3 From Electrons to Interatomic Potentials for Materials Simulations

Ralf Drautz
ICAMS
Ruhr-Universität Bochum

Contents

E. Pavarini and E. Koch (eds.)
Topology, Entanglement, and Strong Correlations
Modeling and Simulation Vol. 10
Forschungszentrum Jülich, 2020, ISBN 978-3-95806-466-9
http://www.cond-mat.de/events/correl20

1 Interatomic potentials for materials simulation

The computation of phase diagrams or mechanical properties of materials needs millions of force evaluations for thousands of atoms, requirements which make these simulations unfeasible with density functional theory (DFT). For many problems in materials science, chemistry, or physics it is essential to simplify the description of the interatomic interaction in order that large and long time atomistic simulations become possible.
The development of interatomic potentials is not a new field. Potentials became important when the first computers were available to carry out atomistic simulations. Until about the 1980's the development of interatomic potentials was largely empirical. The electrons were regarded as a glue that mediates the interaction of the atomic cores and the mathematical modeling of the glue was based on intuition and trial and error. Interatomic potentials for materials that are the focus of this chapter were developed along different strategies than force fields for biology and polymer science. Guidance for the development of interatomic potentials for materials was then obtained from DFT or tight-binding electronic structure methods. The assumption of a constant semi-infinite recursion chain (discussed in Sec. 3.1) leads to the second-moment potentials [1–3]. The square-root embedding function of the Finnis-Sinclair potential [4],

$$E_i = \sqrt{\rho_i} + \frac{1}{2}\sum_j V(r_{ji})\,, \tag{1}$$

is explained from the root mean square width of the second moment (see Sec. 3.4), where $\rho_i = \sum_j \phi(r_{ji})$ is the local density of atomic sites and where ϕ and V are pairwise functions of the interatomic distance between atoms i and j. The observation that the atomic energy is a non-linear function of the charge density [5, 6] also motivated the embedded atom method [7], where instead of the square-root function of the Finnis-Sinclair potential a general, quasi-concave embedding function is used. While the Finnis-Sinclair and embedded atom method potentials compute the densities as a pairwise sum over neighbors, Tersoff included an angularly dependent three-body term for modeling directional bond formation in semiconductors [8, 9], so that the energy is written as a non-linear function that depends on a three-body contribution $\rho_i = \sum_{jk} \phi(\boldsymbol{r}_{ji}, \boldsymbol{r}_{ki})$. Later angular terms were also introduced in the modified embedded atom method [10].
Since then many different potentials were developed, with more complex many-body contributions and improved descriptions of bond formation. From the many developments I will present two strategies:

1. the derivation of interatomic potentials from a systematic coarse graining of the electronic structure,

2. the development of general parametrizations of the many-atom interactions for interpolating reference data.

The bond-order potentials (BOPs) are derived by first simplifying DFT to the tight-binding (TB) approximation. Then local, approximate solutions of the TB models are developed from which

expressions for the effective interaction between atoms are obtained. In this way the analytic BOPs provide a rigorous derivation of interatomic potentials for semiconductors [11, 12] and metals [13, 14], where the BOPs for metals will be discussed in this chapter. At their lowest order of approximation the BOPs recover the Tersoff [8, 9] and Finnis-Sinclair [4] potentials, respectively.

In contrast, the development of more formal, general parametrizations is often based on large numbers of DFT data that enable the application of methods from statistical learning for interpolating the reference data. This has led to the development of machine-learning interatomic potentials, such as neural networks potentials [15] or Gaussian process regression for the Gaussian approximation potentials [16]. The field is very active with many recent developments [17–36]. The machine-learning potentials reproduce DFT reference data sets with excellent accuracy. As machine-learning potentials are not derived or motivated by physical or chemical intuition, the excellent accuracy of the machine-learning potentials comes at the cost of interpretability. Machine-learning potentials employ a descriptor that quantifies the local atomic environment. The atomic energy or other atomic properties are then learned as a non-trivial function of the descriptor by training with reference data. The atomic cluster expansion (ACE) provides a formally complete descriptor of the local atomic environment [34,37] and may be used to compare and re-expand different machine-learning interatomic potentials.

In section 2, I will discuss the derivation of the TB approximation from DFT. In Sec. 3 the moments theorem will be introduced. Several local expansions that implicitly or explicitly exploit the moments theorem will then be summarized, before the analytic BOPs will be introduced. In Sec. 4, I will discuss the ACE for the many-atom expansion of the interatomic interaction.

2 Coarse graining the electronic structure for interatomic potentials

The TB approximation is obtained from a second-order expansion of the DFT functional. I will first discuss the second-order expansion of the DFT energy and then introduce the TB approximation. This section follows closely the review in Ref. [38]. It builds on many earlier developments. Here I highlight a few references only, some of which were key for the development of modern TB, others which provide excellent reviews [39–47].

2.1 Second-order expansion of the density functional

The contributions to the Hohenberg-Kohn-Sham DFT energy functional [48, 49] are given by

$$E = T_S + E_H + E_{XC} + E_{ext}\,, \tag{2}$$

with T_S the kinetic energy of the non-interacting electrons, E_H the Hartree energy, E_{XC} the exchange-correlation energy and E_{ext} the interaction of the electrons with the nuclei. The Coulomb interaction between the cores of the nuclei still needs to be added for the computation of total energies. Next, the eigenstates ψ_n are expanded in basis functions φ_i. In general

the basis functions are non-orthogonal,

$$S_{ij} = \langle \varphi_i | \varphi_j \rangle \quad \text{and} \quad \delta_{ij} = \langle \varphi^i | \varphi_j \rangle \quad \text{with} \quad |\varphi^i\rangle = \sum_j S^{-1}_{ij} |\varphi_j\rangle \,, \tag{3}$$

with the overlap matrix S. The basis function indices may be raised with the inverse of the overlap matrix and lowered with the overlap matrix, which enables a more compact notation in the following. The eigenstates are then written as

$$|\psi_n\rangle = \sum_i c^{i(n)} |\varphi_i\rangle \,, \quad \text{and} \quad c^{i(n)} = \sum_j S^{-1}_{ij} c^{(n)}_j \,, \tag{4}$$

with expansion coefficients $c^{i(n)}$. The matrix elements of the density matrix are given by

$$\rho^{ij} = \langle \varphi^i | \hat{\rho} | \varphi^j \rangle = \sum_n f_n \langle \varphi^i | \psi_n \rangle \langle \psi_n | \phi^j \rangle = \sum_n f_n c^{i(n)} (c^{j(n)})^* \,, \tag{5}$$

with the occupation numbers f_n of the eigenstates ψ_n. Here I take $f_n = 1$ for occupied states below the Fermi level and $f_n = 0$ for empty states above the Fermi energy. The charge density is expressed as

$$\rho(\boldsymbol{r}) = \sum_{ij} \rho^{ij} \varphi_i(\boldsymbol{r}) \varphi^*_j(\boldsymbol{r}) \,. \tag{6}$$

If expressed in eigenstates, the density matrix is diagonal

$$\rho_{nn'} = f_n \delta_{nn'} \,. \tag{7}$$

The DFT energy can be categorized in first-, second-, and higher-order contributions in terms of the density matrix. The kinetic energy of non-interacting electrons is linear in the density matrix,

$$T_S = \sum_n f_n \langle \psi_n | \hat{T} | \psi_n \rangle = \boldsymbol{T\rho} \,, \tag{8}$$

where here and in the following the trace is implicitly included in the matrix products,

$$\boldsymbol{T\rho} = \sum_{ij} T_{ij} \rho^{ji} \,. \tag{9}$$

The matrix elements of $\boldsymbol{T}$ are given by

$$T_{ij} = \langle \varphi_i | \hat{T} | \varphi_j \rangle \,. \tag{10}$$

The external energy that contains the interaction of the electrons with the ionic cores is also written as a first-order term

$$E_{ext} = \int V^{ext}(\boldsymbol{r}) \rho(\boldsymbol{r}) \, d\boldsymbol{r} = \boldsymbol{V}^{ext} \boldsymbol{\rho} \,, \tag{11}$$

with

$$V^{ext}_{ij} = \int \varphi^*_i(\boldsymbol{r}) V_{ext}(\boldsymbol{r}) \varphi_j(\boldsymbol{r}) \, d\boldsymbol{r} \,. \tag{12}$$

The Hartree energy E_H is of second order in the density matrix $\boldsymbol{\rho}$ or the density ρ,

$$E_H = \frac{1}{2}\int \frac{\rho(\boldsymbol{r})\rho(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} d\boldsymbol{r}\, d\boldsymbol{r}' = \sum_{ijkl} \frac{1}{2} J^H_{ijkl}\, \rho^{ij}\rho^{kl} = \frac{1}{2}\boldsymbol{J}^H\boldsymbol{\rho}\boldsymbol{\rho}\,, \tag{13}$$

with the Coulomb integral

$$J^H_{ijkl} = \frac{1}{2}\int \frac{\varphi_i^*(\boldsymbol{r})\varphi_j(\boldsymbol{r})\,\varphi_k^*(\boldsymbol{r}')\varphi_l(\boldsymbol{r}')}{|\boldsymbol{r}-\boldsymbol{r}'|} d\boldsymbol{r}\, d\boldsymbol{r}'\,. \tag{14}$$

The exchange-correlation energy E_{XC} is in general parametrized as a non-linear functional of the density ρ and gradients of the density and therefore the only contribution to the DFT energy that contains terms beyond second order. Here I just write a formal series expansion as

$$E_{XC} = \boldsymbol{V}^{XC}\boldsymbol{\rho} + \frac{1}{2}\boldsymbol{J}^{XC}\boldsymbol{\rho}\boldsymbol{\rho} + \frac{1}{6}\boldsymbol{K}^{XC}\boldsymbol{\rho}\boldsymbol{\rho}\boldsymbol{\rho} + \cdots\,. \tag{15}$$

As the exchange correlation energy summarizes corrections due to many-electron interactions, it is relatively short ranged. For example, while $\boldsymbol{J}^H$ decays as $1/r$ for large separations between the orbitals, we expect that the term $\boldsymbol{J}^{XC}$ is limited to distances of the order of the interatomic separation.
By grouping terms of the same order,

$$\boldsymbol{V} = \boldsymbol{T} + \boldsymbol{V}^{ext} + \boldsymbol{V}^{XC}\,, \quad \boldsymbol{J} = \boldsymbol{J}^H + \boldsymbol{J}^{XC}\,, \quad \boldsymbol{K} = \boldsymbol{K}^{XC}\,, \tag{16}$$

the DFT energy is written as a polynomial expansion in the density matrix

$$E = \boldsymbol{V}\boldsymbol{\rho} + \frac{1}{2}\boldsymbol{J}\boldsymbol{\rho}\boldsymbol{\rho} + \frac{1}{6}\boldsymbol{K}\boldsymbol{\rho}\boldsymbol{\rho}\boldsymbol{\rho} + \cdots\,. \tag{17}$$

2.1.1 Hamiltonian and band energy

The elements of the Hamiltonian matrix are obtained from the derivative of the energy with respect to the density matrix

$$\boldsymbol{H} = \frac{\partial E}{\partial \boldsymbol{\rho}} = \boldsymbol{V} + \boldsymbol{J}\boldsymbol{\rho} + \frac{1}{2}\boldsymbol{K}\boldsymbol{\rho}\boldsymbol{\rho} + \cdots\,. \tag{18}$$

The energy may then be represented as the band energy $E_{band} = \boldsymbol{H}\boldsymbol{\rho}$ and a double-counting contribution,

$$E = \boldsymbol{H}\boldsymbol{\rho} - \frac{1}{2}\boldsymbol{J}\boldsymbol{\rho}\boldsymbol{\rho} - \frac{1}{3}\boldsymbol{K}\boldsymbol{\rho}\boldsymbol{\rho}\boldsymbol{\rho} - \cdots\,. \tag{19}$$

The band energy may be decomposed into contributions from different orbitals simply as

$$E_{band,i} = \sum_j H_{ij}\rho^{ij} \quad \text{and} \quad E_{band} = \sum_i E_{band,i}\,. \tag{20}$$

This is called the *intersite* representation of the band energy as it involves two orbitals i and j. By making use of $\sum_j H_{ij}c^{j(n)} = E_n c_i^{(n)}$, where E_n is the eigenvalue of eigenstate ψ_n, and

inserting in the band energy, one arrives at $E_{band} = \sum_i \sum_n f_n E_n c_i^{(n)} (c^{i(n)})^*$. It is customary to define the local and global density of states

$$n_i(E) = \sum_n c_i^{(n)} (c^{i(n)})^* \, \delta(E_n - E) \,, \quad \text{and} \quad n(E) = \sum_i n_i(E) = \sum_n \delta(E_n - E) \,, \tag{21}$$

such that

$$E_{band,i} = \int^{E_F} E \, n_i(E) \, dE \quad \text{and} \quad E_{band} = \int^{E_F} E \, n(E) \, dE \,, \tag{22}$$

with the Fermi energy E_F. This is called the *onsite* representation of the band energy as it involves only one orbital i.

