PHYSICAL REVIEW B VOLUME 24, NUMBER 12 15 DECEMBER 1981

Nonlinear lattice relaxation by a regional virial theorem

Jan-Åke Schweitz
Institute of Technology, Uppsala University, Box 534, S-751 21 Uppsala, Sweden

Jean-Louis Calais
Quantum Chemistry Group, Uppsala University, Box 518, S-715 20 Uppsala, Sweden

(Received 17 March 1981)

Methods and problems of electronic approaches to the treatment of lattice defect cores are briefly discussed, and the need of a new theoretical framework—less descriptive and more operative—is concluded. Two main objectives are formulated. (1) A new set of basic equations, providing relationships between practical physical parameters like force, torque, momentum, energy, etc., is needed. (2) These should be regional, real-space relationships, applicable to open subsystems like lattice defect cores. Regional virial and hypervirial relations are suggested as one possible set of equations of the desired type, and such relations are derived in a second quantization formalism, well suited for systems of undetermined numbers of electrons. This regional theory is developed within the Schrödinger picture by means of field operators, as well as within the Heisenberg picture in terms of double-time Green's functions (propagators). A connection with trace algebra is discussed, and a starting point for further developments is given. The Kanzaki approach to lattice relaxations is briefly reviewed and compared with the present approach, which is based on a regional virial equation, defined for an iterative series of artificial, intermediate equilibrium states in the relaxation process. This approach has three major advantages: (i) the forces are given a nonlinear representation, (ii) the fundamental equations are defined directly for the defect lattice, and (iii) the virial theorem will automatically be satisfied for the relaxed system.

I. INTRODUCTION

The plastic deformation of monocrystalline metals is mainly caused by formation of and interaction between lattice defects of various types. These effects are extensively treated within the theory of dislocations,[1,2] which supplies qualitative explanations to most features of plasticity. Dislocation theory, however, is largely of a continuum elastic nature and is not valid inside the core of a defect (the strongly distorted region close to the defect center). To obtain a deeper insight into the problems of plasticity in metals, it is necessary to establish a sound physical basis at an atomistic level for the creation and behavior of lattice defects.

A successful, nonempirical representation of crystal imperfections must be based on the electronic origin of cohesion. In point of principle, such an approach demands the solution of the Schrödinger equation for the atomic configuration under consideration. In the traditional solid-state approach,[3,4] the real-space problem is projected into $\vec{k}$ space, where the parts of the problem concerned with energy are explicitly solved. Only in rare cases are the real-space aspects of the problem considered in the form of explicit wave functions. In recent years the interest in surface phenomena and disordered systems have, however, contributed to a stronger emphasis on real-space methods.[5] For semiconductors and insulators the interest in real-space methods has been stronger, although the reciprocal space has been dominating there too.[4]

The formal and computational problems associated with a nonempirical approach are overwhelming, the Hamiltonian of a distorted lattice of realistic atoms being extremely complicated even if lattice vibrations are neglected. To make a solution possible, the problem must be simplified by a series of approximations. First, the many-electron Schrödinger equation must be reduced to a set of effective one-electron equations to be solved by some type of self-consistent field (SCF) scheme. In solid-state theory this is usually achieved by means of a density-functional procedure,[6] but there are also authors who prefer a Hartree-Fock treatment at this level of approximation.[7] Secondly, the picture of the crystal ions is often simplified beyond recognition in order to obtain a manageable description of the interaction between conduction electrons and lattice. Examples of common lattice approximations are uniform background, point-ion lattice, or simple pseudo-ion lattice models.[8,9] Thirdly, the interaction within the conduction gas is often given a simplified description by means of local, electron-statistical approximations.[6,10] Finally, the geometry of the defect must be simulated somehow; by an assumed scattering potential,[4] by iterative schemes of Kanzaki-force type,[3,4] or in some other way.

The consequenses of all these superposed approximations are grave. In spite of the advanced mathematical formalism used, and in spite of a heavy computational burden, the quantitative results (misfit energy, etc.) usually are no better than those obtained from the much simpler, empirical pair-potential approach.[11] Within the conventional set of approximations there does not seem to be a simple way out of this dilemma. Any attempt to improve the results by making the crystal model more realistic seems to lead to a virtually exponential rise in mathematical complexity and computational effort. The reason is the fact that the basic equations are too detailed in their description of a system: In order to obtain a quantitatively correct result for some parameter, we must feed the equations (via their Hamiltonians and boundary conditions) with very detailed information about the system. And even then, the desired result may only be acquired—at best—by an extensive computational effort. The apparently "unuseful" parts of the information are cumbersome to drag along in the calculations, and it may be tempting to reduce them by some kind of approximation. However, such a reduction may be risky; at a later stage it may turn out that a seemingly innocent approximation distorts the "useful" information. In short, the conventional approach is too much descriptive and too little operative.

