Local volume effects in the generalized pseudopotential theory

Guy C. G. Skinner and Anthony T. Paxton

Department of Physics, King’s College London, Strand, London WC2R 2LS, United Kingdom

John A. Moriarty

Lawrence Livermore National Laboratory, Livermore, California 94551-0808, USA

(Received 12 October 2018; published 19 June 2019)

The generalized pseudopotential theory (GPT) is a powerful method for deriving real-space transferable interatomic potentials. Using a coarse-grained electronic structure, one can explicitly calculate the pair ion-ion and multi-ion interactions in simple and transition metals. While successful in determining bulk properties, in central force metals the GPT fails to describe crystal defects for which there is a significant local volume change. A previous paper [J. A. Moriarty and R. Phillips, Phys. Rev. Lett. 66, 3036 (1991)] found that by allowing the GPT total energy to depend upon some spatially averaged local electron density, the energetics of vacancies and surfaces could be calculated within experimental ranges. In this paper, we develop the formalism further by explicitly calculating the forces and stress tensor associated with this total energy. We call this scheme the adaptive GPT (aGPT) and it is capable of both molecular dynamics (MD) and molecular statics. We apply the aGPT to vacancy formation, divacancy binding, and stacking faults in hcp Mg. We also calculate the local electron density corrections to the bulk elastic constants and phonon dispersion for which there is refinement over the baseline GPT treatment. In addition, we demonstrate aGPT-MD simulation through the calculation of thermal expansion in magnesium to 700 K.

DOI: 10.1103/PhysRevB.99.214107

1. INTRODUCTION

Generalized pseudopotential theory (GPT) is a first-principles framework for deriving real-space interatomic potentials in metals and alloys from density-functional quantum mechanics [1,2]. In a basic plane-wave basis, the GPT provides an updated and refined version of second-order pseudopotential perturbation theory, with linear screening and nonlocal, energy-dependent pseudopotentials, that can be applied to $sp$-valent, nearly-free-electron (NFE) simple metals. More generally, in a mixed basis of plane waves and localized atomic $d$ states, the GPT additionally captures both tight-binding (TB) $d$-state interactions and $sp$-$d$ hybridization between the broad NFE $sp$ bands and the narrow TB $d$ bands. The practical challenges of the GPT for pure transition metals have also led to the development of a simplified model GPT (MGPT) [3] which allows for large-scale atomistic simulations in these materials. The GPT and MGPT have been successfully applied to transition-series metals with empty, filled, and partially filled $d$ bands [1–4] to transition-metal alloys [2,5] and, with localized $f$ states in place of the $d$ states, to actinide metals as well [2,4].

For bulk elemental metals, the GPT total energy $E_{\text{tot}}$ is developed in a volume-dependent many-body cluster expansion [1,2], which in its simplest form is truncated at pairwise interactions,

$$
E_{\text{tot}}(\mathbf{R}, \Omega) = NE_{\text{vol}}(\Omega) + \frac{1}{2} \sum_{i,j} v_2(R_{ij}, \Omega),
$$

where $\Omega$ is the atomic volume and the prime on the double summation over ion positions $i$ and $j$ excludes the $i = j$ term. The large volume term $E_{\text{vol}}$ is independent of the positions of the ions, and accounts for most of the equilibrium cohesive energy of the metal, as illustrated in Fig. 1 for Mg. The functional form of the smaller pair potential $v_2$ is also independent of atomic structure, and $v_2(R_{ij}, \Omega)$ accounts for structural energy differences between different configurations of the ions at volume $\Omega$ through its explicit dependence on the ion-ion separation distance $R_{ij} = |\mathbf{R}_j - \mathbf{R}_i|$. The GPT total-energy functional given by Eq. (1) well describes the bulk properties of simple metals (e.g., Mg, Al), pretransition metals with nearly empty $d$ bands (e.g., Ca), late transition metals with nearly filled $d$ bands (e.g., Ni, Cu), and posttransition metals with completely filled $d$ bands (e.g., Zn). For the remaining central transition metals, it is necessary to extend the total-energy expansion in Eq. (1) to include angular-dependent three- and four-ion potentials, which are established, respectively, by the third- and fourth-order moments of the $d$-band density of electronic states. Computationally, the evaluation of the GPT total energy for all metals scales linearly with the number of atoms and is thus an order-$N$ process. For the non- and late-transition elements covered by Eq. (1), however, there is an additional computational overhead relative to short-ranged central-force empirical potentials as a result of the long-ranged screening oscillations in the GPT pair potential $v_2$. Even so, this is not a significant barrier in most applications today, and using modern high-performance computers, large-scale GPT atomistic simulations involving millions of atoms can be routinely performed [2,4].

*tony.paxton@kcl.ac.uk
The structure-independent nature of the pair and multi-ion potentials in the GPT ensures that these potentials are transferable to all ion configurations of the bulk metal, either ordered or disordered. This includes all structural phases of both the solid and the liquid, as well as the deformed solid ordered or disordered. This includes all structural phases of the total energy to an equivalent local electron-density representation, such that Eq. (1) becomes

$$E_{\text{tot}}([\mathbf{R}], n_{\text{val}}) = \sum_i E_{\text{vol}}(\bar{n}_i) + \frac{1}{2} \sum_{i,j} v_2(R_{ij}, \bar{n}_{ij}),$$

(2)

where $\bar{n}_i$ is a simple functional of the average value of the valence electron density $n_{\text{val}}$ on the site $i$, and $\bar{n}_{ij}$ is the arithmetic average $(\bar{n}_i + \bar{n}_j)/2$. For central transition metals, there are corresponding three- and four-ion potential contributions on the right-hand side of Eq. (2). In the perfect crystal with equivalent ion positions, Eq. (2) is an exact transformation and only a redefinition of variables, with all quantities still determined from first principles. The step forward comes in then, as an ansatz, applying Eq. (2) to all ion configurations, including free surfaces and bulk defects. In doing so, one notes from Fig. 1 that, qualitatively, the missing positive formation energy for surfaces and vacancies is indeed now supplied by the volume term, because $\bar{n}_i$ is lower near a surface or vacancy site than at a bulk ion site. Moriarty and Phillips went on to show that good unrelaxed surface energies and vacancy formation energies could thereby be obtained for both the late transition metal Cu and for the central transition metal Mo. In the case of Cu, the local-density corrections were found to be very large, averaging about 70% for both the surface energies and for the vacancy formation energy. In the case of Mo, on the other hand, the corrections were found to be significantly smaller, 30–40% for the surface energies and only 5% for the vacancy formation energy. The physical reason for the latter behavior is that in transition metals the essential local character needed in the total energy is already present to a large degree in the global-volume representation through the $d$ bonding contributions to $E_{\text{tot}}$ provided by the localized $d$-state moments. Thus for central transition metals, one expects that bulk defect energies will be well calculated by either the global-volume or the local-density formulations of the total energy.