2.1.2 Perturbation expansion

Often one is interested in the response of a material to a perturbation. Then instead of expanding the DFT energy about $\rho = 0$, one would like to discuss the energy associated to the deviation of the density from a particular density $\rho^{(0)}(\boldsymbol{r})$ [39, 40]. I re-expand the series Eq. (17) about a reference density matrix $\boldsymbol{\rho}^{(0)}$ such that

$$\boldsymbol{\rho} = \boldsymbol{\rho}^{(0)} + \delta\boldsymbol{\rho} \,. \tag{23}$$

From Eq. (17) one then obtains

$$E = E^{(0)} + \boldsymbol{H}^{(0)} \delta\boldsymbol{\rho} + \frac{1}{2} \boldsymbol{J}' \delta\boldsymbol{\rho} \, \delta\boldsymbol{\rho} + \frac{1}{6} \boldsymbol{K}' \delta\boldsymbol{\rho} \, \delta\boldsymbol{\rho} \, \delta\boldsymbol{\rho} + \cdots \,, \tag{24}$$

where $\boldsymbol{J}'$ and $\boldsymbol{K}'$ refer to the second and third-order expansion coefficients about $\boldsymbol{\rho}^{(0)}$. The Hamiltonian is given by

$$\boldsymbol{H} = \boldsymbol{H}^{(0)} + \boldsymbol{J}' \delta\boldsymbol{\rho} + \frac{1}{2} \boldsymbol{K}' \delta\boldsymbol{\rho} \delta\boldsymbol{\rho} + \cdots \,, \quad \text{with} \quad \boldsymbol{H}^{(0)} = \boldsymbol{V} + \boldsymbol{J}\boldsymbol{\rho}^{(0)} + \frac{1}{2} \boldsymbol{K} \boldsymbol{\rho}^{(0)} \boldsymbol{\rho}^{(0)} + \cdots \,. \tag{25}$$

2.2 Tight-binding approximation

In the TB approximation one takes the view that bond formation takes place when atomic-like orbitals overlap. In practice this means that one builds TB models on a minimal basis of atomic-like orbitals. The one-electron eigenstates are expanded as linear combinations of atomic-orbital type (LCAO) basis functions. Orbital $|i\alpha\rangle$ is located on atom i and has a well defined angular momentum character, so that α comprises $\alpha = n, l, m$. The basis functions are written as

$$\varphi_{i\alpha}(\boldsymbol{r}) = R_{nl}(|\boldsymbol{r} - \boldsymbol{r}_i|) \, Y_l^m(\theta, \phi) \,, \tag{26}$$

where the radial functions R_{nl} depend only on the distance to the position $\boldsymbol{r}_i$ of atom i and Y_l^m are spherical harmonics or real linear combinations of spherical harmonics. Differently from an LCAO basis that is used in DFT, where often several radial basis functions are employed for a given angular momentum, in TB one typically uses only one radial function for each angular momentum and only includes orbitals that are dictated by the chemistry of the problem at hand.

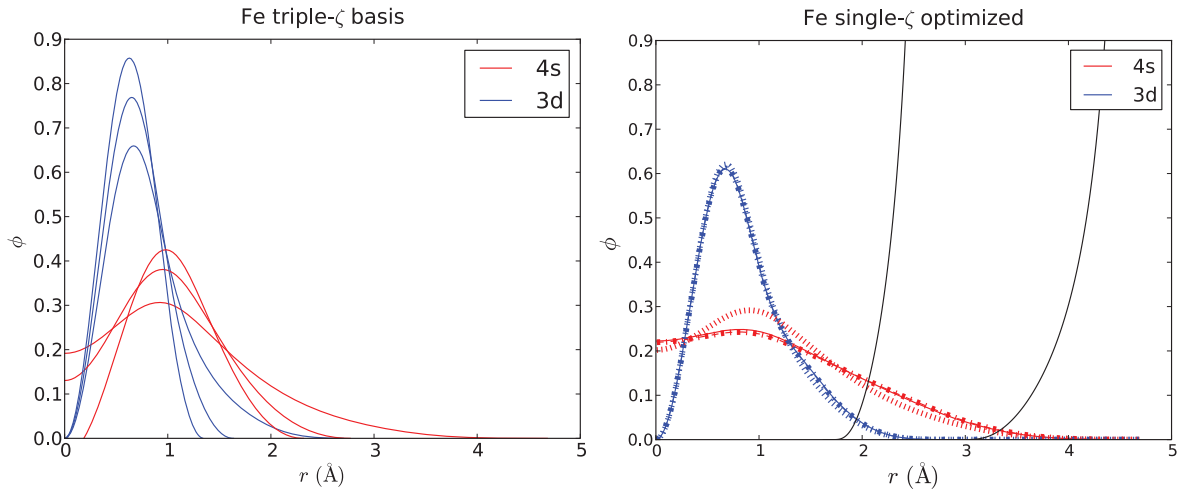

Fig. 1: *Derivation of a single s and a single d radial function for Fe (right) from multiple s and d basis functions (left). Taken from Ref. [50].*

For example, for carbon or silicon four orbitals are used, one s orbital and three p orbitals. This makes the TB approximation a chemically and physically intuitive method for analyzing bond formation in materials.

The TB approximation builds on the perturbation expansion of the Hohenberg-Kohn-Sham functional discussed in the previous section. The reference charge density $\boldsymbol{\rho}^{(0)}$ and the reference Hamiltonian $\boldsymbol{H}^{(0)}$ in Eq. (24) are formally obtained by placing charge neutral, non-magnetic atoms on positions for which the calculation is carried out and then overlapping the charge densities of the atoms. The expansion of the energy Eq. (24) is typically terminated after second order, which implies that the Hamiltonian Eq. (25) is a linear function of the density matrix. As argued before, this should be a good approximation to DFT as only the exchange-correlation energy includes contributions that are of higher than third order and these contributions are partly taken into account in Eq. (24).

The radial functions of the TB orbitals must be modified from the radial functions of a free atom in order that a good representation of the original DFT eigenfunctions may be achieved. In a solid the atomic charge densities contract when the charge densities of neighboring atoms are overlapped and therefore the radial function of the TB orbitals must also contract. Optimal radial functions may be obtained by downfolding from DFT eigenstates [50,51]. Fig. 1 illustrates the downfolding of several radial functions onto a single radial function in Fe.

2.2.1 Parametrized Hamiltonian

Often the Slater-Koster two-center approximation is used to parametrize the Hamiltonian matrix $\boldsymbol{H}^{(0)}$. The matrix element $\langle i\alpha|\hat{H}|j\beta\rangle$ is assumed to depend only on the position of atoms i and j and the orbitals α and β. Clearly, this is a crude approximation and in general the matrix element $\langle i\alpha|\hat{H}|j\beta\rangle$ will depend on other close-by atoms [52], but it is often surprising how much can be achieved with the simple two-center approximation. Fig. 2 shows the Hamiltonian matrix elements for Fe that were obtained from the DFT eigenstates.

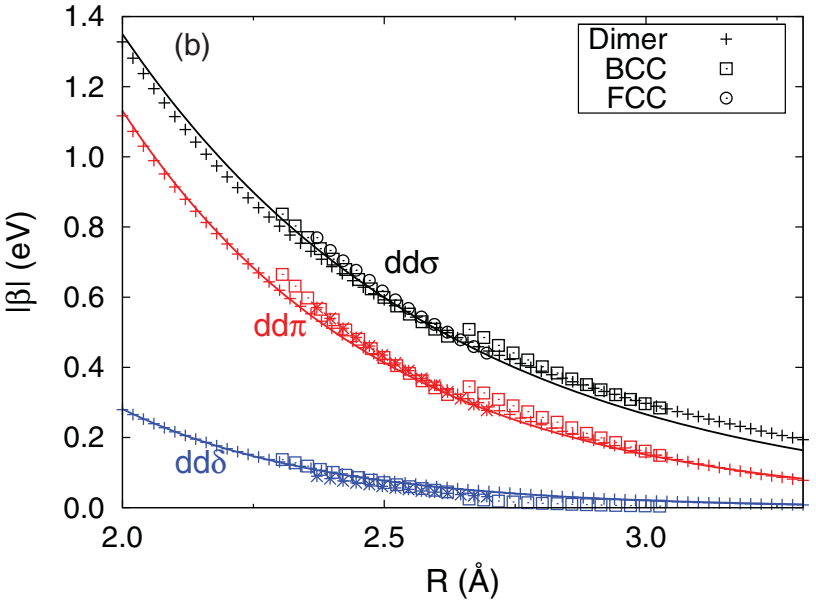

Fig. 2: *Hamiltonian matrix elements for a d-valent orthogonal TB model of Fe in different crystal structures. The solid lines show a fit to an exponential function. Taken from Ref. [50].*

2.2.2 Charge transfer

In TB one frequently assumes that only the diagonal elements of $\delta\boldsymbol{\rho}$ contribute to the second-order term $\frac{1}{2}\boldsymbol{J}\delta\boldsymbol{\rho}\delta\boldsymbol{\rho}$. The diagonal elements are the charges in each orbital

$$q_{i\alpha} = N_{i\alpha} - N_{i\alpha}^{(0)} \,, \tag{27}$$

that correspond to Mulliken charges in a non-orthogonal basis. The index 0 indicates the population of orbital $|i\alpha\rangle$ in a non-magnetic free atom. This approximation has important implications for the structure of the TB model. From Eq. (24) and Eq. (25) the energy is given by

$$E = E^{(0)} + \sum_{i\alpha j\beta} H_{i\alpha j\beta}^{(0)} n_{j\beta i\alpha} + \sum_{i\alpha j\beta} \frac{1}{2} J_{i\alpha j\beta} q_{j\beta} q_{i\alpha} \,, \tag{28}$$

and the Hamiltonian as

$$H_{i\alpha j\beta} = H_{i\alpha j\beta}^{(0)} + J_{i\alpha k\gamma}\, q_{k\gamma} S_{i\alpha j\beta} \,. \tag{29}$$

The modification of $\boldsymbol{H}$ may thus be written as

$$\left(E_{i\alpha} - E_{i\alpha}^{(0)}\right) S_{i\alpha j\beta} \qquad \text{with} \qquad E_{i\alpha} - E_{i\alpha}^{(0)} = J_{i\alpha k\gamma}\, q_{k\gamma} \,. \tag{30}$$

Charge transfer modifies the onsite matrix elements $E_{i\alpha}$ but leaves the rest of the Hamiltonian unchanged from its reference state $H_{i\alpha j\beta}^{(0)}$. Sometimes multipole expansions are used for non-spherical charges with an explicit parametrization of the angular contributions of $\boldsymbol{J}$ in the above equations.

