

密度泛函理论

Density Functional Theory

[美] David S. Sholl Janice A. Steckel 著
李健 周勇 译

国防工业出版社

·北京·

著作权合同登记 图字:军-2014-135号

图书在版编目(CIP)数据

密度泛函理论 / (美) 萧(Sholl, D. S.), (美) 斯特克尔(Steckel, J. A.) 著; 李健, 周勇译. —北京: 国防工业出版社, 2014. 9

书名原文: Density functional theory

ISBN 978-7-118-09737-5

I. ①密… II. ①萧… ②斯… ③李… ④周…
III. ①密度泛函法 IV. ①0414.2

中国版本图书馆 CIP 数据核字 (2014) 第 222524 号

Copyright © 2009 by John Wiley & Sons, Inc. All rights reserved.

Density Functional Theory by DAVID S. SHOLL and JANICE A. STECKEL.

ISBN 978-3-527-41024-8

This translation published under John Wiley & Sons. No part of this book maybe reproduced in any form without the written permission of the original copyrights holder.

Copies of this book sold without a Wiley sticker on the cover are unauthorized and illegal.

本书简体中文版由 John Wiley & Sons, Inc. 授权国防工业出版社独家出版。版权所有, 侵权必究。

※

国防工业出版社出版发行

(北京市海淀区紫竹院南路 23 号 邮政编码 100048)

国防工业出版社印刷厂印刷

新华书店经售

*

开本 880 × 1230 1/32 印张 7 $\frac{7}{8}$ 字数 225 千字

2014 年 9 月第 1 版第 1 次印刷 印数 1—2000 册 定价 68.00 元

(本书如有印装错误, 我社负责调换)

国防书店: (010) 88540777

发行邮购: (010) 88540776

发行传真: (010) 88540755

发行业务: (010) 88540717

译者序

基于量子力学并根据密度泛函理论 (Density Functional Theory, DFT), 可以通过自洽计算来确定材料的几何结构、电子结构、热力学性质和光学性质等材料物性。近年来, 随着计算机技术的飞速发展, 计算能力得以空前提高。此外, 各种密度泛函计算程序也取得了快速的发展, 这类计算已经越来越多地被应用到固体、表面、材料设计、合成、模拟计算、大分子和生物体系等诸多方面的研究中。毫无疑问, 密度泛函理论及计算已成为研究材料各种物理和化学性质非常普遍的手段, 获得了许多突破性的进展, 已经成为计算化学、计算物理和材料模拟的一个重要基础方法和核心技术。

本书的英文原著采用大量计算实例, 对 DFT 计算尤其是平面波 DFT 计算, 进行了深入浅出、简明扼要的介绍。本书不仅介绍了 DFT 的一些重要概念, 更难能可贵的是对 DFT 在隔离态分子、体相固体、表面和界面等方面的实际应用, 进行了系统而全面的讲解。使读者在学习 DFT 计算时, 能把主要精力集中在实际问题的建模求解上, 尽量避免在深奥难懂的量子力学、固体物理等基本理论中花去太多时间。

译者认为, 这是一本关于 DFT 计算非常好的入门学习教材和实际计算时的参考用书, 无论对于 DFT 的初学者, 还是已经能够较为熟练地使用 DFT 方法的研究者, 都可以从这本书中获益良多。这也正是翻译本书的一个主要目的。

本书的文字翻译和校稿、统稿工作是由西安石油大学的李健和周勇共同完成的。本书由西安石油大学优秀学术著作出版基金和焊接技术与工程教学团队建设经费资助出版。在翻译和出版过程中, 得到了国防工业出版社、西安石油大学科技处、西安石油大学材料科学与工程学院等单位 and 牛旭东、张莹、徐学利等人的大力支持与协助, 在此一并

表示感谢。

限于译者的英语和专业知识水平,在本书译文中可能存在着许多问题,请读者不吝指正。

译者
2014.08

前 言

在物理学、化学、材料科学以及其他工程领域中,密度泛函理论(DFT)及其计算已经快速发展成为材料建模模拟的一种“标准工具”。尽管已经有许多书籍和文献详细介绍 DFT 的基础理论,但对读者而言,仍然难以在短时间内掌握上述领域中 DFT 常规运算的实际操作。本书旨在填补此空白,引导读者掌握 DFT 理论的应用方法,而这些应用方法正是数量不断增长的此类科技论文的核心所在。本书每一章都包含了一些计算练习,读者可以通过这些计算积累经验。

为了使读者能够更切实地通读全书,作者力图在简略和详实之间取得平衡,这就不可避免地对某些技术细节只进行了有限的记述。在美国卡内基梅隆大学和佐治亚理工学院,笔者与化工、物理学、化学、材料科学和机械工程专业的研究生和博士后具有多年的教学互动,这对本书的内容选择具有直接影响。每章之后的扩展阅读列表可以为更深入地了解本章内容提供切入点。给出这些阅读列表,原因在于这些文献中具有亮点,而并非在现有海量文献中,对相关所有研究工作的严格引用。

在内容上的一个重要选择是:作者在本书中仅陈述了一种适用于固体和空间延展性材料的 DFT 方法,即平面波 DFT。尽管平面波 DFT 的许多基础理论与化学领域中对隔离分子所使用的其他 DFT 方法具有一定关系,但这两类方法的实际使用仍然存在本质区别。若将两者均涵盖于本书中,势必会极大地增加本书的篇幅。此外,在计算化学领域有关隔离分子的 DFT 计算,也已经有许多实践指导类的资源可供参考。

作者在 DFT 计算研究以及本书的写作过程中,均与许多同行学者进行了长期交流,并得益良多。作者特对以下学者致谢:J. Karl Johnson

(匹兹堡大学)、Aravind Asthagiri (佛罗里达大学)、Dan Sorescu (国际能源技术实验室)、Cathy Stampfl (悉尼大学)、John Kitchin (卡内基梅隆大学) 和 Duane Johnson (伊利诺斯大学)。作者还要感谢 Jeong - Woo Han 对几幅插图的帮助。Bill Schneider (美国圣母大学)、Ken Jordan (匹兹堡大学) 和 Taku Watanabe (佐治亚理工学院) 在本书写作过程中提出了详尽、有价值的意见。本书难免会存在许多错误或不足, 则均应视为作者的责任。

David S. Sholl 谨将此书献给父亲和岳父, 他们对科学的热爱和对世界的探索是作者写作本书的一大动力。Janice A. Steckel 谨将此书献给丈夫和女儿。

DAVID SHOLL

佐治亚理工学院,
亚特兰大, 佐治亚州, 美国

JAN STECKEL

国家能源技术实验室,
匹兹堡, 宾夕法尼亚州, 美国

目 录

第 1 章 何谓密度泛函理论	1
1.1 如何使用本书	1
1.2 DFT 运行范例	2
1.2.1 通过异质催化合成氨	2
1.2.2 微量杂质引发的金属脆化	4
1.2.3 模拟行星形成时的材料性质	5
1.3 薛定谔方程	7
1.4 密度泛函理论——从波函数到电荷密度	10
1.5 交换关联泛函	14
1.6 量子化学的“旅行者”	15
1.6.1 局域化的和空间扩展的函数	16
1.6.2 波函数方法(Wave - function - based Methods)	18
1.6.3 Hartree - Fock 方法	19
1.6.4 Hartree - Fock 方法之后(Beyond Hartree - Fock)	23
1.7 DFT 不能做哪些事	28
1.8 其他领域中的密度泛函理论	30
1.9 如何使用本书(再谈)	30
参考文献	32
扩展阅读	32
第 2 章 简单固体的 DFT 计算	34
2.1 周期性结构、超晶胞和晶格参数	34
2.2 面心立方材料	38

2.3	密排六方材料	40
2.4	晶体结构预测	43
2.5	相变	44
	练习	45
	扩展阅读	46
	附录 计算细节	47
第3章	DFT 计算的基本要素	48
3.1	倒易空间和 k 点	49
3.1.1	平面波和布里渊区	49
3.1.2	k 空间中的积分	52
3.1.3	在布里渊区中选择 k 点	54
3.1.4	金属—— k 空间的特例	58
3.1.5	k 空间的总结	60
3.2	截断能 (Energy Cutoffs)	60
3.2.1	赝势	63
3.3	数值优化	64
3.3.1	一维空间中的优化	64
3.3.2	大于一维的优化	69
3.3.3	关于优化我真正需要知道什么	73
3.4	DFT 总能计算——一个迭代优化问题	73
3.5	几何优化	75
3.5.1	内部自由度	75
3.5.2	具有约束原子的几何优化	78
3.5.3	优化超晶胞的体积和形状	79
	练习	80
	参考文献	81
	扩展阅读	81
	附录 计算细节	82

第 4 章 固体表面的 DFT 计算	83
4.1 表面的重要性	83
4.2 周期性边界条件和板块模型	84
4.3 选定表面计算的 k 点	87
4.4 采用 Miller 指数对表面分类	88
4.5 表面弛豫	95
4.6 表面能的计算	97
4.7 对称和非对称的板块模型	99
4.8 表面重构	100
4.9 表面上的吸附质	105
4.9.1 吸附能的精度	108
4.10 表面覆盖效应	108
练习	111
参考文献	112
扩展阅读	112
附录 计算细节	113
第 5 章 振动频率的计算	115
5.1 隔离态分子	116
5.2 原子集合的振动	119
5.3 表面上的分子	123
5.4 零点能	124
5.5 声子和非定域化振型	130
练习	131
参考文献	131
扩展阅读	131
附录 计算细节	132
第 6 章 使用过渡态理论计算化学过程速率	134
6.1 一维的例子	135

6.2	多维过渡态理论	142
6.3	寻找过渡态	145
6.3.1	弹性带方法	147
6.3.2	微动弹性带方法	149
6.3.3	初始化 NEB 计算	150
6.4	寻找正确的过渡态	153
6.5	将单个速率组合为总动力学	156
6.6	量子效应及其他复杂问题	160
6.6.1	高温/低能垒	161
6.6.2	量子隧穿	161
6.6.3	零点能	162
	练习	163
	参考文献	164
	扩展阅读	164
	附录 计算细节	165
第 7 章	基于 <i>Ab Initio</i> 热力学的平衡相图	167
7.1	体相金属氧化物的稳定性	168
7.1.1	含有无序化的例子——位形熵	173
7.2	金属和金属氧化物表面的稳定性	176
7.3	多元化学势和偶联化学反应	178
	练习	179
	参考文献	180
	扩展阅读	180
	附录 计算细节	181
第 8 章	电子结构和磁特性	182
8.1	电子态密度	182
8.2	局域态密度和原子电荷	189
8.3	磁性	191

练习	194
扩展阅读	195
附录 计算细节	195
第 9 章 从头算分子动力学	197
9.1 经典分子动力学	197
9.1.1 恒定能量的分子动力学	197
9.1.2 正则系综的分子动力学	200
9.1.3 分子动力学的实际应用方面	201
9.2 从头算分子动力学	202
9.3 从头算分子动力学的应用	205
9.3.1 考察结构复杂材料:液体和非晶相	205
9.3.2 考察复杂能量表面	208
练习	211
参考文献	211
扩展阅读	211
附录 计算细节	212
第 10 章 精确度及“标准”计算以外的其他方法	213
10.1 DFT 计算有多精确?	213
10.2 选择一个泛函	219
10.3 物理精度的例子	224
10.3.1 分子体系的基准计算——能量和几何构型	224
10.3.2 分子体系的基准计算——振动频率	226
10.3.3 晶体结构和内聚能	226
10.3.4 吸附能和成键强度	228
10.4 改善电子关联效应处理方式的 DFT + X 方法	229
10.4.1 色散相互作用和 DFT + D	230
10.4.2 自作用误差、强关联电子体系和 DFT + U	232
10.5 线性比例方法的更大尺寸体系和经典力场	234

10.6 结论	235
参考文献	236
扩展阅读	237

第 1 章 何谓密度泛函理论

1.1 如何使用本书

在许多物理学和工程领域,取得科学和技术进步的关键在于能够从原子或分子尺度,理解并调控物质的性质。对于描述原子和分子量子行为的基本关系式(薛定谔方程),密度泛函理论(Density Functional Theory, DFT)是一个非常好的求解方法,而且这些原子和分子结构也具有一定的实际含义。这种方法从量子力学理论前沿少数物理学家和化学家的一种专业科学,迅速成长为化学、物理学、材料科学、化学工程、地质学等领域众多研究者的常用工具。在科学引文索引(Science Citation Index)中,如果以“Density Functional Theory”作为标题名或摘要检索关键词,在 1986 年发表的文章中只检索出不到 50 条文献记录,而在 1996 年和 2006 年发表的文章中则分别检索出 1100 条和 5600 条以上的文献记录。

正如书名所言,本书的主旨是:从实践角度对如何使用 DFT 计算进行介绍。在阅读本书之前,并不需要读者已经完成过这些计算,或者甚至已经掌握了这些计算;作者所假设的是:读者希望明确这些方法可以用来做什么,或者读者能够对某一研究领域运行自己的计算,再或者是读者能够与使用这些方法的同行进行深入交流。

在这里,可以用汽车作为一个有助于解释上述意图的类比。在您学会如何驾驶之前,您对在驾驶汽车之后能够完成更多事情这一结果是具有心理预期的。对于您驾驶一辆汽车来说,最重要的是掌握如何操控汽车的一些基本概念(如需要定期将燃料加入汽车中、遵守基本的交通法规等),并花时间在各种路况下实际驾驶汽车。您并不需要

知道有关燃料喷射器是如何工作的所有细节,以及如何建立一个散热系统来有效地冷却引擎,或者是有关如何制造汽车的其他种种细节。这些细节只在读者试图完成某些特别困难的、与汽车有关的任务时才是重要的,例如独自驾车横跨南极洲。但如果仅仅是驾车穿过小镇到朋友家而后返回自己家,则无需理解这些有关汽车内部的细节。

通过本书,作者希望读者能够学会的是在做自己的计算时,或者是在你所感兴趣物理问题上,理解其他人所完成的计算时,能够学会“驾车穿过小镇”。如果您感兴趣于探索该领域理论方法发展的前沿,并“建造一辆更好的汽车”,作者会为您的热情鼓掌。但是,您也应该继续阅读本章,至少要找到一个必定能取得成功,并有可能获得诺贝尔奖的研究方向,然后还需要深入钻研本章末尾“扩展阅读”所列出的每一本书。

每章的末尾都给出了一系列练习题,大部分练习都是使用所在章节的理论思想进行实际计算。不要只是阅读本书,只有做这些练习才能很快提高知识和能力。因此,作者强烈建议在时间允许的前提下,尽可能地多做这些练习。

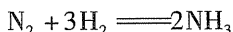
1.2 DFT 运行范例

在进一步定义密度泛函理论之前,我们先来展示一下在某些科学领域中它是如何使用的,而且这也非常有必要。在使用这些方法的科学领域中,已经发表了上千篇文献,本书原著者选择了三个范例。之所以选定这些范例,是因为它们能够表明:使用 DFT 计算对涉及相当广泛领域的复杂科学问题都具有重要作用,而且能够获得实验手段无法得到的信息。

1.2.1 通过异质催化合成氨

第一个范例是关于催化合成氨(NH_3)——一个很重要的工业制程。氨是农业所需化肥的核心组分,每年商业制造的氨超过一亿吨。根据估算,全世界能源的大约 1% 都用于氨的制造。氨合成的核心反

应非常简单,即



为了使该反应向正向发生,需要存在金属催化剂[如铁(Fe)和钌(Ru)等],且反应体系处于高温(大于400℃)和高压(大于100大气压)条件下。尽管早在大约100年前,Haber和其他人就已经确认了这些金属催化剂是非常有效的,但对这些催化剂表面发生的催化反应机理仍然不甚明了,一部分原因是由于催化剂的实际结构很复杂。为了使金属催化剂具有较大的表面积,可以把活性金属颗粒均匀弥散地分布在多孔材料中。这是纳米科技的一个广泛应用,且该应用早于纳米科技这一术语用于材料,并具有令人鼓舞的科学意义。为了理解金属纳米颗粒的化学活性,使用表面原子的局域配位关系对其进行表征是非常有意义的,因为正是这些不同的配位关系导致了不同的化学反应活性。根据表面原子的局域配位关系,可将表面原子分为若干“种类”。纳米颗粒表面通常包含有各种不同种类(不同配位关系)的原子,因此,表面反应活性整体上可以看作是纳米颗粒形状和每种原子反应活性的复杂函数。

以上讨论提出了一个基本问题:能否在金属纳米颗粒的形状尺寸,与作为合成氨反应催化剂的活性之间建立直接联系?如果能够找到这个问题的详细解答,就很有可能合成出更有效的催化剂。目前,对于该问题的详细解答来自于 Honkala 及其合作者^[1]对 Ru 纳米颗粒的 DFT 计算研究。他们的研究结果显示:总的化学反应可看做是由金属催化剂表面上12个单独步骤所完成的,这些步骤的反应速率很大程度上取决于所牵涉金属原子的局域配位关系,其中一个最重要的反应步骤就是在催化剂表面 N₂ 键的断裂。在类似于体相 Ru 的催化剂表面区域(更具体地说,就是平直原子区域),N₂ 键断裂所需的能量很大,意味着此时该反应速率极慢。如果催化剂表面的 Ru 原子形成了常见的表面台阶,则在靠近该台阶边缘处,该反应所需能量大大降低。采用附加 DFT(additional DFT)计算,Honkala 及其合作者预测了 Ru 纳米颗粒表面不同局域配位关系的相对稳定性,该方法可以将颗粒尺寸作为自变量,预测纳米颗粒的具体形状。该预测建立了纳米颗粒表面上,N₂ 键

断裂反应有效位置的数目与 Ru 纳米颗粒直径之间的精确联系。最终,所有这些计算都用来形成一个整体模型:如何把纳米颗粒表面不同种类金属原子处的单个反应速率联系起来,从而定义实际反应条件下的整体反应速率。整个过程中没有任何阶段使用实验数据进行拟合,或者调整模型,因此最终的结果完全是复合催化反应速率的一个纯预测性描述。完成以上所有研究后,Honkala 等人将其预测结果与 Ru 纳米颗粒催化剂在类似于工业反应条件下的实验测试值进行了比较。他们的预测结果在数值上与实验值是非常吻合的。

1.2.2 微量杂质引发的金属脆化

当读者阅读本书时,很可能在你周围就存在大量铜导线。铜是在元件和电器设备之间传递电流的最常用材料。除了其成本较低之外,实际应用中的另一个优点是铜比较软,塑性很好。常见的铜(以及其他)金属零件几乎都是多晶体,也就是这些零件中包含有大量微小的区域,称为晶粒,而每个晶粒都是规则排列的单晶。两个相邻晶粒具有相同的晶体结构和对称性,但它们在空间中的位向有所不同。因此,相比于纯金属所具有的晶体结构,晶粒之间的接触区域存在有大量复杂结构。这些接触区域存在于所有多晶体材料中,称为晶界。

早在 100 年前,人们就已经知道:在铜中加入少量的某种杂质,可以将其从塑性状态变成脆性断裂方式(断裂之前没有塑性变形)。例如,当在铜(Cu)中铋(Bi)的含量低于 $100\mu\text{g/g}$ 时,铜就会发生脆化。对于铅(Pb)和汞(Hg)杂质,也观察到类似的效果。但这一现象是如何发生的呢?定性地说来,当这些杂质引发脆性断裂时,断裂总是倾向于发生在晶界处,因此这些杂质一定是在某个方面极大地改变了晶界性质。当杂质元素 Bi 的含量很低时就会发生脆化现象,这并不是什么难以置信的事情,其原因在于 Bi 几乎不溶于体相 Cu 中。也就是说,Bi 原子更有可能偏聚在晶界处,而不是存在于晶粒内部,因此晶界上 Bi 的浓度要比该材料整体的平均浓度高得多。

能否对 Bi 引起铜的变化进行更为详细的解释呢?正如读者所期望的那样,对这个很多年前就已经观察到的有趣现象,研究者们已经给

出了几种不同的解释。其中:第一类解释把发生这一现象的原因归结于电荷作用。例如,Bi 原子与邻近 Cu 原子形成键合作用,使这些 Cu 原子要比纯 Cu 更硬,从而降低了 Cu 晶格平滑变形的能力。第二类解释是紧邻晶界的杂质原子可能改变了晶界处原子的电子结构,从而弱化晶界处原本的成键。第三种解释则归因于尺寸效应,认为 Bi 原子比 Cu 原子大得多,如果 Bi 原子存在于晶界处,则会分离界面两侧的 Cu 原子,使二者之间的自然间距增大。被拉大的键长将会弱化这些原子键,并使晶界处更容易断裂。第二和第三种解释都涉及到了晶界附近原子键的弱化,但对弱化给出了不同的根本原因。直接采用实验方法区分和辨别这些可能的原因机理是非常困难的。

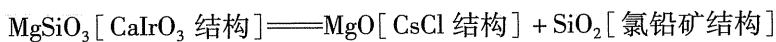
最近,Schweinfest、Paxton 和 Finnis 使用 DFT 计算方法,明确描述了 Bi 是如何脆化铜的,他们的研究题目就表露出了这一结论^[2]。针对纯 Cu 和含有 Bi 原子杂质的 Cu,他们首先使用 DFT 计算分别预测了应力-应变关系。如果上述原子键加强变硬的观点是正确的,那么在添加 Bi 后,金属的弹性模量将会增大。实际上,DFT 计算给出了相反的结果,意味着原子键强化解释是不正确的。另外,在更加困难的一系列计算中,他们直接给出了一些晶界的粘附功,这些晶界均已由实验证实会被 Bi 脆化。与实验观察结果相比较,定性地看来,计算结果具有很好的一致性:晶界的粘附功会由于 Bi 的存在而大幅降低。特别是 DFT 结果可以直接给出晶界原子的电子结构,据此发现,上述有关晶界电荷作用的解释并不是导致脆化的原因。晶界性质的巨大转变几乎全都可以从 Bi 原子所导致的超体积角度加以解释,也就是尺寸效应。该推理表明:对于任何比 Cu 原子尺寸更大的杂质原子,均强烈偏聚于晶界,从而可以使 Cu 脆化。这一描述实际上也适用于上文中提及的 Pb 和 Hg 杂质,这些杂质也已知会使铜发生脆化。

1.2.3 模拟行星形成时的材料性质

为了建立不同大小行星形成原因的详细模型,(相对于其他信息)必须要知道行星内部含有何种矿物质,以及这些矿物质传递热量的效果如何。行星内部所存在的极端条件对通过实验探索这些命题提出了

极大挑战。例如,木星中心的压强超过 40Mbar,温度超过 15000K。正如在 Umemoto、Wentzcovitch 和 Allen^[3]的研究工作中所展现的,DFT 计算能够成为探索极端条件下物质特性的有效工具。硅酸盐矿物对行星的形成非常重要,他们的研究正是围绕一种硅酸盐矿物质(MgSiO_3)的体相性质展开。在常温常压条件下, MgSiO_3 会形成相对常见的钙钛矿晶体结构。在 Umemoto 等人的计算研究之前,人们已知,如果把 MgSiO_3 置于类似于地核—地幔边界处的条件下,它会转变为一种不同的晶体结构,即 CaIrO_3 结构(通常,将每种晶体结构以所发现的具有该结构的第一种化合物来命名,此结构的命名就是该规则的一个例子)。

Umemoto 等人试图了解在比地核—地幔边界处更为严苛的条件下, MgSiO_3 结构会发生什么变化。他们采用 DFT 计算建立了相图,比较了固体 MgSiO_3 几种可能晶体结构所具有的稳定性。所有这些计算均针对体相材料,并且考虑了 MgSiO_3 分解为其他化合物的可能性。计算结果预测当压强在约 11Mbar 时, MgSiO_3 会发生分解,即



在该反应中,用方括号注明了每种化合物的晶体结构。一个有意思的特点是右侧的化合物均并非是它们在常温常压下的稳定结构。例如, MgO 在常温常压下倾向于形成 NaCl 结构(与食盐的晶体结构相同)。 SiO_2 的结构变化也类似,只不过更为复杂:这种化合物从常温常压下到 MgSiO_3 的分解条件下,经历了若干中间结构。这些 MgO 和 SiO_2 结构转变发生的条件,可以在实验室中达到并直接测试,因此能够在 DFT 计算和实验之间建立必要联系。使用 DFT 预测的相变压强与实验观测到的值是一致的,从而表明了这些计算的精确性。

从 Umemoto 等人采用 DFT 计算所预测的该分解反应出发,可以得到一个建立行星形成模型的重要启示。在最简单的层面上,它给出了大行星内部存在何种材料的新信息。例如,通过计算预测,天王星或海王星可能含有 MgSiO_3 ,但木星和土星则没有。在更深的层面上,材料的热力学特性可以用来对一些现象建模,如行星体系内部的对流(Convection)现象。Umemoto 等人推测,对于“高密度土星(Dense - Saturn)”(在太阳系外侧发现的一个类似于土星的行星,其质量大约是地

球的 67 倍),上述分解反应很可能极大地限制其内部的对流现象。

对于类似于上述反应的理论预测,一个合理的疑虑是:很难想象如何把这些计算结果与实验数据相比较并验证。幸运的是,对于在实验条件所能达到的压强下发生的类似反应,DFT 计算也能够给出相应结果。使用这种方法,预测得到 NaMgF_3 也经历了类似于 MgSiO_3 的一系列转变,即在高于常压的某一压强下,从钙钛矿结构转变为后钙钛矿结构(Postperovskite),随后在更高压强下分解为 NaF 和 MgF_2 ^[4]。预测这一分解发生在压强为 0.4Mbar 左右,远低于 MgSiO_3 的平衡分解压强。相对于相变机理的直接实验测试,这些预测意味着一个很好的研究途径,且 DFT 计算结果对研究行星形成也具有很大作用。

我们能够就 DFT 计算对许多科学研究领域所产生的重大影响给出很多类似例子。但愿读者能从这三个范例中明白 DFT 计算对科学认知的重大影响,并得到该研究方法特点的一些启示。对于这三个范例,思考其相通之处是非常有好处的,三者均与固体材料相关,尽管第一个例子主要关注的是固体和气体的界面;在每个例子中,都对由材料原子尺度性能所决定的物理问题,给出了相关信息,而这些工作对于实验研究都是极度困难的;在每个例子中,计算所给出的信息不仅仅是一些理论上的理想状态,而是重在理解不同温度、压强和化学组成下的相关现象,而这些均与实际应用直接相关。

1.3 薛定谔方程

到现在为止,读者应该已经知道密度泛函理论是一个有用和有趣的话题。但它究竟是什么?让我们从实际观察开始,20 世纪最深远的一个科学进展是量子力学,无数实验观测已经证实该理论描述了我们周围宇宙万物的规律,并且具有惊人的正确性。

我们在这里要先回顾一下,在 DFT(及其他形式计算化学)的理论基础中,从量子力学出发的一些重要思想。在这里,我们并不是想要说明 DFT 所采用各种技术的全部推导过程。相反,我们的目的是对 DFT 中那些重要的、最为基本的公式给出一个清晰、简明、介绍性的描述。

对于量子力学整个内容而言,已经有很多极为出色的教材,其中一些已经列在本章之后的扩展阅读部分。

请试想一下,当我们将要描述规则排列原子集合(可以试着比较一个隔离态的孤立分子,以及一个矿物晶体中的有序原子)的性质时,我们希望知道的一个基本信息是这些原子的能量,而更为重要的是,当移动这些原子时,它们的能量将如何变化。为了定义一个原子的位置,我们不仅需要定义它的原子核在哪,也需要定义这个原子的所有电子在哪。在将量子力学应用于原子体系时,一个基本事实是原子核的质量远大于单个电子质量。原子核中的每个质子或中子要比一个电子的质量大 1800 倍。这意味着,通俗地说来,电子对环境变化的响应远比原子核快得多。因此,我们可以把所研究的物理问题分割为两个部分。第一部分,我们固定原子核的位置,求解描述电子运动的方程组。对于在给定原子核势场中运动的一系列电子,我们可以得到能量最低的电子构型,或电子态(State)。最低能量态亦称为电子的基态(Ground State),而将原子核和电子分为各自独立数学问题的是 Born - Oppenheimer 近似。如果我们有 M 个在 $\mathbf{R}_1, \dots, \mathbf{R}_M$ 位置的原子核,那么我们可以把基态能量 E 表示为这些原子核位置的函数,即 $E(\mathbf{R}_1, \dots, \mathbf{R}_M)$ 。这个方程亦称为这些原子的绝热势能面(Adiabatic Potential Energy Surface)。一旦能够计算得到该势能面,我们就能够处理上面所给出的根本性问题——当移动原子时,材料的能量将如何变化?

Schrödinger 方程的一个简单形式(更准确地说,与时间无关的非相对论 Schrödinger 方程),就是大家熟悉的 $H\psi = E\psi$ 。该方程具有精妙的形式,但为了更好地理解它,我们需要定义其中出现的几个量。方程中 H 是哈密尔顿算符(Hamiltonian Operator), ψ 是哈密尔顿量的一套求解,或本征值。这些解中的每一个(ψ_n),对应于相应的本征值 E_n ,这是满足本征值方程的真实值^①。哈密尔顿量的详细定义取决于 Schrödinger 方程所描述的物理体系。已经有几个众所周知的例子,如盒子中的粒子、谐振子等,其哈密尔顿量具有简单的形式,并且

① 函数 ψ 的数值是一些复数,但 Schrödinger 方程的本征值是实数。

Schrödinger 方程也能够被精确求解。我们所感兴趣的多电子与多原子核交互作用体系是更为复杂的情形。在这种情形下, Schrödinger 方程可写成一个更完整的形式, 即

$$\left[-\frac{\hbar^2}{2m} \sum_{i=1}^N \nabla_i^2 + \sum_{i=1}^N V(\mathbf{r}_i) + \sum_{i=1}^N \sum_{j<i}^N U(\mathbf{r}_i, \mathbf{r}_j) \right] \psi = E\psi \quad (1.1)$$

式中: m 为电子质量。

方括号中的三项依次分别是每个电子的动能、每个电子与所有原子核之间的作用能、不同电子之间的作用能。对于我们所选定的哈密顿量, ψ 是电子波函数, 它是 N 个电子每个电子空间坐标的函数, 即 $\psi = \psi(\mathbf{r}_1, \dots, \mathbf{r}_N)$, E 是电子的基态能量^①。基态能量与时间无关, 因此这是与时间无关的 Schrödinger 方程^②。

尽管电子波函数是所有 N 个电子坐标的函数, 但仍可以将 ψ 近似表达为单个电子波函数的乘积, 即 $\psi = \psi_1(\mathbf{r})\psi_2(\mathbf{r}) \cdots \psi_N(\mathbf{r})$ 。该波函数表达式就是所谓的 Hartree 乘积 (Hartree Product), 并且将全波函数近似表达为单电子波函数的这种乘积形式, 也有一定有益的动机。因为每个原子仅有一个原子核, 但含有许多电子, 所以电子数目 N 要比原子核数目 M 大得多。假如研究对象为单个 CO_2 分子, 其全波函数是一个 66 维函数 (22 个电子, 每个电子 3 维); 再假如研究对象为 100 个 Pt 原子的纳米团簇, 其全波函数则超过了 23000 维! 这些数据可以使读者了解到一个概念: 为什么求解实际材料的 Schrödinger 方程让许多顶尖学者困扰了大半个世纪。

当我们再回过头来看一下哈密顿量 H , 情形看起来似乎更糟糕了。从求解方程的观点来看, 哈密顿量 H 中关于电子-电子相互作用能这一项是最为重要的。而该项的数学形式意味着: 为了确定上述单电子波函数 $\psi_i(\mathbf{r})$, 需要同时考虑该电子与其他所有电子有关的波函数。换言之, Schrödinger 方程是一个多体问题。

① 为了更清楚地表述, 此处忽略了电子自旋。在完整表达式中, 每个电子是由三个空间变量及其自旋所定义的。

② 电子动力学是由与时间相关的 Schrödinger 方程所定义的, 即 $i\hbar(\partial\psi/\partial t) = H\psi$, 其中, $i = \sqrt{-1}$, 这清楚地表明了其波函数为复数函数, 并非实数函数。

尽管求解 Schrödinger 方程可以视作量子力学的基本问题,但仍需明白一点,并不能直接观测到基于某套特定坐标的波函数。(理论上)能够测量的量是 N 个电子在某套特定坐标 $(\mathbf{r}_1, \dots, \mathbf{r}_N)$ 出现的概率。这个概率值等于 $\psi^*(\mathbf{r}_1, \dots, \mathbf{r}_N)\psi(\mathbf{r}_1, \dots, \mathbf{r}_N)$, 其中星号表示一个共轭复数。需要进一步注意的是实验上我们通常并不关心材料中哪个电子是电子 1、电子 2...等,而且即便我们关心这一问题,也没法容易地标注这些原子。这意味着在物理上我们感兴趣的物理量其实就是一套 N 个电子位于某套坐标 $\mathbf{r}_1, \dots, \mathbf{r}_N$ 时的概率值。与此概率值密切相关的物理量是空间中某个具体位置上的电荷密度 $n(\mathbf{r})$ 。可以用单电子波函数的形式将其写为

$$n(\mathbf{r}) = 2 \sum_i \psi_i^*(\mathbf{r})\psi_i(\mathbf{r}) \quad (1.2)$$

此式中,针对电子系统所占据的全部单电子波函数进行求和。因此,求和项中的表达式就是在单电子波函数 $\psi_i(\mathbf{r})$ 中一个电子位于 \mathbf{r} 处的概率值。表达式中出现因子 2 是因为电子具有自旋,且泡利不相容原理表明:每个单电子波函数能够被不同自旋的两个电子所占据。这是纯粹的量子力学效应,而在经典物理中并没有与此相类似的地方。上述讨论中的要点在于电荷密度 $n(\mathbf{r})$,它仅是三个坐标的函数,却含有 Schrödinger 方程全波函数解的大量信息(可以从物理上实际观测到全波函数),但是全波函数是 $3N$ 个坐标的函数。

1.4 密度泛函理论——从波函数到电荷密度

密度泛函理论全部建立在由 Kohn 和 Hohenberg 所证明的两个基本数学定理,以及由 Kohn 和 Sham 在 1960 年代中期所推演的一套方程的基础上。其中,由 Kohn 和 Hohenberg 所证明的第一个定理是:从 Schrödinger 方程得到的基态能量是电荷密度的唯一函数。

该定理表明:在基态波函数和基态电荷密度之间,存在一一对应的关系。为了明确这一结果的重要性,需要先弄清“泛函”这一概念。从字面上看,泛函(Functional)与大家熟悉的函数(Fuction)非常近似。函

数的取值是由一个或多个变量所决定的,且由该变量唯一地确定了一个数值。一个简单的例子是单变量函数,如 $f(x) = x^2 + 1$ 。泛函有点类似,但其取值是由一个函数作为变量所决定的,且该函数唯一地确定了泛函的一个数值。例如

$$F[f] = \int_{-1}^1 f(x) dx$$

就是函数 $f(x)$ 的泛函。如果以 $f(x) = x^2 + 1$ 为例计算该泛函的数值,则 $F[f] = \frac{8}{3}$ 。据此,可以将 Hohenberg 和 Kohn 的结果重新表述为:基态能量 E 可以表达为 $E[n(r)]$,其中 $n(r)$ 是电荷密度,这也就是为什么将这一理论称为密度泛函的原因。

另一个表述 Hohenberg 和 Kohn 结果的方式是:基态电荷密度唯一决定了基态的所有性质,包括能量和波函数。这个结果为什么很重要?因为它意味着:可以通过找到含有三个空间变量的电荷密度函数来求解 Schrödinger 方程,而不用求解得到含有 $3N$ 个变量的波函数。在此处“求解 Schrödinger 方程”的意思,更准确地说,是得到基态能量。据此,对于含有 100 个 Pd 原子的纳米团簇而言,上述理论将该问题从 23000 维简化为一个只有 3 维的问题。

然而,尽管第一个 Hohenberg - Kohn 定理严格证明了存在一个可用来求解 Schrödinger 方程的电荷密度泛函,但该理论并没有给出这个泛函的具体形式。第二个 Hohenberg - Kohn 定理给出了这个泛函的一个重要特征:使整体泛函最小化的电荷密度就是对应于 Schrödinger 方程完全解的真实电荷密度。如果已知这个“真实的”泛函形式,那就能够通过不断调整电荷密度直到由泛函所确定的能量达到最小化,并可以找到相应的电荷密度。实际上,该变分原理常用于泛函的近似表达形式。

将 Hohenberg - Kohn 定理所描述的泛函写成单电子波函数 $\psi_i(r)$ 的形式,是一个有益的方法。在式(1.2)中,这些泛函整体上定义了电荷密度 $n(r)$ 。能量泛函可以写为

$$E[\{\psi_i\}] = E_{\text{known}}[\{\psi_i\}] + E_{\text{XC}}[\{\psi_i\}] \quad (1.3)$$

其中,将泛函分开为能够写成简单解析形式的一项 $E_{\text{known}}[\{\psi_i\}]$,以及

所有其他部分 E_{xc} 。标有“known”这一项包含有四方面的贡献,即

$$E_{\text{known}}[\{\psi_i\}] = -\frac{\hbar^2}{m} \sum_i \int \psi_i^* \nabla^2 \psi_i d^3r + \int V(\mathbf{r}) n(\mathbf{r}) d^3r \\ + \frac{e^2}{2} \iint \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r d^3r' + E_{\text{ion}} \quad (1.4)$$

在上式右侧,依次分别是:电子的动能,电子和原子核之间的库伦作用,电子之间的库伦作用,原子核之间的库伦作用。对于 V_{xc} 能量泛函完整表达式中的另一项 $E_{xc}[\{\psi_i\}]$,是交换关联泛函,它所定义的是没有包括在“known”这一项中所有其他的量子力学效应。

现在试想一下,我们可以采用更好的方式表达尚未确切定义的交换-关联能泛函。在寻找总能泛函最小能量解的过程中,都用到了什么?与求解 Schrödinger 方程波函数这一巨大困难相比,目前为止,我们似乎仍没有把握使这项工作变得更加容易。这个难题是由 Kohn 和 Sham 解决的,他们给出了这样一个结果:求解正确的电荷密度可以表示为求解一套方程,而其中每个方程都只与一个电子有关。

Kohn - Sham 方程的表达式为

$$\left[-\frac{\hbar^2}{m} \nabla^2 + V(\mathbf{r}) + V_H(\mathbf{r}) + V_{xc}(\mathbf{r}) \right] \psi_i(\mathbf{r}) = \varepsilon_i \psi_i(\mathbf{r}) \quad (1.5)$$

表面上看来,该方程类似于式(1.1),主要的区别就是 Kohn - Sham 方程中去掉了完全 Schrödinger 方程式(1.1)中的加和符号。这是由于 Kohn - Sham 方程的解是只取决于三个空间变量的单电子波函数 $\psi_i(\mathbf{r})$ 。在 Kohn - Sham 方程左侧,有三个势能项 V 、 V_H 和 V_{xc} 。其中,第一个也出现在完全 Schrödinger 方程式(1.1)中,以及上述总能泛函的“known”部分式(1.4),这个势能定义的是一个电子与所有原子核之间的相互作用。第二个也称为 Hatree 势能,可以写为

$$V_H(\mathbf{r}) = e^2 \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3r' \quad (1.6)$$

这个势能描述的是一个 Kohn - Sham 方程所考虑的单个电子,与该问题中全部电子所产生的总电荷密度之间的库伦排斥作用。Hatree 势能包含了一个所谓的自作用(Self - Interaction)部分,因为 Kohn - Sham 方程中所描述的那个电子也是总体电荷密度的一部分,所以 V_H 中有一

部分是那个电子与其自身的库伦作用。自作用是物理上不存在的,在 Kohn - Sham 方程的完整表达式中,将其合并到单电子方程的交换作用能中进行修正,且这只是 V_{xc} 中的一部分。 V_{xc} 可以在形式上可以表示为交换关联能的“泛函导数”,即

$$V_{xc}(\mathbf{r}) = \frac{\delta E_{xc}(\mathbf{r})}{\delta n(\mathbf{r})} \quad (1.7)$$

与函数的导数相比,这个泛函导数的严格数学定义式似乎有点不易理解,但在概念上,可以把它看作是一个普通的导数。这个泛函导数的写法中用了 δ , 而不是 d , 以凸显它并不是一个普通的导数。

在以上有关 Kohn - Sham 方程的讨论中,如果你感到似乎是陷入了一个循环,那么这个感觉是对的。为了求解 Kohn - Sham 方程,需要确定 Hartree 势能;而为了得到 Hartree 势能,又需要知道电荷密度;但为了找到电荷密度,又必须知道单电子波函数方程;为了知道这些波函数,又必须求解 Kohn - Sham 方程。为了打破这一循环,这个问题通常用迭代算法来处理,其过程简述如下:

(1) 定义一个初始的、尝试性的电荷密度 $n(\mathbf{r})$ 。

(2) 求解由尝试性的电荷密度所确定的 Kohn - Sham 方程,得到单电子波函数 $\psi_i(\mathbf{r})$ 。

(3) 计算由第(2)步 Kohn - Sham 单粒子波函数所确定的电荷密度,即 $n_{KS}(\mathbf{r}) = 2 \sum_i \psi_i^*(\mathbf{r})\psi_i(\mathbf{r})$ 。

(4) 比较计算得到的电荷密度 $n_{KS}(\mathbf{r})$ 和在求解 Kohn - Sham 方程时使用的电荷密度 $n(\mathbf{r})$ 。如果两个电荷密度相同,则这就是基态电荷密度,并可将其用于计算总能。如果两个电荷密度不同,则用某种方式对尝试性电荷密度进行修正,然后再从第(2)步重新开始。

