International Journal of Theoretical Physics, Vol. 29, No. 11, 1990

Cartan Connection and Defects in Bravais Lattices

L. Mistura[1]

Received April 11, 1990

The geometrical approach to a field theory of defects in crystalline solids including both dislocations and intrinsic point defects is developed in the framework of Cartan affine differential geometry. A clear distinction between linear and affine connection is made. The discussion is restricted to the basic kinematic level.

1. INTRODUCTION

That differential geometry, in particular affine differential geometry, provides the natural language to approach the so-called *many-defects problem* in crystalline solids has been widely recognized (Kröner, 1981) since the pioneering work of Kondo (1952) and Bilby *et al.* (1955), which discovered the identity between dislocation density, as defined by Nye (1953), and Cartan (1923) torsion.

Besides dislocations, which are line defects, we can identify in a monoatomic crystal other elementary defects, such as intrinsic point defects (self-interstitials and vacancies). The interactions between dislocations and point defects are important in explaining many properties of crystals. This implies, as noted by Kröner (1988), that a theory of dislocations which does not include a description of point defects cannot be complete. We would like to have a theory of defects to describe such elementary phenomena as the creation of vacancies and interstitials by the nonconservative motion (climb) of dislocations. Such a theory should take into account that dislocations and point defects are of a quite different nature. A point defect is a perturbation of a crystal that can be completely enclosed in a small sphere of a few atomic diameters in dimension, while the presence

[1]Dipartimento Energetica, Università "La Sapienza", 14-00161 Rome, Italy.

0020-7748/90/1100-1207$06.00/0

of a dislocation can be detected by the closure failure of a Burger circuit drawn at an arbitrary distance from the singular line, a property which can be restated in the modern language of the *homotopic classification of defects* (Rogula, 1976; Toulouse and Kléman, 1976; Kléman *et al.*, 1977) by saying that a dislocation is a topologically stable defect, while a point defect is not.

It is the purpose of this paper to present some developments of the geometrical approach to a field theory of defects in crystalline solids including both dislocations and intrinsic point defects. The discussion will be mainly restricted to the basic kinematic level, namely the introduction of a convenient set of so-called internal variables describing the internal mechanical state of the crystal as determinated by the presence of defects.

2. CRYSTALS AS ORDERED MEDIA

In order to talk about defects, one must first explain what one means by a crystal without defects, namely a *perfect crystal.* By a perfect crystal (monoatomic) I mean a nonuniform thermodynamic phase in equilibrium characterized by a triply periodic density distribution. By contrast, a fluid in equilibrium, in the absence of external fields, is characterized by a uniform density distribution. One can describe this situation by means of an *order parameter* in the sense of Landau, namely, a quasithermodynamic variable whose value is zero in the symmetric (fluid) phase and nonzero in the crystalline phase. The complex amplitudes of the Fourier components of the periodic density distribution can be used to this purpose. It is, however, traditional to represent schematically a (monoatomic) perfect crystal by means of a Bravais lattice. So I found it natural to take as order parameter manifold the *space* Λ of all Bravais lattices, which I suggest calling the *Rogula space* because it was introduced by Rogula (1976) in the topological classification of defects. Any Bravais lattice is represented by a single point of Λ and a continuous lattice deformation by a path in Λ.

From a field-theoretic point of view, the choice of Λ as the field manifold, namely the choice of a Bravais lattice as a local value of the order parameter, is admittedly odd. It is not difficult, however, to provide a more traditional, analytical representation of the order parameter space. Indeed, a Bravais lattice is completely determined by a point $P \in E(3)$ of the three-dimensional Euclidean space and three basic lattice vectors $\vec{a}_1$, $\vec{a}_2$, $\vec{a}_3$, so one can identify a Bravais lattice with an *affine moving frame* consisting of the origin P and the triad of basic vectors.

In any region of the material body corresponding to a "good crystal," one can associate to each point a value of the order parameter so that each point which is not in the core of a defect is labeled with a triad of basic vectors and a position vector which fixes the origin of the moving frame.

