The Bain Strain, Lattice Correspondences, and Deformations Related to Martensitic Transformations

J. S. BOWLES AND C. M. WAYMAN

The contribution of E. C. Bain in identifying the correspondence and pure strain involved in ferrous martensitic transformations is reviewed, together with later work which has confirmed Bain's hypothesis. Modern crystallographic theories of martensitic transformations have their origin in Bain's recognition that a pure strain by itself would produce a major accommodation problem between the parent and product phases; thus, additional lattice invariant processes have been introduced. Some aspects of the crystallographic theories are considered and attention is drawn to the increased number of factorizations of a total strain that become possible when sequences of three (or more) strains are considered instead of the usual two. Some implications of the recent double shear theory (and double interface mechanism) and plastic accommodation model are examined. The nucleation of martensitic transformations with respect to the Bain strain is considered briefly.

IN 1924 in a classic paper concerned with "The Nature of Martensite," E. C. Bain introduced the concept that a structural change of a certain kind might occur by means of a homogeneous deformation in which the product phase is derived from the parent by a simple "upsetting" process.[1] It seems to have been implicit in the thinking of the time that the martensitic reaction in steels is diffusionless, as evidenced by Bain's comment that "The outstanding characteristic, perhaps, of martensite is the complete allotropic change from γ to α iron and the completely suppressed precipitation of carbide." Bain also appreciated that the transformation could involve only small relative displacements of neighboring atoms: "It is reasonable, also, that the atoms themselves will rearrange with the greatest possible facility by a method that will require the least temporary distortion; we cannot expect an atom to squeeze through some narrow spacing, for instance. A mode of atomic shift was conceived by the author some months ago and was later discussed with Dr. Ancel St. John who, in the meantime, had thought of the same scheme of things and accepted it as highly probable. It is not surprising that this should have occurred, as it is the only easy method of constructing a body-centered atomic structure from a face-centered cubic crystal." Attention was then drawn to Figs. 1 and 2 (taken from the original in which the figure designations were 3 and 4), and the Bain strain, as we now know it, was discussed. It was pointed out that "There are three axes in the austenite grain on which this change may take place by the upsetting procedure just indicated," and suggested that these different variants must be operative in practice: "Unless this movement is compensated by encountering soon the opposite state of affairs along any given direction a rupture would ensue." This last comment seems to anticipate that ferrous martensite plates should be twinned, which is now known to be the case in a number of transformations.

In a recent communication[2] Bain, commenting on the program of this symposium, stated: "I would like to have a clearer impression of what the characteristics of 'Bain Strain' may be, than I can now envision. I fear it may conceivably relate to the 'upsetting' of the body-centered tetragonal lattice in any F.C.C. austenite to form martensite or ferrite, a rational enough end result, but one very difficult for me now to model, as actually occurring so simply and reasonably as I once imagined."

This question of simplicity vs reality will be considered in the present paper, in addition to other aspects of the Bain strain. Also to be discussed will be lattice correspondences, as implied by homogeneous (Bain) strains. Further, since the homogeneous deformation as specified by the Bain strain can be only a part of the total transformation crystallography for any given martensitic transformation, other required and related deformations such as the "inhomogeneous shear"[3] or "complementary strain"[4] are discussed. In particular, attention is drawn to the increased number of factorizations of the total strain that becomes possible, depending on the number of component strains and their order of occurrence. This, in turn, leads to an examination of recent multiple shear theories and attendant consequences such as a double interface mechanism for the propagation of martensite plates. Finally, some comments are made on the nucleation of martensite, *à propos* the original strain suggested by Bain.

THE BAIN STRAIN AND LATTICE CORRESPONDENCES

Figs. 1 and 2 indicate that for steels, martensite may be derived from austenite by a homogeneous distortion in which a ⟨100⟩ axis of austenite is contracted by about 20 pct to become the martensite *c* axis and two ⟨110⟩ axes in austenite are extended by some 12 pct to become martensite *a* axes. These strains are large by any measure. The "upsetting" process suggested by Bain also indicates that a correspondence between the fcc and bcc lattices exists. In general, a correspondence defines lattice points (and hence lattice vectors) in the parent that are transformed into specified lattice points (vectors) in the

---

J. S. BOWLES is Research Professor of Metallurgy, University of New South Wales, Kensington, Australia. C. M. WAYMAN is Professor, Department of Metallurgy and Mining Engineering, and Materials Research Laboratory, University of Illinois at Urbana-Champaign, Ill.

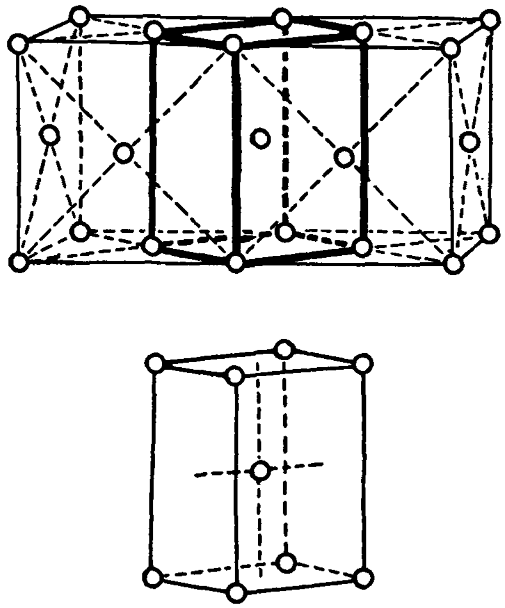

Fig. 1—Bct lattice delineated in austenite (fcc) structure.[1]

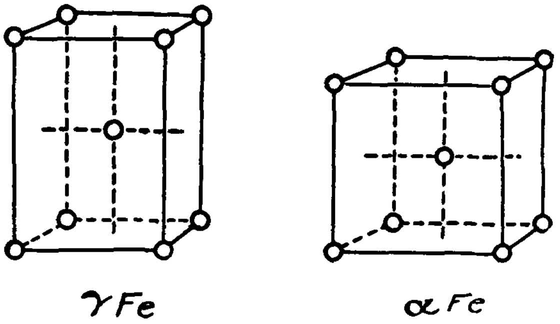

Fig. 2—Bct lattice of austenite and bcc lattice of α iron produced by upsetting the former; drawn to scale for comparison.[1]

martensite. In the transformation depicted in Figs. 1 and 2 the lattice points and atoms are identical since in each phase the structure is the lattice, but more complex correspondences are thought to apply in martensitic transformations in some materials (*e.g.*, uranium[3]). The general crystallographic theory of martensitic transformations deals with lattices and is thus independent of the nature of the atoms which occupy lattice sites.

