{grad}\left(\tilde{\rho}_j^{\mathrm{d}}\right)\right\rangle\right)_{i,j} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4\tilde{\rho}_2^{\mathrm{d}} & 6\tilde{\rho}_3^{\mathrm{d}} \\ 0 & 6\tilde{\rho}_3^{\mathrm{d}} & \frac{9}{2}\left(\tilde{\rho}_2^{\mathrm{d}}\right)^2 \end{pmatrix}.$$

An application of Theorem 4 provides a characterization of the orbit space as the set of all points for which $\left(\left\langle \operatorname{grad}\tilde{\rho}_i^{\mathrm{d}}, \operatorname{grad}\tilde{\rho}_j^{\mathrm{d}}\right\rangle\right)_{i,j}$ is positive semidefinite. Using minors to check positive semidefiniteness, a description of the orbit space can be stated as follows:

$$
\begin{aligned}
\tilde{\rho}_2^{\mathrm{d}} &\geqq 0,\\
\tilde{\rho}_3^{\mathrm{d}} &\leqq \pm\sqrt{\frac{1}{2}(\tilde{\rho}_2^{\mathrm{d}})^3}.
\end{aligned}
$$

The orbit space is visualized in Fig. 4.

4.1. First example: Cubic-to-tetragonal phase transition

To illustrate the ideas presented above, let us construct an energy function for the cubic-to-tetragonal phase transition. Using the usual group-theoretic notation, the cubic symmetry is denoted by O, the tetragonal symmetry is denoted by D_4. Since we studied the representations of the diagonal and the representation of the off-diagonals separately, we have to identify the groups acting on these sets corresponding to the given groups on the whole set. It is easy to see that $[D_4]$ corresponds to the Cartesian product of $[O]$ on the off-diagonals and $[D_4]$ on the diagonals (this also can be read off from the the table in Fig. 7).

We make the ansatz

$$\overline{f}(\rho_1^{\mathrm{o}}, \rho_2^{\mathrm{o}}, \rho_3^{\mathrm{o}}, \tilde{\rho}_1^{\mathrm{d}}, \tilde{\rho}_2^{\mathrm{d}}, \tilde{\rho}_3^{\mathrm{d}}, \theta) := \overline{f}_1(\rho_1^{\mathrm{o}}, \rho_2^{\mathrm{o}}, \rho_3^{\mathrm{o}}, \theta) + \overline{f}_2(\tilde{\rho}_1^{\mathrm{d}}, \tilde{\rho}_2^{\mathrm{d}}, \tilde{\rho}_3^{\mathrm{d}}, \theta)$$

and have to determine $\overline{f}_1, \overline{f}_2$. In particular, this means we ignore all secondary invariants except for the trivial $\rho_7 = 1$. The reason for this is that the orbit space $\rho(V)$ of the entire space can be shown to consist of finitely many layers that project to the primary invariants; different layers are distinguished by the secondary invariants. This is explained in more detail in the example of the cubic-to-orthorhombic transition. Here, we can ignore this subtlety: it can be shown that for the tetragonal case, there is just one layer and hence there is no need to introduce secondary invariants unless this is useful for fitting some parameters.

As seen before (Fig. 3), since the restriction of both O and D_4 on V to the off-diagonals is O, it is immediately obvious that $\overline{f}_1$ has to have a minimizer in the tip of this orbit space. A possible choice is

$$\overline{f}_1(\rho_1^{\mathrm{o}}, \rho_2^{\mathrm{o}}, \rho_3^{\mathrm{o}}, \theta) := \rho_1^{\mathrm{o}}.$$

On the diagonals, D_4 acts as D_4, and Fig. 4 shows that the strata associated with this are the surfaces. A parametrization of this surface is given by

$$\begin{aligned} \tilde{\rho}_1^{\mathrm{d}} &= v, \\ \tilde{\rho}_2^{\mathrm{d}} &= w, \\ \tilde{\rho}_3^{\mathrm{d}} &= \pm\sqrt{\frac{1}{2}w^3}. \end{aligned}$$