This procedure is reminiscent of that of constructing the tangent bundle of a manifold, with the difference that in that case the origin has an absolute meaning (only homogeneous transformations are allowed), while here the origin can be translated (affine tangent space). It is with these considerations in mind that I consider the *Cartan affine connection* (Cartan, 1923) instead of the more familiar *linear connection* traditionally used to describe crystal defects. A precise definition of the affine connection will be given in the next section; here I want to sketch briefly the principle of the topological classification.

The identification of the order parameter of a crystalline phase with an affine moving frame is satisfactory from the local point of view, but the manifold of all possible moving frames is significantly different from the Rogula space of Bravais lattices. Indeed, given a frame, a unique Bravais lattice is identified, conversely to each Bravais lattice there correspond infinitely many affine frames (Figure 1). Neither the origin nor the triad of basic vectors is determined by the lattice. More specifically, the origin can be any point of the lattice and the basic vectors are determined apart from a unimodular linear transformation represented by an integral

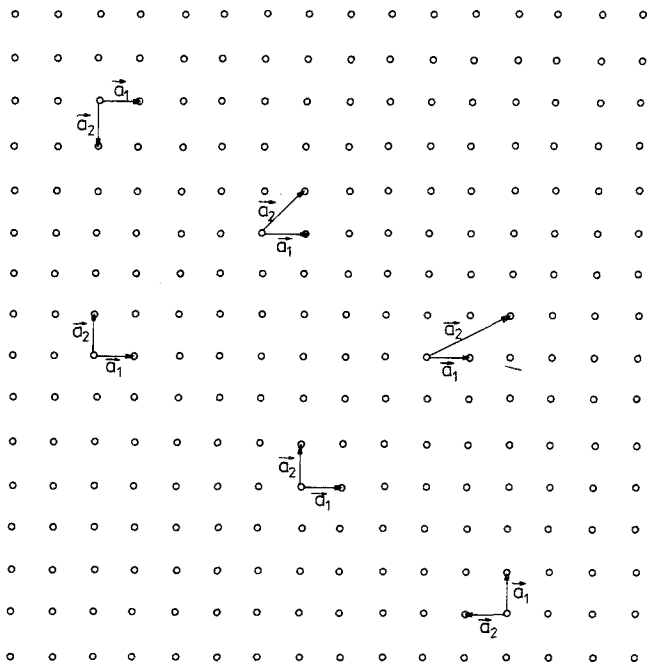

Fig. 1.

matrix. This situation, although formally more complex, is conceptually quite the analog of that found, e.g., in a nematic liquid crystal, to quote the simplest example, where the average orientation of the axis of the molecules is locally described by means of a unit vector $\vec{n}$, but the two opposite orientations of $\vec{n}$ have to be identified. As a consequence, the order parameter space, which is a two-sphere for $\vec{n}$, becomes the projective plane with different global topological properties. This in turn implies the possible existence of disclinations as stable topological defects.

In a crystal we have exactly the same situation: a moving affine frame may be used to identify a local coordinate system in the order parameter space, but to describe correctly the global topological structure of the Rogula space one must take into account the identification of different points imposed by the multivaluedness of the representation. By analogy with the nematic case, it is not difficult to see that it is just this identification procedure that leads to nontrivial topological properties and as a consequence to the possible existence of topologically stable line defects in the form of dislocations (nonuniqueness of the origin of the frame), disclinations (rotational arbitrariness of the basic vectors), and also, as pointed out by Rogula, line defects of shear type associated with lattice arbitrariness of the basic vectors not described by rotations. As mentioned before, an unambiguous result of the homotopic classification of defects is that intrinsic point defects, in the form of vacancies or interstitials, are not topologically stable.

3. LATTICE DEFORMATION AND AFFINE CONNECTION

In a crystal we can distinguish between two types of deformations, a nondissipative one, which is usually called *elastic*, and a dissipative one called *plastic* (Kröner, 1981).