There seems to be little doubt that the Bain correspondence implied by Bain's "upsetting" process is correct for steels. This correspondence can be inferred from known orientation relationships (*e.g.*, Greninger-Troiano, Kurdjumov-Sachs, Nishiyama) but these also served to indicate that the simple strain envisaged by Bain does not alone describe the whole transformation. This follows since the martensite unit cell axes are found to be rotated with respect to the ⟨001⟩ and ⟨110⟩ austenite axes from which the former are generated.* Thus, in a localized region of the product phase the transformation seems to be comprised of the Bain strain and a rotation, this combination being termed the total lattice strain $\mathbf{S}_T$ by Bowles and Mackenzie.[4] In equation form $\mathbf{S}_T = \mathbf{RB}$ where **R** and **B** are matrices which designate a rigid body rotation and the Bain strain respectively. It can be seen from Figs. 1 and 2 that the ⟨100⟩ and ⟨110⟩ austenite axes from which the three ⟨100⟩ martensite axes are generated in Bain's upsetting process are mutually orthogonal axes which remain orthogonal in the martensite lattice, *i.e.*, they are principal axes. Since every lattice correspondence implies a set of principal axes and a pure strain which extends these axes to their final lengths without rotation, it is clear then that the strain identified by Bain is the pure strain component of the martensite transformation in steel. The term "Bain Strain" has thus become a convenient description for the pure strain component of martensitic transformations in general, and of course figures prominently as one of the strain matrices with which the modern crystallographic theories are concerned.

The tetragonality of martensite in steels containing carbon or nitrogen is consistent with the correspondence implied by the Bain strain. Whereas in cubic ferrite the octahedral interstitial sites are occupied at random (*i.e.*, along each of the cube axes), in tetragonal martensite there must be a preferred occupancy of sites, only those octahedral sites along the martensite *c* axes being filled, so producing the observed tetragonality. Even though the carbon atoms in the austenite are distributed at random (along the midpoints of the ⟨100⟩ axes of the austenite unit cell) the carbon atoms in the martensite become ordered automatically as a consequence of the Bain strain. This follows since the fcc austenite may also be referred to a bct unit cell of axial ratio $\sqrt{2}$ and in this cell the carbon atoms are ordered. The interstitial atoms need only remain *in situ* as the Bain strain occurs reducing the axial ratio to a value much closer to unity.

An even more direct experimental verification of the Bain correspondence follows from work on an ordered Fe-Pt alloy containing 25 at. pct Pt.[5] For this material the fcc matrix orders into a Cu_3Au structure during annealing near 650°C. Transformation of the ordered austenite into martensite then occurs during subsequent cooling below room temperature. Using transmission electron microscopy and diffraction patterns, Tadaki and Shimizu[5] showed that martensite (001) and $(\bar{1}\bar{1}1)$ reciprocal lattice sections (both containing superlattice reflections) were in the configurations which would be obtained had the martensite formed by the Bain distortion of the ordered parent phase.

Further verification of the Bain strain was provided recently by studies of the transformation behavior of thin films of iron vapor deposited on the {001} surfaces of copper[6] and nickel[7] films. Up to certain film thicknesses iron deposited on copper and nickel at relatively low temperatures is (pseudomorphically) fcc and the overgrowth is completely epitaxial, the cube planes and directions in Fe/Cu or Fe/Ni being parallel. In both cases continuous iron films are formed in the early stages of growth and are strained to match the substrate lattice. There are important differences, however; iron has a smaller lattice parameter than copper, but a larger one than nickel, so the iron deposits on copper are in tension while those on nickel are in compression.

---

*Taking the Kurdjumov-Sachs relationship, $(111)_A \parallel (011)_M$ and $[\bar{1}01]_A \parallel [\bar{1}\bar{1}1]_M$ as representative, it is easily shown that the martensite axes are rotated about 10 deg about a direction in the austenite which makes an angle of ~70 deg with the ⟨001⟩ contraction axis.

As the iron films increase in thickness, "nuclei" of bcc iron form, presumably martensitically. It is quite clear that the iron-substrate misfit stresses interact with the Bain strain and accordingly determine the orientation of the transformed bcc with respect to its fcc iron matrix. In partially transformed films the parent-product orientation relationship has been measured and the relationships are close to the Pitsch, Kurdjumov-Sachs, Nishiyama, and Bain relations in cases studied, although as might be expected[8] the orientation relationship in thin films is not exactly that which obtains between austenite and martensite in the bulk. The important point, however, is that in each case[7] the orientation relationship is consistent with its being prescribed by the iron-substrate misfit stresses. For iron films in tension on copper the orientations of the transformed regions are such that the expansion axes of the Bain strain lie in the plane of the film, whereas for iron films (in compression) on nickel, variants of the contraction axis of the Bain deformation lie in the film plane. In either situation it is as though film-substrate interaction strains "assist" the principal strains of the Bain deformation. These examples involving particularly simple and well-chosen geometry constitute powerful evidence that the Bain strain in iron under stress occurs.

For a given transformation between two lattices an infinite number of possibilities exist for the correspondence. An example of an alternative correspondence[9] for the fcc to bcc transformation is shown in Fig. 3. In this case, however, the principal strains are larger than those implied by the Bain correspondence and the carbon atoms would not be shifted from their positions in austenite to their preferential *c* axis alignments in bct martensite. To move the carbon atoms to their correct final positions "shuffle" motions (of carbon atoms) would be needed in addition to the homogeneous strain implied by the correspondence of Fig. 3.

When superlattices or preferred arrangements of interstitial atoms are not present to aid the identification of the correct correspondence it must be deduced by other means. The first suggestion that was made[10] for determining the correct correspondence from the infinite number of possibilities that exist for any particular transformation was that the logical correspondence would be the one involving the smallest atomic displacements. Such a selection procedure requires a knowledge of the orientation relationship and is therefore restricted in its application to those cases for which this relationship is known. A later proposal by Lomer[3] avoids this difficulty. He suggested that the correct correspondence is always the one for which the associated principal strains are smallest. This criterion requires no knowledge of the orientation relationship. For the fcc-bcc Bain correspondence it is well-known that the associated principal strains are smaller than for other possible correspondences.