A. Objective 1

We need a new set of basic equations which provide relationships between practical physical parameters like force, torque, momentum, energy, etc. Such a formalism should not require a too detailed description of the physical system in order to be operative. The hypervirial theorems[12-14] seem to provide one "safe" set of reduced relationships of the desired type, In these equations, parts of the unuseful information have been reduced, in a formally exact manner, by averaging. An attempt to use such relationships is presented in this paper.

Hypervirial relationships can be utilized for two different purposes.

(i) As an instrument to test the validity and the accuracy of a calculation.[15,16] For example, it is a well-known fact that the total energy of a system is fairly stable; good energy values can be obtained from rather approximate wave functions or electron distributions. The virial equation, which governs the kinetic-potential balance in the system is very sensitive to perturbations. Therefore, if the different parts of the energy can be determined by some other means, the virial theorem provides a most sensitive possibility to check these results.

(ii) In order to perform direct calculations of physical parameters. Examples of such parameters are the forces of constraint acting on the atoms in an unrelaxed vacancy configuration. If, for any fixed-vacancy geometry, it is possible to calculate good regional mean values of kinetic and potential energy and surface-flux virial, the nonlinear lattice relaxation can be determined from the regional virial equation. This is further developed in Sec. III. The above example points towards a second objective.

B. Objective 2

The set of reduced basic equations should be regional, real-space relationships, applicable to open subsystems like the cores of lattice defects. The traditional $\vec{k}$-space viewpoint that introduction of a lattice defect affects the whole bulk of the crystal (in the sense that the electron-energy levels of the whole crystal are shifted) cannot be directly transferred into real space. In reality, a lattice imperfection affects the distant parts of the crystal no more than a pole in the water outside Gothenburg affects the waves washing the shores of New York.[17] An imperfection in an extended crystal affects only a limited part of that crystal, and it should therefore be possible to handle it within a regional real-space formalism for open subsystems.

To combine the desiderata presented under the headings Objectives 1 and 2 is by no means trivial. The quantum theory of open systems is still in its early stages of development[18] and the program sketched, touches therefore, some very basic theoretical questions. Our own interest is primarily aimed at practical applications. Although all applications will require a number of approximations, we find it desirable to first develop a theoretical framework, in which we are at least aware of the main conceptual difficulties. As a first step in such a program we discuss, in Sec. II, certain aspects of regional hypervirial theorems. Two characteristics of the open system must then be kept in mind. First, there is no wave function or state vector for such a system; it must be described by a density matrix.[19-21] Second, since the subsystem is open, the number of particles in it is not fixed.

In order to illustrate how to use such abstract concepts, we sketch in Sec. IV, an application of a special hypervirial relation—the virial theorem for a subsystem—designed to give a scheme for the calculation of the lattice relaxation around a defect in a crystal.

Bader and collaborators[22-24] have made very important developments towards a quantum mechanics for a subsystem. By limiting themselves to a special kind of subsystems—so-called virial fragments which are limited by surfaces through which the flux of the gradient of the density vanishes—they have been able to contribute in a fundamental way towards one of the central questions of chemistry: what do we mean by an atom in a molecule? This type of consideration has important implications for solids in general and for defect problems in particular. In the present paper it is, however, necessary to work with a subsystem of a more general nature than the virial fragments.

II. REGIONAL HYPERVIRIAL RELATIONS

The term "hypervirial" seems to have been coined by Hirschfelder[12] in a paper dealing with both classical and quantum-mechanical hypervirial theorems. In classical mechanics, the long-time average of the Poisson bracket $\{H, Z\}$ between the Hamiltonian function H and a function Z of the generalized coordinates and momenta satisfying certain conditions, vanishes:

$$\{H,Z\}_{tav}=\left[\sum_{j=1}^{3N}\left(\frac{\partial Z}{\partial q_j}\frac{\partial H}{\partial p_j}-\frac{\partial Z}{\partial p_j}\frac{\partial H}{\partial q_j}\right)\right]_{tav}=0. \tag{1}$$

Hirschfelder proposes the name classical hypervirial theorem for (1) when Z is a homogeneous function of the generalized momenta of degree one, i.e.,

$$Z(q_1, q_2, \ldots, q_{3N}); kp_1, \ldots, kp_{3N} = kZ(q_1, \ldots, q_{3N}; p_1, \ldots p_{3N}) . \tag{2}$$