Often, in an even simpler approximation, only the total charge on each atom $q_i = \sum_\alpha q_{i\alpha}$ is considered and the second-order term takes the form $\sum_{ij} \frac{1}{2} J_{ij} q_i q_j$. Then from Eq. (29) all onsite levels on an atom are shifted in parallel upon charge transfer. In literature it is sometimes incorrectly assumed that this automatically corresponds to a point charge approximation.

2.3 Bond formation in the tight-binding approximation

Following Eq. (28) two different types of bond formation are represented in a TB model, the formation of covalent bonds through the modification of the off-diagonal elements of the density matrix $\delta\rho_{i\alpha j\beta}$ with $i\alpha \neq j\beta$ and ionic interactions driven by charge transfer through the modification of the diagonal elements of the density matrix $\delta\rho_{i\alpha i\alpha} = q_{i\alpha}$. I will discuss the decomposition of the TB energy in physically and chemically intuitive and transparent contributions in the following.

2.3.1 Bond energy

The bond energy summarizes the energy that is stored in the bonds between different atoms

$$E_{bond} = \sum_{i\alpha j\beta} \big(H_{i\alpha j\beta} - E_{i\alpha} S_{i\alpha j\beta}\big) n_{j\beta i\alpha} = E_{band} - \sum_{i\alpha} E_{i\alpha} N_{i\alpha} \,. \tag{31}$$

Differently from the band energy, the bond energy is invariant with respect to a shift of the energy scale. Using Eq. (29) the bond energy is closely related to the linear term in the TB energy

$$\boldsymbol{H}^{(0)}\boldsymbol{n} = E_{bond} + \sum_{i\alpha} E^{(0)}_{i\alpha} N_{i\alpha} \,. \tag{32}$$

Equivalent to the intersite representation of the bond energy above is the onsite representation,

$$E_{bond} = \sum_{i\alpha} \int_{-\infty}^{E_F} \big(E - E_{i\alpha}\big) n_{i\alpha}(E)\, dE \,. \tag{33}$$

The population of the density of states below $E_{i\alpha}$ leads to a negative contribution to the bond energy, i.e., corresponding to the filling of bonding states. Once states above $E_{i\alpha}$ have to be populated, the bond energy decreases, corresponding to a filling of anti-bonding states. The integral over the complete band is zero, $0 = E_{bond} = \sum_{i\alpha} \int_{-\infty}^{\infty} \big(E - E_{i\alpha}\big) n_{i\alpha}(E)\, dE$, which helps to show that the bond energy is always smaller or equal to zero,

$$E_{bond} \leq 0 \,. \tag{34}$$

2.3.2 Promotion energy

When bonds are formed, the onsite levels are re-populated. In the free atom the number of electrons per orbital is denoted by $N^{(0)}_{i\alpha}$. For charge neutral atoms the promotion energy is then written as

$$E_{prom} = \sum_{i\alpha} E^{(0)}_{i\alpha} \big(N_{i\alpha} - N^{(0)}_{i\alpha}\big) \,. \tag{35}$$

In contrast to the bond energy, the promotion energy is strictly positive as the electrons in the free atom occupy the energetically lowest orbitals.

2.3.3 Free atom energy

For the evaluation of the binding energy the energies of the free atoms are subtracted

$$E_B = E^{TB} - E^{TB}_{free\ atoms}\,. \tag{36}$$

From Eq. (28) the TB energy of non-magnetic, charge neutral free atoms is given by

$$E^{TB}_{free\ atoms} = \sum_{i\alpha} E^{(at)}_{i\alpha} N^{(0)}_{i\alpha} - \frac{1}{2}\sum_{i\alpha\beta} J_{i\alpha i\beta} N^{(0)}_{i\alpha} N^{(0)}_{i\beta}\,, \tag{37}$$

where $E^{(at)}_{i\alpha}$ are the eigenstates of the free atom i and the population of the atomic orbitals is equal to the population in the reference state, $N^{(0)}_{i\alpha}$.

2.3.4 Preparation energy

The preparation energy takes into account modifications of the onsite levels when the free atoms are brought together and their charge density is overlapped to the reference charge density $\boldsymbol{\rho}^{(0)}$,

$$E_{prep} = \sum_{i\alpha} \left(E^{(0)}_{i\alpha} - E^{(at)}_{i\alpha}\right) N^{(0)}_{i\alpha}\,. \tag{38}$$

2.3.5 Charge transfer

Charge transfer leads to two further contributions to the energy and a somewhat modified expression for the promotion energy above. Because of the onsite level difference between atoms there is an energy linear in charge, and further the second-order contribution to the TB energy Eq. (28). The two terms are denoted as an electrostatic interaction of charges on different atoms

$$E_{es} = \frac{1}{2}\sum_{ij}^{i\neq j} J_{ij} q_i q_j\,, \tag{39}$$

and an ionic onsite contribution for charging each atom

$$E_{ion} = \bar{E}_i q_i + \frac{1}{2}\sum_i J_{ii} q_i^2\,. \tag{40}$$

The energy $\bar{E}_i$ is obtained as a weighted average of the reference onsite levels on atom i and corresponds to the electronegativity of the atom. The parameter J_{ii} further determines resistance against charge transfer from the charge neutral state.

2.3.6 Repulsive energy

The repulsive energy summarizes all terms that do not explicitly depend on the modification of the density matrix $\delta\boldsymbol{n}$. For this reason E_{prep} is also absorbed in the repulsive energy

$$E_{rep} = -\frac{1}{2}\sum_{i\alpha j\beta}^{j\neq j} J_{i\alpha j\beta} N^{(0)}_{i\alpha} N^{(0)}_{j\beta} + E_{nuc} + E_{prep}\,, \tag{41}$$

where E_{nuc} corresponds to the Coulomb repulsion of the bare atomic cores.

2.3.7 Summary of the energy in the tight-binding approximation

In summary, the second-order expansion of DFT, Eq. (24), cast in a TB binding energy is re-written in the form

$$E_B = E_{bond} + E_{prom} + E_{ion} + E_{es} + E_{rep}\,. \tag{42}$$

The TB expansion suggests a representation of bond formation in the steps summarized in the following table. Typically, the steps 1–3 are repulsive, while step 4 is attractive and drives bond formation.

1	E_{rep}	$\rightarrow$ overlap atomic charge densities
2	E_{ion}	$\rightarrow$ charge atoms
3	E_{prom}	$\rightarrow$ re-populate atomic energy levels
4	$E_{bond} + E_{es}$	$\rightarrow$ chemical and electrostatic interactions

3 The moments theorem and local expansions

For the derivation of interatomic potentials I next turn to local solutions of the TB model. In particular, the moments theorem will allow us to relate the local electronic structure to the local atomic environment, which is critical for analyzing the interaction between atoms.
The moments of the local density of states may be defined as

$$\mu_{i\alpha}^{(N)} = \int E^N n_{i\alpha}(E)\, dE\,. \tag{43}$$

The moments may be used to characterize the density of states. The zeroth moment is just the norm, $\mu_{i\alpha}^{(0)} = 1$. The first moment gives the center of the density of states. From the second moment the root mean square width of the density of states may be obtained and from the third moment its skewness. The fourth moment characterizes the bimodality of the local density of states, etc. If all moments are known, then the density of states may be reconstructed and therefore may be viewed as a function of its moments,

$$n_{i\alpha}(E) = n_{i\alpha}(E, \mu_{i\alpha}^{(0)}, \mu_{i\alpha}^{(1)}, \mu_{i\alpha}^{(2)}, \dots)\,. \tag{44}$$

The idea of reconstructing the density of states from its moments (or equivalent information) is the basis for the different methods that will be discussed in the following. In order that such a reconstruction can be efficient, the moments of the density of states need to be accessible. Here the moments theorem that I will briefly derive in the following provides the critical link. I assume for ease of notation that the basis functions are orthonormal

$$\langle i\alpha | j\beta \rangle = \delta_{ij}\delta_{\alpha\beta}\,. \tag{45}$$

By using the definition of the density of states from Eq. (21) the moments may be rewritten as

$$\begin{aligned}\mu_{i\alpha}^{(N)} &= \int E^N n_{i\alpha}(E)\, dE = \sum_n E_n^N \langle i\alpha|n\rangle\langle n|i\alpha\rangle \int \delta\big(E_n - E\big)\, dE \\ &= \sum_n \langle i\alpha|\hat{H}^N|n\rangle\langle n|i\alpha\rangle = \langle i\alpha|\hat{H}^N|i\alpha\rangle\,, \end{aligned} \tag{46}$$

where I used the completeness of the eigenstates $\hat{1} = \sum_n |n\rangle\langle n|$. A further manipulation enables a geometric interpretation

$$\begin{aligned}\mu_{i\alpha}^{(N)} = \langle i\alpha|\hat{H}^N|i\alpha\rangle &= \sum_{j\beta k\gamma\ldots} \langle i\alpha|\hat{H}|j\beta\rangle\langle j\beta|\hat{H}|k\gamma\rangle\langle k\gamma|\hat{H}\ldots\hat{H}|i\alpha\rangle \\ &= \sum_{j\beta k\gamma\ldots} H_{i\alpha j\beta} H_{j\beta k\gamma} H_{k\gamma\ldots}\ldots H_{\ldots i\alpha}\,, \end{aligned} \tag{47}$$

with a complete basis $\hat{1} = \sum_{i\alpha} |i\alpha\rangle\langle i\alpha|$. The last identity tells us that the N^{th} moment may be obtained from the product of N Hamiltonian matrix elements. The N^{th} moment may therefore be described as the sum of all self-returning hopping path of length N that start and end on the same basis function.

Along the same lines it is straightforward to show that the moments of the spectrally resolved density matrix

$$n_{i\alpha j\beta}(E) = \frac{d\rho_{i\alpha j\beta}}{dE}\,, \quad \text{or} \quad \rho_{i\alpha j\beta} = \int^{E_F} n_{i\alpha j\beta}(E)\, dE\,, \tag{48}$$

may be obtained as

$$\xi_{i\alpha j\beta}^{(N)} = \int E^N n_{i\alpha j\beta}(E)\, dE = \langle i\alpha|\hat{H}^N|j\beta\rangle\,. \tag{49}$$

The N^{th} moment $\xi_{i\alpha j\beta}^{(N)}$ of the spectrally resolved density matrix is thus given by all interference paths of N products of the Hamiltonian matrix that start and end on orbitals $|i\alpha\rangle$ and $|j\beta\rangle$, respectively.

As the Hamiltonian matrix elements depend on the positions of the atoms, the moments theorem relates the atomic structure to the electronic structure. For the reconstruction of the local density of states, the lowest moments contribute basic information on the width and shape of the density of states, while higher moments may be used to reconstruct increasingly finer details. This is intuitive: the matrix elements $H_{i\alpha j\beta}$ decay roughly exponentially with distance between the atoms i and j, which means that the low moments only sample the local environment of an atom and higher moments incorporate information of an increasingly distant neighborhood of orbital $|i\alpha\rangle$.