Even though the principal strains from the Bain correspondence in steels and iron alloys are small compared to those associated with alternative fcc-bcc correspondences, they are large compared with those involved in many nonferrous transformations. For

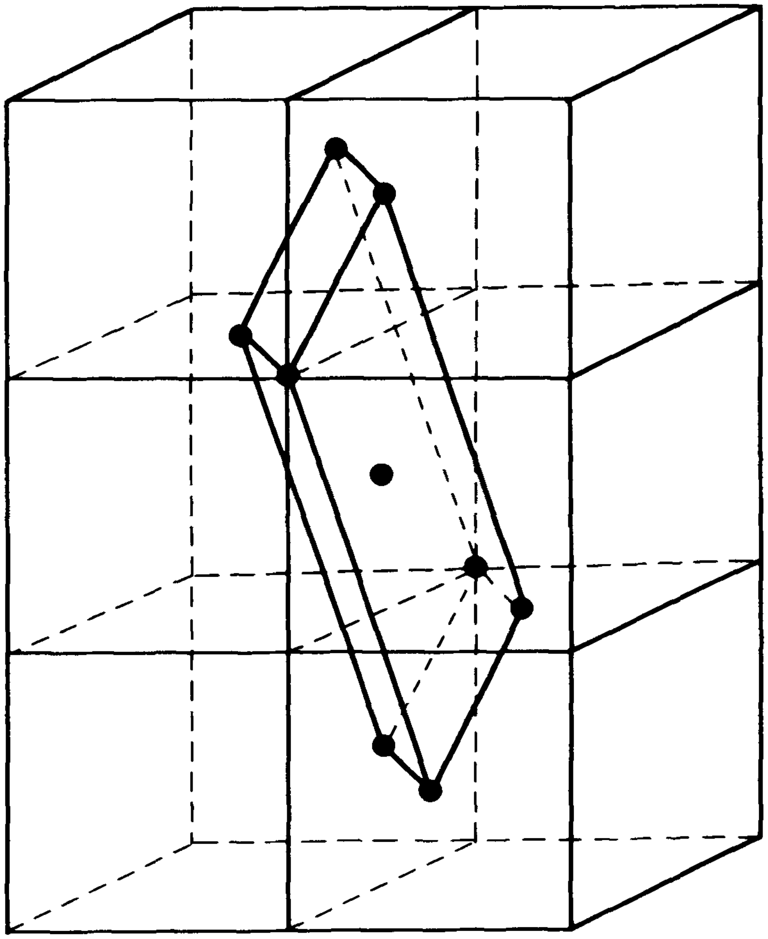

Fig. 3—Alternative fcc-bcc lattice correspondence involving larger principal strains than required by the Bain correspondence. This correspondence does not homogeneously convey interstitial atoms (carbon) to selected positions along the martensite tetragonal axis and thus shuffle motions would be required.[9]

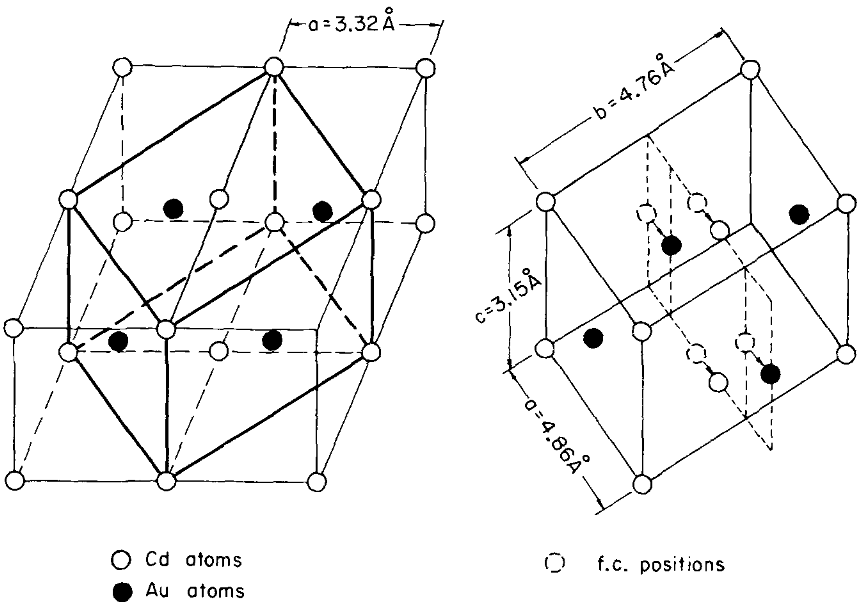

Fig. 4—Cubic-orthorhombic lattic correspondence for the β to β' martensitic transformation in Au-Cd alloys showing both "direct" and "shuffle" atomic movements.

example, in the fcc-bct transformation in In-Tl alloys[11] the strains vary from 1 to 2 pct and are thus an order of magnitude smaller than for steels. As would be expected from theoretical considerations, the observed width of the transformation twins is inversely proportional to the relative magnitude of the principal strains, being as small as 50Å in steels.

In some martensitic transformations all the atomic movements are not prescribed by the homogeneous Bain strain implied by the lattice correspondence. A correspondence must of course relate unit cells containing equal numbers of atoms in the initial and final structures. If the transformation is between structures which both have only one atom in each primitive unit cell, the Bain strain will describe the movement of every atom. If, however, the primitive unit cells of the two structures contain different numbers of atoms, then the correspondence will relate non-primitive cells. In this case, the Bain strain will describe the movement of only a fraction of the atoms and the remainder will undergo shuffling movements. This situation has been considered in detail by Christian,[12,13] who proposed that the definitive feature of a martensitic transformation is the existence of a lattice correspondence, with no diffusive mixing of atoms other than the shuffles needed to complete the structural change.

An example of this behavior is found in the well-known nonferrous martensitic transformation occurring in Au-Cd alloys containing about 47.5 at. pct Cd.[14] The lattice correspondence believed applicable to this transformation is shown in Fig. 4, in which case an fct unit cell is delineated within four CsCl unit cells of the parent phase. The martensitic product is base-centered orthorhombic with cadmium atoms taking the positions $[0,0,0]$ and $[0,\frac{5}{8},\frac{1}{2}]$ and gold atoms the positions $[\frac{1}{2},\frac{1}{2},0]$ and $[\frac{1}{2},\frac{1}{8},\frac{1}{2}]$.[15] If the structure were face-centered orthorhombic, the previous coordinates would be $[0,0,0]$ and $[0,\frac{1}{2},\frac{1}{2}]$, and $[\frac{1}{2},\frac{1}{2},0]$ and $[\frac{1}{2}, 0,\frac{1}{2}]$ respectively. The correspondence in Fig. 4 involves rather small principal distortions[14] (5, 1.4, and 3.5 pct along the orthorhombic *a*, *b*, and *c* axes respectively) but it is to be noted that half the atoms in the martensite unit cell are displaced from face-centered positions. It is implied therefore that in the course of transformation half of the atoms undergo homogeneous movements, as indicated by the lattice correspondence, whereas the other half, in addition to this, undergo shuffles. This situation is not extraordinary in the sense that both homogeneous and shuffling movements are also involved in mechanical twinning. For example, in the case of (301) $[\bar{1}03]$ twinning in tetragonal tin, only half[16] the atoms undergo the strict twinning displacements; the other half undergo shuffle motions to satisfy the requirement that the deformation be lattice invariant.

