PERGAMON

International Journal of Solids and Structures 38 (2001) 967–995

INTERNATIONAL JOURNAL OF SOLIDS and STRUCTURES

www.elsevier.com/locate/ijsolstr

Twinning analyses in the X-ray theory

J.L. Ericksen *

5378 Buckskin Bob Road, Florence, OR 97439, USA

Received 20 April 1999

Abstract

Commonly, expositions of twinning theory combine at least two different kinds of observations: measurements of macroscopic deformation and those made using X-rays. I believe that there is some merit in considering the latter separately because, often, the two kinds do not mesh very well. Here, my aim is to elaborate this and to improve the twinning theory based on a theory of X-ray observations to be described. © 2001 Elsevier Science Ltd. All rights reserved.

Keywords: Continuum theory of crystal defects; Twinning theory; Crystal multilattices

1. Introduction

The research of Zanzotto (1992) made clear that, for many crystals, the commonly used Cauchy–Born rule for relating changes in lattice vectors to macroscopic deformation is not consistent with observations of the changes associated with twinning and phase transitions. Briefly, this is the assumption that the macroscopic deformation gradient, applied as a linear transformation to lattice vectors before deformation, gives a possible set of lattice vectors in the deformed crystal. By arguments that I find convincing, he concludes that, when it is not, elasticity theory is inadequate to deal with these phenomena. He does note that the assumption does seem to be reliable for some kinds of crystals, the Bravais lattices and shape-memory alloys, for which elasticity theory has been used successfully to describe such phenomena. For other kinds of crystals, the assumption sometimes applies, but often fails, and there seems to be no clear pattern in this. When it fails, we have no reliable theory for relating deformation to the lattice vectors and shifts describing the crystal structures, commonly observed using X-ray methods. I thought it desirable to construct some type of theory to describe at least some such observations. To this end, I proposed (Ericksen, 1997) a continuun theory of crystal multilattices [1] dealing only with the X-ray observations, hereafter called X-ray theory. Certainly, it is important to deal with deformations, but I do not know how to do so. Experts do find this very difficult. In my presentation of this theory, I included some brief comments about related twinning theory, introducing an idea not used before in the literature, as far as

* Fax: +1-541-997-6399.

[1] The book by Pitteri and Zanzotto (2000) is a very good reference for relevant information on these and the Bravais lattices.

0020-7683/01/$ - see front matter
PII: S0020-7683(00)00069-X

I know. Here, my purpose is to discuss this kind of theory in some detail, including an illustrative example, an analysis of one of the five twinning modes observed in orthorhombic α-uranium, which involves some subtleties. Also, a partial analysis of another twin in this material is presented to illustrate other points. Generally, my aim is to link better such calculations to the theory of constitutive equations. Twinning analyses in the X-ray theory are somewhat different from what you are likely to find in expositions of twinning theory, so I will elaborate this.

Another of my aims is to adapt to the X-ray theory, the kinds of analyses, based on elasticity theory, that have been used to describe some microstructures involving twinning. Conceptually, these have relied on notions of reference configurations and deformations, concepts which are not involved in the X-ray theory. Also, to relate calculations to X-ray observations, the Cauchy–Born rule is used, and the X-ray theory can be used for cases where it fails or is not relevant, as is the case for some studies of growth twins. Mathematically, such use of elasticity theory involves considering minimizing sequences for energy functionals that do not converge to minimizers, but have useful limits which can be described, using the theory of the probability distributions known as the Young measures. The literature on applications of this to elasticity theory has grown rather large. For readers not familiar with it, I suggest starting with the paper by Ball and James (1992), which covers the basic mathematical theory, some applications and a number of relevant references. Actually, it might be useful to reconsider how the mathematical theory applies to functions defined on bounded domains, to provide a better basis for most calculations of this kind. Some newer ideas and references are covered by Bhattacharya et al. (1994). I will introduce an alternative to the formulation used for elasticity theory which generalizes more easily to the X-ray theory, then briefly indicate how to do the generalization. For this part to be comprehensible, one needs to have some understanding of the basic mathematical theory, and I will not rehash this.

Particularly, in Section 6, I have made a serious effort to understand and explain common practices that I have found confusing and seem likely to be confusing to other theorists. In this, I could not avoid making some guesses, so bear this in mind, in assessing my opinions about this.

2. X-ray theory

Here, I will give a brief summary of my X-ray theory (Ericksen, 1997), restricting the discussion of this to monatomic crystals for simplicity. The configurations are described as n lattices, where n is any positive integer: constitutive equations treat n as fixed, the configurations as variable. An n lattice consists of n identical lattices, translated in different ways relative to each other. The usual idea is that a lattice describes a set of "positions" of identical atoms, representable in the form:

$$n^a \mathbf{e}_a + \text{const.}, \quad n^a \in Z, \tag{2.1}$$

where the (three) vectors $\mathbf{e}_a$ and lattice vectors are a set of linearly independent vectors. The quotation marks indicate that, physically, "position" really means a point describing some averaged location of an atom. For various purposes, it is also important to introduce the reciprocal lattice vectors (dual basis) $\mathbf{e}_a$, satisfying

$$\mathbf{e}^a \otimes \mathbf{e}_a = \mathbf{e}_a \otimes \mathbf{e}^a = \mathbf{1}, \quad \mathbf{e}^a \cdot \mathbf{e}_b = \delta^a_b. \tag{2.2}$$

For an n lattice, one also needs to describe how the different lattices are translated relative to some point in space. With the usual ideas of invariance under translations associated with Galilean invariance, there are some advantages in taking the point to be some position in one of the lattices, using what Pitteri (1985a) has called shifts, a set of vectors denoted by

$$\mathbf{p}_i, \ i = 1, \ldots, \nu = n - 1. \tag{2.3}$$

For any configuration, there are infinitely many ways of choosing these vectors. Commonly, estimates of these are obtained using X-rays, probed over the width of an X-ray beam, wide enough to include a very large number of atoms. For this reason, observations made using electron microscopes reveal more about atomic arrangements near defects which might even be invisible in X-ray observations, for example. This motivated me to consider a continuum theory, treating the indicated vectors as vector fields, functions of position in space. This is what I call the X-ray theory, which is an equilibrium theory. My interest is in providing some theory for certain phenomena outside the range of validity of elasticity theory. Included in this are the discontinuities associated with twinning in some crystals, such as are documented by Zanzotto (1992), changes associated with phase transitions in some crystals, and growth twins. For such phenomena, I believe that it is reasonable to accept some assumptions I made in most cases. There are exceptions for some growth twins, noted in Section 6. I excluded continuous distributions of dislocations, leading to the conclusion that one has an analog of the inverse of the deformation gradient in elasticity theory. That is, there are scalar functions χ^a such that

$$\mathbf{e}^a = \nabla\chi^a. \tag{2.4}$$

This leaves open the possibility of analyzing isolated dislocations, similar to the way this is done in elasticity theory. Generally, twinning involves finite jumps in $\mathbf{e}_a$, $\mathbf{e}^a$ and $\mathbf{p}_i$ across some surface. After pondering observations and common practices, I did and still do consider it to be reasonable to assume that it is possible to choose lattice vectors on the two sides, so that the Burger's vector vanishes for all Burger's circuits intersecting the discontinuity surface, not enclosing other defects, most likely to be dislocation lines. This implies that χ^a can be taken to be continuous across the discontinuity surface. This is an analog of the usual assumption of continuity of the displacement in elasticity theory for twins. Associated with this is the usual kinematic condition of compatibility, which can be put in the form

$$\bar{\mathbf{e}}^a = (\mathbf{1} - \mathbf{n}\otimes\mathbf{a})\mathbf{e}^a, \tag{2.5}$$

where $\mathbf{n}$ is the unit normal to the discontinuity surface and $\mathbf{a}$ is an amplitude vector. Also, $\bar{\mathbf{e}}^a$ and $\mathbf{e}^a$ here represent limiting values from the two sides. This will play an important role in later discussions. For twins in unstressed crystals, it is the common understanding that the volume of a unit cell is the same on both sides. Also, for most if not all mechanical twins and some growth twins, the experience is consistent with the assumption that $\bar{\mathbf{e}}^a$ and $\mathbf{e}^a$ can be selected so as to have the same orientation. With both of these assumptions,

$$\mathbf{a}\cdot\mathbf{n} = 0, \tag{2.6}$$

and Eq. (2.5) is equivalent to

$$\bar{\mathbf{e}}_a = (\mathbf{1} + \mathbf{a}\otimes\mathbf{n})\mathbf{e}_a. \tag{2.7}$$

Another possibility will be described in Section 6. I expect that Eq. (2.5) also applies to twins in samples bearing small loads, which are often associated with metastable equilibria. Then, Eq. (2.6) might well fail. No doubt, Eq. (2.7) looks more familiar to those with some experience in the analysis of mechanical twins. There, the indicated linear transformations are often associated with the macroscopic deformation gradient, a simple shear, and this is, effectively, an application of the Cauchy–Born rule. Similarly, there are conditions on the shifts associated with twins. I will not discuss these in general, but will treat a special case later, in an example, discuss it a bit more, in Section 6.

Typically, mechanical twins do involve a simple shearing deformation like this, in unloaded specimens, with a deformation gradient of the form.

$$\mathbf{F} = \mathbf{1} + \mathbf{b}\otimes\mathbf{n}, \tag{2.8}$$

where $\mathbf{b}$ is perpendicular to $\mathbf{n}$. To compare this with Eq. (2.7), take as a reference configuration with the lattice vectors $\mathbf{e}_a$. Commonly, X-ray observations provide infinitely many possibilities for satisfying

Eq. (2.7), for a given twin and choice of $\mathbf{e}_a$. When the Cauchy–Born rule does apply, one of these satisfies $\mathbf{a} = \mathbf{b}$ and, usually, in Bravais lattices, it is the one with the smallest value of $|\mathbf{a}|$. So, this is a guess often made by workers, to get a value of $\mathbf{F}$ from X-ray observations, although they use other ideas for multi-lattices. When the Cauchy–Born rule fails, none of the possibilities indicated by Eq. (2.7) agrees with Eq. (2.8).

Obviously, one really needs both X-ray observations and measurements of macroscopic deformation to check this. Later, we will consider an illustrative example. To get something like the Cauchy–Born rule, workers often use what Zanzotto (1992) calls sublattices. Briefly, this involves using larger lattice vectors, effectively ignoring some sets of atoms in the real lattices. Empirically, it seems that such sublattices exist for all mechanical twins, when the Cauchy–Born rule fails to apply, although I do not know of a good physical reason for this. In Section 6, I will mention variations on twinning equations used by workers to allow for this. Obviously, it is desirable to test the aforementioned assumptions by analyzing observed twins for which there is some reason to suspect that they might not apply. It seemed to me possible that the examples to be discussed might be of this kind. We will see how analyses of these work out. It is an old difficulty in twinning theory that, often, theoretically possible twins are not observed, and I have no new ideas for trying to remedy this. It is tricky, since we cannot know what might be observed in the future.

One does need to be able to account for mass densities, and I made a rather obvious assumption (Ericksen, 1997) about this, to be mentioned later. Also, I introduced a constitutive function for φ, the Helmholtz free energy per unit mass, of the form

$$\varphi = \hat{\varphi}(\mathbf{e}_a, \mathbf{p}_i, \theta), \tag{2.9}$$

where θ denotes absolute temperature. Given Eq. (2.4), it is really better to use one of the form

$$\varphi = \hat{\hat{\varphi}}(\mathbf{e}^a, \mathbf{p}_i, \theta). \tag{2.10}$$

Obviously, it is a matter of making a change of variables to get one from the other, and I will not belabor converting the analyses I gave to put them in terms of $\hat{\hat{\varphi}}$.

For the analysis of twinning, in particular, one needs to make some assumptions about the invariance group for $\hat{\varphi}$ or $\hat{\hat{\varphi}}$ and, unfortunately, this is a complicated business, only partly because one is dealing with n lattices, where n can be any positive integer. It is a bit easier to discuss this for $\hat{\varphi}$, which better fits conventional thinking, and to ignore the dependence on θ. Certainly, one needs invariance under finite rotations,

$$\hat{\varphi}(\mathbf{R}\mathbf{e}_a, \mathbf{R}\mathbf{p}_i) = \hat{\varphi}(\mathbf{e}_a, \mathbf{p}_i), \quad \mathbf{R} \in SO(3), \tag{2.11}$$

implying that these transformations must map the domain of $\hat{\varphi}$ onto itself. Physically, $\hat{\varphi}$ should be at least differentiable, to be able to calculate the Cauchy and configurational stresses, and entropy. This is not likely to be true, if $\mathbf{e}_1 \cdot \mathbf{e}_2 \wedge \mathbf{e}_3 = 0$ or ∞. As the domain of $\hat{\varphi}$ must be a connected set, we must then have, throughout its domain,

$$\text{either } \mathbf{e}_1 \cdot \mathbf{e}_2 \wedge \mathbf{e}_3 > 0 \text{ or } \mathbf{e}_1 \cdot \mathbf{e}_2 \wedge \mathbf{e}_3 < 0. \tag{2.12}$$

So, for example, any lattice described by $\mathbf{e}_a$ can also be described by $-\mathbf{e}_a$, but the two possibilities cannot both be in the domain of $\hat{\varphi}$. With growth twins, there is real possibility of having enantiomorphic configuarations on the two sides. There is then the possibility that two different "mirror image" constitutive equations are appropriate. This does not necessarily require that their lattice vectors be oppositely oriented, but it allows for this possibility, in constitutive theory. I interpret the different constitutive equations to mean that it is impossible to remove such twins by applying loads. This fits the experience with Brazil twins in quartz, for example. Similarly, one needs to exclude the values of $\mathbf{p}_i$, which let two different atoms occupy the same position. This puts some conditions on the topology of the domain that are not encountered in the

theory of Bravais lattices (1-lattices). I will ignore these, as they are not needed for matters to be discussed here.

