The nonlocal theory of elasticity and its application to interaction of point defects

B. K. D. GAIROLA (STUTTGART)

THE lattice theory from the very beginning takes into account the finite range of internal forces and the discreteness of the crystals. The conventional theory of elasticity is too narrow to accomodate these concepts. In this paper we discuss the basic mathematical concepts of a nonlocal theory of elasticity which has been developed in the last few years as an alternative to the lattice theory. As a typical example of the application of this theory we consider the interaction between defects which takes place through the displacement fields produced by them when the lattice is allowed to relax. The results obtained by means of the nonlocal theory of elasticity are quite different from those of the conventional theory. In particular, we find that the energy of interaction between two spherically symmetric defects in an isotropic medium does not vanish, as it does according to the conventional theory.

Teoria sieci strukturalnych od samego początku uwzględnia skończony zakres sił wewnętrznych i dyskretność kryształów. Konwencjonalna teoria sprężystości jest zbyt wąska, aby z niej wynikały te koncepcje. W niniejszej pracy przedyskutowano podstawowe koncepcje matematyczne nielokalnej teorii sprężystości, która była rozwijana w ciągu ostatnich paru lat jako teoria alternatywna do teorii sieci strukturalnych. Jako typowy przykład zastosowania tej teorii rozważono oddziaływanie między defektami, jakie ma miejsce przez pola przemieszczeń, wywołane defektami, gdy dopuści się relaksację sieci. Wyniki otrzymane za pomocą nielokalnej teorii sprężystości są zupełnie inne niż rezultaty otrzymane z teorii konwencjonalnej. W szczególności wykazano, że energia oddziaływania między dwoma sferycznie symetrycznymi defektami w izotropowym ośrodku nie znika, jak to wynika z teorii konwencjonalnej.

Теория структурных решеток с самого начала учитывает конечный радиус действия внутренних сил и дискретность кристаллов. Конвенциональная теория упругости слишком узка, чтобы с ней вытекали эти концепции. В настоящей работе обсуждены основные математические концепции нелокальной теории упругости, которая развивалась в течение последних несколько лет как теория альтернативная к теории структурных решеток. Как типичный пример применения этой теории, рассмотрено взаимодействие между дефектами, что имеет место через поля перемещений, вызванные дефектами, когда допускается релаксация решетки. Результаты полученные с помощью нелокальной теории упругости совсем другие, чем результаты полученные из конвенциональной теории. В частности показано, что энергия взаимодействия между двумя сферически симметричными дефектами в изотропной среде не исчезает, как это следует из конвенциональной теории.

1. Introduction

A SOLID body is in fact an assembly of discrete particles which are held together by forces of a finite range and are, at least in ideal crystals, arranged in a regular lattice structure. Therefore, in any problem dealing with such bodies the two characteristic parameters which play a role are the range of interaction L and the discreteness length a. Obviously, L can never be less than a, i.e., $L \geqslant a$.

However, in any experimental investigation of the properties of crystal we use such an apparatus for observation which has a finite resolution. Thus, there is always a lower

limit to the length we can observe through the apparatus. We therefore have a third characteristic length which plays a role, namely the observable length λ. While the other two parameters make the theory nonlocal, it is this parameter which determines whether the theory is local or nonlocal. For instance, if λ is so large that it fulfills the condition $\lambda \gg L \gg a$, we obtain a local theory. Naturally, its application is limited to gross phenomena and a number of known effects elude it altogether.

The nonlocal theory which has been proposed independently by various authors [KRUMHANSL (1963), ROGULA (1965), KRÖNER (1965), KRÖNER and DATTA (1966), KUNIN (1966)] incorporates the essential properties of the lattice theory. The purpose of this paper is twofold. First, we discuss briefly the underlying mathematical concepts of the nonlocal theory. All the relevant mathematical results are given with brief explanations so that the reader may follow the arguments without much difficulty. Second, as an example of the application of the nonlocal theory, we consider the problem of interaction of point defects. Our aim is to illustrate that the nonlocal theory yields results quite different from those of the classical theory and are in better agreement with the lattice theory.

2. Nonlocality due to long range interaction

The nonlocal elasticity theory is a field theory in the sense that all properties of solid bodies appear as fields, i.e., as functions of continuous variables, space coordinates. In other words, the continuum here is actually a mathematical continuum. It follows that nonlocality can arise in this continuum theory not only from the long range of interaction but also from the discreteness. This is clear since, in the extreme case, $L = a$, which is a finite quantity.