In the following I will discuss methods for the reconstruction of the band energy from the atomically local neighborhood. These methods were developed originally for linear-scaling DFT or TB. The methods have different starting points, but implicitly or explicitly they all correspond to a reconstruction of the density of states from its moments. For the analytic bond-order potentials we use the moments to derive a hierarchical analytic expansion of the interatomic interaction.

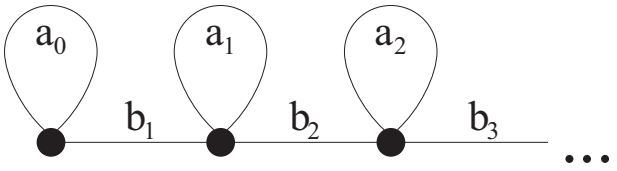

Fig. 3: *Illustration of the semi-infinite recursion chain Hamiltonian.*

3.1 Recursion and numerical bond-order potentials

In the recursion method the Hamiltonian matrix is transformed to tridiagonal form, from which the Green function may be obtained as a continued fraction [53, 54]. Given a starting orbital $|u_0\rangle = |i\alpha\rangle$ first a basis transformation is carried out,

$$b_{n+1}|u_{n+1}\rangle = \big(\hat{H} - a_n\big)|u_n\rangle - b_n|u_{n-1}\rangle\,, \tag{50}$$

with

$$b_n = \langle u_n|\hat{H}|u_{n-1}\rangle\,, \quad \text{and} \quad a_n = \langle u_n|\hat{H}|u_n\rangle\,. \tag{51}$$

The recursion is initialized with $b_0 = 0$ and leads to orbitals $|u_n\rangle$ that have remarkable properties: the resulting Hamiltonian matrix is tridiagonal, i.e., it only has entries on the diagonal and next to the diagonal,

$$\langle u_n|\hat{H}|u_m\rangle = \begin{pmatrix} a_0 & b_1 & 0 & 0 & 0 & 0 & \dots \\ b_1 & a_1 & b_2 & 0 & 0 & 0 & \dots \\ 0 & b_2 & a_2 & b_3 & 0 & 0 & \dots \\ 0 & 0 & b_3 & a_3 & b_4 & 0 & \dots \\ 0 & 0 & 0 & b_4 & a_4 & b_5 & \dots \\ 0 & 0 & 0 & 0 & b_5 & a_6 & \ddots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \ddots \end{pmatrix}, \quad \text{and} \quad \langle u_n|u_m\rangle = \delta_{nm}\,. \tag{52}$$

The tridiagonal Hamiltonian may be viewed as semi-infinite, one-dimensional chain and is illustrated in Fig. 3. The recursion also shows that every Hamiltonian may be represented as a one-dimensional semi-infinite chain with nearest-neighbor interactions.
The moments of the density of states Eq. (43) may be obtained from self-returning paths along the tridiagonal Hamiltonian matrix as

$$\begin{aligned} \mu_{i\alpha}^{(0)} &= 1\,, \\ \mu_{i\alpha}^{(1)} &= a_0\,, \\ \mu_{i\alpha}^{(2)} &= a_0^2 + b_1^2\,, \\ \mu_{i\alpha}^{(3)} &= a_0^3 + (2a_0 + a_1)b_1^2\,, \\ \mu_{i\alpha}^{(4)} &= a_0^4 + b_1^4 + (3a_0^2 + 2a_0a_1 + a_1^2 + b_2^2)b_1^2\,, \\ &\vdots \end{aligned}$$

3.1.1 Green function expansion

The Green function is defined as

$$\hat{G} = \left(E\hat{1} - \hat{H}\right)^{-1} . \tag{53}$$

The matrix elements of the Green function in eigenstates are given by

$$\langle\psi_n|\hat{G}|\psi_m\rangle = \frac{\delta_{nm}}{E - E_n} . \tag{54}$$

This is easily verified by inserting the identify $\hat{1} = \sum_{n'} |\psi_{n'}\rangle\langle\psi_{n'}|$ into

$$\begin{aligned}\delta_{n,m} = \langle\psi_n|\psi_m\rangle &= \langle\psi_n|(E\hat{1}-\hat{H})\hat{G}|\psi_m\rangle \\ &= \sum_{n'}\langle\psi_n|E\hat{1}-\hat{H}|\psi_{n'}\rangle\langle\psi_{n'}|\hat{G}|\psi_m\rangle = \frac{E - E_n}{E - E_m}\,\delta_{nm} .\end{aligned} \tag{55}$$

The Green function matrix elements in basis functions are given by

$$G_{i\alpha j\beta}(E) = \sum_{nm}\langle i\alpha|\psi_n\rangle\langle\psi_n|\hat{G}|\psi_m\rangle\langle\psi_m|j\beta\rangle = \sum_n \frac{\langle i\alpha|\psi_n\rangle\langle\psi_n|j\beta\rangle}{E - E_n} . \tag{56}$$

With the help of the recursion chain Eq. (50) the diagonal elements of the Green function matrix may be expressed in the form of a continued fraction [53],

$$G_{i\alpha i\alpha}(E) = G_{00}(E) = \cfrac{1}{E - a_0 - \cfrac{b_1^2}{E - a_1 - \cfrac{b_2^2}{E - a_2 - \cfrac{b_3^2}{E - a_3 - \cfrac{b_4^2}{E - a_4 - \cfrac{b_5^2}{\ddots}}}}}} . \tag{57}$$

Next the Green function is related to the density matrix by making use of the identity

$$-\frac{1}{\pi}\,\mathrm{Im}\int \frac{1}{E - E_n}\,dE = \int \delta(E-E_n)\,dE , \tag{58}$$

so that comparing to Eq. (21) and Eq. (48) results in the identities

$$n_{i\alpha}(E) = -\frac{1}{\pi}\,\mathrm{Im}\,G_{i\alpha i\alpha}(E) \quad \text{and} \quad n_{i\alpha j\beta}(E) = -\frac{1}{\pi}\,\mathrm{Im}\,G_{i\alpha j\beta}(E) . \tag{59}$$

The Green function may therefore be used for the computation of the band or bond energy. This is used in the numerical bond-order potentials (BOPs) that will be introduced next.

3.1.2 Numerical bond-order potentials

We are interested in the local calculation of the band energy or, for the TB approximation, the bond energy associated to orbital $|i\alpha\rangle$. This is achieved by terminating the recursion expansion of the Green function after a few recursion levels n: The recursion coefficients a_m and b_m for $m > n$ are replaced by a constant terminator

$$a_m = a_\infty\,, \quad b_m = b_\infty \quad \text{for } m > n\,. \tag{60}$$

The terminator in the Green function may be summed analytically and leads to termination of the continued fraction at level n by

$$E - a_{n-1} - T(E)\,, \tag{61}$$

with

$$T(E) = \frac{1}{2}\left(E - a_\infty - \sqrt{(E-a_\infty)^2 - 4b_\infty^2}\right) \tag{62}$$

obtained from the functional equation $T(E) = b_\infty^2/(E - a_\infty - T(E))$ for the ininite continued fraction with constant coefficients.

The corresponding density of states is different from zero only in the interval between $a_\infty - 2b_\infty$ and $a_\infty + 2b_\infty$. A local expansion of the bond energy is now obtained as

$$E_{bond,i\alpha} = -\frac{1}{\pi}\,\mathrm{Im}\int^{E_F} \cfrac{E - E_{i\alpha}}{E - a_0 - \cfrac{b_1^2}{E - a_1 - \cfrac{b_2^2}{\ddots - \cfrac{\ddots}{E - a_{n-1} - T(E)}}}}\,. \tag{63}$$

The integration of the Green function is carried out numerically, therefore the name numerical BOPs.

For the evaluation of forces the off-diagonal elements of the Green function are also required. These are obtained by defining

$$|u_0\rangle = \frac{1}{\sqrt{2}}\Big(|i\alpha\rangle + e^{i\vartheta}|j\beta\rangle\Big)\,, \tag{64}$$

with $\vartheta = \cos^{-1}\lambda$ and therefore

$$G_{00} = \lambda G_{i\alpha j\beta} + \frac{1}{2}\big(G_{i\alpha i\alpha} + G_{j\beta j\beta}\big)\,. \tag{65}$$

A particular termination of the expansion ensures that the onsite and intersite representation of the bond energy are identical [55]. Details of the numerical BOPs are available in Refs. [56–58].

3.2 Kernel polynomial method

In the Kernel Polynomial Method (KPM) [59–61] the density of states is represented as

$$n_{i\alpha}(\varepsilon) = \int K(\varepsilon,\varepsilon')\, n_{i\alpha}(\varepsilon')\, d\varepsilon' \,, \tag{66}$$

where the energy has been rescaled as

$$\varepsilon = \frac{E - a_\infty}{2b_\infty} \,. \tag{67}$$

It is clear that this identity only holds if the kernel fulfills $K(\varepsilon,\varepsilon') = \delta(\varepsilon-\varepsilon')$. In order to achieve an approximate, local representation of $n_{i\alpha}(\varepsilon)$ the kernel is expanded in Chebyshev polynomials of the first kind

$$K(\varepsilon,\varepsilon') = \frac{1}{\pi}\frac{1}{\sqrt{1-\varepsilon^2}}\left(g_T^{(0)} + 2\sum_{n=1}^{n_{max}} g_T^{(n)} T_n(\varepsilon) T_n(\varepsilon') \right) . \tag{68}$$

The factors $g_T^{(n)}$ are chosen in such a way that for every n_{max} the kernel is positive, $K(\varepsilon,\varepsilon') \geq 0$, while it is also as narrow as possible for an efficient convergence to the Dirac delta function when n_{max} is increased. Typically $g_T^{(n)}$ smoothly decays as a function of n from $g_T^{(0)} = 1$ to $g_T^{(n_{max})} = 0$. Fig. 4 illustrates different damping factors.
By inserting the expansion for the Kernel in Eq. (66), an expansion for the density of states is obtained as

$$n_{i\alpha}(\varepsilon) = \frac{1}{\pi}\frac{1}{\sqrt{1-\varepsilon^2}}\left(g_T^{(0)} + 2\sum_{n=1}^{n_{max}} g_T^{(n)} T_n(\varepsilon) \mu_n^T \right) , \tag{69}$$

with the Chebyshev moments

$$\mu_n^T = \int_{-1}^{1} T_n(\varepsilon)\, n_{i\alpha}(\varepsilon)\, d\varepsilon \,. \tag{70}$$

As the Chebyshev polynomials may just be written in powers of ε, the Chebyshev moments are linear combinations of the moments of the density of states, Eq. (21).
The damping factors $g_T^{(n)}$ that ensure a positive expansion of the density of states remove much of the contribution of higher moments. One way to avoid this is to add higher moments that are generated based on a maximum entropy principle [63].