Another complication not encountered in Bravais lattices is one that I glossed (Ericksen, 1997). The domain is likely to include nonessential descriptions in the terminology of Pitteri (1998) and Pitteri and Zanzotto (2000, Chapter 4). Briefly, these are the values of $\mathbf{e}_a$ and $\mathbf{p}_i$ for an n lattice describing a configuration which can also be described as an n' lattice, with $n' < n$. He gave one characterization of these and, more recently, I gave (Ericksen, 1997) a different one. In some cases of interest, but not all, one can pick domains excluding these. Otherwise, it is still not clear how best to deal with those occurring in the domains of functions, so, I will not face up to this difficulty. As candidates for transformations to be included in the invariance group for a function $\hat{\varphi}$, we get some infinite discrete ones, which do form groups. For the lattices, $\mathbf{e}_a$ and $\tilde{\mathbf{e}}_a$ describe the same lattice provided

$$\tilde{\mathbf{e}}_a = m_a^b \mathbf{e}_b, \qquad \mathbf{m} = \|m_a^b\| \in G, \tag{2.13}$$

the group of unimodular matrices of integers, often denoted by $\mathrm{GL}(3, Z)$ or a similar notation. Of course, Eq. (2.12) excludes the possibilities with $\det \mathbf{m} = -1$ although, as I noted (Ericksen, 1997), one could combine these with improper orthogonal transformations. In these and other matrices, my convention is that the lower index always labels rows. To define shifts, we picked, an atom in one of the lattices an an origin, and numbered the lattices in some manner. Picking another such atom and renumbering gives transformations of the form

$$\mathbf{p}_i \to \tilde{\mathbf{p}}_i = \alpha_i^j \mathbf{p}_j + n^a h \mathbf{e}_a, \quad n^a i \in Z, \tag{2.14}$$

where the matrices $\boldsymbol{\alpha} = \|\alpha_i^j\|$ can be interpreted as forming an offbeat representation of the permutation group on n objects. Pitteri (1985a) described a set of generators which, with my convention, consists of

$$\begin{cases} \text{(a) matrices obtained by replacing column in the unit matrix by entries all equal to } -1 \text{ and} \\ \text{(b) matrices obtained by interchanging two columns in the unit matrix,} \end{cases} \tag{2.15}$$

for essential descriptions. He uses a different convention, making his matrices transposes of mine. As I see it, these transformations also apply to those nonessential descriptions, but one should also explore additional implications of their being describable as n' lattices. As I discussed (Ericksen, 1999) in some detail, it suffices to use a proper subset of these, for monoatomic crystals, except for some smaller values of n. For polyatomic crystals, additional restrictions are imposed on the $\boldsymbol{\alpha}'$s.

Now, for Bravais lattices, it has been very useful to consider as a domain one of the neighborhoods of Pitteri (1984, 1985a) because, effectively, these pick out finite subgroups of the infinite discrete groups, which map the neighborhood onto itself, in a nice way, more traditional invariance groups. It is possible for these neighborhoods to be unbounded with respect to applying uniform dilatations to lattice vectors, but most analyses done do not use this flexibility. Unfortunately, there are cases where the configurations of interest cannot be in the same neighborhood, for example, the very common twins in cubic and hexagonal crystals, although it is physically reasonable to take the domains of interest to be bounded for these. So, there is room for some new ideas, to deal with such exceptions, preferably without requiring the use of infinite discrete groups.

Pitteri (1985a) generalized the basic theory of neighborhoods from his earlier version (Pitteri, 1984) for Bravais lattices, but only when they are centered at essential descriptions. Then, the relevant discrete group reduces to a finite group, the lattice group of the center. The lattice group for a nonessential description, described in the same way, is never unique. Elements of lattice groups are sets of integers, of the form

$$(\mathbf{m}, \boldsymbol{\alpha}, \mathbf{q}), \quad \mathbf{q} = \|q_i^a\| \in Z, \tag{2.16}$$

where the matrices $\mathbf{m}$ and α are of the form described above, this set being associated with the equations

$$\begin{cases} \mathbf{Q}\mathbf{e}_a = m_a^b \mathbf{e}_b, \quad \mathbf{Q} \in O(3), \\ \mathbf{Q}\mathbf{p}_i = \alpha_i^j \mathbf{p}_j + q_i^q \mathbf{e}_a. \end{cases} \tag{2.17}$$

One uses these, restricted to be consistent with Eq. (2.12), evaluated at the lattice vectors and shifts describing the configuration chosen as center. It turned out that crystallographers had not worked on these groups, hampering development of this theory. However, progress is being made in understanding them. For example, Pitteri and Zanzotto (1998) constructed a very nice example showing that these can be used to distinguish subtle differences in symmetry that are not by the commonly used space groups and site-symmetry groups. Generally, they can distinguish differences in symmetry missed by the latter groups. Also, Pitteri and Zanzotto (2000, Chapter 4) present calculations of these groups for some special cases. I characterized (Ericksen, 1999) the groups for two lattices and three lattices with lattice vectors of most of the lower symmetry triclinic and monoclinic types.

Adeleke (1999) characterized these for general n lattices with body-centered orthorhombic lattice vectors. The latter two papers explain how the $\boldsymbol{\alpha}$s are a representation of the permutation group on n objects. Also, Parry (1998) treated some low-dimensional lattice groups. I find that it eases some such analyses, if one replaces Eq. (2.17) by the equivalent

$$p_i^a m_a^b = \alpha_i^j p_j^b + q_i^b, \tag{2.18}$$

where

$$\mathbf{p}_i = p_i^a \mathbf{e}_a \Rightarrow p_i^a = \mathbf{p}_i \cdot \mathbf{e}^a. \tag{2.19}$$

This also better fits formulation (2.10), with $\mathbf{p}_i$ replaced by p_i^a. As yet, we know very little about what kinds of phase-transitions, twinning, etc. can be analyzed, using these neighborhoods, for multilattices. Also, there is need to understand how to use neighborhoods centered at nonessential descriptions, in a similar way. Generally, it is not hard to calculate the lattice group for a particular configuration, given an essential description, as I will do for an example discussed later.

I think it obvious that the X-ray theory is still in a very rudimentary state and it is a complicated theory. However, tackling special problems with newer theories usually results in gains in our understanding, and I think it is now feasible to do some of this.

3. An example

In attempting to analyze twin patterns which are or might be observed in a particular material, a good first step is to analyze a single twin that has been observed in it, with constant lattice vectors and shifts on each side. One can find lists of observed twins and other general information on them in standard references such as Barrett and Massalski (1966), Hall (1954), Kelly and Groves (1951), Klassen-Nekliudova (1964) and Reed-Hill et al. (1964). Also, there is a fairly recent review of work on deformation twinning, containing much information and numerous references, by Christian and Mahajan (1995). Bear in mind that observations of new kinds of twins occur from time to time, so views of the subject and tables can and do change, as a result of this. Briefly, deformation twins are produced by suitably loading a crystal, then removing the loads. Transformation twins do not involve loading, occurring naturally in unstressed crystals as a result of various phase transitions involving a change of symmetry. Intuitively, producing deformation twins treats crystals rather roughly, more so than the treatment producing transformation twins, for example. From this, one might guess that other kinds of defects are more likely to accompany deformation twins. However, twinning theory not accounting for this has been quite successful. Some information listed in such tables is not really relevant to the X-ray theory, referring to observations of macroscopic deformation. For an example, I will pick a material for which Zanzotto (1992) concluded that elasticity theory is

inadequate to describe all observed twins. Conceivably, this might suggest that my assumptions do not all apply. The materials he implicates include some hexagonal close-packed crystals. The lattice group for these is maximal, which implies that it is impossible for the two configurations involved in a twin to be in the same Pitteri neighborhood. In this respect, the lower symmetry (orthorhombic) α-uranium he analyzes seems to be more promising, these twins being produced by stress, induced mechanically or thermally. This material is also interesting, because a variety of rather unusual twins are observed in it. It is described as a (monatomic) four lattice by Barrett and Massalski (1966, p. 170), a description used by all workers, as far as I know. The lattice vectors are chosen to be of the form

$$\begin{cases} \mathbf{e}_1 = a\mathbf{i}, \quad \mathbf{e}_2 = b\mathbf{j}, \quad \mathbf{e}_3 = c\mathbf{k}, \\ \mathbf{e}^1 = \mathbf{i}/a, \quad \mathbf{e}^2 = \mathbf{j}/b, \quad \mathbf{e}^3 = \mathbf{k}/c, \end{cases} \tag{3.1}$$

where the vectors **i**, **j** and **k** are orthonormal, and a, b, c are constants, with

$$0 < a < b < c. \tag{3.2}$$

Later, I will say more about relevant values of these and the shifts. Look at the table given by Barrett and Massalski (1966, p. 415), for example and you get a rather standard way of listing twinning information. Some tables include additional entries, use slightly different notations, etc. For α-uranium, they list five twinning modes. The first, which is also the mode most frequently observed, is described as

$$\begin{cases} \text{Twinning} & \text{Twinning} & \text{Second undisorted} & \text{Direction} & \text{Shear} \\ \text{plane}, K_1 & \text{direction}, \eta_1 & \text{plane}, K_2 & \eta_2 & \\ \{1\,3\,0\} & \langle 3\,\bar{1}\,0\rangle & \{1\,\bar{1}\,0\} & \langle 1\,1\,0\rangle & 0.229 \end{cases} \tag{3.3}$$

The labels K_1, η_1, etc. are standard. I will follow the common practice of using the K_1 entry to label twinning modes, so these are the $\{1\,3\,0\}$ twins, or it is the $\{1\,3\,0\}$ mode. The numbers involved are crystallographic indices of the directions mentioned; one can use the numbers as they stand, or use crystallographically equivalent sets of directions. In some tables occurring in the literature, using such numbers as they stand can lead to error. At least check that they imply that K_1 and η_1 as well as K_2 and η_2 are orthogonal, as they must be. Later, I will mention cases in point. The description of twinning elements by Barret and Massalski (1966, p. 411) should make fairly clear the interpretation of these entries, at least for mechanical twins, but I will raise a question about the interpretation of η_1. Here, only the first and, sometimes, the second entries are relevant to the X-ray theory, the others referring to observations of macroscopic deformation. Actually, workers exercise ingenuity, intrying to get best estimates of all entries. I do not think it is worth getting into a lengthy discussion of this. So, allow for the fact that my remarks about such matters are somewhat simplistic. With the obvious difference in how the planes are transformed some workers prefer to use different names for the planes, calling the latter composition planes.

To illustrate some ideas in a simpler way I will make a "lucky guess" that $\mathbf{a}\|\boldsymbol{\eta}_1$ (**a** is parallel to $\boldsymbol{\eta}_1$), the vector with η_1 as components relative to $\mathbf{e}_a$. Workers familiar with twinning analyses might spot that this is not just a guess. Later, I will explain this. A more thoughtful treatment is described in Section 6. Also, I will gloss some subtleties related to twinning theory, discussed in Section 6, that are not important here. In terms of the vectors **a** and **n**, the first two entries give

$$\mathbf{n}\|\mathbf{e}^1 + 3\mathbf{e}^2, \qquad \mathbf{a}\|3\mathbf{e}_1 - \mathbf{e}_2, \tag{3.4}$$

or a pair of directions crystallographically equivalent to these. Ignoring the remaining entries, we have

$$\mathbf{S}_x \overset{\text{def}}{=} 1 + \mathbf{a} \otimes \mathbf{n} = 1 + \alpha(3\mathbf{e}_1 - \mathbf{e}_2) \otimes (\mathbf{e}^1 + 3\mathbf{e}^2), \tag{3.5}$$

where α is some scalar. Here, the $\mathbf{e}_a$ are lattice vectors on one side of the twin. Subscript on $\mathbf{S}$ indicates that I interpret this as a shear to be determined by the X-ray theory and measurements only, although I will

compare it with the value of the shear deformation. Here, $\mathbf{S}_X$ represents a one parameter family of values, parameterized by α.