In even more complex situations such as the β to α transformation in uranium, if the correspondence is deduced from the minimum principal strains criterion,[3] it turns out that the majority of atoms must undergo shuffling motions. Lomer examined some 1600 possible correspondences for the uranium case and concluded that only a few of these resulted in reasonably small principal strains, and even then shuffles were needed for the majority of the atoms.* Although the principal axes of the Bain strain are usually rational, in the case of uranium, the minimum principal strains criterion leads to axes that are not rational. Procedures for determining the principal axes and strains are well-established.[3,4,17]

FURTHER ASPECTS OF THE BAIN STRAIN

As mentioned earlier, Bain recognized that the occurrence of his proposed atomic movements as a homogeneous strain would involve a major accommodation problem and that the misfit between the martensite and the surrounding austenite would be reduced if different variants of the Bain Strain occurred on a sufficiently fine scale. The phenomenological theories of martensite,[4,18] which were developed about 20 years after the publication of Bain's paper, were derived from a detailed consideration of this accommodation problem. These theories have their origin in the conclusion that the homogeneous change of shape accompanying the formation of a martensite plate must be,

---

*In the "best" correspondence only one atom in thirty is placed in its correct position due to the homogeneous distortion.

at least approximately, a strain in which the interface between the martensite and the parent phase is undistorted and unrotated, *i.e.*, the shape change is an invariant plane strain. Since the Bain strain is not an invariant plane strain, the transformation cannot be described as a simple homogeneous distortion of the parent lattice. It is necessary to propose, therefore, that in addition to the pure (Bain) strain, other processes that do not produce any further distortion of the martensite lattice also accompany the transformation and that the resultant of all these strains is an invariant plane strain. The additional processes that have been considered in these theories are slip, twinning, faulting, and the rigid body rotation, **R**, mentioned earlier.

The phenomenological theories are concerned with the relationships that exist between the Bain strain, the shape strain, the rotation, and the shear processes that comprise the so-called inhomogeneous[18] or complementary[4] strain. Some detailed aspects of these theories will be considered later, but it will be sufficient at this stage to indicate the differences in the approaches adopted in the original theories of Wechsler, Lieberman, and Read (WLR), and Bowles and Mackenzie (BM). In the WLR theory the shape strain was derived as the resultant of **R**, **B** and an inhomogeneous shear. In the BM theory[4] the total strain $\mathbf{S}_T = \mathbf{RB}$, which describes the homogeneous lattice strain occurring over a range of a few unit cells, was derived as the resultant of the shape strain and a complementary shear. The two approaches are equivalent.

It was originally proposed by Bowles-Mackenzie that the complementary strain is part of a twinning shear in the martensite.* This hypothesis implies the possibility that oppositely directed complementary shears could produce twin orientations within a martensite plate, and it is now well known that the plates produced in many martensitic transformations are twinned on a very fine scale. The existence of twins in a martensite plate means that there are two martensite orientations and two total lattice strains, $\mathbf{S}_T$. The formation of these twin martensite orientations involves two different Bain strains but, provided that the lattice correspondences for the twins are crystallographically equivalent, it can be shown that the two Bain strains are variants of one another and that the twinning plane is generated from a plane of symmetry in the parent lattice. This result was demonstrated first by Bowles and Mackenzie for Type I twinning but has recently been shown by Ross and Crocker[19] to hold also for Type II twinning. In ferrous martensites containing twins, the twinning plane is $\{112\}_M$, which is generated from $\{110\}_A$, and it is clear from Fig. (1) that the plane $(\bar{1}01)_A$ is symmetrically disposed with respect to the $[100]_A$ and $[001]_A$ directions. These axes become the martensite *c* axes in the two twin orientations in the martensite. Thus, the modern theories confirm Bain's expectation that the contraction of a <100> austenite axis to generate a martensite *c* axis could not continue to occur over a very great distance. The formation of twinned martensite plates in which different variants of the pure (Bain) strain operate alternately on a fine scale to reduce the misfit between the parent phase and the martensite and allow the shape strain to be an invariant plane strain, is a feature of many martensitic transformations.* The twinning plane is apparently always generated from a symmetry plane of the parent lattice but in some cases the twinning system is not that observed during plastic deformation. Examples of this behavior are to be found in a number of bcc to hcp martensitic transformations which produce martensite plates twinned on the $\{10\bar{1}1\}$ plane.[20]

FORMULATIONS OF THE THEORY ON MARTENSITE CRYSTALLOGRAPHY

The notations and focal points originally taken by BM and WLR are different, but algebraically the central equations are identical. In each case the shape strain (surface tilt) is macroscopically homogeneous and an additional inhomogeneous deformation is required. Formally, the complementary strain **C** employed by BM is the matrix inverse of the inhomogeneous shear used by WLR. While certain insights are to be gained by the WLR approach in which the Bain strain is taken as an entity and in effect it is considered what other lattice invariant processes are required to allow a known habit plane to be macroscopically invariant, following the BM type of analysis may form a more convenient basis of departure for the analysis of strain sequences and alternative theoretical factorizations.

In the original BM theory, the invariant line strain, **S**, derived from the total strain $\mathbf{S}_T$ by removing a suitable dilatation, was factorized into two invariant plane strains, a shape strain, $\mathbf{Sh} = \mathbf{I} + m_1\mathbf{d}_1\mathbf{p}_1'$, and a complementary shear $\mathbf{C} = \mathbf{I} + m_2\mathbf{d}_2\mathbf{p}_2'$.* The elements of the component strains refer to their positions before the strain **S** and since **Sh** was assumed to occur first the factorization was written

$$\mathbf{S} = (\mathbf{I} + m_2\mathbf{S}\mathbf{d}_2\mathbf{p}_2'\mathbf{S}^{-1})(\mathbf{I} + m_1\mathbf{d}_1\mathbf{p}_1') = (\mathbf{SCS}^{-1})\,\mathbf{Sh} \qquad [1]$$

where $(\mathbf{SCS}^{-1})$ is an invariant plane strain on the plane $\mathbf{p}_2'\mathbf{S}^{-1}$ and in the direction $\mathbf{Sd}_2$, to which $\mathbf{p}_2'$ and $\mathbf{d}_2$ are transformed in the final lattice.