In conventional quantum mechanics the hypervirial relations associated with a quantity Z are nothing but the diagonal elements (in the energy representation) of Heisenberg's equation of motion for the quantity Z[12-14]

$$\langle \Psi_m, [H, Z]\Psi_m\rangle = 0. \tag{3}$$

An expression of this kind presupposes the existence of a set of wave functions for the pure states of a system with a well-defined number of electrons. The primary purpose of the papers in Refs. 12–16 was to use hypervirial relations as a means of optimizing approxmate wave functions. We also notice that (3) means that the integration is carried out over the whole space. The ordinary virial theorem is the special case of (1) and (3) when

$$Z=\sum_i \vec{r}_i \cdot \vec{p}_i . \tag{4}$$

As pointed out in the introduction, we are interested in relationships referring to partial regions in real space, i.e., to subsystems. These subsystems are allowed to exchange matter and momentum with other parts of the total system; they are so-called open systems. For such open systems there are no pure state wave functions. Instead they are characterized by a density matrix.[18-21]

Schweitz[25] has derived a classical virial theorem for open systems, which in addition to the ordinary terms, also contains a surface-flux virial associated with particles that leave and enter the open system. For the quantum-mechanical analog Schweitz[26] has given a derivation in terms of reduced density matrices. This constitutes one step towards a correct treatment of open systems, since the number of particles in the open *sub*system is not fixed. However, this treatment is incomplete, since it presupposes the existence of a state vector for a larger system to which the open subsystem belongs. The same objection can be made about the generalization to hypervirial theorems.[27] As a logical next step, we now develop a treatment, which ultimately, will eliminate the need for a state vector.

One characteristic of open systems, the undetermined number of particles, makes it attractive to work with second quantization. For our purposes this means that we should study expectation values of operators like

$$\int_\Omega \psi^\dagger(x) z(x) \psi(x) dx , \tag{5}$$

where $\psi^\dagger(x)$ and $\psi(x)$ are fermion creation and annihilation operators, and $z(x)$ is the one-electron operator under consideration. Ω is the volume of the region that is studied. The expression (5) gives an explicit *regional operator*. We need the expectation value of the operator

$$Z=\sum_j z(x_j) , \tag{6}$$

with respect to a state function $\Phi(t)$, i.e.,

$$\langle Z(t)\rangle_\Omega = \int_\Omega \langle \Phi(t) | \psi^\dagger(x) z(x) \psi(x) | \Phi(t)\rangle \, dx . \tag{7}$$

Assuming that the operator z is time independent, we obtain with the time-dependent Schrödinger equation

$$\frac{d}{dt}\langle Z(t)\rangle_\Omega = i \int_\Omega \langle \Phi(t) | [H, \psi^\dagger(x) z(x) \psi(x)] | \Phi(t)\rangle dx . \tag{8}$$

We write the Hamiltonian for the system as

$$H=\int \psi^\dagger(y)[-\tfrac{1}{2}\nabla^2+f(y)]\psi(y)dy + \tfrac{1}{2}\int \psi^\dagger(y)\psi^\dagger(z)g(y,z)\psi(z)\psi(y)dz\,dy \tag{9}$$

and obtain with the anticommutation relations satisfied by the field operators

$$[H, \psi^\dagger(x) z(x) \psi(x)] = \psi^\dagger(x)[e(x), z(x)]\psi(x) + S(x) , \tag{10}$$

where

$$e(x)=-\tfrac{1}{2}\nabla_x^2+f(x)+\bar{g}(x), \tag{11}$$

with

$$\bar{g}(x)=\int\psi^\dagger(y)g(x,y)\psi(y)dy \tag{12}$$

and

$$S(x)=\tfrac{1}{2}\psi^\dagger(x)\nabla_x^2[z(x),\psi(x)]$$
$$-\tfrac{1}{2}[\nabla_x^2,\psi^\dagger(x)]z(x)\psi(x). \tag{13}$$

The volume integral over $S(x)$ can be transformed to a surface integral which can be written as

$$\int_\Omega S(x)dx=i\oint_S[\vec{\mathcal{T}}_{xx'}z(x)\psi^\dagger(x')\psi(x)]\cdot d\vec{S}_x. \tag{14}$$

Here $\vec{\mathcal{T}}_{xx'}$ is the flux density operator

$$\vec{\mathcal{T}}_{xx'}=-\tfrac{1}{2}i(\vec{\nabla}_x-\vec{\nabla}_{x'}), \tag{15}$$

introduced by Schweitz.[26,27] Combining (8) with (10) and (14), and taking the long-time average of (8) we obtain

$$\left(i\int_\Omega\langle\Phi(t)|\psi^\dagger(x)[e(x),z(x)]\psi(x)|\Phi(t)\rangle dx-\oint_S\langle\Phi(t)|\vec{\mathcal{T}}_{xx'}z(x)\psi^\dagger(x')\psi(x)|\Phi(t)\rangle\cdot d\vec{S}_x\right)_{\text{tav}}=0. \tag{16}$$