3.3 Fermi operator expansion

In DFT it is customary to introduce an electronic temperature. This is done in part for practical reasons, as temperature dampens details of the Fermi surface in a metal and leads to faster convergence of the k-space integration over the Brillouin zone as a function of the k-mesh density. For the Fermi operator expansion the electronic temperature provides the starting point for the expansion. With temperature, the band energy and the number of electrons are computed as

$$E_{band} = \int \varepsilon\, f(\varepsilon,\mu)\, n(\varepsilon)\, d\varepsilon \quad \text{and} \quad N = \int f(\varepsilon,\mu)\, n(\varepsilon)\, d\varepsilon \tag{71}$$

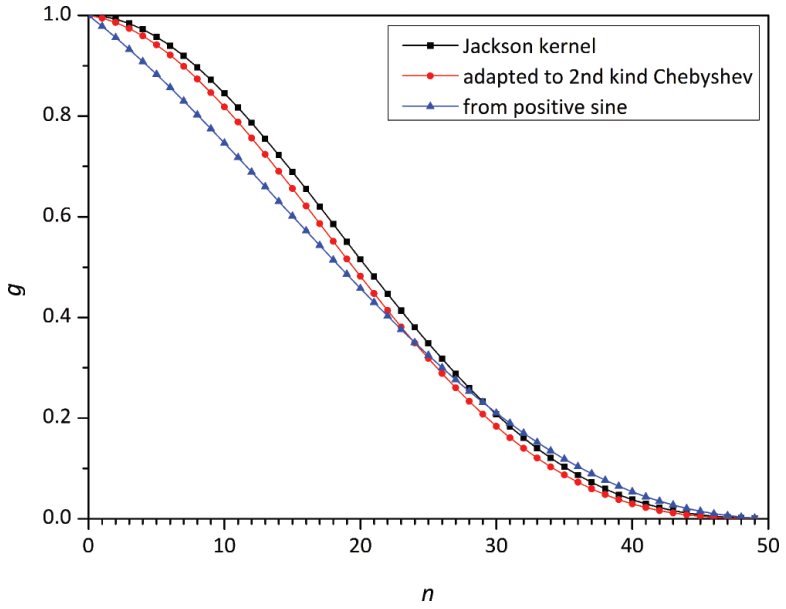

Fig. 4: *Damping factors used in the KPM (Jackson kernel) and analytic BOPs for* $n_{exp} = 50$. *Taken from Ref. [62].*

with the energy scale Eq. (67), where μ is the electron chemical potential and

$$f(\varepsilon,\mu) = \frac{1}{1+\exp\left(\frac{\varepsilon-\mu}{k_B T}\right)}\,, \tag{72}$$

the temperature dependent Fermi-Dirac distribution function and $n(\varepsilon)$ the density of states. At $T = 0$ K the smearing is zero and $f(\varepsilon,\mu)$ corresponds to the Heaviside step function $\Theta(\varepsilon,\varepsilon_F)$ which is one below the Fermi energy ε_F and zero above. In the Fermi operator expansion (FOE) method [64, 65] the density matrix is locally approximated by writing it as

$$\rho_{i\alpha j\beta} = \int f(\varepsilon,\mu)\, n_{i\alpha j\beta}(\varepsilon) d\varepsilon\,, \tag{73}$$

and then expanding f in a polynomial

$$f(\varepsilon,\mu) = \sum_k c_k \varepsilon^k\,. \tag{74}$$

By making use of Eq. (49) the density matrix is written as

$$\rho_{i\alpha j\beta} = \sum_k c_k\, \xi^{(k)}_{i\alpha j\beta}\,, \quad \text{and} \quad E_{band} = \sum_{i\alpha j\beta}\sum_k c_k\, \xi^{(k)}_{i\alpha j\beta} H_{j\beta i\alpha} = \sum_{i\alpha}\sum_k c_k \mu^{(k+1)}_{i\alpha}\,. \tag{75}$$

In practice the Fermi-Dirac distribution function is expanded in Chebyshev polynomials

$$f(\varepsilon) = \frac{1}{\pi\sqrt{1-\varepsilon^2}}\left(\mu_0^T + 2\sum_{n=1}^{n_{max}} \mu_n^T T_n(\varepsilon)\right). \tag{76}$$

with the Chebyshev moments μ_n^T, Eq. (70), and the band energy accordingly. Another representation of the Fermi operator expansion, the rational representation, may be related to the recursion expansion.

3.4 Analytic bond-order potentials

The analytic BOPs combine the recursion expansion and the KPM. I start from a scaled energy Eq. (67) that allows us to work with Chebyshev polynomials that are defined on the interval $[-1,1]$. The analytic BOPs use the Chebyshev polynomials of the second kind as the they are orthogonal with respect to the square root function

$$\frac{2}{\pi}\int_{-1}^{+1} U_n(\varepsilon)\, U_m(\varepsilon)\, \sqrt{1-\varepsilon^2}\, d\varepsilon = \delta_{nm}\,. \tag{77}$$

The square root function is also the density of states of the semi-infinite recursion chain when all matrix elements are identical $a_0 = a_1 = \cdots = a_\infty$ and $b_1 = b_2 = \cdots = b_\infty$. If the reference energy is shifted to $a_0 = 0$, then $b_1^2 = \mu_{i\alpha}^{(2)}$. As $b_1 = b_\infty$ determines the width of the density of states, the bond energy scales as $\sqrt{\mu_{i\alpha}^{(2)}}$ and the Finnis-Sinclair potential, Eq. (1), is immediately obtained. Therefore, by choosing the Chebyshev polynomials of the second kind, the analytic BOPs incorporate the Finnis-Sinclair potential at the lowest order of approximation.
The Chebyshev polynomials of the second kind fulfill the recursion relation

$$U_{n+1}(\varepsilon) = 2\varepsilon\, U_n(\varepsilon) - U_{n-1}(\varepsilon)\,, \tag{78}$$

with $U_0 = 1$ and $U_1 = 2\varepsilon$.
The expansion coefficients for the density of states are obtained by projection,

$$\sigma_{i\alpha}^{(n)} = \int_{-1}^{+1} U_n(\varepsilon)\, n_{i\alpha}(\varepsilon)\, d\varepsilon\,, \tag{79}$$

and the density of states is expressed as

$$n_{i\alpha}(\varepsilon) = \frac{2}{\pi}\sum_n \sqrt{1-\varepsilon^2}\, \sigma_{i\alpha}^{(n)}\, U_n(\varepsilon)\,. \tag{80}$$

The expansion coefficients $\sigma_{i\alpha}^{(n)}$ are computed using the moments theorem Eq. (43) as

$$\sigma_{i\alpha}^{(n)} = \langle i\alpha | U_n(\hat{h}) | i\alpha \rangle\,, \tag{81}$$

with the scaled Hamiltonian

$$\hat{h} = \frac{\hat{H} - a_\infty}{2b_\infty}\,. \tag{82}$$

In practice n_{max} expansion coefficients are computed, which corresponds to evaluating n_{max} moments of the density of states. Thus the expansion Eq. (80) becomes

$$n_{i\alpha}(\varepsilon) = \sum_{n=0}^{n_{max}} \sqrt{1-\varepsilon^2}\, \sigma_{i\alpha}^{(n)} U_n(\varepsilon)\,, \tag{83}$$

or using $\varepsilon = -\cos\phi$,

$$n_{i\alpha}(\varepsilon) = \sum_{n=0}^{n_{max}} \sigma_{i\alpha}^{(n)} \sin(n+1)\phi\,. \tag{84}$$

If a Fourier series expansion is abruptly terminated because only the first n_{max} expansion coefficients are taken into account, this may lead to significant oscillations that are known as Gibbs ringing. In particular, these oscillations may be so large that the expansion of the density of states Eq. (84) may become negative.
The Gibbs ringing may be removed and a strictly positive expansion of the density of states may be enforced by damping the expansion coefficients,

$$n_{i\alpha}(\varepsilon) = \sum_{n=0}^{n_{max}} g_n \, \sigma_{i\alpha}^{(n)} \sin(n{+}1)\phi \, . \tag{85}$$

The damping factors g_n are similar to the damping factors in KPM and decrease monotonically from $g_0 = 1$ to $g_{n_{max}} = 0$ [62, 66], see Fig. 4. They damp oscillations and avoid Gibbs ringing and potentially negative values of the density of states. As the damping factors decrease to zero for n_{max}, they remove most of the contribution from higher moments. Therefore more moments need to be calculated, i.e., n_{max} needs to be increased. As the calculation of the moments is the most time consuming part in the energy and force evaluation, one would like to keep n_{max} as small as possible.
This may be resolved by terminating the expansion Eq. (85). One first evaluates moments up to n_{max} from the Hamiltonian using the moments theorem. Further moments from $n_{max}{+}1$ up to $n_{exp} \gg n_{max}$ are then computed using an estimated model Hamiltonian that has the form of a semi-infinite chain with nearest neighbor bonds only [62]. Because of the simple structure of the semi-infinite chain, only very few matrix elements need to be multiplied and therefore the computation of the moments along the chain is very fast. The resulting expansion takes the form

$$n_{i\alpha}(\varepsilon) = \sum_{n=0}^{n_{max}} g_n \, \sigma_{i\alpha}^{(n)} \sin(n{+}1)\phi + \sum_{n=n_{max}+1}^{n_{exp}} g_n \, \sigma_{i\alpha}^{(n)} \sin(n{+}1)\phi \, . \tag{86}$$

The damping factors decay monotonically from $g_0{=}1$ to $g_{n_{exp}}{=}0$. For $n_{exp}{\gg}n_{max}$ the damping factors for the first few moments are close to one and the contributions of the corresponding moments are hardly affected. This means that the expansion as a function of n_{max} converges quickly to the tight-binding reference. In practice one uses $n_{exp} \approx 20{\times}n_{max}$. This leads to a good quality of the reconstructed DOS already at a small number of calculated moments.
The density of states Eq. (86) may be integrated analytically,

$$E_{bond,i\alpha} = 2b_\infty \sum_{n=0}^{n_{exp}} g_n \, \sigma_{i\alpha}^{(n)} \Big(\hat{\chi}_{n+2}(\phi_F) - \gamma_0 \hat{\chi}_{n+1}(\phi_F) + \hat{\chi}_n(\phi_F) \Big) \, , \tag{87}$$

with $\gamma_0 = (E_{i\alpha}{-}a_\infty)/b_\infty$ and $\hat{\chi}_0 = 0$,

$$\hat{\chi}_1 = 1 - \frac{\phi_F}{\pi} + \frac{1}{2\pi} \sin(2\phi_F) \, . \tag{88}$$

The response functions take the form

$$\hat{\chi}_n(\phi_F) = \frac{1}{\pi} \left(\frac{\sin(n{+}1)\phi_F}{n{+}1} - \frac{\sin(n{-}1)\phi_F}{n{-}1} \right) . \tag{89}$$

Finally, the number of electrons in orbital $N_{i\alpha}$ is obtained as

$$N_{i\alpha} = \sum_{n=0}^{n_{exp}} g_n \, \sigma_{i\alpha}^{(n)} \, \hat{\chi}_{n+1}(\phi_F) \, , \tag{90}$$

with the Fermi phase $\varepsilon_F = -\cos\phi_F$. For the gradients of the bond energy $E_{bond,i\alpha}$ and the number of electrons $N_{i\alpha}$ efficient analytic expressions may be obtained [14, 62]. The analytic BOPs have been extended further to include non-collinear magnetism [14, 67]. The analytic BOPs scale-linearly with the number of atoms. An efficient and parallel implementation is available [68] that enables simulations with millions of atoms.