In the present paper, we take an additional major step and develop the local-density representation of GPT into a robust general method we now call the adaptive GPT or aGPT, which includes not only energies but the forces and stresses needed for atomistic simulation and a much wider treatment of materials properties. The formalism of the aGPT is elaborated in Sec. II, including the averaging required in Eq. (2). For simplicity, this discussion is done in the context of a well-studied simple-metal Mg prototype (see Sec. IB), but the results can be immediately applied to the empty, almost filled, and filled $d$-band metals covered by Eqs. (1) and (2). The averaging is not unique, but it can be optimized, and we have developed a good way to do this that makes calculated properties quite insensitive to the parameters defining the averaging, while at the same time allowing the calculation of smooth derivatives of the averaged quantities. In Sec. III, we discuss the evaluation of the corresponding aGPT forces and stresses, and test the results with calculations of phonons and elastic constants. In this regard, an earlier, simplified form of Eq. (2) was used by Rosenfeld and Stott [8] to resolve the well-known bulk compressibility problem in pseudopotential perturbation theory, as we further discuss below in Sec. IA, and as we use here as an additional fundamental test for the aGPT elastic moduli. Finally, in Sec. IV we apply the aGPT to the calculation of relaxed single vacancies and divacancies as well as to stacking fault energies and thermal expansion in hcp Mg.

A. The bulk compressibility problem

There are two fundamental ways to calculate the bulk modulus of a single crystal using the interatomic pair potentials derived from second-order pseudopotential perturbation theory, or more generally from the GPT. The first method
involves taking the explicit second volume derivative of the total energy given by Eq. (1). This procedure corresponds to a homogenous deformation of the primitive cell of the lattice and produces the so-called static bulk modulus \( B_s \).

The second method calculates the bulk modulus using the long-wavelength (low-\( q \)) behavior of the dynamical matrix, which determines the elastic constants of the material [9]. This produces the so-called dynamic bulk modulus \( B_d \). These two methods are known to disagree over the value of the bulk modulus produced. This discrepancy can be seen immediately to be the result of the absence of explicit volume derivatives in the dynamical matrix. In conventional pseudopotential perturbation theory, the discrepancy is only resolved at fourth order [10], albeit in a computationally challenging and non-transparent manner. It was later shown [8], by allowing the total energy to depend on local electron density as in Eq. (2), that the requisite volume derivatives arise to correct the bulk modulus calculated from the dynamical matrix. In the present context, one can use the accurate value of \( B_s \) calculated from Eq. (1) to test the value of \( B_d \) calculated with the aGPT from Eq. (2).

B. Magnesium prototype and baseline vacancy formation energy

Magnesium is an important lightweight metal whose bulk properties are very well described by the GPT via Eq. (1), making it an excellent prototype material for developing the aGPT. The first-principles pair potentials \( v_2 \) and volume term \( E_{\text{vol}} \) for this metal have been calculated over a wide volume range in connection with detailed studies of the temperature-pressure phase diagram and thermodynamic properties of Mg in the mid 1990s [11], and in subsequent studies of thermoelectricity [12]. The volume term is that displayed in Fig. 1, and the Mg pair potentials used in this paper are the same as those used in Refs. [11,12] except for an improved smooth long-ranged cutoff function discussed in Sec. II. As can be appreciated from Fig. 1, good elementary cohesive properties are predicted, including the cohesive energy, hcp lattice constant, and static bulk modulus. The latter has a value \( B_s = 35.8 \) GPa at the observed equilibrium volume in good agreement with the measured experimental value of 35.2 GPa [13]. The calculated hcp phonon spectrum is in excellent agreement with experiment, as are the high-temperature values of the volume term.

The local volume change associated with a crystal defect occurs to a local change in the valence electron density. We briefly review the treatment of the electron density in the GPT as applied to \( sp \)-valent simple metals. The valence electron density consists of a uniform electron density \( n_{\text{unif}} = Z/\Omega \) (where \( Z \) is the valence) plus small oscillatory and charge-neutral screening and orthogonalization-hole components [1,2]:

\[
n_{\text{val}}(r) = n_{\text{unif}} + \delta n_{\text{scr}}(r) + \delta n_{\text{oh}}(r).
\]

The screening electron density \( \delta n_{\text{scr}} \) arises from first-order pseudopotential perturbation theory which for a simple metal has the form [2]

\[
\delta n_{\text{scr}}(r) = \sum_q S(q) n_{\text{scr}}(q) e^{i q \cdot r},
\]

where \( S(q) = N^{-1} \sum \exp(-i q \cdot R_i) \) is the structure factor and [2]

\[
n_{\text{scr}}(q) = -\left( \bar{\omega}_{\text{ion}}(q) + \frac{4 \pi e^2}{q^2} (1 - G(q)) n_{\text{oh}}(q) \right) \frac{\Pi_0(q)}{\epsilon(q)},
\]

where \( \bar{\omega}_{\text{ion}} \) is a well-defined average value of the ionic pseudopotential over the free-electron Fermi sphere, \( G \) is the exchange-correlation functional, \( \epsilon \) is the dielectric function of the interacting electron gas, and \( \Pi_0 \) is the electron gas polarizability in the Hartree or random phase approximation. Each of these quantities can be directly evaluated in terms of input pseudopotential and electron gas quantities.