To qualify as a twin in an unstressed material, it is generally agreed that some orthogonal transformation applied to the configuration on one side gives the configuration on the other. For most observed mechanical twins, this is a 180° rotation with $\mathbf{n}$ as axis, associated what are called type I or rotation twins. [2] For the occasional exception, it is most often a 180° rotation with $\mathbf{a}$ as axis, associated with what are called type II or reflection twins. For a long time, experts were not convinced that other types of mechanical twins exist in nature. However, recent observations, to be discussed later, seem to establish that there are. Simply, I do not know how well these are accepted by experts. For α-uranium, the types of observed twins are mentioned by Christian and Mahajan (1995, Section 2.8). The twin at hand is compound, meaning that it can be analyzed as either type I or type II. It will be analyzed as type I, using

$$\begin{cases} \mathbf{R} = -1 + 2\mathbf{n} \otimes \mathbf{n}, \\ \mathbf{n} = (\mathbf{e}^1 + 3\mathbf{e}^2)/|\mathbf{e}^1 + 3\mathbf{e}^2|. \end{cases} \tag{3.6}$$

Now, from this, the set

$$\hat{\mathbf{e}}_a = \mathbf{R}\mathbf{e}_a, \quad \hat{\mathbf{p}}_i = \mathbf{R}\mathbf{p}_i \tag{3.7}$$

is a possible set of lattice vectors and shifts on the other side. From Eq. (2.7), the vectors indicated there by $\bar{\mathbf{e}}_a$ are also a possible set of lattice vectors on the same side. Using Eq. (2.13) to put this together, we get the twinning equation

$$\begin{aligned} \bar{\mathbf{e}}_a &= \mathbf{S}_X \mathbf{e}_a = [1 + \alpha(\mathbf{3}\mathbf{e}_1 - \mathbf{e}_2) \otimes (\mathbf{e}^1 + 3\mathbf{e}^2)]\mathbf{e}_a \\ &= m_a^b \hat{\mathbf{e}}_b = m_a^b \mathbf{R}\mathbf{e}_b, \end{aligned} \tag{3.8}$$

with $\mathbf{R}$ given by Eq. (3.6). This is to be solved for possible values of the scalar α and $\mathbf{m} \in G$. This will look familiar to those dealing with twins in Bravais lattices and shape memory alloys, for which the Cauchy–Born rule applies, according to Zanzotto (1992). Then, a possible value of $\mathbf{S}_X$ is taken to be the macroscopic deformation gradient $\mathbf{F}$, as described earlier. However, Eq. (3.8) is based on a different concept and, generally, I believe that it applies when the Cauchy–Born rule fails. To compare with the macroscopic shear deformation, it is convenient to proceed as follows: For any value of α, we can determine a matrix $\boldsymbol{\mu}(\alpha)$ such that Eq. (3.8) holds, with $\mathbf{m}$ replaced by $\boldsymbol{\mu}$. This is just a description of (mixed) components of $\mathbf{R}\mathbf{S}_X$ in the lattice vector basis. A calculation gives

$$\boldsymbol{\mu}(\alpha) = \begin{Vmatrix} 1 - 3\mu & \mu & 0 \\ 6 - 9\mu & 3\mu - 1 & 0 \\ 0 & 0 & -1 \end{Vmatrix}, \tag{3.9}$$

where

$$\mu = 6a^2/(9a^2 + b^2) + \alpha. \tag{3.10}$$

It is easy to check that

$$\det \boldsymbol{\mu} = 1, \quad \boldsymbol{\mu}^2 = 1. \tag{3.11}$$

Clearly, we will have $\boldsymbol{\mu} \in G$ if and only if μ is an integer, call it m. So, Eq. (3.8) will be satisfied, for any integer m, if we take

[2] One should be wary of the fact, pointed out by Zanzotto (1988), that inequivalent definitions of the types occur in the literature, and it is sometimes hard to know which a writer has in mind, if you do not know some common practices, explained in Section 6. I use a slight generalization of the one Zanzotta selects, covering multilattices.

$$\alpha = m - 6a^2/(9a^2 + b^2). \tag{3.12}$$

This type of indeterminacy is familiar to those who do twinning analyses, using X-ray data, to estimate **F**, assuming that the Cauchy–Born rule might apply and, soon, I will explain this. One then cannot really determine the macroscopic deformation gradient uniquely from such X-ray observations. To try to do so, I will use the guess commonly used for Bravais lattices, minimizing the shear magnitude s_X, which is given by

$$\begin{aligned} s_X(m) &= |\alpha\|3\mathbf{e}_1 - \mathbf{e}_2\|\mathbf{e}^1 + \mathbf{e}^2| \\ &= |m(9a^2 + b^2) - 6a^2|/ab \end{aligned} \tag{3.13}$$

by a routine calculation. Here, it might be reasonable to consider the two smallest, as they give shears in opposite directions, they are

$$s_X(0) = 6a/b \quad (\alpha < 0), \tag{3.14}$$

$$s_X(1) = b/a + 3a/b \quad (\alpha > 0). \tag{3.15}$$

Now, from the description of η_2 in Eq. (3.3), one can calculate a similar formula for s the magnitude of the macroscopic shear deformation. This gives a (known) formula

$$s = |b/a - 3a/b| = s_X(1/2). \tag{3.16}$$

This does not agree with Eq. (3.3) for any integer m For our α-uranium, data cited by Barrett and Massalski (1966, p. 170) give

$$b/a = 2.056, \quad (\alpha > 0), \tag{3.17}$$

yielding the value $s = 0.299$ noted in Eq. (3.3). For this ratio, Eq. (3.14) gives 2.92, much too large and in the wrong direction. With Eq. (3.15), we get the right direction, but with the value 3.515 much too large. Certainly, one cannot attribute these large discrepancies to experimental errors. What we are seeing seems to be clear evidence of a failure of the Cauchy–Born rule, but we shall see that it is not.

This analysis applies to any n lattice, at least when the lattice vectors are associated with an essential description. From the four-lattice description of α-uranium presented by Barrett and Massalski (1966, p. 170), I read off shifts, obtained by using as origin the point they label as $0y1/4,\ y = 0.015 \pm 0.005$, getting

$$\begin{cases} \mathbf{p}_1 = (\mathbf{e}_1 + \mathbf{e}_2)/2, \\ \mathbf{p}_2 = -2y\mathbf{e}_2 + \mathbf{e}_3/2, \\ \mathbf{p}_3 = \mathbf{p}_2 + (\mathbf{e}_1 + \mathbf{e}_2)/2. \end{cases} \tag{3.18}$$

This is essentially the same as an example I presented (Ericksen, 1998, Eq. (82)), of a nonessential four-lattice description. It can also be described as a two-lattice, with lattice vectors describing a base-centered orthorhombic lattice. The routines presented there give these as

$$[(\mathbf{e}_1 + \mathbf{e}_2)/2, \mathbf{e}_2, \mathbf{e}_3], \tag{3.19}$$

and one shift, which can be taken as

$$\mathbf{p}_2 = -2y\mathbf{e}_2 + \mathbf{e}_3/2. \tag{3.20}$$

However, analyses to follow work out more neatly if one uses the equivalent set

$$\begin{cases} \tilde{\mathbf{e}}_1 = (\mathbf{e}_1 - \mathbf{e}_2)/2, \\ \tilde{\mathbf{e}}_2 = (\mathbf{e}_1 + \mathbf{e}_2)/2, \\ \tilde{\mathbf{e}}_3 = \mathbf{e}_3, \end{cases} \tag{3.21}$$

a commonly used description of a base-centered orthorhombic lattice, along with an equivalent shift, given by

$$\mathbf{p} = \mathbf{p}_2 + \tilde{\mathbf{e}}_2 = 2y\tilde{\mathbf{e}}_1 + (1-2y)\tilde{\mathbf{e}}_2 + \tilde{\mathbf{e}}_3/2. \tag{3.22}$$

Certainly, experts know that the configuration can be described as a two-lattice, as is clear from the discussion of Christian and Mahajan (1995, Section 2.8), for example, and they compensate for this by allowing half integers in Eq. (3.8). As is noted by Pitteri and Zanzotto (2000, Chapter 4), point and space groups calculated using nonessential descriptions are sometimes only proper subgroups of those obtained using essential descriptions. Here, the two descriptions give the same point groups and space groups. This means that, for most and perhaps all conventional analyses, it really does not matter which description one uses. However, for an analysis to follow, the difference is important. Consider Eq. (3.16), describing the macroscopic deformation. Changing variables to replace the old lattice vectors by the new, we get, with the shear $\mathbf{S}$ described by Eq. (3.16),

$$\begin{cases} \mathbf{S}\tilde{\mathbf{e}}_1 = -\mathbf{R}\tilde{\mathbf{e}}_1 - \mathbf{R}\tilde{\mathbf{e}}_2, \\ \mathbf{S}\tilde{\mathbf{e}}_2 = \mathbf{R}\tilde{\mathbf{e}}_2, \\ \mathbf{S}\tilde{\mathbf{e}}_3 = -\mathbf{R}\tilde{\mathbf{e}}_3, \end{cases} \tag{3.23}$$

and the coefficients on the right now form a unimodular matrix of integers, given by

$$\mathbf{m}_1 = \begin{Vmatrix} -1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{Vmatrix} \tag{3.24}$$

with

$$\det \mathbf{m}_1 = 1, \qquad \mathbf{m}_1^2 = 1. \tag{3.25}$$

The effect is to allow half integers as well as integers in Eq. (3.13) and minimizing s_X in this larger set gives $\mathbf{S}$, which agrees with measurements of the deformation gradient. With the nonessential description, we thus get the poor estimate of minimum shear, and an apparent failure of the Cauchy–Born rule, using the same X-ray observations. With the essential description, the Cauchy–Born rule and minimum shear hypothesis do apply. As was mentioned earlier, workers avoid this trap, by allowing half integers. As a matter of taste, I do not like this dodge, but it works, for this calculation. For the linear transformation $\mathbf{S}$, it does not matter what basis we use to describe it. So,

$$\begin{aligned} &\text{the Cauchy} - \text{Born rule does apply, with the new choice of lattice vectors and shifts} \\ &\text{given by Eqs. (3.21) and (3.22), providing an essential description.} \end{aligned} \tag{3.26}$$

A calculation gives

$$\begin{cases} \mathbf{e}^1 + 3\mathbf{e}^2 = 2\tilde{\mathbf{e}}^2 - \tilde{\mathbf{e}}^1, \\ 3\mathbf{e}_1 - \mathbf{e}_2 = 2(2\tilde{\mathbf{e}}_1 + \tilde{\mathbf{e}}_2) \end{cases} \tag{3.27}$$

and

$$\mathbf{S} = 1 + 2(b^2 - 3a^2)(2\tilde{\mathbf{e}}_1 + \tilde{\mathbf{e}}_2) \otimes (2\tilde{\mathbf{e}}^2 - \tilde{\mathbf{e}}^1)/[9a^2 + b^2)]. \tag{3.28}$$

With the new choice of lattice vectors, the table corresponding to Eq. (3.3) becomes

Twinning plane, K_1	Twinning direction, η_1	Second Undistorted plane, K_2	Direction η_2	Shear
$\{\bar{1}20\}$	$\langle 210\rangle$	$\{100\}$	$\langle 010\rangle$	0.299

(3.29)

As in the previous calculation, the X-ray theory involves an infinite number of shears, now including the one describing the macroscopic shear. As is familiar to experts, this occurs because of the existence of lattice

invariant shears S_L, meaning shears that map one set of lattice vectors to another one for the same lattice. For this reason, they are invisible, in X-ray observations. Here, the important ones are of the form

$$\mathbf{S}_L = 1 + m(3\mathbf{e}_1 - \mathbf{e}_2) \otimes (\mathbf{e}^1 + 3\mathbf{e}^2), \tag{3.30}$$

for the old set of lattice vectors, where m is any integer. For the new set, a look at Eq. (3.27) gives

$$\mathbf{S}'_L = 1 + m'(2\tilde{\mathbf{e}}_1 + \tilde{\mathbf{e}}_2) \otimes (2\tilde{\mathbf{e}}^2 - \tilde{\mathbf{e}}^1), \quad m' = 2m, \tag{3.31}$$

where m' is any integer, justifying the rather common practice of using integers and half integers in Eq. (3.13), as workers do and as I did in Eq. (3.16).

For a satisfactory analysis of these twins, $\hat{\varphi}$ should be invariant under finite rotations and the subgroup of the lattice group for $\tilde{\mathbf{e}}_a$ and $\mathbf{p}$ which is consistent with Eq. (2.12), in particular. The complete lattice group is of order eight, involving all of the orthogonal transformations in the point group for the lattice vectors indicated in Eq. (3.21). Elements corresponding to a central inversion and 180° rotations with axes $\mathbf{e}_2$ and $\mathbf{e}_3$ are, respectively,

$$\left\{ \left\| \begin{matrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{matrix} \right\|, -1, (0,0,0) \right\}, \tag{3.32a}$$

$$\left\{ \left\| \begin{matrix} -0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & -1 \end{matrix} \right\|, 1, (-1,-1,-1) \right\}, \tag{3.32b}$$

$$\left\{ \left\| \begin{matrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{matrix} \right\|, -1, (0,0,1) \right\}, \tag{3.32c}$$

where, interpreted as in Eq. (2.16), these serve as generators of the complete lattice group for this configuration, and the first can be deleted in considering the invariance group for $\hat{\varphi}$. To properly cover the twins considered, we also need to have $\hat{\varphi}$ invariant under the $\mathbf{m}_1$ given by Eq. (3.24), hence under the group generated by it and the two $\mathbf{m}s$ in Eqs. (3.32b) and (3.32c). For these twins to be included in a Pitteri neighborhood, it is necessary that this be a finite group. This can be true only if the three $\mathbf{m}s$ are included in the lattice group of some Bravais lattice. If so, the lattice vectors for the latter, denoted by $\mathbf{c}_a$, are possible candidates for the center of a Pitteri neighborhood. So, one looks at the three equations of the form

$$\mathbf{R}\mathbf{c}_a = m_a^b \mathbf{c}_b, \tag{3.33}$$

for the three indicated $\mathbf{m}s$, the $\mathbf{R}s$ being some unknown rotations, the $\mathbf{c}_a$ also being unknowns. As all satisfy $\det \mathbf{m} = 1$, $\mathbf{m}^2 = 1$, these are all 180° rotations, if Eq. (3.33) can be satisfied. Analyzing this, one finds that the $\mathbf{c}_a$ must satisfy the following conditions:

$$\begin{cases} \mathbf{c}_1 \cdot \mathbf{c}_3 = \mathbf{c}_2 \cdot \mathbf{c}_3 = 0, \\ |\mathbf{c}_1| = |\mathbf{c}_2|, \\ 2\mathbf{c}_1 \cdot \mathbf{c}_2 = -|\mathbf{c}_2|^2, \end{cases} \tag{3.34}$$

identifying these as commonly used lattice vectors for hexagonal lattices. These should be selected to have the same orientation as the lattice vectors $\tilde{\mathbf{e}}_a$. It is easy to determine the three axes of the aforementioned 180° rotations, which are included in the point group for the hexagonal lattice.