At the purely kinematic level the elastic deformation is characterized by the fact that neighboring particles remain neighboring particles and the lattice is dragged along into the new particle configuration as schematically illustrated in Figure 2; it can therefore also be called *lattice deformation.* On the contrary, no change in the lattice structure occurs as a consequence of a plastic deformation (Figure 3), which can be obtained by adding or removing matter so that neighboring particles do not remain neighbors. Because such a deformation can be obtained by processes of slip and climb of dislocations as schematically illustrated in Figure 4, it can also be called *dislocation deformation.* Here the term deformation is used in a more general meaning than usual to include local rotations and translations. The specific term distortion is frequently used in a similar sense.

The existence of these two types of deformations is characteristic of crystalline solids. It is related to the *rigidity* of crystals, namely to the fact

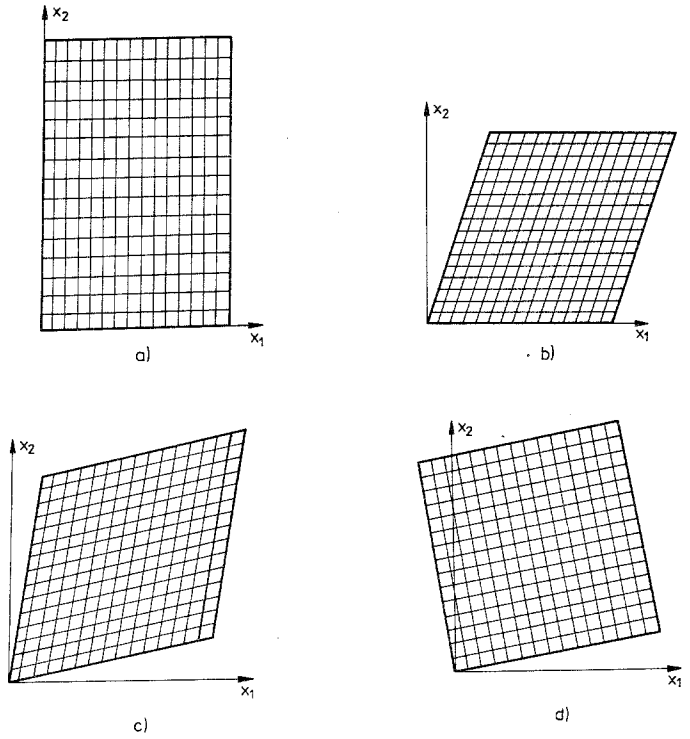

Fig. 2. Elastic deformation.

that the free energy is minimized when symmetry is broken perfectly in phase throughout the sample. The analogy between crystal rigidity and the possibility of superflow in ^{4}He and superconductors was mentioned by London (1961).

Let us consider now a monoatomic crystal in the ideal reference state at 0 K. Introduce a lattice distortion by means of a vector displacement field

$$\vec{u}(x^1, x^2, x^3) = u^\alpha \frac{\delta}{\delta x^\alpha}$$

where the x^α (Greek indices) are space coordinates which one may conveniently assume to be Cartesian and orthogonal. As a consequence, a material point originally at $P(x^1, x^2, x^3)$ goes into P' with coordinates $x^\alpha + u^\alpha$. The coordinates of P' will be identical to those of P in a new coordinate system with origin given by