在这里,虽然省略了整个过程中的一些重要细节(例如,在两个电荷密度之间,二者接近到什么程度可以认为是相同的?采用什么方式才能较好地修正电荷密度?应当怎样确定初始电荷密度?),但仍然能够说明这个迭代方法是如何求解 Kohn - Sham 方程的,这也就是自洽 (Self - Consistent) 求解过程。

1.5 交换关联泛函

先简要回顾一下我们已经学习的部分,我们本想求解 Schrödinger 方程的基态能量,但因为是多体问题,所以非常困难。根据 Kohn、Hohenberg 和 Sham 所给出的漂亮结果,可以通过将能量泛函的能量最小化,得到其基态,并且这可由一套单粒子方程的自洽解所给出。在这个设想中,只有一个关键的难点:求解 Kohn - Sham 方程,必须给定交换关联方程 $E_{XC}[\{\psi_i\}]$ 。正如式(1.3)和式(1.4)所表明的那样,确定 $E_{XC}[\{\psi_i\}]$ 是非常困难的。要知道,在式(1.4)中,已经明确写出了所有“容易”的部分。

实际上,我们并不清楚交换关联泛函的真实形式,尽管 Hohenberg - Kohn 定理肯定它是确定存在的。幸运的是,对于均匀电子气这种情形,该泛函可以直接导出。在此情形下,电荷密度在空间所有点上都是常数,即 $n(\mathbf{r}) = \text{常数}$ 。对于任何真实材料而言,这种情形的意义可能不大,其原因在于:正是由于电荷密度的变化才确定了化学键,也才使材料更有意义。但均匀电子气给出了实际使用 Kohn - Sham 方程的可行方法。为了做到这一点,我们把每个位置的交换关联势能都设定为已知的交换关联势能,这个已知的交换关联能是根据该位置所观测到的电荷密度,由均匀电子气得到的,即

$$V_{XC}(\mathbf{r}) = V_{XC}^{\text{electron gas}}[n(\mathbf{r})] \quad (1.8)$$

这一近似仅仅使用了局域密度来确定近似的交换关联泛函,因此称作局域密度近似(Local Density Approximation, LDA)。LDA 使我们能够完全确定地写出 Kohn - Sham 方程,但需要注意的是:用这些方程所得的结果并不能严格求解真实的 Schrödinger 方程,因为我们并没有使用真实的交换关联泛函。

需要说明的是, LDA 并非是用于求解 DFT 计算的唯一泛函。开发出更接近实际情形的泛函,仍是量子化学领域中一个重要的、活跃的研究方向。我们在本章的开头就说,存在可以获得诺贝尔奖的问题,下面所述就是其中一个:开发出精确表示物质本质的泛函,并将该泛函表达

为对大量原子也能快速求解的数学形式(这个建议似乎有点像 Hohenberg - Kohn 定理, 仅仅说出了某种存在的东西, 而并没有给出任何找到它的线索)。

如若你能够严格求解该问题, 就可能成为一个家喻户晓的名字(至少在学术圈内)。即使如此, 仍存在许多已被广泛采用的近似泛函, 而且这些近似泛函对于许多不同种类的物理问题, 都能给出较好的结果。本书的主要目的就是帮助读者懂得如何使用这些已经存在的泛函来进行计算。在 LDA 之后, 最为广泛知名的一类泛函是局域电荷密度和电荷密度上的局域梯度, 这种方法所给出的就是广义梯度泛函(Generalized Gradient Approximation, GGA)。它吸引人的一点在于: 相比于 LDA, GGA 包含有更多的物理信息, 因此理应更为精确, 然而结果并不总是这样。

因为存在有许多不同的方法, 将电荷密度的梯度信息包含在 GGA 泛函中, 所以存在有许多不同的 GGA 泛函。在与固体相关的计算中, 最为广泛使用的两个泛函是 Perdew - Wang 泛函(PW91)和 Perdew - Burke - Ernzerhof 泛函(PBE), 这两种泛函都是 GGA 泛函。此外, 还开发和使用了十几种其他的 GGA 泛函, 特别是对于隔离分子的计算。对于同一个特定的原子构型, 因为不同的泛函会给出有所不同的结果, 所以有必要说明在某一计算中所使用的具体泛函, 而不是简单标注为“DFT 计算”。

关于 GGA 泛函的描述, 包括了电荷密度和电荷密度的梯度信息, 这表明可以根据更多其他物理信息, 建立更加复杂的泛函。实际上, 可以建立泛函的等级, 随等级升高, 则逐步包含了更多的物理信息。有关泛函等级的更多内容请参阅第 10.2 节。

1.6 量子化学的“旅行者”

除了 DFT 之外, 当读者浏览 Schrödinger 方程的其他数值求解方法时, 可能会很快碰到相当多令人眩晕的缩略词。这种经历有点类似于逗留在一个异国他乡的繁华都市。你也许会觉得这个新城市很美, 而且你确实希望能领略它的各种优秀之处, 但并不打算永久居住在那里。在这

个旅程开始之前,当你尚未踏足这个国家时,你可能会花几年时间去学习她的语言、历史、文化和地理。但对于一个短期逗留者而言,你更感兴趣的可能是与已经去过那里的朋友谈一谈、阅读一些旅行指南、翻阅常用语手册,或者是试着找出几家当地有名的餐馆。本节的目的就是:对量子化学方法进行一个类似于常用语手册或旅行指南的总体介绍。

1.6.1 局域化的和空间扩展的函数

将量子化学计算进行分类的一个有效方法是:根据它们解的函数表达形式来分类。广泛说来,这些方法要么使用了空间局域化函数(Spatially Localized Functions),要么使用了空间扩展函数(Spatially Extended Functions)。作为一个空间局域化函数的例子,图 1.1 显示了这样一个函数,即

$$f(x) = f_1(x) + f_2(x) + f_3(x) \quad (1.9)$$

式中

$$f_1(x) = \exp(x^2)$$

$$f_2(x) = x^2 \exp(x^2/2)$$

$$f_3(x) = \frac{1}{10} x^2 (1-x)^2 \exp(x^2/4)$$

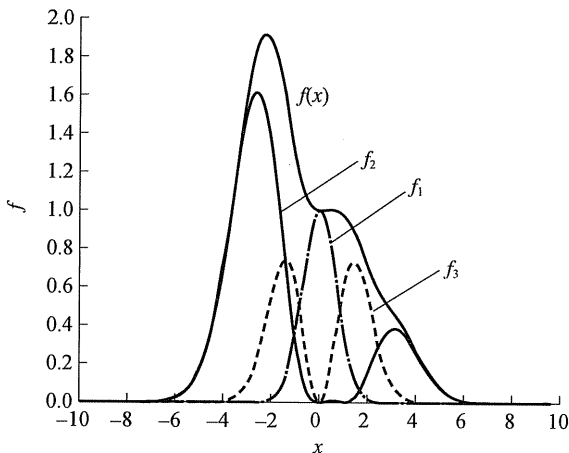


图 1.1 空间局域化函数的示例

图 1.1 中也显示出了 f_1 、 f_2 和 f_3 。当 $|x|$ 数值较大时,所有这些函数都快速地趋近于零。这类函数非常适合于用来表征波函数,或者是表征一个隔离原子的电荷密度。这个例子包含了这样一个思想:可以根据不同的空间范围、不同的对称性等因素,将多个函数合并为一个总体函数。我们也可以通过合并更多的单个函数,在这个最终函数中包含更多的信息。而且,为了建立描述多个原子的函数,只需要针对每个原子,使用一套合适的局域化函数。

由于隔离分子的波函数实际上是在距分子很远处衰减为零的,因此空间局域化函数是一种对隔离态分子的量子化学计算非常有用的架构。但是,假如我们所关注的是体相材料,例如固态硅中的原子或者金属催化剂表面以下的原子,那又会怎样呢? 我们仍然可以使用空间局域化函数来描述每个原子,并将这些函数进行加和来描述整个材料,但这肯定不是前进的唯一途径。一个较好的替换方案是使用周期性函数来描述波函数或电荷密度。图 1.2 是这一思想的一个简单示例,其中有

$$f(x) = f_1(x) + f_2(x) + f_3(x)$$

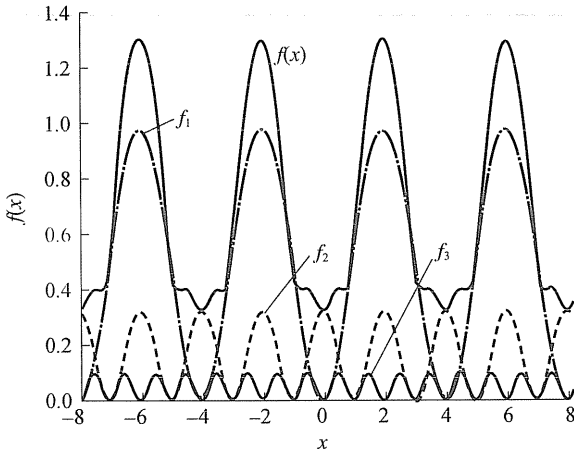


图 1.2 空间周期性函数的示例

式中

$$f_1(x) = \sin^2\left(\frac{\pi x}{4}\right)$$
$$f_2(x) = \frac{1}{3}\cos^2\left(\frac{\pi x}{2}\right)$$
$$f_3(x) = \frac{1}{10}\sin^2(\pi x)$$

最终函数是周期性的,即

$$f(x + 4n) = f(x)$$

式中: n 为任何整数。这种函数对于描述体相材料非常有用,因为至少对于无缺陷的材料而言,电荷密度和波函数实际上都是空间周期性函数。

对于隔离态分子而言,空间局域化函数是很自然的选择,所以在化学领域发展出的绝大多数量子化学方法都是基于这些函数。而物理学家历来对体相材料更感兴趣,在物理领域数值求解 Schrödinger 方程的方法主要是空间周期性函数。注意不要认为其中某个方法是“对的”,而另一个是“错的”,因为这两种方法各有优点和不足。

1.6.2 波函数方法 (Wave – function – based Methods)

对于量子化学计算的第二个基本分类方法是根据所计算的物理量来分类。在前面对 DFT 的介绍中,已经强调了 DFT 是为了计算电子的密度,并不是电子的波函数。然而,有许多计算方法,它们的目的就是得到全电子波函数。这些以波函数为基础的方法具有一个 DFT 计算所不具备的特别优势,那就是该方法具有一个明确的等级体系,假如计算时间无限长,就可以收敛到 Schrödinger 方程的精确解。我们无法在短短几个段落中充分阐述这一方法,在本章末的扩展阅读中列出了几个非常好的相关介绍。DFT 和波函数方法之间的强烈关联性,以及二者在科学上的重要性在 1998 年得以确认:Walter Kohn 发展了 DFT 理论基础,而 John Pople 开创性地开发出了计算原子和分子电子结构的量子化学计算机代码,当年的诺贝尔化学奖被共同授予这两人。有意思的是,这不是诺贝尔化学奖或物理奖第一次颁授给一个数值方法

(更准确地说,是一类数值方法)的开发工作,并不是什么罕见的科学发现。Kohn 在诺贝尔获奖辞中对波函数计算方法和 DFT 计算方法的优点和不足进行了饶有趣味的描述^[5]。

在对波函数方法进行简述之前,应当先对描述波函数的常用方法加以说明。前面我们提到了一个 N 粒子体系的波函数是一个 N 维函数。但确切地说,波函数是怎样的? 因为我们想要用波函数给出一个 N 电子体系的量子力学描述,这些波函数必须满足真实电子所具有的几个数学特征。例如,在相同的物理位置,泡利不相容原理不允许两个电子同时处于相同的自旋^①。我们当然希望在所建立的某种波函数近似形式中,能够保留这些特征。

1.6.3 Hartree – Fock 方法

假如我们想要得到 N 电子体系波函数的近似表达式,假设电子彼此之间没有作用,则电子的哈密顿量就可以写为

$$H = \sum_{i=1}^N h_i \quad (1.10)$$

式中 h_i 描述了电子 i 的动能和势能。如果我们忽略电子 – 电子之间的相互作用,在式(1.1)中所写出的全电子哈密顿量就是这种形式。根据这个哈密顿量,当只写出一个电子的 Schrödinger 方程时,其解满足

$$h\chi = E\chi \quad (1.11)$$

式(1.11)所定义的本征方程称为自旋轨道(Spin Orbitals)。对于每个单电子方程,都存在多个本征函数,这样就定义了一系列自旋轨道 $\chi_j(\mathbf{x}_i)$ ($j=1,2,\dots$),其中 \mathbf{x}_i 是定义电子 i 位置及其自旋态(向上或向下)的坐标矢量。我们将自旋轨道 $\chi_j(\mathbf{x}_i)$ 的能量表示为 E_j ,这对自旋轨道的标注是很有用的。 $j=1$ 的轨道具有最低的能量, $j=2$ 的轨道具有

① 自旋是量子力学特征,在经典力学中不存在。一个电子可以具有两种不同自旋中的一种,向上自旋或向下自旋。对于一个电子态的所有标定必定包含它的位置和它的自旋。泡利不相容原理只应用于自旋态相同的电子。

第二高^①的能量,依次类推。总的哈密顿量是单个电子算符 h_i 的简单相加,它具有以下规律: H 的本征函数是单个电子自旋轨道的乘积,即

$$\psi(\mathbf{x}_1, \dots, \mathbf{x}_N) = \chi_{j_1}(\mathbf{x}_1)\chi_{j_2}(\mathbf{x}_2)\cdots\chi_{j_N}(\mathbf{x}_N), \quad (1.12)$$

这个波函数的能量是自旋轨道能量之和,即 $E = E_{j_1} + \cdots + E_{j_N}$ 。在第 1.3 节,我们已经对 N 电子体系波函数的这种近似(Hatree Product)进行了简要介绍。

很可惜,Hatree Product 并不满足波函数所有的重要准则。因为电子是费米子,如果两个电子相互交换位置,其波函数一定改变符号,这就是众所周知的反对称原理。而交换两个电子并不会改变 Hatree Product 的符号,这是一个严重的缺陷。我们可以使用 Slater 行列式得到波函数更好的近似形式。在一个 Slater 行列式中, N 电子波函数是由一种满足反对称原理的方式,将单电子波函数合并形成的。可以把整体波函数表达为单电子波函数矩阵的行列式。以两个电子为例,可以很清楚地明白其中的道理。对于两个电子而言,Slater 行列式为

$$\begin{aligned} \psi(\mathbf{x}_1, \mathbf{x}_2) &= \det \begin{bmatrix} \chi_j(\mathbf{x}_1) & \chi_j(\mathbf{x}_2) \\ \chi_k(\mathbf{x}_1) & \chi_k(\mathbf{x}_2) \end{bmatrix} \\ &= \frac{1}{\sqrt{2}} [\chi_j(\mathbf{x}_1)\chi_k(\mathbf{x}_2) - \chi_j(\mathbf{x}_2)\chi_k(\mathbf{x}_1)] \end{aligned} \quad (1.13)$$

式中:系数 $\frac{1}{\sqrt{2}}$ 为正交化因子。

该表达式隐性描述了物理上的电子交换,如果两个电子相互交换,则其符号改变。该表达式还有其他优点。例如,它并没有对电子进行区分,而当两个电子具有相同的坐标时,或者当两个单电子波函数相同时,它就消失了。这表明 Slater 行列式满足泡利不相容原理的条件。Slater 行列式能够很容易推广到 N 电子体系,此时它是一个单电子自旋轨道 $N \times N$ 矩阵的行列式。使用 Slater 行列式,可以确保求解 Schrödinger 方程时是包含了交换效应的。但是,为了达到良好的计算

^① 原文如此,似乎应是“第二低”——译者注。

精度,这并不是我们所需要描述的唯一一种电子关联效应。上述讨论似乎显得并没有多大帮助,因为在所要求解的体系中,所有电子都是彼此相互作用的。在如何对电子互作用进行近似处理这一方面,求解 Schrödinger 方程的不同波函数方法是有所不同的。为了搞清所能使用的近似方法的不同类型,需要先仔细考虑一下最简单的方法,即 Hatree - Fock 方法。在 Hatree - Fock 计算和前面我们介绍过的 DFT 计算之间,也存在许多近似之处,因此深入理解 Hatree - Fock 方法可以使我们从不同角度来对待这两个方法。

在 Hatree - Fock (HF) 计算中,固定了原子核的位置,以确定 N 电子互作用体系的波函数。在 HF 计算中的第一部分,是确定所需要求解的方程。对于每个电子, Schrödinger 方程可以写为

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r}) + V_H(\mathbf{r}) \right] \chi_j(\mathbf{x}) = E_j \chi_j(\mathbf{x}) \quad (1.14)$$

左侧的第三项就是与式(1.5)中 Hatree 势能相同的一项,即

$$V_H(\mathbf{r}) = e^2 \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3 r' \quad (1.15)$$

简单说来,这意味着单个电子“感受”其他电子的作用,就如同一种平均的作用,而不是当电子在空间接近时,就能立即感受到所产生的排斥力。将式(1.14)与式(1.15) Kohn - Sham 方程相比较,可以发现二者间的唯一区别是 Kohn - Sham 方程中出现了额外的交换关联势能项。

为了对 HF 方法进行完整说明,我们还必须确定:如何表达上述单电子方程的解? 如何将这些解合并为 N 电子波函数? 在 HF 方法中,可以使用一个 Slater 行列式近似表示全部波函数。这意味着可以得到单电子方程的 N 个最低能量自旋轨道,即 $\chi_j(\mathbf{x}) (j=1, \dots, N)$; 并且,这些自旋轨道的 Slater 行列式也就形成了总的波函数方程。

在具体计算中为了实际求解单电子方程,就必须使用有限的信息来定义自旋轨道,因为无法在计算机上描述任意的连续函数。为了实现这一点,可以定义一个包含有限个方程的基组,用这些方程的加和近似表达真实的自旋轨道。若将这些方程写成 $\phi_1(\mathbf{x}), \phi_2(\mathbf{x}), \dots, \phi_K(\mathbf{x})$, 则可以把自旋轨道近似表达为

$$\chi_j(\mathbf{x}) = \sum_{i=1}^K \alpha_{j,i} \phi_i(\mathbf{x}) \quad (1.16)$$

使用这个表达式,只需要给出 $i = 1, \dots, K$ 且 $j = 1, \dots, N$ 时的扩展系数 $\alpha_{j,i}$, 就可以完全定义 HF 方法中使用的所有自旋轨道。 $\phi_1(\mathbf{x}), \phi_2(\mathbf{x}), \dots, \phi_K(\mathbf{x})$ 这一套函数称为计算时的基组 (Basis Set)。直观来看,所使用的基组越大(增大 K), 则计算精度相应提高,但也会增加求解所需的工作量。类似地,选择与真实材料中实际自旋轨道类型相近似的基函数 (Basis Functions), 也能够提高 HF 计算的精度。正如在第 1.6.1 节所提示的,当所研究的材料类型不同时,这些函数的特点就应有所区别。

现在,已经具备了进行 HF 计算的条件——可以将单个自旋轨道展开的基底,自旋轨道必须满足这些方程,以及一旦确定了自旋轨道,就能够形成最终波函数的描述方式。但是,仍然必须要解决一个关键问题,这个问题也存在于第 1.4 节有关 Kohn - Sham 方程的讨论中。为了得到自旋轨道,我们必须求解单电子方程;为了确定单电子方程中的 Hartree 势能,我们又必须知道电荷密度;为了确定电荷密度,我们又必须确定电子波函数,而这可以根据单个自旋轨道得到。为了打破这一循环,采用迭代方法求解 HF 计算,其过程简述如下:

(1) 通过指定扩展因子 $\alpha_{j,i}$, 建立自旋轨道的一个初始估值

$$\chi_j(\mathbf{x}) = \sum_{i=1}^K \alpha_{j,i} \phi_i(\mathbf{x})。$$

(2) 从自旋轨道的当前估值出发,确定电荷密度 $n(\mathbf{r}')$ 。

(3) 根据第(2)步得到的电荷密度,求解单电子方程,得到自旋轨道。

(4) 如果第(3)步得到的自旋轨道与第(2)步使用的自旋轨道一致(或相容),则它们就是 HF 问题的解;如果并不一致,则必须建立一个自旋轨道新的估值,然后重新从第(2)步开始。

该求解过程非常近似于第 1.4 节中使用 DFT 计算求解 Kohn - Sham 方程的迭代方法。正如在第 1.4 节所讨论的,我们忽略了 HF 计算过程中极为重要的许多细节。其中几个细节是:怎样才能确定两套

自旋轨道是否足够近似,即称作二者一致(或相容)?在第(4)步中应当如何对自旋轨道进行修正,从而使所有计算能够有效收敛到一个解?基底应当要多大才合适?怎样才能形成一个良好的自旋轨道初始估值?怎样才能有效地确定展开系数,从而定义单电子方程的解?我们在这一节主要是对量子化学方法进行一个总体概述,而深入到这些细节问题已经大大超出了我们的初衷,但我们希望读者能思考一下这些问题的合理答案,从而能够懂得:对于具有实际物理意义的材料,如何进行 HF 计算。

1.6.4 Hartree – Fock 方法之后 (Beyond Hartree – Fock)

Hartree – Fock 方法给出了电子交换的精确描述,这意味着当两个或更多个电子彼此交换时, HF 计算得到的波函数具有相同的特点,这就如同完全 Schrödinger 方程的真实解一样。如果 HF 计算能够采用无限大的基底,则所计算得到的 N 电子体系能量就是所谓的 Hartree – Fock 极限 (Hartree – Fock limit)。该能量与真实电子波函数的能量并不相同,这是由于 HF 方法并没有正确地描述两部分电子之间是如何相互影响的。更简单地说, HF 方法并没有考虑电子关联效应。

正如我们在前面部分所提示的,可以直接写出电子关联作用所遵循的物理定律,但除了最简单的体系之外,对于一个任意体系而言,想要准确描述其电子关联作用是很棘手的。为了完成量子化学的计算,电子关联作用所引致的能量是以一种特定方式定义的:电子关联能是 Hartree – Fock 极限和真实(非相对论的)基态能之间的差值。量子化学方法以某种方式改进了 Hartree – Fock 方法所采用的其中一条假设,因此量子化学方法要比 HF 方法更为复杂,而且在近似求解 Schrödinger 方程时,也包含了一些电子关联能。

在 HF 方法之上,更先进的量子化学方法是如何进行改善的呢?不同的方法均有所不同,但它们共同的目的都是把电子关联效应包含进来。电子关联通常也被描述为在波函数中“掺混 (Mixing)”了一些从较低能量轨道被激发(或提高)到较高能级上的电子。可以用单个行列式方法对此进行处理,其中使用单个 Slater 行列式作为参考波函数,并

在该波函数的基础上表示激发态。这类方法通常称为“后 Hartree - Fock (Post - Hartree - Fock)”方法,其中包括 CI (Configuration Interaction)、CC (Coupled Cluster)、MP (Møller - Plesset Perturbation Theory) 和 QCI (Quadratic Configuration Interaction) 等方法。这些方法中每一种都具有多个变量,而它们的名字就表明了各自最重要的细节。例如,CCSD 计算是耦合簇 (Coupled Cluster) 计算,含有单个电子 (S) 的激发,以及电子对 (double-D); 而 CCSDT 计算则还包括了三电子的激发 (triples-T)。Møller - Plesset 微扰理论 (Møller - Plesset Perturbation Theory) 是在零阶哈密尔顿量 (通常是 HF 哈密尔顿量) 中加入了微扰 (关联势能)。在 Møller - Plesset 微扰理论方法中,使用一个数值表征微扰理论的阶次,如 MP2 就是二阶理论。

另一类方法使用了不止一个 Slater 行列式作为参考波函数。在这些计算中,描述电子关联效应的方法与上面列出的方法在某些方面是类似的。这些方法包括多构型自洽场 (MultiConfigurational Self - Consistent Field, MCSCF) 方法,多参考单 (或双) 构型互作用 (Multireference Single and Double Configuration Interaction, MRDCI) 方法,以及 N 电子价态微扰理论 (N -Electron Valence state Perturbation Theory, NEVPT) 方法^①。

对于波函数方法 (Wave - Function - Based Methods) 的分类,可以根据它们在以下两方面的区别来划分,即理论层级和基底。不同理论层级给定了用于描述电子之间交互作用的近似形式,在上文叙述各种不同理论层级的段落中,采用一系列缩略语描述了电子与电子之间的相互作用。有趣的是,评价量子化学计算精确性的一个有效规则就是:“缩略词越长,其理论层级越好”^[6]。对波函数方法进行分类的第二个方面,也是同等重要的方面,就是基底。如同在第 1.6.1 节中有关空间局域函数的简单例子,用三个单独函数加和在一起形成一个总体函数。

① 也许现在就是一个恰当时间,可以对读者进行以下提醒:这部分量子化学的综述有点类似于初到异域都市时的常用语手册或旅行指南。有关于此处所列方法的详细信息可在本章末的扩展阅读中去获取。

如果以这种方式对某个特定函数(如 Schrödinger 方程的解)进行近似,我们总可以用更多的函数进行加和,从而给出更精确的近似表达式。每个方程贡献给总方程的量是由数值系数所决定的,而这些数值系数也必须经过计算才能得到。因此,虽然使用含有更多函数的基底能够对真实解给出更精确的表达式,但也需要花费更多的计算工作量。正如可以使用不同的理论层级那样,形成基底也有许多种不同的方式。

为了说明理论层级和基底的作用,让我们来看看 CH_4 分子的两个特性:C-H 键长和离化能(Ionization Energy)。实验上测得 C-H 键长为 $1.094\text{Å}^{[7]}$,离化能为 12.61eV 。首先,表 1.1 中给出了使用相同基底的四种不同理论层级所计算的结果。该表中,其中三种理论层级是波函数方法,即 HF、MP2 和 CCSD。采用隔离态分子最常用的 DFT 泛函(B3LYP 泛函)得到的 DFT 计算结果也列在表中(在本节最后,我们将讨论这种泛函的特点)。该表也显示了每种计算所需时间除以 HF 计算时间的相对时间。从这一列可以观察到:HF 和 DFT 的计算时间是近似相同的——这是一个通用性的结论。理论层级越高,特别是 CCSD 计算,相比于 HF(或 DFT)所花的计算时间显著增加。

表 1.1 中所有不同理论层级预测的 C-H 键长误差均在 1% 之内。表 1.1 另一个令人振奋的信息是:DFT 预测的 C-H 键长远比用时较长的 CCSD 更为精确。HF 方法预测的离化能误差相当大,而其他三种方法所给出的预测结果较好。更高的理论层级(MP2 和 CCSD)所给出的离化能结果要比 DFT 更加精确。

表 1.1 使用 pVTZ 基底和 4 种不同理论层级计算的 CH_4 分子性质^a

理论层级	C-H/Å	误差百分比	离化能/eV	误差百分比	相对时间
HF	1.085	0.8	11.49	8.9	1
DFT(B3LYP)	1.088	0.5	12.46	1.2	1
MP2	1.085	0.8	12.58	0.2	2
CCSD	1.088	0.5	12.54	0.5	18

^a 所有误差均为与实验值相比较

现在让我们看看相同理论层级,不同基底的计算所预测的 CH_4 性质。表 1.2 列出了使用 B3LYP 泛函的一些 DFT 计算结果,并使用了一系列具有复杂名称的不同基底。不用深入研究具体细节,我们就能大致从名称上了解到 STO-3G 是一个常用的“最小”基底,而 cc-pVDZ、cc-pVTZ 和 cc-pVQZ (D 表示 double, T 表示 triple, 依次类推) 是专门为分子的高效数值计算所开发的一系列常用基底。表中列出了在每个计算中所使用的基底函数的数目,以及与最快计算所对比的相对计算时间。与实验值相比,表 1.2 中所有基底得到的 C-H 键长误差均小于 1%。然而,基底越大,则离化能的结果明显更加精确。

表 1.2 使用 DFT (B3LYP) 和 4 种不同基底计算的 CH_4 分子性质^a

基底	基底函数数量	C-H/Å	误差百分比	离化能/eV	误差百分比	相对时间
STO-3G	27	1.097	0.3	12.08	4.2	1
cc-pVDZ	61	1.100	0.6	12.34	2.2	1
cc-pVTZ	121	1.088	0.5	12.46	1.2	2
cc-pVQZ	240	1.088	0.5	12.46	1.2	13

^a 所有误差均为与实验值相比较,相对时间是与 STO-3G 计算相比

可以从表 1.2 中观察到另一个有意思的现象:两种较大基底 pVTZ 和 pVQZ 的计算结果是完全相同的(至少在该表数据所取的数值精度上是如此)。发生这一现象的原因是:当基底包含有足够的函数,从而能精确描述 Schrödinger 方程的解时,就会出现所谓的“收敛于相应的基底处”。发生这一现象本是一件好事! 但另一个事实是:当基底对某个理论层级(如 DFT)已经足够大的时候,可能对其他更高的理论层级仍没达到足够大。因此,表 1.2 中的数据并不意味着,将 pVTZ 基底用于表 1.1 的 CCSD 计算时能够收敛于该基底。

为了使波函数方法 (Wave-Function-Based Methods) 收敛于 Schrödinger 方程的真实解,必须同时使用一个较高的理论层级和较大的基底。然而,这个办法只有在进行少量原子的计算时才是可行的,因为随着理论层级和基函数数目的增加,计算所需的计算成本也快速增加。例如,对于具有 N 个函数的基底而言,一个常规 HF 计算通常需要

约 N^4 次运算；而一个常规的 coupled - cluster 计算需要约 N^7 次运算。目前,对于改善 HF 和后 HF 计算的上述比例,已经取得了一些进展。但即便是有所改善,对于像 CH_4 这样的一个小分子,从表 1.2 中所采用的包含有几百个基函数的合理基底,仍能感受到这一比例关系问题。根据较高理论层级波函数方法的计算成本,可以看出:对于求解包含 10 ~ 20 个原子的单个有机分子,这些方法是可行的,但对于更大的物理体系而言,就会出现“颇具挑战性”或“计算不可行”这样的结果。

对于量子化学的“简介”到此基本结束。正如本书的标题所言,我们在全书的主要关注点是密度泛函理论计算。此外,我们只关注基于空间周期性函数的方法——所谓的平面波方法,对于绝大多数我们感兴趣的物理材料都与晶体材料有关(而不是隔离态分子),此时都需要选择平面波方法。正如前文所述,不能把基于周期性函数的方法当作是“正确的”,而把基于空间局域函数的方法当作是“错误的”(或者反过来)。掌握的方法越全面,则越有利于解决更多的科学问题,因此能够全部理解这两类方法将是非常有好处的。然而,如果研究对象是除了隔离态分子之外的其他材料,则从平面波 DFT 计算起步就是一个很好的起点。

需要强调的是,在 DFT 计算中也可以使用空间局域函数——表 1.1 和表 1.2 中的结果就是这样一些例子。在使用周期性函数和空间局域性函数的 DFT 计算之间,主要区别或许就是它们通常所用的交换关联泛函。第 1.4 节已经从其所包含的内容方面定义了什么是交换关联泛函——如果以一种简单方式分别写出各个部分的贡献,它其实就是全部能量泛函中所剩余的那一部分。然而,在我们对 HF 方法的讨论中,意味着可以用一种(至少在理论上)精确的方式处理“交换能”。在基于空间局域基底函数的 DFT 计算中,最常用的泛函是“混合”泛函,其中将交换能部分的精确泛函与关联能部分的近似表达混合在一起。B3LYP 泛函是迄今为止使用最广泛的混合泛函,其中:B 表示 Becke,他解决了交换能部分;LYP 表示 Lee、Yang 和 Parr,他们开发出了泛函的关联能这一部分;而 3 表示这些结果混合在一起的特定方式。尽管交换能部分的精确表达形式能够有效地用于空间局域函数基

底的计算,但当使用周期性函数基底时则不行。因此在平面波 DFT 计算通常所使用的泛函中,并没有包含交换能的精确表达。

1.7 DFT 不能做哪些事

掌握以下这一事实是非常重要的:实际的 DFT 计算并不是完全 Schrödinger 方程的精确解。之所以不精确,是由于无法知道 Hohenberg - Kohn 定理所应用的精确泛函形式。因此,无论何时进行 DFT 计算,在 DFT 计算所得能量和 Schrödinger 方程真实基态能量之间,本质上都存在着“测不准”。在许多情形下,除了与实验测试进行仔细比较,没有其他方法可以直接估算这个“测不准”的大小。希望读者在读完本书时能够明白:对于许多物理问题,在有效预测复杂材料性质方面,DFT 计算已经具有足够高的精度。在第 1.2 节中已经给出了几个相关的例子,在第 10 章还将讨论 DFT 计算精度这一复杂问题。

在某些重要情形下,DFT 计算并不能给出良好的物理精确度。下面,对这类情形中最常见的问题进行简要叙述。

(1) 第一种情形是电子激发态,DFT 计算只有有限的精确性。回顾第 1.4 节中有关 Hohenberg - Kohn 定理的描述,就可以大致理解这一点。在两个 Hohenberg - Kohn 定理中,都只是针对基态能量。虽然 DFT 计算肯定可以对激发态进行预测,但需要记住的重要一点是:这些预测(从理论上说)并没有建立在基态性质类似预测的相同基础上。

(2) DFT 计算中人所共知的一个不准确性就是在半导体和绝缘体材料中计算得到的带隙是被低估的。在隔离分子中,单个电子所能达到的能级形成了一个离散集合(通常用分子轨道这一术语来描述)。在晶体材料中,这些能级必须用连续函数来描述,也就是所谓的能带(Energy Bands)。金属和绝缘体的最简单定义与以下问题有关:由泡利不相容原理,一旦所有的低能级能带都占满了,材料中电子所能达到的最高能级。如果下一个电子态在最高已占满态之上,仅仅高了无限小的能量,则该材料就是金属;如果下一个电子态在最高已占满态之上,高出有限的能量,则该材料就不是金属,并且这两个电子态的能量

差称为带隙。通常来说,具有“大”带隙(若干电子伏特带隙)的材料被称为绝缘体,而具有“小”带隙的材料被称为半导体。使用现有泛函的标准 DFT 计算,所得带隙的精度都有限,与实验数据比较,其误差普遍大于 1eV 。关于这一问题,令人难以理解的一点是:即便是前文中精确的 Kohn - Sham 交换关联泛函,也已经证实存在着同样的问题^①。

(3) 另一个 DFT 计算给出不准确结果的情形,与原子和分子之间较弱的范德华(van der Waals, vdW)吸引力有关。为了明白这种吸引力确实存在,只需要考虑一下类似 CH_4 (甲烷)这样的简单分子。甲烷在温度足够低且压强足够高时变为液体,以液体形式来运输甲烷要比气体形式经济得多,这也是全世界液化天然气(Liquefied Natural Gas, LNG)工业的基本立足点。但是,当变为液体时,在 CH_4 分子对之间一定存在某种相互吸引作用,这个相互吸引作用就是范德华作用。在最基本的层面上,它的出现是由于两个分子之间存在关联作用,即:一个分子电荷密度出现短暂波动,而另一个分子对这些波动作出回应,从而其电子能量也出现变化。该解释已经暗示出为什么在 DFT 中描述这些相互作用是很有挑战的,这是因为范德华作用是长程电子关联效应的直接结果。为了从量子力学上精确地计算这些相互作用的强度,必须使用更高级的波函数方法,从而可以系统地处理电子关联作用。目前已经做到了这一点,例如,已经计算了 H_2 分子对之间非常弱的相互作用,对于处在实验上已知的最优几何构型的两个 H_2 分子,其能量要比分开很远距离的同样两个分子低约 0.003eV ^[8]。

(4) DFT 还有另一个更基本的缺陷,而且这对于理解(或者说它也源自于)求解 DFT 数学问题所需的计算成本,是非常重要的。目前的计算通常是对数十个原子进行的,还无法对几百个原子进行计算,但已经被作为一个颇具挑战的研究性问题来考虑。对数千个或更多原子

^① 如何在 DFT 方法中精确地处理这一问题,仍是一个活跃的研究领域,而且正在形成许多有意义的进展。这方面工作的两个代表性例子是 Rinke P, Qteish A, Neugebauer J, et al, Exciting Prospects for Solids: Exact Exchange Based Functional Meet Quasiparticle Energy Calculations. Phys. Stat. Sol. 245, 2008, 929, 和 Uddin J, Peralta J E, Scuseria G E. Density Functional Theory Study of Bulk Platinum Monoxide. Phys. Rev. B, 71, 2005: 155112。

有关的计算也是有可能的,但仅仅限于开发了最先进代码,并且能够使用世界上最大计算机的那一小部分人。从物理角度来看,半径为 $1\mu\text{m}$ 的一滴水含有 10^{11} 数量级的原子。直接对这个数量级的原子进行 DFT 计算,需要在计算技术或代码效率上取得极大进展,而这是无法想象的。因此,每个使用 DFT 计算的人必须清楚地懂得:如何将极少量原子的计算结果与具有物理相关性的真实材料联系起来。

1.8 其他领域中的密度泛函理论

为了叙述的完整性,需要指出,密度泛函理论这一名词并非单指本章中所描述的这些类型的量子力学计算。使用密度泛函解决问题这一思想也已经被用于经典力学中的流体动力学,其中的密度是流体密度,并不是电荷密度,而且基本方程也不是 Schrödinger 方程。认识到这两个截然不同的学术领域对其各自方法使用了相同的名字,会避免产生混淆。否则,你有可能发现自己出现在其他领域的研讨会中。

1.9 如何使用本书(再谈)

本章的开头作了一个有关于学习开车的类比,用来说明本书的目的。既然我们已经介绍了大量有关 DFT 和量子化学计算的学术术语,在这里就能够更准确地说明本书的主题和方法,后续章节主要内容是:将常见泛函的平面波 DFT 计算,用于体相材料、表面、纳米颗粒和分子等相关物理问题。目前,能够执行这些平面波计算的软件代码已经较为普及,我们主要介绍一些具体问题,重点在于如何将这些方法以高效计算的方式用于感兴趣的科学问题中去。

本书读者主要有两类。

第一类读者想要进入的研究领域中,需要每天进行 DFT 计算(或许也包括其他的计算化学或材料模拟)。如果你就是这类读者,应该尽可能多地完成每章末尾的练习,这是非常重要的。完成这些练习只需要适中的计算资源,却包含了每一章所介绍的重点内容。简单说来,

如果你希望进入的研究领域中需要进行计算,那么你就必须实际运行你自己的计算,而不仅仅是查阅其他人的工作。只有尽最大努力地亲自做过之后,才能更好地掌握其中的许多细节。对于这一类读者,建议按章节顺序阅读本书。

第二类读者是那些不需要每天亲自执行计算的人,但他们的研究领域中,DFT 计算已经成为一种“标准”方法。对于这一类人,重要的是明白用来描述 DFT 计算的语言以及 DFT 的优点和不足。这与“标准”实验技术没什么不同,例如 X 射线衍射或扫描电子显微镜。对于庞大数量的研究者而言,这些基本方法的应用知识是不可或缺的,无论他们本人是否亲自使用这些方法。如果你是这类读者中的一员,希望这本书有助于你成为一名 DFT 结果的高级“消费者”。如果你时间有限(如在一次长途飞行中),建议你阅读第 3 章、第 10 章以及第 4 至 9 章中与你最为相关的部分。如果你的航班晚点了^①,阅读一章与你的研究内容并不直接相关的内容——很可能你就会从中找到感兴趣的东西。

作者已经尽可能地缩短本书篇幅,因为我们深信:阅读和理解这样篇幅简短的全书时,显然要比面对(或携带)类似于大城市电话号码本那样的篇幅,更加地吸引人。当然,这也就不可避免地限制了本书主题内容的覆盖范围,特别是对于使用局域化基底(Localized Basis Sets)的 DFT 计算,除了在本章简略介绍之外,并没有说明它的具体细节。此外,也没有对 DFT 理论和建立泛函进行深入说明。正因如此,就谨慎地在本书的标题中出现了“简介”这一单词。读者可以将本书看作是进入 DFT、计算化学和材料模拟这一充满活力“世界”的起点。对每章末尾扩展阅读部分所列文献进行追加阅读,可以使读者不断扩展自己的视野,远超过本书所介绍的内容。

作者把 DFT 计算精度这一重要问题推迟到了第 10 章。在前面的章节中,介绍了 DFT 对一系列不同物理性质的应用。第 10 章的讨论着重说明了无法用非常简单的方式描述这个主题。对于把 DFT 方法

^① 此处原文为:“If (when?) your flight is delayed, . . .”——译者注。

应用于很有挑战性的物理问题,第 10 章也指出了目前正在进行的几个快速发展领域。

参 考 文 献

- [1] Honkala K, Hellman A, Remediakis I N, et al. Ammonia Synthesis from First - Principles Calculations. *Science* 307, 2005 :555.
- [2] Schweinfest R, Paxton A T, Finnis M W. Bismuth Embrittlement of Copper is an Atomic Size Effect. *Nature* 432, 2004 :1008.
- [3] Umemoto K, Wentzcovitch R M, Allen P B. Dissociation of MgSiO_3 in the Cores of Gas Giants and Terrestrial Exoplanets. *Science* 311, 2006 :983.
- [4] Umemoto K, Wentzcovitch R M, Weidner D J, et al. NaMgF_3 : A Low - Pressure Analog of MgSiO_3 , *Geophys. Res. Lett.* 33, 2006 :L15304.
- [5] Kohn W. Nobel Lecture: Electronic Structure of Matter - Wave Functions and Density Functionals. *Rev. Mod. Phys.* 71, 1999 :1253.
- [6] Cramer C J, Roberts J T. Computational Analysis: Y_2K . *Science* 286, 1999 :2281.
- [7] Herzberg G. Electronic Spectra and Electronic Structure of Polyatomic Molecules. Van Nostrand, New York, 1966.
- [8] Diep P, Johnson J K. An Accurate $\text{H}_2 - \text{H}_2$ Interaction Potential from First Principles. *J. Chem. Phys.* 112, 2000 :4465.