It was noted by Bowles and Mackenzie that an alternative factorization of **S** is

$$\mathbf{S} = \mathbf{ShC} \qquad [2]$$

and this order of factors has been used in recent papers as a more convenient representation than that given by [1]. This factorization describes a mathematical sequence of strains in which the complementary strain **C** occurs first as an invariant plane strain on the elements $\mathbf{p}_2'$ and $\mathbf{d}_2$ in the initial lattice, followed by

---

*Without going into detail here, it can be mentioned that an internally twinned martensite is mathematically equivalent to one which is internally slipped. Furthermore, the inhomogeneous shear aspect of the total transformation can as well be referred to the parent phase. That is, algebraically a "predeformed" austenite followed by the Bain deformation is equivalent to a martensite formed by the Bain mechanism followed by an inhomogeneous shear. Both situations provide for a macroscopically undistorted plane of contact, or habit plane.

*When the martensite is not internally twinned (*e.g.*, "lath" martensite in low-carbon steels) a single correspondence relation and Bain deformation applies to an entire plate. In general, the reason(s) for a particular operative substructure (inhomogeneous shear) in a given material are still rather obscure.

*The complementary shear is a lattice strain but does not contribute to the macroscopic shape change because its effects are annulled by an equal and opposite lattice invariant deformation. In the various equations m, **d** and **p**$'$ respectively refer to the magnitude of the strain, the direction of displacement, and the normal to the plane which is invariant; **I** denotes the unit (3×3) matrix.

the shape strain $\mathbf{Sh}$. When $\mathbf{Sh}$ occurs second it is, of course, no longer the same lattice strain, for it is occurring on the plane and direction that have been displaced to $\mathbf{p}_1'$ and $\mathbf{d}_1$ in the final lattice; the macroscopic shape change it describes is, however, the same as before because the macroscopic elements of the strain have not changed.

In the WLR analysis[18] the approach was somewhat different and the shape strain $\mathbf{Sh} = \mathbf{I} + m_1\mathbf{d}_1\mathbf{p}_1$ was regarded as the resultant of the lattice strain $\mathbf{S}$ and a lattice invariant shear, $\mathbf{C}^{-1} = \mathbf{I} - m_2\mathbf{d}_2\mathbf{p}_2'$ so that

$$\mathbf{Sh} = \mathbf{SC}^{-1}.$$

This relationship is clearly the same as [2] and the alternative representation,

$$\mathbf{Sh} = (\mathbf{SC}^{-1}\mathbf{S}^{-1})\ (\mathbf{S})$$

is equivalent to [1]. In this case, since both $(\mathbf{C}^{-1})$ and $(\mathbf{SC}^{-1}\mathbf{S}^{-1})$ are lattice invariant, $\mathbf{S}$ represents the same lattice strain whether it occurs first or second. The strain $(\mathbf{SC}^{-1}\mathbf{S}^{-1})$ has rational elements in the final lattice if the elements of $\mathbf{C}^{-1}$ are rational in the initial lattice.

An alternative expression for the factorization described by Eq. [1] can be obtained by substituting $\mathbf{S} = \mathbf{ShC}$ to give*

$$\begin{aligned}\mathbf{S} &= (\mathbf{ShCSh}^{-1})\ (\mathbf{Sh})\\ &= (\mathbf{I} + m_2\mathbf{Shd}_2\mathbf{p}_2'\mathbf{Sh}^{-1})\ (\mathbf{I} + m_1\mathbf{d}_1\mathbf{p}_1')\end{aligned} \qquad [3]$$

In this form the strain $(\mathbf{ShCSh}^{-1})$ is clearly an invariant plane strain on the plane and direction to which $\mathbf{d}_2$ and $\mathbf{p}_2'$ are displaced by the shape strain, $\mathbf{Sh}$. The distinction between this description of the "second" strain and that given above is trivial for the case of factorization into two strains, but it assumes more significance when the factorization of $\mathbf{S}$ into three strains is considered. The need to consider factorization of the total lattice strain into three strains has arisen from the failure of the original theories to explain the full crystallography of certain martensitic transformations, in terms of two component strains. The dilatation of about $1\frac{1}{2}$ pct which was originally suggested[4] to account for (225) ferrous martensite crystallography has now been shown by direct measurement to be absent[22] and evidence that there may be more than one complementary shear occurring in some transformations is accumulating.[23-25] Thus recently attempts have been made to explain martensite crystallography in terms of a complementary strain that is not a simple shear. Theories have been developed independently by Acton and Bevis[26] and Ross and Crocker[27] for the case where the complementary strain consists of two shears, and Bowles and Dunne[28] have described a model of the (225) transformation in which the complementary strain is even more complex.

Table I. Equivalent Representation of $S_T = ShS_2S_3$

BM Expressions	WLR Expressions
$S_T = ShS_2S_3$	$Sh = S_TS_3^{-1}S_2^{-1}$
$S_T = Sh(S_2S_3S_2^{-1})S_2$	$Sh = S_T(S_3^{-1}S_2^{-1}S_3)S_3^{-1}$
$S_T = (ShS_2S_3S_2^{-1}Sh^{-1})ShS_2$	$Sh = (S_TS_3^{-1}S_2^{-1}S_3S_T^{-1})S_TS_3^{-1}$
$S_T = (ShS_2S_3S_2^{-1}Sh^{-1})(ShS_2Sh^{-1})Sh$	$Sh = (S_TS_3^{-1}S_2^{-1}S_3S_T^{-1})(S_TS_3^{-1}S_T^{-1})S_T$
$S_T = (ShS_2Sh^{-1})ShS_3$	$Sh = (S_TS_3^{-1}S_T^{-1})S_TS_2^{-1}$
$S_T = (ShS_2Sh^{-1})(ShS_3Sh^{-1})Sh$	$Sh = (S_TS_3^{-1}S_T^{-1})(S_TS_2^{-1}S_T^{-1})S_T$

Note that still more sequences can be derived, *e.g.*, $Sh = S_TS_2^{-1}(S_2S_3S_2^{-1})$.[21]