However, let us first consider only the long range of interaction without referring to the discreteness. In this case the strain energy W is a generating functional which, for small strains, can be expanded in a functional series. For the sake of simplicity let us assume that the initial state is stress-free and the strains ϵ are so small that the terms of order higher than second can be neglected. Then, we can write

$$(2.1) \qquad W[\epsilon] = \frac{1}{2} \int \epsilon(\mathbf{x}) \ldots \mathbf{C}(\mathbf{x}, \mathbf{x}') \ldots \epsilon(\mathbf{x}') dv\, dv' = \frac{1}{2} \langle \epsilon | \mathbf{C} | \epsilon \rangle,$$

where $|\epsilon\rangle$ is a vector in the function space and $\mathbf{C}$ an integral operator. The stress $|\sigma\rangle$ is then defined as a variational derivative:

$$(2.2) \qquad |\sigma\rangle = \frac{\delta W}{\delta|\epsilon\rangle} = \mathbf{C}|\epsilon\rangle.$$

If the medium is homogeneous, the kernel of $\mathbf{C}$ will depend only on the difference $\mathbf{x}-\mathbf{x}'$. Hence, in this case (2.2) is a convolutions equation which, in the usual notation, reads

$$(2.3) \qquad \sigma(\mathbf{x}) = \int C(\mathbf{x}-\mathbf{x}') \ldots \epsilon(\mathbf{x}') dv',$$

or

$$\sigma = C \divideontimes \epsilon.$$

In the above derivation we have used the idealization that a field variable such as the strain tensor ϵ can be defined at every point of space. In reality the field at a point cannot be measured; only its averages are observable as this may be inferred from the analyses of BOHR and ROSENFELD (1933, 1950). Measurements are made by operating with a testing body. As a result of interaction between the field and the testing body the averaging process depends on both. Therefore, we obtain averages or "smeared fields" such as

$$(2.4) \qquad \epsilon(\varphi) = \int \epsilon(\mathbf{x}) \cdot \varphi(\mathbf{x}) dv = \langle \epsilon|\varphi \rangle,$$

where $\varphi(\mathbf{x})$ are smooth functions called averaging functions or test functions. They may be thought of as a weight function representing the influence of the testing body. The continuous linear functional $\epsilon(\varphi)$ is called a distribution or a generalized function.

There are several types of test function spaces and accordingly one can define a variety of distributions (see, e.g., SCHWARTZ 1966 or GELFAND-SCHILOW 1964–1968). The choice of the proper space of distributions should be dictated by the physical requirements of the theory. There are a number of ways in which we could approach this problem but the simplest method by far is to say that, since there is no discreteness, there is also no restriction on the Fourier spectrum of the field variables. In other words, the support of the distributions in Fourier space can be as small or as large as we please. (The support of a distribution ϵ is the smallest closed set of points outside of which ϵ vanishes.) It is, therefore, desirable to choose the narrowest space of test functions so that we have the widest space of distributions. This is the space of test functions of compact support usually denoted by D. The corresponding dual space of distributions is denoted by D'. An example of such distributions is the function $e^{|\mathbf{x}|}$.

The convolution of the two distributions ϵ and C is defined by the rule

$$(2.5) \qquad \langle \epsilon \ast C|\varphi \rangle = \langle \epsilon|\overline{C} \ast \varphi \rangle,$$

where φ is a test function and

$$(2.6) \qquad \overline{C}(\mathbf{x}-\mathbf{x}') \equiv C(\mathbf{x}'-\mathbf{x}).$$

However, as a result of the symmetry properties of C we have

$$(2.7) \qquad C(\mathbf{x}-\mathbf{x}') = C(\mathbf{x}'-\mathbf{x}).$$

Hence, (5) may be written

$$(2.8) \qquad \langle \epsilon \ast C|\varphi \rangle = \langle \epsilon|C \ast \varphi \rangle.$$

If $\varphi \in D$, then $C \ast \varphi$ should also belong to D. Hence, if a sequence $\{\varphi_m\}$ of the functions $\varphi(\mathbf{x}) \in D$ $(m = 1, 2, ...)$ converges to zero in the sense of D, then $C \ast \varphi_m$ should also converge to zero in D. This is possible if, and only if, the supports of the functions $C \ast \varphi_m$ all lie within a certain bounded region. Therefore, it is not possible to define, in general, a convolution for two arbitrary distributions. We have to impose some condition on the distribution C so that it can form a convolution with the distribution ϵ. This condition is that C be a distribution with compact support. (A closed and bounded set is called compact.) Thus, the convolution, the Eq. (2.3) has a meaning if this condition is fulfilled. Here, it should be added, a solution of this equation exists which may be put in the form

$$(2.9) \qquad \epsilon = C^{-1} \ast \sigma,$$

where C^{-1} is the inverse of C defined by

$$(2.10) \qquad\qquad C \ast C^{-1} = I\delta.$$

Here, I is the unit tensor of the fourth rank and δ is the Dirac distribution. This follows from the fact that the space of distributions with compact support is a commutative convolution algebra with δ as the unit element.