$$\vec{e}_0 \equiv \vec{u}$$

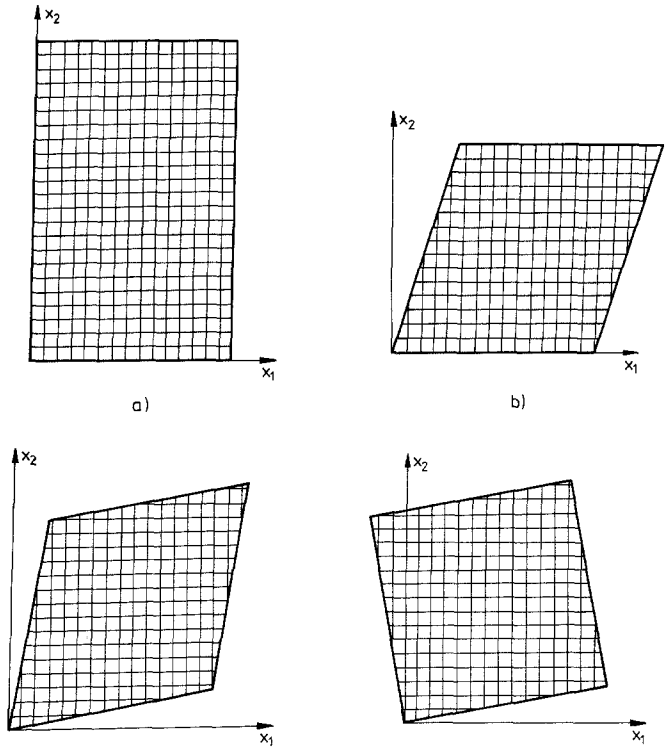

Fig. 3. Plastic deformation.

and new axes (see Figure 5), with unit vectors given by

$$\mathbf{e}_i = \frac{\partial}{\partial x^i}\left[(x^\alpha + u^\alpha)\frac{\partial}{\partial x^\alpha}\right] = \left[\delta_i^\alpha + \frac{\partial u^\alpha}{\partial x^i}\right]\frac{\partial}{\partial x^\alpha} \equiv B_i^\alpha \frac{\partial}{\partial x^\alpha}$$

These last two equations define local affine frames at each point of the crystal in the distorted state which are crystallographic. The frames at two

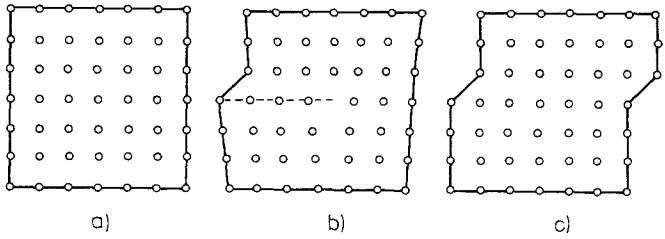

Fig. 4. Taylor illustration of plastic deformation due to dislocation motion.

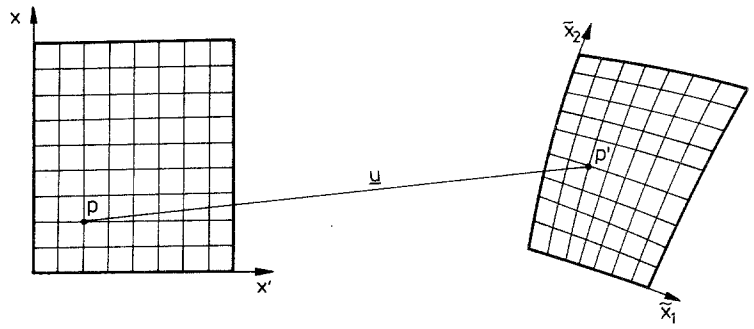

Fig. 5.

neighboring points are related by

$$d\mathbf{e}_0 = du^\alpha \frac{\partial}{\partial x^\alpha}$$

$$d\mathbf{e}_i = dB_i^\alpha \frac{\partial}{\partial x^\alpha}$$

Introducing the matrix B_α^J inverse to B_i^α ($B_i^J B_\alpha^J = \delta_i^J$), one gets

$$d\mathbf{e}_0 = du^\alpha B_\alpha^J \mathbf{e}_J \equiv w^J \mathbf{e}_J$$

$$d\mathbf{e}_i = dB_i^\alpha B_\alpha^J \mathbf{e}_J \equiv w_i^J \mathbf{e}_J$$

which are, in Cartan notation, the basic equations defining an affine connection. The assignment of the vector-valued differential form (w^J) and of the matrix-valued form w_i^J completely describes the distortion process. It is necessary at this point to give a precise definition of an affine connection as introduced by Cartan. Admittedly there is some confusion in the literature about the meaning of the term "affine connection." Traditionally, the terms "linear connection" and "affine connection" have been used interchangeably. We need instead a clear distinction. Briefly, an affine connection on a manifold M is a connection on the bundle of affine frames over M, while a linear connection is a connection on the bundles of linear frames over M (Kobayashi, 1963).