This is one form of the hypervirial theorem associated with the operator $z(x)$ for an open system. The first- and second-order reduced density matrices associated with the state $\Phi(t)$ are

$$\gamma(x|x')=\langle\Phi(t)|\psi^\dagger(x')\psi(x)|\Phi(t)\rangle, \tag{17}$$
$$\Gamma(x_1,x_2|x_1',x_2')=\langle\Phi(t)|\psi^\dagger(x_1')\psi^\dagger(x_2')\psi(x_2)\psi(x_1)|\Phi(t)\rangle.$$

Combining (16) and (17) we retrieve the form of the hypervirial theorem given in Ref. 27.

Another modification is obtained by going over to the Heisenberg picture

$$\Phi(t)=\exp(-iHt)\Phi^H,$$
$$A(x,t)=\exp(iHt)A(x)\exp(-iHt), \tag{18}$$

and by expressing $\langle Z\rangle_\Omega$ in terms of the one-particle propagator[28]

$$G(x,t;x',t')=\begin{cases}-i\langle\Phi^H|\psi(x,t)\psi^\dagger(x',t')|\Phi^H\rangle & \text{for } t>t'\\ +i\langle\Phi^H|\psi^\dagger(x',t')\psi(x,t)|\Phi^H\rangle & \text{for } t'>t.\end{cases} \tag{19}$$

This gives

$$\frac{d}{dt}\langle Z(t)\rangle_\Omega=-i\int_{\substack{\Omega\\ t'\to t+0}}z_1\left(\frac{\partial}{\partial t}+\frac{\partial}{\partial t'}\right)G(x_1,t;x_1',t')dx_1, \tag{20}$$

where we use the convention that z_1 operates only on the unprimed variable x_1. A set of straightforward but tedious operations analogous to those leading to (16) then allow us to write the hypervirial theorem in the form

$$\left(\int_{\substack{\Omega\\ t'\to t+0}}[e_1(t),z_1]G(x_1,t;x_1',t')dx_1\right.$$
$$\left.+i\oint_{\substack{S\\ t'\to t+0}}[\vec{\mathcal{T}}_{11'}z_1G(x_1,t;x_1',t')]\cdot d\vec{S}_1\right)_{\text{tav}}=0. \tag{21}$$

The two forms, (16) and (21), of the hypervirial theorem offer interesting prospects for more realistic applications than was possible with the special form derived in Ref. 27.

The surface term in (21) can be rewritten as a sum of two volume integrals

$$Z_s=i\oint_{\substack{S\\ t'\to t+0}}[\vec{\mathcal{T}}_{11'}z_1G(x_1,t;x_1',t')]\cdot d\vec{S}_1$$
$$=\tfrac{1}{2}\int_{\substack{\Omega\\ t'\to t+0}}[\nabla_1^2,z_1]G(x_1,t;x_1',t')dx_1$$
$$+\tfrac{1}{2}\int_{\substack{\Omega\\ t'\to t+0}}z_1(\nabla_1^2-\nabla_{1'}^2)G(x_1,t;x_1',t')dx_1. \tag{22}$$

Since

$$\nabla_1^2-\nabla_{1'}^2=(\vec{\nabla}_1+\vec{\nabla}_{1'})\cdot(\vec{\nabla}_1-\vec{\nabla}_{1'})$$
$$=(\vec{\nabla}_1-\vec{\nabla}_{1'})\cdot(\vec{\nabla}_1+\vec{\nabla}_{1'}), \tag{23}$$

the last integral in (22) vanishes identically if

$$(\nabla_1\pm\nabla_{1'})G\equiv 0. \tag{24}$$

This happens (with the plus sign) for translationally invariant systems, for which

$$G(x_1,x_1')=G_0(x_1-x_1'). \tag{25}$$

The free-electron gas offers an example of (25). In the special case of the virial theorem ($z_1=\vec{x}_1\cdot\vec{p}_1$) the surface virial then reduces to

$$Z_s^{(v)}=-2\langle T\rangle_\Omega, \tag{26}$$

in complete agreement with the classical result for an ideal gas.[26]