It is easily shown that in factorizing a total lattice strain $\mathbf{S}_T$ into a shape strain $\mathbf{Sh}$ and two other strains $\mathbf{S}_2$ and $\mathbf{S}_3$, expressions analogous to Eqs. [1] and [3] can be written. Thus,

$$\mathbf{S}_T = \mathbf{ShS}_2\mathbf{S}_3 \qquad [4]$$

and

$$\mathbf{S}_T = (\mathbf{ShS}_2\mathbf{S}_3\mathbf{S}_2^{-1}\ \mathbf{Sh}^{-1})\ (\mathbf{ShS}_2\mathbf{Sh}^{-1})\ (\mathbf{Sh}) \qquad [5]$$

If $\mathbf{S}_2$ and $\mathbf{S}_3$ are also invariant plane strains [5] becomes

$$\begin{aligned}\mathbf{S}_T = {}&(\mathbf{I} + m_3\mathbf{ShS}_2\mathbf{d}_3\mathbf{p}_3'\mathbf{S}_2^{-1}\ \mathbf{Sh}^{-1})\ (\mathbf{I} + m_2\mathbf{Shd}_2\mathbf{p}_2'\mathbf{Sh}^{-1})\\ &(\mathbf{I} + m_1\mathbf{d}_1\mathbf{p}_1')\end{aligned}$$

where $\mathbf{d}_2$ and $\mathbf{p}_2'$ and $\mathbf{d}_3$ and $\mathbf{p}_3'$ are the elements of the strains $\mathbf{S}_2$ and $\mathbf{S}_3$ in the initial lattice. In addition to these, however, many other equivalent phenomenological descriptions become possible when three component strains are considered. Other factorizations of $\mathbf{S}_T$ into sequences of lattice strains, in accordance with the BM and WLR approaches, are shown in Table I. It is evident that the strains $\mathbf{S}_2$ and $\mathbf{S}_3$, provided they are "modified" suitably, can be considered to occur in any order both before and after $\mathbf{Sh}$ to produce the same $\mathbf{S}_T$. If $\mathbf{S}_2$ and $\mathbf{S}_3$ are simple shears, then the "modified" versions are shears on the planes and directions to which the shear plane and direction have been displaced by preceding strains. It should be noted, however, that a simple interchange in the order of $\mathbf{S}_2$ and $\mathbf{S}_3$ is not permissible.*

If the WLR approach is used and the shape strain $\mathbf{Sh}$ is described as the resultant of $\mathbf{S}_T$ and two lattice invariant strains $\mathbf{S}_2^{-1}$ and $\mathbf{S}_3^{-1}$, then the second set of expressions shown in Table I can be derived. These two sets of expressions reveal the relationship to one another of the various current theories involving three strains. In the Ross-Crocker and Acton-Bevis theories the strain $\mathbf{S}_3^{-1}$ is considered to be part of the twinning shear and $\mathbf{S}_2^{-1}$ is taken to be a supplementary shear, $\mathbf{Su}$, for which various specific systems are suggested. These theories are presented in terms of the first expression in the right hand column of Table I, but any of the other expressions could have been used. The plastic accommodation model of Bowles and Dunne[28] is derived from a consideration of an "accommodation" strain, $\mathbf{P}$, occurring in the austenite at the edge of an advancing (225) plate. As they have pointed out, however, the model is equivalent to a hypothesis that the complementary strain is composed of the twinning shear* followed by the inverse of the accommodation strain $\mathbf{P}$, *i.e.*, $\mathbf{S}_3$ is considered to be part of the twinning shear and $\mathbf{S}_2 = \mathbf{P}^{-1}$. Thus the main difference between the current theories lies in the hypotheses

---

*These expressions are quite general and do not depend on the component strains being invariant plane strains. Any strain $\mathbf{S} = \mathbf{S}_1\mathbf{S}_2$, where $\mathbf{S}_1$ and $\mathbf{S}_2$ are not necessarily invariant plane strains can be written as

$$S = (SS_2S^{-1})\ (S_1) = (S_1S_2S_1^{-1})(S_1)$$

and the factorization given recently by Dunne and Wayman[21] can be written as

$$S = (S_2)\ (S^{-1}S_1S) = (S_2)\ (S_2^{-1}S_1S_2).$$

*This is only possible if $(\mathbf{p}_2'\mathbf{d}_3)\mathbf{d}_2\mathbf{p}_3' = (\mathbf{p}_3'\mathbf{d}_2)\mathbf{d}_3\mathbf{p}_2'$.

*Bowles and Dunne also consider the possibility that S_3 is a combination of a shear on $(111)_F$ with the twinning shears.

made about $\mathbf{S}_2$. Ross and Crocker and Acton and Bevis propose that $\mathbf{S}_2$ is a simple shear, while Bowles and Dunne propose that it is an invariant line strain resulting from multiple slip.

Further accurate experimental measurements are needed to test the current theories, but on the evidence at present available it seems that neither the double shear theories nor the plastic accommodation models are tenable in their present form. Dunne and Wayman[21] have made a detailed assessment of the double shear theories and they conclude that although a shear in the twinning direction is a major component of the complementary strain, the remaining part is neither a simple rational shear as required by the double shear theories nor is it the result of multiple rational shears as required by the plastic accommodation model.

MECHANISMS AND PHYSICAL IMPLICATIONS

Although the original crystallographic theories were phenomenological, in later modifications there has been an increasing tendency to attribute real physical significance to the component strains. The first and most significant development in this direction was the application by Bullough and Bilby[29] of the concept of surface dislocations to the martensite interface. In the BB theory, the boundary between the parent phase and the martensite is considered to consist of an array of dislocations which accommodates the mismatch between the two lattices. When this interface moves normal to itself, generating the martensite lattice, the motion of the dislocations produces the lattice invariant (complementary) strain, and the macroscopic shape change produced is an invariant plane strain. Such an interface is glissile and can move without changing its structure provided that the lattice invariant (complementary) strain is a simple shear.

The surface dislocation concept provides an elegant mechanism for growth of martensite plates normal to the habit plane but does not reveal how the habit plane interface is created and, of course, is not applicable when the complementary strain involves two unrelated shears as in the double shear theories.[26,27] In this case, the interface would need to contain two sets of dislocations and ordinarily would not be glissile.

