

基于数值原子轨道基组的第一性原理计算软件ABACUS

刘晓辉 陈默涵 李鹏飞 沈瑜 任新国 郭光灿 何力新

Introduction to first-principles simulation package ABACUS based on systematically improvable atomic orbitals

Liu Xiao-Hui Chen Mo-Han Li Peng-Fei Shen Yu Ren Xin-Guo Guo Guang-Can He Li-Xin

引用信息 Citation: *Acta Physica Sinica*, 64, 187104 (2015) DOI: 10.7498/aps.64.187104

在线阅读 View online: <http://dx.doi.org/10.7498/aps.64.187104>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn/CN/Y2015/V64/I18>

您可能感兴趣的其他文章

Articles you may be interested in

金衬底调控单层二硫化钼电子性能的第一性原理研究

Tuning the electronic property of monolayer MoS₂ adsorbed on metal Au substrate: a first-principles study

物理学报.2015, 64(18): 187101 <http://dx.doi.org/10.7498/aps.64.187101>

GGA+U 的方法研究 Ag 掺杂浓度对 ZnO 带隙和吸收光谱的影响

GGA+U study on the effects of Ag doping on the electronic structures and absorption spectra of ZnO

物理学报.2015, 64(15): 157101 <http://dx.doi.org/10.7498/aps.64.157101>

N-F 共掺杂锐钛矿二氧化钛(101)面纳米管的第一性原理研究

N-F co-doped in titanium dioxide nanotube of the anatase (101) surface: a first-principles study

物理学报.2015, 64(14): 147103 <http://dx.doi.org/10.7498/aps.64.147103>

第一性原理研究稀土掺杂 ZnO 结构的光电性质

First principles study on the electronic and optical properties of ZnO doped with rare earth

物理学报.2013, 62(4): 047101 <http://dx.doi.org/10.7498/aps.62.047101>

虚晶近似法研究 AlN-Al₂O₃ 固溶体系的力学性能和电子结构

Investigation of the electronic and mechanical properties of Al₂O₃-AlN solid solution by virtual crystal approximation

物理学报.2012, 61(16): 167101 <http://dx.doi.org/10.7498/aps.61.167101>

专题: 计算物理的近期发展和应用

基于数值原子轨道基组的第一性原理 计算软件 ABACUS*

刘晓辉¹⁾²⁾ 陈默涵³⁾ 李鹏飞¹⁾²⁾ 沈瑜⁴⁾ 任新国^{1)2)†}
郭光灿¹⁾²⁾ 何力新^{1)2)‡}

1) (中国科学技术大学, 中国科学院量子信息重点实验室, 合肥 230026)

2) (中国科学技术大学, 量子信息与量子科技前沿协同创新中心, 合肥 230026)

3) (普林斯顿大学机械和航空航天工程系, 美国新泽西州 08544)

4) (中国科学技术大学, 网络信息中心, 超级计算中心, 合肥 230026)

(2015年7月2日收到; 2015年9月3日收到修改稿)

随着超级计算机硬件和数值算法迅速发展, 使得目前利用密度泛函理论研究上千个原子体系的电子能带和结构等性质变得可行. 数值原子轨道基组由于其基组较小和局域等特性, 可以很好地与电子结构计算中的线性标度算法等的新算法结合, 用来研究较大尺寸的物理体系. 本文详细介绍了一款中国科学技术大学量子信息重点实验室自主开发的基于数值原子轨道基组的第一性原理计算软件 Atomic-orbital Based *Ab-initio* Computation at UStc. 大量的测试结果表明: 该软件具有很好的准确性和较高的并行效率, 可以用于包含 1000 个原子左右的系统的电子结构和原子结构的研究以及分子动力学模拟计算.

关键词: 第一性原理, 数值原子轨道, 计算软件

PACS: 71.15.Ap, 71.15.Mb

DOI: 10.7498/aps.64.187104

1 引言

基于密度泛函理论^[1] (density functional theory, DFT) 的第一性原理方法在凝聚态物理、材料科学、化学和生物等研究领域变得越来越重要. 随着超级计算机的迅速发展和数值算法的提高, 利用密度泛函理论研究包含上千个原子的复杂体系的电子能带、结构和动力学性质变得可行. 对于这些体系, 之前广泛使用的平面波基组方法由于自身扩展性问题, 效率将大大降低; 而此时的局域基组方法, 如原子轨道方法, 将是更好的选择.

对于 Kohn-Sham 理论^[2] 下的第一性电子结构计算, 选取原子轨道作为基组有几个优点: 首

先, 原子轨道基组的数目远远小于平面波或者实空间格点的基组数目; 其次, 原子轨道是严格局域的, 因此可以很好地与电子计算中的线性标度算法^[3] 或者其他优于 $O(N^3)$ 的算法结合起来. 例如, Lin 等^[4,5] 最近发展的极点分解和选择性求逆方法 (pole expansion and selected inversion), 利用局域基组下哈密顿矩阵的稀疏性, 可以使求解 Kohn-Sham 方程的标度达到 N^α ($\alpha \leq 2$), 其中 N 是体系中的原子数目; 第三, 原子轨道具有球对称性, 这对于有些物理性质的计算很重要, 比如在研究铁电材料时, 要得到正确的结果, 基组的对称性必须要高于晶格的对称性.

在量子化学界, 解析的高斯轨道已经很好地

* 科技部重大研究计划 (批准号: 2011CB921200)、国家自然科学基金 (批准号: 11374275, 11374276) 和中国科学院战略性先导科技专项 (B类) (批准号: XDB01030100).

† 通信作者. E-mail: renxg@ustc.edu.cn

‡ 通信作者. E-mail: helx@ustc.edu.cn

发展了几十年,而数值原子基组在计算凝聚态物理领域也使用得越来越普遍.最近几年,国外开发了几款基于原子轨道基组的第一性原理软件,如 SIESTA [6], Openmx [7], FHI-aims [8] 等. 这些软件都是为了做大规模 DFT 计算而设计的. 但是与平面波算法不同,原子基组的构造必须很好地保证基组的精确性和可移植性,并且基组的质量也应该是可以系统性提高的. Chen, Guo, 和 He (CGH) 提出了一个方案 [9,10], 可以为 DFT 计算构建可系统性提高的、全面优化的原子轨道基组. 在 CGH 基组生成的方案的基础上,我们在中国科学技术大学量子信息重点实验室,从零开始开发了一款第一性原理计算软件包,命名为“Atomic-orbital Based *Ab-initio* Computation at UStc (ABACUS)” [11]. 这款软件以数值原子轨道基组作为优先选择,同时也提供了平面波基组的方法. 双重基组的特性对检查原子轨道计算结果的准确性非常有帮助. 在这两种方法中,我们都使用了 Unified Pseudopotential Format (UPF) 格式的模守恒赝势 [12]. UPF 格式是另一款第一性原理计算软件 Quantum Espresso [13] 赝势的标准格式. 这种格式赝势可以有多种来源,如利用赝势生成软件 Opium [14] 得到. 目前在 ABACUS 实现了多种密度泛函,其中最主要的包括局域(自旋)密度近似(L(S)DA)和广义梯度近似(GGA). GGA 采用的是 Perdew, Burke, 和 Ernzerhof (PBE) 提出的 GGA-PBE 泛函 [15]. 此外, ABACUS 实现了由 Grimme 提出的 DFT 半经验范德华(van der Waals, vdW)修正(DFT-D2) [16]. 同时,我们还在发展更高级的泛函,比如杂化泛函 [17] 等. 这些结果将在后续的工作中报导. 在 LDA 和 GGA 层面上, ABACUS 可以完成常规的电子结构计算、结构优化和分子动力学计算.