扩 展 阅 读

本书每章末都列有扩展阅读的文献资源,利用这些文献列表,有助于了解本章中提及或简略叙述的相关主题,作者并没有详尽地列出所有文献(究其原因,可以从现存与“量子力学”或类似内容有关的海量教科书这一点得到答案)。

在已完成的关于量子力学的许多著作中,下列书目可以作为读者对本章中所遇到基本概念进行回顾学习的一个良好起点:

Atkins P W, Friedman R S. *Molecular Quantum Mechanics*. Oxford University Press, Oxford, UK, 1997.

McQuarrie D A. *Quantum Chemistry*. University Science Books, Mill

Valley, CA, 1983.

Ratner M A, Schatz G C. Introduction to Quantum Mechanics in Chemistry. Prentice Hall, Upper Saddle River, NJ, 2001.

Simons J, Nichols J. Quantum Mechanics in Chemistry. Oxford University Press, New York, 1997.

以下文献介绍了 DFT 具体细节:

Koch W, Holthausen M C. A Chemist's Guide to Density Functional Theory. Wiley - VCH, Weinheim, 2000.

Martin R M. Electronic Structure: Basic Theory and Practical Methods. Cambridge University Press, Cambridge, UK, 2004.

Parr R G, Yang W. Density - Functional Theory of Atoms and Molecules. Oxford University Press, Oxford, UK, 1989.

了解量子化学计算方法在 DFT 之外,所应用的广泛范围,可以参考以下文献:

Foresman J B, Frisch A. Exploring Chemistry with Electronic Structure Methods. Gaussian Inc., Pittsburgh, 1996.

Szabo A, Ostlund N S. Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory. Dover, Minneola, NY, 1996.

Young D. Computational Chemistry: A Practical Guide for Applying Techniques to Real World Problems. Wiley, New York, 2001.

对能带理论进行简要综述的一本书是:

Sutton A P. Electronic Structure of Materials. Oxford University Press, Oxford, UK, 1993.

对这方面进行更深入介绍的两个传统的参考文献是:

Ashcroft N W, Mermin N D. Solid State Physics. Saunders College Publishing, Orlando, 1976.

Kittel C. Introduction to Solid State Physics. Wiley, New York, 1976.

如果想要了解流体动力学中的 DFT, 一个不错的参考文献是:

Ted Davis H. Statistical Mechanics of Phases, Interfaces, and Thin Films. Wiley VCH, 1995.

第 2 章 简单固体的 DFT 计算

本章主要介绍如何使用 DFT 计算预测固体的最基本物理特性——晶体结构。使用 DFT 计算是本章的重点内容,为了更清楚地理解这一点,将略过实际计算过程中的一些技术细节,例如收敛性测试。这并非意味着这些细节无关紧要,恰恰相反,这些细节对计算的可靠性是至关重要的。为了在本章中完成这一目的,假设我们现在已经能够使用 DFT 程序计算某些原子集合的总能了。如果读者希望演练本章内的例子,则请参考本章末对这类计算工作所给出的一些具体细节。

2.1 周期性结构、超晶胞和晶格参数

读者可能对晶体结构的概念已非常熟悉了,这一概念是学习使用平面波 DFT 计算的核心,我们将从最简单的晶体结构开始讲起。首先需要做的是定义纯金属晶格中所有原子的位置,想象一下:在一个边长为 a 的立方体中,在每个顶角上放置一个金属原子,如果采用笛卡儿坐标系,就可以用 $\mathbf{r} = (n_1 a, n_2 a, n_3 a)$ 来定义所有原子的位置,其中 n_1, n_2, n_3 为三个任意整数。这种晶格称为简单立方结构,实际上有一种元素的原子正是按照这种结构排列的,这就是钋。在具体定义一个简单立方晶体结构的金属时,需要哪些信息呢? 其实只有一个,就是晶格常数 a 。

我们现在来定义一个原子集合,用它来表示一个简单立方材料,并将其用于 DFT 计算。更准确地说,我们需要定义一套原子,并且在每个方向都重复这套原子,则建立了一个完整的三维晶体结构。对于下面的第一个范例,我们分成两部分来讲解,这样处理是很有好处的;但

在实际上并不需要将这项工作分成两个部分。第一部分工作是,先定义一个体积,使其在所有方向都重复并占满整个空间。对于简单立方金属而言,显而易见的是该体积为一个边长为 a 的立方体,其中一个顶角位于 $(0,0,0)$,而该顶角的三条棱边分别指向三维空间中的 x 、 y 、 z 坐标方向。第二部分工作是,定义该体积中所包含原子的位置。对于我们现在所考察的简单立方晶格,在立方体中只包含一个原子,可以将其放置在 $(0,0,0)$ 。将这两部分工作综合起来,就定义了一个简单立方元素的晶体结构。通常,把定义该晶胞体积的矢量,以及晶胞中的原子位置统称为超晶胞,而定义超晶胞是 DFT 计算中最基本的输入和初始化工作。

其实,以上定义简单立方超晶胞的方案并不是唯一的。例如,还可以使用一个边长为 $2a$ 并且包含有 4 个原子的立方体,4 个原子分别放置在 $(0,0,0)$ 、 $(0,0,a)$ 、 $(0,a,0)$ 和 $(a,0,0)$ 。在空间内重复这个超晶胞所定义的简单立方结构,与上述使用较小体积超晶胞时是完全相同的。但有一点区别是很明显,那就是在上面的第一种方案中,所包含的原子数最少。这种很具吸引力的超晶胞称为“原胞 (Primitive Cell)”。

根据原胞的定义,回顾上述简单立方结构的建模过程,其实并非一定要把原子放置在超晶胞的 $(0,0,0)$ 位置,也可以放置在 $(a/2, a/2, a/2)$,或者是 $(0,0, a/2)$,甚至是 $(0.647a, 0.2293a, 0.184a)$ 。所有这些位置都只在超晶胞中放置了一个原子,并利用超晶胞在空间中的重复性构建了简单立方结构。从美观的角度来看,将原子放置在 $(0,0,0)$ 似乎是最好的选择,而 $(a/2, a/2, a/2)$ 也是一个不错的选择,因为此时原子位于超晶胞的中心。需要记住的一点是:上述任何一个放置位置在数学上都是等效的,因此它们在用于实际计算时也具有完全相同的等效性。

现在大家已经懂得如何定义一个用于 DFT 计算,且具有简单立方晶体结构的超晶胞了。在本章开始时,我们就假设大家已经会使用 DFT 程序计算一些原子集合的总能。那么,如何运用这些总能计算去确定自然界中那些简单立方金属的晶格常数呢?较为可取的方法是考

察总能与晶格常数的函数关系,即 $E_{\text{tot}}(a)$ 。图 2.1 所示为这类计算的一个典型结果。关于这些计算(以及本章内其他计算)的细节列于本章末的附录中。

图 2.1 曲线的形状非常简单,在最低点处所对应的 a 值记为 a_0 。如果存在有 a 值大于或小于 a_0 的情形,则晶格参数将不断变化为 a_0 ,而其总能也会随之不断降低。由于自然界中物质能量总是倾向于最小化,所以,根据这些计算可以直接给出一个物理预测,即可由 DFT 计算预测该简单立方材料的晶格常数为 a_0 。

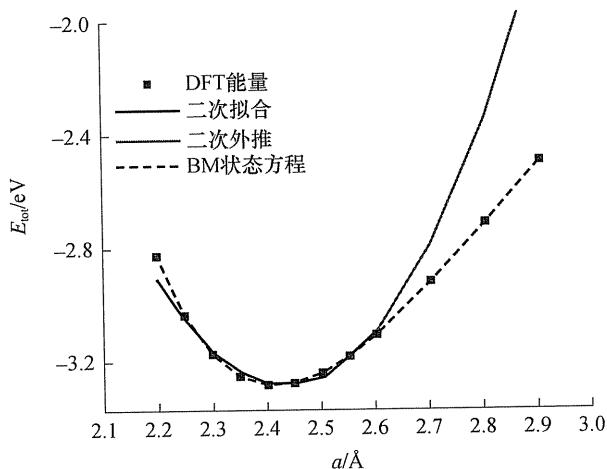


图 2.1 简单立方晶体结构下 Cu 的总能 $E_{\text{tot}}(a)$ 与其晶格常数 a 的函数关系,实心方形符号表示的是 DFT 计算结果,而三条线为使用图例中所标明拟合方法的拟合线。

为了能从总能计算中得到 a_0 数值,考察一下 $E_{\text{tot}}(a)$ 的函数形式是非常有必要的,最简单的表达形式是把总能写成截断的泰勒展开式为

$$E_{\text{tot}}(a) \approx E_{\text{tot}}(a_0) + \alpha(a - a_0) + \beta(a - a_0)^2 \quad (2.1)$$

式中: $\alpha = dE_{\text{tot}}/da|_{a_0}$, $\beta = d^2E_{\text{tot}}/da^2|_{a_0}$ 。

根据定义,如果 a_0 是对应于最低能量的晶格常数,则 $\alpha = 0$ 。这表

明可以将数据拟合为

$$E_{\text{tot}}(a) \approx E_0 + \beta (a - a_0)^2 \quad (2.2)$$

式中: E_0 、 β 和 a_0 为拟合参数。

图 2.1 中黑色实线是在 a 从 2.25 ~ 2.6 Å 范围内对计算数据拟合后的曲线, 由该拟合曲线可预测 $a_0 = 2.43$ Å。

尽管上述等式中把 β 作为拟合参数, 但它实际上具有直接的物理含义。材料的平衡态体模量 B_0 可定义为 $B_0 = Vd^2 E_{\text{tot}}/dV^2$, 即平衡晶格常数时的二次导数。将这个表达式与上述泰勒展开式比较, 可以发现 $B_0 = \frac{2}{9}(1/a_0)\beta$ 。也就是说, 对曲线进行上述拟合给出了平衡晶格常数和平衡体模量两个数值。图 2.1 中实线所得的体模量为 0.641 eV/Å³。其单位是由 DFT 计算结果直接给出的, 但与宏观数据相比较时较为麻烦。将该数据单位转化为常用单位, 则所得到的 B_0 预测值为 103 GPa。

式(2.2)给出了 $E_{\text{tot}}(a)$ 和 a 的简单关系, 但该等式对于晶格常数在平衡值附近较小范围内变化时并不适用。这一点可由图 2.1 直接看到, 图示的灰色线为 $a > 2.6$ Å 时的二次拟合曲线。这个范围的 DFT 数据并没有用在上述黑色实线的拟合中。图示清楚地显示出: 随着晶格常数的增加, 二次拟合曲线大大偏离了 DFT 数据。其本质原因在于 $E_{\text{tot}}(a)$ 整体上并非是晶格常数的简单二次方程。如果进一步进行更详细的数学处理, 可以让 E_{tot} 和 a 有关的状态方程适用于更广泛的晶格常数范围内。一个著名的例子是各向同性固体的 Birch - Murnaghan 状态方程, 即

$$E_{\text{tot}}(a) = E_0 + \frac{9V_0 B_0}{16} \left\{ \left[\left(\frac{a_0}{a} \right)^2 - 1 \right]^3 B'_0 + \left[\left(\frac{a_0}{a} \right)^2 - 1 \right]^2 \left[6 - 4 \left(\frac{a_0}{a} \right)^2 \right] \right\} \quad (2.3)$$

式中: a_0 为平衡晶格常数; V_0 为单个原子的平均体积; B_0 为在压强 $P = 0$ 下的体模量; $B'_0 = (\partial B / \partial P)_T$ 。

为了将图 2.1 中的数据应用于该方程, 把 a_0 、 B_0 、 B'_0 和 E_0 看做是

拟合参数。使用该方程将全部 DFT 数据点拟合后的曲线是图 2.1 中的虚线。从图示可见,使用该状态方程可以更精确地拟合 DFT 数据。根据该拟合曲线,预测 Cu 在简单立方晶体结构下的 a_0 为 2.41\AA , $B_0 = 102\text{GPa}$ 。需要指出的是,这些预测值与前面用 $E_{\text{tot}}(a)$ 简单二次拟合模型得到的数值非常近似,原因在于二者所得晶格参数 a_0 彼此相近,因此方程式(2.3)可以简化为方程式(2.2)。

2.2 面心立方材料

上面讨论的简单立方晶体结构是最简单的晶体结构,但是,除了铯之外,没有其他元素以该晶体结构存在,因此,这种晶体结构实际上的重要性非常有限。元素周期表中更为常见的晶体结构是面心立方结构 (Face - Centered - Cubic, FCC)。对于该晶体结构,可以看做是边长为 a 立方体的四个顶角和每个面心均有一个原子,而这种立方体充满了整个空间。对于定义 FCC 材料的超晶胞,可以在简单立方晶格的基础上,在 $(0,0,0)$ 、 $(0, a/2, a/2)$ 、 $(a/2, 0, a/2)$ 和 $(a/2, a/2, 0)$ 位置放置原子。读者可以画出结构图,从而验证以上叙述。

与上面简单立方材料超晶胞的定义不同, FCC 金属的超晶胞含有 4 个独立原子。这暗示我们,可能这并不是它的原胞。画出草图,在“自然”结构中,清点边角原子周围的相邻原子个数,然后在面心原子周围重复这一清点,这样更加重了上述怀疑。画图清点表明:两种原子都有 12 个以相同几何构型排列的相邻原子,并且与任意一个原子之间 (以及 12 个邻近原子之间) 的距离全都相同。

在上述 FCC 结构中,将原点处的原子和三个相邻面的面心原子进行连接,就可以定义 FCC 金属的原胞。也就是说,定义晶格矢量为

$$\mathbf{a}_1 = a\left(\frac{1}{2}, \frac{1}{2}, 0\right) \quad \mathbf{a}_2 = a\left(0, \frac{1}{2}, \frac{1}{2}\right) \quad \mathbf{a}_3 = a\left(\frac{1}{2}, 0, \frac{1}{2}\right) \quad (2.4)$$

如果在下列等式的位置 \mathbf{r} 处放置原子,这些矢量就定义了 FCC 晶格为

$$\mathbf{r} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3 \quad (2.5)$$

式中: n_1 、 n_2 和 n_3 为整数。

图 2.2 对此给出了示意图,根据这些定义,在这个晶胞中有一个原子,因此,这一定是原胞。而且,由一个参数 a 就完全定义了晶体的几何构型。FCC 金属最近邻原子的间距是 $a/\sqrt{2}$ 。另外,可以观察到第三个特点,即原胞的晶格矢量不是正交的(即 $\mathbf{a}_i \cdot \mathbf{a}_j$ 均不为零)。

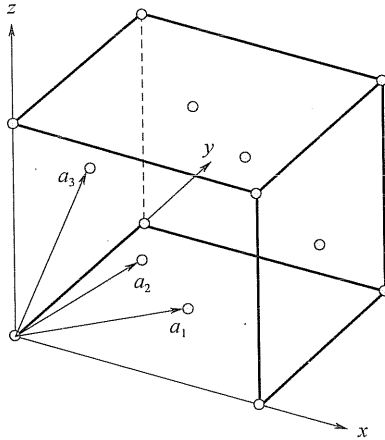


图 2.2 式(2.4)所定义 FCC 金属晶格矢量的示意图

图 2.3 给出了计算得到的 FCC Cu 总能与其晶格常数之间的函数关系。该曲线的形状类似于 Cu 在简单立方晶体结构下的相应曲线(图 2.1),但相比于简单立方结构的最小能量,FCC 结构的最小能量更低。这一现象可以给出一个简单的物理解释:FCC 晶格结构的 Cu 要比简单立方结构更为稳定。这个结论并不奇怪,因为 Cu 实际上就是 FCC 金属,而且,令人感到兴奋的是,DFT 计算结果与物理现实是一致的。还可以把计算得到的晶格常数与实验结果进行比较,从图 2.3(DFT 数据的 Birch - Murnaghan 拟合曲线),可以预测 Cu 的 $a_0 = 3.64\text{\AA}$, $B_0 = 142\text{GPa}$ 。实验上,Cu 的晶格常数是 3.62\AA , $B_0 = 140\text{GPa}$,对于这两个物理量,预测值都与实验值非常接近(尽管并不相等)。

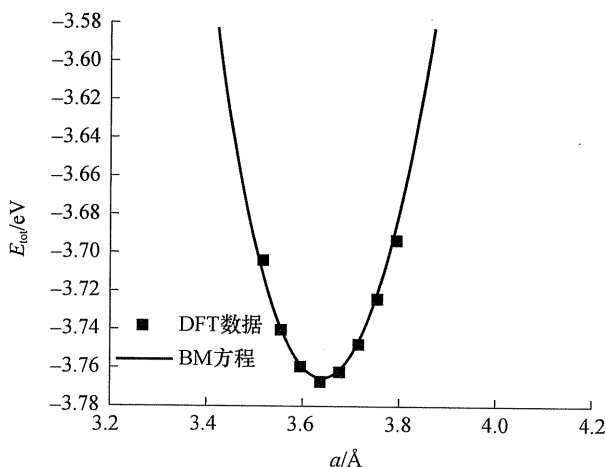


图 2.3 FCC 晶体结构下 Cu 的总能 E 与其晶格常数 a 的函数关系。数据点是 DFT 计算结果, 虚线^①为 Birch - Murnaghan 状态方程。

2.3 密排六方材料

为什么许多金属倾向于 FCC 晶体结构? 其实这并不难理解: 在空间中堆积硬球, 没有哪种结构比 FCC 结构能够产生更高的密度。这一事实的数学证明, 即著名的 Kepler 猜想 (Kepler Conjecture), 在过去几年间刚刚被发现。然而, 存在另一种堆积方式, 具有与 FCC 堆积完全相同的密度, 即密排六方 (Hexagonal Close Packed, HCP) 结构。作为 DFT 应用的第三种周期性结构, 下面对 HCP 金属进行讨论。

相比于已经讨论过的简单立方或面心立方结构, HCP 金属的超晶胞稍微有点复杂。可以使用以下晶格矢量定义其超晶胞为

$$\mathbf{a}_1 = (a, 0, 0) \quad \mathbf{a}_2 = \left(\frac{a}{2}, \frac{\sqrt{3}a}{2}, 0 \right) \quad \mathbf{a}_3 = a(0, 0, c) \quad (2.6)$$

^①原文如此, 但似乎应为实线——译者注。

并将两个原子分别放入超晶胞的笛卡儿坐标 $(0,0,0)$ 和 $(a/2, a/2\sqrt{3}, c/2)$ 位置。可以确信的是,如果只考虑矢量 \mathbf{a}_1 和 \mathbf{a}_2 以及放置在超晶胞 $(0,0,0)$ 位置的原子,会形成一个在 $x-y$ 平面内的六方阵列。还需要留意的是,定义这个超晶胞需要两个未知参数 (a 和 c),而前文中简单立方和 FCC 结构均只需要一个晶格常数。对于理想的硬球堆垛, $c = \sqrt{8}/3a = 1.633a$,而且所有相邻原子对之间的间距都是相同的。然而,在真实的 HCP 金属中,常常观测到晶体存在小的扭曲畸变。例如,对于钪,实验测得 $c/a = 1.59$ 。

利用上述 HCP 超晶胞的定义,可以较容易地说明一个更重要的、常用在周期性几何中定义原子坐标的概念。在上述定义中,矢量确定了超晶胞的形状(晶格矢量),而原子位置由三维的笛卡儿坐标给出。但是,如果用笛卡儿坐标系中定义的晶格矢量来表示原子坐标,这样做通常更为方便。总之,可以将超晶胞中原子 j 的位置写为

$$\mathbf{r}_j = \sum_{i=1}^3 f_{j,i} \mathbf{a}_i \quad (2.7)$$

这个表达式能够在超晶胞中确定每一个原子位置,即对于所有的 i 和 j 都存在 $0 \leq f_{j,i} \leq 1$,而这些系数就称为原子在超晶胞中的分数坐标(Fractional Coordinates)。对于每个原子,分数坐标通常写作向量的形式。例如,在上面定义的 HCP 结构中,两个原子的分数坐标分别是 $(0, 0, 0)$ 和 $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ 。采用这样的定义方式,就只在定义超晶胞晶格矢量时才需要用到晶格参数。采用一组晶格矢量和一组分数坐标定义超晶胞,是目前描述任意超晶胞最方便的方法,在本书后文中均使用这种表示方式。如果不是全部,则至少也是大多数常用 DFT 计算软件,都允许(或要求)其使用者用这种表示方式定义超晶胞。

使用 DFT 计算,可以直接预测 Cu 在简单立方和 FCC 晶体结构下的晶格常数。所做的只是一系列总能与晶格常数 a 函数关系的计算。而 HCP 结构具有两个独立的参数 (a 和 c),这就使这一过程更为复杂。绝大多数 DFT 软件能够自动地处理与此类似的多变量问题,但是,现在假设我们只懂得如何使用计算软件,每次只计算一个晶格体积和几

何构型的总能。在这种情形下,一个解决办法就是固定 c/a 的数值,然后进行一系列以 a 为变量,总能为函数的计算,图 2.4 显示了这些计算的结果。该图中,数据点的连线就很能说明问题。这些计算表明:相比于硬球堆积模型,Cu 在 c 轴上的扭曲畸变很小。更重要的是,相比于 FCC 结构,HCP 结构下 Cu 的最低总能要大近 0.015eV(参见图 2.3),因此计算结果与实际观测是一致的:Cu 是 FCC 金属,而不是 HCP 金属。在这两种晶体结构之间,所预测的能量差值很小,即 1.4kJ/mol。这并不合理,因为两种结构在许多方面都非常近似。

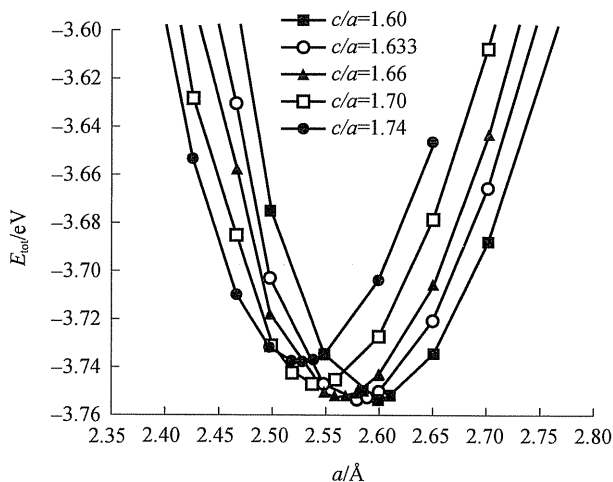


图 2.4 HCP 晶体结构 Cu 的总能 E_{tot} 与 c/a 比值间的函数关系。每个数据点都是 DFT 计算结果,连接数据点的直线给出了直观信息。

对 HCP 结构 Cu 的计算应该(至少)能使读者思考两件事情。第一个是 DFT 计算的数值精确性。能否使用这些计算结果,仅仅依据 1kJ/mol 的能量差值,就能可靠地判定两种晶体结构的稳定性? 是什么导致二者能量相差约 1kJ/mol? 对于这个问题的回答,与数值求解特定原子集合 DFT 这一复杂数学问题的能力是密切相关的。在后续章节中,有很多有关数值收敛性这类重要主题的讨论。第二个方面,我们应当懂得:确定两个晶格参数的计算工作量,比只需确定一个晶格参数

的工作量大得多。给出这个例子的另一部分原因还在于,想要使读者体会到这样一个事实:对于一个存在很多自由度的问题,无论这些自由度是超晶胞内的原子位置,还是与晶格矢量有关的晶格参数,每次只改变一个参数从而对其总能进行最小化这一做法是难以持续的。在讨论了数值收敛之后,下一章讲述对于复杂原子构型,在不需要系统地改变每一个自由度的情形下,如何高效地计算得到能量最小值。

2.4 晶体结构预测

根据前文中的计算结果,似乎可以说,我们已经预测了 Cu 的晶体结构,但是这并不完全正确。更准确地说,应当是:预测得到 FCC 结构的 Cu 比 HCP 结构或简单立方结构的 Cu 更加稳定。仅仅根据我们的计算,无法排除这样一种可能:Cu 事实上可能形成某种我们尚未考察过的其他晶格。对于 Cu 来说,这个想法完全是迂腐而学究的,因为根据实验我们已经知道 Cu 是 FCC 金属。因此,我们才能叙述成:计算结果与实验得到的晶体结构完全一致,而且预测了 FCC 结构 Cu 的晶格常数。

再从另一个角度来说明这一点,试想你被要求预测二钪化钾 (Y_2K) 的晶体结构,并且没有任何相关实验数据。你可以使用 DFT 方法,考察 AB_2 化学配比物质的所有已知晶体结构,将每种结构的 Y_2K 进行总能最小化。这远非一个简单任务,因为已知的 AB_2 化学配比物质具有超过 80 种的不同晶体结构,而且其中某些结构还非常复杂。在这给出一个例子, $NiMg_2$ 、 $NbZn_2$ 、 $ScFe_2$ 、 $ThMg_2$ 、 $HfCr_2$ 和 UPt_2 全都是 C36 六方 Laves 相有序结构,它们的原胞中含有 106 个原子。即便是你完成了所有这些看似“英勇”的计算任务,你也无法肯定 Y_2K 是否会形成另一种从未被发现的新晶体结构。这些讨论表明了:为什么确定一个新化合物的晶体结构仍是一个有意义的科学探索。所传达出的主要信息就是:DFT 非常适合于在一系列可能的结构范围内预测晶体结构的总能,但在缺少实验数据的情形下,仅仅依据这些计算,无法充分地预测新的晶体结构。

2.5 相变

以上对于 Cu 晶体结构的分析全都基于一个合理的思路,即能量最低的晶体结构更有可能在自然界中存在。这个想法是正确的,但在确定一个材料的能量时需要格外小心,从而使其更精确。更准确地说应该是:优先存在的晶体结构具有较低的 Gibbs 自由能,即 $G = G(P, T)$ 。Gibbs 自由能可以写为

$$G(P, T) = E_{\text{coh}} + PV - TS \quad (2.8)$$

式中: E_{coh} 、 V 和 S 分别为材料的内聚能 (Cohesive Energy) (将材料分离成隔离态原子集合所需的能量)、体积和熵。

如果比较两种可能的晶体结构,则所需计算的就是两种结构下 Gibbs 自由能的改变量,即

$$\Delta G(P, T) = \Delta E_{\text{coh}} + P\Delta V - T\Delta S \quad (2.9)$$

在固体中,表达式的前两项要比最后一项来自熵的贡献大得多,因此有

$$\Delta G(P, T) \approx \Delta E_{\text{coh}} + P\Delta V \quad (2.10)$$

在进行 DFT 计算时,表达式右侧的两个量都在能量上有其各自的简单含义。两种结构内聚能的差值就是 DFT 总能的差值。对于能量曲线上的任意一点,压强可以定义为

$$P = \frac{\partial E_{\text{coh}}}{\partial V} \quad (2.11)$$

在 2.1 节 ~2.3 节,对于 DFT 能量与晶格参数之间函数关系曲线图的最小值,可以将其预测为所考察晶体结构的最优晶格参数。根据式(2.8)和(2.11)可以看出,这个预测的更准确描述应该是:当 $P = 0$ 和 $T = 0$ 时最小值所对应的就是最优晶格参数。

对于式(2.10),有意思的一个结果是:如果 $\Delta E_{\text{coh}} = -P\Delta V$,则两种具有不同内聚能的晶体结构可能具有相同的 Gibbs 自由能。将该条件与式(2.11)相比较可以看出,如果两种结构的 $\Delta E_{\text{coh}} - V$ 函数关系图中具有相同的正切函数值,则这两种结构就满足上述条件。如图 2.5 所示,在 $P = 0$ 时最优晶体结构是结构 1,且由其晶格参数所确定的体积

为 V_0 。沿着曲线从 $V/V_0 = 1$ 的最小值处向着 A 点移动,这就等效于对材料增大压强。当达到 A 点所对应的压强值时,结构 1 与结构 2 具有相同的正切函数值。在较高的压强下,结构 2 的 Gibbs 自由能要比结构 1 的低,因此,根据该图可以预测在这个压强处将会发生一个压力诱发的相变,且相变使内聚能变大,但在自由能中,增大的内聚能被体积的减少给平衡掉了。

许多固体都会发生压力诱发相变,这是人所共知的。例如,体相 Si 在常温常压下具有金刚石结构,但当压强增大到 100kbar 左右时,就会转变为 β -锡结构。图 2.5 显示了如何使用本章中的 DFT 计算信息,来预测是否存在压力诱发相变。在 1.2 节中曾经提及一个启发性例子,对类似 MgSiO_3 这样的矿物质,在地质学上进行了相关预测,本质上就是使用了上述思想。

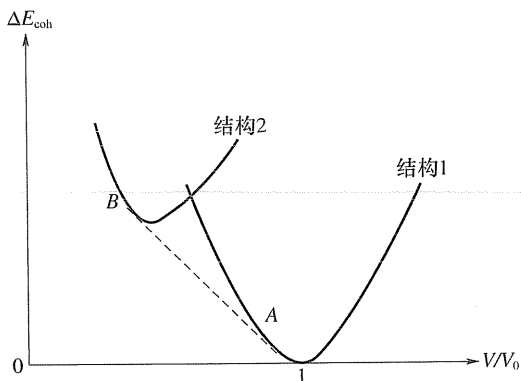


图 2.5 在材料的两种晶体结构之间压强诱发相变的示意图

练 习

为了完成这些练习,需要具备能够执行平面波 DFT 计算的程序代码,并懂得如何使用该程序代码计算一个超晶胞的总能。正如在后续章节将会看到的那样,对于任何一个 DFT 程序代码,都需要你提供许多输入参数。关于这一点,作者建议:如果可能的话,尝试使用默认的

输入参数完成下列练习,或者由一个富有经验的同事来为你设定。可能你希望了解这些输入参数的具体含义,在后续章节中将会满足你这个好奇心。下列这些练习旨在培养熟练建立体相材料超晶胞的能力,并对预测这些材料的晶体结构进行实践。

1. 通过计算确定 Pt 在简单立方、FCC 或 HCP 晶体结构中,更倾向于哪一种? 将 DFT 预测得到的最优晶格参数与实验观测值进行比较。

2. Hf 在实验上观测到是 HCP 金属, $c/a = 1.58$ 。计算并预测 Hf 的晶格参数,并与实验观测值比较。

3. 许多化学配比为 AB 的固体都表现为 CsCl 结构。该结构中, A 原子为简单立方排布, B 原子位于 A 原子立方体的中心。确定 CsCl 结构的晶格矢量和分数坐标,并使用这一结构预测 ScAl 的晶格常数。

4. AB 化合物的另一种常见结构是 NaCl 结构。该结构中,沿简单立方结构的任意一个轴, A 和 B 原子交替形成简单立方晶格。预测 ScAl 在这种结构下的晶格参数,并与上一题的结果进行比较,应该可以得到 ScAl 不倾向于形成 NaCl 结构。

扩展阅读

在几乎所有有关固体物理或材料科学的教材中,都讲述了晶体结构的最基本概念及其表示方法。两部著名教材分别是: Kittel C. *Introduction to Solid State Physics*. Wiley, New York, 1976 和 Ashcroft N W, Mermin N D. *Solid State Physics*. Saunders College, Orlando, FL, 1976.

对于 AB 化合物晶体结构的详细信息,以及对晶体结构各种定义方式的典范性综述,可以参阅: David Pettifor. *Bonding and Structure of Molecules and Solids*. Oxford University Press, Oxford, UK, 1995.

关于 Kepler 猜想的有趣历史,可以参阅: Szpiro G G. *Kepler's Conjecture: How Some of the Greatest Minds in History Helped Solve One of the Oldest Math Problems in the World*. Wiley, Hoboken, NJ, 2003.

有关怎样将量子化学用于预测气相 Y_2K 的性质,请参阅: Cramer C J, Roberts J T. *Science* 286. 1999:2281. 在阅读理解该文献时,请注意其

出版时间较早。

附录 计算细节

对于每章中所给的具体计算结果,将会在类似于此处的附录中给出这些计算的相关细节。

本书中所有计算结果都是使用 Vienna *ab initio* Simulation Package (VASP)^①完成的。尽管作者喜欢使用 VASP 并极其感谢许多人多年来在开发该软件过程中所投入的不懈努力,但这并非意味着对该软件包进行正式推荐。VASP 是几个常用平面波 DFT 计算软件中的一个,其他常用软件也可以用于本书所叙述的所有计算。选择何种软件是读者自己的权利,并受到一些因素影响。例如,所在机构是否具有某个软件的有效许可,以及读者和你们的合作者对其中某种程序是否具有前期使用基础。作者在使用 VASP 时,应尽可能只使用在其他所有平面波 DFT 程序代码中也都有的功能(或特点),从而避免读者在使用任意主流软件重复计算本书范例时的不便。

除非特别说明,本书所有计算都是使用 Perdew - Wang 91 广义梯度近似泛函。简单地说,就是 PW91 - GGA 计算。除非特别说明,在倒易空间中均使用 Monkhorst - Pack 方法确定 k 点。本章每一节中的部分计算细节列于下面:

2.1 节 图 2.1 中立方 Cu 的计算使用了含有 1 个 Cu 原子的立方超晶胞,截断能(Cutoff Energy)为 292eV, k 点为 $12 \times 12 \times 12$ 。

2.2 节 图 2.3 中 FCC 结构 Cu 的计算使用了含有 4 个 Cu 原子的立方超晶胞,截断能(Cutoff Energy)为 292eV, k 点为 $12 \times 12 \times 12$ 。

2.3 节 图 2.4 中 HCP 结构 Cu 的计算使用了含有 2 个 Cu 原子的原胞超晶胞(Primitive Supercell),截断能(Cutoff Energy)为 292eV,倒易空间中使用 Γ 中心网格(Γ - Centered Grid)确定的 k 点为 $12 \times 12 \times 8$ 。

① 有关该计算软件包的信息可以参阅 <http://cms.mpi.univie.ac.at/~vasp>。

第 3 章 DFT 计算的基本要素

在第 2 章中,我们有意忽略了 DFT 计算的许多必要细节,以便更好地阐明这些计算中所探讨的物理量。当初学这些方法时,这种做法是有益的;然而,从长远来看,这显然不是一个好办法。在本章中,我们将深入到一些具体细节,从而明确以下两种“计算”之间的区别:一种是能够给出可靠物理信息的 DFT 计算,另一种是那些仅仅只是耗费时间、占用计算机的所谓计算。

收敛 (Convergence) 是我们已经重复多次的一个重要概念。当执行 DFT 计算时(或者当与其他从事类似计算的人交流时),可能总被问起计算是否已经收敛了^①。那么收敛是什么意思呢?要想回答这一问题,需要简要回顾一下第 1 章中有关 DFT 的描述。DFT 所确定的原子构型基态电荷密度,就是由一组复杂数学方程的解所给出的。为了在计算机上实际求解这一问题,必须进行一系列数值近似:多维空间中的积分必须近似为该函数在有限点集合上的积分;本应表示为无穷多个加和项的解,也必须被截断为有限个加和等。对于这类数值近似中,都能够在使用更多计算资源的前提下,得到越来越接近真实值的求解,这一过程就称为收敛。对于采用了特定交换关联泛函的 DFT 所描述的数学问题,“收敛良好的”计算能够给出与真实解尽可能接近的数值解。

数值收敛这一概念,与 DFT 计算是否精确描述了物理现实这一问题,二者是截然不同的。DFT 所定义数学问题与完整 Schrödinger 方程并不完全相同(因为无法给出更好的关联泛函的精确表达形式)。这就意味着,DFT 问题的精确解并不是 Schrödinger 方程的精确解,而

^① 对任何一个理论计算研究者,这都是一个万全的问题:“请问你计算结果的收敛精度是多高?”而对于实验研究者而言,相同的问题是:“在你的实验中有可能存在外界污染?”