It seems logical that the support of C should be compact because, after all, it is determined by the range of interaction which is finite. If it were not finite, the convolution integral (2.3) might diverge. Of course, divergent concolution integrals do occur in the quantum field theory. There have been many attempts to give a finite and well-defined meaning to these integrals (see, for instance, BOGOLIUBOV and PARASIUK, 1957; AKHIEZER and BEREZTEZKI, 1965). Such a situation may also occur in a stellar body reaching a critical mass.

However, a sharp cut-off in the range of interaction does not seem to be very reasonable from the physical point of view. The known interactions which are of relevance here attenuate only gradually and not suddenly. In this case the distribution C with compact support is not satisfactory. This undesirable feature of the continuum theory could be removed if we considered some suitably larger space of test functions. From the mathematical point of view, however, there does not seem to be any logical necessity for this. The necessity arises, as we shall see in the next section, in the case of a discrete medium.

3. Nonlocality in a discrete medium

A discrete medium is most simply characterized by the fact that all the variables in the problem have their Fourier spectrum truncated at a certain value of the wave vector. This means that the corresponding distributions have a compact support in the Fourier space. This support, in fact, is the Brillouin zone. Such distributions can also be written in the form of a product of a distribution with arbitrary large support and a function δ_B defined by

$$(3.1) \qquad\qquad \delta_B(\mathbf{k}) = \begin{cases} 1 \text{ for } \mathbf{k} \in \text{Brillouin zone,} \\ 0 \text{ for } \mathbf{k} \notin \text{Brillouin zone.} \end{cases}$$

For instance,

$$(3.2) \qquad\qquad \epsilon(\mathbf{k}) = \epsilon'(\mathbf{k})\,\delta_B(\mathbf{k}).$$

The multiplication of a distribution $\epsilon'(\mathbf{k})$ with a function $\delta_B(\mathbf{k})$ is defined by the rule

$$(3.3) \qquad\qquad \langle \epsilon'(\mathbf{k})\,\delta_B(\mathbf{k})|\varphi(\mathbf{k})\rangle = \langle \epsilon'(\mathbf{k})|\delta_B(\mathbf{k})\varphi(\mathbf{k})\rangle.$$

If we put $\psi(\mathbf{k}) = \delta_B(\mathbf{k})\varphi(\mathbf{k})$, we can write this equation as

$$(3.4) \qquad\qquad \epsilon'(\psi(\mathbf{k})) = \langle \epsilon'(\mathbf{k})|\psi(\mathbf{k})\rangle.$$

Now, the following question arises: what is the nature of the test function $\psi(\mathbf{x})$ and the distribution $\epsilon(\psi)$. Clearly enough $\psi(\mathbf{k}) \in D$ and its inverse Fourier transform is

$$(3.5) \qquad\qquad \psi(\mathbf{x}) = \int \psi(\mathbf{k})\,e^{i\mathbf{k}\cdot\mathbf{x}}dv.$$

If we replace $\mathbf{x}$ by a complex variable $\mathbf{z} = \mathbf{x} + i\mathbf{s}$, we get

$$(3.6) \qquad \psi(\mathbf{z}) = \int \psi(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{z}} dv_k = \int \psi(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{x} - \mathbf{k} \cdot \mathbf{s}} dv_k.$$

The function $\psi(\mathbf{z})$ is an entire function in the complex hyperplane. The holomorphy of $\psi(\mathbf{z})$ follows from the fact that $\psi(\mathbf{k})$ has a compact support. Such functions satisfy the inequality

$$(3.7) \qquad |z^p \psi(\mathbf{z})| = \left| \int D^p \psi(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{z}} dv_k \right| \leqslant \Lambda_p e^{a|\mathbf{s}|}.$$

Here, Λ_p and a are constants which may depend on ψ; z^p, and $D^p \psi(\mathbf{k})$ are short notations for the expressions

$$z^p = z_1^{p_1} z_2^{p_2} z_3^{p_3}$$

and

$$D^p \psi(\mathbf{k}) = \frac{\partial^{|p|} \psi(k)}{\partial k_1^{p_1} \partial k_2^{p_2} \partial k_3^{p_3}}$$

with

$$|p| = p_1 + p_2 + p_3.$$

One can show (see GELFAND SCHILOW, vol. II, p. 154) that function $\psi(\mathbf{z})$ satisfies the more general inequality

$$(3.8) \qquad |z^p D^q \psi(\mathbf{z})| \leqslant \Lambda_{p,q} e^{|as|}.$$

The left-hand side of this inequality is called the Schwartz norm.