With the $\mathbf{R}s$ thus defined, one also needs to define a shift $\bar{\mathbf{p}}$, for the hexagonal configuration, transforming so as to match the second and third entries in Eq. (3.32), and to be such as to preserve the invariance associated with Eq. (3.24). Also, it is preferable that the description be essential, as we know so

little about neighborhoods centered at nonessential descriptions. However, for this special situation, one could deal with this, if it were necessary, but it is not. Represent the shift by

$$\bar{\mathbf{p}} = \bar{p}^a \mathbf{c}_a. \tag{3.35}$$

The requirement that this fit the lattice group elements (3.32b) and (3.32c) yields

$$\bar{p}_1 + \bar{p}_2 = 1, \quad \bar{p}_3 = 1/2. \tag{3.36}$$

The condition to be avoided, that the description be nonessential, is that the shift is equivalent to one with components all equal to zero or a half. To include a lattice group element involving $\mathbf{m}_1$, we need to have

$$\left\{ \begin{array}{l} -\bar{p}^1 = \pm\bar{p}^1 + n^1, \\ -\bar{p}^1 + \bar{p}^2 = \pm\bar{p}^2 + n^2, \\ -\bar{p}^3 = \pm\bar{p}^3 + n^3, \end{array} \right. \tag{3.37}$$

where the n^a are integers and the same choice of sign must be used in the three entries. Here, I used Eq. (2.18). For the upper sign, one gets only nonessential descriptions, so we try the lower. The conditions do not determine the $\bar{p}^a$ uniquely, leaving room for using some equivalent shifts. Whichever one uses, one gets an hexagonal close-packed configuration. I prefer a standard choice,

$$\bar{\mathbf{p}} = \mathbf{c}_1/3 + 2\mathbf{c}_2/3 + \mathbf{c}_3/2. \tag{3.38}$$

This defines the corresponding lattice group element as

$$\left\{ \left\| \begin{array}{ccc} -1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{array} \right\|, -1, (0,1,0) \right\}, \tag{3.39}$$

the complete lattice group for the center being that for an hexagonal close-packed configuration, as described here. This makes it seem likely that the α-uranium configurations considered can be included in such a neighborhood, and I do believe this. I will not try to prove it, but will give a plausibility argument. For example, experts know that the configuration can be viewed as a distorted hexagonal one. Frank (1953), an ingenious person, got results similar to mine, by noticing a similarity between α-uranium and zinc. I got it by a rather routine calculation and will point out implications of this concerning constitutive theory, assuming my belief is correct. I note that, for this analysis, it is at best awkward to use the four-lattice description. So, this is one of various examples illustrating why it is better to use essential descriptions, unless there is a very good reason not to do so.

Let us try to construct a path joining the α-uranium configuration first considered to an hexagonal close-packed configuration, with

$$|\mathbf{c}_1| = |\mathbf{c}_2| = |\tilde{\mathbf{e}}_1| = |\tilde{\mathbf{e}}_2|. \tag{3.40}$$

On this, we wish to have the lattice groups of all configurations on the path be exactly that of the α-uranium configuration, except for the center, which will have a larger lattice group, of course. Denote by $\mathbf{f}_a$ and $\mathbf{q}$ the lattice vectors and shift for any configuration on the path. We require these to satisfy

$$\left. \begin{array}{l} |\mathbf{f}_1| = |\mathbf{f}_2| = |\tilde{\mathbf{e}}_1|, \mathbf{f}_1 \cdot \mathbf{f}_3 = \mathbf{f}_2 \cdot \mathbf{f}_3 = 0, \\ \mathbf{q} = \lambda\mathbf{f}_1 + (1-\lambda)\mathbf{f}_2 + \mathbf{f}_3/2, \end{array} \right\} \tag{3.41}$$

where λ is a parameter. Also, think of the $\mathbf{f}_a$ continuous functions λ, so, we have a path beginning at $\lambda = 2y = 0.03$ and ending at $\lambda = 1/3$. At the beginning, using the data on α-uranium, one can calculate the angle determined by $\tilde{\mathbf{e}}_1$ and $\tilde{\mathbf{e}}_2$, which is 128°, roughly. So, the angle determined by $\mathbf{f}_1$ and $\mathbf{f}_2$ needs to decrease from this value to 120° at the other end, and consider it to decrease monotonically, for simplicity. A problem could arise if this path contained a nonessential description. Now, Eq. (3.41) describes one at

$\lambda = 1/2$, for example, but there are none in the λ-interval of interest. Also, Eq. (3.14) is enough to guarantee that the lattice groups on the path have that at the beginning as a subgroup, at least. A problem could arise if the path passed through a configuration with a larger lattice group, but a calculation indicates that this does not happen. From this, I conclude that such a path is a connected set fitting the description of an hexagonal close-packed neighborhood. It is not really necessary to assume that Eq. (3.40) holds, but it makes the reasoning easier. To proceed, I assume that this inclusion applies. It should be noted that the paths and center need not be equilibrium configurations, because this requires that $\partial\hat{\varphi}/\partial\mathbf{p} = 0$ which might or might not be satisfied, except at the α-uranium configuration. Here, I interpret observations as implying that the latter is an equilibrium configuration.

In molecular theories of elasticity, workers follow Born (1923), using molecular models to determine $\hat{\varphi}$, solving the indicated equation or the obvious analog for n lattices for the shifts in terms of lattice vectors. With this, one can reduce $\hat{\varphi}$ to a function of lattice vectors. Then, the Cauchy–Born rule is used to reduce this to a function of $\mathbf{F}$. Of course, this is a dubious assumption, which might explain some failures of such theory. My X-ray theory does without the molecular models and the Cauchy–Born rule but, by design, it has a very similar structure. The exposition of Born's theory by Stakgold (1950) is clearer and more sound than the original.

Now, the idea is to assume that the domain of $\hat{\varphi}$ is such an hexagonal neighborhood, and that $\hat{\varphi}$ is invariant under this lattice group and SO(3). Also, it is to have a minimizer at the orthorhombic configuration described by Eqs. (3.21) and (3.22). Taking the orbit of this under the hexagonal group then gives three minimizers, called variants. Taking the orbit of these under SO(3) gives three disjoint infinite sets of minimizers. As will be familiar to those involved in the theory of microstructures, the number three is obtained by dividing the order of the hexagonal group by that of the orthorhombic group, by elementary group theory. For the twins considered above, configurations from two different such packages are used. That is, we apply the lattice group element (3.39) and the rotation $\mathbf{R}$ given by Eq. (3.6) to the vectors $\tilde{\mathbf{e}}_a$ and $\mathbf{p}$ to describe its twin, as indicated by

$$\tilde{\mathbf{e}}_a \to (m_1)_a^b \mathbf{R}\tilde{\mathbf{e}}_b, \tag{3.42}$$

used in Eq. (3.8), and

$$\mathbf{p} \to \mathbf{R}(-\mathbf{p} + \tilde{\mathbf{e}}_2). \tag{3.43}$$

Note that the procedure picks out one of the infinitely many equivalent shifts for the twin. In the jargon used by workers by workers in this area, we have predicted a definite kind of shuffling. Workers use various kinds of reasoning to estimate shuffling in various kinds of twins, but may seem not to relate this to any definite theory of constitutive equations, as far as I can tell. As we use only two of the three variants, it might be possible to use the third to describe more complicated patterns of twins coexisting in one crystal, but I will not pursue this. While this provides some basis for analyzing patterns of modes of the $\{1\,3\,0\}$ kind, it seems most unlikely, on the face of it, that it will be possible to include the other observed twinning modes in this neighborhood. Later, I will do a partial analysis of one such mode, eliminating any doubt about this, finding that it alone cannot be contained in any such neighborhood. Certainly, it will be very difficult to construct a theory to deal with all of the modes. However, we have good clues for constructing a theory to explore the effect of small loads on $\{1\,3\,0\}$ twins, for example. I presented (Ericksen, 1997) general equilibrium equations etc. needed for this. I do not mean to imply that it would be an easy matter to find an appropriate constitutive function $\hat{\varphi}$. So, we have barely scratched the surface, in constructing a theory of twinning, for this material. Later, we will add a little to this.

Earlier, I described taking $\mathbf{a} \| \boldsymbol{\eta}_1$ as a guess, but I also mentioned that these twins are compound. This means that they could also have been treated as type II, with the direction of $\mathbf{a}$ determined by the axis of rotation, something that could be determined from X-ray observations. Now, what is η_1? In most expositions, it is described as representing crystallographic indices of the vector $\mathbf{b}$ in Eq. (2.8). I believe that, in

twinning tables, η_1 can always be regarded as representing the axis of rotation for type II and compound twins. Here, this makes $\mathbf{a}\|\boldsymbol{\eta}_1$. After reading Section 6, you should understand my reasons for saying this. For the two prescriptions to agree, one should always have $\mathbf{b}\|\mathbf{a}$, for such twins, and I know of no way to justify assuming this, theoretically. I do not know of any observations of such mechanical twins indicating failure of these directions to be parallel, so the assumption that they are seem to have some status, empirically. Theoretically, I do not like using the same name for things that are different, conceptually. I note that, by itself, the usual description really implies that one cannot determine η_1, using only X-ray observations, as there is no reliable way of relating the latter to deformation. However, workers sometimes do, by accepting the other description, for this, as is clear from the work of Cahn (1953), for example.

If you look at various twinning tables, you will see three kinds of entries. For some twins, including those at hand, you will find entries presented as precise integers. This might make you wonder, since they represent experimental data, which are always subject to some error. For other entries, you will some comment to the effect that ertain entries are irrational. Then, the table might or might not list an approximation to the entries, using some set of integers. Certainly, workers know about approximating irrational numbers by rationals, and that, given an experimental estimate of some number associated with theory, one cannot determine whether the exact value is irrational or rational. To make sense of this, one needs to have some understanding of how workers make such decisions. In Section 6, I will explain this, as I understand it.

Do not think that this is a typical example of twinning analyses for multilattices. These twins are unusual in more than one way. As the basic atomic arrangement is not that of a Bravais lattice, it is at least somewhat exceptional to have the Cauchy–Born rule apply, as was mentioned earlier. Various type I and type II twins are not compound and this can complicate the analysis of them, as will be illustrated in Section 6. Even for Bravais lattices, there are many twins that cannot be included in any Pitteri neighborhood, as was mentioned earlier. With all of these nice features, we got a good correlation with the macroscopic deformation, using an hypothesis which is ad hoc, but is quite successful, for Bravais lattices. For these twins, elasticity theory might be adequate, but there are the other modes in this material, to which this theory does not apply, according to Zanzotto (1992). Workers trying to deal with deformation in other crystals are generally interested in trying to understand how all atoms move in this deformation, and we produced a particular estimate of shuffling which workers might accept, as a reasonable guess about this for these twins. Generally, such workers deal with twins that do not work out as happily, trying various hypotheses, with limited success, to correlate X-ray observations with measurements of deformation. I will not deal with these issues. For those interested in them, I suggest reading the discussion by Christian and Mahajan (1995) of them, which includes relevant references. In Section 6, I will analyze another mode in α-uranium, to illustrate some of the kinds of complications that can occur, in dealing with X-ray observations of twins.

4. Microstructures: elasticity theory

For elasticity theory, the idea that has been used is to fix some reference configuration for a crystalline body, and to consider minimizing sequences for the strain energy, or Helmholtz free energy functional, with sequences involving twinning deformations. Commonly, the constitutive equation is considered to be restricted to one of the Pitteri (1984) neighborhoods, [3] making this invariant under a finite material symmetry group G_M, essentially some point group, and the body is considered to be unloaded. As was mentioned earlier, there are practical difficulties involved in treating twins that cannot be included in such a neigh-

[3] For a slightly more general result and simpler treatment of such neighborhoods for Bravais lattices, cf. Ball and James (1992). Largely, thermoelasticity theory treats multilattices as Bravais lattices.

borhood, although one could do so, in principle. Essentially no progress has been made, in analyzing microstructures in such cases. Also, the energy density is assumed to have a set of minima generated by one or two, the orbit of these generated by SO(3) and G_M. We discussed an analog of this, in our example. Most of the interest has been in Martensitic transformations, where fine-scale twinning microstructures occur naturally when Austenite transforms to Martensite, in crystals bearing no loads. However, the method can be used for twins not of such transformation types, at least when they fit into some neighborhood.

In such endeavors, one is trying to relate theory to observations of twin microstructures. Experimentally, one then has information concerning the region actually occupied by the twinned specimen, in the Martensitic phase, the crystallographic orientation and arrangements of the twins, and, often, some information about the parent untwinned Austenitic specimen, commonly taken as a reference configuration.