后者其实才是我们最感兴趣的。这一问题就是 DFT 的物理精度,它是极重要的,但相比于数值收敛来说,完全说明这个问题是非常复杂的。关于 DFT 计算的物理精度这一问题,将在第 10 章讲述。

在本章中,首先说明完成良好收敛 DFT 计算需要些什么。总之,因为无法非常确定地得到 DFT 数值问题的精确解,所以通常需要合理地比较 DFT 结果与物理现实之间的一致性(或其中的不足)。

3.1 倒易空间和 k 点

我们通常习惯于在三维物理空间中定义原子位置,但是,为了对数值收敛性进行说明,首先就会把我们带入到所谓的倒易空间。与倒易空间有关的概念是固体物理的基本知识,这些概念被某些人看得颇为神秘,而许多物理学家却并没有留意到这一点。在这里,我们并不打算对这些概念给出详细说明。本章末的扩展阅读中,列出了许多正式的固体物理教材,都对这些内容给出了极其详细的介绍。对于如何把实际 DFT 计算与倒易空间联系起来,在本章中将就我们认为最重要的概念进行介绍,尤其是关注这些概念与数值收敛之间的关系。

3.1.1 平面波和布里渊区

在第 2 章中强调了 we 感兴趣的是把 DFT 计算应用于在空间中周期性排列的原子集合。采用晶格矢量 \mathbf{a}_1 、 \mathbf{a}_2 和 \mathbf{a}_3 ,我们定义了空间中周期性重复的晶胞,即超晶胞。如果在这个周期性系统中求解 Schrödinger 方程,它的解一定满足所谓的 Bloch 定理(Bloch's Theorem)的基本特点,该原理认为:这个解可以表达为若干项的连加和,即

$$\phi_{\mathbf{k}}(\mathbf{r}) = \exp(i\mathbf{k} \cdot \mathbf{r})u_{\mathbf{k}}(\mathbf{r}) \quad (3.1)$$

式中: $u_{\mathbf{k}}(\mathbf{r})$ 在空间中具有与超晶胞相同的周期性,也就是对于任意整数 n_1 、 n_2 和 n_3 有 $u_{\mathbf{k}}(\mathbf{r} + n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3) = u_{\mathbf{k}}(\mathbf{r})$ 。该定理意味着可以用每一个 \mathbf{k} 值分别独立地求解 Schrödinger 方程。我们刚才是从 Schrödinger 方程的解这一角度进行说明的,但该定理也适用于由 Schrödinger 方程这些解所衍生出的其他物理量,例如电荷密度。

从许多方面来说,使用 \mathbf{k} 求解 DFT 数学问题要比使用 \mathbf{r} 更为方便。因为函数 $\exp(\mathbf{ik} \cdot \mathbf{r})$ 称为波函数,所以采用这一计算思想的方法常称为波函数计算。矢量 \mathbf{r} 的空间称为实空间,而矢量 \mathbf{k} 的空间称为倒易空间(或简称为 k 空间)。在本书后文所讨论的计算中,都是围绕着倒易空间展开的,因此,有必要介绍一下 k 空间的几个特点。

在实空间中可以用晶格矢量 \mathbf{a}_1 、 \mathbf{a}_2 和 \mathbf{a}_3 的表达式来确定位置,同样地,在虚空间中也可以定义三个矢量用来确定位置,也就是所谓的倒格矢(Reciprocal Lattice Vectors) \mathbf{b}_1 、 \mathbf{b}_2 和 \mathbf{b}_3 ,并且,如果 $i = j$ 则 $\mathbf{a}_i \cdot \mathbf{b}_j$ 等于 2π ,否则等于 0。这就意味着

$$\mathbf{b}_1 = 2\pi \frac{\mathbf{a}_2 \times \mathbf{a}_3}{\mathbf{a}_1 \cdot (\mathbf{a}_2 \times \mathbf{a}_3)} \quad \mathbf{b}_2 = 2\pi \frac{\mathbf{a}_3 \times \mathbf{a}_1}{\mathbf{a}_2 \cdot (\mathbf{a}_3 \times \mathbf{a}_1)} \quad \mathbf{b}_3 = 2\pi \frac{\mathbf{a}_1 \times \mathbf{a}_2}{\mathbf{a}_3 \cdot (\mathbf{a}_1 \times \mathbf{a}_2)} \quad (3.2)$$

有关这种计算的一个简单例子就是第 2 章中所讨论的简单立方晶格。对于实空间晶格矢量,一个很自然的选择就是对于所有 i ,都有 $|\mathbf{a}_i| = a$ 。可以验证得到,对于所有 i ,倒格矢满足 $|\mathbf{b}_i| = 2\pi/a$ 。对于该体系而言,晶格矢量和倒格矢都定义了立方体,而前者的边长为 a ,后者的边长为 $2\pi/a$ 。

在第 2 章中,提到了可以用晶格矢量 $\mathbf{a}_i = a$,或者是用晶格矢量 $\mathbf{a}_i = 2a$ 来定义简单立方超晶胞。选择第一个则超晶胞中只有一个原子,是简单立方材料的原胞;选择第二个则超晶胞中有 8 个原子。两种选择定义的是同一种材料。如果使用第二种选择,则倒格矢将会变为 $\mathbf{b}_i = a/\pi$ 。通过这个例子可以看出一个非常重要的一般性结论:实空间中的晶格矢量越大,则倒易空间中所对应的倒格矢越小。

倒格矢所定义的三维形状,并非总与实空间中超晶胞的形状完全相同。对于 FCC 原胞而言,在第 2 章中给出为

$$\mathbf{a}_1 = a \left(\frac{1}{2}, \frac{1}{2}, 0 \right) \quad \mathbf{a}_2 = a \left(0, \frac{1}{2}, \frac{1}{2} \right) \quad \mathbf{a}_3 = a \left(\frac{1}{2}, 0, \frac{1}{2} \right) \quad (3.3)$$

此时,倒格矢为

$$\mathbf{b}_1 = \frac{2\pi}{a} (1, 1, -1) \quad \mathbf{b}_2 = \frac{2\pi}{a} (-1, 1, 1) \quad \mathbf{b}_3 = \frac{2\pi}{a} (1, -1, 1) \quad (3.4)$$

可以再次看到:倒格矢的长度与实空间中晶格矢量的长度具有反比关系。图 3.1 中画出了这些实空间晶格矢量和倒格矢。

前文中介绍了原胞的概念,就是在完全定义一个无限扩展的周期性材料时,含有必要的最少原子个数的超晶胞。更一般地看,也可以把原胞看作是含有所需信息的最小体积单元。用所谓的 Wigner - Seitz 晶胞可以更准确地说明这一概念。此处,本书并不深入到具体细节,仅仅进行一个简单说明:无论是对于实空间矢量,还是对于倒易晶格矢量,都可以非常容易地定义 Wigner - Seitz 晶胞。用更简单的语言来说,我们可以在倒易空间定义一个原胞。由于倒易空间的原胞具有许多特点,就专门对其进行命名,即布里渊区(Brillouin Zone, BZ)。布里渊区在材料能带理论中具有核心意义,对布里渊区中几个特别重要的点,也单独进行了命名。其中,最重要的点就是 $k=0$ 处,在 k 空间中这个位置称为 Γ 点。为了完全说明实空间和倒易空间中长度的联系,将 BZ 的体积记作 V_{BZ} ,而根据 Wigner - Seitz 方法在实空间中建立的原胞体积记作 V_{cell} ,则二者间的联系为

$$V_{BZ} = \frac{(2\pi)^3}{V_{cell}} \tag{3.5}$$

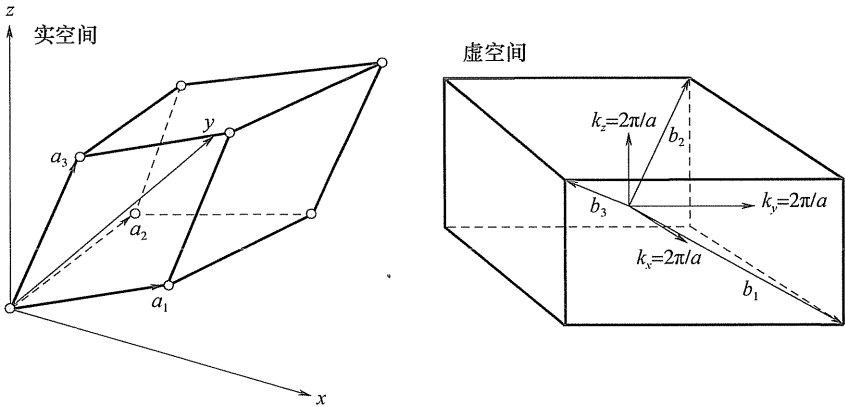


图 3.1 FCC 原胞实空间和倒易空间中晶格矢量示意图。实空间的图中,圆圈表示原子。倒易空间的图中,以立方体的中心处为原点画出了边长为 $4\pi/a$ 的基矢。

3.1.2 k 空间中的积分

对于平面波 DFT 计算而言,布里渊区为什么这么重要? 答案很简单:在实际的 DFT 计算中,所需要进行的大量工作可以归纳为求算下述形式的积分值,即

$$\bar{g} = \frac{V_{\text{cell}}}{(2\pi)^3} \int_{\text{BZ}} g(\mathbf{k}) d\mathbf{k} \quad (3.6)$$

该积分的主要特点是:它是在倒易空间中所定义的,并对布里渊区中所有可能的 \mathbf{k} 值进行了积分。我们暂时不去说明这个积分从哪儿得来的,现在要考虑的是如何求得它的数值。

在我们开始求算式(3.6)的积分值之前,先来看一个数值求算 $\int_{-1}^1 f(x) dx$ 的简单任务。可能读者记得:该积分可以看作是在 $[-1, 1]$ 区间内 $f(x)$ 曲线下所围的面积。这表明了求解该积分近似值的一个简单方法,首先将区间断开为相等的若干段,然后估算在每一段上,该曲线下方的面积;估算方法是将若干区间段端点之间的曲线视作直线,即将每个区间段内的曲线裁弯取直,将整个曲线看作由一段段直线所组成。这样就可以利用梯形图解法(Trapezoidal Method)进行估算,即

$$\int_{-1}^1 f(x) dx \approx \frac{h}{2} \left[f(-1) + f(+1) + 2 \sum_{j=1}^{n-1} f(x_j) \right] \quad (3.7)$$

式中: $x_j = -1 + jh$, 而 $h = 2/n$ 。

作为一个简单测试,可以对 $\int_{-1}^1 (\pi x/2) \sin(\pi x) dx$ 进行求算。该积分可以被分割为若干部分来精确求解,而选择这个函数是因为其积分值恰好为 1。使用梯形图解法对该测试积分求算,结果列于表 3.1 中。结果在我们的意料之中:在 x 取值空间中所采用的步长越小,则求算得到该函数的数值也就越精确。表 3.1 表明:只有当 n 值大于 5 时,积分的求算结果才较好,此时精度为 1%。

表 3.1 使用梯形图解法和 Legendre 求积法对积分

$$\int_{-1}^1 (\pi x/2) \sin(\pi x) dx = 1 \text{ 的近似求解}$$

N	梯形图解法	Legendre 求积法
2	0.6046	1.7605
3	0.7854	0.8793
4	0.8648	1.0080
5	0.9070	0.9997

所使用的梯形图解法具有两个特征:在 $f(x)$ 估算位置之间,间距是均匀相等的;每一个 $f(x)$ 的估算值都使用了相同的权重(除了两端的端点)。其实这两个特点都不是必须的(或者说不是我们所期望的)。还有一类精致的积分方法称为高斯求积(Gaussian Quadrature)法,其形式为

$$\int_{-1}^1 f(x) dx \approx \sum_{j=1}^n c_j f(x_j) \quad (3.8)$$

式中:积分点 x_j 对应于正交多项式的根,而权重 c_j 与这些多项式的积分有关。对于在 $[-1, 1]$ 区间上的积分,这种方法称为 Legendre 求积(Legendre Quadrature)。下面给出一个具体例子,当 $n=3$ 时,式(3.8)中的权重和积分点分别是 $x_1 = -x_3 = 0.775967$, $x_2 = 0$, $c_1 = -c_3 = 0.555555$ 和 $c_2 = 0.88888$ 。将这种方法用于上述测试积分,结果列于表 3.1 中。与梯形图解法形成鲜明对比,这种方法的结果随着 n 增大快速收敛于正确结果:一旦 $n > 3$,则该积分的误差就小于 1%。

上述对一维函数进行数值积分的例子可以总结为以下三个要点,而且可以应用于多维积分:

(1) 通过将积函数在一系列离散点上进行估值,并在每一点上使用合适权重,再对这些函数值加和,可以近似求算该积分。

(2) 随着加和所使用离散点个数的增加,这种数值方法可以给出更精确的结果。当所使用的点非常多,并对其取极限时,这些数值方法

就收敛于积分的精确解。

(3) 在近似求解函数的积分时,对位置和权重的不同选择,会极大地影响该数值方法收敛到精确结果的速度。

3.1.3 在布里渊区中选择 k 点

对于类似式(3.6)的积分,因为所消耗的 DFT 计算工作量非常大,所以对如何快速估算这些积分进行认真细致地研究,就并不奇怪了。最为广泛使用的办法是由 Monkhorst 和 Pack 在 1976 年提出的,绝大多数 DFT 计算软件都提供了基于该方法选定 k 点的选项。使用该方法,所需做的仅仅是确定在倒易空间每个方向上都使用多少个 k 点。对于每个晶格矢量长度都相等的超晶胞的计算,其倒格矢的长度也相等,因此在每个方向上使用相同数量的 k 点就是很自然的了。如果在每个方向都使用了 M 个 k 点,则将该运算标记为使用了 $M \times M \times M$ 个 k 点。

从以上有关数值积分的一般性讨论中,可以明确的是:如果 $M > N$,则相比于使用了 $N \times N \times N$ 个 k 点的计算,使用 $M \times M \times M$ 个 k 点的计算应该能够给出更精确的结果。但是,实际上应该怎样选择,以及选择多少个 k 点呢? 回答这个问题要比读者所想象的更为复杂和困难,如表 3.2 所列。这个表中的数据是用与 2.2 节中计算 FCC 结构 Cu 相同的计算方法得到的,其晶格常数是根据 2.2 节中的计算确定的,即晶格常数为 3.64\AA 。表 3.2 中的每个计算都是使用 Monkhorst - Pack 方法所确定的 $M \times M \times M$ 个 k 点。正如在 2.2 节中所介绍的,这些计算使用了含有 4 个原子的 FCC 立方超晶胞,而不是其原胞。使用该表中的部分数据作图,如图 3.2 所示。

首先看一下表 3.2 所列(图 3.2)能量值。当 $M > 8$ 时,可见总能(几乎)与 k 点数无关,正如我们所期望的那样,即 k 空间内的所有计算都数值收敛良好。更确切地说,在这个区间上,随 M 增大而导致的能量变化小于 0.003eV 。然而,对于较少的 k 点,能量随 k 点的变化仍然很大。这清楚地表明:较少的 k 点无法给出收敛良好的结果。

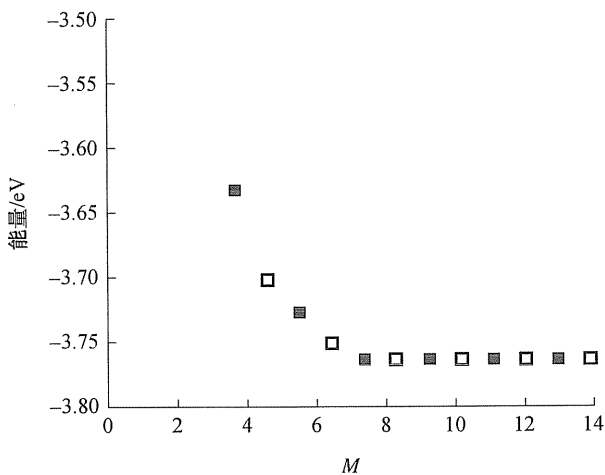


图 3.2 当使用 $M \times M \times M$ 个 k 点时计算得到的体相 Cu 总能(平均到每个原子)与 M 之间的函数关系。偶数和奇数 M 的计算结果分别用空心和实心数据点表示。

表 3.2 采用 Monkhorst - Pack 方法使用 $M \times M \times M$ 个 k 点计算 FCC 结构 Cu 的总能结果

M	E/eV	在 IBZ 中的 k 点数量	τ_M/τ_1
1	1.8061	1	1.0
2	3.0997	1	1.1
3	3.6352	4	2.3
4	3.7054	4	2.6
5	3.7301	10	5.2
6	3.7541	10	6.0
7	3.7676	20	10.4
8	3.7671	20	11.2
9	3.7680	35	16.9
10	3.7676	35	17.1
11	3.7662	56	31.2
12	3.7665	56	28.5
13	3.7661	84	40.0
14	3.7659	84	39.7

表 3.2 中最后一列是计算总能的时间除以 $M = 1$ 时计算时间的比值。请注意,达到我们刚才所确定的收敛结果所需时间,要比仅用一个 k 点的计算时间延长了至少 20 倍。一个令人奇怪的特征是:如果 M 是奇数,则对于 M 和 $M + 1$,二者的计算时间接近于相等。其发生原因是:这些计算利用了 FCC 理想晶体中所存在的大量对称性。这些对称性意味着:倒易空间的积分并不需要使用全部布里渊区来估算,而是可以使用一部分区域来估算这个积分。而根据对称性,不需要进行任何近似,就可以用这部分区域扩展充满整个布里渊区。这个 k 空间中缩小的部分区域就是不可约布里渊区 (Irreducible Brillouin Zone, IBZ)。对于理想 FCC 晶体等对称性很高的材料,使用 IBZ 可以极大地简化在 k 空间中积分所需要的计算工作量。例如,在本例中,如果采用 $10 \times 10 \times 10$ 的 Monkhorst - Pack 布里渊区取样,在 k 空间中只有 35 个独立的点位于 IBZ 内(如果在计算时没有使用任何对称性,则将会有 1000 个点)。

表 3.2 列出了每个计算在 IBZ 中的 k 点数量。将其与表中所列计算时间相比较,可以说明为什么奇数和偶数 M 的计算时间基本相同——因为它们在 IBZ 中具有相同的独立 k 点数。其发生的原因是:在 Monkhorst - Pack 方法中,使用奇数 M 包括了在 IBZ 边界上的一些 k 点(如 Γ 点),而偶数 M 只给出了 IBZ 内部的 k 点。这意味着:当使用少量 k 点时,使用偶数 M 要比使用奇数 M 的收敛性稍好一些,而二者的计算工作量相同^①。当然,最好能保证你的计算相对于 k 点已经收敛得很好了,在做到这一点的前提下,偶数和奇数 k 点之间的区别就不那么重要了。

为了表明对称性对减少 DFT 计算工作量具有多大的帮助,我们使用四原子超晶胞,并将每个原子从它原本的 FCC 晶格位置移动一个微小距离,然后重复了表 3.2 中的部分计算。这些位移并不大,仅仅将最近邻原子间距改变了 $\pm 0.09 \text{ \AA}$,但是改变了该体系的所有对称性,此时 IBZ 中的 k 点数量是 $M^3/2$ 。这些计算结果列于表 3.3,表中也列出了 ΔE ,即有对称性和没有对称性计算之间的能量差。

尽管表 3.3 中各个计算所用时间与 IBZ 中 k 点数量存在紧密联

① 某些情形下,例如考察电子结构时,将 Γ 点的 k 点包括在内是很重要的。

系,但在 k 空间中计算的收敛性仍与全部布里渊区的 k 点密度有关。如果我们把表 3.3 和图 3.2 的相关数值进行比较,可以看到:当 $M = 8$ 时非对称性体系的计算,似乎是该表中唯一一个已经适度收敛的数值。对于非对称体系而言,如果需要达到更高的能量收敛性,则需要采用更大数量的 k 点继续进行计算。

表 3.3 破坏对称性后 FCC Cu 不同的总能计算结果^a

M	E/eV	$\Delta E/\text{eV}$	在 IBZ 中的 k 点数量	τ_M/τ_1
1	1.8148	0.009	1	1.0
2	3.0900	0.010	4	2.1
3	3.6272	0.008	14	5.6
4	3.6969	0.009	32	12.3
5	3.7210	0.009	63	21.9
6	3.7446	0.010	108	40.1
7	3.7577	0.010	172	57.5
8	3.7569	0.010	256	86.8

^a采用 Monkhorst - Pack 方法使用 $M \times M \times M$ 个 k 点进行计算

可能读者已经注意到,在表 3.3 中,随着 k 点数量增大,能量差 ΔE 要比总能 E 收敛得更快。这非常有好处,相比于两种构型各自的绝对总能值,我们更感兴趣二者间的能量差,它具有更多的物理意义。但为什么会这样呢?对于某个原子构型的积分值,在任何一个 k 点集合上所得到的数值估算值都与该积分的真实值间存在一定系统误差。如果比较两种结构近似的原子构型,则可以做出以下预期:二者的这种系统性数值误差也是非常近似的。这意味着,当求算这两种状态之间的能量差时,就抵消掉了它们的系统误差,计算得到的能量差就要比相应的总能收敛得更加精确。需要强调并加以理解的是,这个试探性的论断建立在如下基础上:两种原子构型需要“足够近似”,才能让二者所使用的 k 点也非常近似。例如,在表 3.3 中,我们特意选择了两种原子构型,二者的区别仅仅是超晶胞中的原子位置存在很小的微扰,因此完全可以应用上述论断。然而,当比较晶体结构完全不同的两种材料时,上述论断就不再合理。

很多时候,采用晶格矢量长度不同的超晶胞是非常有好处的。一个明显的例子就是采用具有以下晶格矢量的超晶胞,对体相 Cu 进行计算,即

$$\mathbf{a}_1 = a(1,0,0) \quad \mathbf{a}_2 = a(0,1,0) \quad \mathbf{a}_3 = a(0,0,4) \quad (3.9)$$

该超晶胞中含有 16 个原子,而在上述立方超晶胞中,只有 4 个原子。在对这个扩大的超晶胞进行计算时,我们应当使用多少个 k 点,才能与表 3.2 立方超晶胞结果的收敛精度相同呢? 一个最基本的规则就是:在倒易空间中采用相似的 k 点密度,则其计算结果的收敛精度也类似。根据该规则,采用 $8 \times 8 \times 2$ 个 k 点就可以得到合理的收敛精度,其中的三个数字分别指沿倒易晶格矢量 \mathbf{b}_1 、 \mathbf{b}_2 和 \mathbf{b}_3 方向的 k 点数量。根据倒格矢的定义,可以验证得到:采用上述 k 点设置方法,定义了倒易空间每个方向都具有相同的 k 点密度。

3.1.4 金属—— k 空间的特例

在前面介绍数值积分这一概念时,是以 $\int_{-1}^1 (\pi x/2) \sin(\pi x) dx$ 这一积分为例进行说明的。在这之前,一直没有提到这个例子的一个特点:它是一个连续函数的积分。这是一个很有用的数学特征,在对这类函数开发数值计算方法时,可以快速收敛到积分的精确解,但是对于 k 空间积分的 DFT 计算而言,并非总是存在这种数学特征。这类情形一个特别重要的例子就是金属,金属的其中一个定义就是:在金属中,布里渊区可以分为电子已占满的和未占满的两个区域。 k 空间中将这两个区域分割开来的表面称为费米面(Fermi Surface)。这一情形的 k 空间积分计算是非常复杂的,因为被积函数从非零值不连续地变化为零(在费米面上)。在计算这些积分时,如果没有采用其他措施,就需要使用大量 k 点才能给出收敛精度较高的结果。

在所有材料中,金属是相当重要的一类。因此,为了改善刚才提到的收敛较慢这一问题,开发出了许多有效的计算方法,这里对其中两个最为著名的方法进行说明。第一个方法称为四面体(Tetrahedron)方法,其思想是:使用一个离散的 k 点集合来定义一组完全充满倒易空间

的四面体,并使用内插法,在四面体的每一点上确定被积函数值。最简单的情形是在每个四面体内部使用线性内插法,一旦完成了数值内插,则在 k 空间所有位置上,该被积函数都具有简单形式,从而就可以使用整个空间来估算该积分,而并非仅仅使用那些初始的离散点。Blöchl 开发出了该方法的一个版本,并使用超越线性内插法的其他内插方法。这是目前最为广泛使用的四面体方法(参见扩展阅读)。

对于金属的不连续积分,另一个较为不同的方法是模糊化方法(Smearing Methods)。该方法的思想是:通过“模糊化(Smearing)”处理其不连续性,强迫这些函数进行连续积分。一个模糊化函数(Smearing Function)的例子是 Fermi - Dirac 函数,即

$$f\left(\frac{k - k_0}{\sigma}\right) = \left[\exp\left(\frac{k - k_0}{\sigma}\right) + 1 \right]^{-1} \quad (3.10)$$

图 3.3 给出了该函数对于几个不同 σ 的曲线形状。可以看到,当 $\sigma \rightarrow 0$ 时,该函数趋近于一个台阶函数(Step Function),在 $k = k_0$ 时从 1 不连续变化到 0。模糊化方法的思想就是,将任何台阶函数替换为光滑连续函数,例如 Fermi - Dirac 函数,从而就可以使用通常的方法对其积分。理想情况下,应当使用某些方法将计算结果进行外推并得到最终极限值,从而使模糊化得以消除(对于 Fermi - Dirac 函数,就是 $\sigma \rightarrow 0$)。

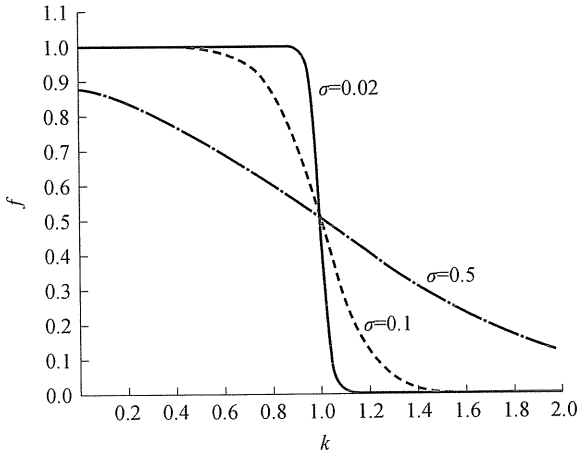


图 3.3 $k_0 = 1$ 及不同 σ 数值下 Fermi - Dirac 函数[式(3.10)]的曲线

Methfessel 和 Paxon 提出了一个广泛使用的模糊化 (Smearing) 方法, 相比于上述例子中的简单方法, 他们使用了更为复杂的 Smearing 函数表达式, 但仍然用单一参数 σ 表征 (参见扩展阅读)。

3.1.5 k 空间的总结

如果想要完成一个有意义的计算, 那么如何在 DFT 计算中处理倒易空间, 对这个问题做出明智选择是非常重要的。因此, 读者应当仔细阅读以下的总结部分, 并在开始做具体计算之前再次回顾一下。关于如何得到良好收敛结果的重要思想包括:

(1) 在对感兴趣的体系进行大量 DFT 计算之前, 应当先考察计算结果对于 k 点数量的收敛性。

(2) 应该给出计算中所使用的 k 点数量, 因为如果不这样做, 就很难对这些结果进行再现。

(3) 增大超晶胞的体积减少了达到收敛时所需要的 k 点数量, 因为实空间体积的增加对应于倒易空间体积的减少。

(4) 如果计算中涉及不同体积的超晶胞, 并需要对其结果进行比较, 则在倒易空间中选定 k 点时, 需要使不同超晶胞倒易空间中的 k 点密度大体相同, 这是使这些计算在 k 空间中具有类似收敛精度的有效方法。

(5) 理解了为什么对称性能够减少实际计算所需的 k 点数量, 则有助于理解单个计算 (Individual Calculations) 需要多长时间。但是, 整体的收敛性是由全布里渊区 (Full Brillouin Zone) 中的 k 点密度所决定的, 而不仅仅是 IBZ 中的 k 点数量。

(6) 对于金属, 必须使用合适的近似方法, 才能更精确地处理 k 空间问题。

3.2 截断能 (Energy Cutoffs)

以上有关 k 空间的长篇讨论是从 Bloch 定理开始的, 该定理告诉我们: 对于超晶胞 Schrödinger 方程的解, 具有如下形式, 即

$$\phi_k(\mathbf{r}) = \exp(i\mathbf{k} \cdot \mathbf{r}) u_k(\mathbf{r}) \quad (3.11)$$

式中: $u_k(\mathbf{r})$ 在空间中是周期的, 且具有与超晶胞相同的周期性。现在来更仔细地看看这部分问题, $u_k(\mathbf{r})$ 具有周期性, 这意味着它可以用一系列特殊的平面波函数展开为

$$u_k(\mathbf{r}) = \sum_{\mathbf{G}} c_{\mathbf{G}} \exp[i\mathbf{G} \cdot \mathbf{r}] \quad (3.12)$$

式中: 加和项是对由 $\mathbf{G} = m_1 \mathbf{b}_1 + m_2 \mathbf{b}_2 + m_3 \mathbf{b}_3$ (其中 m_i 为整数) 所定义的所有矢量加和。在倒易空间中, 由这些 \mathbf{G} 矢量所定义的矢量集合也是确定的。对于任意实空间晶格矢量 \mathbf{a}_i , 则有 $\mathbf{G} \cdot \mathbf{a}_i = 2\pi m_i$ (可以根据倒易晶格矢量的定义, 对该说法进行验证)。

将上述两个方程合并, 则有

$$\phi_k(\mathbf{r}) = \sum_{\mathbf{G}} c_{k+\mathbf{G}} \exp[i(\mathbf{k} + \mathbf{G})\mathbf{r}] \quad (3.13)$$

根据这个表达式, 即便是在 k 空间的一个点上估算求解, 也需要对无限多 \mathbf{G} 的可能取值进行加和。对于实际计算而言这并不是很好。然而, 形如式(3.13)的函数, 可以非常简单地表示 Schrödinger 方程的解, 这些解中所包含的动能是

$$E = \frac{\hbar^2}{2m} |\mathbf{k} + \mathbf{G}|^2$$

可以合理地预期: 相比具有更高能量的解, 具有较低能量的解在物理上更重要。因此, 通常将上述无穷加和项截断为仅包括动能低于某个数值的部分, 即

$$E_{\text{cut}} = \frac{\hbar^2}{2m} G_{\text{cut}}^2$$

则无限加和项简写为

$$\phi_k(\mathbf{r}) = \sum_{|\mathbf{G}+\mathbf{k}| < G_{\text{cut}}} c_{k+\mathbf{G}} \exp[i(\mathbf{k} + \mathbf{G})\mathbf{r}] \quad (3.14)$$

对于不同的 \bar{k} 值, 该表达式所包含的项数也有所不同。

上面的讨论已经表明, 无论何时进行 DFT 计算, 都需要再多定义一个参数——截断能 E_{cut} 。相比于 k 点而言, 可以用许多方法, 更加容易地确定该参数。如果使用者没有特别给出, 绝大多数软件都将会使用一个合理的默认值。如同 k 点一样, 说明计算中所用的截断能是一

个良好的做法,以便其他人能够较容易地重算你的结果。图 3.4 显示了 FCC Cu 的能量随计算所用截断能变化的收敛关系曲线。

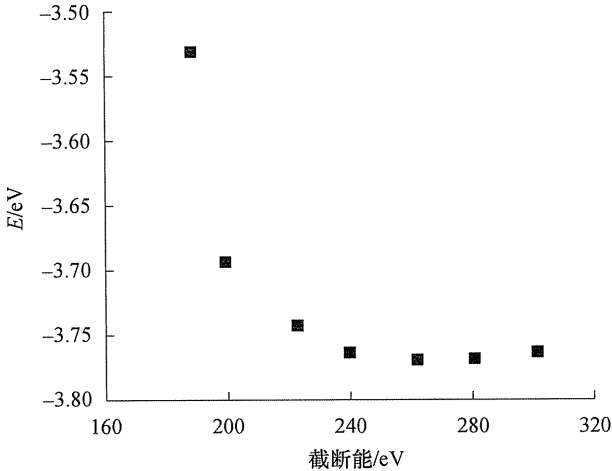


图 3.4 在 3.64\AA 晶格常数下使用 $12 \times 12 \times 12$ 个 k 点计算得到 FCC Cu 单个原子能量与截断能的函数关系图,按照与图 3.2 类似的能量坐标轴比例绘制。

如果不加思考盲目地接受截断能的默认值,这是不明智的,但这也是个常见情形。在多数情况下,会对超晶胞中的每一种元素指定一个默认截断能,且超晶胞中所有元素的最大截断能被用作整个体系的截断能。如果使用这种方法,计算有序体心立方 (Body - Centered Cubic, BCC) 铜钯合金与体相钯、铜之间的能量变化,即 $\text{Pd}_{(s)} + \text{FCC Cu}_{(s)} \rightarrow \text{BCC CuPd}_{(s)}$ 。可以分别求出每种固体收敛良好的能量值,进而计算出这个能量变化(本章末的练习题 7 就是该问题)。为了进一步说明上述问题,先假设 Pd 和 Cu 的截断能满足 $E_{\text{Pd}} > E_{\text{Cu}}$,如果使用默认的截断能,则在计算 FCC Pd 和 BCC CuPd 的能量时使用了 E_{Pd} 截断能,但在计算 FCC Cu 的能量时使用了 E_{Cu} 截断能。这样,就在计算得到的能量差中引入了一个系统性误差,而为了消除这个误差,可以使用更高的截断能 E_{Pd} 来计算 FCC Cu 的能量。在这里,需要记住的关键一点是:无论

何时,只要是对多体系 DFT 计算得到的能量值进行比较,都需要在所有的计算中使用完全相同的截断能。

3.2.1 赝势

上述讨论指出了—个事实:必须使用较大的截断能,才能把实空间中振荡尺度较小的平面波考虑进来。这是一个需要解决的问题,因为,对于原子内部结合紧密的芯区电子,恰恰是与这类振荡特征的波函数相联系的。然而,从物理观点来看,对于确定化学键以及材料的其他物理性质,芯区电子并不是特别重要,这些性质主要是由结合较不紧密的价电子所主导的。在平面波方法最初的发展时期,就已经很清楚:在计算中,如果以某种方式对芯区电子的性质进行近似处理,且采用该处理方式能够减少计算中所必须的平面波数量,那么就会产生极大的好处。

为了减小由芯区电子所导致的计算负担,最重要的方法就是使用赝势(Pseudopotentials)。从概念上说,赝势就是把某个芯区电子集合所产生的电荷密度,替换为符合真实离子实某些重要物理和数学特性的圆滑电荷密度。在后续所有计算中都采用这种近似的方法,将芯区电子的特性看作定值,这就是冻结核心近似(Frozen Core Approximation)。不使用该近似的计算称为全电子计算(All - Electron Calculations),它的使用范围较小。赝势是根据理想情况下的隔离原子开发得到的,但所得到的赝势能够可靠地用于处理任何化学环境中的这种原子,并不需要对其赝势进行调整。这一良好特性称为赝势的可移植性(Transferability)。现有的 DFT 计算程序大都提供了赝势库,包含了周期表中的每个(或者至少是绝大多数)元素。

对于每种原子的特定赝势,都详细确定了一个在使用该泛函进行计算时所需使用的最小截断能。需要较高截断能的赝势称为“硬的”,而计算效率更高且具有较低截断能的赝势称为“软的”。使用最普遍的赝势定义方法是基于 Vanderbilt 的研究工作;这些赝势就是超软赝势(Ultra Soft Pseudo Potential, USPP)。正如其名字所显示的那样,这些赝势所需的截断能显著低于其他赝势方法。

使用 USPP 的一个缺点就是:建立每种原子的赝势时,都需要确定

一系列经验参数。目前常用的 DFT 计算软件中,只含有那些经过仔细推演并经过严格测试的超软赝势,但对于某些元素,DFT 计算软件给出了几种具有不同软度的超软赝势。另一种冻结原子核法在某种程度上避免了超软赝势的缺点,这就由 Blöchl 最先提出的投影缀加平面波 (Projector Augmented - Wave, PAW) 方法,后来又由 Kresse 和 Joubert 将其改编用于平面波计算。Kresse 和 Joubert 对小分子和有序扩展固体,分别采用 USPP、PAW 和全电子计算方法进行了大量比较,对许多计算对象,所得到的结果都是完全相同的;而更重要的是,这些结果与全电子计算取得了良好的一致性。在具有较大磁矩的材料中,或者是原子的电负性具有较大差别时,PAW 方法所给出的结果要比 USPP 方法更加可靠。

3.3 数值优化

到目前为止,我们在本章中所有使用 DFT 方法求得某个原子构型的电荷密度和总能有关数值收敛的介绍都有局限性。如果读者回去看看第 2 章的例子,就会记得我们曾经对一种简单材料(体相 Cu),比较了一系列总能计算,从而预测其最稳定的晶体结构和晶格常数。其中的一个计算,即 HCP 晶体结构 Cu 的计算,说明了一个晶体结构可能具有一个以上的自由度,要找到该结构的最低能量构型,就必须非常系统地逐个考察这些自由度的单调变化。

为了能够在实际操作中实现数值收敛的 DFT 计算,我们需要借助于一些方法,从而能有效地处理具有多个自由度的问题。这方面的问题就是数值优化。如同理解倒易空间一样,充分理解优化的核心概念对是否能够有效地使用 DFT 方法也是非常重要的。在本章的后续内容中,我们将主要介绍这一概念。

3.3.1 一维空间中的优化

我们先来看一个在运行 DFT 计算时非常关键的问题,而且在其他领域也存在许多类似的衍生问题。一个简单的起始点就是以下数学问

题:找到 $f(x)$ 的一个极小值,其中 $f(x)$ 是由 x 单值确定的某个函数。假设 $f(x)$ 是一个“光滑”函数,也就是说该函数的导数是存在的并且是连续的。可以认为 $f(x)$ 所表征的是一个非常复杂的情形,且无法用代数方法对其求解。换言之,我们必须使用计算机来求得该问题的数值解。请注意,我们所要求的是“一个”极小值,并非“所有”极小值,也并非是该函数所能取得的最低值。这意味着我们试图要求解一个局域优化问题,而非一个全局优化问题。

但愿读者能够很容易地明白:为什么这个数学问题(或者与此问题紧密关联的其他问题)在DFT计算中是如此重要。对于第2章中我们确定FCC金属的晶格常数这一问题,就可以将其作为一个最小化问题。如果我们并不知道超晶胞中 N 个原子的精确排布构型,那么这个超晶胞的总能就可以写作是 $E(\mathbf{x})$,其中 \mathbf{x} 是定义原子位置的一个 $3N$ 维矢量。找到这些原子在物理上最为优先的排布构型就相当于将这个能量函数最小化^①。

我们现在来考虑一下上面所提到的那个一维问题——找到 $f(x)$ 的一个极小值。如果我们能够找到满足 $f'(x) = df/dx = 0$ 这一条件的 x 值,那么这个点就确定了该函数的一个极大值或者极小值点。因为可以非常容易地区分一个极大值点和一个极小值点(或者是观察二者的二阶导数,或者是用该点附近的函数值来进行试值评估),所以我们能够把上述问题重新定义为:找到一个满足 $f'(x) = 0$ 的 x 值。下面我们两种数值方法来求解这一问题(能够求解该方法的方法有很多种,但这两种方法可以说明所有这些方法中都含有的核心本质概念)。

第一种方法是对分法(或二分法, Bisection Method), 这种方法的思路多少有点像是在漆黑雨夜中沿着一条陌生的小区马路,开车寻找这条街的一个门牌地址。此时,你应该先试着确定这条街上的两点,而你所要去的地址是处于这两点之间,然后在这两点之间仔细寻找。更具体地说,我们先在 $x_1 = a$ 和 $x_2 = b$ 两点之间去寻找,并且这两点上 $f'(x)$ 的正负号应该是相反的,即 $f'(x_1)f'(x_2) < 0$ 。只要 $f'(x)$ 是光滑的,那

① 正如我们将在3.4节所看到的那样,在DFT计算的其他方面,最小化也是非常重要的。

么我们一定能在 a 和 b 之间找到某一点使得 $f'(x) = 0$ 。我们并不知道这一点在哪,因此我们先对二者的中点 $[x^* = (x_1 + x_2)/2]$ 考察一下 $f'(x^*)$ 。此时我们可能会看到两种情况:如果 $f'(x_1)f'(x^*) < 0$,那么我们就知道使 $f'(x) = 0$ 成立的 x 值处于 x_1 和 x^* 之间;如果 $f'(x^*)f'(x_2) < 0$,那么使 $f'(x) = 0$ 成立的 x 值就处于 x^* 和 x_2 之间。如果想要很清晰地理解上述讨论,可以画个图来看一下:当区间 $[a, b]$ 内只存在唯一一个 $f'(x) = 0$ 的地方时,就恰好只会出现这两种情形中的某一种(绝不会两者都发生,也绝不会二者都不发生)。在这两种情形下,我们都通过使用因子 2,缩短了所要搜寻的区间大小。对于所得到的新区间,可以重复上述所有步骤。这个计算每重复一次,就会生成一个新的 x^* 值,如果我们足够长时间地重复这个计算,这些 x^* 值就会越来越接近于我们所要寻找的 $f'(x) = 0$ 那一点。

下面给出一个简单例子,我们用对分法寻找 $f(x) = e^{-x} \cos x$ 的极小值,该函数在 $x = 2.356194490 \dots$ 处取得极小值。如果从 $x_1 = 1.8$ 和 $x_2 = 3$ 两点开始,接下来我们会生成以下一系列近似解: $x^* = 2.4, 2.1, 2.25, 2.325, 2.3625, 2.34375, \dots$ 。随着我们不断地重复计算,这些近似解越来越接近于真实解,尽管我们仅仅应用了 6 次该算法,就已经得到了精确到两位有效数字的解。

上述问题的另一个方法是牛顿法(Newton's Method),该方法的思路如图 3.5 所示。如果我们定义 $g(x) = f'(x)$,那么根据泰勒展开式,则有 $g(x+h) \approx g(x) + hg'(x)$ 。该表达式忽略了 h 的高阶项,因此,只有当 h 值“很小”时,这个表达式才是精确的。当我们在一些 $g(x) \neq 0$ 的位置上考察 $g(x)$ 时,就可以根据上述近似方程,得到该函数为零时的一个良好估测值: $x^* = x + h = x - g(x)/g'(x)$ 。因为该泰勒表达式仅仅只是一个近似, x^* 并不能精确地确定 $g(x^*) = 0$ 的位置。如同我们在对分法中所做的那样,就必须重复这个计算。对于牛顿法而言,我们就是从前一步的估算值出发,继续重复下一个计算。

使用牛顿法从 $x = 1.8$ 开始寻找 $f(x) = e^{-x} \cos x$ 的极小值,可以生成以下一系列近似解: $x^* = 1.8, 2.183348, 2.331986, 2.355627, 2.356194, \dots$ (对于每次迭代都取了七位有效数字)。和对分法一样,

这一系列近似解也不断收敛到该函数极小值所对应的 x 真实值。在对分法和牛顿法之间,二者的显著区别就是:解的收敛速度有多快。可以通过定义一个表达式 $\varepsilon_i = |x_i - x|$ 来表征它们的收敛速度,其中 x_i 是在应用了 i 次数值方法后所得到的 x 真实解的近似值。图 3.6 描绘了这两个算法的结果,通过比较二者哪一个能够更快地获得求解,可以明显地看到:牛顿法极大地优于对分法。

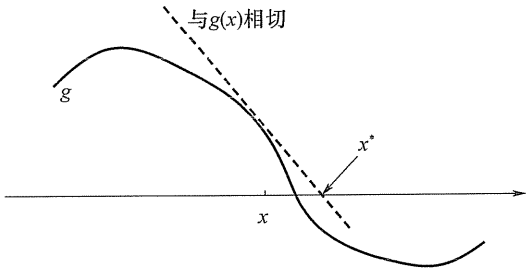


图 3.5 牛顿法对函数 $g(x)$ 的一次迭代示意图

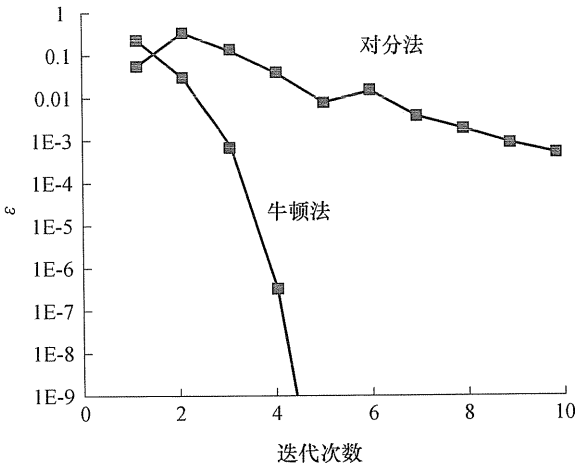


图 3.6 对分法和牛顿法计算的收敛性分析

这些数值方法中,还有一个我们没有讨论到的重要方面。两种方法都能够在长时间运算后生成一系列越来越接近于真实解的近似解。

那么,我们如何来判断何时停止迭代计算呢?在实际应用中,我们并不知道真实解,因此我们无法使用 $\varepsilon_i = |x_i - x|$ 来测试当前估值有多精确。为了找到一个替换方案,我们必须定义一个终止标准,而该标准只决定于我们所计算的一系列近似解。常用的方法就是:在连续两次迭代之间,如果二者的差值小于某个容限值时(即 $|x_{i+1} - x_i| < \delta$),就终止迭代计算。对上述例子,如果我们都取 $\delta = 0.01$,就将在 6 次对分法迭代后终止计算,并给出 $x \approx 2.34375$;或者是在 3 次牛顿法迭代后终止计算,并给出 $x \approx 2.355628$ 。如果选择一个更小的容限值,即 $\delta = 0.001$,将会在 11 次对分法(或 4 次牛顿法)迭代后终止,并得到近似解为 $x \approx 2.3554688$ ($x \approx 2.3561942$)。

上面的例子显示了数值优化的几个普遍性特点,而且充分领会这些特点是极为重要的。这些特点包括:

(1) 这些算法都是迭代的,因此它们无法给出精确解;相反地,它们提供了一系列精确解的近似值。

(2) 使用这些算法时,都必须提供一个真实解的初始估值。对于如何选择这些初始估值,这些算法本身并没有给出任何引导信息。

(3) 所运算的迭代次数是由容限值参数所控制的,这个容限值估测了当前值与真实解的接近程度。

(4) 使用不同的容限值,或者使用不同的初始估值,对任何一种算法进行多次重复的迭代计算,会生成多个(通常)彼此接近的最终近似解,但是它们并不完全相同。

(5) 使用不同算法,则收敛于真实解的速度也是相当不同的。因此,选择一个合适的算法能够极大地降低所需的迭代次数。

对于这些算法,在上述例子中还有其他一些普遍性特点没有显示出来,但这些特点也同等重要,这些特点有:

(1) 如果我们仅仅应用一次计算,那么这些方法无法告诉我们所考察的函数是否存在多个极小值。使用多个不同的初始估值,并应用多次该计算方法,可以生成多个极小值,但即便如此,这些方法也无法给出足够的信息来证明已经找到了所有可能的极小值。我们上面例子中所使用的函数是特别选定的,因为它具有多个极小值。本章末的练

习 1 就要求读者探讨这个概念。

(2) 对于大多数方法,无法保证该方法能从任意一个初始估值收敛于真实解。本章末的练习 2 就给出了一个具体例子来说明这一概念。

3.3.2 大于一维的优化

我们现在回到前面提到的问题:通过不断改变超晶胞中的原子位置,对超晶胞中一组原子的总能 $E(\mathbf{x})$ 进行最小化。这个例子将会凸显出我们在试图求解多变量优化问题时的一些复杂难题,而这些难题在我们已经讨论过的一维情形下并不存在。如果我们定义 $g_i(\mathbf{x}) = \partial E(\mathbf{x}) / \partial x_i$, 那么最小化 $E(\mathbf{x})$ 就相当于找到一组原子位置,使得 $g_i(\mathbf{x}) = 0$ 对于 $i = 1, \dots, 3N$ 都同时成立。

把一维的对分法推广应用于求解这个多维问题并不是很容易。但是,可以把牛顿法推广应用于这个情形。一维的牛顿法是根据泰勒展开式推导得到的,对于多维问题也可以用相同的思路来处理。如果结果中含有 $3N \times 3N$ 的导数矩阵 \mathbf{J} , 其元素为 $J_{ij} = \partial g_i / \partial x_j$ 。请注意,该矩阵的元素是我们所真正感兴趣函数 $E(\mathbf{x})$ 的二阶偏导数。牛顿法通过下式定义了一系列迭代,即

$$\mathbf{x}_{i+1} = \mathbf{x}_i - \mathbf{J}^{-1}(\mathbf{x}_i) \mathbf{g}(\mathbf{x}_i) \quad (3.15)$$

这看起来很类似于我们对一维情形所推导的结果,但这个表达式在实际中被更多地加以应用。考虑一种情形:我们有 10 个不同原子,则在我们所考虑的问题里就有 30 个不同变量。在每次迭代中,我们都必须计算矩阵 \mathbf{J} 的 55 个不同元素;其他 45 个矩阵元素也就随之而定,因为该矩阵是对称的。然后,我们必须求解含有一个 30×30 矩阵的线性代数问题。可以与一维问题比较一下,可见在这个计算中每一步所含有的计算工作是相当大的。可以相当肯定的是,通过计算机执行所有这些多维问题的计算工作,才是更加可行的。上述讨论的主要意义就是:随着问题变量数目的增大,处理多维优化所需要的数值计算工作量也极大增加。

一维和多维问题的另一个区别是与寻找多个解有关的。如果我们

想要找到一维函数 $f(x)$ 在区间 $[a, b]$ 内的所有极小值,理论上我们可以使用该区间内的许多个初始估值,重复应用牛顿方法,从而可以建立一个不同极小值的列表。虽然我们无法严格证明这个方法所建立的列表是完全完备的;但是,对于一个简单函数而言,通过进行多次计算,从而确信无法再通过增加计算次数找到新的极小值,做到这一点也并不很困难。但这一思路无法应用于 30 维的问题,因为必须要考虑极其大量可能的初始估值,才能对可能的解空间进行完全取样。因此,对于 DFT 计算,如果要给出除了确定某个函数是否存在一个极小值之外的其他信息,在多维空间对该问题进行优化总是极其困难的。虽然存在着能够严格求得某个函数全局最小值的算法,但是这些算法都需要进行特别大量的计算工作。就目前所知,还没有人试图在 DFT 计算中合并使用这类算法。

到这里,似乎是可以对以上有关优化方法的讨论进行总结了,因为我们已经定义了能够快速收敛到多维问题最优解的一种方法(牛顿法)。但是,牛顿法不能应用于本节开始所设定的 DFT 问题中去!要想用牛顿法对超晶胞中一组原子的总能 $E(\mathbf{x})$ 进行最小化,就需要计算形如 $\partial^2 E / \partial x_i \partial x_j$ 的二阶导数矩阵。然而,在平面波 DFT 方法中很难直接估算能量的二阶导数,并且大多数计算软件也并不想要执行这类计算。其中的问题不仅是牛顿法在数值计算上的效率不高——更直接的原因是:在实际计算中,估算这些在该方法中所需要用到的函数值是不通的。因此,我们必须寻找其他方法来最小化 $E(\mathbf{x})$ 。关于这一问题,下面我们简单讨论两种最常用的数值方法:拟牛顿法(Quasi-Newton Method)和共轭梯度法(Conjugate-Gradient Method)。在这两者中,我们只是简单说明一些主要概念。有关这两种方法的详细讨论可以参考列于本章末扩展阅读中的相关文献。

拟牛顿法的实质就是把式(3.15)替换为

$$\mathbf{x}_{i+1} = \mathbf{x}_i - \mathbf{A}_i^{-1}(\mathbf{x}_i) \mathbf{g}(\mathbf{x}_i) \quad (3.16)$$

式中: \mathbf{A}_i 为一个矩阵,其近似于雅可比矩阵(Jacobian Matrix)。在计算过程中该矩阵也会由每次迭代不断更新,并且其更新形式为

$$\mathbf{A}_i = \mathbf{A}_{i-1} + F[\mathbf{x}_i, \mathbf{x}_{i-1}, \mathbf{G}(\mathbf{x}_i), \mathbf{g}(\mathbf{x}_{i-1})] \quad (3.17)$$

我们把这种方法称为“Quasi - Newton Methods”，而不是“Quasi - Newton Method”，这是由于有许多种不同函数都可以用作该表达式中的函数 F 。关于函数 F 的细节信息并非我们所讨论的重点，我们应该注意的是：这个更新步骤使用了该方法当前和前一次迭代的相关信息。这一点与我们前面所介绍的所有方法都不同，上面的方法中只用当前迭代的信息来生成下一次新的迭代。如果你对此稍加考虑，就会意识到该等式只是告诉我们：当已经完成该方法的若干次迭代后，如何向前进行下一次迭代。关于如何实现这一复杂步骤已经超出了我们在这里的讨论范围，但我们仍然可以看出：在使用拟牛顿法时，只有在进行了 4 或 5 次最小值迭代后，才能考察该方法对于真实解的收敛程度。

用共轭梯度方法对 $E(\mathbf{x})$ 进行最小化是基于一个完全不同于拟牛顿法的思路。我们使用下面的一个简单例子来说明该思路：计算二维函数 $E(\mathbf{x}) = 3x^2 + y^2$ 的极小值，其中 $\mathbf{x} = (x, y)$ 。读者也许能够立即看出该问题的解：该函数在 $\mathbf{x} = (0, 0)$ 位置上恰好有一个极小值。我们试着从 $\mathbf{x}_0 = (1, 1)$ 开始通过迭代来进行求解。因为我们想要最小化该函数，所以应该考察一下沿哪个方向会使该函数减小。向量微积分的基本结果就是：函数 $E(\mathbf{x})$ 降低最快的方向就是沿着该函数梯度 $\nabla E(\mathbf{x})$ 负值所定义的方向。这表明我们应该定义并生成一个新的估算值，即

$$\mathbf{x}_1 = \mathbf{x}_0 - \alpha_0 \nabla E(\mathbf{x}_0) = (1 - 6\alpha_0, 1 - 2\alpha_0) \quad (3.18)$$

这个表达式告诉我们在空间内沿着特定的一条直线去寻找，但并没有告诉我们在这一条线上下一次迭代应该放在哪里。这类方法就是所谓的最速下降法 (Steepest Descent Methods)。理想情况下，我们将选择步长 α_0 ，因为此时 $E(\mathbf{x}_1)$ 要比我们用任何其他步长所得到的结果都要小。对于上述例子，如果我们选择 $\alpha_0 = \frac{5}{28}$ ，只需进行少量代数运算就显示出 $E[\mathbf{x}_1(\alpha_0)]$ 已经最小化了。

现在我们可以重复这一过程，进行第二次迭代。这一次，我们想要沿着由 $E(\mathbf{x})$ 梯度在 $\mathbf{x}_1: \nabla E(\mathbf{x}_1) = \left(-\frac{3}{7}, \frac{9}{7}\right)$ 位置上所定义的方向进行搜索。之所以这么详细地定义这一方向，是为了突出下面这一结

果,即

$$\nabla E(\mathbf{x}_0) \cdot \nabla E(\mathbf{x}_1) = 0 \quad (3.19)$$

在上述二维空间中,这两个搜寻方向是彼此垂直的。用更为通用的数学语言来说,这两个搜寻方向是正交的。这并不是一个仅仅对于我们所定义的特定问题才会出现的巧合;这是最速下降法在式(3.18)所定义的直线搜寻问题能够得到最优解时的一个普遍特点。

现在我们就能够定性描述共轭梯度法是如何对一个通用函数 $E(\mathbf{x})$ 进行最小化的了,其中 \mathbf{x} 是一个 N 维向量。我们从初始估值 \mathbf{x}_0 开始,第一次迭代选择沿着由 $\mathbf{d}_0 = -\nabla E(\mathbf{x}_0)$ 所确定的方向,于是有 $\mathbf{x}_1 = \mathbf{x}_0 - \alpha_0 \mathbf{d}_0$ 。与上面的简单例子有所不同,如何选择最优(或者甚至是较好的) α_0 值这一问题是无法精确求解的。因此,我们使用某种近似方法选择步长,最简单的方法就是对几个可能的步长分别测试 $E(\mathbf{x}_1)$, 并选择其中最优者。

如果我们已经得到了最优步长,接下来的搜寻方向就是当前方向的正交方向。总的说来,正是因为我们无法对每个直线搜寻方向都生成最优步长,才会有 $\nabla E(\mathbf{x}_0) \cdot \nabla E(\mathbf{x}_1) \neq 0$ 。共轭梯度的核心思想就是定义新的搜寻方向时,一定要保证新方向与原有的搜寻方向是正交的。为了做到这一点,我们可以定义

$$\mathbf{d}_1 = -\nabla E(\mathbf{x}_1) + \frac{(\nabla E(\mathbf{x}_1) \cdot \mathbf{d}_0)}{(\mathbf{d}_0 \cdot \mathbf{d}_0)} \mathbf{d}_0 \quad (3.20)$$

该定义表达式中使用 $\nabla E(\mathbf{x}_1)$ 作为新搜寻方向的一个估计值,但是舍弃了该向量中可以投影在原有搜寻方向 \mathbf{d}_0 上的那一部分。这一过程称为正交化,并且新形成的搜寻方向是原有方向的共轭方向(这也正是这类方法的命名原因)。我们现在就可以在 $\mathbf{x}_2 = \mathbf{x}_1 - \alpha_1 \mathbf{d}_1$ 方向内搜寻不同的步长,进行第二次迭代。

在生成共轭梯度方法的第三次迭代时,我们需要用 $-\nabla E(\mathbf{x}_2)$ 来估算搜寻方向,但需要保证该方向同时与 \mathbf{d}_0 和 \mathbf{d}_1 正交,这一思路在随后的迭代步骤中也不断重复。我们无法使这一过程无限地继续下去,因为在一个 N 维空间中,我们只能使一个矢量至多正交于其他 $(N-1)$ 个矢量。因此,对于这个定义精巧的算法,当完成不到 N 次迭代之后,我

们就不得不重新开始确定搜寻方向这一过程。

共轭梯度法是一种强大而稳健的算法。在实际问题中使用这一算法涉及到一些选择问题。例如,在每次迭代中如何选定步长,应该多久就重新开始搜寻方向的正交化过程。对于拟牛顿方法,可以仅仅通过估算函数值以及函数的一阶导数,就能够对函数进行最小化。

3.3.3 关于优化我真正需要知道什么

我们已经讲述了有关数值优化方法的许多内容,如果读者对这些方法并不熟悉的话,那么就需要把所有这些细节全部凝练成几个关键的结论。这并不仅仅是一个数学上的归纳,因为高效地数值优化是进行高效 DFT 计算的核心所在。我们强烈推荐读者回顾并重读 3.3.1 节中对数值优化方法所总结的 7 个特点。你或许会认为这些(谨向 Stephen Covey 致歉^[2])是“高效率优化计算工作者的 7 个习惯(seven habits of effective optimizers)”。这些特点不仅适用于一维优化,也适用于多维优化。我们还进一步推荐你们开始实际地运行 DFT 计算,从而可以更深入理解这 7 个思想中每一点的主要特征。

3.4 DFT 总能计算——一个迭代优化问题

最基本的 DFT 计算类型就是计算在空间中处于预设位置上的一组原子的总能。我们在第 2 章中已经给出了许多这类计算的结果,但是并未说明这些计算实际上是如何运行的。本节的目的就是说明这类计算在很多方面都类似于我们上面所讨论的优化问题。

正如我们在第 1 章所讨论的那样,DFT 计算的主要目的就是得到体系基态构型的电荷密度 $\rho(\mathbf{r})$ 。电荷密度是用 Kohn - Sham 方程解 $\psi_j(\mathbf{r})$ 定义为

$$\rho(\mathbf{r}) = \sum_j \psi_j(\mathbf{r})\psi_j^*(\mathbf{r}) \quad (3.21)$$

Kohn - Sham 方程为

$$-\frac{\hbar}{2m}\nabla^2\psi_j(\mathbf{r}) + V_{\text{eff}}(\mathbf{r})\psi_j(\mathbf{r}) = \varepsilon_j\psi_j(\mathbf{r}) \quad (3.22)$$

式中： $V_{\text{eff}}(\mathbf{r})$ 为有效势场(Effective Potential)。

造成这些方程非常难以直接求解的一个实际原因就是：有效势本身也是 $\rho(\mathbf{r})$ 的一个复杂函数。

我们前面对于数值优化的介绍表明了求解该问题的一个有效的通用策略，即尝试着对该问题进行迭代求解。我们先来估测一下全体电荷密度，然后用其作为试验值再去确定有效势。然后用这个有效势数值求解 Kohn - Sham 方程，从而确定了一个新的电荷密度。如果新的电荷密度和旧的电荷密度并不匹配，那么就还没有完成整个问题的求解。接着，将旧的和新的电荷密度以某种方式合并起来，给出一个新的电荷密度试验值。然后使用这个新的电荷密度试验值，确定一个新的有效势，并进而得到一个更新的电荷密度，如此反复。如果计算成功完成，这个迭代过程将会给出一个自洽解。上述描述忽略了所有各种数值细节，而为了使 DFT 程序在数值上更加高效，这些细节是非常重要的。

这个迭代过程与我们前面已经讨论过的、更为通用的优化问题之间存在着相似性，这表明：关于如何使用 DFT 计算能量，“高效率优化工作者的 7 个习惯”也许会给我们一些有益的启示，特别是当我们需要思考如何开始迭代过程的时候。在缺少其他信息的前提下，初始的电荷密度可以近似为每个原子隔离态的电荷密度。这是大多数 DFT 计算软件所通常使用的初始化方法。但是，如果能够得到一个更好的初始化近似，我们就能够更快地（更少的迭代次数）找到自洽解。因此，如果我们已经计算过非常类似于当前原子构型的另一个情形的电荷密度，那个计算所得电荷密度就可能是一个较好的初始近似值。正是因为这个原因，把大量计算所得的电荷密度及相关信息存储下来备用，在随后开始另一个类似计算时就有可能很有用的。

我们也需要思考一下，如何终止迭代计算。有时无法方便地直接比较电荷密度的两个解，并确定二者的近似程度，尽管这是确定我们是否已经找到自洽解的最直接的测试方法。另一个较为容易解释的方法就是在每次迭代之后都计算与电荷密度对应的能量。毕竟这才是我们在求解过程中最终感兴趣的物理量。如果我们的迭代计算不断收敛，

那么在连续两次迭代之间的能量差就会趋近于零。这意味着,可以在两次迭代能量差的数量级小于一个合适容限值时,就终止该迭代过程。绝大多数 DFT 计算软件都对该参数定义了一个合理的默认值,但是对于高精度计算而言,就应该使用更低的数值。计算中所使用的这类迭代算法具有一个普遍性特征:一旦找到了一个较好的电荷密度近似值,它们就会快速地收敛。这意味着即便把收敛容限值(该参数定义了自洽计算的终点)减小若干个量级,也常常只会产生数量较少的额外迭代步骤。

在从这一节继续向下讲述之前,需要读者能够明白你所能获得的各种 DFT 计算软件:①使用什么算法来求解自洽 Kohn - Sham 问题;②你如何能验证某个计算的结果已经达到了自洽解。

3.5 几何优化

3.5.1 内部自由度

目前,我们只是描述了以下情形的 DFT 计算:超晶胞中每个原子的位置都是固定的。如果我们只是想要预测简单固体的晶格常数,这就已经足够了;但是,对于除此之外的其他问题,上述方法则只会给我们一个观察整个世界的有限视角!例如,试想一下,假如我们所感兴趣的是与氮有关的一组反应,那么必定有这样一件事是我们想要知道的:就是气相 N_2 分子的能量,因为如果要描述任何从气态 N_2 开始的总反应,就可能需要知道这一信息。为了计算这个能量,我们需要找到使其分子能量最小化的 N_2 几何构型。由于这是一个非常简单的分子,其实这个任务也就意味着必须确定其键长。那么,我们如何使用基于周期性超晶胞的 DFT 计算来完成这一任务呢?

为了模拟一个气相分子,我们需要建立一个几乎全部是真空的超晶胞。一个简单方法就是先建立一个边长为 L 的立方超晶胞,然后把两个 N 原子放在这个超晶胞的分数坐标 $(0,0,0)$ 和 $(+d/L,0,0)$ 位置上。只要 L 显著大于 d 的长度,这个超晶胞就表示了键长为 d 的一个

隔离态 N_2 分子^①。现在我们就用拟牛顿法,或者是共轭梯度法,通过最小化上述超晶胞的总能(在最小化过程中,固定超晶胞的尺寸和形状,但是允许两个原子的分数坐标自由变动),就可以得到 N_2 分子 DFT 优化后的键长。此外,我们必须定义一个终止标准,用以决定是否这些迭代计算已经收敛到最小值。因为在我们正在寻找的最终原子构型中,两个原子上的力均为零,所以我们设定当两个原子上力的数量级小于 $0.01\text{eV}/\text{\AA}$ 时,就终止迭代计算。在这个例子中,由于对称性使然,在两个原子上力的数量级始终是完全相同的,当然两个原子上的力分别指向相反的方向。这个力的标准为什么是合理的呢?如果我们把一个原子的位置稍稍改变一个很小的量 (Δr),那么由于这一变化所导致的总能变量就可以估算为 $|\Delta E| \approx |\mathbf{F}| |\Delta r|$,其中 $|\mathbf{F}|$ 是作用于原子上的力。如果在两个原子上的力都小于 $0.01\text{eV}/\text{\AA}$,那么,当把任何一个原子移动 0.1\AA 时(相对于化学键长,这是相当大的一个距离),则总能的变化都小于 0.001eV ,这是一个很小的能量。我们只是通过这一讨论告诉读者,用力的数量级作为计算的终止标准是合理的。

我们必须要进行的第二个选择是:在初始放置两个原子时,二者应该距离多远。这一选项似乎并不很重要,因为我们将要进行能量的最小化,并且假定只有一个键长真正确定了这个能量最小值。但是,对于键长初始估值的选择,也会在很大程度上影响该计算能否顺利完成。例如,我们用共轭梯度优化法,使用两个不同的原子间初始距离,分别进行计算。当我们把两个原子的初始间距设为 1\AA 时,我们的计算进行了 11 次共轭梯度迭代之后,力已经满足了上述收敛标准,并且最终所得键长为 1.12\AA 。实验上观测到 N_2 的键长为 1.10\AA 。如同我们在第 2 章所考察的体相固体的晶格参数一样,这个 DFT 结果也不是精确的,但是在 DFT 优化几何构型和实验结果之间的差值很小。

在第二次的计算中,初始原子间距为 0.7\AA 。此时,经历了 25 次共轭梯度迭代之后,两个原子间距为 2.12\AA ,这更加糟糕,并且该算法还

① 只要这个超晶胞足够大,该分子的周期性影像对其总能的影响就很小,尤其是对于类似于 N_2 这样没有偶极矩的分子。

没有收敛到一个能量极小值。为什么会这样？让我们来思考一下我们的初始状态，当两个原子间距为 0.7\AA 时，就相当于在物理上对其键长进行了极大压缩（相比于 N_2 分子键长的正常值）。这就意味着存在着一个极大的排斥力，把两个原子相互推开。这会产生很大影响，因为在我们的优化方法中，要根据某个位置上的能量导数估测总能的变化有多快。数值优化方法识别到两个原子想要彼此移开，这就迫使其采用一个较大的初始步长，从而将两个原子分隔开一个更大的距离。

然而，相比于我们所给出的初始估值，两个原子现在所处的位置（在某种意义上）就是对分子最低能量状态给出的一个更糟糕近似；而且，该计算也就无法回复到原有状态并找到真正的极小值。简言之，我们的计算非常不幸地失败了，因为所使用的初始几何构型在化学上并不是真实而合理的。关于怎样从一些细节上判断是否已经出现了这类失误，是与优化算法的具体细节有直接关系的；但是，对于任何一种几何优化方法，如果对你所研究的原子几何构型使用了很差的初始估值，那么就能够（并很可能会）发生这种失误。在这里所应该记取的一个重要经验就是：花费较大计算工作量，对原子坐标给出一个良好的初始估值，可以极大地提高 DFT 计算速度，并且在很多情形下，对于计算能否收敛到一个能量极小值也具有决定性影响。

对于优化超晶胞中原子位置的第二个例子，我们来优化一个 CO_2 分子的几何构型。如果再次使用一个边长为 L 的立方超晶胞，我们可以通过下列方法创建一个 CO_2 分子：把 C 原子放置在分数坐标 $(0,0,0)$ ，而两个 O 原子放置在 $(+d/L,0,0)$ 和 $(-d/L,0,0)$ 位置上。从 $d = 1.3\text{\AA}$ 的初始状态开始优化该分子的能量，并采用与 N_2 分子计算相同的终止标准，最终得到优化后的 C-O 键长为 1.17\AA ，且 O-C-O 键角为 180° 。

这个 CO_2 的结果似乎是很合理的，但我们能否相信这个结果呢？我们在定义终止标准时，所使用的是作用于原子上力的数量级。让我们来考察一下这个原子构型中作用在一个 O 原子上的力，如果我们把这个力写作 $\mathbf{f} = (f_x, f_y, f_z)$ ，那么根据对称性，无论我们对 d 怎样取值，都有 $f_y = f_z = 0$ 。这说明：在能量最小化过程中，分子的几何构型被反复

更新, C-O 键长将会变化, 但是 O-C-O 键角将始终固定在 180° , 也就是我们在初始构型中所给定的角度值。换种方式来说, 在上述计算中, 键角并非一个真正的收敛结果, 这是我们在分子结构的初始估值中, 所施加对称性的一个必然结果。

优化这个分子另一个更可靠的方法就是: 选择一个初始键角, 并使力的两个分量不会仅仅由于对称性而消失为零。当我们把以下构型作为出发点, 就能很容易地做到这一点: C 原子放置在分数坐标 $(0, 0, 0)$, 两个 O 原子放置在 $(+a/L, b/L, 0)$ 和 $(-a/L, b/L, 0)$ 位置上。从该结构出发 (初始态的 $a = 1.2 \text{ \AA}$, $b = 0.1 \text{ \AA}$, 初始 O-C-O 键角为 107.5°), 对该分子的能量进行最小化之后, 得到的收敛结果是 C-O 键长为 1.17 \AA , 而 O-C-O 键角为 179.82° 。这个几何构型与我们从直线分子结构出发时得到的结果非常接近。我们不可能期望这两个构型完全相同; 因为它们是采用迭代优化方法, 通过我们所设置的收敛标准, 在一旦满足“足够接近”于真实解时就立即停止计算的状态。根据这两个计算结果可以很合理地归纳出一个结论: 我们的 DFT 计算预测 CO_2 是一个直线形分子 (O-C-O 键角为 180°)。实验上已知 CO_2 是一个直线形分子, 且 C-O 键长为 1.20 \AA , 因此 DFT 预测结果与试验结果具有良好的一致性。

能够高效地对原子集合总能进行最小化, 也许是大多数 DFT 计算的核心问题。在继续下面的内容之前, 读者应该重读这一节, 目的是总结一下我们在考察两个简单例子时所确认过的那些陷阱。对于为什么会存在这些陷阱, 以及如何能够发现或者避免这些陷阱, 对这些问题形成较强的直觉将会在将来的计算中节省很多不必要的工作量。

3.5.2 具有约束原子的几何优化

优化一部分原子的位置, 而保持其他原子位置固定不动, 在上述前提下实现一个超晶胞能量的最小化, 这对于许多种计算都是很有意义的。下一章中我们对固体表面的相关计算, 就是这类情形的一些具体例子。相比于无约束的优化问题, 牵涉到约束的优化问题通常都在数值计算上更具挑战性。令人高兴的是, 对于基于力的几何优化 (例如

上文所描述的方法),可以很容易地扩展到超晶胞中一个或多个原子被约束的情形。对于有约束的体系,在优化计算的每次迭代过程中,只有那些未约束原子的位置会被不断更新。如果这类计算的终止标准是基于每个原子上力的数量级,那么在确定终止标准时,就只考虑了那些未约束原子。

3.5.3 优化超晶胞的体积和形状

上述计算可以允许超晶胞中原子的位置发生变化,但是超晶胞的尺寸和形状是固定不变的。但在第2章所介绍的计算中,我们改变了超晶胞的尺寸来确定几种不同体相固体的晶格常数。读者可能已经懂得:既然数值优化方法能够使我们优化原子位置,从而应该也可以将其扩展用以优化超晶胞的尺寸。我们不对这些计算的细节进行深入说明——读者应该阅读你所使用的DFT计算软件的说明文件,从中找到如何使用该软件来精确地运行这类计算。我们只在这里给出一个例子,在第2章中我们通过分别对晶格参数 a 和 c 的许多不同取值,进行许多次单独计算,从而试图找到Cu在HCP晶体结构中的晶格常数(可回顾图2.4所示内容)。一个更容易地完成这个计算任务的方法就是:先创建一个晶格参数 a 和 c 较为合理的HCP Cu初始超晶胞,然后优化该超晶胞的体积和形状直至总能最小化^①。如果在运行该计算时,与第2章计算HCP Cu时所使用的 k 点数量等计算参数都一样,所得到的优化后的结构应该是 $a = 2.58\text{\AA}$, $c/a = 1.629$ 。这些数值与第2章的计算结果是完全一致的,但是这里的结果显然更为精确。而且,相比于生成第2章的那些数据,我们得到这些结果所需的计算工作量更少。

^① 对于允许超晶胞体积变化的计算,存在着一个微妙而敏感的难题,这就是:在超晶胞形状和体积变化所导致的力中,有可能包含着系统数值误差,除非 k 点数量和截断能都很大。由于这一效应所引发的人为虚拟应力称为Pulay应力(参见扩展阅读)。减小这一效应的常用方法就是在对超晶胞体积进行优化时,将截断能提高30%~50%。这样就必须使用标准截断能对最终优化后的结构重新计算总能,这样才是一个完整的计算。

练习

1. 在第2章的练习中我们建议读者进行几个材料的计算,包括立方和 FCC 结构下的 Pt,以及 CsCl 结构的 ScAl。重复这些计算,并在这次计算时提出相关证据,用以表明计算结果已经在 k 空间取样和截断能上收敛得很好。

2. 我们已经说明了如何使用对分法和牛顿法,找到 $f(x) = e^{-x} \cos x$ 的一个极小值。如上所述,使用不同的初始预估值,应用这两个方法找到相同的极小值。这对图 3.6 所示收敛特点产生了怎样的变化? 该函数具有多个极小值,使用牛顿法找到至少两个极小值。

3. 只有初始估值足够接近于真实解,才能保证牛顿法能够收敛。为了能够明白这一点,试着用牛顿法找到 x 值,使得 $g(x) = \tan^{-1} x = 0$ 。此时,牛顿法应写作 $x_{i+1} = x_i - (1 + x_i^2) \tan^{-1} x_i$ 。采用以下初始估值: $x_0 = 0.1$, $x_0 = 1$ 和 $x_0 = 10$,并探讨该方法的优异程度。

4. 我们没有对多维牛顿法定义一个计算终止标准,你会如何定义这样一个标准呢?

5. 使用类似于 3.5.1 节的方法,优化 H_2O 和氰化氢 (HCN) 的几何构型。HCN 是一种剧毒气体,尽管如此,因为它可用作许多工业制程中的化学前驱体,所以产量也很大。在计算中,可以通过每种分子从多种不同初始几何构型都收敛到同一种最终状态,来确认你已经考虑了所有的自由度。将你计算的几何构型与实验数据相比较。

6. 在第2章的练习中,我们建议读者计算六方 Hf 的晶格常数 a 和 c 。使用一个优化该超晶胞体积和形状的方法重复这一计算。你的计算结果与第2章中较为费力的计算结果是否是一致的? 相比于理想的硬球堆积值, c/a 的扭曲偏离程度有多大?

7. 运行必要的计算,用以估算从 FCC Pd 和 FCC Cu 形成有序 BCC CuPd 的能量差。有序合金可以定义为一个 BCC 晶体,其中 Cu 原子处于每个顶角,而 Pd 原子处于立方体中心(或者是反过来)。已知当 Pb 和 Cu 两种原子按照 1:1 的化学配比相混合时,该有序合金是一种更

倾向于在低温下存在的晶体结构。根据这一结论,你所试图计算的能量差的正负号是怎样的?为了计算这个能量差,你需要优化每种材料的晶格常数,并且需要注意应该如何选择截断能和 k 点。

参 考 文 献

- [1] Kresse G, Joubert D. From Ultrasoft Pseudopotentials to the Projector Augmented - Wave Method. *Phys. Rev. B* 59, 1999 : 1758.
- [2] Covey S R. *The Seven Habits of Highly Effective People*. Simon & Schuster, New York, 1989.

扩 展 阅 读

倒易空间的概念、布里渊区,以及类似内容是所有固体物理教材必不可少的重要内容。在第 1 章和第 2 章的扩展阅读中都列出了一些相关书目。对能级、能带、 k 空间和能带结构进行了清晰介绍的另一个文献是: Hoffmann R, Janiak C, Kollmar C. *Macromolecules* 24. 1991 : 3725.

对于学习了解有关优化的高效数值方法,一个极好的文献是: Press W H, Teukolsky S A, Vetterling W T, et al. *Numerical Recipes in C + + : The Art of Scientific Computing*. Cambridge University Press, Cambridge, UK, 2002。该书有多个种版本,并且有用其他计算机语言编写的相同内容。

对共轭梯度法进行了深入讨论的一个文献是: Shewchuk J R. *An Introduction to the Conjugate Gradient Method Without the Agonizing Pain* [EB/OL]. <http://www.cs.cmu.edu/~quake-papers/painless-conjugategradient.pdf>。

Methfessel 和 Paxton 模糊化方法 (Smearing Method) 可参考: Methfessel M, Paxton A T. *Phys. Rev. B* 40. 1989 : 3616.

对于四面体方法 (tetrahedron methods) 的详细信息,请参见: Blöchl P E, Jepsen O, Andersen O K. *Phys. Rev. B* 49. 1994 : 16223.

关于赝势和 PAW 方法的发展历史和详细信息,请参见: Kresse G,

Joubert D. Phys. Rev. B59. 1999:1758, 以及其中的参考文献。

当在改变超晶胞体积的情形下计算超晶胞中的力时, 对于 Pulay 应力以及涉及到有限平面波基底和 k 点的相关复杂问题, 更多信息请参见 Francis G P, Payne M C. J. Phys. Condens. Matter2. 1990:4395.

附录 计算细节

本章的所有计算都使用 PW91 GGA 泛函。

3. 1. 3 和 3. 1. 2 节 体相 Cu 计算使用了含有 4 个 FCC 结构 Cu 原子的立方超晶胞, 截断能为 292eV。使用了 Methfessel - Paxton smearing, smearing width 为 0. 1eV。

3. 5. 1 节 对于 N_2 和 CO_2 分子的计算, 使用了边长为 10Å 的立方超晶胞, 使用了 Monkhorst - Pack 方法对倒易空间取样, 所设置的 k 点为 $3 \times 3 \times 3$ 。这些计算中的截断能为 358eV。

第 4 章 固体表面的 DFT 计算

4.1 表面的重要性

表面对于许多领域而言,在技术上都是非常重要的,如催化、界面、气体分离薄膜、半导体制造等。理解表面的原子几何构型和电子结构是很有意义的,例如,已经在表面结构及其催化活性之间建立了二者之间普遍存在的关系。有关催化剂的研究中,DFT 计算的一个重要应用领域就是:对减少汽车 CO、NO_x、SO_x 排放的三效催化剂(Three - Way Catalysts)进行技术改进,并且降低其成本。这些催化剂不仅实现了将碳氢化合物和 CO 进行氧化这一重要目标,并且同时能够减少 NO_x。传统意义上,这些催化剂都与贵金属(如铂)有关,并且在特定条件下也常常会发生催化剂中毒效应。对于催化金属、沸石和氧化物之间的相关反应机理,DFT 方法能够给出非常有用的信息,并且可以指导催化转换器(Catalytic Converters)的设计,以改善其效率、降低成本。

在一些科研项目中,表面科学实验和 DFT 计算经常成功地结合在一起。DFT 计算常常与超高真空表面科学的实验技术相结合,例如:扫描隧道显微镜(Scanning Tunneling Microscopy, STM)、程序升温脱附(Temperature Programmed Desorption)、X 射线衍射(X - Ray Diffraction)、X 射线光电子能谱(X - Ray Photoelectron Spectroscopy),从而可以确定金属、金属氧化物、纳米颗粒、碳化物和硫化物的表面结构。例如,在 1959 年,通过对 Si(111) 表面进行电子衍射实验,表明该表面具有复杂对称性,并且形成表面的原子排列与体相 Si 晶体结构完全不同。由于 Si 在微电子技术中的重要性,详细了解其表

面结构是非常有必要的。直到 1992 年,才完成了一个很大的超晶胞进行了 DFT 计算,从而对 $\text{Si}(111) - (7 \times 7)$ 表面重构进行了有效性测试,并且该测试结果迄今为止都获得广泛认可。这些计算不仅明确了该表面中每个原子的位置,并且可以用来模拟该表面的 STM 成像,有助于解释实验所得图像(这些实验图像虽然完美,但却很复杂)。

对于另一种与该技术相关的材料,也将 DFT 结果与一系列实验方法相结合,从而对其表面化学进行深入理解,这就是二氧化钛(TiO_2)。二氧化钛是一种重要的材料,可用于颜料、氧传感器、抗菌剂以及金属催化剂的载体。二氧化钛还可以用作光伏电池中的光敏剂,也可以用作光电解电池中的电极涂层,提高水的电解分解效率。DFT 研究已经与许多不同的实验技术相结合,用于表征不同原子、分子物质与表面之间的键合,以及表面自身的特征。对于许多过程而言,氧原子在表面上的消失位置是非常重要的,而 DFT 计算不仅能够有助于明确该位置,而且有助于解释该现象背后的物理意义。

在本章中,主要说明如何将 DFT 计算用于考察固体表面。首先,介绍一些定义洁净表面原子结构所必需的概念,然后对有关气体 - 表面这样的界面计算给出一些范例。

4.2 周期性边界条件和板块模型

当研究对象的是一个表面时,理想的模型应该是在两个方向上无限大,而在沿表面法向上有限大的一长条材料。为了在建模时实现这一点,很自然地可以在两个方向上利用周期性边界条件的优势,但在第三个方向上则行不通。虽然已经有相关计算软件能够实现这一点,但在研究表面时,更加普遍使用的计算程序是在所有三个坐标方向上都应用了周期性边界条件,本章中也只对这种方法进行讨论。其基本概念如图 4.1 所示,超晶胞在竖直方向上,只有一部分区域内含有原子。在超晶胞下部,这些原子在 x 和 y 方向充满了整个超晶胞,但在这些原子的上方,也就是在超晶胞的顶部区域则是空的,因此这种模型称为

板块(Slab)模型。当该超晶胞在所有三个方向上都重复时,就定义了一系列堆垛在一起的,并由空白区域所分隔开的固体材料板块,如图 4.2 所示。在 z 方向上分隔开周期性板块图像的空白区域称为真空区(Vacuum Space)。当使用这种模型时,一定需要具有足够大的真空区,以便材料的电荷密度在真空区逐渐消失为零,从而使该板块的顶面与相邻板块的底面之间不存在实质的相互作用。图 4.3 所示为从真空区内观察使用这种方法所定义材料中的原子视图,可以注意到:从这个角度观察,超晶胞实际上定义了两个表面,一个是上表面,一个是下表面。

假设对 FCC 金属(例如 Cu)的表面进行计算,应当如何建立图 4.1 所描述的板块模型呢?使超晶胞的晶格矢量分别与笛卡儿坐标系的 x 、 y 和 z 轴一致是非常方便的,其中 z 轴与超晶胞中的表面法向一致。对于 FCC 金属,其晶格常数与常规晶胞中立方体的边长相等。超晶胞的矢量应该是

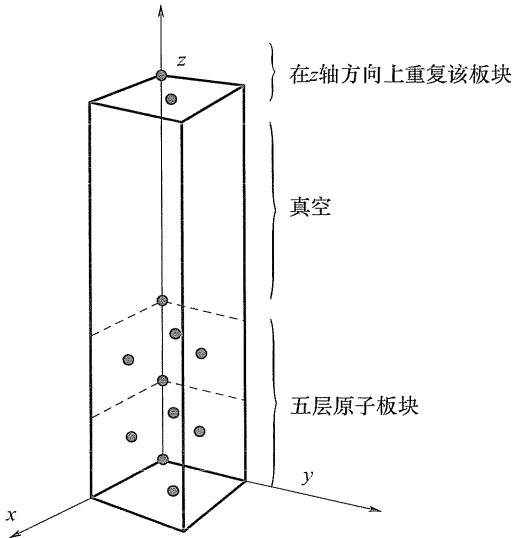


图 4.1 使用所有 3 个方向上都具有周期性边界条件的超晶胞定义固体表面材料模型

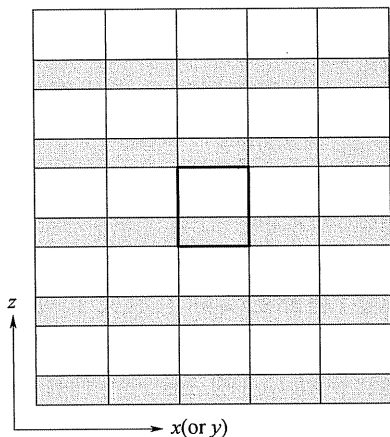


图 4.2 图 4.1 中超晶胞所定义材料的二维示意图(阴影区域表示原子所占据的空间区域;白色区域表示真空)

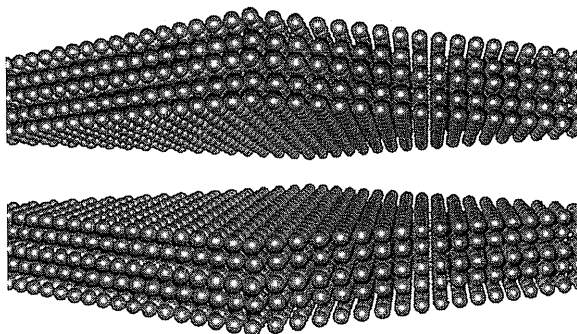


图 4.3 在完全周期性计算中所使用的五层原子板块的表面模型视图(对于该视图,其超晶胞与图 4.1 中类似,并在 x 和 y 方向上重复了 20 次,而在 z 方向上重复了 2 次)

$$\mathbf{a}_1 = a(1, 0, 0) \quad \mathbf{a}_2 = a(0, 1, 0) \quad \mathbf{a}_3 = a(0, 0, 5) \quad (4.1)$$

式中: a 为晶格常数。

图 4.1 中原子的分数坐标是

0.0,0.0,0.0	0.5,0.5,0.2
0.5,0.5,0.0	0.0,0.5,0.3
0.0,0.5,0.1	0.5,0.0,0.3
0.5,0.0,0.1	0.0,0.0,0.4
0.0,0.0,0.2	0.5,0.5,0.4

如上所述,则定义了含有5层独立原子的板块,且在 z 方向上,两层原子间距为 $a/2$,而真空区宽度为 $3a$ 。对于该超晶胞,板块中的每层原子均含有两个独立原子。在定义 α_3 时,数值5具有一定的随意性。该矢量必须足够大,从而保证在超晶胞中含有足够的真空区;但真空区越大则所需的计算量越大。实际计算时,需要对若干个具有不同真空区宽度的模型,检验真空区中的电荷密度,从而判定需要多少真空区才是“足够”的。理想的情况是在真空区中电荷密度接近于零。如果读者想要做该测试,则需要注意:在改变 α_3 长度的同时,原子分数坐标的 z 值也需要相应地调整,以保证材料中相邻两层原子之间是正确的物理间距(如果不能很好地理解这一点,请回顾一下分数坐标的定义)。