We shall now consider the function $\psi(\mathbf{z})$ only for the real argument $\mathbf{x}$. By putting $\mathbf{s} = 0$ in the inequality (3.8) we find that the Schwartz-norm of the test function $\psi(\mathbf{r})$ satisfies the condition

$$(3.9) \qquad |x^p D^q \psi(\mathbf{x})| \leqslant \Lambda p,q.$$

This means that $\psi(\mathbf{x})$ is a rapidly decreasing function. More exactly, $\psi(\mathbf{x})$ is a smooth function which, together with its derivatives, approaches zero more rapidly than any power of $\frac{1}{|\mathbf{x}|}$ as $|\mathbf{x}| \to \infty$. The space of these test functions is usually denoted by S and the distributions which are defined as the linear functionals on this space are called tempered distributions; the space of these distributions is denoted by S'. With the help of Riesz's theorem in functional analysis (see, e.g., KANTOROWITSCH and AKILOW, 1964, p. 168) one can easily show that the tempered distributions have the form

$$(3.10) \qquad \epsilon'(\psi) = \sum_{m \leqslant n} \int (1 + |\mathbf{x}|)^n \mathbf{f}_m(\mathbf{x}) D^m \psi(\mathbf{x}) dv,$$

where $m \leqslant n$ means $m_1 \leqslant n_1$, etc. m_1, n_1, etc. are whole numbers and $\mathbf{f}_m(\mathbf{x})$ are measurable and essentially bounded functions.

Since $\psi(\mathbf{x})$ is rapidly decreasing it can easily be seen that the functions $\psi_n(\mathbf{x})$ defined by

$$(3.11) \qquad \psi_n(\mathbf{x}) = (1 + |\mathbf{x}|)^n \psi(\mathbf{x})$$

also belong to S. Therefore, by means of Hahn-Banach theorem one can show that tempered distributions can be represented as a finite sum of derivatives of continuous func-

tions growing at infinity not faster than some polynomial. (The term tempered is used because of this slow growth).

Referring back to (3.3) we may note that this multiplication rule is meaningful only if $\delta_B(\mathbf{k})\varphi(\mathbf{k})$ and $\varphi(\mathbf{k})$ belong to the same space of test functions. Hence, the test function $\varphi(\mathbf{x})$ is an element of S. It follows that distribution $\epsilon = \epsilon' \ast \delta_B$ is again a tempered one. From (3.1) it can easily be shown that

$$(3.12) \qquad\qquad \delta_B \ast \delta_B = \delta_B$$

and

$$(3.13) \qquad\qquad \epsilon \ast \delta_B = \epsilon' \ast \delta_B \ast \delta_B = \epsilon' \ast \delta_B = \epsilon$$

so that δ_B acts as a unit element in the convolution algebra. Although the support of $\delta_B(\mathbf{x})$ is the whole space one could define an effective support with the dimensions of not more than double the lattice spacings. This is obvious from the following two examples. For a spherical Brillouin zone

$$(3.14) \qquad\qquad \delta_B(\mathbf{x}) = \frac{1}{2\pi^2}\left[\frac{\sin k_B|\mathbf{x}|}{|\mathbf{x}|^3} - \frac{k_B\cos k_B|\mathbf{x}|}{|\mathbf{x}|^2}\right],$$

where k_B is the radius of the B.z. sphere, and for a cubic Brillouin zone

$$(3.15) \qquad\qquad \delta_B(\mathbf{x}) = \prod_i^3 \frac{\sin k_{Bi}x_i}{\pi x_i a^3}\ .$$

Thus, we have proved the following theorem:

THEOREM. *If the Fourier transform of a distribution has a compact support in Fourier space, it is a tempered distribution of finite algebraic growth for* $|\mathbf{x}| \to \infty$.

The space of these distributions is a commutative convolution algebra with δ_B *defined by* (3.1) *as the unit element.*

Our discussion is a highly simplified version of a more complicated mathematical situation. The above theorem is actually a converse of the one known as the Paley-Wiener-Schwartz theorem (see, e.g., HÖRMANDER, 1963, p. 22 or YOSIDA, 1971, p. 162).

After this discussion we can see that, according to the convolution rule, both C ($= C' \ast \delta_B$) and ϵ in (2.3) should be considered as tempered distributions. However, to give a meaning to this convolution integral the growth at infinity of ϵ' should be "matched" by the decay of C'. As the growth of ϵ' is slow at infinity, it should be sufficient to require that the decay of C' at infinity be faster than any power of $\frac{1}{|\mathbf{x}|}$, i.e., it should decrease rapidly enough as $|\mathbf{x}| \to \infty$.

In physics a "rapidly enough" decrease is usually considered to be an exponential decrease. For instance, we can put $C' = Cg_L$, where

$$(3.16) \qquad\qquad g_L(\mathbf{x}) = \exp(-|\mathbf{x}|^2/L^2)/(\sqrt{\pi}L)^2.$$

Hence,

$$(3.17) \qquad\qquad C = Cg_L \ast \delta_B.$$

Obviously, the sphere of radius L is the effective support of this function. (Note that despite its appearance the function g_L does not belong to S because the sequence $\{g_L\}$

does not converge to a function $\varphi \in S$ which is a necessary condition for the space S to be closed.) However, a serious difficulty arises with this distribution. Its inverse C^{-1} does not belong to S'. Hence, some additional condition must be specified so that a solution of the form (2.9) can exist.