3.5 Examples for the analytic bond-order potentials

3.5.1 Bond order

The density matrix is also called the bond order, a factor of two is usually between the two quantities in non-magnetic systems. I transform the orbitals $|i\alpha\rangle$ and $|j\beta\rangle$ in a new basis of bonding and anti-bonding dimer states to analyze bond formation

$$|+\rangle = \frac{1}{\sqrt{2}}\Big(|i\alpha\rangle + |j\beta\rangle\Big) \quad \text{bonding state} \, , \tag{91}$$

$$|-\rangle = \frac{1}{\sqrt{2}}\Big(|i\alpha\rangle - |j\beta\rangle\Big) \quad \text{anti-bonding state} \, . \tag{92}$$

The density matrix may then be obtained from the difference of the number of electrons in bonding and anti-bonding states

$$\rho_{i\alpha j\beta} = \langle i\alpha|\hat{\rho}|j\beta\rangle = \frac{1}{2}\big(N_+ - N_-\big) \, , \tag{93}$$

with $N_+ = \langle +|\hat{\rho}|+\rangle$ and $N_- = \langle -|\hat{\rho}|-\rangle$ and $N_{i\alpha i\alpha} + N_{j\beta j\beta} = (N_+ + N_-)$ the number of electrons in the bond. If we assume a non-magnetic calculation and take into account spin degeneracy, then $0 \leq N_{i\alpha i\alpha} \leq 2$ and the same for $N_{j\beta j\beta}$. This allows one to put limits on the density matrix

$$|\rho_{i\alpha j\beta}| \leq N_{i\alpha j\beta} \quad \text{and} \quad |\rho_{i\alpha j\beta}| \leq 2 - N_{i\alpha j\beta} \, , \tag{94}$$

with $N_{i\alpha j\beta} = (N_{i\alpha i\alpha} + N_{j\beta j\beta})/2$. Fig. 5 shows the density matrix for close packed transition metals as a function of band filling.

3.5.2 Structural stability

The analytic BOPs may be employed for the analysis of the structural stability of different phases. To this end one compares the structures at the same repulsive energy, following the structural energy difference theorem [44]. Fig. 6 shows the structural energy differences for a number of close packed phases.

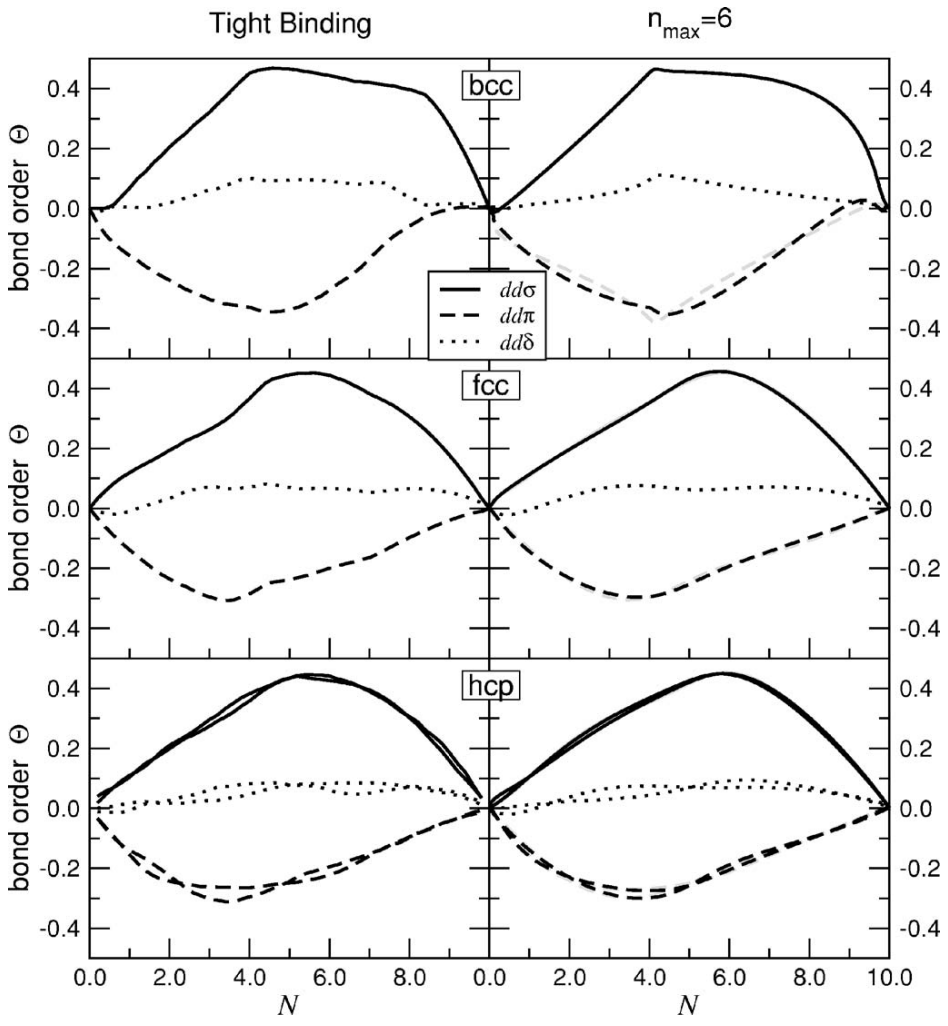

Fig. 5: *Bond order as a function of band filling for close-packed transition metals. The TB approximation is compared to BOPs. Taken from Ref. [13].*

3.5.3 Ti phase diagram

Fig. 7 shows the free energy differences between competing phases in Ti. The phase diagram predicted from analytic BOPs is in very good agreement with experiment and DFT, despite the fact that the energy differences between the competing phases are of the order of only a few meV.

3.5.4 Parametrization

For the parametrization of the analytic BOPs, first the Hamiltonian matrix elements are obtained from downfolding on DFT wavefunctions [50, 51], see Figs. 1 and 2. The Hamiltonian is then parametrized and together with the repulsive energy fitted to reproduce DFT reference data. Software that largely automatizes the parametrization procedure is available [68, 71].

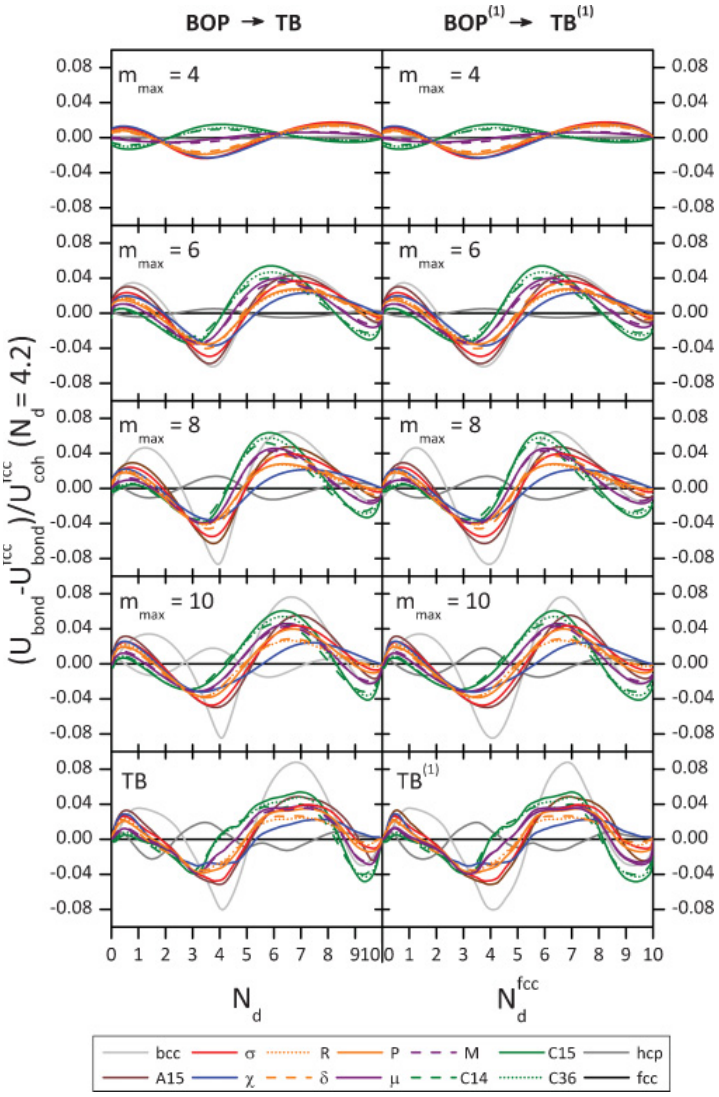

Fig. 6: *Structural energy differences for a number of close-packed phases (left panel).* m_{max} *indicates the moment at which the expansion was terminated, the lowest panels show the TB reference. The right panel shows a first order expansion. For details Ref. [69], from which this figure was also taken.*

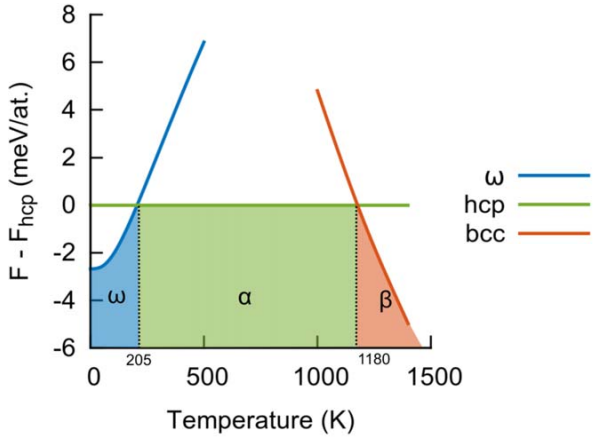

Fig. 7: *Free energy differences as computed for Ti with analytic BOPs. Taken from Ref. [70].*

4 Many atom expansions

In the first part of this chapter I discussed the derivation of simplified electronic structure models from DFT. A local expansion of the TB energy then led to explicit interatomic potentials. While a lot of insight can be gained from the analysis of bond formation in the BOPs and robust parametrizations may be achieved with few parameters, the accuracy and transferability of the BOPs are also limited by the coarse-graining approximations from DFT to TB and BOP.

In the past years an important focus in the field of atomistic modeling was the parametrization of DFT reference data with very high accuracy, i.e., errors of less than a few meV. The TB and BOP expansions introduced in the previous sections are not competitive here, as the approximations made down the coarse-graining hierarchy from DFT to TB to BOPs introduce errors that are larger than a few meV. As I will discuss in the following, one can develop models that incorporate some of the spirit of TB for obtaining meV accurate parametrizations of DFT reference data.

For the parametrization of large DFT datasets with very high accuracy typically methods that are rooted in machine-learning are employed, for example, neural network potentials that are based on neural networks [15] or Gaussian process regression in the Gaussian approximation potentials [16]. All machine-learning methods have in common that the target property, for example, the atomic energy, is obtained as a complex, non-linear function of some mathematical descriptions of the local atomic environment. The detailed mathematical structure of the descriptors are mostly obtained by intuition. One may view the empirical mathematical structure of the descriptors as the Achilles' heel of machine-learning interatomic potentials and a formally complete descriptor of the local atomic environment is desirable. The atomic cluster expansion [34, 37] achieves a formally complete description of the local atomic environment and will be introduced in the following.