在这篇文章中,我们将介绍软件的主要特色和已经实现的功能,简单介绍软件的核心算法;最后,给出一些软件使用的例子.

2 ABACUS 程序简介

我们小组自 2006 年起,从零开始发展了第一性原理软件 ABACUS. 我们从发展成熟的平面波算法入手,借助平面波的计算结果来生成数值原子轨道,然后在此基础上发展了基于数值原子轨道的程序,可以用来研究较大的物理体系. ABACUS

主要由 C++ 语言编写而成,现已发展成为一套以数值原子轨道或平面波为基矢的、高度并行的、基本功能齐全的第一性原理软件包. 截至目前共有 117538 行源代码. 另外, ABACUS 包含了一套可以用来生成数值原子轨道的程序,共有 13611 行源代码.

目前 ABACUS 的主要功能包括: 可以进行电子结构自洽迭代(包括加入共线电子自旋)计算、原子结构优化以及分子动力学计算; 密度泛函上可以选择 LDA 或者 GGA, 还可以进行包含范德华力修正的 DFT-D2 及加外电场的计算. ABACUS 采用模守恒赝势和周期性边界条件,可以对晶格对称性、布里渊区的 k 点对称性、电荷密度对称性以及力的对称性进行分析,可以处理只含 Γ 点或多 k 点的情况. 求解 Kohn-Sham 方程时采用的是共轭梯度算法(平面波基组)或者直接对角化方法(数值原子轨道基组); 电荷密度混合使用的主要是 Pulay mixing 和 Pulay-Kerker mixing 等算法.

3 程序核心算法介绍

3.1 基于数值原子轨道的 Kohn-Sham 方程

3.1.1 算法简介

在第一性原理计算领域,平面波基组方法已经比较成熟,有多款应用广泛的软件. 但是在处理较大尺寸、含有原子数较多的体系时,平面波算法收敛较慢,效率不高. 而使用数值原子轨道基组,由于基组数较小以及轨道在实空间局域的特点,可以大幅提高计算速度;另一方面,数值原子轨道也适合于线性标度的算法 [3], 所以受到越来越多的关注. 一般来说,一个原子的数值原子轨道可以写成径向轨道与球谐函数的乘积的形式,即

$$\phi_{\mu}(\mathbf{r}) = f_{nl\zeta}(r)Y_{lm}(\hat{\mathbf{r}}), \quad (1)$$

其中 $\mu = \{I, l, m, \zeta\}$, I 是原胞中的原子指标, l 是轨道角动量, m 是磁量子数, ζ 是给定 l 的数值轨道的数目. 系统的电子本征波函数可以表示为数值原子轨道的线性叠加,即

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} \sum_{\mu} c_{n\mu,\mathbf{k}} e^{i\mathbf{k}\cdot\mathbf{R}} \phi_{\mu}(\mathbf{r} - \tau_I - \mathbf{R}), \quad (2)$$

其中 n 是电子能带指标, \mathbf{k} 是布洛赫波矢, N 是玻恩-冯·卡门周期性边界条件下超胞中原胞的数目.

$\phi_\mu(\mathbf{r} - \tau_I - \mathbf{R})$ 是原胞 \mathbf{R} 中第 I 个原子的原子轨道。将 (2) 式代入 Kohn-Sham 方程,

$$\left\{ -\frac{\hbar^2}{2m}\nabla^2 + V_{\text{ext}}(\mathbf{r}) + V_{\text{H}}[\rho](\mathbf{r}) + V_{\text{xc}}[\rho](\mathbf{r}) \right\} \psi_{n\mathbf{k}}(\mathbf{r}) = \varepsilon_{n\mathbf{k}} \psi_{n\mathbf{k}}(\mathbf{r}), \quad (3)$$

可以得到矩阵方程

$$(\mathbf{H} - \varepsilon\mathbf{S})\mathbf{C} = 0, \quad (4)$$

其中, \mathbf{H} , \mathbf{S} 分别表示哈密顿量矩阵和交叠矩阵; \mathbf{C} 是本征波函数, 即

$$\begin{aligned} \mathbf{H}_{\mu\nu} &= \langle \phi_\mu | \mathbf{H} | \phi_\nu \rangle, \\ \mathbf{S}_{\mu\nu} &= \langle \phi_\mu | \phi_\nu \rangle, \\ \mathbf{C} &= (c_{n1}, c_{n2}, \dots)^T. \end{aligned} \quad (5)$$

方程 (3) 中, 等式左边括号中的几项依次为电子的动能、电子感受的外势、Hartree 势和交换关联势。要求解上述矩阵方程, 首先需要构建哈密顿量矩阵和交叠矩阵。其中哈密顿量矩阵包括动能矩阵、非局域赝势矩阵和局域势(赝势的局域部分, Hartree 势和交换关联势) 矩阵。在我们的程序中, 动能矩阵、非局域赝势矩阵和交叠矩阵的矩阵元由双中心积分得到, 而局域势矩阵元由实空间格点积分得到, 因此双中心积分和格点积分是数值原子轨道方法中的两个核心部分, 它们的计算精度和效率尤为重要。

3.1.2 双中心积分

数值原子轨道作为基组的其中一个优点就是交叠矩阵、动能矩阵和非局域赝势矩阵可以用双中心积分高效、精确地计算^[6]。交叠矩阵元一般可以写成如下形式,

$$\mathbf{S}_{\mu\nu}(\mathbf{R}) = \int \phi_\mu^*(\mathbf{r}) \phi_\nu(\mathbf{r} - \mathbf{R}) d\mathbf{r}, \quad (6)$$

这可以进一步写成,

$$\begin{aligned} \mathbf{S}_{\mu\nu}(\mathbf{R}) &= \sum_{l=0}^{2l_{\text{max}}} \sum_{m=-l}^l S_{l_\mu m_\mu, l_\nu m_\nu, lm}(\mathbf{R}) \\ &\quad \times G_{l_\mu m_\mu, l_\nu m_\nu, lm} Y_{lm}(\hat{\mathbf{R}}). \end{aligned} \quad (7)$$

这里, l_μ (l_ν) 和 m_μ (m_ν) 分别为轨道的角动量子数、磁量子数。上式中,

$$\begin{aligned} S_{l_\mu m_\mu, l_\nu m_\nu, lm}(\mathbf{R}) \\ = 4\pi i^{-l} \int_0^\infty j_l(kR) f_\mu(k) f_\nu(k) k^2 dk, \end{aligned} \quad (8)$$

其中 $f_\mu(k)$ 和 $f_\nu(k)$ 分别是两个原子轨道径向函数的一维傅里叶变换, 即

$$f_\mu(k) = \sqrt{\frac{2}{\pi}} (-i)^{l_\mu} \int_0^\infty r^2 j_{l_\mu}(kr) f_\mu(r) dr. \quad (9)$$