The orthogonalization-hole component arises from the difference between the valence electron density constructed from the one-particle pseudo-wave-functions and the valence electron density constructed from the “true” one-particle wave functions. For the nonlocal, energy-dependent Austin-Heine-Sham (AHS) pseudopotential [14] used in the GPT, there exists an exact transformation between the one-particle pseudo- and true wave functions which can be exploited to obtain the exact orthogonalization-hole density. The orthogonalization-hole contribution to \( n_{\text{val}} \) in Eq. (4) has the form [2]

\[
\delta n_{\text{oh}}(r) = \left( \frac{Z^*}{Z} - 1 \right) n_{\text{unif}} + \sum_i n_{\text{oh}}(r - R_i),
\]

where \( Z^* \) is an effective valence occupation (\( Z \geq Z^* \)) and \( n_{\text{oh}} \) is a localized hole density. For a simple metal, \( n_{\text{oh}} \) is confined to the inner-core region of the site \( i \), but both \( Z^* \) and
$n_{oh}$ depend on the properties of the pseudopotential. For the nonlocal, energy-dependent AHS pseudopotential used in the GPT, we have [2]

$$Z^* = Z + \frac{2\Omega}{(2\pi)^3} \int dk \langle \rho_c(k) \Theta^*(k - k_f) \rangle$$

and

$$n_{oh}(r) = \frac{2\Omega}{(2\pi)^3} \int dk \left[ \langle \rho_c(k) | \rho_c(k) \rangle - \langle \rho_c(k) | \rho_c(r) \rangle \right] - \langle \rho_c(k) | \Theta^*(k_f - k) \rangle,$$

where $\Theta^*$ is a Heaviside step function that ensures that the integral is over just the free-electron Fermi sphere and $\rho_c$ is the inner-core projection operator:

$$\rho_c = \sum_i |\phi_i\rangle \langle \phi_i|.$$

The valence electron density $n_{val}$ can equivalently be written as a superposition of self-consistently screened pseudatom densities $n_{pa}$:

$$n_{val}(r) = \sum_i n_{pa}(r - R_i).$$

The precise form of the pseudatom density $n_{pa}$ can readily be derived from Eqs. (4), (5), and (7). We do this by inserting the full form of the structure factor $S(q)$ into Eq. (5) and then adding the $q = 0$ term to the summation over $q$ to account for the net uniform density $Z n_{uni}/2$. Finally, we convert the summation over $q$ to an integral and infer that the single-site pseudatom density is given by

$$n_{pa}(r, \Omega) = \frac{\Omega}{(2\pi)^3} \int dq n_{scr}(q) e^{iq \cdot r} + n_{oh}(r).$$

The calculated GPT radial pseudatom density $u_{pa}(r) = 4\pi r^2 n_{pa}(r)$ for Mg at the experimental room-temperature atomic volume $\Omega = 156.8$ a.u. is shown in Fig. 2 and compared with the corresponding free-atom density for the valence $3s$ and $3p$ electrons. In this calculation, and all those subsequent, the exchange-correlation functional $G(q)$ is taken to be the analytic expression developed by Ichimaru and Utsumi [15] referenced to the exchange-correlation energy of Vosko et al. [16]. While the discussion in this section has been limited to $sp$-valent simple metals, the extension to empty, filled, and partially filled $d$-band metals covered by Eqs. (1) and (2) does not alter the subsequent discussion.

**B. Implementing the aGPT**

To connect the GPT valence electron density with the aGPT total energy in Eq. (2), we spatially average the GPT valence electron density $n_{val}$ about the site $i$ using an arbitrary normalized distribution function $f_w$. For a bulk crystal with equivalent ion sites, the spatially averaged electron density $\bar{n}_i$ about a site $i$ is constrained to be the uniform valence electron density $n_{uni}$. Combining the two equivalent valence electron density formulations in Eqs. (4) and (11) yields the bulk constraining equation:

$$\bar{n}_i \equiv n_{uni} = \sum_j \bar{n}_{pa}(R_{ij}, \Omega) - \delta \bar{n}_{oh} - \delta \bar{n}_{scr}.$$

Here the bar over the densities refers to an averaging with respect to some distribution function $f_w$, i.e.,

$$\bar{n}_{pa}(R_{ij}, \Omega) = \int dr f_w(r - R_i) n_{pa}(r - R_j, \Omega).$$

with both $\delta \bar{n}_{scr}$ and $\delta \bar{n}_{oh}$ having similar forms. Typically, this averaging smooths out the long-range screening oscillations. As a result of the bulk constraining equation, the aGPT preserves the bulk total energy for any given crystal structure with equivalent ion sites.

The first step toward developing a practical aGPT scheme for describing defects or surfaces is to make the approximation that

$$\bar{n}_i = \sum_j \bar{n}_{pa}(R_{ij}, \Omega) - \delta \bar{n}_{oh} - \delta \bar{n}_{scr}$$

can be applied generally. Furthermore, the spatially averaged local electron density $\bar{n}_i$ can be broken down into an effective on-site contribution $\bar{n}'_i = \bar{n}_{pa}(R_{ii}, \Omega) - \delta \bar{n}_{oh} - \delta \bar{n}_{scr}$ and an off-site or background component $\bar{n}_{b}'$, where

$$\bar{n}'_i = \sum_{j \neq i} \bar{n}_{pa}(R_{ij}, \Omega).$$

We make an additional assumption that the on-site density is constant $\bar{n}'_i \equiv \bar{n}_a$ and as a result only the background density $\bar{n}_{b}'$ is site dependent. Under these assumptions, we may calculate the on-site density $\bar{n}_a$ using the bulk constraining equation in Eq. (13). In practice, this amounts to first calculating $\bar{n}_a$ for an ideal bulk crystal prior to calculating the total energy for the surface or defective crystal. For certain $d$-band metals, e.g., Cu, there may be $s$-$d$ transfer between the surface and the bulk [7]. In which case, all of the densities $\bar{n}_i$, $\bar{n}'_i$, and $\bar{n}_{b}'$ must
be scaled by a factor $Z_i/Z$ to account for this, where $Z_i$ is an effective $sp$ occupation on the site $i$. This quantity would have to be determined self-consistently.