This condition can be found through mathematical reasoning if we consider the distribution on the space of test functions $\varphi(z)$ as defined earlier and use the theory of analytic functions. However, we shall opt here for the shorter way of physical reasoning. The interaction between the particles forming a solid body are, in reality, of two kinds: short range repulsive interaction and long range attractive interaction. Here, by the term range we mean effective range without a sharp cut-off. Since the repulsive interaction is much stronger than the attractive one we can assume that

$$(3.18) \qquad C = C_0\,\delta_B + C_1 g_L \divideontimes \delta_B,$$

where $C_0 \gg C_1$ and L is the effective range of attractive interaction. The inverse of C is easily calculated and is given by the convergent series

$$(3.19) \qquad C^{-1} = C_0^{-1}\delta_B - C_0^{-1} \cdot\cdot\ C_1 \cdot\cdot\ C^{-1}g_L \divideontimes \delta_B + \ \dots \ .$$

It belongs to S'. On the other hand, it would be incorrect to put

$$(3.20) \qquad C = (C_0 g_l + C_1 g_L) \divideontimes \delta_B$$

where l is the effective range of repulsive interaction, because then $C^{-1} \notin S'$.

In the extreme case when the interaction range L is of the order of the lattice spacing a we have

$$(3.21) \qquad C = C_0\,\delta_B + C_1 g_a \divideontimes \delta_B.$$

Both terms on the right-hand side of (3.21) are nonlocal. Therefore, in an absolute sense, there is no local law in a discrete medium.

Let us now use the observable length λ as the measuring unit. This means that we introduce the dimensionless variable $\mathbf{y} = \mathbf{x}/\lambda$ or $\mathbf{x} = \lambda\mathbf{y}$ or, in Fourier space, the dimensionless variable $\mathbf{k}' = \lambda\mathbf{k}$. Then, in accordance with the rule of similarity transformation of the distributions, we have

$$(3.22) \qquad \langle f(\lambda\mathbf{y})|\varphi(\mathbf{y})\rangle = \langle f(\mathbf{y})|\lambda^{-3}\varphi(\mathbf{y}/\lambda)\rangle,$$

or

$$\langle f(\mathbf{k}'/\lambda)|\varphi(\mathbf{k}')\rangle = \langle f(\mathbf{k}')|\lambda^3\varphi(\lambda\mathbf{k}')\rangle.$$

It can be seen that as $\lambda \to \infty$, the support of $\delta_B(\mathbf{k}')$ becomes infinite, i.e., $\delta_B \to \delta$. On the other hand, g_L becomes $g_{L/\lambda}$ defining a delta sequence $\{g_{L/\lambda}\}$ which converges to δ function. Hence, in the limit $\lambda \to \infty$ we have

$$(3.23) \qquad C = (C_0 + C_1)\,\delta$$

which yields a local material law. If $\lambda \gg a$ but L is comparable to λ, we obtain the nonlocal material law of the previous section but without a sharp cut-off, under the form of

$$(3.24) \qquad C = C_0\,\delta + C_1 g_L.$$

We may thus see that the most general correct form is given by (3.18). It is also clear that such a medium is not, in general, isotropic.

10*

4. Equation of equilibrium and its solution in terms of Green's function

If a body force per unit mass is also present we can introduce another quantity $\overline{w}$ which is a functional of the displacement $\mathbf{u}$. Its variation is then given by

$$(4.1) \qquad \delta\overline{w} = \langle\mathbf{F}|\delta\mathbf{u}\rangle,$$

where

$$(4.2) \qquad |\mathbf{F}\rangle = \frac{\delta\overline{w}}{\delta|\mathbf{u}\rangle}.$$

Then, according to Hamilton's principle,

$$\delta w + \delta\overline{w} = \langle\sigma|\delta\epsilon\rangle + \langle\mathbf{F}|\delta\mathbf{u}\rangle = 0.$$

Since the operation δ commutes with the partial differentiation, we can write

$$(4.3) \qquad \delta\epsilon = \text{Def}\,\delta\mathbf{u},$$

where

$$(4.4) \qquad (\text{Def})_{ijk} = \delta_{ik}\frac{\partial}{\partial x_j} + \delta_{jk}\frac{\partial}{\partial x_i}.$$

Then, using partial integration and the nonlocal material law, we finally obtain the following inhomogeneous convolution equation

$$(4.5) \qquad \mathbf{D}\divideontimes\mathbf{u} + \mathbf{F} = 0,$$

where

$$(4.6) \qquad \mathbf{D} = \nabla_\mathbf{x}\cdot C\cdot\nabla_{\mathbf{x}'} = -\nabla_\mathbf{x}\cdot C\cdot\nabla_\mathbf{x}.$$

Thus, $\mathbf{D}$ is a self-adjoint operator.