我们希望板块模型能够很好地模拟一个真实表面的重要特征。上述讨论中的板块模型包括有FCC金属的5层原子。当然,一个真实表面是一块材料的最边缘部分,且除非在某些极端条件下,该材料至少几微米厚,而这与上述5层表面模型完全不同。这就提出了下列问题:多少层原子才是足够的?通常而言,原子层数越多越好,但所用原子层数越多,则一定会使计算时间延长。怎样判定原子层数已经达到足够收敛了,这一点决定于所研究材料的物质本性和所需计算的内容。如同真空区宽度一样,这个问题也可以通过如下方法找到答案:计算并观察某些特性(如表面能或吸附能)随着表面原子层数的增加是如何变化的。后文中会就此给出一些具体例子。在选定模型总的原子层数时,通常是在计算成本和物理精度之间取得折中和平衡。

4.3 选定表面计算的 k 点

在第3章中,讨论了在布里渊区上积分时如何选择 k 点。对于板

块计算,以及其他与体相计算具有相同特点的情形,都可以用 Monkhorst - Pack 方法选定 k 点。为了进一步直观地了解这一点,可以考察一下与上述超晶胞相关的倒易晶格矢量。根据等式(3.2),可得 $|\mathbf{b}_1| = |\mathbf{b}_2| = 2\pi/a$, 且 $|\mathbf{b}_3| = 2\pi/5a$, 因为超晶胞在实空间中较“长”的尺寸,在倒易空间中则较“短”。这表明在倒易空间中求解积分的近似解时,在 \mathbf{b}_3 方向上并不需要其他两个方向上那么多的 k 点。对于该超晶胞,使用一个 $M_1 \times M_1 \times M_2$ 的 k 点网格。其中,对于该几何构型而言, $M_1 > M_2$ 是较为合适的。这一点与第3章中 HCP 金属选定 k 点的讨论是类似的。

超晶胞中较长的尺寸包括了真空区,这一事实引入了另一个有关于 k 点选择的有益经验。如果真空区足够大,则在距离板块边缘很短的位置上,电荷密度就消失为零。这意味着:在 \mathbf{b}_3 方向上只使用一个 k 点就有可能得到精确的结果^①。因此,对于板块计算常常使用 $M \times N \times 1$ 的 k 点网格。其中,所选定的 M 和 N 能够在表面模型的晶面内,对其 k 空间进行足够精确地取样。对于上述讨论中 $|\mathbf{b}_1| = |\mathbf{b}_2|$ 的例子, M 和 N 是相等的,并符合超晶胞的对称性。

4.4 采用 Miller 指数对表面分类

对于图 4.3 所示表面,在物理上可以想象为将体相晶体劈裂 (Cleave) 后所形成的。如果认真思考一下真实晶体劈裂,则可以沿着很多不同晶面对晶体进行劈裂,并且这些不同的劈裂方向也恰恰表明了不同的表面,表现为表面上不同的原子排列方式。这意味着需要采用一种标注方法,用于定义所讨论的表面。

图 4.4 给出了 Cu 表面的晶面与体相晶体结构的对应关系。该晶面的朝向可以用与该晶面相垂直的法向矢量来描述^②。从图 4.4 中可

① 发生这一现象的原因是由于在该坐标上电子能带结构的分散度极小。

② 该晶面通常表示为圆括号中的三个指数 (hkl); (在立方晶系中) 与该晶面相垂直的法向方向表示为方括号中的指数 $[hkl]$ 。

见,这个矢量的一个有效选择就是 $[0,0,1]$,另一个有效选择是 $[0,0,2]$ 或 $[0,0,7]$,以此类推,因为这些矢量彼此平行。

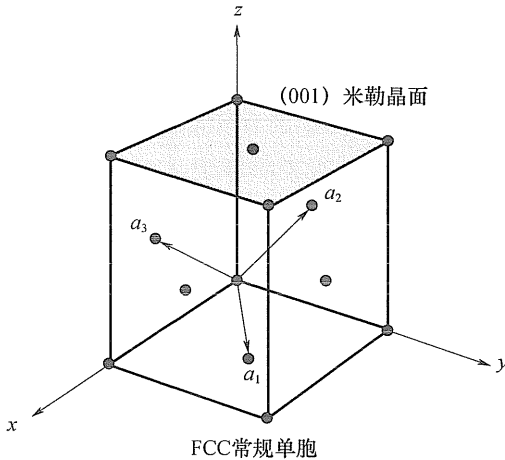


图 4.4 FCC 金属的常规单胞及其(001)米勒晶面

为了确定应该使用这些矢量中的哪一个,通常需要先确定该晶面与材料原胞(Primitive Cell)或常规单胞(Conventional Cell)(二者均可选用)三轴的截交点。将这些截交点的倒数同乘以某个因子,从而将每个倒数都化为整数,并使每个倒数都尽可能地小,最终所得的数值就称为该表面的 Miller 指数。如图 4.4 所示,该晶面与常规单胞的 z 轴相交于 1(以晶格常数为单位),而与 x 和 y 轴并不相交,则这些截交点的倒数是 $(1/\infty, 1/\infty, 1/1)$,于是该表面可以表示为(001)。这些指数并不需要同乘以某个因子,因此图 4.4 所示表面就称为(001)表面。

如果定义该表面垂直方向的矢量需要加个负号,则其 Miller 指数的相应分量顶上也要加一个横杠。根据这个表示方法,图 4.4 中立方体的底面所定义的表面也是(001)表面。可以肯定的是,图 4.4 中立方体的其他 4 个表面是(100)、(010)、 $(\bar{1}00)$ 和 $(0\bar{1}0)$ 表面。对于许多简单材料而言,不需要严格区分这 6 个表面,根据体相晶体的对称性,

这 6 个表面都具有相同的结构^①。

图 4.5 中显示了面心立方材料的另一种表面,该晶面与 x 、 y 和 z 轴分别截交于 1、1 和 1,则这些截交点的倒数是 $1/1$ 、 $1/1$ 和 $1/1$,因此,该表面就是(111)表面。这也是一个非常重要的表面,对于 FCC 材料的所有 Miller 指数表面而言,该表面上的原子层具有最高的原子密度。也正是如此,这个表面在平衡态的真实晶体中具有重要的意义。上述定性的讨论表明:在 Cu 的真实多晶体中,Cu(111)表面应该占据了该晶体整个表面的绝大部分。

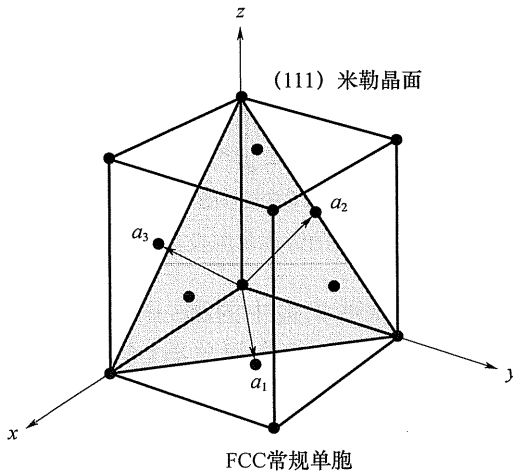


图 4.5 与图 4.4 类似的(111) Miller 晶面

图 4.6 显示了 Cu(110)表面。该(110)表面的晶面与 x 和 y 轴均截交于 1,而与 z 轴并不相交。这些截交点的倒数是 $(1/1, 1/1, 1/\infty)$,则该表面可以表示为(110) Miller 晶面。

图 4.7 显示了 FCC(001)、(111)和(110)表面的俯视图。这些视图给出了每个表面的不同对称性。(001)表面具有四重对称性(Fourfold Symmetry),(111)表面具有三重对称性(Threefold Symmetry),而

^① 这 6 个表面可以用波形括号指数 $\{100\}$ 表示。

(110)表面具有二重对称性(Twofold Symmetry)。这三种 FCC 表面全是原子尺度的平直表面,即:任一种表面上的每个原子都具有相同的配位关系,且在表面法向上的坐标也相同。这些表面综合在一起称为 FCC 材料的低指数表面。其他晶体结构也具有自身的低指数表面,但与 FCC 结构不同的是:其他结构的低指数表面往往具有不同的 Miller 指数。例如,对于 BCC 材料,其原子密度最高的表面是(110)表面。

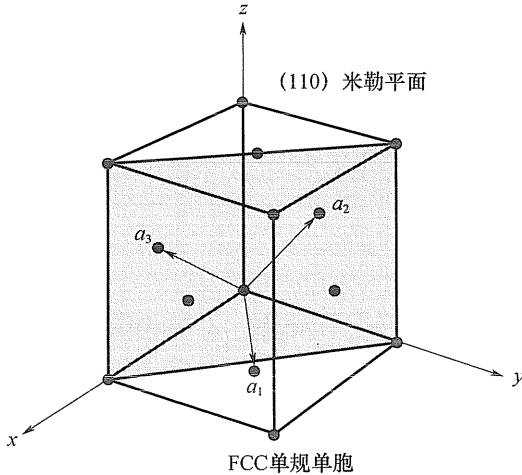


图 4.6 与图 4.4 类似的(110)Miller 晶面

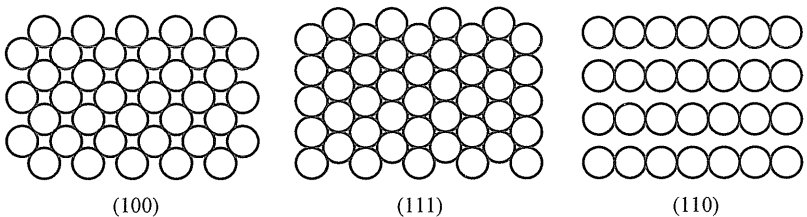


图 4.7 FCC 金属(100)、(111)和(110)表面的俯视图(只显示了顶层的部分原子)

上文仅仅给出了低指数表面的一些示例,这些表面的稳定性更高,因此对于实际情形而言是很重要的。这些表面对于 DFT 计算也很方便,因为定义低指数表面的超晶胞只需要适当的原子数。然而,还有其他许多有意思的和具有重要特性的表面。这里只给出一个例子,

图 4.8 所示是 Cu(322) 表面的俯视图和侧视图。该表面的许多局部区域看起来都很像(111)表面,并且这些区域均由一个原子高度的台阶所分隔开来。相比于表面上的其他原子,这些位于台阶边缘处的原子具有较低的配位数,这往往导致其具有更高的反应活性。在催化合成氨时,台阶边缘的反应活性原子就扮演着至关重要的作用,这一点已经在第 1.2 节中的第一部分进行了说明。尽管 Cu(322) 结构要比低指数表面更为复杂,但在该表面的晶面上仍然是周期性的,因此,仍然可以使用超晶胞来描述该表面。

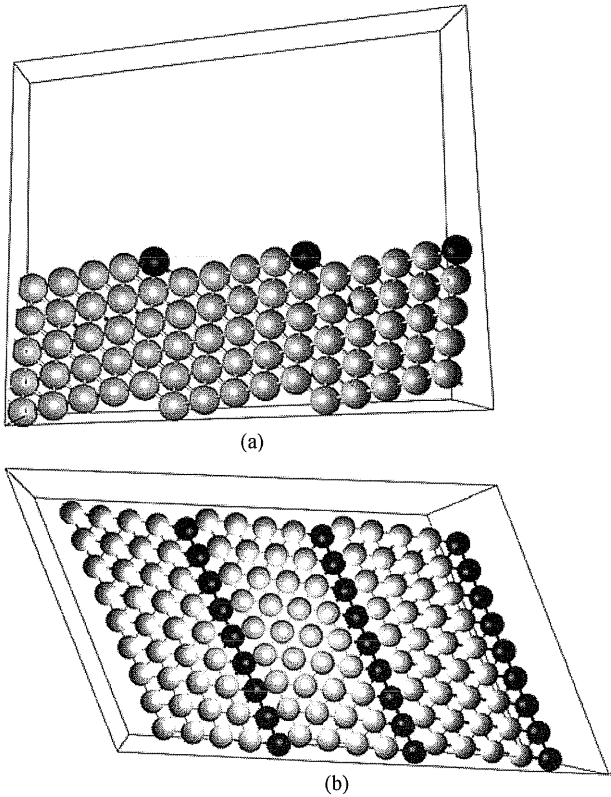
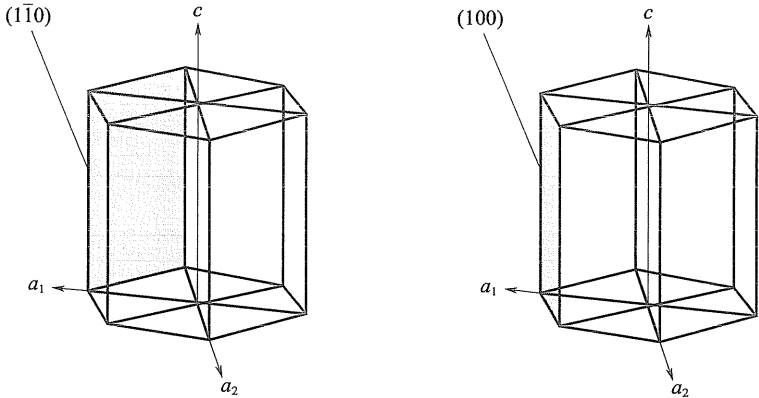


图 4.8 Cu(322) 的俯视图和侧视图(形成台阶边缘的 Cu 原子都用深灰色表示)
(a) 俯视图; (b) 侧视图。

在上面提到了其他晶体结构可能具有与 FCC 结构不同的 Miller 指数。在密排六方 (Hexagonal Close - Packed, HCP) 结构中,小球的排布规则是:以一个密排原子层作为基面 (Basal Plane),其中每个原子均由其他 6 个原子所包围;添加后续原子层时,小球交替地放置在基面的三重对称孔穴位置上。如此添加到第三层原子时,小球正好是在基面小球的正上方,从而得到 HCP 结构(如果第三层原子添加在孔穴位上,而没有恰好位于基面的正上方,则得到了 FCC 结构)。

在垂直于基面的方向上,密排六方 (HCP) 材料具有六重对称轴。使用三轴对称性来定义这种结构的 Miller 指数是不符合其对称关系的。如图 4.9 所示,两个晶面的对称性是等效的,然而它们的 Miller 指数并没有反映出这种关系。这一点很令人遗憾,因为 Miller 指数非常有用的一个原因就是:等效晶面具有相似的指数。另一个不足之处就是在这个例子中 $[hkl]$ 方向并不是垂直于 (hkl) 晶面。



截距值: 1, -1, ∞
倒易空间值: 1, -1, 0

截距值: 1, ∞, ∞
倒易空间值: 1, 0, 0

图 4.9 在 HCP 固体中使用三轴体系标注晶面(这两个晶面的对称关系是等效的,而其三轴指数却并不相同)

一个更好的办法就是对 HCP 固体使用四轴和四指数体系。在图 4.10 中给出了几个例子,Miller 指数的获得也如前文所述,即采用该晶面与四个坐标轴截交点的倒数来表示。采用这一体系,上述讨论中的两

个等效晶面就变为(1100)和(1010)。这样一来,由六重对称轴所导致的六个等效晶面可以标注为{1100}晶面系。使用四轴体系, $[hkil]$ 方向也与 $(hkil)$ 晶面相垂直,这一点与立方固体中使用三轴体系是相同的^①。

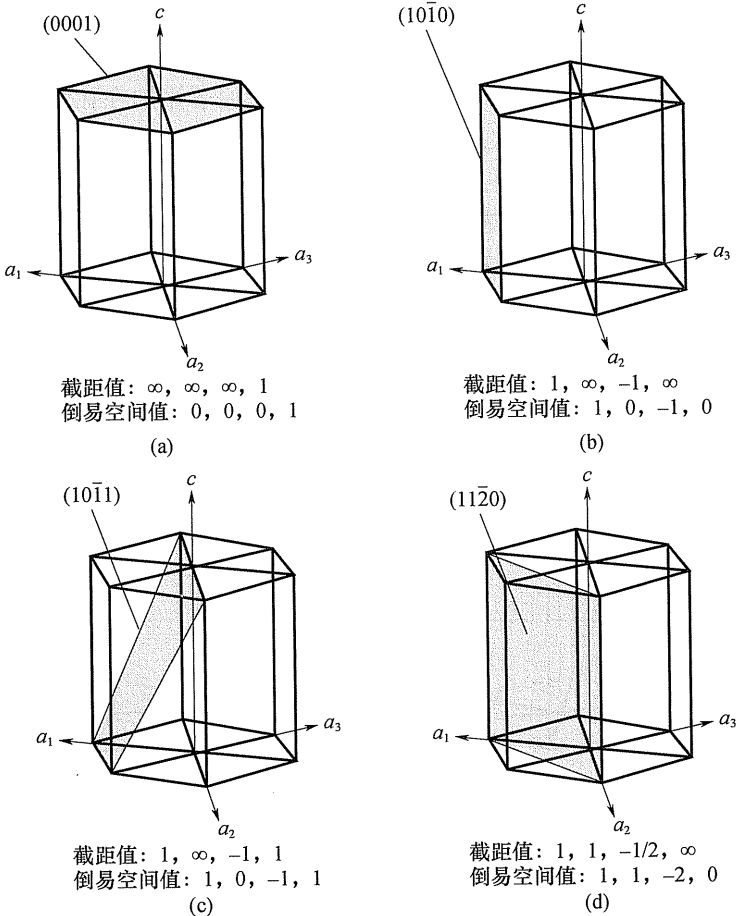


图 4.10 对 HCP 材料使用四轴和四指数体系的 Miller 指数示例

① 因为四个坐标轴并不是完全线性无关的,在四指数体系中的前三个指数之和总是为零(这可以作为一个常规校验方法)。

4.5 表面弛豫

在上述例子中,为了建立五层原子的板块模型,将原子放置在了板块模型中。原子的位置是理想的,是 FCC 材料中的体相原子位置。在体相的 FCC 金属中,任何两个相邻原子的间距一定是相等的。但是,对于靠近材料表面的原子层而言,则没有理由保持相同的原子层间距。相反地,因为与体相中的原子相比,表面处原子的配位数减少了,可以自然地想到:靠近表面处的原子层间距应该与体相中有所不同。这一现象称为表面弛豫(Surface Relaxation),而后续表面计算中的一个目标就是对该弛豫进行表征。

在对五层板块模型的表面进行 DFT 计算时,图 4.11 中给出了该模型在弛豫前后的原子位置。左侧是原始的板块模型,其中的原子按体相位置排列;而右侧是顶部三层原子发生弛豫后的板块模型。表面弛豫意味着弛豫后的表面要比初始表面(理想表面)的能量更低。在对能量进行最小化后(能量与超晶胞中原子位置具有函数关系),就能够得到弛豫后的表面几何构型。可以把板块下部的原子看作是材料的体相部分,并对这些底部原子层中的原子进行约束,使其固定于理想的体相位置上。这样一来,该计算就是与原子位置呈函数关系的超晶胞总能的最小化,其中仅允许顶部几层原子移动,正如第 3.5.2 节所述。这就产生了如图 4.11(b)所示的结构。实际上,表面中原子移动的量级仅为 0.1 \AA 。在该图中,夸张地表示了顶部原子层的弛豫量级。

在垂直于表面方向上,表面的突然终止会引发表面上原子配位关系的剧烈变化。这常常会导致第一层和第二层的原子层间距减小。最外层和次外层的距离可以表示为 d_{12} ,而其弛豫改变量表示为 δd_{12} 。弛豫改变量通常表示为体相材料层间距离的百分比,负值表示两层原子间距缩短,而正值表示两层原子间距扩展。在其他几个原子层之间的改变量表示为 δd_{23} 、 δd_{34} ,依次类推。表 4.1 列出了 Cu(100)和(111)表面的层间距离弛豫,其中采用不同的原子层数建立板块模型。在每个情形下,都将最下方的两层原子固定在体相位置,而允许其他所有原子层弛豫。

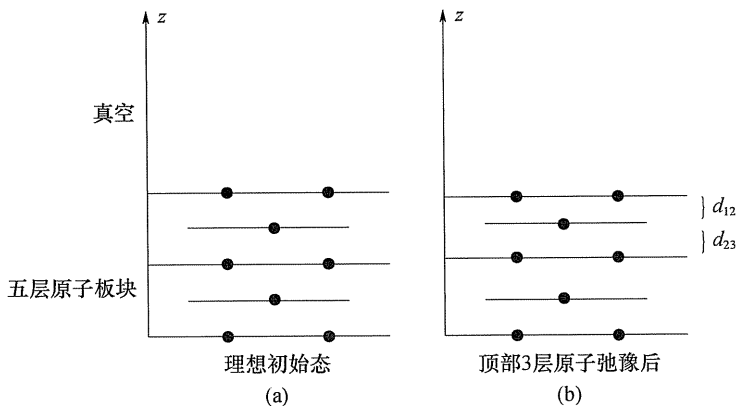


图 4.11 板块模型中表面原子弛豫的示意图(允许顶部三层原子弛豫,而底部的两层原子保持在理想的体相位置上)

表 4.1 使用 DFT 方法计算 Cu(100)和(111)表面层间弛豫与板块厚度的函数关系

表面层间弛豫 板块厚度	Cu(001)			Cu(111)		
	$\delta d_{12}(\%)$	$\delta d_{23}(\%)$	$\delta d_{34}(\%)$	$\delta d_{12}(\%)$	$\delta d_{23}(\%)$	$\delta d_{34}(\%)$
5 层	-3.84	-0.50	-0.53	-0.61	-0.08	+0.08
6 层	-1.93	+0.83	+0.37	-0.64	-0.11	+0.27
7 层	-2.30	+0.55	-0.25	-0.56	-0.04	-0.32
8 层	-2.14	+0.85	+0.00	-0.59	-0.32	+0.51
实验值	-2.0 ± 0.5^a	$+1.0 \pm 0.7^a$			-0.7 ± 0.5^b	

*a*摘自参考文献[1],测试温度为 305K;
*b*摘自参考文献[2]

Cu(100)表面层间距离的弛豫改变量更为显著,而在密排的Cu(111)表面上则稍弱。对于这两个表面,计算得到的层间距离弛豫改变量,均随着模型中所包含原子层数的增加而更加接近于实验结果。使用8层的板块模型,对于Cu(100)计算得到的弛豫改变量 δd_{12}

是 -2.14% , 或约为 0.036\AA 。这处于实验值 -2.0 ± 0.5 的误差范围内。

类似地, 对于 $\text{Cu}(111)$ 表面, DFT 估算的 δd_{12} 是 -0.59% , 或约为 0.012\AA ; 这一数值与实验值非常接近。考虑到这些位置改变量已经如此之小, 因此 DFT 能够很好地再现这些实验测试。

在表面弛豫的讨论中, 仅仅考虑了垂直于表面方向上的弛豫。这对上述计算的描述是正确的, 因为板块模型在表面的晶面内是完全对称的。因此, 在该表面的晶面内作用于每个原子上力的分量恰巧为零, 从而在弛豫过程中, 原子并不会在这个晶面内移动位置。对于真实表面而言, 这一点在物理上也是正确的; 而并不是仅仅存在于计算中的人为假象。如果将超晶胞中该晶面内的部分, 看作是这些类似于体相的原子层的复制, 那么就可以明确一点: 必须采用体相晶格间距来定义超晶胞, 正如等式(4.1)所定义的那样。这一讨论强调了表面计算的另一个尚未提及的特点, 即: 必须对所用晶格常数的数值进行选择。最好的方法是: 使用与表面计算相同的交换关联泛函, 先对体相材料进行精确收敛的 DFT 计算, 并采用该计算所得到的晶格常数。使用任何其他的晶格常数(例如, 实验测得的晶格常数)会导致材料模型中出现人为应力(Artificial Stress), 并会导致表面原子在物理上产生虚假弛豫(Spurious Relaxation)。

4.6 表面能的计算

如前文所述, 可以通过沿某个晶面劈裂体相材料来建立表面。表面能(σ)就是劈裂体相晶体所需的能量。表面能计算方法的思想 and 依据就是: 可以把这个切断过程所需的能量等效为新形成两个表面所需的能量(假设该过程是可逆过程)。这就意味着可以根据 DFT 板块计算来确定表面能, 即

$$\sigma = \frac{1}{A} [E_{\text{slab}} - nE_{\text{bulk}}] \quad (4.2)$$

式中: E_{slab} 为表面板块模型的总能量; E_{bulk} 为体相材料中单个原子或分子的能量; n 为板块模型中的原子或分子个数; A 为表面板块模型中的总

面积(包括顶面和底面)^①。

在宏观角度上,表面能通常表示为 J/m^2 这一单位。在 DFT 计算中,表面能更自然地定义为电子伏特/平方埃 ($\text{eV}/\text{\AA}^2$),而二者间可以方便地换算,即: $1\text{J}/\text{m}^2 = 16.02\text{eV}/\text{\AA}^2$ 。

式(4.2)所定义的表面能是两个量之间的差值,而这两个量的计算方式稍有不同。对于表面部分,通常使用相对较大的超晶胞,包含有一个真空区,并且使用相对较少的 k 点。对于体相部分,则恰好相反。那么,怎样做才能保证二者的理论计算差值不会影响表面能结果?遗憾的是,对于这个问题并没有唯一正确的方法;尽管经过一系列努力(相对于板块模型的原子层数、 k 点、截断能、超晶胞尺寸等因素进行测试,确保这两个能量中的每一个都具有良好的收敛性),可以把上述问题最小化。

例如,DFT 计算的铜表面能如表 4.2 所列,其中所使用的板块模型与表 4.1 相同。Cu(111)的表面能比 Cu(100)要低,意味着 Cu(111)要比 Cu(100)更稳定(或更加“类似于体相”)。这一点与第 4.4 节中的观点是一致的,即:简单材料的最稳定表面通常是那些原子密度最大的表面。通过将该表面能计算结果与一个实验观测值相比较,可以清楚地看到上述 DFT 计算能够给出该表面能较为精确的数值。

表 4.2 DFT 计算所得 Cu(100)和(111)

表面能与板块厚度的函数关系^a

单位: $\text{eV}/\text{\AA}(\text{J}/\text{m}^2)$

板块模型	$\sigma, \text{Cu}(001)$	$\sigma, \text{Cu}(111)$
5 层	0.094(1.50)	0.087(1.40)
6 层	0.097(1.55)	0.089(1.43)
7 层	0.098(1.57)	0.089(1.43)
8 层	0.096(1.53)	0.091(1.46)
实验值		0.114(1.83) ^b

^a底部的两层原子都约束在它们的体相位置,而允许其他原子层弛豫;
^b摘自参考文献[3]

① 该定义中忽略了表面能中熵的贡献。

在表 4.1 和表 4.2 的结果中,一个显著的区别就是:前者表明必须包含 8 层原子才能使弛豫改变量 δd_{12} 的结果收敛得很好;后者表明用 4 层原子的模型就可以得到表面能的合理估算值。这揭示了一个重要的思想:“收敛性”并不是一个可以同等地应用于所有物理特性的概念。它更有可能是指:经过一系列计算,对某个物理量已经达到“良好收敛”(给出了一个数值稳定的结果,且如果采用更高的计算精度,该结果并不会改变),但可能并没有达到另一个物理量的收敛范围。在原子层间距弛豫的计算中,仅为 0.02\AA 的层间距改变量就会导致 δd_{12} 为 1%。这意味着:对于确定表面模型中原子的精确几何构型而言,这个物理量是个非常敏感的指标。相反地,根据表 4.2 中的表面能结果,表明这个物理量(表面能)对表面层数的精确几何构型是较不敏感的。这些讨论对下面这个问题给出了一个范例:为什么最好是相对于一个特定物理量达到良好收敛状态,而不仅是泛泛所称的良好收敛。

4.7 对称和非对称的板块模型

先回顾一下图 4.1 中所显示的 FCC 材料 5 层板块模型,其中在该模型进行弛豫时,仅选择顶部的 3 层原子可以弛豫,而保持下部的两层原子固定在体相位置上。这个方法所定义的是一个非对称性板块模型;一侧的原子层可以弛豫,模拟表面;而另一侧的原子层保持固定,模拟体相区域。非对称性板块模型的一个重要特点是:该表面在随后的研究过程中会产生一个偶极子(Dipole)。如果想要研究原子或分子在表面上的吸附,这一点就尤为重要。例如,可以想象一下如果考察氟在一个表面上的吸附,这样一个电负性原子吸附后所产生的偶极子是很可观的。当将周期性边界条件应用于包含有这个非零偶极子的超晶胞时,这个偶极子与其上方和下方的其他影像之间形成静电作用,从而会对总能产生一部分贡献。在周期性体系中,产生这个静电作用是数学上的一个正确反映,但如果试图真正描述一个隔离态表面,这个静电作用其实在物理上是虚拟的(Spurious)。许多 DFT 计算程序都已经采用了相关方案,消除了对非零偶极子模型应用周期性边界条件所引发的

人为虚拟静电场。通常的做法是,能够计算得到偶极子,然后应用一个相等且相反的偶极子,从而校正真空区的局域电势。

另一个方法就是采用对称性模型来描述表面。在对称性模型中,板块的中心处存在一个镜面。中心处原子层的原子通常是固定在体相结构,而允许上方和下方的原子层弛豫。对称性模型的一个优点就是:表面上所产生的任何偶极子都会自动地抵消掉。然而,这样做也需要一定的成本,因为相比于非对称性板块,在对称性板块中通常需要包含更多的原子层数。图 4.12 中显示了一个具有 9 层原子的对称性板块,请注意在这个例子中,在板块的每一端都允许三层原子弛豫。而对比一下前文中非对称性板块模型,只在 5 层原子板块的一侧允许三层原子弛豫。因此,为了在计算中允许三层原子弛豫,就需要在对称性模型中包含有 9 层原子,而非对称模型只需要 5 层原子。

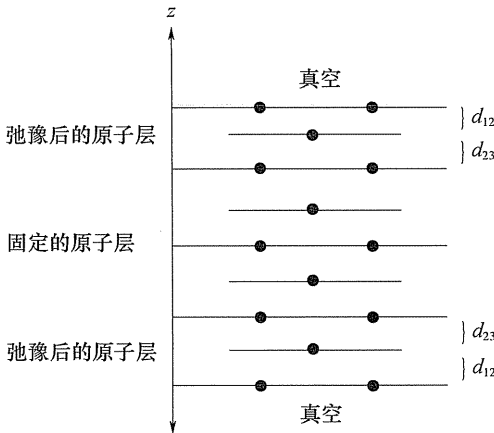


图 4.12 包含有 9 层原子的对称性表面板块模型示意图。其中,中间 3 层原子固定不动,而外侧的原子层可以弛豫,在板块底面和顶面上的原子层间距 (以及所有其他特性) 必须完全相同

4.8 表面重构

在本章中提到了由于存在表面弛豫,形成表面的原子与体相材料中的原子并不相同。然而,需要着重说明的是:如果仅仅允许表面进行

弛豫,也许并不能由 DFT 计算得到足够精确的信息。例如,在许多表面所经历的重构(Reconstruction)中,表面原子形成了新键^①。

表面重构的一个非常明显的例子是 Si(100) 表面。硅、碳和某些其他第 IV 主族材料的体相是金刚石结构。在这种结构中,每个原子都与它的 4 个最近邻原子以四面体排布成键。如果沿着(001) Miller 晶面劈裂硅,则在表面的硅原子具有两个最近邻原子,而非 4 个,如图 4.13 所示。图示的材料表面已经完成了严格定义,可以像之前对 Cu 表面所用的方法那样,对其表面原子进行 DFT 弛豫。图 4.13 中的图像正是由这种计算所得到的。

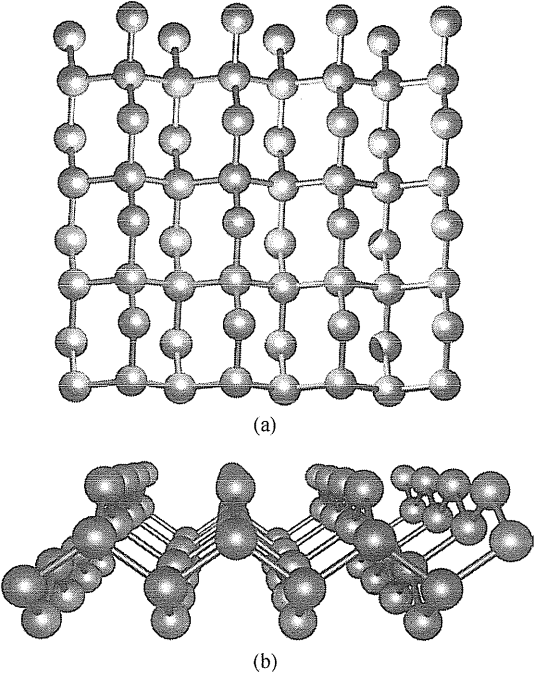


图 4.13 以 Si 的体相结构为封端的未发生重构 Si(001) 的(a)俯视图和(b)侧视图

^① 另一个描述表面重构的方法是:在重构过程中,相对于(弛豫后的)体相封端表面的对称性,表面原子的对称性发生了某种改变。

从化学的观点来看, Si(001) 的体相封端并不完全合理, 因为每个表面硅原子都有对应于一个未配对电子的两个悬空键。可以想象得到, 表面原子应该倾向于向着某个位置移动, 从而使未配对电子与其他未配对电子相互配对形成新键, 尽管这样做会使表面上现有的原子键产生新的应变, 从而耗费一些能量。这个例子中, 连接悬空键所得到的能量相当大, 从而真实而自然地发生这一现象。表面上的原子通过配对进入到在能量上更为稳定的状态。图 4.14 显示了重构后的 Si(100)

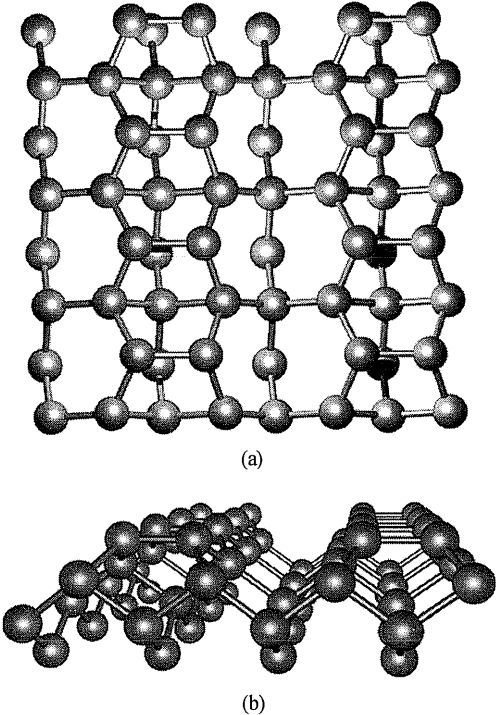


图 4.14 以二聚体排列重构的 Si(001) 的 (a) 俯视图和 (b) 侧视图
(图中所用坐标取自于完全弛豫的 DFT 计算)

表面的俯视图和侧视图,其中相邻表面原子彼此移动得更加接近,并形成新键。图 4.13 和图 4.14 均为完全弛豫后的 DFT 结构,两者之间的能量差值是约 0.7eV ,这个能量值是相当大的!在表面上形成长排的配对原子称为二聚体排列(Dimer Rows),这些独特的排列在 Si(001)的 STM 图像中也可以清楚地观察到。

一旦明白了表面重构在自然界中确实存在,则由该论断所引出的关键点就是:如果在计算中使用 Si 的体相封端作为初始构型,则图 4.14 所示的 Si(001) 重构并不会自发地出现。因为用优化程序的计算语句来描述的话,重构的表面所定义的是全局性的表面能量最低点,而体相封端表面会产生一个不同的局域性能量最小值;无法用一组 DFT 计算“证实”所研究的表面没有发生表面重构。这一情形类似于第 2.4 节中“预测”晶体结构的讨论。

接下来,作为一个警诫性的例子,再来介绍另一个 Si 表面,即(111)表面。使用与上文相同的化学推演方法,不难确信这个表面也会以某种方式发生重构。如果着眼于“预测”该重构的细节信息,可以将表面上的悬空键以各种可能的组合方式连接在一起,进行一系列的计算。在经过许多繁复的这类计算之后,哪一种结构中每个表面原子所具有的平均能量最低,这种结构就是所得到的“预测”结果。可以先对 Si(111)的单个单胞(Single Unit Cell)进行计算,如果想要做得更多更好,可以尝试使用更多的单胞进行计算。实际上,在 7×7 个单胞中(参见图 4.15),Si(111)的稳态重构包含着错综复杂的原子排列。为了表明该结构是正确的,DFT 计算具有重要的作用。但是,这也仅是在收集到大量实验数据的基础上,并且需要用实验数据来指导计算。

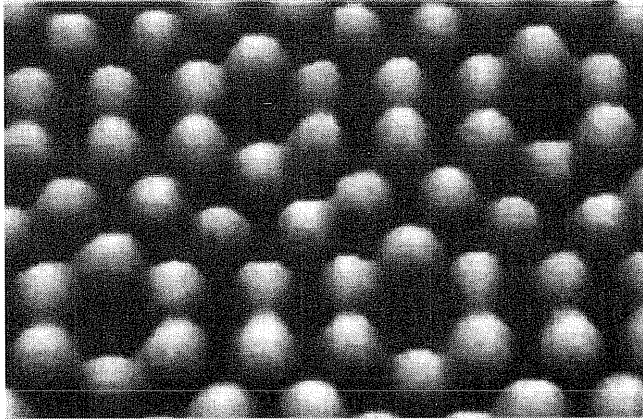
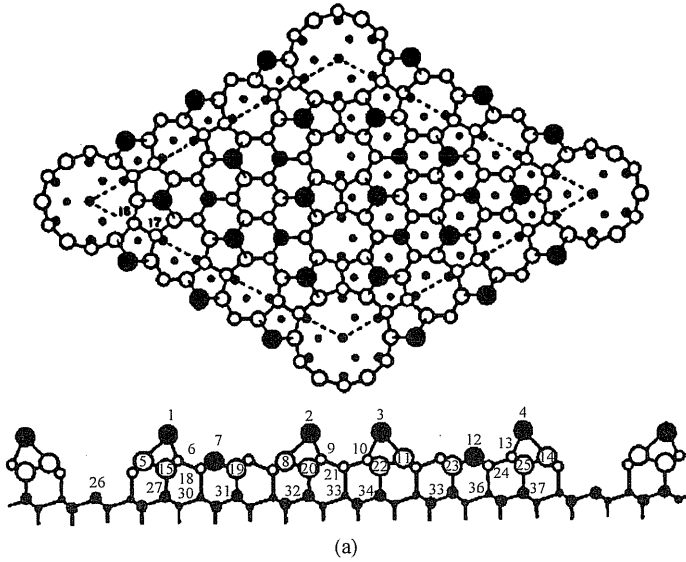


图 4.15 Si(111)7×7 重构图与 STM 图像

(a) 在 Si(111) 上的 Si(111)7×7 重构图; (b) 重构的 STM 图像。[重印许可来自于 Brommer K D, Needels M, Larson B E 和 Joannopoulos J D, *Ab Initio Theory of the Si(111) (7-7) Surface Reconstruction: A Challenge for Massively Parallel Computation*, *Phys. Rev. Lett.* 68 (1992), 1355 (Copyright 1992 by the American Physical Society).]