The trace algebra of Löwdin[20] suggests still another form which might be suitable for a regional hypervirial theorem. In that formalism, the physical situation is described by a system operator Γ of unit trace. The expectation value of an operator Z is given by

$$\langle Z\rangle = \mathrm{Tr}(Z\Gamma)\,. \tag{27}$$

If $\Gamma^2 = \Gamma$ we have a pure state, i.e., there is a wave function. It is important to note, however, that the trace algebra formalism also covers the general situation $\Gamma^2 \neq \Gamma$ when there is no wave function. The time evolution of the system operator is governed by the Liouville equation

$$\frac{d\Gamma}{dt} = [M, \Gamma]\,, \tag{28}$$

where the operator M is related to the evolution operator S by

$$M = \frac{dS}{dt} S^{-1}\,. \tag{29}$$

Provided we can define a suitable regional version of (27), i.e.,

$$\langle Z\rangle_\Omega = \mathrm{Tr}_\Omega(Z\Gamma)\,, \tag{30}$$

which is consistent both with the trace algebra and with the statistics of the particles under consideration, we obtain with (28) (if Z is time independent),

$$\frac{d}{dt}\langle Z\rangle_\Omega = \langle [Z, M]\rangle_\Omega\,, \tag{31}$$

which can serve as a starting point for further developments.

From the derivation of the two main results (16) and (21) of this section, one might get the impression that we still need a state vector for the system under consideration. The same derivations can, however, be carried through to final results of the same form, with the expectation value (7) and the propagator (19) defined instead by means of a density operator.[29] We have thus obtained a regional hypervirial theorem expressed in terms of quantities that allow both a general density- (system-) operator description of the region under consideration and an undetermined number of particles. Further studies are needed in order to connect this form of the hypervirial theorem with the trace algebra formulation.

III. NEW APPROACH TO LATTICE RELAXATIONS

As mentioned in the introduction, virial relationships can be used not only as instruments to test the validity of given results, but also as a tool for direct calculations of physical parameters. Here, we suggest a method, according to which, a regional virial equation can be used to determine the geometrical relaxation close to a lattice defect. This method also provides an application of regional relationships to open systems with indefinite numbers of electrons.

The contribution from distortions caused by lattice defects is of importance in discussing various phenomena, e.g., interaction between defects, binding energy of defect clusters, x-ray diffraction, electrical resistivity, and deep energy levels in semiconductors. Experimentally, it is very difficult or even impossible to obtain detailed, quantitative information about these distortions. We must, therefore, resort to theoretical considerations in order to obtain such information. Basically, this issue is of a complicated, many-body character. Some 25 years ago, Kanzaki[30] suggested a simplified treatment of the lattice relaxations near point defects. His method, the so-called lattice-statics method, has been extensively applied, and is still one of the main methods for the study of relaxation phenomena.[3,4] Below, we give a short outline of the Kanzaki approach.

Kanzaki's method is based on the Born-Oppenheimer crystal approximation, and the fundamental equations are defined for a distorted, *perfect* lattice, i.e., all points are occupied by normal atoms, which, however, have been shifted from their true equilibrium positions into a simulated, relaxed point-defect geometry. The system is held in an artificial state of equilibrium by a set of external forces of constraint, which are assumed to be known. For a lattice of that type, Kanzaki expands the crystal energy E in a Taylor series in terms of the displacement vectors $\vec{\xi}_l$ (l refers to the numbering of atoms). This series is truncated after the second-degree terms, whereby the forces become linear functions of the displacement

$$\vec{\mathrm{K}}_m = \sum_l \vec{\xi}_l \cdot \underline{A}_{lm}\,. \tag{32}$$

m and l refer to the numbering of atoms, $\vec{\mathrm{K}}_m$ is the external force of constraint (Kanzaki force) acting on the mth atom, and the elements of the matrices $\underline{A}_{lm}$ are given by

$$A_{lm}^{\alpha\beta} = \left(\frac{\partial^2 E}{\partial x_l^{(\alpha)} \partial x_m^{(\beta)}}\right)_0\,, \quad \alpha, \beta = 1, 2, 3\,. \tag{33}$$

The subscript "0" indicates that the derivatives are to be taken at undisplaced lattice positions, i.e., $\underline{A}_{lm}$ are materials constants. Assuming $\vec{\mathrm{K}}_m$ and $\underline{A}_{lm}$ to be known, the corresponding equilibrium displacements can be determined by an iterative procedure. Since the force constants are characteristic of the perfect lattice, important simplifications can be achieved by means of a normal coordinate analysis. Here we suggest an alternative, nonlinear approach, based on the complete regional virial equation of the system. We adopt the Born-Oppenheimer picture, but do not linearize the theory by some approximation, and we do not study a distorted perfect lattice. Instead, we consider directly a distorted *defect* lattice. Thus, in the case of a monovacancy, one atom is missing at some lattice site.