We define Green's function $\mathbf{G}$ by

$$(4.7) \qquad \mathbf{D}\divideontimes\mathbf{G} = \mathbf{1}\delta_B,$$

where $\mathbf{1}$ is the usual unit tensor of second rank. As a boundary condition we assume that $\mathbf{G}$ vanishes at infinity. Then $\mathbf{G}$ is given by

$$(4.8) \qquad \mathbf{G} = |\mathbf{D}|^{-1}\divideontimes\gamma,$$

where $|\mathbf{D}|$ is the determinant of $\mathbf{D}$ defined by

$$(4.9) \qquad |\mathbf{D}| = \frac{1}{6}\mathbf{D}\times\times\mathbf{D}\cdot\cdot\mathbf{D}$$

and γ is the cofactor

$$\gamma = \frac{1}{2}\mathbf{D}^T\times\times\mathbf{D}^T.$$

However, it would be difficult to calculate $\mathbf{G}$ in this way. The nonlocal form (3.18) of C on the other hand makes it possible for us to use a perturbation method. The operator $\mathbf{D}$ can be put in the form

$$(4.10) \qquad \mathbf{D} = \mathbf{D}_0 + \mathbf{D}_1,$$

where

$$(4.11) \qquad \begin{aligned} \mathbf{D}_0 &= \nabla_\mathbf{x}\cdot C_0\cdot\nabla_{\mathbf{x}'}\delta_B. \\ \mathbf{D}_1 &= \nabla_\mathbf{x}\cdot C_1\cdot\nabla_{\mathbf{x}'}g_L\divideontimes\delta_B. \end{aligned}$$

We introduce Green's tensor $\mathbf{G}_0$ which satisfies the equation

(4.12) $$\mathbf{D}_0 \divideontimes \mathbf{G}_0 = \mathbf{1}\delta_B.$$

Naturally, it is much simpler to calculate $\mathbf{G}_0$.

We can then write (4.7) either in the form

(4.13) $$\mathbf{G} = \mathbf{G}_0 - \mathbf{G}_0 \divideontimes \mathbf{T} \divideontimes \mathbf{G}_0,$$

where the "T-matrix" is given by

(4.14) $$\mathbf{T} = \mathbf{D}_1(\delta_B + \mathbf{G}_0 \divideontimes \mathbf{D}_1)^{-1}$$

or in the form of Dyson's equation

(4.15) $$\mathbf{G} = \sigma_0 - \mathbf{G}_0 \divideontimes \mathbf{D}_1 \divideontimes \mathbf{G}.$$

Any of these expressions may readily be expanded so that

(4.16) $$\mathbf{G} = \mathbf{G}_0 - \mathbf{G}_0 \divideontimes \mathbf{D}_1 \divideontimes \mathbf{G}_0 + \mathbf{G}_0 \divideontimes \mathbf{D}_1 \divideontimes \mathbf{G}_0 \divideontimes \mathbf{D}_1 \divideontimes \mathbf{G}_0 - \ldots .$$

This is the usual Neumann series. The simplest Green's function that we can calculate is that for an isotropic medium though, this is not realistic. Considering only the first two terms of the perturbation series (4.16), $\mathbf{G}$ can be written as $\mathbf{G} = \mathbf{G}' \divideontimes \delta_B$ with

(4.17) $$\mathbf{G}'(\mathbf{x}-\mathbf{x}') = -\frac{1}{8\pi\mu_0} [\mathbf{1}\nabla_x \cdot \nabla_x |\mathbf{x}-\mathbf{x}'| - \alpha \nabla_x \nabla_x' |\mathbf{x}-\mathbf{x}'|$$

$$+ \beta\mathbf{1} \int \frac{g_L(\mathbf{x}-\mathbf{x}'')}{|\mathbf{x}''-\mathbf{x}'|} \, dv'' - \nu \int g_L(\mathbf{x}-\mathbf{x}'')\nabla_{x'}\nabla_{x'} \frac{1}{|\mathbf{x}''-\mathbf{x}'|} \, dv'';$$

(4.18)
$$\alpha = \frac{\lambda_0 + \mu_0}{\lambda_0 + 2\mu_0}, \qquad \beta = \frac{\mu_1}{\mu_0},$$

$$\nu = \frac{1}{\lambda_0 + 2\mu_0}\left[\frac{\mu_1}{\mu_0}(\lambda_0 + \mu_0) - \frac{\mu_0(\lambda_1 + \mu_1)}{\lambda_0 + 2\mu_0} + \frac{\mu_1(\lambda_0 + \mu_0)}{\lambda_0 + 2\mu_0}\right],$$

where λ_0, μ_0 and λ_1, μ_1 are Lamé's constants corresponding to C_0 and C_1.