4.9 表面上的吸附质

截至目前,本章只对洁净表面进行了讨论。也就是说,表面的化学成分与构成该表面的体相物质完全相同。许多有意思的事情都会发生在表面上,当然会发生其他化学成分出现在这些表面上,现在就对这种情形进行讨论。本节只讨论原子或分子化学吸附于表面上这类情形——这可以并不严格地定义为:在原子或分子与表面之间存在某种化学键。

例如,可以运行一些计算来理解 H 原子是怎样结合在 Cu(100)表面上的,这个金属表面早在本章最开始就已经提到了。许多有关金属的化学反应,都可以形成单个的 H 原子(而不是其他化合物)。因此,为了对这类反应进行描述,就必须了解这些原子在表面上的特点。或许我们所应该问的最简单问题是:在表面上 H 到底倾向于处于何处? 如果对 Cu(100)表面来讨论这个问题,则存在几个具有特殊对称性的位置,直观上看来,这些位置就很可能成为 H 的潜在结合位置(这些位置如图 4.16 所示)。

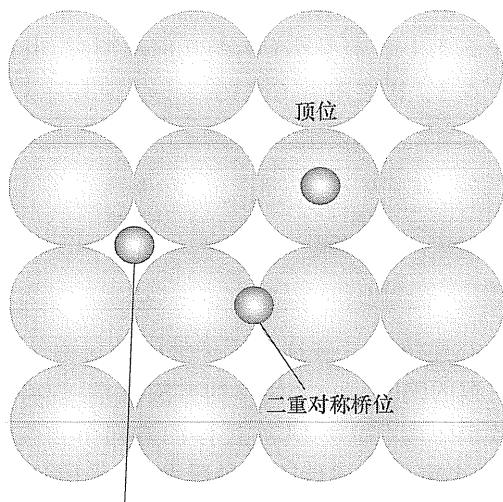


图 4.16 H(浅色小球)在 Cu(100)上可能的高对称结合位置示意图
Cu(100)表面最顶部的原子层显示为较大的深色球

读者或许能够猜到应该如何用计算来描述 H 在图 4.16 中的每个位置。先定义一个超晶胞包含有 Cu(100)的板块模型,将一个 H 原子放置在相对接近于表面顶层的某个图示位置处,然后最小化体系的能量,并允许几个原子层和 H 原子的位置发生弛豫。对类似于这样的一组计算,可以简单地比较其总能,因为它们总能的差值也就相当于在表面不同位置移动 H 原子消耗(或得到)的能量。表 4.3 给出了这组计算的结果。这些计算预测了 H 在四重空穴位更加稳定,但是,二重桥位的能量仅高了 0.08eV。因为在室温下热能的典型单位是 $k_b T \sim 0.03\text{eV}$,这些结果表明:至少在某些情形下,H 可能在这两种位置上都存在。

此外,对于这些结果,有必要验证一下根据该体系实际上的局域能量最低值所得到的原子构型。为了说明这一点,以图 4.16 中顶位(On-Top Site)上的力矩为例。对于正好位于表面 Cu 原子上方的 H 原子,根据对称性,在平行于金属表面的平面内,作用在这个 H 原子上的力必定为零。这意味着,在优化计算过程中,H 原子将会始终位于 Cu 原子的正上方,虽然是将其从这个对称位置移开后有可能降低其能量。事实上,对于图 4.16 所示的全部三种高对称性位置,这一论断都完全适用。在第 3.5.1 节中考察 CO₂的几何构型时,也存在类似情形。

表 4.3 H 吸附在 Cu(100)上的计算结果^a

H 的初始位置	相对于空穴位的吸附能	相对于 H ₂ (g)的吸附能	是否局域最小值?
四重空穴位	0	-0.19	是
二重桥位	0.08	-0.11	否
顶位	0.57	+0.38	否

^a 所有能量单位是电子伏特(eV)

可以通过小心谨慎地打破这些原子坐标的对称性,从而避免这类由于对称性而引致的陷阱。一个简单的方法就是:先将 H 原子沿任意方向移动一个微小距离(如 0.2Å),且移动方向不再与表面上的对称方向相一致,然后再重复上述计算。最后发现,采用这种计算方法,如果 H 原子初始位置是在上述任一高对称性位置偏离了 0.2Å,则在弛豫后 H 原子都会位于四重孔穴位置,即便是其初始位置更靠近顶位或桥位。这表明顶位和桥位并不是该体系的能量最低点。

到此为止,可以由计算得到在 Cu(100)表面上 H 原子所倾向处于的位置。另一个简单但很重要的问题是:在这个表面上,相比于其他一些位置,这些原子处于优先位置上的倾向到底有多强烈?通常,可以通过计算物质在表面上的吸附能来回答这个问题。对于吸附能,可以对其定义为:

$$E_{\text{ads}}^{\text{atomic}} = E_{\text{H/surf}} - E_{\text{H(g)}} - E_{\text{surf}} \quad (4.3)$$

这里,右侧的三项分别是表面吸附了 H 之后的总能、在气相中单个 H 原子自身的总能以及洁净表面的总能。简而言之,这个物理量是将 H 原子从表面上拉脱开来,并使其进入气相这一过程所需的能量。这个定义简单易用,但在化学上是不符合自然规律的,因为 H 原子几乎不可能长时间存在。一个更具物理含义的量就是:将 H 原子从表面上拉脱开来并形成气相 H₂分子的能量这一过程所增加(或减少)的能量,即

$$E_{\text{ads}} = E_{\text{H/surf}} - \frac{1}{2}E_{\text{H}_2(\text{g})} - E_{\text{surf}} \quad (4.4)$$

这个吸附能与式(4.3)所定义的数值是完全不同的,因为在 H₂分子中存在相当大的键能。H 在 Cu(100)表面上的吸附能,以及根据式(4.4)所定义的吸附能分别列于表 4.3 的左侧和右侧,前者大致相当于后者的一半。负值表明:相比于洁净表面和一半 H₂分子,吸附态的总能更低。可以观察到,四重对称位置在热力学上更加优先。尽管表面上的另外两种吸附位置并不是 H 原子的局域最小值,但仍然能使用式(4.4)描述它们的能量。桥位的结果是负值,但也仅仅只是如此而已。顶位的结果较大且为正值,这意味着:即使这个位置是局域最小值,当该表面与气相 H₂达到平衡状态后,能够在这个位置上观察到的 H 原子数量也是极其稀少的。

由上述概念自然地联想到,可以用这个方法预测分子在表面上的吸附。如果想要描述氢氧根(OH)在 Cu(100)表面的吸附,需要对该物质在表面上一系列可能的结合位置,分别单独地进行许多计算。与上述原子态 H 的计算不同,对氢氧根的计算不仅需要考虑物质在表面的结合位置,也需要考虑 O-H 键相对于表面的位向关系。与原子吸附相类似,对于简单物质(如 OH)的计算,常见的失误之处就是在计算中仅仅考虑了高对称性构型,例如 O-H 键垂直于表面的构型。随着

吸附于表面的分子尺寸增大,所有可能的分子吸附方式的数量就会急速增大。在许多情形下,分子都含有较多的原子,如果不参考实验数据,就很难完全确定最优吸附构型。

4.9.1 吸附能的精度

对于 H 在 Cu(100) 上的结合,已经采用两种方式进行了表征,也就是分别相对于形成气相 H_2 的吸附能(表 4.3)以及不同表面状态之间的相对吸附能。在这两类物理量之间,使用 DFT 计算所能得到的精确性是差别很大的。第 1 章中已经说到,由于交换关联泛函本身是一个近似,相对于 Schrödinger 方程真实解,DFT 计算结果存在系统(但是未知)的误差。然而,非常大量的计算显示:如果在两个 DFT 计算之间比较“化学近似”态,则在这两个计算中,二者的系统误差也是近似的。换句话说,使用 DFT 方法,相比于绝对能量,可以更加精确地计算得到化学近似态的相对能量(两个近似态的能量差值)。

对这个普适原理,表面吸附就可以给出一个很好的实例。可以计算 H 在四重和二重对称位置之间的相对能量,由于二者是化学近似的,使用相对能量对其进行描述肯定是合理的。因此,能够预期到这个相对能肯定具有更高的数据精度。对于吸附于金属表面的一个 H 原子,以及气相 H_2 分子中的一个原子,对这两者进行比较,则不能认为它们是化学近似的。因此,对于使用 DFT 方法计算等式(4.4)所定义吸附能中的误差,应该要比不同表面状态之间相对能的误差更大。在第 10 章有关 DFT 计算精度更加详细的讨论中,也会对这一点再次说明。

4.10 表面覆盖效应

在表 4.3 中,列出了计算得到的 H 在 Cu(100) 表面的吸附能(相对于气相 H_2 分子)。然而,当给出这个数值时,我们忽略了一个对吸附能具有潜在重要贡献的因素,这就是表面覆盖效应。因为使用了周期性边界条件,将一个吸附质放置在超晶胞中,自然也就意味着:每个吸附质都能在与它相邻的每个超晶胞内“观察到”一个它的复制品。这

一现象的本质可以归结为:在建立带有一个吸附质的表面超晶胞时,实际上必须描述为一个周期性的、规则性重复的表面覆盖层。

这里存在有两个问题。首先,在超晶胞计算中的吸附质一定具有长程规则排布样式,该样式是与超晶胞一样不断重复的。在周期性边界条件下,不可能建立吸附质的任何一种真实的无序排列模型。幸运的是,在自然界中,晶体表面上的吸附质大多是长程有序的,特别是在较低的温度下,因此能够用 DFT 计算模拟许多情形下的真实体系。

第二个问题是,超晶胞的尺寸决定了吸附质之间的距离。如果超晶胞较小,则定义了一个具有高密度(或高覆盖率)吸附质的表面。如果超晶胞较大,则定义了一个具有较低密度吸附质的表面。当每个表面原子都存在一个吸附质时,则称该吸附层的覆盖率为 1 单分子(原子)层。如果平均两个表面原子存在一个吸附质,则覆盖率为 0.5ML,依次类推^①。

不难想到,对于覆盖层的对称性,已经形成了一套命名体系。这里并不介绍所有这些命名背后的理论规则,只是给出一些具有代表性的范例。

在图 4.17 中,给出了几个 H 吸附在 Cu(100)上的例子。先从其中的三个示例开始介绍:从图 4.17(a)到(b)再到(d),覆盖率从 1ML 降低到 0.5~0.125ML,这三个覆盖层的命名分别是(1×1)、c(2×2)和 c(4×4)。观察一下图中这些超晶胞的边长,并比较一下沿着每个边长方向吸附质的数量,可见以这种方式命名这些晶胞似乎是合乎逻辑的。这些命名名称中的字母“c”表示“中心的”,意思是在这个超晶胞中心处,以及边角处各有一个吸附质原子。在本例中 H 原子位于孔穴位,但需要说明的是,它们也可以全部放在顶位或桥位,或其他位置,而覆盖层仍然以完全相同的方式来命名。比较图 4.17(b)和(c),可以看到二者均描述了 c(2×2)覆盖层,二者间的区别就是图 4.17(c)中的超晶胞较小,因而所需要的计算成本较小;给出图 4.17(b)中的超晶胞就是为了表明:为什么对这种覆盖层使用 c(2×2)这个名称。类似的关系也存在于图 4.17(d)和(e)之间。

^① 对于表面上不止一种表面原子时,例如合金或化合物的表面,就不能采用该术语清楚地进行定义。

现在回顾一下这个问题:不同的覆盖率是否会使计算得到的吸附能有所不同。分别针对几个不同的覆盖率,使用 DFT 计算得到 H 原子在 Cu(100)上孔穴位的吸附能,如表 4.4 所列。可见,在吸附 H 原子之间的空间越大,则吸附能将有所降低。对于 $c(4 \times 4)$ 覆盖层,是图示模型中覆盖率最低的,它的吸附能也最低,因此是最优先的。 (1×1) 覆盖层中 H 原子占据着(100)表面上的每个孔穴位,相对于气相 H_2 ,在热力学上是不优先的。对于大多数计算,不能够忽略相邻吸附质的影响作用。这个简单例子指出:描述一个原子或分子在表面上的吸附能,重要的是先确定覆盖率和吸附质的对称性,从而最终完全定义所需考虑的体系。

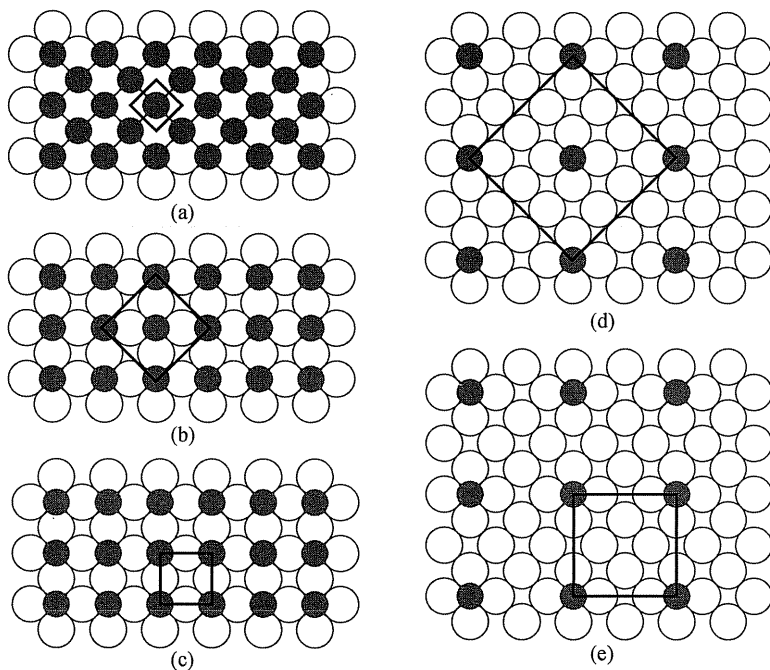


图 4.17 不同覆盖率的 H 吸附于 Cu(100)上的示意图

(a)覆盖率为 1ML 的 $c(1 \times 1)$ 序列;(b)覆盖率为 0.5ML 的 $c(2 \times 2)$ 序列;(c)与 (b)相同的覆盖层,但所构建的超晶胞较小;(d)覆盖率为 0.125ML 的 $c(4 \times 4)$ 序列;(e)与(d)相同的覆盖层,但所构建的超晶胞较小。(Cu 原子为空白圆圈,吸附 H 原子为灰黑色圆圈,黑色线所表示的就是超晶胞)

表 4.4 H 在 Cu(100) 上 3 种不同覆盖层的吸附能

	(1 × 1)	c(2 × 2)	c(4 × 4)
H 覆盖率	1.00ML	0.50ML	0.125ML
相对于 $\frac{1}{2}\text{H}_2$ 的 $E_{(\text{ads})}$	0.08	0.11	0.19

图 4.18 给出了在 FCC(111) 表面上的一些覆盖层范例。这些覆盖层名称中的字母“p”表示覆盖层单胞是“原胞(Primitive)”。需要注意,对于密排(111)表面上完全覆盖的单原子(分子)层覆盖率,与较为开放的(100)表面相比,在能量上是较不优先的,因为吸附质会被迫非常靠近。

练 习

在所有这些练习中,请仔细思考与表面原子层数、 k 空间取样等方面有关的数值收敛性。当前几个练习中获得经验后,如果能够清楚地解释前几个练习中是如何得到收敛性信息的(这些信息与后面的练习也是相关的),则在后面的计算中,就不需要进行过多的收敛性测试。

1. 针对 FCC 金属的(100)、(110)和(111)表面,建立适合于进行计算的超晶胞。为了考察每种表面的表面弛豫,每个表面单胞需要多大的尺寸?

2. 将练习 1 扩展为计算 Pt(100)、Pt(110)和 Pt(111)的表面能。

3. 从实验上已知 Pt(110)会重构形成所谓的“Missing - Row Reconstruction”,在这种重构中,来自于(2 × 1)表面单胞中顶层的交替行(Alternate Rows)消失了。使用 Pt(110)的(2 × 1)表面单胞定义一个超晶胞,计算其未重构的和重构之后的表面能。在比较时,为什么是基于二者的表面能来比较,而不是简单地比较两个超晶胞的总能? 所得计算结果是否与实验观测是一致的? 采用类似的计算预测对于 Cu(110)是否存在类似的重构?

4. 运行计算,确定原子态 O 在 Pt(111)表面上较为优先的结合位

置。其中,采用覆盖率为 0.25 和 0.33ML 的有序化覆盖层。随着表面覆盖率增大,吸附能是增加还是减小?

5. 类似于练习 4 运行计算,考察氢氧根基团(OH)在 Pt(111)上的吸附。在其优先的吸附构型中,OH 键与表面法线之间会形成怎样的倾斜角度? 如果对所有可能的倾斜角度进行充分计算,则这些计算能够提供什么数值论据?

6. 由于水分子在 Pt 表面上分解,才会使得 OH 基团出现在该表面上。为了估算这个过程在能量上是否具有正向发生倾向,计算洁净 Pt(111)表面和气相 H₂O 分子的能量,以及 H 原子、OH 基团共同吸附于 Pt(111)表面的能量,并比较这两个能量。

参 考 文 献

- [1] Fowler D E, Barth J V. Structure and Dynamics of the Cu(001) surface Investigated by Medium - Energy Ion Scattering. Phys. Rev. B 52,1995;2117.
- [2] Lindgren S Å, Walldén L, Rundgren J, et al. Low - Energy Electron Diffraction from Cu (111): Subthreshold Effect and Energy - Dependent Inner Potential; Surface Relaxation and Metric Distances between Spectra. Phys. Rev. B 29,1984;576.
- [3] Boer F R de Boom R, W C M Mattens, et al. Cohesion in Metals. North - Holland, Amsterdam, 1988.

扩 展 阅 读

以下两本书给出了晶体表面标注方法的详细信息和物理现象: Hudson J B. Surface Science: An Introduction. Butterworth - Heinemann, Boston, 1992, 和 Somorjai G A. Introduction to Surface Chemistry and Catalysis. Wiley, New York, 1994.

与表面过程有关的 DFT 计算,已经有大量的相关文献,这里列出一些具有历史意义和代表性的例子。对于 Si(111)重构的详细信息,参见: Brommer K D, Needels M, Larson B E, et al. Phys. Rev. Lett. 68. 1992:1355.

关于将 DFT 计算和实验相结合,用以理解 TiO_2 表面与氧相互作用的文献,参见:Henderson M A, Epling W S, Perkins C L, et al. *J. Phys. Chem.* B103. 1999:5328, 及其参考文献。有关于使用 DFT 方法考察材料对改善和控制汽车污染的文献是:Schneider W F, Li J, Hass K C. *J. Phys. Chem.* B105. 2001:6972.

对于表面能计算,以及 LDA、GGA、几个“后 DFT(post - DFT)”方法精度的讨论,参见:Alfé D, Gillan M J. *J. Phys. : Condens. Matter*18. 2006:L435.

对于在表面上考察具有许多自由度的分子这一挑战性工作的文献,参见:Mae K, Morikawa Y. *Surf. Sci.* 553. 2004:L63.

对于表面的 DFT 计算中有关讨论偶极子校正的文献,参见:Neugebauer J, Scheffler M. *Phys. Rev.* B46. 1992:16067.

附录 计算细节

本章所有的计算都使用 PBE GGA 泛函。对于与 Cu 表面有关的计算,截断能为 380eV,且 Methfessel - Paxton 方案采用的 Smearing Width 为 0.1eV。对于与 Si 表面有关的计算,截断能为 380eV,且 Gaussian Smearing 为 0.1eV。在倒易空间中使用 Monkhorst - Pack 设定 k 点。对于所有表面计算,使用 DFT 优化后的晶格参数来定义超晶胞在表面晶面内的尺寸。

4.5 节 采用非对称板块模型考察了表面弛豫,模型原子层厚度为 5、6、7 或 8 层,且底部的两层原子固定在体相位置上,而其他原子均可以自由弛豫。对 Cu(100),超晶胞具有 $c(2 \times 2)$ 表面对称性,每层包含 2 个原子。对于 Cu(111),采用 $(\sqrt{3} \times \sqrt{3})R30^\circ$ 表面单胞,且每层有 3 个原子。所有模型在沿着表面法线方向上,都包含最小为 23Å 的真空。所有计算中使用的 k 点网格为 $6 \times 6 \times 1$ 。

4.6 节 在第 4.5 节中的表面能计算,需要用到体相 Cu 的能量计算。该体相计算是采用优化后的晶格参数建立 FCC 原胞,并以 $11 \times 11 \times 11$ 的 k 点网格计算的。

4.8 节 采用 4 层的 $c(4 \times 4)$ 非对称性板块模型,对 Si(001) 表面重构进行模拟。底部的两层原子固定在体相位置。倒易空间的取样是 $11 \times 11 \times 11$ 的 k 点网格,真空层为 17 \AA 。

4.9 节 采用 4 层的非对称性板块模型,考察了 Cu(100) $c(4 \times 4)$ - H 覆盖层,其中底部的两层原子固定在体相位置,而 k 点网格为 $6 \times 6 \times 1$ 。

4.10 节 在考察不同覆盖度下, H 在 Cu(100) 上吸附的计算中,所使用的是 4 层非对称的 $c(1 \times 1)$ 、 $c(2 \times 2)$ 和 $c(4 \times 4)$ 板块模型,其 k 点网格分别为 $16 \times 16 \times 1$ 、 $12 \times 12 \times 1$ 和 $6 \times 6 \times 1$ 。H₂ 的计算是在一个边长为 20 \AA 的盒子中进行的,优化后的 H₂ 键长为 0.75 \AA 。

第 5 章 振动频率的计算

在前面的章节中,已经学习了如何使用 DFT 计算优化分子、体相固体和表面结构。在许多方面,这些计算都可以提供令人满意的计算结果,因此,能够通过这些计算来预测非常广泛的材料性质。但是,读者目前所学习过的内容,似乎都证实了一个直接针对 DFT 计算的常见批判意见:该方法中的温度为绝对零度。这意味着:在这些计算所给出的材料性质中,原子被局限在“平衡态”或“最低能量”的位置上,在经典力学上,这就对应于在 0 K 下的材料性质^①。这个批判意见暗示着:了解一些材料在 0 K 下的信息也许是一件有意思的事情,但是真实情形全都发生在较高的温度下。

后续的大部分章节都旨在使读者明白:DFT 计算怎样才能对非零温度下的材料给出有用信息。首先,在经典力学的描述中,将材料降温至 0K 时材料中的原子会发生弛豫,从而使材料能量最小化,因此可以将能量最小状态下的原子坐标称为平衡位置。如果将材料升温到一个非零温度,会发生一个最简单的事情(再次从经典力学的角度来看):材料中的原子会在它们的平衡位置周围振动。从更加精确的量子力学观点来看,即便是在 0 K 下,原子在平衡位置周围的振动也会表现为零点能的形式,从而对材料的能量有所贡献。在多数情形下,可以使用光谱学对这些振动进行实验测量,而且人们通常更感兴趣的是这些振动的频率。在本章中,主要介绍如何使用 DFT 方法来计算振动频率。

^① 这在经典力学中是正确的,但在量子力学中,更准确的说法是:原子被局限在它们的能量最小值处。当在第 5.4 节中讨论零点能(Zero Point Energy)时,还会再次说明这个概念。

5.1 隔离态分子

先从一个简单的例子开始,求解一个隔离态 CO 分子的振动频率。更确切地说,是对分子中两个原子之间的化学键进行伸缩 (Stretching)。为方便起见,假设这个化学键在空间中是沿着 x 方向的,则键长可以定义为 $b = x_c - x_o$,其中 x_c 和 x_o 是两个原子的位置。对于分子的能量,可以用泰勒展开式在平衡键长 (b_0) 周围展开为

$$E = E_0 + (b - b_0) \left[\frac{dE}{db} \right]_{b=b_0} + \frac{1}{2} (b - b_0)^2 \left[\frac{d^2E}{db^2} \right]_{b=b_0} + \dots \quad (5.1)$$

在展开式中,因为是在能量最低点上,所以一阶导数恰好为零。因此,对于平衡键长 b_0 很小的位移而言, $E = E_0 + \alpha/2 (b - b_0)^2$,其中 $\alpha = [d^2E/db^2]_{b=b_0}$ 。这个方法忽略了泰勒展开式中的高阶项,称为谐函数近似 (Harmonic Approximation)。

在这种近似中,怎样才能使原子位置移动呢? 将 C 原子核看作是服从牛顿定律的经典粒子,则 $F_c = ma_c$,其中 $F_c = -\partial E/x_c$, $a = d^2x_c/dt^2$ 。可以对 O 原子核写出类似的表达式。根据代数关系,整体键长的运动方程为

$$\frac{d^2b(t)}{dt^2} = -\alpha \left(\frac{m_c + m_o}{m_c m_o} \right) (b(t) - b_0) \quad (5.2)$$

该方程的解是 $b(t) = b_0 + a \cos \omega t$,其中 a 是任意常数,且

$$\omega = \sqrt{\alpha \frac{m_c + m_o}{m_c m_o}}$$

这表明键长振动时的振动频率为

$$\nu = \frac{\omega}{2\pi} = \frac{1}{2\pi} \sqrt{\alpha \frac{m_c + m_o}{m_c m_o}}$$

为了使用 DFT 方法计算 CO 的振动频率,首先必须找到分子能量最小化时的键长,所需计算的其他信息就只是 $\alpha = [d^2E/db^2]_{b=b_0}$ 。然而,对于与原子位置有关的能量二阶导数解析表达式,平面波 DFT 方

法通常并不能对其进行估算。但是,可以使用有限差分近似得到二阶导数令人满意的估算值,即

$$\left(\frac{d^2E}{db^2}\right)_{b_0} \approx \frac{E(b_0 + \delta b) - 2E(b_0) + E(b_0 - \delta b)}{(\delta b)^2} \quad (5.3)$$

这个表达式在 $\delta b \rightarrow 0$ 时是完全精确的,读者可以用式(5.1)泰勒表达式对右侧展开进行验证。

综上所述,可以使用 DFT 计算方法,利用有限差分近似,估算气相 CO 分子在伸缩振型下的频率。根据式(5.3),采用一系列不同的有限差分位移(δb)进行计算,结果图 5.1 所示和如表 5.1。相对来说,计算所得振动频率对于位移区间(如 $\delta b = 0.005 \sim 0.04 \text{ \AA}$)是不敏感的。然而,对于更大和更小的位移,情形就发生了巨大变化。理解这种行为的本质是最为重要的。对于较大的 δb ,计算的不准确性主要是由式(5.3)中的导数项所引发的。这个表达式只有当 $\delta b \rightarrow 0$ 时才是完全精确的,因此当 δb 较大时得到了不精确的结果,也就并不奇怪了。

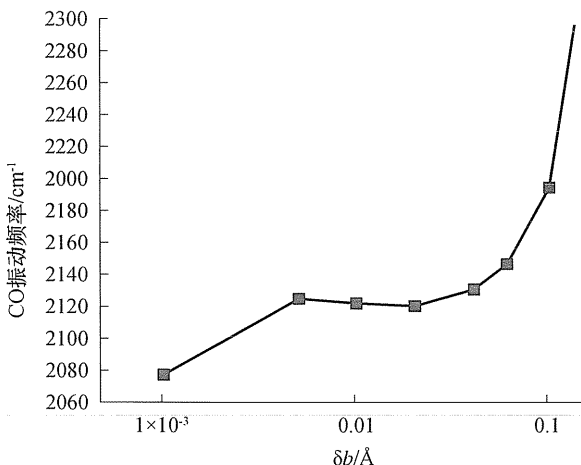


图 5.1 采用 DFT 方法计算的气相 CO 振动频率与有限差分位移(δb , 单位为埃)之间的函数关系图

图中文字: CO vibrational frequency; CO 振动频率。

表 5.1 在若干不同的有限差分位移下,采用 DFT 计算得到的气相 CO 振动频率

$\delta b/\text{\AA}$	振动频率/ cm^{-1}
0.001	2078
0.005	2126
0.010	2123
0.020	2122
0.040	2132
0.060	2148
0.100	2197
0.200	2433
0.500	4176

如表 5.1 所列,当 $\delta b = 0.001 \text{\AA}$ 时振动频率的不准确性似乎更加令人惊讶,因为更小的 δb 取值本应该会使式(5.3)向着精确值的极限方向移动。但是,只有当我们能够精确估算所使用的能量值时,该表达式才是精确的。为此,可以把式(5.3)改写为

$$\left(\frac{d^2 E}{db^2}\right)_{b_0} \approx \frac{\Delta E(\delta b) + \Delta E(-\delta b)}{(\delta b)^2} \quad (5.4)$$

式中: $\Delta E(x) = E(b_0 + x) - E(b_0)$ 。但是,对于极小的 δb 数值,式(5.4)中的能量差就变得非常小,也就无法用 DFT 方法进行精确地估算。例如,对于 CO 分子,从其平衡位置下将其键长改变 0.001\AA ,则分子能量的变化大约只是 0.00002eV 。如果回顾一下第 3 章中有关收敛性的详细讨论,就会明白:想要在计算中对这样小的能量差进行精确度量,将是极富挑战性的。

CO 分子的例子表明:在对振动频率进行有限差分计算时,所使用的位移既不能太大,也不能太小。实际操作中,较好的方法是选择使能量差处在 $0.01 \sim 0.10 \text{eV}$ 量级上的位移,因为不需要超乎寻常的谨慎,就可以精确地计算这些能量差。对于大多数材料,这些能量差对应于有限差分位移的范围是 $0.03 \sim 0.1 \text{\AA}$ 。不管使用哪种 DFT 计算方法,当

开始计算一种新材料的振动特性时,先进行一系列计算,验证数值结果与 δb 函数关系的相容性,是非常明智的做法。

既然已经理解了如何使用 DFT 方法得到 CO 收敛良好的振动频率,就可以将这个结果与实验数据相互比较。实验上得到气相 CO 的伸缩频率(Stretching Frequency)为 2143cm^{-1} ,比 DFT 结果高了大约 20cm^{-1} 。DFT 方法所计算的振动频率具有较宽范围,这样的计算结果是非常典型的。关于 DFT 计算结果与真实振动频率之间的差值,一方面是由于对振动进行了谐函数处理,另一方面是由于 DFT 方法求解 Schrödinger 方程时并非本质上的严格求解。在第 10 章对 DFT 方法精确性更为广泛的讨论中,会再次谈及这个问题。

5.2 原子集合的振动

在第 5.1 节的计算中,只说明了有关振动的一个最简单例子,也就是隔离态双原子分子的振动。如果分子与其他原子存在相互作用,则其振动又会产生什么变化呢?例如,如果 CO 分子已经在一个金属表面上产生了化学成键,为了描述这种情形,可以把上述振动频率的概念引申扩展到一种更为广泛的情形,即:多原子的集合也会发生振动。

首先,定义一个 N 原子体系,并将其笛卡儿坐标写作一个含有 $3N$ 个元素的矢量 $\mathbf{r} = (r_1, \dots, r_{3N})$ 。如果处在 \mathbf{r}_0 时是这些原子能量的最低点,那么可以方便地定义一个新的坐标 $\mathbf{x} = \mathbf{r} - \mathbf{r}_0$ 。这些原子能量关于 \mathbf{r}_0 的泰勒二阶展开式为

$$E = E_0 + \frac{1}{2} \sum_{i=1}^{3N} \sum_{j=1}^{3N} \left[\frac{\partial^2 E}{\partial x_i \partial x_j} \right]_{\mathbf{x}=0} x_i x_j \quad (5.5)$$

上述表达式中的一阶导数项均为零,其原因与式(5.1)的原因相同。如果继续定义

$$H_{ij} = \left[\frac{\partial^2 E}{\partial x_i \partial x_j} \right]_{\mathbf{x}=0} \quad (5.6)$$

这些倒数所定义的 $3N \times 3N$ 矩阵就是海塞矩阵。按照对 CO 分子在一维空间中所做的那样,就可以得到关于这些原子的一个经典力学运动

方程,其中与第 i 个坐标有关的力记作 F_i 。这个力对应于这个坐标的加速度为 $F_i = m_i(d^2x_i/dt^2)$,其中 m_i 是第 i 个坐标的原子质量。这个力同时与 N 个原子的能量有关,即 $F_i = -\partial E/\partial x_i$ 。如果用矩阵形式来表示,上述分析所得到的运动方程就是

$$\frac{d^2\mathbf{x}}{dt^2} = \mathbf{A}\mathbf{x} \quad (5.7)$$

式中:矩阵 \mathbf{A} 的元素为 $A_{ij} = H_{ij}/m_i$ 。这个矩阵称为质量加权的海塞矩阵(Mass - Weighted Hessian Matrix)。

经过上述推导得到的运动方程具有对应于矩阵 \mathbf{A} 本征矢量的一组特殊解。该矩阵的本征矢量就是与其本征值 λ 满足 $\mathbf{A}\mathbf{e} = \lambda\mathbf{e}$ 的矢量 \mathbf{e} 。如果全体原子运动的初始条件是 $\mathbf{x}(t=0) = a\mathbf{e}$ (其中 a 为任意常数)和 $d\mathbf{x}(t=0)/dt = 0$,则原子坐标就可以简单表达为 $\mathbf{x}(t) = a\cos(\omega t)\mathbf{e}$,其中 $\omega = \sqrt{\lambda}$ 。也就是说,如果这些原子的初始位移(相对于能量最小值)是沿着本征矢量 \mathbf{A} 的指向,那么位移也将始终沿着同一个本征矢量的指向,并且其幅度是正弦变化的,而频率是由本征值确定的。运动方程的这些特殊解称为正态振型(Normal Modes)。

总的来说,质量加权海塞矩阵 \mathbf{A} 具有 $3N$ 个本征矢量: $\mathbf{e}_1, \dots, \mathbf{e}_{3N}$ 。更为重要的是,式(5.7)运动方程的通解是这些正态振型的线性加和,即

$$\mathbf{x}(t) = \sum_{i=1}^{3N} \left[a_i \cos(\omega t) + b_i \sin(\omega t) \right] \mathbf{e}_i \quad (5.8)$$

式中: a_i 和 b_i 为由各个原子的初始位置和速度所唯一确定的常数。

这意味着这些正态振型并不仅仅是运动方程的特殊解,对于所有可能的初始条件,都能用这些解有效地表征原子运动。如果能够完整地得到所有正态振型,就可以将它们视为对这些原子振动的完整描述。

为了在 DFT 计算中确定出所有正态振型,最主要的任务就是计算海塞矩阵的所有元素。如同在一维空间计算 CO 分子那样,对于海塞矩阵中出现的二阶倒数,可以用有限差分近似来估算,即

$$H_{ij} = \left(\frac{\partial^2 E}{\partial x_i \partial x_j} \right)_{x=0} \approx \frac{E(\delta x_i, \delta x_j) - 2E_0 + E(-\delta x_i, -\delta x_j)}{\delta x_i \delta x_j} \quad (5.9)$$

式中: $E(\delta x_i, \delta x_j)$ 为当所标注坐标为非零值时的全体原子的能量。一

且计算得到了海塞矩阵的所有元素,就可以由其得到质量加权的海塞矩阵 A , 然后可以根据该矩阵的 $3N$ 个本征值得到 $3N$ 个正态振型频率 $\nu = \sqrt{\lambda_i/2\pi}$ 。对于特定原子运动所引致的某个特定振动频率, 可以由对应于该频率的本征矢量来确定。

根据上述讨论, 可以预测: 气相 CO 分子的频率计算会产生 $3N = 6$ 个本征值, 因为一个隔离态的 CO 分子具有 6 个自由度(每个原子都有三个空间坐标), 该结论是正确的。对于前文中没有考虑到的其他 5 个振型, 可以通过分析一个隔离分子完整的海塞矩阵来确定。这样就得到了由刚性平移(Rigid - Body Translation)所确定的三个振型, 可以把这三个振型看作是在分子键长不变的前提下, 沿着 x 、 y 和 z 轴的刚性移动。以这样的方式移动分子时, 并不改变分子的能量, 从而使得其海塞本征值(在理论上)为零。因为在球坐标中确定一个方向时, 需要两个独立角度, 从而存在两个独立的旋转振型。同样地, 对应于这两个振型的本征值(在理论上)也为零。也就是说, 存在有三个平移振型和两个旋转振型, 这并非只有当前所考察的双原子分子才是如此; 对于任何隔离态的两个或多个原子集合, 当允许所有原子自由移动时, 都存在这 5 个正态振型。

在实际计算中由于使用有限差分近似来计算海塞矩阵, 因此在 DFT 计算中上述平移和旋转振型所对应的本征值并不严格为零。当有限差分位移为 $\delta b = 0.04 \text{ \AA}$ 所对应的 CO 分子正态振型如表 5.2 所列。该表中给出了对于每种振型计算得到的频率以及对应于每种振型的本征矢量。在计算中, CO 分子是沿着 z 轴对正的。

表 5.2 DFT 方法计算气相 CO 的正态振型^a

正态振型	频率/ cm^{-1}	原子	本征矢量			振型类型
			$x/\text{ \AA}$	$y/\text{ \AA}$	$z/\text{ \AA}$	
1	2132	C	0.00	0.00	0.76	CO 拉伸
		O	0.00	0.00	-0.65	
2	59	C	-0.65	0.00	0.00	旋转
		O	0.73	0.00	0.00	

(续)

正态振型	频率/cm ⁻¹	原子	本征矢量			振型类型
			<i>x</i> /Å	<i>y</i> /Å	<i>z</i> /Å	
3	56	C	0.00	-0.69	0.00	旋转
		O	0.00	0.70	0.00	
4	0.10 <i>i</i>	C	0.00	0.00	-0.65	平移
		O	0.00	0.00	-0.76	
5	4.5 <i>i</i>	C	0.00	0.71	0.00	平移
		O	0.00	0.70	0.00	
6	7.9 <i>i</i>	C	0.75	0.00	0.00	平移
		O	0.66	0.00	0.00	

a 所列本征矢量的分量单位为 Å, 精确到 2 位有效数字, 对于非零值用粗体加以强调。振型 4、5 和 6 的频率为虚频

在表 5.2 中, 通过观察本征矢量, 可以看出振型 1 中, C 和 O 分子在向彼此相反的方向移动, 从而可以视为 CO 分子的拉伸。类似地, 振型 2 和 3 的本征矢量分别对应于在 $x-z$ 和 $y-z$ 平面内的旋转, 对应于这两种振型的频率并不为零, 大约为 $50 \sim 60 \text{ cm}^{-1}$ 。振型 4、5 和 6 分别对应于沿 z 、 y 和 x 方向的平移, 与旋转振型一样, 这些振型计算所得的频率也不为零。实际上, 每一种振型都存在少量的虚频 (Imaginary Frequency)。对于 2 至 5 振型, 计算所得频率和正确的理论结果 (其频率应严格为零) 之间存在误差的原因是: 通过有限差分表达式以及采用具有有限数值精度的计算方法来求算海塞矩阵, 不可避免地会产生少量的数值不准确。对于平面波 DFT 计算 (或者是对于使用有限差分计算正态振型的任何其他形式量子化学), 结果是具有普遍性的。对于任何特定的原子集合而言, 在计算之前先了解一下存在有多少零频振型 (Zero-Frequency Modes), 这是非常重要的。只要知道了零频振型的数目, 那么通常就能很容易地在计算出的正态振型中识别出这些振型。

5.3 表面上的分子

在讨论了计算原子集合的振动频率之后,进一步来看看更有意义的问题:计算吸附在表面上的 CO 分子的振动频率。下面给出一个简单的例子,来考察一下 CO 吸附在 Cu(001)表面上的情形。实验上已经知道:CO 分子在 $c(2 \times 2)$ 结构(该结构请参见图 4.17)的 Cu(001)表面上形成层状有序排列^[1]。在该结构中,CO 分子与表面上金属原子之间的比例为 1:2。其 DFT 计算结果非常近似于第 4 章中所讨论的,对于这种有序结构,顶位是优先吸附位置,其中:C 原子与 Cu 原子相结合,而 CO 分子的位向是沿着表面法向。对于吸附后的 CO 分子,计算得到的 CO 键长是 1.158\AA ,相比于(计算得到的)气相键长要大了 0.015\AA 。这个平衡键长的伸长现象也反映在分子的振动频率上。

在气相和表面之间,分子几何构型的变化会对分子的振动产生怎样的影响?如果你猜测在表面上 CO 分子的振动频率会降低,那么这就是正确答案。在 CO 吸附到表面后,CO 键会变长;相比于气相中的分子,较长(较弱)的键则以较低的频率振动。从 DFT 计算得到的 Cu(001)表面上 CO 分子的正态振型(如表 5.3 所列^①),也可以看出这一效应。在计算海塞矩阵时,只包括了吸附后分子的自由度;也就是说,在所有的有限差分计算中,表面中金属原子的位置是固定在它们的平衡位置上的。该表中的主要结果就是:在表面上 CO 的拉伸频率计算值为 2069 cm^{-1} ;这与气相 CO 的数值 2131 cm^{-1} 相比是显著降低的。这一现象与实验也是一致的,已经从实验上发现表面上 CO 的拉伸频率是 2086 cm^{-1} ,相比于实验上观测的气相频率要低 57 cm^{-1} 。

在第 5.2 节,说明了气相分子的平移和旋转并不会导致分子能量的变化,因此从理论上说也就导致了频率为零的正态振型。这一描述

① 原文中为表 5.2——译者注。

对于表面上的分子就不再成立了,在平移分子(表面不动)时会由于分子与表面的相互作用而导致分子能量发生改变。这一现象可从表 5.3 中的振型 2~6 观察到。与气相结果一致的是,最大本征值的本征矢量对应于 CO 的拉伸振型(振型 1)。然而,第二大的本征值却远非是零,而是 322cm^{-1} ,这个振型对应于 CO 分子沿 z 方向的平移。仔细思考一下,就能够理解这个振型可以更好地描述为在碳原子和表面铜原子之间进行拉伸。表 5.3 中其他 4 个振型可以很好地描述为受到阻碍(Frustrated)的旋转和平移,但是这些描述都是示意性的,而远非精确描述。此时,在 $x-z$ 和 $y-z$ 平面内的旋转则对应于 Cu-C-O 角度的弯曲度,这就是表 5.3 中的振型 3 和 4。振型 5 和 6 分别与沿 y 和 x 方向的平移是最为接近的。由于本例中的对称性,振型 3 和 4 的频率是相等的,振型 5 和 6 的频率也是如此(在数值精确度内)。

5.4 零点能

上述有关振动的讨论全部是在谐函数近似的前提下进行的。在使用这一近似方法时,每个振动的振型都可以看作是由谐振子(Harmonic Oscillator)所定义的。一维谐振子的势能可以写为

$$E = E_0 + \frac{k}{2}x^2 \quad (5.10)$$

式中: x 为所选定的坐标,而 $x = 0$ 就是平衡位置。在经典力学中,谐振子的行为特性是非常简单的,而且在前两节中,已经使用其行为特性来对振动加以理解。然而,更准确的描述应当是用量子力学来描述该现象。巧合的是,谐振子是可以简单方式就完全求解的少数几个量子力学问题中的一个。有关求解细节在任何量子力学教材中都能找到,其中一些列在本章的最后。在这里,主要说明与 DFT 计算有关的一个特定量子力学现象,也就是零点能(Zero-Point Energy)。

表 5.3 与表 5.2 类似,但是 CO 分子吸附于 Cu(001)的 $c(2 \times 2)$ 表面结构上^a

正态振型	频率/cm ⁻¹	原子	本征矢量			振型类型
			$x/\text{Å}$	$y/\text{Å}$	$z/\text{Å}$	
1	2069	C	0.00	0.00	0.77	CO 拉伸
		O	0.00	0.00	-0.64	
2	322	C	0.00	0.00	0.64	平移
		O	0.00	0.00	0.77	
3	279	C	0.26	0.87	0.00	旋转
		O	-0.12	-0.40	0.00	
4	279	C	-0.87	0.26	0.00	旋转
		O	0.40	-0.12	0.00	
5	46	C	0.08	0.41	0.00	平移
		O	0.18	0.89	0.00	
6	38	C	-0.41	0.08	0.00	平移
		O	-0.89	0.18	0.00	

*a*CO 分子与 z 方向对齐,非零值用粗体加以强调

式(5.10)定义的谐振子所能取得的最低能量是什么? 应用经典力学,答案很简单:就是在 $x = 0$ 时的平衡态,零动能和势能 E_0 。然而,量子力学答案则并不简单,这是由于 Heisenberg 测不准原理 (Uncertainty Principle) (大致)说明:不能够以任意精度同时获知粒子的位置和动量。因为经典最低能量状态同时精确地确定了谐振子动量和位置(为零),所以这并非是一个合理的量子力学状态。对于谐振子,最低的量子力学能量可表示为

$$E = E_0 + \frac{h\nu}{2} \quad (5.11)$$

式中: h 为 Planck 常数; ν 为谐振子(经典的)振动频率。在这个能量和经典最小能量之间的差值称为零点能。量子力学谐振子所存在的其他能级是 $E_n = E_0 + \left(n + \frac{1}{2}\right)h\nu$,其中 n 为可能取值的整数。

试想一下,根据一个原子集合及其所确定的正态振型集合,可以分别独立地确定每个振型的零点能,则该原子集合所能取得的最小能量为

$$E = E_0 + \sum_i \frac{h\nu_i}{2} \quad (5.12)$$

式中: E_0 为 DFT 计算得到的能量; ν_i 为正态振型的频率。在本章的开头,就已经讨论过是怎样把 DFT 结果看作是提供了 0 K 下的信息,其中的原子和分子都具有各自的最低能量。现在能明白:为了更准确地给出这一描述,还应当包括式(5.12)所定义的零点能。在实际计算中,这个方法并不常用。忽略零点能的一个原因是为了讲究实效;相比于仅仅对 DFT 能量的最小化计算,计算正态振型的频率时,所耗费的计算成本要大得多!