A natural way to define a function $\overline{f}_2$ would be to locate the positions of the cubic and the tetragonal phase (indicated in Fig. 5), and to define a (temperature dependent) function in a neighborhood of these positions. Such a function could be a quadratic polynomial, with the coefficients being determined by the elastic moduli. Splines are a natural choice to extend these functions to the entire domain. A certain care has to be taken to ensure that the splines model an energy barrier without introducing additional minimizers. Since the explicit derivation and representation of such a spline is lengthy, we will here use another approach to define the function $\overline{f}_2$.

The key idea in this example is to reduce the problem of defining $\overline{f}_2$ to an essentially one-dimensional problem, which is very well understood and where it is easy

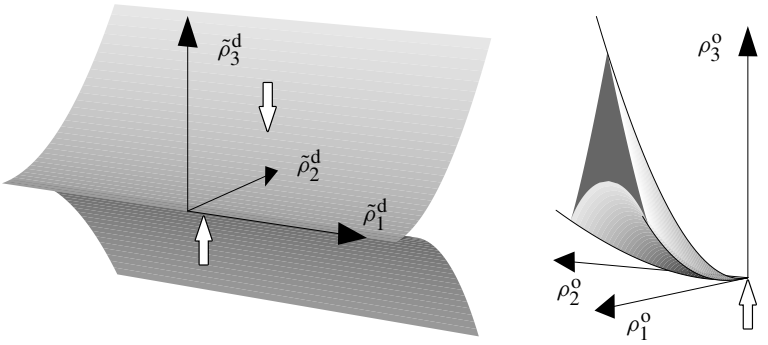

Fig. 5. The location of minimizers for a cubic phase (arrow pointing to the axis $\tilde{\rho}_1^{\mathrm{d}}$ and arrow in the right picture), and location of minimizers for a tetragonal phase (arrow pointing to the surface in the left picture and arrow in the right picture).

to construct temperature-dependent functions with wells at given positions. We first reduct it to a two-dimensional problem by defining a function $f(v) := (v - v_0)^2$, which has minimizers exactly in the plane $v = v_0$ (we need $v_0 > 0$ since, according to Lemma 2, $v_0 = 0$ is not in the image of $\mathrm{Sym}(n, \mathbb{R})^+$). To get, within this plane, minimizers only at the positions marked by the arrows in Fig. 5, we use a one-dimensional Landau energy, where we quotient out the one-dimensional symmetry. This results for first phase transitions in $g(w) := \left(\theta - \theta_c + \frac{1}{4}\right) w - \frac{1}{2}w^2 + \frac{1}{3}w^3$, while second-order phase transitions can be described by $\tilde{g}(w) := (\theta - \theta_c)\, w + \frac{1}{3}w^3$. To move the minimum to the upper surface, we introduce $h(x) := (\theta - \theta_c)x := \frac{\sqrt{2}}{2}(\theta - \theta_c)w^{3/2}$. Putting things together, a possible choice is

$$\overline{f}_2(v, w, x) := f(v) + \left\{ \begin{matrix} g(w) \\ \tilde{g}(w) \end{matrix} \right\} + h(x) \text{ for transitions of } \left\{ \begin{matrix} \text{first} \\ \text{second} \end{matrix} \right\} \text{ order.}$$

It is a calculus exercise to check that this function, defined on the orbit space, has exactly one minimizer.