5. Interaction between point defects

In the lattice theory the term point defect is generally applied to vacancies, interstitials and substitutional atoms. When two point defects are introduced into a crystal they can interact in various ways. However, the interaction which takes place through the displacement fields produced by defects when the lattice is allowed to relax is always present. It is this very interaction which we are now considering. The knowledge of this interaction is important in determining, for instance, the formation energy and also for the discussion of X-ray scattering. We shall exclude interstitials from our consideration because the displacements they produce are so large that a linear theory is not permissible. Besides, they have a much more complicated structure.

In the classical continuum theory these defects are characterized by a dipole type of force distributions (size effect) or δ-type inhomogeneities (modulus effect). In the most

general continuum theory as exposed above, the defects cannot be represented in either of these ways because a strictly localized model is not feasible unless λ is very large in comparison with the lattice spacing and the defect interacts only with a few neighbours. Here, we shall characterize the defect by a nonlocal force distribution because it is general enough to include most of the effects. The inhomogeneity model has already been considered by KOSILOVA, KUNIN and SOSNINA (1968).

Let the force density distribution representing a defect at x be $F(x-x')$. It should satisfy the equilibrium conditions

$$(5.1) \qquad \int F(x-x')\,dv' = 0, \quad \int F(x-x') \times x'\,dv' = 0.$$

From the above it follows that

$$(5.2) \qquad F = \nabla_x \cdot P,$$

where P is a symmetric tensor distribution of second rank. Obviously, it plays the role of a potential function. The form of P and hence of F is determined by the interaction between the impurity atom and the atoms of the host lattice or, in the case of a vacancy, by the interaction between the atoms themselves. The support of P, in the case of an impurity atom, may be of an order different from the range of interaction between the atoms of the host lattice, but for a vacancy it will be of the same order. Therefore, we shall represent P in the form

$$(5.3) \qquad P = A g_l \divideontimes \delta_B,$$

where A is a constant tensor and g_l is a rapidly decreasing function with an effective support of the order l. It is not too much of a restriction to assume that $\{g_l\}$ is a delta sequence so that P reduces to $A\delta$ in the long wave limit ($\lambda \to \infty$).

Consider now two defects at x and x', respectively, which are represented by the force density distributions

$$(5.4) \qquad F_1 = \nabla_x' \cdot P_1, \quad F_2 = \nabla_{x'} \cdot P_2,$$

where

$$(5.5) \qquad P_1 = A_1 g_{l_1} \divideontimes \delta_B, \quad P_2 = A_2 g_{l_2} \divideontimes \delta_B.$$

The interaction energy of the two defects can be written as

$$(5.6) \qquad U = -F_1 \divideontimes u_2,$$

where u_2 is the displacement due to the second defect. It is given by

$$(5.7) \qquad u_2 = -G \divideontimes F_2.$$

Hence, we obtain

$$(5.8) \qquad U = F_1 \divideontimes G \divideontimes F_2.$$

Using the perturbation series (4.16) for G we get

$$(5.9) \qquad U = F_1 \divideontimes G_0 \divideontimes F_2 - F_1 \divideontimes G_0 D_1 \divideontimes G_0 \divideontimes F_2 + \dots.$$

Let us now turn to the simplest case when the medium is isotropic and the defect is spherically symmetric. However, as we have already pointed out, this is an unrealistic case. First of all, even in the long wave approximation there are only a few lattices which

are isotropic or nearly isotropic (W, diamond and Al). Secondly, the spherical symmetric model of the defect does not correspond very well to actual lattice defects because the highest symmetry a lattice defect can have is cubic symmetry. But this case is still instructive because the results reflect qualitative differences between the classical theory and the nonlocal theory. More realistic models will be considered in another paper.

Therefore, we put

(5.10) $$\mathbf{P}_1 = A_1 \mathbf{1} g_{l_1} \overline{\ast} \delta_B, \qquad \mathbf{P}_2 = A_2 \mathbf{1} g_{l_2} \overline{\ast} \delta_B.$$

The distributions g_{l_1} and g_{l_2} are of the same type of exponential functions as g_L in (3.16), and l_1, l_2 define the effective extensions of the defects. These functions have the simple property that

$$g_{l_1} \overline{\ast} g_{l_2} = g_{l_1 + l_2}.$$

In view of (5.9) the interaction energy of the two defects is now given by

(5.11) $$U = M g_{l_1 + l_2} \overline{\ast} \delta_B - N g_{l_1 + l_2 + L} \overline{\ast} \delta_B + \ldots,$$

where

$$M = \frac{A_1 A_2}{\lambda_0 + 2\mu_0},$$

$$N = \frac{A_1 A_2}{\mu_0} \left(v - \frac{\mu_1}{\mu_0} \right).$$

In real crystals the Brillouin zone is a polyhedron. However, because of our simplifying assumption we can take it here as a sphere of the radius k_B. Carrying out the convolution over this sphere we obtain