Following the custom in elasticity theory, workers have used material coordinates as independent variables, so, the twin planes, etc. are described as certain directions in the reference configuration. To compare with observations, one then needs to map these to the actual configuration, to correct the values of angles between differently oriented twin planes, for example. For this, my impression is that what is really used is the identification of relevant crystallographic directions in the reference and deformed configurations provided by the Cauchy–Born rule. Certainly, this hypothesis is used in an important way, to relate X-ray observations of the crystallographic orientations of twin planes etc. to the calculations, and to select G_M. So, one must do something different when the rule fails to apply, one of my reasons for proposing the X-ray theory. Commonly considered minimizing sequences involve increasing numbers of such planes with the distance between parallel planes approaching zero in the limit, the deformation gradient $\mathbf{F}$ undergoing finite jumps across these. Rather obviously, such values of $\mathbf{F}$ will not converge pointwise to a value of $\mathbf{F}$, the basic reason why a minimizer is not obtained in the limit. However, this is one kind of limit which can be described, using the theory of Young measures, permitting one to do some useful calculations. With this theory, one can also describe sequences not only involving twins, such as the Austenite–Martensite interfaces studied by James and Kinderlehrer (1989), for example. Briefly, this describes the kind of theory I would like to adapt to the X-ray theory.

Roughly my idea is to interchange the roles played by the spatial coordinates and material coordinates. We are interested in comparing calculations with observations of a specimen with microstructure, occupying some region Ω in space. Instead of fixing a reference configuration, fix Ω. Again roughly, the idea is to consider the various material bodies that might be able to be in the observed configuration. So, instead of the usual deformation, we consider the inverse, maps of Ω to other domains, of the form

$$\mathbf{x} = \hat{\mathbf{x}}(\mathbf{y}). \tag{4.1}$$

Here, $\mathbf{x}$ and $\mathbf{y}$ denote the material and spatial coordinates, or coordinate-free equivalents, respectively. From the practice used in the previous view, I will borrow the assumption that these functions are included in $W^{1,\infty}$: it seems to me that arguments favoring this fit equally well with either of the two procedures. To use the spatial coordinates as independent variables, write the relevant energy in the form

$$E = \int_{\Omega} \rho w \, \mathrm{d}v, \tag{4.2}$$

putting the constitutive equation for w in the form

$$w = \hat{w}(\mathbf{F}^{-1}), \tag{4.3}$$

where $\mathbf{F}$ is again the usual deformation gradient: include a dependence on temperature, if you like. Clearly, one can take a constitutive equation for the energy per unit reference volume and transform it to get $\hat{w}$, or vice versa. I see no real difficulty in defining Pitter's neighborhoods to fit either formulation, for example. Roughly, the idea is to consider minimizing E, with the constraint that the total mass M be fixed. Here,

$$M = \int_{\Omega} \rho \, \mathrm{d}v, \tag{4.4}$$

ρ being the mass density. In the customary approach, one takes care of this in a trivial way, by fixing the reference mass density ρ_0, and the material region. With the new procedure, we use

$$\rho = \rho_0 \det \mathbf{F}^{-1}, \tag{4.5}$$

with ρ_0 a fixed constant. For the spaces and sequences of interest, $\det \mathbf{F}^{-1}$, along with $\mathbf{F}^{-1}$ and $\operatorname{adj} \mathbf{F}^{-1}$ are weakly continuous, making ρ weakly continuous, in particular. Briefly, this means that, in the limit, one can use the Young measures to calculate the mass of subregions of Ω. Or, one can do this, using the weak limits. There are possible reasons to prefer one to the other, too technical to discuss here. One does need to bear in mind that sequences considered should respect the condition that M is fixed. In typical calculations involving only twins, one considers the twinning equation, which can be put in the form

$$\mathbf{F}_2 = \mathbf{R}\mathbf{F}_1\mathbf{H} = (1 + \mathbf{b} \otimes \mathbf{n})\mathbf{F}_1, \quad \mathbf{R} \in \mathrm{SO}(3), \quad \mathbf{H} \in G_M, \quad \mathbf{b} \cdot \mathbf{n} = 0, \tag{4.6}$$

or the equivalent

$$\mathbf{F}_2^{-1} = \mathbf{H}^{-1}\mathbf{F}_1^{-1}\mathbf{R}^{\mathrm{T}} = \mathbf{F}_1^{-1}(1 - \mathbf{b} \otimes \mathbf{n}), \quad \mathbf{b} \cdot \mathbf{n} = 0, \tag{4.7}$$

where $\mathbf{F}_1^{-1}$ and $\mathbf{F}_2^{-1}$ are values of $\mathbf{F}^{-1}$ for some pair of minimizers of $\hat{w}$, $\mathbf{n}$ is the unit normal to a twin plane and $\mathbf{b}$ is the amplitude vector referred to in Ref. (2.8). Of course, using such minimizers picks out a particular value of M. With the usual understandings about material symmetry and the assumption that the Cauchy–Born rule applies, this is compatible with the twinning equation used in Eq. (3.8). This gives the same value of ρ for the two deformations, so it is easy to deal with Eq. (4.4), in such cases, and not hard for sequences involving a mix of Austenite and Martensite, for example. Essentially, Eq. (4.6) is the kinematic condition of compatibility, enabling one to construct piecewise homogeneous maps of the form (4.1), with $\mathbf{x}$ continuous, $\mathbf{F}^{-1}$ undergoing a finite jump from $\mathbf{F}^{-1}$ to $\mathbf{F}_2^{-1}$ across planes with normal $\mathbf{n}$, fitting $W^{1,\infty}$. Perhaps this is enough to indicate how one can redo calculations in the literature, using this procedure. I will not argue that, for elasticity theory, this procedure is better than the usual one. It does have one little advantage, in avoiding transforming descriptions obtained in the reference configuration. My reason for considering it is pragmatic, to be able to adapt such techniques to the X-ray theory, and only the latter is suitable for this.

5. Microstructures: X-ray theory

Recall Eq. (2.4) and the fact that, by specializing the choice of lattice vectors a bit, one can arrange that χ^a is continuous across twin planes. For example, referring to the lattice vectors in Eq. (3.8), we could use as $\mathbf{e}^{-a}$ and $\mathbf{e}^a$ those values involved in our example. Mathematically, $\mathbf{e}^a = \nabla\chi^a$, then, has essentially the same properties as the $\mathbf{F}^{-1}$ considered before, and there is an analogous twinning equation for these. The other vectors are shifts $\mathbf{p}_i$, $i = 1 \ldots n-1$, for an n lattice. As was the case in our example, these also can suffer finite discontinuities across twin planes, which can be analyzed, using Eq. (3.7). If you like, you can adjust this, using Eq. (2.14) on either side, as we did in our example, there to better fit the descriptions to a neighborhood. So, in place of Eq. (4.1), we have an energy density function, the form described in Eq. (2.10) being more appropriate than is that described in Eq. (2.9).

As an analog of Eq. (4.5), I proposed (Ericksen, 1997) using

$$\rho = k|\det \|\nabla\chi^a\|\|, \tag{5.1}$$

where k is a positive constant, again with essentially the same features as Eq. (4.5). This can exclude some variable continuous distributions of point defects, for example vacancies, a remark that also applies to the way mass is accounted for in elasticity theory.

For χ^a, the obvious analog of what is done in elasticity theory is to consider it to be in the function space $W^{1,\infty}$. For $\mathbf{p}_i$, the finite jumps in it are akin to those occurring in $\nabla\chi^a$, fitting the function function space L^∞. For this combination, the general theory of Young measures, etc. is available. In principle, one can then proceed as before, considering minimizing sequences for

$$E = \int_\Omega \rho\hat{\hat{\varphi}}\,\mathrm{d}v, \tag{5.2}$$

with

$$M = \int_\Omega \rho\,\mathrm{d}v \tag{5.3}$$

held fixed, as before.

In much of this kind of work in elasticity theory, workers do not really use the function $\hat{w}$, merely the assumption that it has certain minimizers. Using Pitteri's neighborhoods makes it rather easy to generate an orbit of minimizers that are also in the neighborhood. For the cases involving twins that cannot be included in one such neighborhood, such as one to be encountered in an example, one could use observations to estimate some minimizers, take their orbits under SO(3), and possibly use other general ideas of invariance to enlarge the list. In Section 6, I will suggest a possible strategy for this, which is speculative and only loosely defined, involving two modes in α-uranium. At least tacitly, one then assumes that such minimizers are in the domain of $\hat{\varphi}$. It is then a matter of analyzing microstructures which can be constructed, using these, and comparing them with observations. This is one way of testing some of the basic ideas used in the theory, as well as finding techniques that should still be useful when the theory becomes better developed.

For the analyses of microstructures that have been done, using elasticity theory, elementary analyses of twins are prerequisite. For the X-ray theory, the example treated earlier is the only such twinning analysis I have examined, although another will be treated later. Certainly, workers have done other twinning analyses, employing similar ideas. However, for me, at least, it is not really trivial to fit these to the X-ray theory, so, I need to take a hard look at each one. Given this, I think it better not to include an illustrative microstructure analysis, but to focus on the more elementary parts of twinning theory.

Finally, I note that the discussion of Christian and Mahajan (1995, Section 3.3) suggests that, when the element K_1 is irrational, as is the case in examples considered later, the "twin plane" is not really a plane, but a kind of saw tooth surface, consisting of very finely spaced steps, too fine to be observed with X-rays, at least. In a macroscopic limit, they are reasonably considered as infinitely fine. In such a limit, the crinkled surfaces should approach the observed plane, but I would expect the direction of the normals to have a limit described by Young measures. They make somewhat similar remarks about cases where η_1 is irrational. I do not fully understand the implications of these thoughts. However, they do suggest that there might be some interesting microstructure problems, exploring these ideas. Of course, anyone interested in this should look at all of the ideas concerning this which are presented by the aforementioned authors.

6. General theory of twinning

Here, my aim is to elaborate the more elementary parts of the general theory of twinning, according to the X-ray theory. This is designed to be compatible with twinning equations used by workers analyzing type I and type II twins, with information obtained from X-ray observations, as well as some permitting analyses of more general kinds of twins. As the atoms are arranged in different ways in the two regions separated by a twin plane, the twin plane direction K_1 can be determined, with the inevitable experimental error, using X-ray methods, it being relatively easy to do so. At least in principle, one can also determine how the atoms are arranged in both regions although, in practice, this can be difficult. We have seen some reasons why it is

better to use essential descriptions, and there are others. However, workers often use nonessential descriptions, so, one might well need to learn how to recognize these and do a translation of the description much like that done for the example considered earlier. The discussion to follow assumes that essential descriptions are used.

It is a very common understanding that, for such a pair of arrangements to be called a twin, in an unloaded crystal, one requirement is that the two arrangements can be related by some orthogonal transformation $\mathbf{Q}$, not necessarily unique. That is, if $\mathbf{e}_a$ or $\mathbf{e}^a$ and $\mathbf{p}_i$ describe the configuration on one side, then,

$$\hat{\mathbf{e}}_a = \mathbf{Q}\mathbf{e}_a, \hat{\mathbf{e}}^a = \mathbf{Q}\mathbf{e}^a \text{ and } \hat{\mathbf{p}}_i = \mathbf{Q}\mathbf{p}_i \tag{6.1a, b, c}$$

represent possible values of these vectors on the other side. In some writings, it is not entirely clear that the authors intend to require an equivalent of Eq. (6.1a), but I am fairly sure that they do. An expert on X-ray observations informs me that some slight departures from Eq. (6.1a,b,c) are tolerated, in practice, but I will ignore this. Obviously, that Eq. (6.1a,b,c) holds can be shown to be possible or not, if one knows how the atoms are arranged. Of course, one can use Eqs. (2.13) and (2.14) to get other descriptions, as we did in the example. If $\mathbf{Q}$ belongs to the point group for the n lattice, so Eq. (2.17) is satisfied for some values of the integers involved, the two configurations are the same, so this possibility is excluded. However, for n lattices, this does not always exclude the possibility that $\mathbf{Q}$ is in the point group for the lattice vectors only.

Another assumption is tacit in twinning tables. That is, they list just one entry for K_1, for example, although, since two different values of $\mathbf{e}^a$ occur, this direction might have different indices on the two sides. It is easy to see that, if they match for one choice of these two sets of vectors, one can make them different, by introducing an equivalent set on one side, leaving that on the other as is. So, this presumes special choices of these, the obvious possibility being to use pairs related as in Eq. (6.1a,b,c). Conditions obtaining from this are discussed by Zanzotto (1988, Note 2). I will satisfy the condition in a different way, which is less restrictive in this respect, but is more restrictive in others. For purposes of discussion, I'll take the conditions described as minimal requirements for a surface discontinuity to be called a twin. However, at the end, I will explain why this is not completely consistent with practice.