4.1 Atomic cluster expansion

The atomic cluster expansion (ACE) provides a complete descriptor for the local environment of an atom [34, 37]. Each atom i has a configuration that in the simplest case of an elemental material and excluding charge transfer or magnetism is fully characterized by the distance vectors to all neighboring atoms j, $\boldsymbol{r}_{ji} = \boldsymbol{r}_i - \boldsymbol{r}_j$, and where $\boldsymbol{r}_i$ and $\boldsymbol{r}_j$ are the positions of the atoms i and j, respectively. The collection of all the distance vectors is the configuration

$$\boldsymbol{\sigma} = (\boldsymbol{r}_{1i}, \boldsymbol{r}_{2i}, \boldsymbol{r}_{3i}, \dots)\,. \tag{95}$$

A scalar product between functions $f(\boldsymbol{\sigma})$ and $g(\boldsymbol{\sigma})$ is defined as

$$\langle f | g \rangle = \int d\boldsymbol{\sigma}\; f^*(\boldsymbol{\sigma})\, g(\boldsymbol{\sigma})\, \omega(\boldsymbol{\sigma})\,, \tag{96}$$

where $\omega(\boldsymbol{\sigma})$ is a weight function. Next complete basis functions that depend only on a single bond are introduced

$$\langle \phi_v(\sigma) | \phi_u(\sigma) \rangle = \delta_{vu} \,, \tag{97}$$

$$\sum_v [\phi_v(\boldsymbol{r})]^* \phi_v(\boldsymbol{r}') = \delta(\boldsymbol{r} - \boldsymbol{r}') \,. \tag{98}$$

For establishing a hierarchical expansion furthermore $\phi_0 = 1$, which may be understood as an atom without properties, i.e., the vacuum state.
A cluster α with K elements contains K atoms $\alpha = (j_1, j_2, \ldots, j_K)$, where the order of entries in α does not matter and indices are pairwise different $i \neq j_1 \neq j_2 \neq j_K$. The vector $\nu = (v_0; v_1, v_2, \ldots, v_K)$ contains the list of single-atom basis functions in the cluster, and only single-atom basis functions with $v > 0$ are considered in ν. A cluster basis function is then given by

$$\Phi_{\alpha\nu} = \phi_{v_1}(\boldsymbol{r}_{j_1 i}) \, \phi_{v_2}(\boldsymbol{r}_{j_2 i}) \ldots \phi_{v_K}(\boldsymbol{r}_{j_K i}) \,. \tag{99}$$

The orthogonality and completeness of the single-atom basis functions transfers to the cluster basis functions

$$\langle \Phi_{\alpha\nu} | \Phi_{\beta\mu} \rangle = \delta_{\alpha\beta} \delta_{\nu\mu} \,, \tag{100}$$

$$1 + \sum_{\gamma \subseteq \alpha} \sum_{\nu} [\Phi_{\gamma\nu}(\boldsymbol{\sigma})]^* \Phi_{\gamma\nu}(\boldsymbol{\sigma}') = \delta(\boldsymbol{\sigma} - \boldsymbol{\sigma}') \,, \tag{101}$$

where α is an arbitrary cluster and the right-hand side of the completeness relation is the product of the relevant right-hand sides of Eq. (98). The expansion of an element of the energy of atom i may therefore be written in the form

$$E_i = J_0 + \sum_{\alpha\nu} J_{\alpha\nu} \Phi_{\alpha\nu}(\boldsymbol{\sigma}) \,, \tag{102}$$

and the expansion coefficients $J_{\alpha\nu}$ obtained by projection

$$J_{\alpha\nu} = \langle \Phi_{\alpha\nu} | G(\boldsymbol{\sigma}) \rangle \,. \tag{103}$$

Writing the expansion Eq. (102) explicitly in single-atom basis functions leads to

$$\begin{aligned} E_i = & \sum_{j}^{i \neq j} \sum_{v_1} J_{v_1} \phi_{v_1}(\boldsymbol{r}_{ji}) + \frac{1}{2} \sum_{j_1 j_2}^{i \neq j_1 \neq j_2} \sum_{v_1 v_2} J_{v_1 v_2} \phi_{v_1}(\boldsymbol{r}_{j_1 i}) \phi_{v_2}(\boldsymbol{r}_{j_2 i}) \\ & + \frac{1}{3!} \sum_{j_1 j_2 j_3}^{i \neq j_1 \neq j_2 \ldots} \sum_{v_1 v_2 v_3} J_{v_1 v_2 v_3} \phi_{v_1}(\boldsymbol{r}_{j_1 i}) \phi_{v_2}(\boldsymbol{r}_{j_2 i}) \phi_{v_3}(\boldsymbol{r}_{j_3 i}) + \ldots \,. \end{aligned} \tag{104}$$

This may be rewritten in a slightly different way with unrestricted sums and updated expansion coefficients

$$\begin{aligned} E_i = & \sum_{j} \sum_{v_1} c_{v_1} \phi_v(\boldsymbol{r}_{ji}) + \frac{1}{2} \sum_{j_1 j_2} \sum_{v_1 v_2} c_{v_1 v_2} \phi_{v_1}(\boldsymbol{r}_{j_1 i}) \phi_{v_2}(\boldsymbol{r}_{j_2 i}) \\ & + \frac{1}{3!} \sum_{j_1 j_2 j_3} \sum_{v_1 v_2 v_3} c_{v_1 v_2 v_3} \phi_{v_1}(\boldsymbol{r}_{j_1 i}) \phi_{v_2}(\boldsymbol{r}_{j_2 i}) \phi_{v_3}(\boldsymbol{r}_{j_3 i}) + \ldots \,, \end{aligned} \tag{105}$$

where only i is excluded from the summations over $j_1, j_2, \ldots$. The expansion Eq. (105) is identical to Eq. (104), with expansion coefficients c_ν that are different from the expansion coefficients J_ν in Eq. (104). The expansion coefficients c_ν are simple functions of J_ν that may be obtained by taking into account that products of basis functions of the same argument may be expanded into linear combinations of single basis functions, for example, $\phi_{v_1}(\boldsymbol{r}_{ji})\,\phi_{v_2}(\boldsymbol{r}_{ji}) = \sum_v a_v \phi_v(\boldsymbol{r}_{ji})$, *etc.*, such that the self-interactions are removed by an appropriate modification of a lower-order expansion coefficient. The relation between J_ν and c_ν is also outlined in Ref. [72]. I next introduce the atomic density

$$\rho_i(\sigma) = \sum_j^{j\neq i} \delta(\sigma - \sigma_j)\,, \tag{106}$$

and the atomic base that is obtained as

$$A_v = \langle \rho_i | \phi_v \rangle = \sum_j^{j\neq i} \phi_v(\boldsymbol{r}_j)\,. \tag{107}$$

This allows us to rewrite the expansion Eq. (105) in the form

$$E_i = \sum_\nu c_\nu \boldsymbol{A}_\nu \quad \text{with} \quad \boldsymbol{A}_\nu = A_{v_1} A_{v_2} A_{v_3} \ldots \tag{108}$$

where the index ν collects the required indices $v_1, v_2, v_3, \ldots$ from Eq. (105).
By construction the expansion is invariant with respect to permutation of identical atoms. The change from Eq. (104) to Eq. (105) further means that the atomic expectation values of the many-atom correlation functions $\Phi_{\alpha\nu}$ may be expressed exactly by products of expectation values of single-bond basis functions. This enables a very efficient implementation as the effort for evaluating the many-atom correlation functions scales linearly with the number of neighbors irrespective of the order of the correlation functions.
The expansion Eq. (107) is general, which also means that in general it is not invariant under rotation. Rotationally invariant expansions may be obtained as outlined in the following. First, the single-bond basis functions are chosen as basis functions of the irreducible representations of the rotation group. In practice, this corresponds to linear combination of atomic orbitals (LCAO) basis functions as in TB, Eq. (26),

$$\phi_{inlm}(\boldsymbol{r}) = R_{nl}(|\boldsymbol{r} - \boldsymbol{r}_i|) Y_l^m(\theta, \phi)\,. \tag{109}$$

Rotationally invariant products are obtained with the help of generalized Clebsch Gordan coefficients

$$B_\nu = \sum_{\boldsymbol{m}} \begin{pmatrix} \boldsymbol{l} \\ \boldsymbol{L} \end{pmatrix}_N A_{n_1 l_1 m_1} A_{n_2 l_2 m_3} A_{n_3 l_2 m_3} \ldots\,, \tag{110}$$

with $\boldsymbol{l} = (l_1, l_2, l_3, \ldots)$, $\boldsymbol{m} = (m_1, m_2, m_3, \ldots)$ and $\boldsymbol{L}$ are intermediate angular momenta that arise from products of the spherical harmonics [34, 37, 72–74]. The expansion for the energy may then be represented as

$$E_i = \sum_\nu c_\nu \boldsymbol{B}_\nu\,. \tag{111}$$

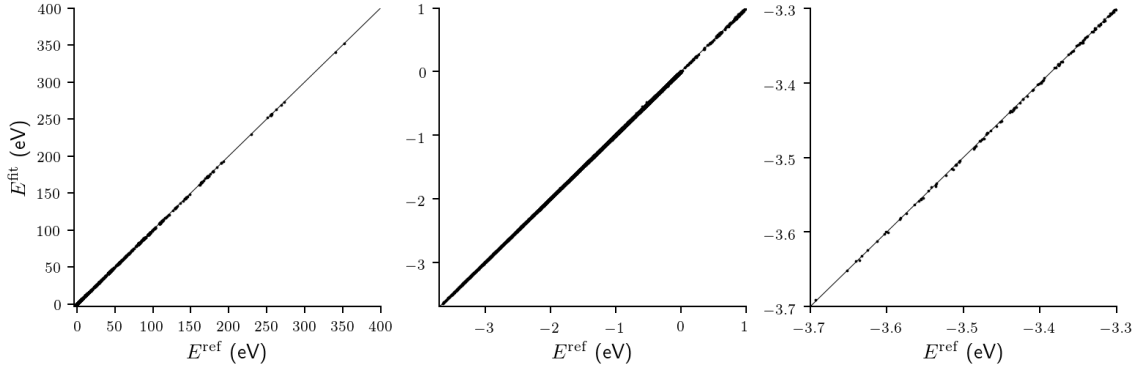

Fig. 8: *Comparison of predictions from ACE to DFT reference. The energies are compared over three orders of magnitude. The graph shows many thousand data points and outliers are visible in particular. Taken from Ref. [34].*

4.2 Relation to other descriptors

The completeness of the ACE allows one to make contact with other frequently used descriptors. Here I list a few popular descriptors and machine-learning potentials that may be rewritten in the form of an ACE [34, 37]:

- Steinhardt parameters: the Steinhardt parameters [75] are frequently employed for structure classification.
- Symmetry functions: neural network potentials use 2-body and 3-body functions, called symmetry functions, as descriptors for the atomic environment [15].
- Smooth overlap of atomic positions (SOAP): the SOAP descriptor [76] is employed, for example with the Gaussian approximation potential (GAP) [16]. A tensorial version of SOAP is also available [27].
- Spectral neighbor analysis potential (SNAP): the SNAP employs the SOAP descriptor with hyperspherical harmonics [21].
- Moment tensor potentials (MTP): the MTPs [22] provide an expansion of the interatomic energy in terms of Cartesian tensors.

4.3 Parametrization

Compared to TB and BOP models, the ACE provides less physical insight but greater flexibility for the accurate parametrization of arbitrary interatomic interactions. This means that a larger number of reference data needs to be available for fitting the ACE expansion coefficients c_ν. Fig. 8 shows the comparison of the ACE predictions for copper compared to DFT reference data. More than 50000 DFT total energy calculations were used for the parametrization of the ACE. The reference data set comprises many different bulk crystal structures, with and without defects, at various volumes and deformations in addition to small clusters with 2 to 25 Cu atoms.

Knowledge and insight gained from TB and BOP can be used for the parametrization of ACE. For the Cu example, the ACE was built on the Finnis Sinclair potential Eq. (1). However, instead of computing the density ρ_i and the repulsion $\sum_j V_{ij}/2$ from pairwise functions, density and repulsion were represented by a general ACE expansion. The non-linear square-root embedding function helps to converge the ACE faster as compared to a linear expansion as the non-linear dependence of the bond energy on coordination is immediately taken into account. Furthermore, the radial functions R_{nl} may be related to the minimal basis radial functions of TB models.

5 Summary and conclusions

This chapter exemplifies two strategies for obtaining interatomic potentials in materials science. The derivation of BOPs from DFT encompasses a systematic coarse-graining of the electronic structure that provides insight into bond formation and is amenable to physical and chemical interpretation. In contrast, ACE is a formal many-atom expansion that is flexible to model bond formation accurately but with many fitting parameters, so that large numbers of DFT reference data are required.