$G_{l_\mu m_\mu, l_\nu m_\nu, lm}$ 是 Gaunt 系数,

$$\begin{aligned} G_{l_\mu m_\mu, l_\nu m_\nu, lm} \\ = \int_0^\pi \sin(\theta) d\theta \int_0^{2\pi} Y_{l_\mu m_\mu}(\theta, \phi) \\ \times Y_{l_\nu m_\nu}(\theta, \phi) Y_{lm}(\theta, \phi) d\phi. \end{aligned} \quad (10)$$

非局域赝势的一般形式为^[18]

$$\hat{V}_\alpha^{\text{NL}} = \sum_{l=0}^{l_{\text{max}}} \sum_{m=-l}^l \sum_{n=1}^{n_{\text{max}}} |\chi_{\alpha l m n}\rangle \langle \chi_{\alpha l m n}|, \quad (11)$$

其中 α 表示元素种类, $|\chi_{\alpha l m n}\rangle$ 是非局域投影子, l, m, n 分别表示角动量、磁角动量和投影子序号。因此非局域赝势的矩阵元也可以按上述求交叠矩阵的方法来做。

我们可以按同样的方法来计算动能矩阵, 这时只需要把 $S_{l_\mu m_\mu, l_\nu m_\nu, lm}(\mathbf{R})$ 换成

$$\begin{aligned} T_{l_\mu m_\mu, l_\nu m_\nu, lm}(\mathbf{R}) \\ = 2\pi i^{-l} \int_0^\infty j_l(kR) f_\mu(k) f_\nu(k) k^4 dk \end{aligned} \quad (12)$$

即可。

利用双中心积分来计算交叠矩阵元、动能矩阵元和非局域赝势矩阵元, 可以把三维积分简化成一维积分, 从而大幅提高计算的效率。由于 $S_{l_\mu m_\mu, l_\nu m_\nu, lm}(\mathbf{R})$ 和 $T_{l_\mu m_\mu, l_\nu m_\nu, lm}(\mathbf{R})$ 只与原子间距相关, 所以在给定的具有严格截断的局域轨道下, 我们只需要在程序的一开始一次性地把它们在足够密的一维格点上计算并存储, 那么在接下来的计算中就可以利用插值高效并精确地求出交叠矩阵和动能矩阵。同样的方法也适用于计算非局域赝势矩阵。由于双中心积分的效率很高, 因此双中心积分的占用时间与下面介绍的三维格点积分相比可以忽略。

3.1.3 实空间格点积分

计算哈密顿量的局域势矩阵元时需要计算形如下式的三中心积分:

$$V_{\mu,\nu}^{\text{loc}} = \langle \phi_\mu(\mathbf{r} - \mathbf{R}_\mu) | V^{\text{loc}}(\mathbf{r}) | \phi_\nu(\mathbf{r} - \mathbf{R}_\nu) \rangle, \quad (13)$$

其中, $\phi_\mu(\mathbf{r} - \mathbf{R}_\mu)$ 和 $\phi_\nu(\mathbf{r} - \mathbf{R}_\nu)$ 分别是中心在 \mathbf{R}_μ 和 \mathbf{R}_ν 的原子轨道; $V^{\text{loc}}(\mathbf{r})$ 是 Kohn-Sham 方程

中的局域势, 包含Hartree势 $V^H(\mathbf{r})$, 交换关联势 $V^{xc}(\mathbf{r})$ 和局域赝势 $V^L(\mathbf{r})$,

$$V^{\text{loc}}(\mathbf{r}) = V^H(\mathbf{r}) + V^{xc}(\mathbf{r}) + V^L(\mathbf{r}). \quad (14)$$

目前尚没有在实空间解析处理三中心积分的方法. 这里, 我们对局域势在实空间格点上直接做积分(离散化求和), 即

$$V_{\mu\nu}^{\text{loc}} = \sum_{\mathbf{r}_i} \phi_{\mu}(\mathbf{r}_i) V^{\text{loc}}(\mathbf{r}_i) \phi_{\nu}(\mathbf{r}_i). \quad (15)$$

在实空间格点做积分, 其精度由格点精细程度决定. 实际操作中, 可以通过调节描述平面波的能量截断值来选择格点精细度, 进而调节计算精度.

3.1.4 求解Kohn-Sham方程

构建哈密顿量矩阵和交叠矩阵后, 接下来需要求解矩阵本征值方程(4). 这里使用的是直接矩阵对角化的方法. 我们在2010年与中国科学院网络中心的赵永华研究员合作, 采用他们自主开发的一款高性能实对称矩阵和厄米矩阵特征值求解器的软件包HPSEPS(High Performance Symmetric Eigenproblem Solver)^[19], 实现了矩阵本征值问题的高效并行求解, 将可以计算的体系提高到1000个原子. 直接对角化方法的时间复杂度为 $O(N^3)$, N 是体系中原子的个数.

此外, 我们在2011年和劳伦斯伯克利的Lin和Yang合作, 将ABACUS和Lin发展的极点分解方法和选择性求逆方法(Pole EXpansion and Selected Inversion (PEXSI) technique)^[4]结合起来. 这种方法不直接求解能量本征值和波函数, 而是直接算出密度矩阵, 其时间复杂度对一维系统为 $O(N)$, 二维系统为 $O(N^{1.5})$, 三维系统为 $O(N^2)$, 低于传统的 $O(N^3)$ 标度. 这个算法的另一个优点在于并行度高. 传统的Scalapack等对角化程序的并行度只有几百个核到一千核左右, 而极点分解加选择性求逆的方法有两层并行: 第一层是极点之间的并行, 相当于把费米算符展开, 然后对展开后的每一项进行计算; 第二层并行在极点展开之后, 每一层极点之间的稀疏矩阵的操作也可以再并行. 最终两层并行的算法可以轻易地扩展到几万个核.

3.2 原子基组生成方法

前面提到, 我们使用的局域基组是原子轨道, 每一个原子轨道都是径向轨道和球谐函数乘积的形式. 对于轨道的径向部分, ABACUS提供了不同

元素的基组文件. 用户也可以根据需要自己生成基组文件. ABACUS自带一套生成轨道的工具, 它是根据2010年Chen, Guo和He提出的方案^[9]实现的. 该方案的优点是在统一的算法下系统地增加轨道基矢的数目.

在我们的方案中, 原子轨道可以写成:

$$\phi_{\mu}(\mathbf{r}) = f_{ll}(r) Y_{lm}(\hat{\mathbf{r}}). \quad (16)$$

其中径向轨道用一组球贝塞尔函数的线性组合表示:

$$f_{\mu l}(r) = \begin{cases} \sum_q c_{\mu q} j_l(qr), & r < r_c, \\ 0, & r \geq r_c, \end{cases} \quad (17)$$

$j_l(qr)$ 是球贝塞尔函数, q 的选择要满足 $j_l(qr_c) = 0$, 其中 r_c 表示径向轨道的截断值. 之所以选择球贝塞尔函数来拟合原子轨道的径向部分, 是因为它近似是一组正交基矢量, 而且它具有严格截断的性质.