To avoid this difficulty, Ross and Crocker proposed that the boundary consists of two closely-spaced parallel interfaces, each containing one set of dislocations and each producing one of the lattice invariant shears. They drew attention to the fact that such a double interface description implies a real physical significance for the shears in their analysis and for the order in which they occur. Ross and Crocker believed that the region between the two interfaces would have the structure of a single crystal of martensite and pointed out that this agreed with their conclusion that the two most satisfactory mechanisms they had found both involved the supplementary shear, $\mathbf{Su}$, preceding the twinning shear. They thus interpreted the transformation twins as forming in this martensite between the two interfaces. This mechanism, however, implies that the untwinned martensite lattice can be generated by the passage of the first interface, which contains only the single set of dislocations that produce the supplementary shear. In other words, it is being implied that $\mathbf{S}_T = \mathbf{ShSu}$, which is not correct; the reason for the development of the double shear theories was the failure to discover any simple shear, $\mathbf{Su}$, which, combined with an invariant plane strain on the (225) habit plane, would generate the martensite lattice in its observed orientation relative to austenite.

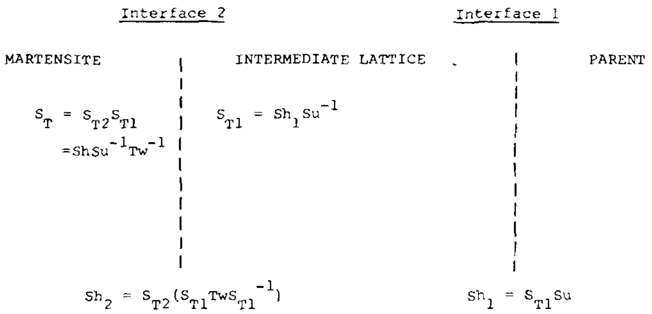

Fig. 5—Sequence of strains implied by a double interface mechanism when $\mathbf{Sh} = \mathbf{S}_T\mathbf{Tw}\mathbf{Su}$. Interfaces 1 and 2 produce the shape changes $\mathbf{Sh}_1$ and $\mathbf{Sh}_2$ respectively. The total lattice strains which generate the intermediate lattice from the parent and the martensite from the intermediate lattices are $\mathbf{S}_{T1}$ and $\mathbf{S}_{T2}$ respectively.

It is clear, therefore, that if a double interface mechanism does occur, the first interface cannot generate the final martensite lattice and probably not even the total shape change. The sequence of events required by a double interface mechanism in which a supplementary shear occurs at the first interface is shown in Fig. 5.

This sequence amounts to not one transformation, but two consecutive transformations, each occurring by a surface dislocation mechanism. Such a mechanism is possible only if both transformations have the same habit plane, and so far it has not been proved that the component shape strains $\mathbf{Sh}_1$ and $\mathbf{Sh}_2$ can have the same habit plane.

An interesting possibility that arises from Fig. 5 is $\mathbf{Sh}_2 = \mathbf{I}$. This could happen if Interface 1 generated an intermediate structure that could be transformed to martensite by a shear on the martensite twinning system. The second interface would then simply be the boundary between the finally twinned martensite and the intermediate lattice, and this interface would have no associated shape change.

The acceptability of the double interface mechanism depends finally, of course, on whether or not it can be established that two independent lattice invariant shears are involved in the transformation. As mentioned earlier, Dunne and Wayman's results indicate that this does not seem to be possible, and if such is the case, then it will be necessary to examine other kinds of mechanisms. A possible modification of the surface dislocation mechanism is suggested by the ideas developed by Bowles and Dunne in the derivation of the plastic accommodation model. They considered the propagation of a (225) martensite plate in the direction of the close-packed line lying in the habit plane and concluded that the decrease in interatomic distance along this line would need to be accommodated by deformation of the austenite in front of the growing plate. They pointed out that for this plastic accommodation strain to be a homogeneous lattice invariant strain plastic deformations associated with each increment of growth of the plate must be localized at the tip (or edge) of the growing plate.

Such a plastic strain differs from the lattice invariant strain produced by a surface dislocation in that the dislocations producing it must move away from the edge of the plate into the austenite at an angle to the habit plane. This suggests that the moving interface (which might not necessarily be parallel to the habit plane) may be able to maintain a constant form only by periodically generating dislocations which move away into the austenite.

If the movement of the interface involves the generation of dislocations which move away into the parent phase, these dislocations would have to be normal stable dislocations of the parent phase. The inhomogeneous strains that are generated by the movement of an interface normal to itself without changing its structure are not restricted in this way however, and they need not necessarily be processes that normally occur during plastic deformation. An interesting example of this is provided by the bcc to hcp transformation in titanium and its alloys. These transformations produce plates that are twinned on a $\{10\bar{1}1\}$ plane that is generated from a $\{110\}$ symmetry plane in the parent phase so that both orientations have equivalent correspondences.[20] This mode of twinning is not observed when bcc titanium is deformed plastically. The extension of a twinned product as the interface advances into the parent phase without changing its structure does not involve the occurrence of a twinning shear in the final lattice. In a double interface mechanism it may involve the shearing of a transition product in opposite directions to form twins but this in no way implies that the shearing process should be a favored twinning shear in the final lattice. This distinction between transformation twinning and deformation twinning suggests that in assessing which "deformation modes" are to be regarded as "reasonable" or "physically realizable" inhomogeneous strains for a martensitic transformation, conventional criteria do not necessarily apply.*

MARTENSITE NUCLEATION

Although the phenomenological theory of martensitic transformations deals with initial and final states and does not pretend to indicate actual atomic paths, the simplicity and probable utility of the Bain strain cannot be denied. Various investigators have suggested "physical" means according to which the austenite-martensite structural change might occur. Kurdjumov and Sachs,[30] for example, employed two consecutive shears to generate martensite, the first consisting of a $(1\bar{1}2)[\bar{1}11]_A$ shear followed by another on $(1\bar{1}2)[\bar{1}11]_M$.* Although this early mechanism is not consistent with all of the observed crystallographic features of the transformation, their work was the first example of using shears, a physically realistic entity, to effect a structural change. Zener[31] later pointed out that a half-faulted fcc lattice is in a near-bcc configuration and suggested that elastic instabilities may be responsible for the particular atomic movements in martensitic transformations. A number of subsequent "shear" models for structural changes have appeared in the literature. Kelly and Nutting[32] reactivated the Kurdjumov-Sachs mechanism, with some modification, in an attempt to account for the formation of untwinned lath martensite in stainless and low carbon steels.