The next step toward a practical aGPT implementation is to specify the form of the distribution function $f_w$ in Eq. (14). We choose $f_w$ to correspond to a sigmoid function

$$f_w(r) = \begin{cases} N^{-1} & r < R_a \\ N^{-1}(1 + \alpha((r/R_c) - 1)^3)e^{-\alpha((r/R_c) - 1)^2} & r \geq R_a, \end{cases}$$

which is the sigmoid function that is typically used in the GPT to truncate the pairwise interaction [17] albeit with a different value of the Gaussian width $\alpha$. For large values of $\alpha$, this corresponds to an average over a sphere of radius $R_a$. The normalization $N$ of the distribution function $f_w$ is given by

$$N = V_w + \frac{8\pi}{\alpha}R_a^3 + \frac{5\pi^{3/2}}{2\alpha^{3/2}}R_a^3 + \frac{3\pi^{3/2}}{\sqrt{\alpha}}R_a^3,$$

which in the limit $\alpha \to \infty$ is the volume of a sphere of radius $R_a$. The two parameters $\alpha$ and $R_a$ represent the only parameters in this form of the aGPT. The Gaussian width $\alpha$ is chosen such that the radial derivatives of the spatially averaged pseudoatom density are smooth. If the radial derivatives were not smooth, then there would be an unphysically large change in the forces as the interatomic separation changes from less than $R_a$ to greater than $R_a$ and vice versa. In the rest of this paper, we choose $\alpha = 25$ which produces a spatially averaged pseudoatom density with smooth derivatives over a wide range of averaging sphere radii. We have a certain amount of freedom in choosing a value for $R_a$ since physical properties of interest do not seem strongly dependent on $R_a$. We choose the optimum of $R_a$ to be that which reproduces the GPT volume-conserving elastic constants most closely. These issues will be discussed further in Secs. III and IV. While other normalized distribution functions have been trialed, none represented an improvement on the sigmoid function.

The resulting spatially averaged pseudoatom density is shown in Fig. 3. For values of $R_a$ in the range $R_a/R_WS \in [1, 2]$, where $R_WS = (3\Omega/4\pi)^{1/3}$ is the Wigner-Seitz radius, the spatially averaged pseudoatom density looks like a Gaussian. A function of this type was employed in the empirical approach taken previously [8,18]. For larger values of $R_a$ in the range $R_a/R_WS \in [3, 4]$, the resulting spatially averaged pseudoatom density is almost flat over the first two neighbor shells.

The spatially averaged pseudoatom density is smoothly truncated to ensure force continuity during molecular dynamics. If we denote $R_0$ and $R_c$ as the cutoff onset and final termination, respectively, then our approach is to replace $\bar{n}_{pa}$ by a polynomial whose value and derivatives exactly match $\bar{n}_{pa}$ at $R_0$ and whose derivatives are precisely zero at $R_c$. This polynomial can be found using Hermite interpolation [19] which finds an $(nm-1)$ interpolating polynomial given knowledge of the function and $m-1$ derivatives at $n$ points. For our purposes, we choose $m = 3$, $n = 2$ and $R_c - R_0 = 0.5R_WS$.

![FIG. 3. The spatially averaged pseudoatom density $\bar{n}_{pa}$ (red) is calculated for Mg at the experimental room-temperature atomic volume $\Omega = 156.8$ a.u. The radial derivatives (green and blue) of the spatially averaged pseudoatom density were calculated using Lagrange interpolation polynomials. All of these quantities are calculated at $\alpha = 25$ and for two values of the averaging sphere radius $R_a = 1.8R_WS$ in (a) and $R_a = 3.4R_WS$ in (b).](image-url)

### III. Total Energy Derivatives

#### A. Forces and force constants

The force $F_{ia}$ on the atom $i$ describes how the total energy changes with respect to an infinitesimal shift in its position $R_{ia}$. As the ion-ion potential is self-consistently screened, we can ignore any change in electron screening [20]:

$$F_{ia} = -\frac{\partial E_{tot}}{\partial R_{ia}}(\{R\}, \Omega).$$

The force in the GPT involves only radial derivatives of the screened ion-ion interaction whereas the aGPT force will involve contributions from density derivatives of both $E_{vol}$ and $v_2$. It is instructive to decompose the force into three parts,

$$F_{ia} = F^{[I]}_{ia} + F^{[II]}_{ia} + F^{[III]}_{ia},$$

where the second term is the force due to the radial derivatives of $v_2$, the first and third components are the force due to the density derivatives of $E_{vol}$ and $v_2$, respectively. The first component can be written

$$F^{[I]}_{ia} = -\frac{\partial E_{vol}}{\partial \bar{n}_{ia}}\frac{\partial \bar{n}_{ia}}{\partial R_{ia}} - \sum_{j\neq i} \frac{\partial E_{vol}}{\partial \bar{n}_{ij}} \frac{\partial \bar{n}_{ij}}{\partial R_{ia}},$$

where $\partial E_{vol}/\partial \bar{n}_{ia}$ is shorthand for the density derivative evaluated at $\bar{n}_i$. We can write the derivatives of the spatially averaged local electron density, noting that the on-site density $\bar{n}_i$ does not contribute, as

$$\frac{\partial \bar{n}_i}{\partial R_{ia}} = \sum_{j\neq i} \frac{\partial \bar{n}_{pa}}{\partial R_{ij}} \frac{R_{ia}}{R_{ij}}$$

and

$$\frac{\partial \bar{n}_{ij}}{\partial R_{ia}} = \frac{\partial \bar{n}_{pa}}{\partial R_{ij}} \frac{R_{ia}}{R_{ij}}.$$
where $R_{ji}$ is the $\alpha$ component of the difference between position vectors $\mathbf{R}_i - \mathbf{R}_j$, and $R_{ji}/R_{ij}$ are the direction cosines. Equation (21) can be in more symmetric form:

$$F_{i\alpha}^{[I]} = \sum_{j \neq i} \left( \frac{\partial E_{\text{vol}}}{\partial R_{ij}} + \frac{\partial E_{\text{vol}}}{\partial R_{ji}} \right) \frac{\delta R_{ja}}{\delta R_{ij}} R_{ji}. \quad (24)$$

The second component of the force looks similar to the GPT force. However, it is only equal to the GPT force in the bulk. It is given by

$$F_{i\alpha}^{[II]} = \sum_{j \neq i} \frac{\partial v_2}{\partial R_{ij}} (R_{ij}, \bar{n}_{ij}) \frac{R_{ji}}{R_{ij}}. \quad (25)$$