我们定义了溢出函数 S ^[20]作为衡量轨道质量的标准,

$$S = \frac{1}{N_n} \sum_{n=1}^{N_n} \langle \psi_n | 1 - \hat{P} | \psi_n \rangle, \quad (18)$$

其中 ψ_n 是由平面波计算出来的参考系统的本征态, N_n 表示系统波函数的数目, 投影算符 \hat{P} 由非正交的数值原子轨道张成,

$$\hat{P} = \sum_{\mu\nu} | \phi_{\mu} \rangle \mathbf{S}_{\mu\nu}^{-1} \langle \phi_{\nu} |, \quad (19)$$

这里 ϕ_{μ} 是数值原子轨道, $\mathbf{S}_{\mu\nu}$ 是它们的交叠矩阵. 径向轨道的形状可由最小化溢出函数 S 得到. S 的取值范围为0到1, 数值越接近0, 表示这时数值原子轨道张成的空间越接近参考系统波函数张成的希尔伯特空间.

在优化轨道时, 我们从一组随机的球贝塞尔函数的系数 $\{c_{\mu,q}\}$ 开始, 通过模拟退火算法得到一组系数使得溢出函数 S 最小. 参考体系可由用户根据需要自由选择. 我们发现不同键长的双原子dimer是很好的参考系统. 值得注意的是, 优化之后的轨道形状有时会出现振荡的情况^[9]. 接下来我们通过降低轨道的动能得到光滑的数值原子轨道.

为了满足不同的需要, ABACUS允许分步产生所需的原子轨道. 每一步用户都可以设置所需要的轨道角动量和每个轨道角动量上径向轨道的数目. 当第一层的轨道优化完以后, 溢出函数会下降到某个值. 第二层的轨道则会在第一层轨道的基础

上, 进一步降低溢出函数. 在这两步之间, 我们会做一个将系统波函数和上一层所得局域轨道正交化的操作, 使第二层的轨道与第一层的轨道正交, 从而保证每一步得到的新轨道不会与上一步得到的局域轨道占据同样的希尔伯特空间. 重复以上步骤即可得到任意所需的原子基组.

利用这个方案, 我们测试了一系列的体系, 都得到了很好的结果, 说明我们的数值原子轨道精度和可移植性都很好. 作为例子, 我们在图 1 显示了 Si 元素 DZP 轨道的径向函数. 生成基组时选取的能量截断值为 50 Ry, 轨道截断为 10.0 Bohr. 我们选取 5 个不同键长 (3.30, 3.78, 4.25, 5.19, 7.08 Bohr) 的 Si dimer 为参考系统. Si 的 DZP 轨道包含 2 个 s, 2 个 p 和 1 个 d 径向轨道. 我们分两步产生这些轨道: 第一步生成一个 s 和一个 p 轨道; 第二步生成一个 s, 一个 p 和一个 d 轨道. 实际操作中, 如何选择每一步生成什么轨道是灵活的, 我们发现在轨道数量一定的条件下, 不同的轨道生成方法对计算结果的精度影响不大.

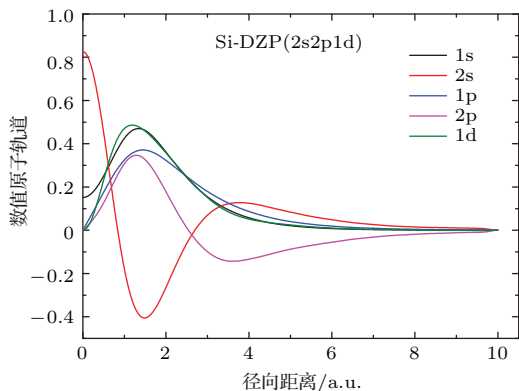


图 1 Si 元素 2s2p1d 轨道的径向函数
Fig. 1. The radial wave functions of Si 2s2p1d orbitals.

图 2 显示了用不同大小的数值原子轨道基组得到的 Si 金刚石结构的总能量随晶格常数的变化, 并与平面波基组得到的结果对比. 结果显示, 随着基组数目的增加, 数值原子轨道基组的计算结果越来越靠近平面波基组的结果. 对于 Si 的金刚石结构, DZP 基组计算得到的晶格常数已经很接近平面波的结果. 我们对大量元素的原子基组做了广泛的测试, 包括分子键长、解离能、振动频率; 固体的晶格常数、体弹性模量、结合能; 表面重构和缺陷的形成能等. 更多的结果可以在文献 [9,21] 中找到. 测试结果表明我们的原子基组具有良好的精度和可移植性. 一般来说, 对于电子结构、晶体结构的研

究, DZP 基组是不错的选择. 我们的测试表明增加基组数目可以系统地提高计算的精度, 但计算量会随之增大.

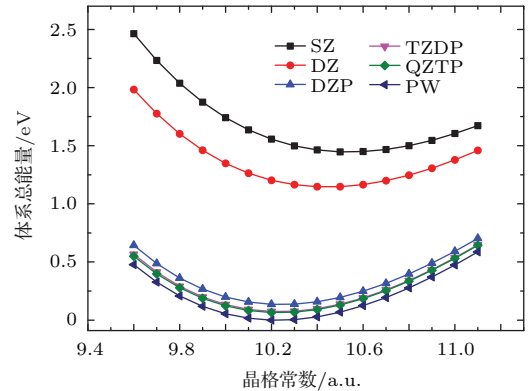


图 2 不同数值原子轨道基组以及平面波基组下 Si 金刚石结构的总能量随晶格常数的变化
Fig. 2. The total energies of bulk silicon as functions of lattice constant calculated by different basis sets.

3.3 ABACUS 并行方案

我们在不同层面对 ABACUS 实现了并行计算从而提高计算效率. 虽然 ABACUS 是一款以局域基组为特色的第一性原理计算软件, 但它还自带传统平面波基组的计算功能. 这两种基矢量的相关程序已经全部实现并行化.

3.3.1 布里渊区内的 k 点并行

对于周期性体系, 根据布洛赫定理和密度泛函理论, 系统的电子波函数满足 Kohn-Sham 方程 (3). 在对第一布里渊区离散化之后可以得到一组 k 点, 对于不同的 k 点, 方程之间是独立的, 所以可以对 k 点并行. 这部分是在程序的初始部分完成的, 程序根据计算时的 k 点数目和所用资源数目, 将 k 点划分成 N 个不同的 POOL, 每个 POOL 同时进行计算, 最后把 N 个 POOL 的计算结果收集起来. 对于 k 点的并行目前只在平面波基组方法中得到了实现. 在局域轨道基组里由于最大的计算量往往来自于对角化哈密顿矩阵, 所以我们暂时并没有对 k 点进行并行.

3.3.2 平面波并行

对于布里渊区里的每一个 k 点, 在求解 Kohn-Sham 方程的过程中, 涉及一些物理量在实空间和倒空间的傅里叶变换, 如电荷密度和局域势等. ABACUS 采用了 FFTW 软件包 [22]. FFTW 软件包具有可移植性强, 支持分布式存储 MPI 并行, 效

率高等优点. ABACUS 内部并没有使用并行的 FFTW, 而是实现了对串行 FFTW 的并行, 这样做的好处是可以使得在倒空间平面波的分布更均匀. 在实空间, 则是首先根据计算体系所需的能量截断值和原胞大小, 得到实空间的三维格点作为快速傅里叶变换 (FFT) 格子; 然后根据计算所用核数沿所模拟的系统的 z 方向对 FFT 格子划分, 实现 FFT 的实空间并行.