In the theory of the linear connection the tangent spaces $T_P(M)$ at each point P of a manifold M are obviously regarded as vector spaces. To define an affine connection, one must regard each tangent space as an affine space, namely as a space of points. The difference between a vector space and an affine space is that in the vector space the origin is fixed, and has an absolute meaning, while in the affine space it can be varied. Therefore an affine frame at a point P of a manifold is a frame of the tangent affine

space at P, namely a point 0 together with a basis of the associated tangent space at P. The point 0 is the *origin of the frame.* An affine connection [called *generalized affine connection* by Kobayashi and Nomizu (Kobayashi, 1963)] on M is an infinitesimal connection on the principal fiber bundle of affine frames on M. An affine connection in a manifold M defines, for each curve $P(t)$, $0 \leq t \leq 1$, on M the parallel displacement of the affine tangent space along the curve. By considering two infinitesimally neighboring points, one gets the basic equations (5) defining an affine connection in Cartan notation.

A linear connection on M can be regarded as a special kind of affine connection where $w^i = \Theta^i$ and Θ^i are the dual forms to the e_i, namely $\Theta^i(e_j) = \delta^i_j$.

The torsion τ and curvature Ω of an affine connection are defined by

$$\tau \equiv D(w^i e_i) \equiv (dw^i) e_i - w^i \wedge de_i = (dw^i + w^i_J \wedge w^J) e_i$$

$$\Omega_i \equiv D(w^J_i e_J) \equiv (dw^J_i + w^J_h \wedge w^h_i) e_J \equiv \Omega^J_i e_J$$

One can easily verify that in the case of the lattice distortion defined above, both torsion and curvature vanish. This is due to the fact that the elastic distortion is realized by means of a displacement vector field u. This is analogous to a potential superflow in ^{4}He ($v_s \simeq \text{grad}\, \varphi$ and therefore rot $v_s = 0$).

However, as in helium, in the presence of many vortices, one can describe the situation assuming the existence of the superflow, at least below a certain critical velocity, namely the existence of v_s but with rot $v_s \neq 0$, so that v_s is no longer the gradient of the phase of the condensate wave function, so in crystals one can imagine more general lattice distortion processes where one can still define an infinitesimal connection but without assuming the existence of a vector displacement field u. The simplest generalization is given by the equations

$$d\mathbf{e}_0 = \beta^\alpha \frac{\partial}{\partial x^\alpha}$$

$$d\mathbf{e}_i = dB^\alpha_i \frac{\partial}{\partial x^\alpha}$$

where β^α cannot be written as du^α, but we still assume

$$B^\alpha_i = \delta^\alpha_i + \beta^\alpha_i$$

so that we have a definite relationship between change in the basic vectors (rotation and strain) and shifting of the origin. The above pair of equations can also be written in Cartan form with $w^J = \beta^\alpha B^J a$ and $w^J_i = dB^\alpha_i \, B^J_\alpha$. With

these expressions for the connection forms, it is easily verified that the torsion is different from zero and it is given by

$$\tau = dB^{\alpha} B^{i}_{\alpha} e_i = d\beta^{\alpha} \frac{\delta}{\delta x^{\alpha}}$$

while the curvature Ω vanishes (*teleparallelism*).

The nonzero value of the torsion is a measure of the dislocation density in the same sense that the nonzero value of rot v_s in superfluids is a measure of the density of vortex lines, while the vanishing value of the curvature corresponds to the fact that, even in the presence of a large number of dislocations, while long-range positional correlation is lost, there still remains long-range orientational correlation.