As might be expected from this, there are, in the literature, different definitions of twins and I have not tried to locate all of these. Those that I have inspected impose Eq. (6.1a,b,c) and some other conditions, depending on the definition. Some and perhaps all exclude some observed configurations called twins. To do common kinds of twinning analyses, one needs more than the minimal requirements noted above. Roughly, one aim is to assume as little as possible, consistent with this desideratum. Another is to sort out information which is relevant to the X-ray theory. Here, I will present my view of two classes which seem to me to be interesting, from this perspective, for what are considered to be unstressed crystals. For obvious reasons, these emphasize statements that can be verified by X-ray observations alone, although these are not the only kinds of observations of interest. One class deals with a subset of twins I will call *generalized type* I *twins*, defined as follows:

$$\begin{cases} \text{These satisfy Eq. (6.1), as interpreted above,} \\ K_1 \text{ is rational,} \\ \text{There is a choice of reciprocal lattice vectors } \bar{\mathbf{e}}^a \text{ equivalent to } \hat{\mathbf{e}}^a \text{ such that} \\ \bar{\mathbf{e}} = (\mathbf{1} - \mathbf{n} \otimes \mathbf{a})\mathbf{e}^a, \\ \bar{\mathbf{e}}^a = m_b^a \hat{\mathbf{e}}^b = m_b^a \mathbf{Q}\mathbf{e}^b, \quad \mathbf{m} \in G, \end{cases} \tag{6.2a–c}$$

where $\mathbf{n}$ is the unit normal to the twin plane and $\mathbf{a}$ is some vector. Here, Eq. (6.2a) just repeats Eq. (2.5), discussed earlier. The second equation merely mathematizes the assumption of equivalence. As is pointed out by Zanzotto (1988, Note II), it is always possible to pick lattice vectors so that two are parallel to the plane, which can help simplify analyses. It follows from these assumptions that

$$\det \|1 - \mathbf{a} \otimes \mathbf{n}\| = 1 - \mathbf{a} \cdot \mathbf{n} = \pm 1, \tag{6.3}$$

so, with the upper sign

$$\mathbf{a} \cdot \mathbf{n} = 0 \tag{6.4}$$

and with the lower

$$\mathbf{a} \cdot \mathbf{n} = 2. \tag{6.5}$$

I note that Eq. (6.4) includes the possibility that $\mathbf{a} = 0$ which is worth bearing in mind, for growth twins, in particular Dauphiné twins in quartz. For these, the simplified model used by James (1987) seems to do quite well. Unlike the Brazil twins, mentioned in Section 2, these can be removed by mechanical treatments, as was discovered during World War II and later reported by Thomas and Wooster (1951). For this reason, it does seem sensible to use the same constitutive equation for both configurations.

I interpret assumption (b) as

$$\mathbf{n} = \mathbf{K}_1/|\mathbf{K}_1|, \quad \mathbf{K}_1 = k_a \mathbf{e}^a, \tag{6.6}$$

where k_a are relatively prime integers. Then, using Eq. (6.2a), (6.4)–(6.6), we have

$$k_a \bar{\mathbf{e}}^a = \mathbf{K}_1 - \mathbf{a} \cdot \mathbf{K}_1 \mathbf{n} = \mathbf{K}_1 - \mathbf{a} \cdot \mathbf{n} \mathbf{K}_1 = \pm \mathbf{K}_1, \tag{6.7}$$

which is commonly interpreted as acceptable matching of the two values of K_1. This does not assume that $\mathbf{e}^a$ and $\bar{\mathbf{e}}^a$ are related by an isometry, as was discussed above, although it achieves the matching of indices. For this, it is not necessary that the k_a be integers, and this will be important when we consider the second class. However, a useful result follows when they are. Suppose that Eq. (6.5) applies. With the k_a being relatively prime integers, there are integers l^a such that

$$k_a l^a = 1, \tag{6.8}$$

by elementary number theory. Then,

$$\hat{\mathbf{m}} = \|\delta^a_b - 2k_b l^a\| \in G, \quad \text{with} \quad \hat{\mathbf{m}}^2 = 1, \quad \det \hat{\mathbf{m}} = -1, \tag{6.9}$$

so we can introduce equivalent reciprocal lattice vectors given by

$$\bar{\bar{\mathbf{e}}}^a = \hat{m}^{a-b}_b \mathbf{e} = \hat{m}^a_b (1 - \mathbf{n} \otimes \mathbf{a}) \mathbf{e}^b, \quad \mathbf{a} \cdot \mathbf{n} = 2. \tag{6.10}$$

By a routine calculation, this gives

$$\bar{\bar{\mathbf{e}}}^a = (1 - \mathbf{n} \otimes \bar{\mathbf{a}}) \mathbf{e}^a, \tag{6.11}$$

where

$$\bar{\mathbf{a}} = \mathbf{a} - 2|\mathbf{K}_1| l^a \mathbf{e}_a \Rightarrow \mathbf{n} \cdot \bar{\mathbf{a}} = 0. \tag{6.12}$$

Similarly, if Eq. (6.4) holds, we can transform it in a similar way to have Eq. (6.5) apply, so, the two versions are equivalent, in this sense, when K_1 is rational. This has its limits. There are type II twins which are in this class, the compound twins. For these, the directions of $\mathbf{a}$ or $\bar{\mathbf{a}}$ can be obtained from X-ray observations, as was mentioned in Section 3, implying that $\bar{\mathbf{a}}$ and $\mathbf{a}$ are not always physically equivalent. However, one can still use the idea to transform twinning equations using Eq. (6.4) to equivalents using Eq. (6.5), or vice versa, which might be helpful, for theoretical studies. It is a common notion that compound twins can be analyzed as type I and as type II. As was mentioned in footnote 1, this is subject to interpretation, as different definitions of the types are in the literature. Properly interpreted, this is true, as far as the lattice vectors are concerned, as is discussed by Zanzotto (1988). Ponder what this presumes about shifts, and you should see that this might not always be true for n lattices. However, my experience is that, in practice, when K_1 and η_1 are reported as rational, it does mean that they can be analyzed either way and,

soon, I will explain what I see as the reasons for this. Those accustomed to use the mathematical definitions of rational and irrational numbers should be aware of the fact that this is not exactly what workers in this area mean by the words. To them, it seems to be like picking a number at random from some interval, when you do not know in advance that it must be rational. Since the rationals are only a countable subset of the real numbers, the number picked will be irrational, almost certainly.

In the remainder of this paper, I will consider only cases for which Eq. (6.4) holds. Then, Eq. (6.2a) is equivalent to

$$\bar{\mathbf{e}}_a = (1 + \mathbf{a} \otimes \mathbf{n})\mathbf{e}_a, \tag{6.13}$$

used earlier and, with Eq. (6.1b), we get what looks like the more standard twinning equation

$$\bar{\mathbf{e}}_a = \bar{m}_a^b \mathbf{Q}\mathbf{e}_b = (1 + \mathbf{a} \otimes \mathbf{n})\mathbf{e}_a, \quad \bar{\mathbf{m}} = \mathbf{m}^{-1}. \tag{6.14}$$

However, when the Cauchy–Born rule fails, this equation can still be used, but it cannot be satisfied for $\mathbf{a} = \mathbf{b}$, the vector indicated in Eq. (2.8). As will be described later, various workers then use a variation on Eq. (6.14), with **a** replaced by **b** From this view, my procedure is not conventional. Some workers interested in mechanical twins use arguments about shearing deformations to motivate using Eq. (6.14). To me, this muddies the water. If the reasoning were sound, Eq. (6.14) should be satisfied when we take as $1 + \mathbf{a} \otimes \mathbf{n}$ the corresponding deformation gradient and, often, it is not. Also, those deformations are clearly irrelevant to growth twins. In my way of doing it, those shears need not be considered. So, Eq. (6.14) is not reliable, as an equation relating to such shears, my reason for interpreting it differently. I will try to make clear that the decisions as to whether twinning elements are rational or irrational are based on predictions obtained from Eq. (6.14) or some similar equation. As to how well Eq. (6.14), properly interpreted, can do in describing crystallographic features of twins, I will say that it looks very promising, as a tool for analyzing X-ray observations. However, I believe that too little has been done to use and test it, so I am trying to pull together ideas for doing better with this. Understandably, workers concerned with mechanical twins use various bits of theory to try to master these. I find no fault with this, but it can make it harder to see the faults and virtues of Eq. (6.14).

Consider the very common cases for which X-ray observations are consistent with a type I description, with

$$\mathbf{Q} = \pm\mathbf{R}, \quad \mathbf{R} = -\mathbf{1} + 2\mathbf{n} \otimes \mathbf{n}. \tag{6.15}$$

While analysis of Eq. (6.14) for these is very familiar to those who have done twinning calculations [4], I will belabor it, to make some points. First, note that, if it is satisfied with $(\mathbf{R}, \bar{\mathbf{m}})$ it is also with $(-\mathbf{R}, -\bar{\mathbf{m}})$ although, in general, only one of these will be consistent with Eq. (6.1c). However, for (monatomic) one and two lattices, both hold when one does, and this includes our example.

Taking the former, consider the equivalent

$$\mathbf{H}\mathbf{e}_a = \bar{m}_a^b \mathbf{e}_b, \quad \mathbf{H} \stackrel{\text{def}}{=} \mathbf{R}(\mathbf{1} + \mathbf{a} \otimes \mathbf{n}), \tag{6.16}$$

and verify that

$$\begin{cases} \mathbf{H}^2 = 1, \quad \det \mathbf{H} = 1 \Rightarrow \bar{\mathbf{m}}^2 = 1, \quad \det \bar{\mathbf{m}} = 1, \\ \mathbf{H}^T\mathbf{n} = \mathbf{n} \Rightarrow \bar{m}_a^b k_b = k_a. \end{cases} \tag{6.17}$$

It seems to be consistent with experience that type I twins always have K_1 rational, as is required for Eq. (6.16) to have any solutions. However, with the inevitable experimental errors, measurements cannot really

[4] See Zanzotto (1988, Note 1) and Pitteri (1985b), for example.

confirm or contradict this. Thus, necessarily, workers rely on some theory to decide this, using equations more or less like Eq. (6.16). The conditions on $\bar{\mathbf{m}}$ require that, for some set of integers l^a satisfying Eq. (6.9),

$$\bar{m}^b_a = -\delta^b_a + 2l^b k_a, \tag{6.18}$$

one point being that X-ray observations of type I twins that are not compound can yield values of k_a, but not of l^a or of the direction of the vector **a**. So, here, we can use Eq. (6.4) and regard it as physically equivalent to Eq. (6.5). Also, if Eq. (6.9) is satisfied by l^a, it is also satisfied by

$$\bar{l}^a = l^a + r^a, \quad r^a k_a = 0, \tag{6.19}$$

where the r^a must be integers, of course. For any choice of these integers, there is a vector **a** such that Eq. (6.14) is satisfied, given by

$$\mathbf{a} = 2(\mathbf{n} - |\mathbf{K}_1| l^a \mathbf{e}_a) \Rightarrow \mathbf{a} \cdot \mathbf{n} = 0. \tag{6.20}$$

From Eq. (6.20), it is easy to see that replacing l^a by $\bar{l}^a$ amounts to adding in a lattice invariant shear, not detectable in X-ray observations. So, understandably, theory gives us this ambiguous estimate of **a**. Here, K_1 is the only twinning element used, although one does need more information to determine that a twin is of type I. Excepting compound twins, I do not see how one could determine other elements from X-ray observations alone, without adding some hypotheses, for twins in this class: sometimes, workers do add such hypotheses, to estimate other elements.

Now, how is η_1 determined? According to most expositions, it describes the direction of the vector **b** in Eq. (2.8). I have come to believe that this is not the only interpretation used, in practice. Try the following experiment: Select any twinning table, and ignore all entries for which η_1 is described as rational. I will deal with these, later. This will leave you with a much shorter list: if there are none left, try another table. Now, select one such mode. It is rather likely that this will give no quantitative information for η_1. To understand this, it might help to consider a case history. Cahn (1953) did pioneering work in determining twinning elements for the modes he observed in α-uranium, and he used some ingenuity in doing so. For example, the data in Eq. (3.3) are his, but we are ignoring these. However, his observations of $\{1\,1\,2\}$ twins are relevant. He concluded that these are type I twins, with η_1 irrational, but was unable to get quantitative estimates of it. One might think it a simple matter to measure the direction of **b** but, in practice, it can be very difficult and, for him, it was not feasible to get good data of this kind. To infer that η_1 is irrational, he used an indirect argument. He also observed "$\{1\,\bar{1}\,2\}$" twins, to be analyzed later and concluded that these two modes are conjugate, using theory relating to this to draw this conclusion. This gives enough information to enable you to make a theoretical estimate of η_1, if you want to. Thus, these experiments really gave no information about η_1, and this seems not to be an extremely unusual difficulty. Often, if one mode is observed, its conjugate is not, as seems to be the case for the $\{1\,3\,0\}$ twins, for example. Thus, one cannot always use the reasoning based on this, which helped Cahn. So, an answer to the question posed is that η_1 is not always determined and, if it is, it might be by a theoretical estimate, or by some experiment.

The set of integers r^a can always be represented parametrically, with two arbitrary integers as parameters. In our example, we saw only one. A calculation shows that, had this not been compound, or had we not noticed this, the shear in Eq. (3.31) should have been replaced by

$$1 + [m'(2\tilde{\mathbf{e}}_1 + \tilde{\mathbf{e}}_2) + r\tilde{\mathbf{e}}_3] \otimes (2\tilde{\mathbf{e}}^2 - \tilde{\mathbf{e}}^1), \tag{6.21}$$

where r is an arbitrary integer. A calculation shows that $r = 0$ for the minimum shear. So, in this case, using this would cancel the error of omission, since it is compound. In other cases involving type I twins, it certainly is better practice to determine whether the twin is compound. For compound twins, the relevant r^a in Eq. (6.19) reduce to a one parameter family, as in our example.

For the X-ray theory, I believe that Eq. (6.16) is reliable, for locating the possible energy wells associated with the twins considered. However, only a small number of these are likely to be relevant, physically.