3.3.3 实空间计算的并行

在使用数值原子轨道基组时, 除了前面提到的局域势矩阵元, 电荷密度以及求力中涉及局域势的项也需要由实空间格点积分得到. 直接在实空间格点做积分, 计算量较大, 是算法中最耗时的部分之一. 对于这几个涉及实空间空间操作的部分, ABACUS 利用前面做 FFT 并行时划分的实空间格点, 实现了对实空间格点积分的并行, 计算效率因此得以大大提高.

3.3.4 对角化

对于数值原子轨道基组, 在求解矩阵本征值方程时, ABACUS 通过调用软件包 HPSEPS 直接做矩阵对角化. HPSEPS 软件包已经实现了高效的并行. 它采用 2D 块划分的方式, 将矩阵的数据结构分布在二维处理器网格上, 然后进一步并行求解. 在 ABACUS 中, 2D 块的划分是在程序初始化时完成的. 程序按照 HPSEPS 的要求把计算资源数目划分成 2D 块, 然后把矩阵元信息分配到每一个 2D 块. 之后在计算矩阵元时, 每一个 2D 块会根据之前分到的信息判断该矩阵元是否归其所有. 当每一个 2D 块都得到各自的矩阵元后即调用 HPSEPS 包直接求解矩阵本征值和本征函数.

我们完成了对各个耗时部分的优化. 现在 ABACUS 具有较高的并行效率和较快的计算速度. 在图 3 中我们给出了随着体系中原子数的增加, 一个电子步中 ABACUS 最耗时的两部分: 构建哈密顿矩阵和矩阵对角化及总的计算时间的变化曲线. 在图 3 中, 计算所选体系分别为含有 16, 64, 128, 192, 256 和 512 个 Ti 原子的六角密堆积 (hcp) 结构, 所选能量截断为 100 Ry. 所采用的局域轨道基组为 DZDP 基组, 每个原子有 18 个原子轨道, 轨道截断半径为 8.0 Bohr. 我们用了 16 个核来进行测试. 图 3 的时间统计结果显示, 在系统少于 256 个原子 (原子基组数少于 4608) 时, 格点积分占了计算的主要部分, 对角化时间较少. 但随着系统尺寸增大,

格点积分的时间随原子数 N 线性增长, 而对角化的时间是 N^3 增长, 并逐渐占据主导地位. 结果显示, 在系统较大时对角化占了绝大部分时间, 并且随原子数增多对角化所用时间增长非常快, 逐渐成为计算的瓶颈. 如果需要计算更大的系统, 可以使用 PEXSI 算法.

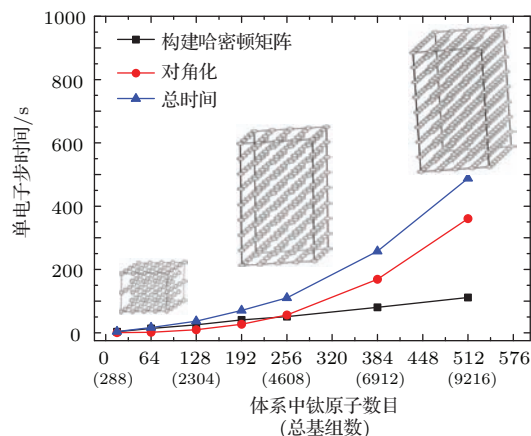


图 3 构建哈密顿矩阵、对角化和单步电子迭代总时间随系统大小的变化

Fig. 3. The times to construct Hamiltonian matrices, diagonalization and total time per electronic step as functions of system sizes.

4 电子结构算法

电子结构计算是第一性原理计算软件的核心, 具体包括电子波函数求解、电荷密度混合、电荷密度插值等. 前文介绍了在 ABACUS 中如何构建矩阵元和求解本征值和本征波函数, 接下来将介绍 ABACUS 中使用的电荷密度混合和电荷密度插值方法.

4.1 电荷密度混合

在密度泛函理论框架下, 需要通过自洽迭代的方法来求解 Kohn-Sham 方程. 首先从一个初始的电荷密度出发, 代入求解 Kohn-Sham 方程, 得到新的电荷密度, 然后由新的电荷密度出发, 再次求解 Kohn-Sham 方程, 直至输出的电荷密度和输入的电荷密度的差别小于给定的收敛值. 在电子密度自洽迭代的过程中, 如何得到新的电荷密度, 会很大程度上影响计算的收敛速度, 进而影响计算效率. 为了稳定地收敛电荷密度, 一般需要将输出电荷密度和输入电荷密度做混合, 将得到的混合电荷密度作为下一步的输入电荷密度. 常见的电荷密度混合方法有 plain mixing, Pulay mixing^[23], Kerker

mixing^[24] 和 Pulay-Kerker mixing^[25] 等. 以上几种方法都已在 ABACUS 中得到实现. 下面简单介绍一下 Pulay mixing 和 Pulay-Kerker mixing 两种混合方法.

4.1.1 Pulay 算法

使用 Pulay mixing 的方法, 要存储在自洽迭代过程中一系列的输入电荷密度 ρ_{in} 和电荷密度残余量 $\mathbf{R}[\rho_{in}]$,

$$\mathbf{R}[\rho_{in}] = \rho_{out}[\rho_{in}] - \rho_{in}, \quad (20)$$

ρ_{out} 是每步的输出电荷密度. 假设当前最佳的输入电荷密度 ρ_{in}^{opt} 是之前 N 步输入电荷密度的一个线性组合,

$$\rho_{in}^{opt} = \sum_{i=1}^N \alpha_i \rho_{in}^i. \quad (21)$$

要保证电荷密度守恒, 线性组合系数 α_i 必须满足归一化条件,

$$\sum_{i=1}^N \alpha_i = 1. \quad (22)$$

这个最佳的输入电荷密度应当使当前残余量矢量的模最小, 即 $\langle \mathbf{R}[\rho_{in}^{opt}] | \mathbf{R}[\rho_{in}^{opt}] \rangle$ 最小. 这里我们再次做出一个近似

$$\mathbf{R}[\rho_{in}^{opt}] = \mathbf{R} \left[\sum_i \alpha_i \rho_{in}^i \right] = \sum_i \alpha_i \mathbf{R}[\rho_{in}^i]. \quad (23)$$

定义不同残余量之间的重叠矩阵为

$$\mathbf{A}_{ij} = \langle \mathbf{R}[\rho_{in}^j] | \mathbf{R}[\rho_{in}^i] \rangle, \quad (24)$$

利用变分法, 可以得到最佳的线性组合系数为

$$\alpha_i = \frac{\sum_j (\mathbf{A}^{-1})_{ji}}{\sum_{kj} (\mathbf{A}^{-1})_{kj}}, \quad (25)$$

其中, \mathbf{A}^{-1} 表示矩阵 \mathbf{A} 的逆. 假设当前的电子迭代步数是第 m 步, 则第 $m+1$ 步的输入电荷密度为

$$\rho_{in}^{m+1} = \rho_{in}^{opt} + \beta \mathbf{R}[\rho_{in}^{opt}], \quad (26)$$

这里 β 是电荷密度混合参数. 对于绝缘体来说, Pulay mixing 是一种很好的电荷密度混合方法, 一般在 10 个电子步左右就可以实现电荷密度的收敛.