(5.12) $$U(\mathbf{x} - \mathbf{x}') = M g_{l_1 + l_2}(\mathbf{x} - \mathbf{x}') R_e \mathrm{erf} \left[\frac{k_B (l_1^2 + l_2^2)^{1/2}}{2} + \frac{i |\mathbf{x} - \mathbf{x}'|}{(l_1^2 + l_2^2)^2} \right]$$

$$- N g_{l_1 + l_2 + L}(\mathbf{x} - \mathbf{x}') R_e \mathrm{erf} \left[\frac{k_B (l_1^2 + l_2^2 + L^2)^{1/2}}{2} + \frac{i |\mathbf{x} - \mathbf{x}'|}{(l_1^2 + l_2^2 + L^2)^2} \right]$$

$$- \frac{\sin k_B |\mathbf{x} - \mathbf{x}'|}{\pi^2 |\mathbf{x} - \mathbf{x}'|} |\exp[-k_B (l_1^2 + l_2^2)/4] \left[\frac{N \exp(-k_B^2 L^2/4)}{l_1^2 + l_2^2 + L^2} - \frac{M}{l_1^2 + l_2^2} \right].$$

At large distances $|\mathbf{x} - \mathbf{x}'|$ U exhibits an oscillatory behaviour.

In the expression (5.11) we may notice an interesting fact. The interaction potential U of the two defects is of the same form as the potential **P** of the defects. This is not unusual. In the electron theory of metals, as it is well known, such a behaviour is also manifested by the interaction of vacancies in metals, or the interaction of charged impurities ([1]). This interaction is oscillatory just like the effective interaction between ions in metals.

([1]) This is one case where the defect could be considered as being spherically symmetric. The reason for this is that charged impurity is screened by the conduction electrons. According to the Fermi-distribution of the conduction electron the Fermi-surface represents a sharp boundary between the occupied and unoccupied electron-states in the Fourier-space. Hence, the potential **P** of the defect has a compact support in the Fourier-space. This support is — the Fermi-sphere.

An oscillatory interaction between defects could also be obtained by putting $P = A\mathbf{1}\delta_B$. We then have

$$(5.13) \qquad\qquad U = (M-N)\delta_B.$$

That this is oscillatory can be seen in the expression (3.14) or (3.15) for δ_B. However, this is not a correct model because it would yield zero values for $\mathbf{F}$ as well as U at all lattice points.

Since $\{g_{l_1+l_2}\}$ and $\{g_{l_1+l_2+L}\}$ are δ-sequences it can be seen that in the long wave limit U reduces to the result of the local theory:

$$(5.14) \qquad\qquad U = (M-N)\delta.$$

On the other hand, if L and λ are comparable but very large in comparison with l_1, l_2, we obtain the very simple result

$$(5.15) \qquad\qquad U = M\delta - Ng_L.$$

Acknowledgment

References

1. A. I. Akhiezer, W. B. Berestezki, Quanten-Elektrodynamik, Frankfurt/Main 1962.
2. N. N. Bogoliubov, O. S. Parasiuk, Acta Math., 97, 227, 1957.
3. N. Bohr, L. Rosenfeld, Danske Mat.-Fys. Meddr, 12, no. 8, 1933.
4. N. Bohr, L. Rosenfeld, Phys. Rev., 78, 794, 1950.
5. I. M. Gel'fand, G. E. Shilov, Generalized Functions, New York-London 1964–1968.
6. L. Hörmander, Linear Partial Differential Operators, Berlin-Göttingen-Heidelberg 1963.
7. L. W. Kantorowitsch, G. P. Akilow, Funktionalanalysis in Normierten Raumen, Berlin 1964.
8. V. G. Kosilova, I. A. Kunin, E. G. Sosnina, Fiz. Tverdogo Tela, 10, No. 2, 1953.
9. E. Kröner, Intern. Conf. on Fracture, 1, 1, 1965.
10. E. Kröner, B. K. Datta, Z. Phys., 196, 203, 1966.
11. J. A. Krumhansl, in Lattice Dynamics, Copenhagen 1963, Ed. R. F. Wallis, Pergamon Press, Oxford 1965.
12. I. A. Kunin, PMM 30, 642, 1966.
13. D. Rogula, Bull. Acad. Polon. Sci., Série Sci. Techn., 13, 7, 1965.
14. L. Schwartz, Théorie des Distributions, 2nd edition, Paris 1966.
15. K. Yosida, Functional Analysis, 3rd edition, Berlin-Heidelberg-New York 1971.

INSTITUT FÜR THEORETISCHE UND ANGEWANDTE PHYSIK
UNIVERSITÄT STUTTGART.