A further consequence of the fact that $\Omega = 0$ is the conservation law (Bianchi identity)

$$D\tau = \Omega^{j}_{i} \wedge w^{i} = 0$$

This last equation shows that in the description given so far dislocations have an autonomous life: there is no possibility to describe their interaction either with the elastic field (phonons) or with point defects. The next step will be to explore the metric properties of a crystal.

4. DIFFERENTIAL GEOMETRY OF INTRINSIC POINT DEFECTS

As I have already mentioned, intrinsic point defects, such as vacancies and interstitials, are not topologically stable defects. Technically, this is due to the fact that the second homotopy group of the order parameter space of a crystal vanishes. This remark, however, does not help too much in trying to understand how to describe point defects, apart from the obvious conclusion that if they are not revealed by an observer who knows only topology, one must introduce further geometrical structures.

But point defects differ from dislocations in another important respect besides topological stability: they can be in thermal equilibrium. This means that, in the absence of dislocations or grain boundaries, one does not need to introduce extra variables to describe their presence in a crystal, because their average number is completely determined, in local equilibrium states, by the value of thermodynamic variables (temperature and pressure). The situation changes drastically in the presence of dislocations or grain boundaries, because of the interaction between these singularities and point defects. Figure 6 illustrates, for example, the creation of a vacancy and an interstitial through the interaction with the crystal surface. In these non-equilibrium situations one needs to introduce extra variables describing the internal mechanical state of the crystal due to the presence of point defects

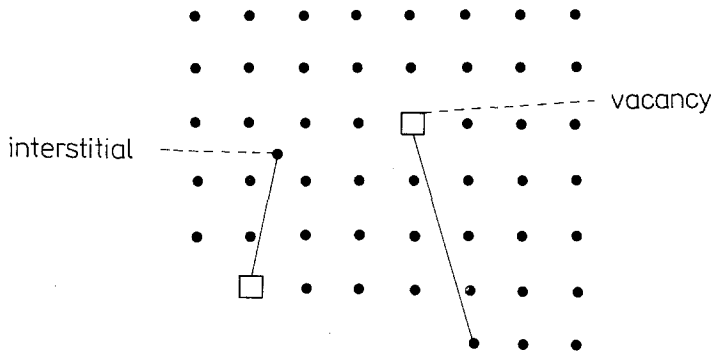

Fig. 6.

and their interaction with dislocations. It is therefore natural to try to extend to point defects the geometrical picture already available to describe dislocations. In this respect there is a suggestion by Kröner that point defects represent disturbances in the metric structure of the crystal.

In order to understand how this can happen, it is convenient to introduce an *internal observer.* This internal observer has no information on how to measure distances between atomic positions from the external, Euclidean space, in which the crystal is embedded. He has, however, all the notions of affine geometry and so he can move parallel to himself along lattice lines. He can measure distances along lattice lines by counting atomic steps, so that vacant sites will not be counted, while interstitials on lattice lines will be counted as regular atomic positions. He will not be able, however, to compare distances along different lattice lines. Nonetheless, because the atomic fraction of point defects, even at the melting point, is relatively small, it is natural to conceive that, sooner or later, the internal observer will discover the possibility of comparing distances along different lattice lines by means of Pythagoras's theorem. Indeed, with reference to a plane simple square lattice, as in Figure 7 the internal observer will find that the sum of the squares of the atomic steps along orthogonal directions has a constant ratio to the square of the atomic step along the diagonal. Although we can imagine that the observer had at the beginning the same religious repulsion toward $\sqrt{2}$ as did Pythagoras and his school, the validity of Pythagoras's theorem would finally be accepted. However, in a "bad" region of the crystal, where vacancies or self-interstitials are present along lattice lines, the internal observer will find deviations from the validity of Pythagoras's theorem. As these deviations can be detected by making a closed contour along lattice lines (see Figure 7), it is natural to look for a one-form to describe the presence of point defects.