We now proceed to set up the regional virial equation of the system. First of all, we note that all interactions among the electrons and the atomic nuclei are of the type r^{-1}. For the *relaxed true* equilibrium configuration, the regional virial equation can be expressed as

$$2\langle T_{\mathrm{rel}}\rangle_\Omega + \langle U_{\mathrm{rel}}\rangle_\Omega + Z_{s\,\mathrm{rel}} = 0\,. \tag{34}$$

The first term is twice the average kinetic-energy content of a volume Ω containing the defect, the second term is the corresponding average potential energy, and the last term is the average surface-flux virial, taken over the surface s enclosing Ω. The last term may be evaluated from Eq. (22), for instance, with $z_1 = \vec{x}_1 \cdot \vec{p}_1$. The evaluation of the quantities in (34) can actually be made in several different ways as illustrated in the previous section. In Eq. (34), as always in the following, time averages are tacitly understood. Next, we shall define a similar relationship for some *intermediate* stage of the relaxation process. It is important to remember that a virial theorem is valid only for a state of equilibrium. Therefore, for an intermediate stage, we create an artificial equilibrium state by introducing external forces of constraint ($\vec{F}_l$) to fix the atoms in their displaced positions ($\vec{r}_l$). The virial equation of this state must be completed with the contribution from the constraint forces to the virial sum

$$2\langle T\rangle_\Omega + \langle U\rangle_\Omega + Z_s + \sum_l \vec{r}_l \cdot \vec{F}_l = 0\,. \tag{35}$$

In the last term, the summation runs over all atoms contained in Ω. In the following we shall assume that the regional averages $\langle T\rangle_\Omega$, $\langle U\rangle_\Omega$, and Z_s can be determined for any given set of nuclear positions $\vec{r}_l$ by some independent method. Such a method may, for instance, be based on a density-functional or a Green's-function approach. Thus, we consider the first three terms of Eq. (35) to be known functions (explicit or implicit) of the nuclear positions $\vec{r}_l$. Consequently, the total energy content of Ω is also known:

$$\langle E\rangle_\Omega = \langle T\rangle_\Omega + \langle U\rangle_\Omega\,. \tag{36}$$

The forces $\vec{F}_l$, required to keep the crystal in an artificial state of equilibrium, must be equal but oppositely directed to the forces exerted by the crystal on the various atoms

$$\vec{F}_l = +\vec{\nabla}_l \langle E\rangle_\Omega\,, \tag{37}$$

where $\vec{\nabla}_l$ is the gradient with respect to the nuclear position $\vec{r}_l$. Hence, for any given set of $\vec{r}_l$, the forces $\vec{F}_l$ are known. Equation (37) is based on the assumption that the relaxation affects only the energy content of the observed volume Ω, so that the energy of the "outside region" is unaffected. This is true if Ω is chosen sufficiently large. It has been shown that the creation of a vacancy causes a uniform contraction of the whole crystal (including the outside region), but this contraction affects the energy to negligible if any extent.[31]

Combining Eqs. (35) and (36), the total energy can be expressed as

$$\langle E\rangle_\Omega = -\Big(\langle T\rangle_\Omega + Z_s + \sum_l \vec{r}_l \cdot \vec{F}_l\Big)\,. \tag{38}$$

Equation (38) holds for every state of equilibrium, artificial ($\vec{F}_l \neq 0$) as well as true ($\vec{F}_l = 0$). For both types of equilibria, the *total* force acting on every atom must vanish. In the artificial case, this is the result of an exact balance between nonvanishing external and internal components. The true equilibrium state is the special case when the external and internal force components vanish individually, i.e., the conditions

$$\vec{F}_m = \vec{\nabla}_m \langle E\rangle_\Omega = 0 \tag{39}$$

must be fulfilled for all atoms. Applying these conditions to Eq. (38), the following set of equations is obtained:

$$\vec{\nabla}_m(\langle T\rangle_\Omega + Z_s) + \sum_l \vec{\nabla}_m(\vec{r}_l \cdot \vec{F}_l) = 0\,. \tag{40}$$

Assuming rotation-free forces ($\vec{\nabla}_m \times \vec{F}_l \equiv 0$ for all m and l), the last sum of Eq. (40) can be expanded:

$$\sum_l \vec{\nabla}_m(\vec{r}_l \cdot \vec{F}_l) = \vec{F}_m + \sum_l (\vec{r}_l \cdot \vec{\nabla}_m)\vec{F}_l\,, \tag{41}$$

where

$$\vec{F}_m = \vec{\nabla}_m \langle E\rangle_\Omega = \vec{\nabla}_m(\langle T\rangle_\Omega + \langle U\rangle_\Omega)\,. \tag{42}$$