4.1.2 Pulay-Kerker 算法

对于含有长程电荷密度涨落的体系, 如金属体系和原胞中含有大量原子的周期性体系, 在电子迭

代过程中会出现电荷密度振荡、很难收敛的情况. Kerker^[24] 在 1981 年提出可以将电荷密度残余量乘上一个预处理的矩阵,

$$\lambda(\mathbf{G}) = \frac{|\mathbf{G}|^2}{|\mathbf{G}|^2 + |\mathbf{G}_0|^2}, \quad (27)$$

其中 \mathbf{G} 为系统的倒格矢, 即对于长程、低能的电荷密度变化, 预处理矩阵只会混入一小部分的残余量, 这样可以有效抑制长程的电荷振荡现象, 加快收敛速度. 把 Kerker 算法与 Pulay 混合算法结合起来就可以得到 Pulay-Kerker 电荷密度混合算法^[25],

$$\rho_{in}^{m+1} = \rho_{in}^{opt} + \lambda(\mathbf{G})\beta \mathbf{R}[\rho_{in}^{opt}]. \quad (28)$$

目前 ABACUS 已经实现了 Pulay-Kerker mixing 的功能, 对于一些较大的金属体系, 电荷收敛速度得到很大的提升. 下面给出 ABACUS 用数值原子轨道基组计算含有 200 个 Ti 原子体系时, 分别使用 Pulay 与 Pulay-Kerker 电荷混合算法的收敛速度对比. 结果如图 4 所示, 横坐标表示电子迭代的步数, 纵坐标表示的是每个电子步输出电荷密度 ρ_{out} 与输入电荷密度 ρ_{in} 的相对误差, 即

$$\frac{\sum_{r_i} |\rho_{out}(r_i) - \rho_{in}(r_i)|}{n_{elec}}, \quad (29)$$

r_i 表示实空间格点, n_{elec} 表示体系中电子的个数. 我们看到对于 200 个 Ti 原子的体系, 在使用 Pulay 混合的算法时, 电荷密度振荡严重, 收敛很慢, 到 40 步还未收敛到预设的精度; 而使用 Pulay-Kerker 混合时, 电荷密度在 12 步就收敛到了 10^{-6} , 收敛速度得到很大提高. 对于越大的体系, Pulay-Kerker 算法的优势越明显.

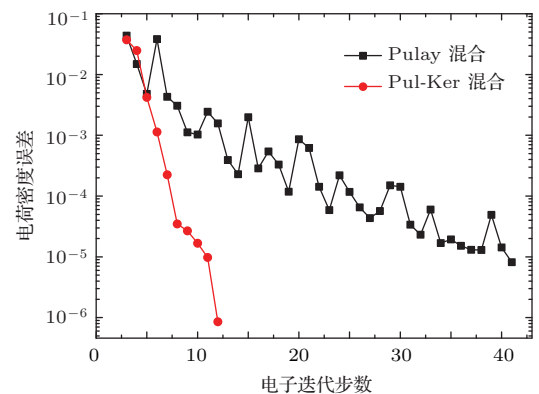


图 4 计算 200 个 Ti 原子的超限时 Pulay 和 Pulay-Kerker 电荷密度混合方法收敛速度对比

Fig. 4. Comparison of the charge density convergence using Pulay and Pulay-Kerker mixing methods for a 200 Ti system.

4.2 电荷密度插值

在做结构优化或者分子动力学计算时, 需要计算大量结构的基态电荷密度. 每次得到新的结构构型后, 如何选取一个好的初始输入电荷密度, 在很大程度上也会影响计算效率, 即所需收敛电荷密度的步数. 常用的提高收敛速度的方法是对电荷密度进行外插值. 对于绝大多数系统, 其电荷密度可以写成其原子电荷密度的叠加 $\rho_{\text{at}}(t)$ 与一个小的变化量 $\delta\rho(t)$ 之和, 即

$$\rho(t) = \rho_{\text{at}}(t) + \delta\rho(t) \quad (30)$$

$\rho(t)$ 是 t 时刻自洽的电荷密度, $\rho_{\text{at}}(t)$ 是 t 时刻原子电荷密度的叠加. 同理, $t + dt$ 时刻的电荷密度可以表示成

$$\rho(t + dt) = \rho_{\text{at}}(t + dt) + \delta\rho(t + dt). \quad (31)$$

每一时刻的原子电荷密度之和可以快速地计算出来, 而 $\delta\rho$ 可以由插值方法得到. 常用的电荷密度插值方法有 [26]:

1) 简单插值

如果相邻离子步之间原子位置变化不大, 可以通过电荷密度的原子插值得到下个离子步的初始输入电荷密度

$$\delta\rho(t + dt) = \delta\rho(t), \quad (32)$$

经过一系列的测试, 我们发现对于一般的结构弛豫计算, 电荷密度的原子插值是很好的选择;

2) 一阶插值

在一些计算中, 比如分子动力学计算, 可能要尝试电荷密度的一阶插值, 甚至是二阶插值方法, 一阶插值的方法要用到前面两个离子步的电荷密度信息,

$$\delta\rho(t + dt) = \delta\rho(t) + [\delta\rho(t) - \delta\rho(t - dt)]; \quad (33)$$

3) 二阶插值

二阶插值的方法要用到前面三个离子步的电荷密度信息,

$$\begin{aligned} \delta\rho(t + dt) = & \delta\rho(t) + \alpha[\delta\rho(t) - \delta\rho(t - dt)] \\ & + \beta[\delta\rho(t - dt) - \delta\rho(t - 2dt)], \end{aligned} \quad (34)$$

其中系数 α 和 β 由不同时刻的原子位置决定. 详细算法参见文献 [26].

对于分子动力学计算, 在一般情况下电荷密度的二阶插值效果较好. 在图 5 中我们对比了用

ABACUS 做含有 224 个原子的橄榄石 (Mg_2SiO_4) 结构在 2200 K 的分子动力学研究时, 几种不同插值方法的效果. 图中横坐标是分子动力学的步数, 时间步长为 1 fs; 纵坐标为每一步分子动力学进来时的第一个电子步的输出电荷密度与输入电荷密度的相对误差, 即 (29) 式. 我们发现在做分子动力学计算时, 电荷密度的一阶、二阶插值效果要明显优于简单插值方法. 在若干分子动力学步后, 使用简单插值方法, 初始电荷密度只能精确到 10^{-2} ; 而使用二阶插值方法, 初始电荷密度能够精确到 10^{-3} , 大大加快了计算收敛速度.

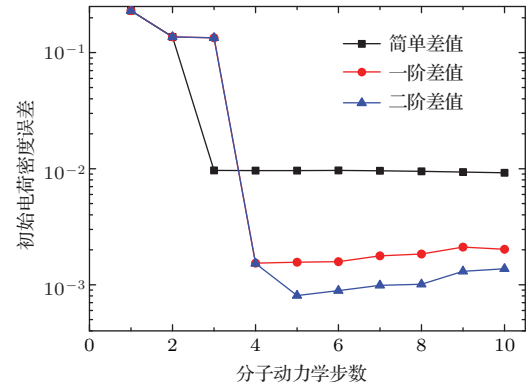


图 5 对比分子动力学计算时, 不同电荷密度插值算法得到的初始电荷密度的精度

Fig. 5. The initial charge density errors in MD simulations using different charge extrapolation methods.