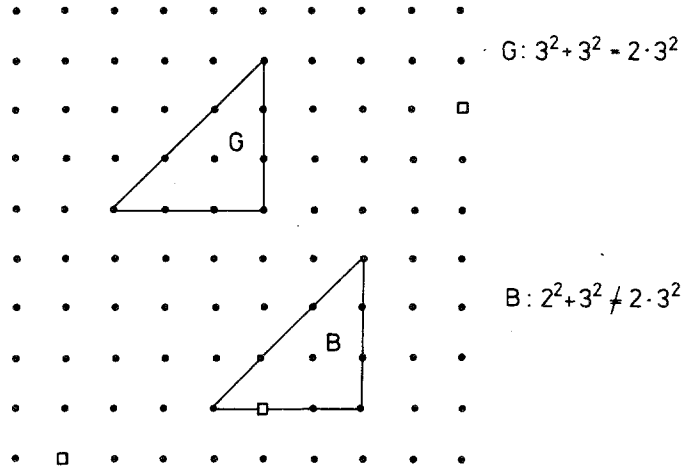

Fig. 7. A simple square lattice with vacancies.

The internal observer decides to measure distances between atomic positions as far as possible in agreement with the external, Euclidean observer, namely by putting

$$g_{iJ} = (\mathbf{e}_i, \mathbf{e}_J) = B_i^\alpha B_J^\beta \delta_{\alpha\beta} = B_i^\alpha B_J^\alpha \tag{1}$$

Then the 1-form

$$Q_{iJ} = dg_{iJ} - (d\mathbf{e}_i, \mathbf{e}_J) - (\mathbf{e}_i, d\mathbf{e}_J) \tag{2}$$

vanishes if there are no point defects, because one can compare distances along different directions. This remains true even in the presence of dislocations because

$$d\mathbf{e}_i = w_i^J \mathbf{e}_J = B_\alpha^J \, dB_i^\alpha \mathbf{e}_J \tag{3}$$

The simplest generalization of this connection law to allow for the presence of point defects ($Q_{iJ} \neq 0$) is

$$d\mathbf{e}_i = (wB_i^\alpha + dB_i^\alpha)\frac{\partial}{\partial x^\alpha} = (w\delta_i^J + B_\alpha^J \, dB_i^\alpha)\mathbf{e}_J \tag{4}$$

This implies, according to equations (1) and (2),

$$Q_{iJ} = -2wg_{iJ} \tag{5}$$

It is interesting to observe that with the new connection law (4), $d\tau$ (the covariant exterior derivative of the torsion) no longer vanishes, because the curvature is not zero:

$$dd\mathbf{e}_i = \Omega_i^J \mathbf{e}_J = dw \, \mathbf{e}_i \tag{6}$$

REFERENCES

Bilby, B. A., Bullough, R., and Smith, E. (1955). *Proceedings of the Royal Society*, **AR 31**, 263.
Cartan, E. (1923). *Annales de l'Ecole Normale Supérieure*, **40**, 325.
Kléman, M. L., Michel, L., and Toulouse, G. (1977). *Journal de Physique Lettres*, **38**, L-195.
Kobayashi, S., and Nomizu, K. (1963). *Foundations of Differential Geometry*, Vol. I, Interscience, New York.
Kondo, K. (1952). In *Proceedings 2nd Japan National Congress on Applied Mechanics*, Tokyo.
Kröner, E. (1981). Continuum theory of defects, in *Physique des défauts* (Les Houches, Session XXXV, 1980), R. Balian *et al.*, eds., North-Holland, Amsterdam.
Kröner, E. (1988). Point defects and metricity in Bravais crystals.
London, F. (1961). *Superfluids*, Vol. II, revised ed., Dover, New York.
Nye, J. F. (1953). *Acta Metallurgica*, **1**, 153.
Rogula, D. (1976). Large deformations of crystals, homotopy and defects, in *Trends in Applications of Pure Mathematics to Mechanics*, G. Fichera, ed., Pitman, London.
Toulouse, G., and Kléman, M. (1976). *Journal de Physique Lettres*, **37**, L-149.