The final component, which contains an additional neighbor sum, is given by

$$F_{i\alpha}^{[III]} = \frac{1}{2} \sum_{j \neq i} \frac{\partial v_2}{\partial R_{ij}} \left( \frac{\delta R_{ja}}{\delta R_{i\alpha}} + \frac{\delta R_{ja}}{\delta R_{j\alpha}} \right) + \frac{1}{4} \sum_{j \neq i} \sum_{k \neq j} \frac{\partial v_2}{\partial n_{jk}} \left( \frac{\delta n_{ja}}{\delta R_{i\alpha}} + \frac{\delta n_{ja}}{\delta R_{j\alpha}} \right). \quad (26)$$

The bulk force constant matrix $A_{j\alpha\beta}$ will largely be the same as for the GPT. However, there will be small contributions from the density derivatives of $E_{\text{vol}}$ and $v_2$. These additional contributions require further neighbor summations. These third- and fourth-order terms can be necessary to capture the phonon dispersion at certain $\mathbf{q}$-points in the Brillouin zone, in particular for Be [21]. Despite this, the phonon dispersion will be dominated by the bulk GPT force constant matrix. However, deviations in the band structure in the low $\mathbf{q}$ limit are expected and correspond to changes in the elastic constants. The phonon dispersion was calculated for mechanically unstable bcc Mg at the equilibrium atomic volume in Fig. 4. This crystal structure was chosen as a representative example due to the presence of the imaginary frequencies along the $\mathbf{q}$-point path from $\Gamma$ to $N$. There is also scientific interest in this particular phase. When Mg is alloyed with Li, the bcc phase is stabilized and the alloy becomes ductile. In addition, the phonon dispersion relation for thermodynamically stable hcp Mg at the equilibrium atomic volume and $c/a$ ratio is shown in Fig. 5. The dispersion relations were calculated numerically with the code ALAMODE [22], using supercells that were extended by $6 \times 6 \times 6$ (for aGPT/GPT) and $3 \times 3 \times 3$ (for DFT). The density functional theory (DFT) results were calculated using the FP-LMTO method [23] with the same lattice parameters as the aGPT/GPT. The Brillouin zone integrations were performed with Methfessel-Paxton sampling [24] and $30 \times 30 \times 30 \mathbf{q}$-point subdivisions. The exchange-correlation functional was taken in the local-density approximation using the correlation function of Perdew and Wang [25]. The local density approximation was used since it is closest to the treatment of exchange and correlation within the GPT. There is good agreement with the DFT data and the aGPT/GPT.

### B. Stress tensor

Molecular dynamics simulations that sample an isobaric ensemble require a barostat to match the external pressure to the internal pressure [26] $P^{\text{int}} = -\sum_{\alpha} F_{\alpha\alpha}/d$ where $\sigma$ is the internal stress tensor and $d$ is the dimension of the cell. This matching is effectively the equilibrium condition, i.e., the time average of the internal pressure is the external pressure. In such simulations, only the lattice parameter $a$ is dynamic. This constraint is slightly artificial if the crystal has multiple lattice parameters as in the case of hexagonal crystals. Relaxing this constraint requires that we now sample an isostress ensemble [27] where the internal stress tensor is matched to an external stress tensor. The stress tensor is defined as the infinitesimal change in total energy as a result of an infinitesimal strain [28]

$$\sigma_{\alpha\beta} = \frac{1}{V} \left( \frac{\partial E_{\text{tot}}}{\partial \varepsilon_{\alpha\beta}} \right)_{\varepsilon_{\alpha\beta}=0}, \quad (27)$$
LOCAL VOLUME EFFECTS IN THE GENERALIZED …

PHYSICAL REVIEW B 99, 214107 (2019)

where the prefactor of inverse volume \( V^{-1} \) is required by dimensional analysis. The application of a strain changes the lattice vectors \( h \) in the following way:

\[
\delta h_{\alpha\beta} \rightarrow \tilde{h}_{\alpha\beta} = \sum_\gamma (\delta_{\alpha\gamma} + \varepsilon_{\alpha\gamma}) h_{\gamma\beta}, \tag{28}
\]

where \( \delta \) is the Kronecker delta. Since the lattice vectors act as basis vectors for the position vectors of the atoms, a strain transforms the ion at site \( i \) to a new position \( \tilde{R}_i \), i.e.,

\[
R_{ia} \rightarrow \tilde{R}_{ia} = \sum_\beta h_{\alpha\beta} S_{i\beta} = R_{ia} + \sum_\beta \varepsilon_{\alpha\beta} h_{\gamma\beta} S_{i\gamma}, \tag{29}
\]

where \( S_{ia} \) is the position of site \( i \) in a fractional coordinate system. After application of the strain, the Cartesian distance between sites \( i \) and \( j \) is given by

\[
\tilde{R}_{ij} = R_{ij} + \frac{1}{2R_{ij}} \sum_\alpha \sum_\beta \delta G_{\alpha\beta} S_{i\alpha} S_{j\beta}, \tag{30}
\]

where \( G_{\alpha\beta} \) is the metric tensor of the unstrained crystal. By denoting the second term as \( \delta G_{\alpha\beta} \) and expanding Eq. (29) about \( \delta G = 0 \), we find

\[
\tilde{R}_{ij} = R_{ij} + \frac{1}{2R_{ij}} \sum_\alpha \sum_\beta \delta G_{\alpha\beta} S_{i\alpha} S_{j\beta}. \tag{32}
\]

This Taylor expansion allows us to explicitly evaluate the derivative of the interatomic separation \( R_{ij} \) with respect to strain:

\[
\frac{\partial R_{ij}}{\partial \varepsilon_{\alpha\beta}} = \lim_{\varepsilon_{\alpha\beta} \to 0} \frac{\tilde{R}_{ij} - R_{ij}}{\varepsilon_{\alpha\beta}} = \frac{R_{ji} R_{ij}}{R_{ij}}. \tag{33}
\]

Turning to the aGPT stress tensor, we make a decomposition of the stress tensor \( \sigma \) along the same lines as for the force:

\[
\sigma_{\alpha\beta} = \sigma_{\alpha\beta}^{[1]} + \sigma_{\alpha\beta}^{[II]} + \sigma_{\alpha\beta}^{[III]}, \tag{34}
\]