The last sum of Eq. (41) can be written as

$$\sum_l (\vec{r}_l \cdot \vec{\nabla}_m)\vec{F}_l = \sum_l \vec{r}_l \cdot \underline{B}_{lm}\,, \tag{43}$$

where the elements of the matrices are

$$B_{lm}^{\alpha\beta} = \frac{\partial^2 \langle E\rangle_\Omega}{\partial x_l^\alpha \partial x_m^\beta}\,, \quad \alpha, \beta = 1, 2, 3\,. \tag{44}$$

Inserting Eqs. (42) and (43) into Eq. (41) and then Eq. (41) into Eq. (40), we obtain

$$\vec{\nabla}_m(2\langle T\rangle_\Omega + \langle U\rangle_\Omega + Z_s) + \sum_l \vec{r}_l \cdot \underline{B}_{lm} = 0 \tag{45}$$

Equations (45) are the fundamental equations which can be utilized in an iterative scheme to determine the true equilibrium positions. The procedure is straightforward: for some initial set of nuclear positions $\vec{r}_l$ the gradient terms $\vec{\nabla}_m(\ldots)$ and the matrices B_{lm} are evaluated. Then a new set of $\vec{r}_l$

is determined from the solution of Eq. (45); new gradient terms and matrices are evaluated, and so on. The iteration goes on until the constraint forces $\vec{F}_m$ [Eq. (42)] vanish or are small enough.

The algebraic problem to solve in every iteration cycle is the linear system of Eqs. (45). It should be noted that, although the mathematical problem at hand is a linear one, the physical model is nonlinear: The forces $\vec{F}_m$ may be nonlinear functions of the displacements.

The convergence properties of the suggested iterative procedure remain to be investigated. The linearity of the mathematical problem leads us to believe that the convergence should be no worse than in the Kanzaki theory. In fact, the convergence may be even more rapid, since the force "constants" $\underline{B}_{lm}$ are redetermined in every cycle and therefore tend to "follow" the crystal in the relaxation process.

Assume that the volume Ω, which we have chosen to consider, contains N atoms. The task of solving the system Eqs. (45) then is, in principle, equivalent to the problem of finding the inverse of a $3N\times 3N$ matrix. Utilizing modern numerical methods and computer techniques, this problem can be adequately handled, even for fairly large matrices. Furthermore, for a given type of defect in a given lattice structure, the number of $(\vec{r}_l)$ unknowns in Eq. (45) can be considerably reduced by symmetry considerations, the size of the matrix being accordingly reduced.

Close to the defect the forces are most certainly nonlinear. Farther from the defect (but not very far), a linear approximation is probably acceptable, and we may replace the functional matrices $\underline{B}_{lm}$ in Eq. (45) with the constant matrices $\underline{A}_{lm}$ of Eq. (33), whereby the numerical problem is further simplified.

For a simple crystal model, such as a point-ion lattice immersed in a quasi-free-electron gas, the surface virial may be approximated according to Eq. (26) if Ω is chosen large enough:

$$Z_s = -2\langle T\rangle_\Omega \,. \tag{46}$$

Thus, in this approximation, the surface-flux virial is directly linked to the kinetic-energy content of the considered volume. Z_s is affected by the presence of a defect via the average formation of T over Ω. Elimination of Z_s in Eq. (45) results in a simplified set of fundamental equations

$$\vec{\nabla}_m\langle U\rangle_\Omega + \sum_l \vec{r}_l\cdot\underline{B}_{lm} = 0 \,. \tag{47}$$

Formally, Eq. (47) resembles the Kanzaki relations of Eq. (32). There are, however, several important differences: $\vec{\nabla}_m\langle U\rangle_\Omega$ does not represent the total external force of constraint; the matrix $\underline{B}_{lm}$ is not a materials constant; the parameters $\vec{r}_l$ of Eq. (47) are not displacement vectors; and the parameters $\langle U\rangle_\Omega$ and $\underline{B}_{lm}$ are not associated with a perfect lattice.

Obviously, the present approach [represented by Eqs. (45) or (47)] demands more detailed previous knowledge about the system than is required by the Kanzaki method [represented by Eq. (32)]. This is quite natural, since the present theory deals with a more complicated nonlinear problem. The suggested method has three major advantages over Kanzaki's method: (i) the forces are given a nonlinear representation, (ii) the fundamental equations are defined directly for the defect lattice, and (iii) the virial theorem will automatically be satisfied for the relaxed system. The price to be paid in return for these advantages is increased computational difficulties, mainly due to the fact that the elegant and effort-saving method of normal coordinates, applied by Kanzaki, cannot be utilized in the present theory.