The energy function is of course defined as the sum of $\overline{f}_1$ and $\overline{f}_2$. Spelled out in coordinates, it reads for a first-order transition,

$$\begin{aligned}
\Phi(e_1, \ldots, e_6, \theta) = {} & e_4^2 + e_5^2 + e_6^2 + \frac{1}{3}(e_1 + e_2 + e_3 - 1)^2 \\
& + \left(\theta - \theta_c + \frac{1}{4}\right)\left(\frac{1}{2}(e_1 - e_2)^2 + \frac{2}{3}\left(\frac{1}{2}e_1 + \frac{1}{2}e_2 - e_3\right)^2\right) \\
& - \frac{1}{2}\left(\frac{1}{2}(e_1 - e_2)^2 + \frac{2}{3}\left(\frac{1}{2}e_1 + \frac{1}{2}e_2 - e_3\right)^2\right)^2 \\
& + \frac{1}{3}\left(\frac{1}{2}(e_1 - e_2)^2 + \frac{2}{3}\left(\frac{1}{2}e_1 + \frac{1}{2}e_2 - e_3\right)^2\right)^3 \\
& + \frac{\sqrt{2}}{2}(\theta - \theta_c)\left(\frac{1}{2}(e_1 - e_2)^2 + \frac{2}{3}\left(\frac{1}{2}e_1 + \frac{1}{2}e_2 - e_3\right)^2\right)^{\frac{3}{2}}. \quad (11)
\end{aligned}$$

Of course, many other functions are possible. Fig. 5 shows again the location of minimizers of a tetragonal phase; any function with minimizers precisely at these locations would be a possible choice.

It should be pointed out that the function defining the minimizers on the orbit space is by no means restricted to the class of polynomials (indeed, the function above involves a root). This is a major difference to energy functions constructed by using the polynomial ansatz derived from Landau theory, and we believe this is useful to fit elastic moduli and other parameters. In fact, it can be show that every smooth (C^∞) function invariant under a compact Lie group can be constructed in the way described above [24]; with some modifications, this is also true for C^k functions [21, 22].

Authors using a polynomial function often neglect to verify that there are no additional minimizers. That can be a nontrivial task (see [28]). In the example discussed above, it is obvious that the function has minimizers exactly at the correct positions.

4.2. Second example: Cubic-to-orthorhombic phase transition

We want to discuss the cubic-to-orthorhombic transition as the second example, since this transition will prove to be more subtle. This can be seen as follows. The symmetry group representing the orthorhombic phase

$$\begin{pmatrix} \alpha & 0 & \gamma \\ 0 & \beta & 0 \\ 0 & 0 & \alpha \end{pmatrix}$$

with $\alpha \neq \beta$ is $D_2^{(2)}$. It is easy to see that this group acts on the diagonal elements as D_4, and as $D_2^{(2)}$ on the off-diagonals (see the table in Fig. 7).

It is easy to mimic the construction in the tetragonal case, as far as the primary invariants are concerned. Indeed, exactly the same methods used to get a minimizer in a surface of the orbit space of the diagonals can be applied to create a minimizer in the upper edge of the orbit space of the off-diagonals. This ensures that only phases represented by these positions in the two orbit spaces have minimizers at a given temperature. Fig. 6 shows the location of minimizers for an orthorhombic phase.

However, in this case, it is not enough to consider the primary invariants. We claim that the stratum of this symmetry group might be foliated over the projection on the two separate orbit spaces. The foliation could consist of up to three leaves. To see this, first compute the orbit length of $D_2^{(2)}$, given by $|\Gamma x| = \frac{|\Gamma|}{|\Gamma_x|}$. Here, $\Gamma = O$ and $\Gamma_x = D_2^{(2)}$. Consequently, the orbit length in this case is 6. Now imagine the secondary invariants were irrelevant. Then, the group would split as a cross product over the two separate spaces for the diagonals and the off-diagonals. The orbit length of the restrictions of the group to these two spaces could be computed as before; the orbit length of the group as the cross product would be the product of the orbit lengths. Here, $D_2^{(2)}$ acts on the diagonals as D_4, resulting