The second term in the decomposition takes the form of a virial stress tensor,

\[
\sigma_{\alpha\beta}^{[II]} = \frac{1}{2V} \sum_{i,j} F_{ij\alpha} R_{ij\beta}, \tag{35}
\]

where \( F_{ij\alpha}^{[II]} \) is defined as the force on ion \( i \) due to ion \( j \):

\[
F_{ij\alpha}^{[II]} = \sum_{j \neq i} F_{ij\alpha}^{[III]}. \tag{36}
\]

Including the volume dependence of the spatially averaged pseudopotential density will mean that the first contribution to the total stress tensor cannot be written as a virial,

\[
\sigma_{\alpha\beta}^{[I]} = \frac{1}{V} \sum_i \frac{\partial E_{\text{vol}}}{\partial \tilde{n}_i} \frac{\partial \tilde{n}_i}{\partial \varepsilon_{\alpha\beta}}, \tag{37}
\]

where, using the identity \( \partial \Omega / \partial \varepsilon_{\alpha\beta} = \Delta \varepsilon_{\alpha\beta} \), we have

\[
\frac{1}{V} \frac{\partial \tilde{n}_i}{\partial \varepsilon_{\alpha\beta}} = \frac{1}{N} \left( \frac{\partial \tilde{n}_i}{\partial \Omega} + \sum_{j \neq i} \frac{\partial \tilde{n}_i}{\partial \Omega} R_{ij} \right) \frac{\partial \Omega}{\partial \varepsilon_{\alpha\beta}} \tag{38}
\]

This form can be inserted into Eq. (37) and made more explicitly symmetric in \( i \) and \( j \):

\[
\sigma_{\alpha\beta}^{[II]} = \frac{1}{N} \sum_i \frac{\partial E_{\text{vol}}}{\partial \tilde{n}_i} \frac{\partial \tilde{n}_i}{\partial \varepsilon_{\alpha\beta}} + \frac{1}{2N} \sum_{ij} \left( \frac{\partial E_{\text{vol}}}{\partial \tilde{n}_i} + \frac{\partial E_{\text{vol}}}{\partial \tilde{n}_j} \right) \frac{\partial \tilde{n}_i}{\partial \varepsilon_{\alpha\beta}} \frac{\partial \tilde{n}_j}{\partial \varepsilon_{\alpha\beta}} + \frac{1}{2V} \sum_{ij} \left( \frac{\partial E_{\text{vol}}}{\partial \tilde{n}_i} \frac{\partial \tilde{n}_i}{\partial R_{ij}} \frac{\partial R_{ij}}{\partial \varepsilon_{\alpha\beta}} + \frac{\partial E_{\text{vol}}}{\partial \tilde{n}_j} \frac{\partial \tilde{n}_j}{\partial R_{ij}} \frac{\partial R_{ij}}{\partial \varepsilon_{\alpha\beta}} \right). \tag{39}
\]

We calculate the elastic constants numerically by approximating the derivative:

\[
C_{\alpha\beta\gamma\delta} = \left( \frac{\partial \sigma_{\alpha\beta}}{\partial \varepsilon_{\gamma\delta}} \right)_{\varepsilon_{\alpha\beta}=0}. \tag{40}
\]

Since the elastic constants are extremely sensitive to minor changes in the potential, we choose to approximate the derivative in Eq. (40) using a central difference method whose error is of quartic order in the strain parameter. These results are shown alongside GPT and experimental values in Table I. The dynamical bulk modulus \( B_d \) is calculated by using combinations of volume-dependent elastic constants. The aGPT values \( B_d = 35.9 \text{ GPa} \) in Table I are in excellent agreement with the static bulk modulus \( B = 35.8 \text{ GPa} \) that was calculated from derivatives of the equation of state.

The elastic constants can be used to find an optimum value of the averaging sphere radius \( R_a \). In particular, the volume-conserving elastic constants should be equivalent in the GPT

<table>
<thead>
<tr>
<th>[GPa]</th>
<th>( C_{11} )</th>
<th>( C_{12} )</th>
<th>( C_{13} )</th>
<th>( C_{33} )</th>
<th>( C_{44} )</th>
<th>( C_{66} )</th>
<th>( B_d )</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPT</td>
<td>73.2</td>
<td>27.8</td>
<td>24.6</td>
<td>63.6</td>
<td>19.5</td>
<td>22.7</td>
<td>40.5</td>
</tr>
<tr>
<td>aGPT (( R_a = 1.8 ))</td>
<td>63.5</td>
<td>25.5</td>
<td>20.6</td>
<td>62.7</td>
<td>19.5</td>
<td>19.0</td>
<td>35.9</td>
</tr>
<tr>
<td>aGPT (( R_a = 3.4 ))</td>
<td>63.3</td>
<td>25.3</td>
<td>21.0</td>
<td>61.6</td>
<td>18.9</td>
<td>19.0</td>
<td>35.9</td>
</tr>
<tr>
<td>Exp. [13]</td>
<td>59.4</td>
<td>23.6</td>
<td>21.4</td>
<td>61.6</td>
<td>16.4</td>
<td>16.9</td>
<td>35.2</td>
</tr>
</tbody>
</table>

![Image](214107-7)
and aGPT. The difference arises as a result of the approximations and assumptions made in the aGPT formalism. With reference to our calculated aGPT elastic constants in Table I, a smaller cutoff radius $R_a = 1.8R_{WS}$ better reproduces the volume-conserving GPT elastic constants. Also note from Table I that the compressibility problem is removed, with both the GPT and aGPT values of the dynamic bulk modulus $B_d$ in good agreement with the static value $B_s = 35.8$ GPa.