When it is feasible to include some in a Pitteri neighborhood, it seems to be a good rule to use these. However, there are many cases for which this is impossible. So, we need, but do not yet have a good alternative, for such cases. Certainly, this is a road block that needs to be surmounted, to develop a good theory of twin patterns and related microstructures. Later, I will mention a few thoughts about this.

From Eq. (6.20) follows another point, that, for type I twins that are not compound, Eq. (6.2a) is redundant, and observations agree with the view that all of these are generalized type I twins. That is, given a rational K_1 and any set of lattice vectors, Eq. (6.16) can always be satisfied. For compound twins, when K_1 and η_1 are determined by X-ray observations, one should confirm that these are compatible with Eq. (6.2a). If not, I would conclude that, most likely, there is either some fault in the experiments or in the interpretation of them, but I am biased. Shortly, I will indicate how η_1 has been determined using X-ray observations, for some twins. Most twins are either of type I or can be described as being both of type I and of type II, so most observed twins are in this class.

Now, I turn to the second class, consisting of what I will call *generalized type* II *twins*. In particular, this covers the less common observations of type II twins with K_1 considered to be irrational. By private communication, Richard James informs me that these are in fact rather common in copper based shape-memory alloys. As was mentioned before, the Cauchy–Born rule seems to apply to all shape-memory alloys. Examples of such twins in other kinds of crystals seem to be rare. To define the class, simply replace statement (b) in Eq. (6.2a–c) by

$$(\mathrm{b})'\ K_1 \text{ is irrational.} \tag{6.22}$$

This allows for the possibility of observing examples not of type II and, later, I will mention a recently discovered example. One does run into rather different problems in analyzing the generalized type II twins, so there is some reason to put them in a separate category.

Before, I mentioned the interpretation of η_1 commonly found in expositions. However, in practice, I find that another one is also used, which is not obviously equivalent. For the following discussion, I will take at face value the following statement by a well-known experimentalist, Cahn (1953), as describing how η_1 can be determined, for type II twins:

> "For a twin of the second kind, the orientations of parent and twin are related by a rotation of 180° about η_1 as axis."

Certainly, he did use this, in estimating η_1 in cases where he could not get good measurements of deformation. So, what is to be determined experimentally is this axis, something that can be determined using X-ray observations, at least in principle. After pondering this and other bits of evidence, I concluded that the best way to mathematize this is as follows: Proceed by satisfying Eq. (6.14) with

$$\mathbf{Q} = \pm\mathbf{R}, \quad \mathbf{R} = -\mathbf{1} + 2\mathbf{v}\otimes\mathbf{v}, \quad \mathbf{v} = \mathbf{a}/|\mathbf{a}|, \tag{6.23}$$

this being what I take as a definition of type II twins, for the X-ray theory. Except for allowing $\mathbf{Q} = -\mathbf{R}$, it is just my interpretation of the quotation above. In this, I also include compound twins. As suggested by the quotation, I will consider η_1 as describing the direction of $\mathbf{a}$. For analyzing the twinning equation, it is enough to consider the upper sign, as before. I still use Eq. (6.4), which seems to fit the examples observed. This differs from the previous case in that $\mathbf{v}$ and $\mathbf{n}$ can be obtained from X-ray observations, with some experimental errors. Here, it is known [5] that for Eq. (6.14) to be soluble, it is necessary that

$$\mathbf{v} = \boldsymbol{\eta}_1/|\boldsymbol{\eta}_1|, \quad \boldsymbol{\eta}_1 = t^a\mathbf{e}_a, \tag{6.24}$$

[5] See Zanzotto (1988), for example.

where the t^a are relatively prime integers. That is, η_1 must be rational. With this or some similar guide, workers will pick the numbers to fit the data to within the errors in these. It is also known that the equation does not require K_1 to be rational. The practice is to call it irrational unless there is some good reason to believe that it is rational. The likely reason is that it can be described equally well as a type I twin. Then, the equations require that both K_1 and η_1 be rational. I believe that this is the real reason why these elements are always reported this way, for what are called compound twins. In twinning tables, I have not yet found an example of a type I twin which is not compound, for which η_1 is reported as rational. So, my impression is that η_1 is judged to be irrational for any type I twin that is not compound. This makes it a pretty safe bet that, if these entries are reported as rational, it implies that the twins can be analyzed as type I or as type II. As I said, I am interpreting η_1 as representing the axis of rotation. As was noted before, from the X-ray observations alone, there is no reliable theory for determining the vector $\mathbf{b}$ in Eq. (2.8). In particular, Eq. (6.14) is not always satisfied with $\mathbf{a} = \mathbf{b}$. So, if we insist that η_1 represents the direction of $\mathbf{b}$, how do workers conclude that η_1 is rational, in these cases? Certainly, workers accept this conclusion and, later, I will explain how they use a variation on the twinning equation mentioned earlier, to obtain such conclusions. For these twins, one could assume that $\mathbf{b}\|\mathbf{a}$ weaker than $\mathbf{a} = \mathbf{b}$ which seems to apply to twins observed in α-uranium, at least, avoiding an inconsistency with the idea that η_1 represents the shear direction. I just do not know whether twins are observed, which violate this assumption. I think that failure of this assumption is a theoretical possibility, and that it might avoid confusion, if we were to use different notations for the two interpretations. Of course, this assumption is not really relevant to the X-ray theory, but one interpretation of η_1 is. Until someone proves me wrong, I will assume that the values reported in tables are consistent with my definition. I think it clear that X-ray observations should be used to determine the direction of this axis and, if only such observations are available, this is how η_1 would be determined, in practice. So, it seems to me to be a good definition for the X-ray theory, at least. One could make the condition that $\mathbf{b}\|\mathbf{a}$ part of the definition of type II twins. However, experimentalists can only check this approximately, with a margin of error which is not always so small. Also, this does not make sense for growth twins, or for twins for which it is not known how they formed. So, I do not like this idea.

From Eq. (6.4), there is some scalar α, such that

$$\mathbf{a} = \alpha \boldsymbol{\eta}_1. \tag{6.25}$$

Then, one question is whether one can always choose α so that Eq. (6.14) is satisfied, with $\mathbf{Q} = \mathbf{R}$, given by Eq. (6.23) and $\mathbf{a}$ by Eq. (6.25). This reduces to the question of whether one can always find relatively prime integers s_a satisfying

$$s_a t^a = 1 \tag{6.26}$$

such that

$$\alpha\mathbf{n} + 2\boldsymbol{\eta}_1/|\boldsymbol{\eta}_1|^2 = 2\mathbf{s} \stackrel{\text{def}}{=} 2s_a\mathbf{e}^a \tag{6.27}$$

is satisfied for some value of α. It is not hard to show that one can pick lattice vectors, integers t^a and a unit vector $\mathbf{n}\perp\boldsymbol{\eta}_1$ for which this is impossible. So, one could argue that, hypothetically, there are type II twins for which Eq. (6.14) is not satisfied. Of course, with my definition of type II, this is impossible. I doubt that workers using another would accept the possibility that one will encounter a realization of this, in nature.

This does make it desirable to explore a realistic example. For α-uranium, there are the type II "$\{1\bar{7}2\}$" twins, with $\eta_1 = \langle 3\,1\,2\rangle$, the quotation marks indicating that this is a rational approximation to K_1, considered to be irrational. These data were produced by Cahn (1953), so η_1 represents the axis of rotation. The entries for these two elements agree with Cahn's and listings in tables presented by Hall (1954) and Klassen-Nekliudova (1964), for example, but the table given by Barrett and Massalski (1966) gives K_1 = "$\{1\,7\,2\}$" and the same η_1. This is not consistent with the fact that these directions are orthogonal. The table

presented by Christian and Mahajan (1995) also does this and puts the entries for η_1 and η_2 in the wrong places.

Now, it is easy to show that, for this mode, the Cauchy–Born rule fails to apply. Essentially, the basic idea is what I used to get (3.16). For this, one uses twinning elements describing shear. In Zanzotto (1992), a list of linear transformations for α-uranium, it is H_2 that is associated with this mode. He uses the usual four-lattice description. Apply this to these lattice vectors and express the result as a linear combination of the lattice vectors. Do it again, using the essential description. For the Cauchy–Born rule to apply, at least the latter coefficients should be integers, forming an element of G. In both cases, one finds that they are rational numbers, but not integers. Actually, various workers consider the twinning equation to be generalized, allowing such rational numbers as well as integers, this being the variation on the twinning equation mentioned earlier. Here, the equation is phrased as one involving the shear deformation. This does not change conclusions about which twinning elements are rational. Essentially, this is a way of describing the observations supporting the view, mentioned in Section 2, that one can always find a sublattice to which the Cauchy–Born rule applies. If you perform the above calculations, you should be able to find the sublattices for the two descriptions, and determine whether these are the same or different. Here, what I am doing is unconventional, assuming that Eqs. (6.14) and (6.23) apply, with an unconventional interpretation of these, despite the failure of the Cauchy–Born rule. I do believe that this is sound. Of course, I concede that it is possible that someone could find clear evidence that my belief is wrong.

Those interested in constitutive theory need to be aware of this, and try to take it into account. When I considered this, I concluded that a sensible form of constitutive equations for the energy density is what I used for the X-ray theory, so, I am interested in learning what can be done with it. My view is that, if identical atoms somehow exchange positions, this does not affect the energy. Of course, using the X-ray theory does not preclude the common practice of introducing other hypotheses to relate the shear deformation to lattice vectors and shifts. Simply, I do not see any good way of relating these to a sound theory of constitutive equations. This is an important and challenging open problem, in need of a good solution.

I will start by taking the approximation as exact. The indices refer to the four-lattice description described in Section 3. Converting them to the essential description used there, I get

$$\mathbf{K}_1 = 4\tilde{\mathbf{e}}^1 - 3\tilde{\mathbf{e}}^2 + 2\mathbf{e}^3, \tag{6.28a}$$

$$\boldsymbol{\eta}_1 = \tilde{\mathbf{e}}_1 + 2\tilde{\mathbf{e}}_2 + \tilde{\mathbf{e}}_3, \tag{6.28b}$$

$$\mathbf{s} = s_a\tilde{\mathbf{e}}^a, \tag{6.28c}$$

with the integers s_a satisfying

$$s_a t^a = \boldsymbol{\eta}_1 \cdot \mathbf{s} = s_1 + 2s_2 + s_3 = 1. \tag{6.29}$$

Then, Eq. (6.27) gives three equations for α: one can use the information on the lattice vectors given in Section 3 to calculate the necessary entries. By routine calculation, I get, as conditions for these to admit a solution for α,

$$2(s_1 + s_2) - s_3 = x^{\mathrm{def}} = (12a^2 - 4c^2)/z, \tag{6.30}$$

and

$$s_1 - 2s_2 = y^{\mathrm{def}} = (3a^2 - b^2 - 8c^2)/z, \tag{6.31}$$

where

$$z = 9a^2 + b^2 + 4c^2. \tag{6.32}$$

It is easy to pick numbers a, b and c such that x is not very close to an integer, for example, so these equations are violated. However, using data presented by Barrett and Massalski (1966, p. 170), I calculate that, approximately,

$$x = -0.001, \quad y = -1.001, \tag{6.33}$$

quite close to $x = 0$, $y = -1$. So, interpret this as an error in what we took for K_1, but take the values of s_a suggested by this, which are

$$s_1 = 4n - 1, \quad s_2 = 1 - 3n, \quad s_3 = 2n, \tag{6.34}$$

where n is any integer, and t^a can be read off from Eq. (6.28b). Again, this would amount to adding lattice invariant shears of a particular kind, invisible to X-rays, if the starting estimate of K_1 had satisfied the twinning equation. For any choice of n, one can solve the twinning equation for $\mathbf{n}$ and α, thereby getting an infinite number of possibilities for both. It is not hard to check that all these directions are very close to that given by Eq. (6.28a), when n is very large, and that α is very large when n is. Define κ_a by

$$\alpha\mathbf{n} = \kappa_a \mathbf{e}^a \Rightarrow \kappa_1 + 2\kappa_2 + \kappa_3 = 0. \tag{6.35}$$

As an example of numbers obtained for n small, I calculate that

$$n = 0 \Rightarrow \kappa_3/\kappa_1 = 0.5005, \tag{6.36}$$

also quite close to the starting value of 1/2. From this, it is pretty clear that, while different values of n will give different values of $\mathbf{n}$, they are not very different, although the values of α can differ considerably. Of course, values of these depend on values of a, b and c, which are subject to some experimental error, changes of temperature, etc. There is no obvious reason why the relevant combinations of these should be rational numbers, so K_1 is regarded as irrational. From this exercise, it does seem pretty clear, on the face of it, that workers have done calculations somewhat similar to mine, to draw their conclusions about K_1 being irrational. Given these calculations, I seriously doubt that experimentalists can get X-ray data sufficiently accurate to pick out a particular value of n. However, I would be happy to be proven wrong about this. Even so, I would like to see some theoretical reason for picking one. Earlier, I noted that, for this mode, in particular, the Cauchy–Born rule fails to apply, so we cannot expect to get a reliable determination of n, by using shear data. Another theoretical possibility is to try to determine it so that the twin is contained in some Pitteri neighborhood. This involves satisfying three equations like Eq. (3.33), the difference being that the $\mathbf{m}_1$ used there is replaced by

$$\mathbf{m}_2 = \| - \delta^b_a + 2s_a t^b \| \Rightarrow \mathbf{m}_2^2 = 1, \quad \det \mathbf{m}_2 = 1. \tag{6.37}$$

However, it is not hard to show that there are no lattice vectors satisfying these conditions for any permissible values of the s_a, so

these twins cannot be included in any Pitteri neighborhood. (6.38)

At least for the present, I do not see another possibility that is easy to assess. However, there is another speculative line of thought that seems to me to be promising enough to mention, although it would take some hard work to firm it up and assess it.