The TB approximation is obtained from DFT as a second-order expansion with respect to the density matrix of overlapping spherical atomic charge densities. A physically transparent mechanism of bond formation is obtained by grouping the terms in the expansion in bond energy and electrostatic interactions that drive the formation of covalent, polar and ionic bonds, the charging of the atoms and the promotion of electrons out of their atomic state, and a repulsive energy that keeps the atoms apart.

The moments theorem allows one to relate the atomic structure to the electronic structure. I introduced several methods that implicitly or explicitly make use of the moments theorem to reconstruct the bond energy from the local atomic environment. The recursion method first transforms the Hamiltonian to tridiagonal form, which enables a continued fraction representation of the Green function. The numerical BOPs make use of the recursion method for a local construction of the bond energy. The KPM expands the kernel into Chebyshev polynomials for a local expansion of the bond energy. Closely related is the FOE that starts from the expansion of a temperature dependent electronic broadening function. Finally, the analytic BOPs provide explicit analytic expressions for energies and forces that may also be used to analyze bond strengths and structural stability. The analytic BOPs reproduce the DFT energies very well and enable the simulations of phase diagrams in good agreement with experiment. Software for simulations with millions and for the efficient parametrization of new models is available.

The ACE provides a formal many-atom expansion that enables parametrizations with arbitrary accuracy given that sufficient reference data is available. The completeness of the ACE also means that other descriptors and machine-learning potentials may be represented in the form of an ACE. The parametrization of the ACE was discussed briefly and illustrated for copper.

Acknowledgments

I thank Isabel Pietka for carefully reading the manuscript and helpful comments and corrections.

References

[1] J. Friedel in J.M. Ziman (Ed.): *Electrons* (Pergamon, London, 1969), *Physics of Metals*, Vol. 1

[2] F. Ducastelle, J. Phys. (Paris) **31**, 1055 (1970)

[3] R.P. Gupta, Phys. Rev. B **23**, 6265 (1981)

[4] M.W. Finnis and J.E. Sinclair, Philos. Mag. A **50**, 45 (1984)

[5] J.K. Nørskov and N.D. Lang, Phys. Rev. B **21**, 2131 (1980)

[6] M.J. Puska, R.M. Nieminen, and M. Manninen, Phys. Rev. B **24**, 3037 (1981)

[7] M.S. Daw and M.I. Baskes, Phys. Rev. B **29**, 6443 (1984)

[8] J. Tersoff, Phys. Rev. Lett. **56**, 632 (1986)

[9] J. Tersoff, Phys. Rev. B **38**, 9902 (1988)

[10] M.I. Baskes, Phys. Rev. B **46**, 2727 (1992)

[11] D.G. Pettifor and I.I. Oleinik, Phys. Rev. Lett. **84**, 4124 (2000)

[12] D.G. Pettifor and I.I. Oleinik, Phys. Rev. B **65**, 172103 (2002)

[13] R. Drautz and D.G. Pettifor, Phys. Rev. B **74**, 174117 (2006)

[14] R. Drautz and D.G. Pettifor, Phys. Rev. B **84**, 214114 (2011)

[15] J. Behler and M. Parrinello, Phys. Rev. Lett. **98**, 146401 (2007)

[16] A.P. Bartók, M.C. Payne, R. Kondor, and G. Csányi, Phys. Rev. Lett. **104**, 136403 (2010)

[17] S. Manzhos and T. Carrington, Jr., J. Chem. Phys. **125**, 084109 (2006)

[18] M. Rupp, A. Tkatchenko, K.-R. Müller, and O.A. von Lilienfeld, Phys. Rev. Lett **108**, 058301 (2012)

[19] Z. Li, J.R. Kermode, and A. DeVita, Phys. Rev. Lett. **114**, 096405 (2015)

[20] J. Behler, Int. J. Quantum Chem. **115**, 1032 (2015)

[21] A. Thompson, L. Swiler, C. Trott, S. Foiles, and G. Tucker, J. Comp. Phys. **285**, 316 (2015)

[22] A.V. Shapeev, Multiscale Model. Simul. **14**, 1153 (2016)

[23] N. Artrith, A. Urban, and G. Ceder, Phys. Rev. B **96**, 014112 (2017)

[24] A. Takahashi, A. Seko, and I. Tanaka, Phys. Rev. Materials **1**, 063801 (2017)

[25] T.D. Huan, R. Batra, J. Chapman, S. Krishnan, L. Chen, and R. Ramprasad, NPJ Comput. Mater. **3**, 37 (2017)

[26] S. Chmiela, A. Tkatchenko, H.E. Sauceda, I. Poltavsky, K.T. Schütt, and K.-R. Müller, Science Advances **3**, e1603015 (2017)

[27] A. Grisafi, D.M. Wilkins, G. Csányi, and M. Ceriotti, Phys. Rev. Lett. **120**, 036002 (2018)

[28] M.A. Wood and A.P. Thompson, J. Chem. Phys. **148**, 241721 (2018)

[29] T.T. Nguyen, E. Székely, G. Imbalzano, J. Behler, G. Csányi, M. Ceriotti, A.W. Götz, and F. Paesani, J. Chem. Phys. **148**, 241725 (2018)

[30] D. Dragoni, T.D. Daff, G. Csányi, and N. Marzari, Phys. Rev. Materials **2**, 013808 (2018)

[31] T. Bereau, R.A. DiStasio Jr., A. Tkatchenko, and O.A. von Lilienfeld, J. Chem. Phys. **148**, 241706 (2018)

[32] A. Kamath, R.A. Vargas-Hernández, R.V. Krems, T. Carrington Jr., and S. Manzhos, J. Chem. Phys. **148**, 241702 (2018)

[33] M. Rupp, O.A. von Lilienfeld, and K. Burke, J. Chem. Phys. **148**, 241401 (2018)

[34] R. Drautz, Phys. Rev. B **99**, 014104 (2019)

[35] M.J. Willatt, F. Musil, and M. Ceriotti, J. Chem. Phys. **150**, 154110 (2019)

[36] C. van der Oord, G. Dusson, G. Csányi, and C. Ortner, Mach. Learn.: Sci. Technol. **1**, 015004 (2020)

[37] R. Drautz, Phys. Rev. B **102**, 024102 (2020)

[38] R. Drautz, T. Hammerschmidt, M. Cak, and D.G. Pettifor, Modelling Simul. Mater. Sci. Eng. **23**, 074004 (2015)

[39] J. Harris, Phys. Rev. B **31**, 1770 (1985)

[40] W.M.C. Foulkes and R. Haydock, Phys. Rev. B **39**, 12520 (1989)

[41] A.P. Sutton, M.W. Finnis, D.G. Pettifor, and Y. Ohta, J. Phys. C **21**, 35 (1988)

[42] M. Elstner, D. Porezag, G. Jungnickel, J. Elsner, M. Haugk, T. Frauenheim, S. Suhai, and G. Seifert, Phys. Rev. B **58**, 7260 (1998)

[43] M.W. Finnis: *Interatomic Forces in Condensed Matter* (Oxford University Press, 2003)

[44] D.G. Pettifor: *Bonding and Structure in Molecules and Solids* (Oxford University Press, 1995)

[45] A. Sutton: *Electronic Structure of materials* (Clarendon Press, Oxford, 1993)

[46] A.T. Paxton in [77], p. 145

[47] T. Hammerschmidt and R. Drautz in [77], p. 229

[48] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964)

[49] W. Kohn and L.J. Sham, Phys. Rev. **140**, A1133 (1965)

[50] G.K.H. Madsen, E.J. McEniry, and R. Drautz, Phys. Rev. B **83**, 184119 (2011)

[51] A. Urban, M. Reese, M. Mrovec, C. Elsässer, and B. Meyer,
Phys. Rev. B **84**, 155119 (2011)

[52] E.J. McEniry, R. Drautz, and G.K.H. Madsen,
J. Phys.: Condens. Matter **25**, 115502 (2013)

[53] C. Lanczos, J. Res. Natl. Bur. Stand. **45**, 225 (1950)

[54] R. Haydock in H. Ehrenreich, F. Seitz, and D. Turnbull (eds.)
Solid State Physics (Academic Press, New York, 1980) Vol. 35, p. 215

[55] M. Aoki, Phys. Rev. Lett. **71**, 3842 (1993)

[56] A.P. Horsfield, A.M. Bratkovsky, D.G. Pettifor, and M. Aoki, Phys. Rev. B **53**, 1656 (1996)

[57] A.P. Horsfield, A.M. Bratkovsky, M. Fearn, D.G. Pettifor, and M. Aoki,
Phys. Rev. B **53**, 12694 (1996)

[58] M. Aoki, D. Nguyen-Manh, D.G. Pettifor, and V. Vitek, Prog. Mat. Sci. **52**, 154 (2007)

[59] R.N. Silver, H. Röder, A.F. Voter, and J.D. Kress, J. Comp. Phys. **124**, 115 (1996)

[60] A.F. Voter, J.D. Kress, and R.N. Silver, Phys. Rev. B **53**, 12733 (1996)

[61] A. Weiße, G. Wellein, A. Alvermann, and H. Fehske, Rev. Mod. Phys. **78**, 275 (2006)

[62] B. Seiser, D.G. Pettifor, and R. Drautz, Phys. Rev. B **87**, 094105 (2013)

[63] R.N. Silver and H. Röder, Phys. Rev. E **56**, 4822 (1997)

[64] S. Goedecker and M. Teter, Phys. Rev. B **51**, 9455 (1995)

[65] S. Goedecker, Rev. Mod. Phys. **71**, 1085 (1999)

[66] R.N. Silver and H. Röder, Int. J. Mod. Phys. C **5**, 735 (1994)

[67] M.E. Ford, D.G. Pettifor, and R. Drautz, J. Phys.: Condens. Matter **27**, 086002 (2015)

[68] T. Hammerschmidt, B. Seiser, M.E. Ford, A.N. Ladines, S. Schreiber, N. Wang, J. Jenke, Y. Lysogorskiy, C. Teijeiro, M. Mrovec, M. Cak, E.R. Margine, D.G. Pettifor, and R. Drautz, Comput. Phys. Commun. **235**, 221 (2019)

[69] B. Seiser, T. Hammerschmidt, A.N. Kolmogorov, R. Drautz, and D.G. Pettifor, Phys. Rev. B **83**, 224116 (2011)

[70] A. Ferrari, M. Schröder, Y. Lysogorskiy, J. Rogal, M. Mrovec, and R. Drautz, Modelling Simul. Mater. Sci. Eng. **27**, 085008 (2019)

[71] A. Ladines, T. Hammerschmidt, and R. Drautz, Comp. Mater. Sci. **173**, 109455 (2020)

[72] G. Dusson, M. Bachmayr, G. Csányi, R. Drautz, S. Etter, C. van der Oord, and C. Ortner, in preparation (2020)

[73] A.P. Yutsis, I.B. Levinson, and V.V. Vanagas: *The Theory of Angular Momentum* (Israel Program for Scientific Translations, Jerusalem, 1962)

[74] D.M. Brink and G.R. Satchler: *Angular Momentum* (Clarendon Press, Oxford, 1968)

[75] P.J. Steinhardt, D.R. Nelson, and M. Ronchetti, Phys. Rev. B **28**, 784 (1983)

[76] A.P. Bartók, R. Kondor, and G. Csányi, Phys. Rev. B **87**, 184115 (2013)

[77] J. Grotendorst, N. Attig, S. Blügel, and D. Marx (Eds.): *Multiscale Simulation Methods in Molecular Sciences* NIC Series Vol. 42 (Forschungszentrum Jülich, 2009)