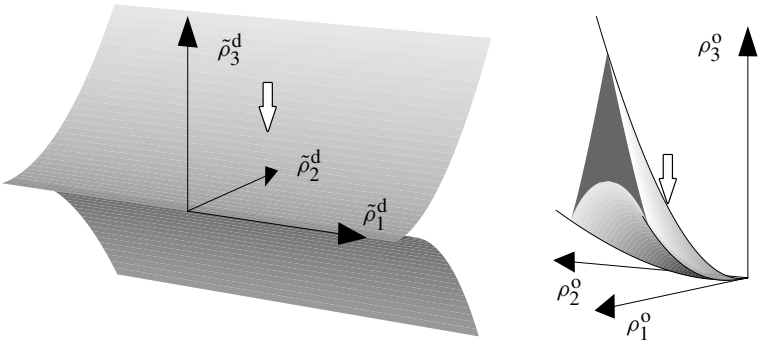

Fig. 6. The location of minimizers for an orthorhombic phase, marked by arrows.

in a first orbit length of $\frac{24}{8} = 3$. On the off-diagonals, $D_2^{(2)}$ acts as $D_2^{(2)}$, giving an orbit length of 6 as for the entire group. Therefore, the group acting as the cross product has an orbit length of 18. Since $D_2^{(2)}$ has an orbit length of 6, the secondary invariants have to distinguish points that are identified in the cross product. One way to visualize this is to say that the real orbit space consists of a number of leaves over the space generated by the primary invariants. Obviously, the quotient of the orbit length of the group acting as cross product and the group itself limits the number of leaves. Here, we find that there could be $\frac{18}{6} = 3$ leaves. If we ignore the secondary invariants, we end up with several non-conjugated minimizers, all of them projecting to the same point in the cross product of the two orbit space. How do we single out just one phase and its variants? There a comparatively simple way. First, we consider the element that should be the energetically stable configuration. Since it is assumed to be orthorhombic, it is in the fixed-point space of $D_2^{(2)}$, with coordinates $(a, b, a, 0, 0, c)$, where $a \neq b$; see Fig. 2. The orbit of this element is $(a, b, a, 0, 0, \pm c)$, $(a, a, b, \pm c, 0, 0)$, $(b, a, a, 0, \pm c, 0)$. We just computed that there are $18 - 6 = 12$ more elements that are mapped to the same point in the product of the two orbit spaces. It is easy to see that they are of the form $(a, b, a, \pm c, 0, 0)$, $(a, b, a, 0, \pm c, 0)$, $(a, a, b, 0, \pm c, 0)$, $(a, a, b, 0, 0, \pm c)$, $(b, a, a, \pm c, 0, 0)$, $(b, a, a, 0, 0, \pm c)$. We have to find secondary invariants that separate the latter elements from the former ones. A short inspection shows that any secondary invariant except the trivial $\rho_7 = 1$ and ρ_{12} does this. For example, ρ_8 maps the "right" elements, such as $(a, b, a, 0, 0, c)$, to $2ac^2$, while all "wrong" elements, for example $(a, b, a, c, 0, 0)$, are mapped to $(a + b)c^2$. The same applies to ρ_9, ρ_{10}, ρ_{11}, since only the exponents change. So, if we construct a function in ρ_8, say, that has a minimizer exactly at the position ac^2, where a and c are functions of the lattice parameters, and carry out the construction of functions on the two separate orbit spaces analogously to the cubic-to-tetragonal case and add the function defined on ρ_8, then the entire function, composed with the Hilbert map, will have exactly the right minimizers and no other minimizers.

This example is also interesting from a group theoretic point of view. In the pre-image of the point in the product space the orthorhombic phase is mapped to, there are points of orbits of different isotropy groups. Therefore, we only find one

leaf for $D_2^{(2)}$, instead of the maximal number 3 computed before. So far, we have no clear understanding as to when this phenomenon occurs – it is possible to construct similar situations, but we are not aware of a really low-dimensional example; neither are we able to give a precise formula for the number of leaves instead of an upper bound.