IV. CONCLUSION

In his monograph on the theory of defects in solids Stoneham[4] states that, since the virial theorem (in its conventional form) is not valid for a subsystem, it is not particularly useful in defect theory. In other papers one of us[25-27] has tried to remedy that situation. In the present paper we have derived a regional hypervirial theorem by different procedures, which correspond to a more realistic description of a subsystem than those used in Ref. 27. Both the virial theorem and other special cases of the hypervirial theorem are, therefore, now open to applications in defect theory.

As a first application we have proposed an iteration scheme for calculating nonlinear lattice relaxation around a defect. This scheme is based on a regional virial theorem and it presupposes that we can calculate the kinetic- and potential-energy contents of the chosen region as well as the surface virial of the limiting surface. Although the final equations are linear, one definite advantage of the procedure is that the forces can be nonlinear functions of the displacements. The method can be expected to converge faster than the Kanzaki method.

On the basic theoretical side, very interesting problems present themselves, as sketched in Sec. II. A representation must be found which accounts both for an undetermined number of particles in the open system and for the fact that such a system is not described by a wave function. A combination of trace algebra and second quantization would seem to provide an interesting possibility.

Applications should be made at two levels. On one hand, hypervirial theorems for different operators should be exploited in order to formulate relationships between "practical" physical quantities. On the other hand, such relations, like the one developed on the basis of a regional virial theorem, should be used for actual quantitative calculations.

ACKNOWLEDGMENT

Jean-Louis Calais is financially supported by the Swedish Natural Research Council.

---

[1] F. R. N. Nabarro, *Theory of Crystal Dislocations* (Oxford University Press, Oxford, 1967).

[2] J. P. Hirth and J. Lothe, *Theory of Dislocations* (McGraw-Hill, New York, 1968).

[3] W. Jones and N. H. March, *Theoretical Solid State Physics* (Wiley-Interscience, London and New York, 1973), Vols. 1 and 2.

[4] A. M. Stoneham, *Theory of Defects in Solids* (Clarendon, Oxford, 1975).

[5] *Solid State Physics*, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1980), Vol. 35.

[6] See, e.g., N. D. Lang, Solid State Phys. 28, 225 (1973) and references therein.

[7] See the papers by T. C. Collins and F. E. Harris, in *Quantum Theory of Polymers*, edited by J-M André, J. Delhalle, and J. Ladik, NATO Advanced Study Institute, Series C, No. 39 (Reidel, Dordrecht and Boston, 1978).

[8] N. H. March and J. S. Rousseau, Cryst. Lattice Defects 2, 1 (1971).

[9] N. H. March, J. Phys. F 3, 233 (1973).

[10] J. C. Slater, *Quantum Theory of Molecules and Solids* (McGraw-Hill, New York, 1974), Vol. IV.

[11] R. A. Johnson, J. Phys. F 3, 295 (1973).

[12] J. O. Hirschfelder, J. Chem. Phys. 33, 1462 (1960).

[13] J. O. Hirschfelder and C. A. Coulson, J. Chem. Phys. 36, 941 (1962).

[14] C. A. Coulson, Q. J. Math. (Oxford) 16, 279 (1965).

[15] S. T. Epstein and J. O. Hirschfelder, Phys. Rev. 123, 1495 (1961).

[16] J. C. Y. Chen, J. Chem. Phys. 40, 615 (1964).

[17] J.-L. Calais, Int. J. Quantum Chem. 18, 607 (1980).

[18] E. B. Davies, *Quantum Theory of Open Systems* (Academic, New York, 1976).

[19] U. Fano, Rev. Mod. Phys. 29, 74 (1957).

[20] P-O Löwdin, Int. J. Quantum Chem. 12, Suppl. 1, 197 (1977).

[21] F. J. Belinfante, Int. J. Quantum Chem. 17, 1 (1980).

[22] S. Srebrenik, R. F. W. Bader, and T. T. Nguyen-Dang, J. Chem. Phys. 68, 3667 (1978).

[23] R. F. W. Bader, S. Srebrenik, and T. T. Nguyen-Dang, J. Chem. Phys. 68, 3680 (1978).

[24] R. F. W. Bader and T. T. Nguyen-Dang, Adv. Quantum Chem. (in press).

[25] J.-Å. Schweitz, J. Phys. A 10, 507 (1977).

[26] J.-Å. Schweitz, J. Phys. A 10, 517 (1977).

[27] J.-Å. Schweitz, Int. J. Quantum Chem. 18, 811 (1980).

[28] S. Raimes, *Many-Electron Theory* (North-Holland, Amsterdam and London, 1972).

[29] J. Linderberg and Y. Öhrn, *Propagators in Quantum Chemistry* (Academic, New York and London, 1973).

[30] H. Kanzaki, J. Phys. Chem. Solids 2, 24 (1957).

[31] J-Å Schweitz and O. Vingsbo, J. Phys. F 7, 1 (1977).