Shear models have also led to stacking fault nucleation hypotheses.[33] A major difficulty with this approach is the problem of determining whether a fault caused the event or was simply a consequence, *i.e.*, the faults observed in association with plates of martensite may have arisen because the austenite is deformed by the growing martensite.[34]

It is commonly accepted that structural imperfections such as dislocations may play a role in the nucleation of martensite. For example, the strained region in the vicinity of a dislocation in the austenite may find itself in a near martensitic (*i.e.*, bcc) configuration.[35] Classical nucleation theory simply does not apply to the nucleation of martensitic transformations. There are a number of difficulties in applying the "critical particle size" and activation barrier approach; not the least is the simple fact that in a number of systems martensitic transformations occur at very low temperatures where thermal activation cannot be appreciable. One well-known nucleation theory[36] holds that preformed embryos are frozen-in and become critical at low temperatures, but there are difficulties with this approach.

It has been recognized that kinematic paths for phase transitions may be related to lattice vibration modes.[37] A different approach to the martensite nucleation dilemma has been suggested by Clapp,[38] who envisions that standing phonon waves cause atomic displacements. According to the model, planar defects such as free surfaces and grain boundaries enhance the formation of standing waves. In essence the phonon model provides for coordinated atomic motions out of the equilibrium positions and this is a means for providing the collective movements implied by the original Bain model. A major weakness of the phonon model is that the extent of collective atomic excursions is on the order of a few percent, and there are presently no provisions for strains as high as 20 pct, as demanded by the Bain strain for steels. Nevertheless, classical nucleation models failing and frozen-in embryos subject to criticism, the phonon nucleation approach to phase transformations may be quite promising, and at the least is a fresh departure from tradition.

ACKNOWLEDGMENTS

One of us (C.M.W.) is grateful for the support of the U.S. Army Research Office (Durham) and the U.S. Atomic Energy Commission. The sponsorship by these agencies has led to a number of findings and conclusions, some of which have been expressed in this paper.

REFERENCES

1. E. C. Bain: *Trans. AIME*, 1924, vol. 70, p. 25.
2. E. C. Bain: private communication.
3. W. M. Lomer: *Inst. Metals Monogr. No.* 18, London, 1955, p. 243.
4. J. S. Bowles and J. K. Mackenzie: *Acta Met.*, 1954, vol. 2, pp. 129, 138, 224.
5. T. Tadaki and K. Shimizu: *Trans. Japan Inst. Metals*, 1970, vol. 11, p. 44.
6. J. W. Matthews and W. A. Jesser: *Phil. Mag.*, 1969, vol. 20, p. 999.
7. G. H. Olsen and W. A. Jesser: *Acta Met.*, 1971, vol. 19, p. 1299.
8. W. Pitsch: *Phil. Mag.*, 1959, vol. 4, p. 577.

---

*$\{10\bar{1}2\}$, $\{11\bar{2}1\}$, $\{11\bar{2}2\}$, and $\{11\bar{2}4\}$ twin modes have been reported, the first two of these being predominant.[20]

*These shears do not produce the correct structure and some atomic readjustments are required.

9. C. M. Wayman: *Adv. Mater. Res.*, 1968, vol. 3, p. 147, ed. H. Herman, John Wiley and Sons Inc., New York.
10. M. A. Jaswon and J. A. Wheeler: *Acta Cryst.*, 1948, vol. 1, p. 216.
11. M. W. Burkhart and T. A. Read: *Trans. AIME*, 1953, vol. 197, p. 1516.
12. J. W. Christian: *Iron Steel Inst.*, London, 1965, Spec. Rep. 93, p. 1.
13. J. W. Christian: *Inst. Metals Monogr. No.* 33, London, 1969, p. 129.
14. D. S. Lieberman, M. S. Wechsler, and T. A. Read: *J. Appl. Phys.*, 1955, vol. 26, p. 473.
15. A. Ölander: *Z. Kristal*, 1932, vol. 83, p. 145.
16. E. O. Hall: *Twinning and Diffusionless Transformations in Metals*, Butterworths, London, 1954.
17. C. M. Wayman: *Introduction to the Crystallography of Martensitic Transformations*, The Macmillan Company, New York, 1964.
18. M. S. Wechsler, D. S. Lieberman, and T. A. Read: *Trans. AIME*, 1953, vol. 197, p. 1503.
19. N. H. D. Ross and A. G. Crocker: *Scripta Met.*, 1969, vol. 3, p. 37.
20. Z. Nishiyama, M. Oka, and H. Nakagawa: *Trans. Japan Inst. Metals*, 1966, vol. 7, p. 174.
21. D. P. Dunne and C. M. Wayman: *Acta Met.*, 1971, vol. 19, p. 425.
22. P. Krauklis and J. S. Bowles: *Acta Met.*, 1969, vol. 17, p. 997.
23. N. F. Kennon and J. S. Bowles: *Acta Met.*, 1969, vol. 17, p. 373.
24. K. Shimizu, M. Oka, and C. M. Wayman: *Acta Met.*, 1971, vol. 19, p. 1.
25. S. Jana and C. M. Wayman: *Met. Trans.*, 1970, vol. 1, p. 2825.
26. A. F. Acton and M. Bevis: *Mater. Sci. Eng.*, 1969-70, vol. 5, p. 19.
27. N. H. D. Ross and A. G. Crocker: *Acta Met.*, 1970, vol. 18, p. 405.
28. J. S. Bowles and D. P. Dunne: *Acta Met.*, 1969, vol. 17, p. 677.
29. R. Bullough and B. A. Bilby: *Proc. Phys. Soc. B*, 1956, vol. LXIX, p. 1276.
30. G. V. Kurdjumov and G. Sachs: *Z. Phys.*, 1930, vol. 64, p. 325.
31. C. M. Zener: *Elasticity and Anelasticity of Metals*, University of Chicago Press, 1948.
32. P. M. Kelly and J. Nutting: *J. Iron Steel Inst.*, 1961, vol. 197, p. 199.
33. See for example the general discussion on Martensite Nucleation by C. L. Magee in: *Phase Transformations*, American Society for Metals, Cleveland, 1970.
34. K. Shimizu, M. Oka, and C. M. Wayman: *Acta Met.*, 1970, vol. 18, p. 1005.
35. A. R. Entwisle: in *Discussion at AIME Symp. on Ferrous Martensite*, Las Vegas, 1970.
36. L. Kaufman and M. Cohen: *Progr. Metal Phys.*, 1958, vol. 7, p. 165, Pergamon Press.
37. M. J. P. Musgrave: *Crystal Acoustics*, Holden-Day Inc., San Francisco, 1970.
38. P. Clapp: Research Report, Ledgemont Laboratory, Kennecott Copper Co., Cambridge, Mass., 1971.