4.3. Geometry of the orbit space

The orbit-space method presented above provides an intuitive, since geometric, way of constructing energy functions. For a specific phase transformation, the key point in the construction is the identification of points in the strata of the orbit space corresponding to stable phases. As soon as they are found, any function defined on the orbit space with minimizers precisely in the designated spots will, composed with the Hilbert map, be an energy function with the correct symmetry breaking. Usually, we would also try to fit parameters, such as elastic moduli. There are several ways of doing this. We found it useful to define the function f on the orbit space in a piecewise manner. One way would be to define a function locally in the neighborhood N of the minimizers. If this function had no other minimizers in its domain of definition and matched the parameters prescribed for the minimizers, we could solve the Laplace equation on the complement of N in the orbit space. With a suitable choice of boundary values, this function would have no additional minimizers. See [7] for an example of this approach. Another way to define f would be to use splines or Bézier curves.

Therefore, the main remaining technical difficulty is the identification of the location of the minimizers. For the reader's convenience, we present in Fig. 7 a table collecting the relevant information. With this information at hand, the construction of the energy function can be carried out essentially as in the two examples discussed above. All data in the table is easy to compute; we use the formulas listed in Section 4.2 to compute the orbit length, the number of preimages of primary invariants and the maximal number of leaves. Displayed are the nontrivial groups (strata) Γ', their restriction Γ'^d to the diagonal, Γ'^o as the restriction to off-diagonal elements, the orbit length o of the entire group Γ', the number of preimages of the primary invariants p (= the orbit length of Γ', acting on the cross product of the space of diagonals and the space of off-diagonal), and the maximal number of leaves of this stratum, l.

Whenever a layer consists only of one leaf, the construction can be carried out analogously to the cubic-to-tetragonal case. The secondary invariants $\rho_7, \ldots, \rho_{12}$ can be ignored, and the geometric approach to work with the two separate orbit spaces will suffice. The table shows that there are only two more cases with several layers. Both cases can be treated exactly as the orthorhombic case, even using the same secondary invariant.

5. Discussion

Rather than constructing specific energy functions for martensitic phase transitions, this paper aims at presenting a general approach, using the orbit space of the

Γ'	Γ'^d	Γ'^o	o	p	l
$[D_4]$	$[D_4]$	$[O]$	3	3	1
$\left[D_2^{(1)}\right]$	$\left[D_2^{(1)}\right]$	$[O]$	6	6	1
$\left[D_2^{(2)}\right]$	$[O]$	$\left[D_2^{(2)}\right]$	6	6	1
$\left[D_2^{(2)}\right]$	$[D_4]$	$\left[D_2^{(2)}\right]$	6	18	3
$\left[\mathbb{Z}_2^{(\alpha^2)}\right]$	$\left[D_2^{(1)}\right]$	$\left[D_2^{(2)}\right]$	12	36	3
$[D_3]$	$[O]$	$[D_3]$	4	4	1
$\left[\mathbb{Z}_2^{(b)}\right]$	$[D_4]$	$[D_3]$	12	12	1
$\left[\mathbb{Z}_2^{(b)}\right]$	$[O]$	$\left[\mathbb{Z}_2^{(b)}\right]$	12	12	1
$\left[\mathbb{Z}_2^{(b)}\right]$	$[D_4]$	$\left[\mathbb{Z}_2^{(b)}\right]$	12	36	3

Fig. 7. The nontrivial groups (strata) Γ', their restriction Γ'^d to the diagonal, Γ'^o as the restriction to off-diagonal elements, the orbit length o of the entire group Γ', the number of pre-images of the primary invariants p (= the orbit length of Γ', acting on the cross product of the space of diagonals and the space of off-diagonal), and the maximal number of leaves of this stratum, l.

symmetry group. In the theory of Higgs potentials, the connection between symmetry breaking, energy functions and the orbit space has been known for some time (see the work by ABUD & SARTORI [1, 2]). For phase transitions in crystals, however, the situation is usually more complicated, since the orbit space is high-dimensional.

Computer algebra can greatly facilitate the computations. We feel that software packages such as Singular [12] or GAP [23] can ease some computations in continuum mechanics, making it unnecessary to rely on invariants given in the literature.