4.3 Smearing 占据

当体系的电子结构为无能隙或能隙较小时, 体系的自洽过程会出现电荷密度振荡, 收敛很慢甚至不收敛. 这时如果对电子分布函数做适当的处理, 会使自洽过程更容易收敛, 这被称为电子占据的 smearing. 目前 ABACUS 实现的电子占据方式有 fixed 占据、Gaussian 占据和 methfessel-Paxton 占据.

4.3.1 Gaussian 分布函数

对于一般的体系, 可以使用高斯分布函数

$$f(\varepsilon) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(\varepsilon - E_F)^2}{2\sigma^2}\right), \quad (35)$$

E_F 是费米能, σ 表示展宽. 高斯函数对于一般的小能隙体系效果都很好.

4.3.2 Methfessel-Paxton 分布函数

Methfessel-Paxton 分布函数可以更精确地对费米面取点 [27], 高阶的展开可以得到较好的基态收敛结果, 加速收敛过程.

$$S_0(x) = \frac{1}{2}(1 - \operatorname{erf} x), \quad (36)$$

$$S_N(x) = S_0(x) + \sum_{n=1}^N A_n H_{2n-1}(x) e^{-x^2}, \quad (37)$$

其中, $x = (\varepsilon - E_F)/\sigma$, H_n 是 Hermite 多项式, A_n 是展开系数, N 表示展开的阶数. 对于金属体系, 建议使用 Methfessel-Paxton 分布函数.

5 ABACUS 计算实例

在过去几年中, 我们一直致力于发展和优化 ABACUS 软件. 目前该软件已经可以有效地研究一千原子以内的物理体系的电子结构、原子构型和包含数百个原子体系的分子动力学.

下面我们给出应用 ABACUS 计算的几个例子.

1) 含有 3 个空位的 Ti 原子体系的结构弛豫

在本例中我们研究了 Ti 金属 hcp 结构中的空位缺陷. 我们从包含 200 个 Ti 原子的超胞中挖去 3 个相邻的 Ti 原子, 形成空位, 并研究空位的基态结构. 我们采用 Ti 原子的 2s2p2d 轨道基组, 能量截断为 100 Ry, 轨道截断为 10.0 Bohr. 我们采用共轭梯度法优化结构. 图 6 展示了系统弛豫时能量和原子受力随优化步数收敛的情况, 横坐标表示结构弛豫步数, 左边的纵坐标表示体系能量, 右边的纵坐标表示每个弛豫步原子受力的最大值. 弛豫过程总共只用了 32 个离子步, 原子受力就收敛到了 0.01 eV/Å. 可见对于结构弛豫计算, ABACUS 具有很好的收敛性.

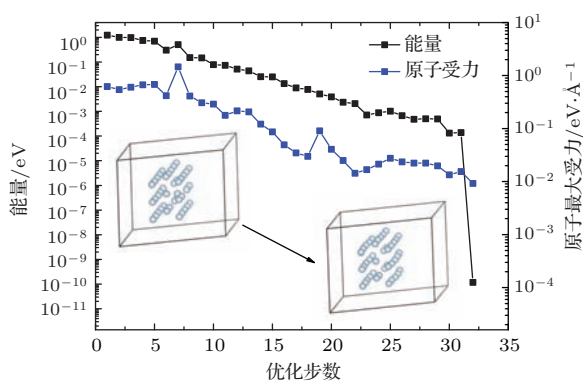


图 6 结构弛豫过程中体系能量和原子受力的变化曲线
Fig. 6. The total energies and maximum forces as functions of structure relaxation steps.

2) 蛋白质结构的能态密度

我们应用 ABACUS 研究了一个模型蛋白质 $C_{78}H_{102}N_{19}O_{19}$ 的电子能态密度, 并与平面波程序

Quantum Espresso^[13] 使用同样赝势得到的结果对比. 模型蛋白的结构如图 7 中插图所示, 包含 C, H, N, O 四种元素, 原子数分别为 78, 102, 19, 19. 在 ABACUS 计算中, 我们采用 DZP 轨道, 能量截断值为 100 Ry, 轨道截断值为 10 Bohr. 我们看到 ABACUS 计算得到的能态与 Quantum Espresso 所得几乎完全一致.

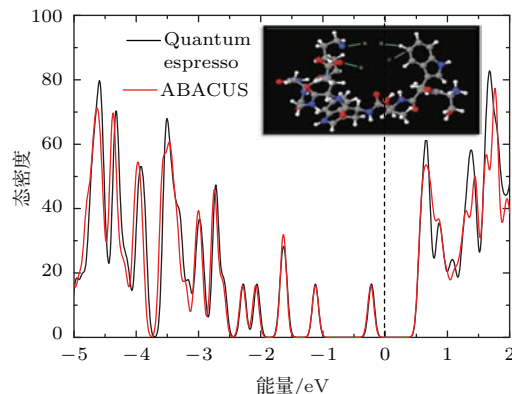


图 7 对比由 Quantum Espresso 和 ABACUS 分别得到的蛋白质结构的态密度, 费米能取为 0
Fig. 7. Comparison of the density of states of a model protein calculated by Quantum Espresso and ABACUS. The energy of Fermi level is set to zero.

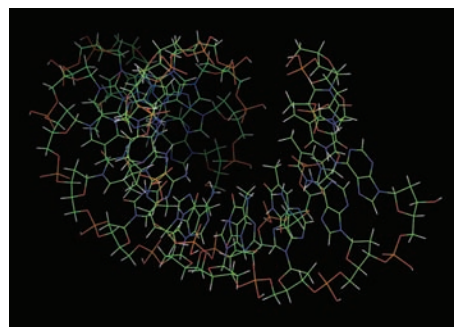


图 8 含 12 对碱基对、788 个原子的 DNA 片段弛豫后的结构
Fig. 8. The structure of a DNA piece containing 12 base pairs, 788 atoms after structure optimization.

3) 含有 12 对碱基对的 DNA 分子片段的结构

在本例中我们研究了一段长度为 12 对碱基对的 DNA 双螺旋片段, 对其进行了结构优化的计算. 该片段为 A 型构象, 序列为 5'-AAAAAAAAAAAA-3', 初始构型由 Supercomputing Facility for Bioinformatics & Computational Biology, Indian Institute of Technology 的 Jayaram 教授研究提供的工具^[28]生成. 该 DNA 片段中含有 C, H, O, N 和 P 五种元素, 原子个数分别为 240, 302, 84, 140 和 22 个, 总共包含 788 个原子. 我们采用 DZP 基组, 能量截断值为 50 Ry, 轨道截

断值为8.0 Bohr, 采用共轭梯度法优化结构至原子受力收敛到0.05 eV/Å. 结构优化时我们做了范德瓦耳斯力修正. 图8展示了优化后的结构. 优化后的结构的碱基轴升(即相邻两个碱基对间的距离)为0.24 nm, 与优化前的0.26 nm相比更接近实验值0.23 nm^[29]; 优化后的小沟宽度为10.4 Å, 与文献报道的A型DNA小沟宽度为10—11 Å符合得很好^[30].