IV. RESULTS

A. Vacancy formation energy

A vacancy is the primary test case for the aGPT, since it is the simplest defect for which there is considerable local volume change. As a consequence, a large amount of the energy required to create a vacancy is not captured by the GPT and other methods based upon second-order pseudopotential perturbation theory. The vacancy formation energy is usually defined as the energy required to remove one atom to infinity and replace it at the surface. The vacancy formation energy $E_{1f}$ can be written without approximation [6] as

$$E_{1f} = \lim_{N_e \to \infty} \left[ E_{\text{tot}}(N_e, 1) - \left( \frac{N_a - 1}{N_a} \right) E_{\text{tot}}(N_a, 0) \right],$$

where $N_a$ is the number of sites and $E_{\text{tot}}$ is a function of both the number of atoms and number of vacancies. The term in the brackets can be evaluated at finite $N$ and then extrapolated into the thermodynamic limit $N \to \infty$. Provided that the atomic positions are relaxed and we are using the bulk equilibrium lattice parameters, it is unnecessary to relax the lattice parameters for the vacancy cell. This is because the large correction to the vacancy formation energy is $-P\Omega_{1f}$, where $\Omega_{1f}$ is the misfit or vacancy formation volume.

We have calculated the relaxed vacancy formation energy in hcp Mg at the experimentally observed atomic volume $\Omega = 156.8$ a.u. and $c/a = 1.62$. In our calculations, the atomic volume is kept constant, which means that the removal of an atom gives rise to a contraction of the lattice. The vacancy formation energy is calculated at multiple values of $N$ and extrapolated to the thermodynamic limit. In addition, we also calculate the misfit volume $\Omega_{1f}$ using the following formula [7]:

$$\frac{\Omega_{1f}}{\Omega_0} = -B_d \frac{\partial E_{1f}}{\partial \Omega},$$

where $B$ is the bulk modulus as calculated in Sec. III. These results are given in Table II and compared to GPT and experimental vacancy formation energies. The vacancy formation energy was calculated for $N_a \in \{54, 128, 250, 432\}$ and then extrapolated to infinity. The extrapolated vacancy formation energy is around 1% less than the vacancy formation energy for $N_a = 432$.

The divacancy binding energy was also calculated for hcp Mg using the following formula:

$$E_{2NN}^{2b} = 2E_{1f} - E_{2NN},$$

where $E_{2NN}^{2b}$ is the divacancy formation energy for a vacancy at the origin and a vacancy in the $i$th neighbor shell. The divacancy formation energy was calculated using an analogous expression to Eq. (41). The ordering of the first- and second-nearest neighbors is dependent on the $c/a$ ratio in hcp crystals. In Mg, the $c/a$ ratio is less than the ideal value which means that the first nearest neighbor lies at a distance less than the lattice parameter $a$. We make a nearest-neighbor definition along the same lines as Usugiu et al. [29]. The divacancy binding energy compiled in Table III converges more slowly with $N_a$ than the vacancy formation energy. In addition, the divacancy binding energy converges more slowly for the aGPT than it does for the GPT. As such, the divacancy binding energy was calculated for larger values of $N_a \in \{250, 432, 686, 1024\}$. Both the aGPT and the GPT are under bound over the first two neighbor shells relative to DFT. We note, however, that the divacancy binding energy is the difference between two quantities with unknown error bars. Therefore, it is unclear whether the underbinding of the aGPT is a deficiency of the method.

B. Stacking fault energies

Information about the plastic behavior of a metal can be inferred from a calculation of the stacking fault energies and the profile of the $\gamma$ line. In particular, the stacking fault energy controls the dissociation width of dislocations into partial dislocations. This in turn controls the ability of a dislocation to cross slip and limits easy glide. The $\gamma$ line is defined in the following manner. An infinite crystal is partitioned into two subcristals with their interface being some crystallographic plane. One half of the crystal is moved relative to the other along some crystallographic direction until the crystal has been translated by an integer multiple of the lattice vectors. The $\gamma$ line is the relative energy change during this process, normalized by the area of the crystallographic plane.

<table>
<thead>
<tr>
<th>$\text{[eV]}$</th>
<th>$E_{2NN}^{2b}$</th>
<th>$E_{2NN}^{2b}$</th>
<th>$E_{2NN}^{2b}$</th>
<th>$E_{2NN}^{2b}$</th>
<th>$E_{2NN}^{2b}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPT</td>
<td>+0.02</td>
<td>+0.02</td>
<td>−0.01</td>
<td>+0.00</td>
<td>+0.00</td>
</tr>
<tr>
<td>aGPT ($R_a = 1.8$)</td>
<td>+0.01</td>
<td>+0.01</td>
<td>−0.05</td>
<td>−0.03</td>
<td>−0.02</td>
</tr>
<tr>
<td>aGPT ($R_a = 3.4$)</td>
<td>+0.00</td>
<td>+0.01</td>
<td>−0.02</td>
<td>−0.01</td>
<td>−0.01</td>
</tr>
<tr>
<td>DFT [29]</td>
<td>+0.06</td>
<td>+0.07</td>
<td>−0.01</td>
<td>+0.01</td>
<td>+0.01</td>
</tr>
</tbody>
</table>

TABLE II. Vacancy formation energies calculated for hcp Mg. The vacancy formation energy $E_{1f}$ was calculated using both the GPT and the aGPT for two values of the averaging sphere radius at the GPT equilibrium atomic volume $\Omega = 156.8$ a.u. and $c/a = 1.62$. The DFT values [29] were calculated at the zero-temperature lattice parameters (excluding zero-point phonons).
Practical computations pose several challenges for this procedure. All the approaches begin by choosing a supercell whose lattice vectors $\{a_1, a_2\}$ define the crystallographic plane over which the slip occurs. For instance, in the basal plane of the hcp structure, these can be represented as the Cartesian vectors $a_1 = [1, 0, 0]$ and $a_2 = [-1/2, \sqrt{3}, 0]$. The supercell is extended $n$ times in the $a_3$ direction such that there are $n$ unit cells. The definition of $a_3$ is not unique and it need not be perpendicular to the crystallographic plane. In fact, the only requirement on $a_3$ is that it connects to an atom which is out of the crystallographic plane. There are a number of ways to create the stacking fault. One such method is the so-called “slab” method [31] whereby the stacking fault is created by moving atoms relative to each other at the approximate center of the supercell. With periodic boundary conditions, the slab method creates an additional stacking fault at the boundary of the supercell with the periodic image. Another method, which we employ, creates the fault by tilting the out-of-plane lattice vector $a_3 \to a_3 + \alpha t$ where $t$ is some integer combination of the in-plane lattice vectors and $\alpha$ is a real number in the interval $[0,1]$. The “tilt” method creates only one stacking fault per supercell whereas the slab method creates two. Thus, with the tilt method there is faster convergence with the number of unit cells $n$.