For the cubic-to-tetragonal transformation, the construction is based on simple geometric arguments, making it possible to construct a variety of energy functions. In particular, we are not restricted to the class of polynomials. This will facilitate the task of fitting parameters. Additionally, this degree of freedom can be used to give examples of energy functions with given growth conditions at infinity, as commonly required by existence and uniqueness theorems for equations of thermo(visco)elasticity with nonconvex energies.

The case of a cubic-to-orthorhombic transition proves to be more complicated. It requires a certain understanding of the complete orbit space in $\mathbb{R}^{12}$. However, we wanted to show that this can be achieved with comparatively simple arguments. A more systematic approach would be to choose invariants such that the stratum for a given symmetry phase is contained in a vector space. A discussion of this idea can be found in [11].

These two examples were not only chosen because of their relevance in material science. The methods needed to find energy functions in these cases are likely to show typical difficulties, as well as the relative simplicity of this approach.

It should be pointed out that only Section 4 applies specifically to phase transitions in crystals; the method is applicable to a variety of problems. The main requirement is that the symmetry group is a compact Lie group.

A drawback of the method discussed in this paper is that, as seen in Section 4, the number of variables can increase. For a problem in the six-dimensional strain space, we have a six-dimensional fundamental domain. The orbit space, however, is embedded in $\mathbb{R}^{12}$, see Theorem 5 (one component is trivial; it is possible to embed the orbit space in $\mathbb{R}^9$ with one trivial component). For the cubic-to-tetragonal transformation, this was not a problem, since the restriction to the 6 primary invariants suffices for most applications. For the cubic-to-orthorhombic transition, this is not true. As demonstrated in Section 4, the construction of an energy function has to take one secondary invariant into account.

Acknowledgements. The results described here improve those of my PhD thesis [29], written under the supervision of JÜRGEN SCHEURLE and KARL-HEINZ HOFFMANN. I thank both of them for the support and advice. MATTHIAS RUMBERGER taught me a lot about orbit spaces. I am very grateful for the discussions with him. Some of the results reported here were obtained while I held a postdoctoral position in the Division of Engineering and Applied Science, California Institute of Technology. I thank KAUSHIK BHATTACHARYA for useful discussions. Financial support of the Deutsche Forschungsgemeinschaft, SFB 438, and the US Air Force Office of Scientifice Research, MURI grant F49620-98-1-0433, is gratefully acknowledged.

References

1. ABUD, M., SARTORI, G.: The geometry of orbit-space and natural minima of Higgs potentials. *Phys. Lett. B* **104**, 147–152 (1981)
2. ABUD, M., SARTORI, G.: The geometry of spontaneous symmetry breaking. *Ann. Physics* **150**, 307–372 (1983)
3. ATIYAH, M.F., MACDONALD, I.G.: *Introduction to commutative algebra*. Addison-Wesley Publishing Co., Reading, Mass.-London-Don Mills, Ont., 1969
4. CHEVALLEY, C.: Invariants of finite groups generated by reflections. *Am. J. Math.* **77**, 778–782 (1955)
5. CONTI, S., ZANZOTTO, G.: *Reconstructive phase transformations, maximal Ericksen-Pitteri neighborhoods, dislocations and plasticity in crystals*. Technical Report 42/2002, Max-Planck Institut für Mathematik in den Naturwissenschaften, Leipzig, 2002
6. COX, D., LITTLE, J., O'SHEA, D.: *Ideals, varieties, and algorithms*. second edition, New York: Springer-Verlag, 1997. (An introduction to computational algebraic geometry and commutative algebra)
7. DONDL, P., ZIMMER, J.: Modeling and simulation of martensitic phase transitions with a triple point. Submitted to *J. Mech. Phys. Solids*
8. EISENBUD, D.: *Commutative algebra*. New York: Springer-Verlag, 1995. (With a view toward algebraic geometry)
9. FALK, F., KONOPKA, P.: Three-dimensional Landau theory describing the martensic phase transformation of shape-memory alloys. *J. Phy. – Condensed Matter* **2**, 61–77 (1990)
10. FIELD, M.: *Lectures on bifurcations, dynamics and symmetry*. Volume **356** of Pitman Research Notes in Mathematics Series. Longman, Harlow, 1996
11. GATERMANN, K.: *Computer algebra methods for equivariant dynamical systems*. Berlin: Springer-Verlag, 2000
12. GREUEL, G.-M., PfiSTER, G., SCHÖNEMANN, H.: Singular 2.0. A Computer Algebra System for Polynomial Computations, Centre for Computer Algebra, University of Kaiserslautern, 2001. http://www.singular.uni-kl.de
13. HOCHSTER, M., EAGON, J.A.: Cohen-Macaulay rings, invariant theory, and the generic perfection of determinantal loci. *Am. J. Math.* **93**, 1020–1058 (1971)