6 结 论

从2006年起, 中国科学技术大学量子信息重点实验室开始发展一套具有完全自主知识产权的第一性原理软件, 本文介绍了该软件的主要特点和算法. 该软件以数值原子轨道为基组, 可以高效地研究包含1000原子左右的物理系统的电子、结构和动力学等性质. 我们为该软件发展了一套可以系统提高的、同时具有高精度与高可移植性的数值原子轨道产生方案. 该软件为研究表面、界面、无序、缺陷以及生物大分子等复杂的物理体系提供了有力的工具.

感谢中国科学院网络中心赵永华研究员对本项目的大力支持. 本项目的部分计算在中科院超算中心和中国科技大学超算中心完成.

参考文献

- [1] Hohenberg P, Kohn W 1964 *Phys. Rev.* **136** B864
- [2] Kohn W, Sham L J 1965 *Phys. Rev.* **140** A1133
- [3] Goedecker S 1999 *Rev. Mod. Phys.* **71** 1085
- [4] Lin L, Lu J F, Car R E W 2009 *Phys. Rev. B* **79** 115133
- [5] Lin L, Chen M H, Yang C, He L X 2013 *J. Phys.: Condens. Matter* **25** 295501
- [6] Soler J M, Artacho E, Gale J D, García A, Junquera J, Ordejón P, Sánchez-Portal D 2002 *J. Phys.: Condens. Matter* **14** 2745
- [7] Ozaki T 2003 *Phys. Rev. B* **67** 155108
- [8] Blum V, Gehrke R, Hanke F, Havu P, Havu V, Ren X G, Reuter K, Scheffler M 2009 *Comput. Phys. Commun.* **180** 2175
- [9] Chen M H, Guo G C, He L X 2010 *J. Phys.: Condens. Matter* **22** 445501
- [10] Chen M H, Guo G C, He L X 2011 *J. Phys.: Condens. Matter* **23** 325501
- [11] <http://abacus.ustc.edu.cn/> [2015-7-2]
- [12] Hamann D R, Schlüter M, Chiang C 1979 *Phys. Rev. Lett.* **43** 1494
- [13] Giannozzi P, Baroni S, Bonini N, Calandra M, Car R, Cavazzoni C, Ceresoli D, Chiarotti G L, Cococcioni M, Dabo I, Dal Corso A, Fabris S, Fratesi G, de Gironcoli S, Gebauer R, Gerstmann U, Gougoussis C, Kokalj A, Lazzeri M, Martin-Samos L, Marzari N, Mauri F, Mazzarello R, Paolini S, Pasquarello A, Paulatto L, Sbraccia C, Scandolo S, Sclauzero G, Seitsonen A P, Smogunov A, Umari P, Wentzcovitch R M 2009 *J. Phys.: Condens. Matter* **21** 395502
- [14] <http://opium.sourceforge.net> [2015-7-2]
- [15] Perdew J P, Burke K, Ernzerhof M 1996 *Phys. Rev. Lett.* **77** 3865
- [16] Grimme S 2006 *J. Comput. Chem.* **27** 1787
- [17] Heyd J, Scuseria G E, Ernzerhof M 2003 *J. Chem. Phys.* **118** 8207
- [18] Kleinman L, Bylander D M 1982 *Phys. Rev. Lett.* **48** 1425
- [19] <http://www.sccas.cas.cn/yhfw/rjzy/zml/>
- [20] Portal D S, Artacho E, Soler J M 1995 *Solid State Communications* **95** 685
- [21] Li P F, Liu X H, Chen M H, Lin P Z, Ren X G, Lin L, He L X 2015 *Comput. Mater. Sci.* (in press)
- [22] Frigo M, Johnson S G 2005 *Proc. IEEE* **93** 216
- [23] Pulay P 1980 *Chem. Phys. Lett.* **73** 393
- [24] Kerker G P 1981 *Phys. Rev. B* **23** 3082
- [25] Kresse G, Furthmüller J 1996 *Comput. Mater. Sci.* **6** 15
- [26] Alfè D 1999 *Comput. Phys. Commun.* **118** 31
- [27] Methfessel M, Paxton A T 1989 *Phys. Rev. B* **40** 3616
- [28] Arnott S, Campbell-Smith P J, Chandrasekaran R 1976 *In Handbook of Biochemistry and Molecular Biology* (Cleveland: CRC Press) pp411–422
- [29] Wang J Y, Zhu S G, Xu C F 2002 *Biology Chemistry* (Beijing: Higher Education Press) p489
- [30] Chattopadhyay B, Mukherjee M 2011 *J. Phys. Chem. B* **115** 1760

SPECIAL ISSUE—Recent developments and applications of computational physics

Introduction to first-principles simulation package ABACUS based on systematically improvable atomic orbitals*

Liu Xiao-Hui¹⁾²⁾ Chen Mo-Han³⁾ Li Peng-Fei¹⁾²⁾ Shen Yu⁴⁾ Ren Xin-Guo^{1)2)†}
Guo Guang-Can¹⁾²⁾ He Li-Xin^{1)2)‡}

1) (Key Laboratory of Quantum Information, University of Science and Technology of China, Hefei 230026, China)

2) (Synergetic Innovation Center of Quantum Information and Quantum Physics, University of Science and Technology of China, Hefei 230026, China)

3) (Department of Mechanical and Aerospace Engineering, Princeton University, Princeton, New Jersey 08544, USA)

4) (Network Information Center, Supercomputing Center, University of Science and Technology of China, Hefei 230026, China)

(Received 2 July 2015; revised manuscript received 3 September 2015)

Abstract

With the rapid development of supercomputers and the advances of numerical algorithms, nowadays it is possible to study the electronic, structural and dynamical properties of complicated physical systems containing thousands of atoms using density functional theory (DFT). The numerical atomic orbitals are ideal basis sets for large-scale DFT calculations in terms of their small base size and localized characteristic, and can be mostly easily combined with linear scaling methods. Here we introduce a first-principles simulation package “Atomic-orbital Based *Ab-initio* Computation at UStc (ABACUS)”, developed at the Key Laboratory of Quantum Information, University of Science and Technology of China. This package provides a useful tool to study the electronic, structural and molecular dynamic properties of systems containing up to 1000 atoms. In this paper, we introduce briefly the main algorithms used in the package, including construction of the atomic orbital bases, construction of the Kohn-Sham Hamiltonian in the atomic basis sets, and some details of solving Kohn-Sham equations, including charge mixing, charge extrapolation, smearing etc. We then give some examples calculated using ABACUS: 1) the energy orders of B20 clusters; 2) the structure of bulk Ti with vacancies; 3) the density of states of a model protein; 4) the structure of a piece of DNA containing 12 base pairs, 788 atoms. All results show that the results obtained by ABACUS are in good agreement with either experimental results or results calculated using plane wave basis.

Keywords: density functional theory, numerical atomic basis, first-principles package

PACS: 71.15.Ap, 71.15.Mb

DOI: 10.7498/aps.64.187104

* Project supported by the Chinese National Fundamental Research Program (Grant No. 2011CB921200), the National Natural Science Foundation of China (Grant Nos. 11374275, 11374276), and “Strategic Priority Research Program (B)” of the Chinese Academy of Sciences (Grant No. XDB01030100).

† Corresponding author. E-mail: renxg@ustc.edu.cn

‡ Corresponding author. E-mail: helx@ustc.edu.cn