Now, I return to the speculation. First, Cahn (1953), in his work on determining twinning elements for α-uranium, presented a number of photographs of patterns of twins he observed, including some involving the modes considered in our examples. Mostly, these involve differently oriented twins. For example, he found that "$\{1\bar{7}2\}$" twins can intersect each other, also that they can intersect $\{1\,3\,0\}$ twins. So, these are among the patterns that the X-ray theory should treat. I have not tried to collect other information of this kind that is available, but this is enough to supply some motivation for the discussion.

Now, in common practices, I observe an intuitive prejudice concerning analyses of twins. Roughly, it is a common notion that interactions involving different energy wells will involve only wells which are very close to each other. In one way or another, various workers use this idea, rather successfully, on the whole. In our first example, we used two versions, one being the minimum shear assumption. The other is more of a topological nature. With twins that can be included in a neighborhood, it has become routine to use this to locate neighboring wells, this being the second version used in the first example. Other variations are used, in dealing with deformation. Mathematically, it is at best unclear that such neighborhoods always include the wells closest to one, and it would not surprise me if someone produced a counterexample, with a reasonable interpretation of "closest". Pragmatically, it has been a successful selection criterion, for cases to which it applies.

For the $\{1\,3\,0\}$ twins, we used this rule to get the three variants discussed before and experience suggests that the corresponding wells are likely to suffice, to describe their role in patterns of twins. Clearly, this idea is of no help, in deciding which wells to use, to try to analyze patterns involving the "$\{1\,\bar{7}\,2\}$" twins, in particular, and we know that these can also involve $\{1\,3\,0\}$ twins. Roughly, what the indicated thoughts suggest to me is selecting the "$\{1\,\bar{7}\,2\}$" wells closest to those of the $\{1\,3\,0\}$ variants, and rejecting the rest. As a general strategy, I like this way of getting rid of most or all of the ambiguities associated with that arbitrary integer, etc. The difficulty is that it is not obvious exactly how best to accomplish this, or to forecast how well it might do, in delivering wells needed for satisfactory analyses. Let us think a bit more about this, in general terms. For one variant, we did determine that infinite set of solutions of the twinning equations. Using Eqs. (2.14) and (6.1a,b,c), one can calculate all possible shifts, to complete these solutions. We should also determine the corresponding sets for the other two variants. This is routine, a matter of applying transformations to the first set. If we like, we can take orbits of these under SO(3). At present, I have not firmly decided whether to do so, but I lean toward it. Either way, one should take each $\{1\,3\,0\}$ variant, look at the entire list of "$\{1\,\bar{7}\,2\}$" possibilities, and pick out the one(s) closest to it. Here, we have another uncertainty, "closest" being subject to interpretation. There are various possible norms that could be used for this. For reasons not yet very clear to me, I think that some are better than others, so I am not ready to make a definite proposal for this. I do not mean to exclude the possibility that some more topological interpretation might be best, but I have no concrete suggestions for this. Obviously, one needs to make definite decisions about these uncertainties, and determine the results. I believe that it should be possible to do so, but the best way to confirm this is to produce the results. So, my proposal is rather vague and speculative. Granted a successful outcome, one should take the orbit of the selected descriptions under SO(3).

Suppose that we have done all this. Then, we have a collection of energy wells, to be used in trying to analyze patterns involving one or both of these modes. There is no guarantee that the procedure will deliver the wells needed to properly analyze the observed patterns, so this needs to be explored. For this reason also, my proposal is speculative. The common practice of using only the wells included in a Pitteri neighborhood is subject to the same reservation but, in practice, it has worked well, giving us some reason to be optimistic about this. For whatever it is worth, Frank (1953) used a somewhat similar idea, to compare twins in α-uranium with those in zinc. If the procedure works well for these modes, one could try including other observed modes, in a similar way. So, with some luck, this could become a useful partial theory of twins in α-uranium. If I had an idea which I thought would be more likely to be successful than this, I would have discussed it instead. For this, it seems necessary to assume that $\hat{\varphi}$ is invariant under an infinite group, unless someone finds a clever trick to evade this. However, for typical analyses, it is not necessary to specify this function, although one is likely to need it in the future. As I have not found a good way of selecting that arbitrary integer etc., this is, for me, a problem not yet solved. For this reason, I have referred to my analysis of these twins as being only a partial one.

Whether or not the suggested procedure is successful, I have barely scratched the surface in constructing a useful theory of twinning in this material. Here, my aim is more to illustrate the issues arising in realistic

cases, and to describe the tools we have for dealing with X-ray observations of twins. The two quite different examples seem to me to be good for this. I think it worth work out additional realistic examples, to better illustrate all of the kinds of difficulties that do arise, and need to be dealt with, perhaps by creating other kinds of tools.

For twins with K_1 irrational, it is not immediately obvious how this should be defined. As best as I can estimate the consensus of opinion about this, it is to use ratios of components of $\mathbf{n}$. So, for our example, we might use

$$K_1 = \{1, \kappa_2/\kappa_1, \kappa_3/\kappa_1\}. \tag{6.39}$$

A different kind of example of a generalized type II twin occurs in orthoniobate. These are transformation twins, associated with a second-order phase transition. I have not studied the literature on these, except that I did look carefully at the analysis of observed microstructures by Jian and James (1997), which agrees well with the experimental data. This alone is rather good evidence that these are neither of type I nor of type II, with K_1 and η_1 both being irrational. This is the first such twin to be observed, as far as I know.

This is one of those nice cases where the Cauchy–Born rule applies and the twins can be included in a neighborhood. The former is expected, this being a shape-memory material. The latter is rather obvious from the fact that these twins are associated with a second-order phase transition. Then, it is not really necessary to use the X-ray theory, thermoelasticity theory being adequate for analyzing these. This is what Jian and James (1997) do. Of course, this theory ignores the shifts. I do not really doubt that they are arranged properly and, otherwise, these twins do qualify as generalized type II twins that are not of type II. I have not tried analyzing these, using the X-ray theory, but would expect to encounter ambiguities similar to those encountered in the last example, associated with lattice invariant shears. It might be worthwhile to do so, to understand how the nice features in this case enable one to eliminate these. Or, perhaps, they are not completely eliminated, but leave us with inconsequential ambiguities.

There is the common idea that, if K_1 and η_1 are rational, they can be treated as either type I or type II. As I explained before, this is, essentially, a tautology, if you recognize how it is decided that they are rational. This seems to cover the observations of generalized type I twins, but I have not studied the mathematical possibility of including others, for which K_1 should be judged to be rational, by similar reasoning. I should say that, in practice, some workers say that calling K_1 irrational really means that one needs rather large integers k_a to match the measurements, within experimental error. Apart from the fact that this seems to require some subjective judgment, and makes the theoretical distinction fuzzy, I do not really object to this. Perhaps, some of them will not like my explanation of this. I believe that the last α-uranium example is fairly typical, illustrating the kinds of issues which will arise, when one tries to analyze type II twins with K_1 irrational. Another such twin of this kind, also observed in α-uranium, involves an additional ambiguity. With $\eta_1 = \langle 5\,1\,2\rangle$, it is either the "$\{1\,9\,7\}$" or the "$\{1\,7\,\bar{6}\}$" mode. Some writers use one, some the other, and a few persons mention both. Of course, rational approximations are not unique, but these seem to me not to be very close to each other. Some workers mention that the rational approximation for the "$\{1\,\bar{7}\,2\}$" mode is unusually good. I have not tried analyzing these, so am not sure what problems are created by the added ambiguity. Generally, some twins of this kind might be simpler than others, because they are contained in a Pitteri neighborhood and/or conform to the Cauchy–Born rule, for example.

Finally, it is fair to ask whether there are things called twins which are not in either of my classes. I know of no observations of mechanical twins of this kind, but there are some such examples of growth twins, as is rather clear from the discussion by Zanzotto (1988). For example, he mentions cases of growth twins in some materials, for example alum, which do not even meet what I took as minimal requirements: they involve inequivalent crystallographic planes on the two sides of the twin plane. Thus, they cannot be listed in the usual way, in twinning tables. Frankly, I do not see how to phrase a general definition of a twin which includes such cases and also excludes the grain boundaries in polycrystals, for example, and rather different

kinds of theory are used for the latter. So, I would not call these twins, but I seem to be alone in this. I do not really object to having a third category of twins, allowing for any not in either of my classes. It is only that less theory is available for these. It would be nice for theorists if workers would agree on a general definition of twins, but I am not very optimistic about this.

This covers my thoughts on what might be viewed as elementary twinning analyses, according to the X-ray theory. Hopefully, this is enough to make clear how well these fit common practices, and to give some idea of what the X-ray theory can and cannot do. For nonspecialists, I have tried to point out practices which seem to me to be somewhat confusing.

Acknowledgements

I thank Richard James, Mario Pitteri and Giovanni Zanzotto for some very helpful suggestions, and Marion Ericksen, for her help in getting the typing done.

References

Adeleke, S.A., 1999. On the classification of monoatomic crystal multilattices. submitted for publication.

Ball, J.M., James, R.D., 1992. Proposed experimental tests of a theory of fine microstructures and the two-well problem. Philosophical Transactions of the Royal Society of London 333A, 389–450.

Barrett, C.S., Massalski, T.B., 1966. Structure of Metals. Third edn., McGraw-Hill, New York.

Bhattacharya, K., Firoozye, N.B., James, R.D., Kohn, R.V., 1994. Restrictions on microstructures. Proceedings of the Royal Society of Edinburgh 124A, 843–878.

Born, M., 1923. Atomtheorie des festen Zustandes. Second edn., Teubner, B.G., Leipzig.

Cahn, R.W., 1953. Plastic deformation of alpha-uranium; twinning and slip. Acta Metallurgica 1, 49–67.

Christian, J.W., Mahajan, S., 1995. Deformation twinning. Progress in Materials Science 39, 1–157.

Ericksen, J.L., 1997. Equilibrium theory for X-ray observations. Archive for Rational Mechanics and Analysis 139, 181–200.

Ericksen, J.L., 1998. On non-essential descriptions of crystal multi-lattices. Journal of Mathematics and Mechanics of Solids 3, 363–392.

Ericksen, J.L., 1999. On groups occurring in the theory of crystal multi-lattices. Archive for Rational Mechanics and Analysis 148, 145–178.

Frank, F.C., 1953. A note on twinning in alpha-uranium. Acta Metallurgica 1, 71–74.

Hall, E.O., 1954. Twinning and diffusionless transformations in metals. Butterworths, London.

James, R.D., 1987. The stability and metastability of quartz. In: Antman, S., Ericksen, J.L., Kinderlehrer, D., Müller, I. (Eds.), Metastability and Incompletely Posed Problems. Springer, New York, pp. 147–175.

James, R.D., Kinderlehrer, D., 1989. Theory of diffusionless phase transformations. In: Rascle, M., Serre, D., Slemrod, M. (Eds.), Lecture Notes in Physics, vol. 344. pp. 51–84.

Jian, L., James, R.D., 1997. Prediction of microstructure in monoclinic $LaNbO_4$ by energy minimization. Acta Materialia 45, 4271–4281.

Kelly, A., Groves, G.W., 1970. Crystallography and crystal defects. Addison-Wesley, Reading, MA.

Klassen-Nekliudova, M.V., 1964. Mechanical twinning of crystals. Consultants Bureau, New York.

Parry, G.P., 1998. Low-dimensional lattice groups for the continuum mechanics of phase transitions in crystals. Archive for Rational Mechanics and Analysis 145, 1–22.

Pitteri, M., 1984. Reconciliation of local and global symmetries of crystals. Journal of Elasticity 14, 175–190.

Pitteri, M., 1985a. On (v+1)-lattices. Journal of Elasticity 15, 3–25.

Pitteri, M., 1985b. On the kinematics of mechanical twinning in crystals. Archive for Rational Mechanics and Analysis 88, 25–57.

Pitteri, M., 1998. Geometry and symmetry of multi-lattices. International Journal of Plasticity 14, 139–157.

Pitteri, M., Zanzotto, G., 1998. Beyond space groups: the arithmetic symmetry of deformable multi-lattices. Acta Crystallographica A54, 359–373.

Pitteri, M., Zanzotto, G., 2000. Continuum models for phase transitions and twinning in crystals. CRC/Chapman and Hall, London, submitted for publication.

Reed-Hill, R.E., Rogers, H.C., Hirth, J.P., 1964. Deformation Twinning. Gordon and Breach, New York.

Stakgold, I., 1950. The Cauchy relations in a molecular theory of elasticity. Quarterly Journal of Applied Mathematics 8, 169–186.

Thomas, L.A., Wooster, W.A., 1951. Piezocrescence-the growth of Dauphine twinning in quartz under stress. Proceedings of the Royal Society of London A 208, 43–62.

Zanzotto, G., 1988. Twinning in minerals and metals: remarks on the comparison of a thermoelastic theory with some available experimental results: Notes I and II. Atti Accademia Nazionale dei Lincei, Rend. Fis., 82, 725–742 and 743–756.

Zanzotto, G., 1992. On the material symmetry group of elastic crystals and the Born rule. Archive for Rational Mechanics and Analysis 121, 1–36.