If a crystal has a stacking fault, the atoms will relax to minimize the interatomic forces that were created by the fault. Using the original Vitek description [31] of the $\gamma$ line, only out-of-plane relaxations are allowed. If such restrictions were not in place then the atoms would relax to either the equilibrium positions or the stable stacking fault up to some strain due to the finite supercell. In certain crystallographic planes and for certain elements, notably the Pyramidal II plane for Mg [32], both the stable stacking fault energy and stacking fault vector calculated using the Vitek method are not very close to the fully relaxed values. Along these planes, if the entire $\gamma$ line is desired then it is necessary to remove the restrictions on in-plane relaxations away from the fault itself [33] or using a nudged elastic band method. The aGPT $\gamma$ line was calculated using the Vitek method for hcp Mg along four directions in four crystallographic planes in Fig. 6 for both the GPT and aGPT. In general, we find that there is agreement between the GPT and aGPT at the stable stacking fault. However, for the unstable stacking fault the aGPT improves upon the GPT relative to the DFT results of Yin et al. [32].

### C. Finite temperature lattice parameters

While we expect the aGPT to apply well to finite temperature, thanks to good agreement with the GPT and DFT harmonic phonon band structure in Figs. 4 and 5, it is important to assess its ability to describe anharmonic effects too. We have looked at thermal expansion, since it is not well captured by quasiharmonic lattice dynamics. For instance, Althoff et al. [11] calculated the thermal expansion coefficient $\beta$ in the quasiharmonic approximation with the GPT and found that there was a discrepancy of roughly 33% between the quasiharmonic values and experimental values. However, close agreement to experiment was found when anharmonic corrections were added in. In This paper, we have calculated the volume in hcp Mg at finite temperature with fixed $c/a$ ratio using molecular dynamics and the stochastic thermostat and barostat of Bussi et al. [34]. For both the GPT and aGPT, we ran eight simulations with different initial velocities corresponding to separate draws from the Maxwell-Boltzmann distribution at three temperatures and zero pressure. These simulations ran for 40 000 time steps, after 40 000 time steps for equilibration, and a time step of 0.1 fs. The aGPT calculation was performed with 512 atoms whilst the GPT calculation was performed with 2000 atoms. The results of this calculation are plotted in Fig. 7. In addition, we estimate the thermal expansion coefficient by regressing the volume on the temperature. For the GPT at 500 K, we calculate $\beta = 7.56 \times 10^{-5}$ K$^{-1}$ and for the aGPT, we calculate $\beta = 7.17 \times 10^{-5}$ K$^{-1}$. These results are in excellent agreement with the previous results of Althoff et al. [11] and experimental values [35].

### V. CONCLUSIONS

We have demonstrated that it is possible to include local volume effects by modifying the GPT so it now depends on a spatially averaged local electron density. In particular, we have developed the aGPT formalism to the extent that it is now possible to do molecular statics and dynamics. To this end, we calculated the vacancy formation energy in hcp Mg at the equilibrium lattice parameters. The aGPT relaxed vacancy formation energy significantly improves upon the relaxed GPT vacancy formation energy relative to the experimentally observed value. In addition, the aGPT provides improved stacking fault energies for hcp Mg.

The computational cost in time of the aGPT is greater than that of the GPT. This is a result of the additional neighbor loop in the calculation of the forces. Provided that the neighbor table maker is linear scaling $O(N)$, for instance using a linked...
FIG. 7. Calculation of the hcp Mg thermal expansion at fixed axial ratio by molecular dynamics. The volume for hcp Mg was averaged over a canonical isobaric isothermal (NPT) ensemble using the GPT and aGPT at three temperatures. The graph shows the change in atomic volume relative to the zero temperature value. The error bars for the 95% confidence interval were found using a bias-free estimator of the variance.

list, both the GPT and aGPT are $O(N)$. Relative to empirical potentials, the major computational cost is due to the long range cut-off in both the GPT and aGPT. This can be demonstrated by considering a short-ranged empirical pair potential whose neighbor cutoff is roughly 1/3 that of the GPT (i.e., it runs over the first handful of neighbor shells); we would expect the GPT to be approximately $3^3 \approx 27$ times slower. Furthermore, we expect the aGPT to be $27N_c$ slower than the GPT where $N_c$ is the number of atoms in a linked-list block.

Bulk properties such as phonon dispersion and elastic constants were also calculated as fundamental tests of the aGPT. The inclusion of the spatially averaged local electron density modifies the bulk phonon dispersion. This is a result of the additional derivatives of the electron density that appear in the expression for the force constant matrix. The elastic constants are also modified although the volume-conserving elastic constants should be the same as for the GPT. It is only the assumptions and approximations in the aGPT that make them differ. Thus, we can use the volume-conserving elastic constants to find an optimum value for the averaging sphere radius $R_c$, which is the free parameter in the aGPT. This constraint would appear to favor near-neighbor values of $R_c$, for instance $R_c \approx 1.8R_W$.

The aGPT can be used to accurately calculate self-diffusion and defect-defect interactions in elemental metals. However, further work needs to be done on extending the aGPT to alloys to study solute diffusion or solute-defect interactions. We plan to use the aGPT to further study vacancies and, in particular, the high-temperature deviation from Arrhenius behavior [36]. All the results presented in this paper were calculated using our in-house Fortran codes. There is a planned future project to incorporate the aGPT into LAMMPS [37].

ACKNOWLEDGMENTS

The authors would like to thank Prof. Mike Finnis for thought-provoking discussions, Prof. Bill Curtin for providing comments on an early version of the paper, and Dr. Zhaoxuan Wu for providing the DFT stacking fault data. G.C.G.S. acknowledges support from the UK EPSRC under the Doctoral Training Partnership (DTP). A.T.P. acknowledges the support of the UK EPSRC under the grant Designing Alloys for Resource Efficiency (DARE), Grant No. EP/L025213/1. The work of J.A.M. was performed under the auspices of the US Department of Energy by Lawrence Livermore National Laboratory under Contract No. DE-AC52-07NA27344.