14. JÄNICH, K.: *Differenzierbare G-Mannigfaltigkeiten*. Berlin: Springer-Verlag, 1968
15. JARIĆ, M.V., MICHEL, L., SHARP, R.T.: Zeros of covariant vector fields for the point groups: invariant formulation. *J. Physique* **45**, 1–27 (1984)
16. LANDAU, L.D.: On the theory of phase transitions. In: HAAR, D.T. (ed.), *Collected papers of L. D. Landau*. New York: Gordon and Breach Science Publishers, 1967
17. LYUBARSKIUı, G.Y.: *The application of group theory in physics*. New York: Pergamon Press, 1960
18. PITTERI, M., ZANZOTTO, G.: *Models for Phase Transitions and Twinning in Crystals*. CRC Press, Boca Rota, 2002
19. PROCESI, C., SCHWARZ, G.: Inequalities defining orbit spaces. *Invent. Math.* **81**, 539–554 (1985)
20. PROCESI, C., SCHWARZ, G.W.: The geometry of orbit spaces and gauge symmetry breaking in supersymmetric gauge theories. *Phys. Lett. B* **161**, 117–121 (1985)
21. RUMBERGER, M.: *Symmetrische dynamische Systeme: Differenzierbarkeit und linearisierte Stabilität*. PhD thesis, Universität Hamburg, Fachbereich Mathematik, 1997
22. RUMBERGER, M.: Finitely differentiable invariants. *Math. Z.* **229**, 675–694 (1998)
23. SCHÖNERT, M. *et al.*: *GAP – Groups, Algorithms, and Programming*. Lehrstuhl D für Mathematik, Rheinisch Westfälische Technische Hochschule, Aachen, Germany, fifth edition, 1995 http://www-gap.dcs.st-and.ac.uk/ gap/
24. SCHWARZ, G.W.: Smooth functions invariant under the action of a compact Lie group. *Topology* **14**, 63–68 (1975)
25. SMITH, G.F., RIVLIN, R.S.: The strain-energy function for anisotropic elastic materials. *Trans. Am. Math. Soc.* **88**, 175–193 (1958)
26. STURMFELS, B.: *Algorithms in invariant theory*. Vienna: Springer-Verlag, 1993
27. WEYL, H.: *The classical groups*: *Their invariants and representations*. Fifteenth printing, Princeton Paperbacks, Princeton, NJ: Princeton University Press, 1997
28. WORFOLK, P.A.: Zeros of equivariant vector fields: algorithms for an invariant approach. *J. Symbolic Comput*. **17**, 487–511 (1994)
29. ZIMMER, J.: *Mathematische Modellierung und Analyse von Formgedächtnislegierungen in mehreren Raumdimensionen* (*Mathematical Modeling and Analysis of Shape Memory Alloys in Several Space Dimensions*). PhD thesis, Technische Universität München, 2000

Max Planck Institute for Mathematics in the Sciences
Inselstr. 22,
04103 Leipzig,
Germany
e-mail: zimmer@mis.mpg.de

(*Accepted May 14, 2003*)
Published online February 11, 2004 –