令人高兴的是,对于 DFT 计算总能时不包括零点能,在物理上也有一个好的理由。在许多情形下,即便是当总的零点能已经相当大了,在所研究的两个状态之间,零点能的改变量也是相对较小的。可以用 CO 在 Cu(100) 上吸附为例进行说明:如果忽略零点能,分子的吸附能是

$$E_{\text{ads}} = (E_{\text{CO}} + E_{\text{Cu}}) - E_{\text{CO/Cu}} \quad (5.13)$$

如第 4 章所述,该表达式包括了三个独立的 DFT 计算结果:气相 CO 的能量 E_{CO} , 洁净 Cu 表面的能量 E_{Cu} , 以及吸附了 CO 分子后 Cu 表面的能量 $E_{\text{CO/Cu}}$ 。对于这个例子,所得到的 $E_{\text{ads}} = 0.71\text{eV}$ 。如果假设只有 CO 分子产生了振动,并将零点能包含进来,则该吸附能为

$$E_{\text{ads}} = (E_{\text{CO}} + E_{\text{Cu}}) - E_{\text{CO/Cu}} + \frac{h\nu_{\text{CO}}}{2} - \sum_{i=1}^6 \frac{h\nu_i}{2} \quad (5.14)$$

该表达式右侧的零点能项来自于气相(吸附态)分子的一(六)振型。包含有零点能的最终结果是 $E_{\text{ads}} = 0.65\text{eV}$, 与上面的结果相比,差别仅约为 0.06eV ^①。与纯粹的吸附能相比较,由零点能所产生的贡献并不大。单纯对于零点能而言,绝大部分来自于表面上具有非零频率的

① 尽管由零点能所产生的吸附能改变量不是很大,但在很多情形下将其忽略也是很不明智的。例如,当在若干可能的吸附位置中,试图确定哪个位置在能量上是优先的,尤其是当不同位置间的(经典)能量差类似于零点能的校正值时。

5 种振型(如式(5.14)所示),而并非来自于 CO 伸缩振型。在吸附能的零点能校正量中,与改变 CO 伸缩频率(在气相和吸附分子之间)有关的能量仅为约 0.004eV。

定性地来说,对于与轻原子有关的计算,零点能效应会变得更加重要。这意味着:特别是在与 H 原子有关的问题中,零点能也许会变得尤为重要。例如,可以比较一下在 Cu(100)表面上一个 H 原子的能量,和体相 Cu 中一个 H 原子的能量。在 FCC 金属内部,可以考虑两种空隙位置,如图 5.2 所示分别为六重对称的八面体位置和四重对称的四面体位置。原子态 H 在这些空隙位置上相对于原子态 H 在 Cu(100)表面上的能量差为

$$\Delta E = (E_{\text{H/Cu}(100)} + E_{\text{bulk}}) - (E_{\text{Cu}(100)} + E_{\text{H/bulk}}) \quad (5.15)$$

式中: E_{bulk} 和 $E_{\text{Cu}(100)}$ 分别为 DFT 计算的含有体相 Cu 和 Cu(100)板块的超晶胞能量; $E_{\text{H/bulk}}$ 和 $E_{\text{H/Cu}(100)}$ 是上述模型中分别引入一个 H 原子之后的超晶胞能量。根据该定义, ΔE 负值意味着:相比于表面位置,所考察的体相间隙位置在能量上更为优先。

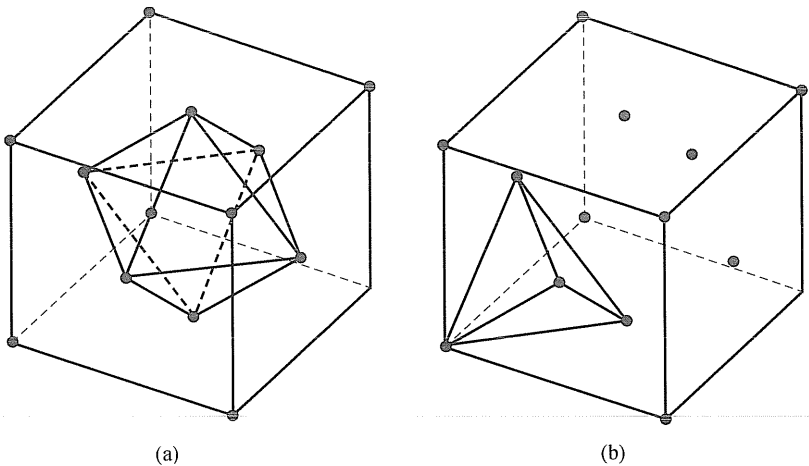


图 5.2 FCC 金属中八面体和四面体空隙位置示意图
(八面体位置是由六个面心原子相连接所形成的,而四面体位置是由四个最近邻原子连接后形成的)

上述计算的结果汇总于表 5.4 中,其中并没有包括零点能。这些数据意味着,体相中八面体位置在能量上稍比表面位置更为优先;而体相中四面体位置则较不优先。对于表面位置和体相八面体位置而言,计算所得零点能具有相似的数量级,因此,即使将零点能包括进来,这两个位置之间的能量差也只会发生稍小变化。原子 H 在体相四面体位置的零点能要比其他两个位置大得多。这在物理上意味着:H 原子在四面体位置中所受的限制要比在八面体位置中的限制更大,从而使四面体位置上的 H 原子具有更高的振动频率。如果想要比较四面体位置和其他两种位置之间的能量差,就必须清楚地意识到:在这个比较中,将零点能包含进来是更加重要的。

表 5.4 根据式(5.15)的定义,H 在 Cu(100)上和 H 处于体相 Cu 间隙位置之间的相对能量差

单位:eV

位置	相对能量差(不含零点能)	零点能	相对能量差(包含零点能)
Cu(100)表面	0	+0.12	0
体相八面体	-0.05	+0.14	-0.03
体相四面体	+0.12	+0.21	+0.21

对于零点能,最后需要说明的是:上述所有零点能都对应于原子集合在 0 K 下所形成的最低能量状态。当在非零温度下考察相同原子集合时,了解其在该特定温度下的平均能量常常是非常有用的。如果继续使用谐函数近似来描述每一个正态振型,那么这个对所有能级加和后的平均能量是量子统计力学所能得到的为数不多的几个结果之一,且可以写作一个简单的闭型(Closed Form)形式。对于式(5.10)所定义的每一个谐振子,在温度 T 下的平均能量为

$$\bar{E}(T) = \frac{h\nu [\exp(\beta h\nu/2) + \exp(-\beta h\nu/2)]}{2 [\exp(\beta h\nu/2) - \exp(-\beta h\nu/2)]} \quad (5.16)$$

式中: $\beta = 1/k_B T$ 。该表达式在 $T \rightarrow 0$ 时简化为式(5.11),因此它与上述已经讨论过的 0 K 下的结果是具有一致性的。在相反的极限中,对于极高的温度则有 $\bar{E}(T) \rightarrow kT$,这也就是经典极限,其中谐振子的平均能

量与其频率无关。

式(5.16)的一个含义就是:当考虑零点能时,在原子集合两种构型之间的能量差是与温度有关的^①。在低温极限中,能量差值是在包括每种构型的零点能后所给出的。在足够高的温度下,每一个正态振型都达到了经典极限,于是在两个构型之间的能量差值也仅仅是经典能量的差值(不需要考虑零点能)。这个概念在图 5.3 中进行了进一步阐释:在体相 Cu 中,H 原子分别处于八面体和四面体位置的能量差是与 T 相关的。当温度处于约 200K 以下时,这个能量差就是零点能的校正值,即 0.24eV(如表 5.4 所列)。在高温下,能量差值缓慢地降低,平滑地向着极高温下的经典极限(0.17eV)靠近。

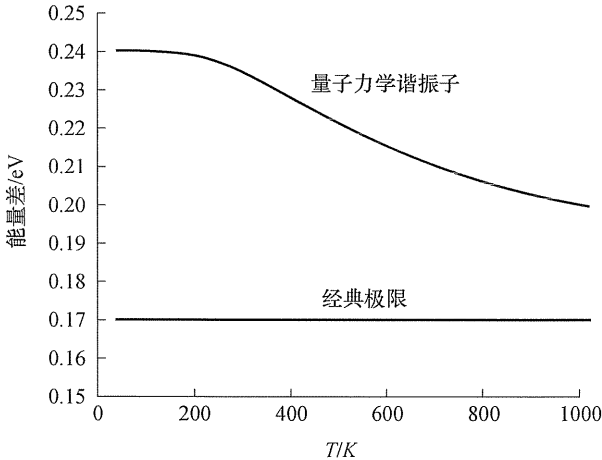


图 5.3 针对 H 原子的每个正态振型,根据式(5.16)计算单个 H 原子在体相 Cu 的八面体和四面体位置之间的能量差(该能量差是与温度相关的。该能量差的低温极限和经典极限可参见表 5.4)

^① 这是采用量子力学描述谐振子的结果。在纯粹的经典力学描述中,每一个谐振子的平均能量都只是 $k_B T$,因此在两个构型之间的能量差是与温度 T 无关的。

5.5 声子和非定域化振型

上述讨论中有关振动的例子都是处理一个局域化的原子集合,或者是一个气态分子,或者是一个吸附在表面上的分子。读者由前面几章的学习可能已经领会到平面波 DFT 计算的优点就在于:能够以一种非常自然的方式,将材料在空间中展开,如体相固体。表征体相材料的振动态就是所谓的声子。类似于局域化体系的正态振型,可以把声子看作是:对振动原子集合进行经典描述后的特殊解,而且对于这些原子从任何初始状态所最终形成的振动,都可以用这些解与其他声子的线性组合来描述。与分子的正态振型有所不同,声子在空间上是非局域化的,并且涉及到在空间上严格周期性排列的无限多的原子集合。一个分子的正态振型是由一组离散的振动所定义的,而材料的声子是由一个连续频率区间内的连续声子谱(Spectrum Of Phonons)所定义的。当描述声子时,重点关注的是具有特定振动频率的声子数目,也就是说,是振动态密度(Vibrational Density Of States)。分子振动在描述分子结构和特性时具有核心作用,类似地,声子态密度对于描述固体的许多物理特性也具有核心作用。在所有固体物理的教科书中都肯定存在与此相关的内容,本章最后列出了其中的部分书目。

使用 DFT 方法预测声子态密度,在概念上类似于求解局域正态振型的过程。在计算中,先使原子在其平衡位置附近发生很小的位移,从而可以对体系的海塞矩阵使用有限差分近似,如同式(5.3)。关于将该信息代入到声子态密度的数学方法,现在已经发展得很完善了,只不过要比第 5.2 节中所叙述的类似内容更加复杂。然而,在大多数常见 DFT 计算软件包中,该计算过程仍不能作为一个常规选项被获取(但是对某些研究小组,这些计算已被广泛使用)。如果读者对此还有进一步的探索兴趣,可以参考本章最后的扩展阅读。

练习

1. 使用 DFT 计算确定气相 N_2 的振动频率。将计算结果与实验值进行比较。计算结果与有限差分近似中所使用的位移之间具有怎样的关系？

2. 使用 DFT 计算确定气相氨 NH_3 的振动频率。将计算结果与实验值进行比较,并说明每个频率所对应的振动类型。

3. 氢原子在 $Cu(111)$ 上可以结合在两种完全不同的三重对称位置上,即 FCC 位置和 HCP 位置。使用 DFT 方法计算这两个位置的经典能量差值。然后,计算 H 在每个位置上的振动频率,其中需要假定吸附后 H 原子的正态振型。当考虑了零点能后,两个位置之间的能量差会发生怎样的变化？

4. 在上一个练习中,假设了 H 的振动是与金属表面原子无关的。对于该近似计算的精度,有一个粗略的检查方法:就是允许 H 原子与三个最近邻金属原子自由移动,并求算其正态振型,该计算涉及到 12 个正态振型。在一个三重对称位置上运行该计算,确认出由 H 原子运动所主导的三个振型。将这些振型的频率与之前简单计算的结果相比较,二者有多大不同？

参考文献

Ryberg R. Carbon - Monoxide Adsorbed on $Cu(100)$ Studied by Infrared Spectroscopy. Surf. Sci. 114, 1982; 627 - 641.

扩展阅读

量子力学中对谐振子的处理在所有有关量子力学的教科书中都有所表述。几个较好的例子是:

Atkins P W, Friedman R S. Molecular Quantum Mechanics. Oxford U-

niversity Press, Oxford, UK, 1997.

McQuarrie D A. Quantum Chemistry. University Science Books, Mill Valley, CA, 1983.

Ratner M A, Schatz G C. Introduction to Quantum Mechanics in Chemistry. Prentice Hall, Upper Saddle River, NJ, 2001.

Simons J, Nichols J. Quantum Mechanics in Chemistry. Oxford University Press, New York, 1997.

有关声子的概念描述了在扩展态固体中的振动,是固体物理的核心主题。该领域的两本经典著作是:

Ashcroft N W, Mermin N D. Solid State Physics. Saunders College Publishing, Orlando, 1976.

Kittel C. Introduction to Solid State Physics. Wiley, New York, 1976.

关于采用 DFT 计算确定体相材料声子态密度,一个简明扼要的综述是:

Ackland G J. J. Phys. Condens. Matter 14, 2002; 2975 - 3000.

附录 计算细节

本章所有的计算都使用 PBE GGA 泛函和平面波基底,其中截断能为 380eV 。对于所有的表面计算,使用 DFT 优化后的晶格参数来定义超晶胞在表面晶面内的尺寸。

5.1 节 有关气相 CO 伸缩振动的计算是在一个大约为 $10 \times 10 \times 10\text{\AA}$ 的盒子里进行的,且 CO 分子与 z 轴对正。盒子的三个尺寸有意做得稍有不同^①,这是为了避免产生一个完全对称的盒子。采用了 Gaussina Smearing 方法,宽度为 0.1eV 。只在 Γ 点对倒易空间取样。每个离子在每个方向上都进行了一个微小的或正或负的位移。如本章所述,在有限差分计算中,对原子位移使用了一系列数值。

5.2 节 除了所使用的有限差分为 0.04\AA 之外,有关气相 CO 伸

① 原文如此——译者注。

缩振动的计算与 5.1 节相同。

5.3 节 在考察吸附于 Cu(100) 表面的 CO 振动时,使用了具有 4 层原子的非对性板块模型,并在正态振型计算之前,将最下端的两层原子固定在体相位置,允许其他所有原子弛豫。该超晶胞具有 $c(2 \times 2)$ 表面对称性,每层含有 2 个金属原子。计算中使用的 k 点网格是 $6 \times 6 \times 1$ 。采用 2 阶的 Methfessel - Paxton 方法,Smearing Width 为 0.1eV。

5.4 节 在考察 H 在 Cu(100) 表面的吸附时,使用了具有 4 层原子的非对称性板块模型,将最底端的两层原子固定在体相位置,允许其他所有原子弛豫。 k 点网格是 $6 \times 6 \times 1$,采用 2 阶的 Methfessel - Paxton 方法,smearing width 为 0.1eV。在考察 H 在间隙位置的吸附时,使用了包含有 64 个金属原子的 FCC Cu 超晶胞。在几何优化过程中,允许超晶胞中所有的 65 个原子自由弛豫,但按照纯 Cu 优化后的数值固定了超晶胞的尺寸。关于确定间隙位置上 H 振动频率的计算,如本章正文所述,使用了有限差分法,并且在所计算的两次位移中,每次位移每个 H 原子均为 0.04Å。

第 6 章 使用过渡态理论 计算化学过程速率

在很多情况下,化学反应速率在技术和科学上都是非常重要的。在第 1 章中,曾经简要地讨论过在由 N_2 和 H_2 合成氨 (NH_3) 的反应中使用异质催化剂。对于这个反应过程中的高效催化剂,也许最基本的特点就是:相比于其他可能的催化剂,能够极大地提高净反应速率。在完全不同的另一个领域内,关于在北极上方导致臭氧洞形成的大气化学,则包含了一系列很复杂的、在极地平流层云团中的冰晶上所发生的化学反应。对于提出有关大气臭氧消耗的理论描述而言,了解这些反应的速率是非常关键的。还可以给出另一个例子,固体材料的辐射损害常常会导致形成局域缺陷,例如晶格空位。为了预测这些缺陷对材料稳定性的长期影响,就必须知道这些缺陷在材料中的移动速率(伴随其他情形)。

读者现在已经对 DFT 计算具有足够的经验,应该能够想象到如何对上述三种情形应用 DFT 计算。例如,对于可能存在于固体材料中的各种晶格缺陷,可以用 DFT 计算确定它们之间的相对能量差。类似地,可以用 DFT 计算确定在冰晶表面反应分子的平衡位置,这也可以视为对极地平流云团的模拟。但这些计算只给出了最低能量状态的信息。在第 5 章已经讲过,这些计算只是给出了 0 K 下的信息,而即便对于极地平流层云团,这个温度也有些太冷了!

在本章中,主要讲述如何用 DFT 方法计算化学过程的速率。这里谨慎地使用了“过程(Process)”这个单词,从而不仅包括了化学反应(从过程的起点出发,能够在过程结束时得到不同的化合物),而且也包括了其他情形,例如固体中缺陷的移动(起点和终点可以视为在化学上是完全相同的)。在这两种情形中,考察这些过程的一个有

效方式就是由 N 原子集团所决定的总能表面 (Overall Energy Surface), 即 $E(\mathbf{R}_1, \dots, \mathbf{R}_N)$ 。总的来说, 这个能量面具有多个极小值: $E_1, E_2, E_3 \dots$, 所需考虑的最基本过程就是 N 原子构型从一个极小值点 E_i 移动到另一个极小值点 E_j , 而没有经过其他任何极小值点。本章最主要的任务就是说明如何用 DFT 方法精确地确定这类过程的速率。

6.1 一维的例子

在本章所有内容中, 都是以一个特定的化学过程为例, 并试图确定该过程的速率, 即一个 Ag 原子在 Cu(100) 表面上的扩散。这个例子的优点就是非常易于想象, 并且对于许多具有更多物理或技术含义的化学过程而言, 也具有这个例子所包含着的特点。

图 6.1 所示为 Ag 原子在 Cu(100) 上的吸附。将 Ag 原子能量具象化的一个有效方法就是: 把该原子固定在表面晶面 ($x-y$ 平面) 的位置上, 然后让 Ag 原子沿表面法线 (z 方向) 向表面移动, 从而使能量最小化, 并允许所有的表面原子弛豫。经过这些步骤, 就能够得到一个二维的能量面 $E(x, y)$, 从而可以在这个表面上确定出 Ag 原子的最小能量。

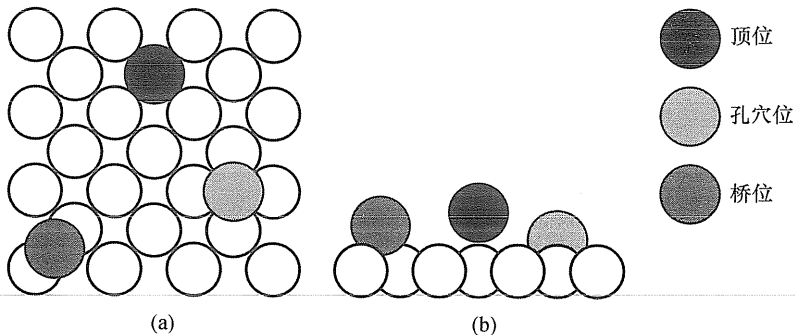


图 6.1 Ag 吸附在 Cu(100) 表面上的示意图, 分别显示了 Ag 原子吸附在顶位、四重对称孔穴位和桥位 (a) 俯视图; (b) 侧视图

该能量面如图 6.2 所示,在表面上有三种类型的临界点,分别对应如图 6.1所示的高对称性表面位置。四重对称位置是这个表面上唯一的能量极小值,桥位是表面上的一阶鞍点(First - Order Saddle Points);从桥位向四重对称位移动可以降低 Ag 原子的能量,但是,从桥位向顶位移动则会增加该能量。顶位是二阶鞍点(Second - Order Saddle Points),因为此时,在表面平面内向任何方向移动都会降低 Ag 原子能量,而沿着表面法线向上或向下移动都会增加改点的能量。

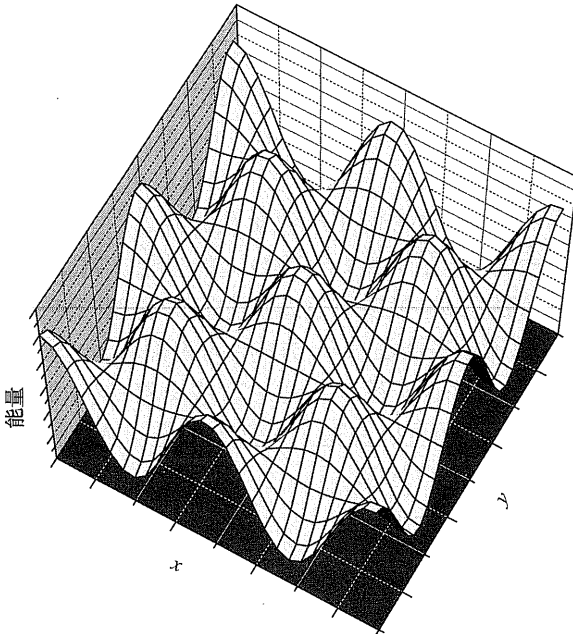


图 6.2 Ag 原子在 Cu(100)上的二维能量面 $E(x,y)$ 示意图
(极小值点是四重对称的表面位置)

当 Ag 原子从一个孔穴位移动到另一个孔穴位时,存在无数可能的移动轨迹,但对于理解其过渡态速率而言,具有特殊意义的一条轨迹就是:在两个位置间沿着能量变化最小的路径运动。从图 6.2 可以看到,这条路径会通过两极小值之间的桥位鞍点,把这条特殊路径称为该

过程的最小能量路径 (Minimum Energy Path)。图 6.3 显示了在两个四重对称位置之间,一个 Ag 原子在 Cu(100)上沿着最小能量路径移动时的能量变化。为了更好地理解这个结果,可以比较一下这个 Ag 原子所必须经历的能量变化范围,以及温度 T 下单个原子所可能具有的典型能量值。根据统计力学可以得到一个重要结论:当原子集合在温度 T 下处于平衡态时,则对于每个自由度,该体系所能取得的平均能都是 $k_B T/2$,式中: k_B 为玻耳兹曼常数。在室温下 $k_B T/2$ 等于 0.013eV 。而根据图 6.3 所示计算结果可以确定:单个 Ag 原子从一个孔穴位移动到桥位时,所需克服的能垒是 0.36eV ,这要比平均热能的数值大约高出 30 倍。因此,从图 6.3 可以合理地推断出:在室温下,一个具有“典型”能量值的 Ag 原子,更有可能发现它非常靠近四重对称位这个能量最低点位置。

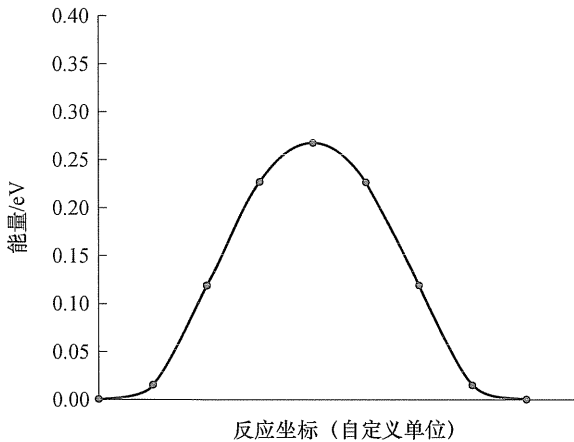


图 6.3 在 Cu(100)上两个四重对称位置之间沿着最小能量路径移动时 DFT 计算得到的 Ag 原子能量。所有能量均为相对于 Ag 在四重孔穴位的能量差。反应坐标 (reaction coordinate) 是在 $x-y$ 平面内连接两个相邻能量极小值点的直线

图 6.3 中能量曲线只是表征了与 Ag 原子位置的关系,在真实的金属表面上,体系内原子之间会由于彼此碰撞不断交换能量,并且每个原子都会受到其他所有原子瞬时力的作用,从而发生位置移动。这意味

着:Ag 原子时常会得到比平均热能(Typical Thermal Energy)多很多的富余能量,使其有可能越过图 6.3 中所示的能垒。对于化学过程所需能量要比平均热能大很多的情形,可以用过渡态理论(Transition State Theory, TST) 计算该过程的总速率(Net Rate)。

图 6.4 显示了一维能量曲线的示意图,与图 6.3 中结果具有相同的形状。图 6.4 是根据反应坐标 x (Reaction Coordinate) 所绘制的, x 就是沿着连接 $x=A$ 和 $x=B$ 两个极小值路径所测得的距离。将两个极小值分隔开的位于 x^\ddagger 的鞍点称为过渡态(Transition State)。可以将图 6.4 中过渡态左侧(右侧)的所有点都称为状态 A(状态 B)。

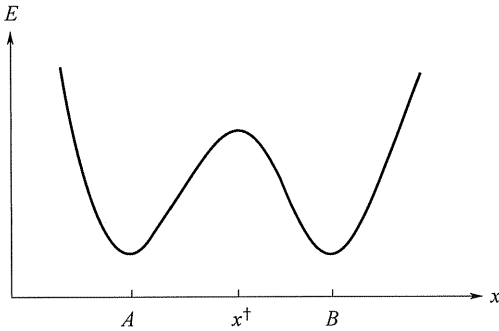


图 6.4 一维能量曲线示意图,其中显示了被一个过渡态所分隔开的两个极小值点

对于从状态 A 到状态 B,过渡态理论给出的跳动速率(Rate of Hopping)为

$$k_{A \rightarrow B} = \left(\frac{1}{2}\right) \times (\text{跳动原子的平均热速度}) \times (\text{原子出现在 } x = x^\ddagger \text{ 位置上的概率}) \quad (6.1)$$

这里出现二分之一因子的原因是:对于原子出现在过渡态的所有可能形式中,只有其中一半的速度是朝向右侧运动的(另一半是朝向左侧运动的)。对于该表达式最后一项所提到的概率,当材料在温度 T 下处于热力学平衡态时,可根据相对于状态 A 所有可能的原子位置来确定。此时,观察到原子位于任意位置的概率可以写为

$$p(x) \propto \exp\left[-\frac{E(x)}{k_B T}\right] \quad (6.2)$$

由于因子 $1/k_B T$ 已经出现多次, 可以令 $\beta = 1/k_B T$ 。

使用这一标记方法之后, 则有

$$\text{原子出现在 } x = x^\dagger \text{ 位置上的概率} = \frac{e^{-\beta E(x^\dagger)}}{\int_A dx e^{-\beta E(x)}} \quad (6.3)$$

其中, 积分项是对应于 $x = A$ 附近所有可能的最小值位置上进行的积分。在温度 T 下处于热力学平衡态时, 材料中每个原子的速度都遵守所谓的 Maxwell - Boltzmann 分布。该分布中一个原子的速度可以精确计算为

$$\text{跳动原子的平均热速度} = \sqrt{\frac{2}{\beta \pi m}} \quad (6.4)$$

式中: m 为原子质量。综合以上表达式, 则有

$$k_{A \rightarrow B} = \frac{1}{2} \sqrt{\frac{2}{\beta \pi m}} \frac{e^{-\beta E^\dagger}}{\int_A dx e^{-\beta E(x)}} \quad (6.5)$$

这里, 我们定义 $E^\dagger = E(x^\dagger)$ 。

上述方程并没有太多帮助, 因为在分母中出现了一个积分。但是, 可以根据下面的思路对其进行简化: 对于位于 $x = A$ 最小值附近且具有典型热能 (Typical Thermal Energy) 的原子, 仅仅从最小值处偏离了很小的数值。因此, 可以根据第 5 章分析讨论振动时的一个概念, 使用泰勒展开式将靠近最小值处的能量展开。这样一来, 就可以给出能量的近似表达为

$$E(x) \approx E_A + \frac{k}{2}(x - x_A)^2 \quad (6.6)$$

在第 5 章中已经说明 k 是与原子在势能最低点振动频率有关的, 其振

动频率为 $\nu = \frac{1}{2\pi} \sqrt{\frac{k}{m}}$ 。这就意味着

$$\int_A dx e^{-\beta E(x)} \approx e^{-\beta E_A} \int_A dx e^{-\beta k(x-x_A)^2/2} \quad (6.7)$$

由于等式右侧积分项中的函数会随着 $|x - x_A|$ 增大而急剧减少, 则有

$$\int_A dx e^{-\beta E(x)} \approx e^{-\beta E_A} \int_{-\infty}^{+\infty} dx e^{-\beta k(x-x_A)^2/2} = \sqrt{\frac{2\pi}{\beta k}} e^{-\beta E_A} \quad (6.8)$$

这也就意味着总速率常数 (Overall Rate Constant) 可以写成非常简单的形式, 即

$$k_{A \rightarrow B} = \nu \exp\left(-\frac{E^\ddagger - E_A}{k_B T}\right) \quad (6.9)$$

根据式(6.6)中所采用的能量近似方式, 该结果也称为谐波过渡态理论 (Harmonic Transition State Theory, HTST)。该速率值只与两个物理量有关, 二者均可在能量面上用一种简单方法确定: ν 是势能最低点处原子的振动频率, 而 $\Delta E = E^\ddagger - E_A$ 是在能量最低点和过渡态之间的能量差。 ΔE 也称为该过程的活化能, 总的过程速率 $k_{A \rightarrow B} = \nu \exp(-\Delta E/k_B T)$ 也就是对于化学反应速率非常著名的 Arrhenius 表达式^①。

大致估算式(6.9)中的前置因子并不是一件很困难的事情。一个“典型的”原子振动具有的周期是 0.1 ~ 1ps, 因此每秒有 $10^{12} \sim 10^{13}$ 次振动。这就意味着: 在式(6.9)中, $\nu = 10^{12} \sim 10^{13} \text{ s}^{-1}$ 就是个合乎情理的估计值, 而不需要进行其他任何计算。将其与前面所讨论的 Ag 原子在 Cu(100) 表面的例子相比较, Ag 原子在两个四重对称位置之间沿反应坐标的振动频率为 $\nu = 1.94 \times 10^{12} \text{ s}^{-1} = 1.94 \text{ THz}$, 这与上述估算值具有良好的一致性。根据计算得到的 Ag 在 Cu(100) 上跳动的活化能 (0.36eV), 以及 Ag 原子的振动频率, 可以用式(6.9)预测 Ag 原子的跳动速率 (Hopping Rate)。表 6.1 列出了在几个不同温度下所得到的速率结果。对于这些结果, 最主要的特点以及所有活化过程 (Activated Processes) 的本质特点就是: 当温度的改变范围相对很小时, 就会对该速率造成若干数量级的变化。这一特点是由式(6.9)中的活化能所导致的, 而不是由于振动频率所导致的。因此, 通常都采用简单的估算值 $\nu = 10^{12} \sim 10^{13} \text{ s}^{-1}$, 而不是进行运算来获得一个更为精确的 ν 值。

① 以瑞典化学家 Svante Arrhenius 的名字命名, 虽然他以及格的成绩获得了博士学位, 但其中的研究工作后来却使他赢得了诺贝尔奖。在 1896 年, Arrhenius 是第一个提出大气中不断增加的 CO₂ 含量与全球气温之间存在关联的科学家。

表 6.1 根据式(6.9)计算 Ag 在 Cu(100)上的跳动速率

温度 /K	速率 /s ⁻¹
1500	0.97
300	1.4 × 10 ⁶
450	1.5 × 10 ⁸

图 6.5 显示了振动频率和活化能对于跳动速率的影响作用。图中实线显示的是用三个振动频率估算值计算得到的一维 TST 结果,这三个频率的其中一个是通过 DFT 计算得到的频率,另外两个是“简单的”估测值,即 10^{12} s^{-1} 和 10^{13} s^{-1} 。虚线显示的是使用了 DFT 计算频率之后得到的结果,但在 DFT 计算所得活化能中加上或减去了 0.05eV。相比于改变振动频率所引起的跳动速率变化,改变活化能所导致的变化是更加显著的。这个图中所给出的主要信息就是:在应用过渡态理论进行计算时,应该使活化能的计算尽可能地精确;而在计算振动频率时,可以不用太过于精确。

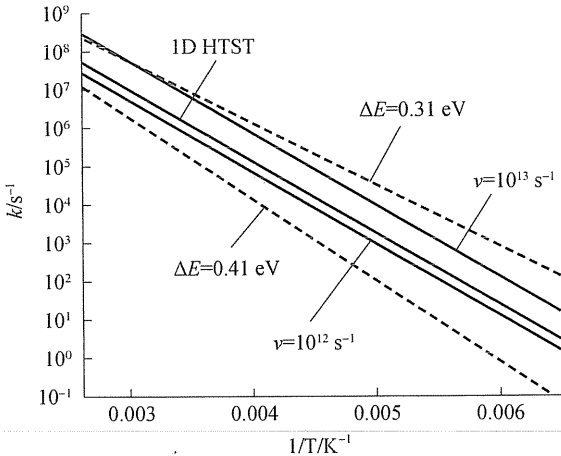


图 6.5 使用一维简谐过渡态理论(1D HTST)预测的 Ag 在 Cu(100)上的跳动速率。其他两条实线显示的是使用 DFT 计算的活化能($\Delta E = 0.36 \text{ eV}$)和预估的前置因子(10^{12} 或 10^{13} s^{-1})所预测的速率。两条虚线显示的是使用 1D HTST 方法,采用 DFT 算得的前置因子($\nu = 1.94 \times 10^{12} \text{ s}^{-1}$),而活化能变化 $\pm 0.05 \text{ eV}$ 时所预测的结果

6.2 多维过渡态理论

上述有关 Ag 在 Cu(100) 上移动的结果并不能描述该过程的所有细节,因为在上面的讨论中把这个问题简化成了一维最小能量路径,原本就是一个近似。如果 Ag 原子能够得到足够多的能量,并越过最小能量路径上的过渡态,那么它也可能以类似的能量处于该路径之外的其他位置上。为了说明这个效应,需要使用多维空间的过渡态理论。幸运的是,该理论的最终表达形式几乎与一维过渡态理论一样简单。

将式(6.5)进行多维推广后的形式是

$$k_{A \rightarrow B} = \frac{1}{2} \sqrt{\frac{2}{\beta \pi m}} \frac{\int_{x=x^\ddagger} \mathbf{d}\mathbf{r} e^{-\beta E(\mathbf{r})}}{\int_A \mathbf{d}\mathbf{r} e^{-\beta E(\mathbf{r})}} \quad (6.10)$$

式中: \mathbf{r} 为所有粒子坐标的矢量; x 为用于定义反应坐标的一个坐标值。可以采用 6.1 节中所使用的谐函数近似,对其中的两个积分项进行近似处理,从而将上式改写为更为易用的形式。对于这个分式,将能量面在能量最小值 $\mathbf{r} = \mathbf{r}_A$ 附近展开为

$$E = E_A + \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \left[\frac{\partial^2 E}{\partial r_i \partial r_j} \right]_{\mathbf{r}_A=0} (r_i - r_{i,A})(r_j - r_{j,A}) \quad (6.11)$$

之前在第 5 章中已经看到过这个表达式,在描述振动型式时就是从它开始的。在第 5 章中发现,可以用一个很自然的方法来考察这类谐势能表面 (Harmonic Potential Energy Surface),这个方法就是定义该体系的正态振型,所具有的振动频率是 $\nu_i (i = 1, \dots, N)$ 。这些正态振型就是运动方程的特殊解,其表达形式为

$$(\mathbf{r} - \mathbf{r}_A) \sim \cos(2\pi\nu_i t) \mathbf{e}_i \quad (6.12)$$

式中: \mathbf{e}_i 为定义第 i 个正态振型方向的矢量。

对于式(6.10)中的分式,可以使用完全相同的思路。将能量在过渡态附近以泰勒方式展开到二阶项,即

$$E = E^\dagger + \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N \left[\frac{\partial^2 E}{\partial r_i \partial r_j} \right]_{r=r^\dagger} (r_i - r_i^\dagger)(r_j - r_j^\dagger) \quad (6.13)$$

在这里之所以没有出现一阶导数项,是因为过渡态是能量面的一个临界点,因此在过渡态上所有的一阶导数均为零。如同在第5章所做的那样,这个在能量面上的谐函数近似可以用正态振型进行分解。这涉及到需要对有二阶导数所定义的质量加权海塞矩阵进行计算,并找到该矩阵的 N 个本征值。

下面考虑一个具体的例子:一个 Ag 原子从 Cu(100) 上的四重对称位跳出;其中,只考虑定义 Ag 原子位置所需要的三个坐标,即 $N = 3$ 。使用第5章的方法可以计算得到,当 Ag 原子处于最低能量(位于表面的四重对称位置上),以及处于过渡态(桥位)时海塞矩阵的本征值。在这两个情形下,表面 Cu 原子的坐标在寻找临界点时是可以弛豫的,而在计算海塞矩阵时是固定不动的。为了从物理上理解这些本征值,不妨回忆一下:正态振型的频率是由 $\nu_i = \sqrt{\lambda_i/2\pi}$ 所确定的。表 6.2 列出了根据该表达式计算的能量最低点和过渡态的频率。

这些频率值最为突出的一个特点就是:在过渡态的频率中有一个虚频(与该振型对应的本征值是负值)。如果把这个频率写作 $\nu_3 = i\alpha$,则在式(6.12)中该正态振型解的形式为

$$(r - r^\dagger) \sim \cos(2\pi i\alpha t) e_3 = \frac{1}{2}(e^{+2\pi i\alpha t} + e^{-2\pi i\alpha t}) e_3 \quad (6.14)$$

这个解意味着:沿着 e_3 所定义方向,从过渡态出发的任何位移都会随着时间呈指数性增加,这与实频率振型的情况(对应于正弦振荡)并不相似^①。更直接地说,过渡态是势能表面上在所有方向中除了唯一一个方向都能取得最小值的一个点。而在能量最低(或极小值)点上,表 6.2 中的三个频率就全部都是实数,但是根据对称性,对应于表面平面内运动的两个频率值本应该完全相等,而计算所得频率实际上相差

^① 与振荡的正态振型解不同,式(6.14)仅短时有效,因为它所定义的轨迹不久就从过渡态移开了足够远,从而就使得式(6.13)中截断的泰勒展开式不再有效。

了 0.05THz,这显示出此类计算所能取得的数值精确度。

表 6.2 Ag 原子在 Cu(100)上四重
对称位置和桥位的振动频率

单位:THz

四重对称位(能量极小值点)	桥位(过渡态)
3.36	3.60
1.94	2.17
1.89	1.12i

虽然在上面的例子中只考虑了 3 个自由度,但是其中的主要特点对任何过渡态都是如此。具体说来,如果对于 N 个自由度体系中的过渡态,是由 $(N-1)$ 个实频振型和一个虚频振型所表征的;而一个极小值点具有 N 个实频振型。对式 (6.8) 进行推演的思想可以应用到多维速率表达式(6.10)中,采用谐函数近似则有

$$k_{A \rightarrow B} = \frac{\nu_1 \times \nu_2 \times \cdots \times \nu_N}{\nu_1^+ \times \cdots \times \nu_{N-1}^+} \exp\left(-\frac{\Delta E}{k_B T}\right) \quad (6.15)$$

式中: ΔE 为活化能,它与一维理论中的定义相同; ν_i 为对应于能量最低点的振动频率; ν_j^+ 为对应于过渡态的实振动频率。

等式(6.15)是本节的主要结论。即使不能熟练地掌握其推导过程,也应当将式(6.15)记住。这个结果说明:如果能够得到最低点和过渡态的能量,以及对应于这些能量状态的振动频率,那么就可以计算与该过渡态有关的过程速率。虽然是式(6.15)中的活化能是由最小能量路径所定义的,但这个总的过渡态理论速率也考虑了来自于其他更高能量轨迹的贡献。

对于在 Cu(100)表面上跳动的 Ag 原子,如果在式(6.15)中使用表 6.2 中的数据,则可以得到其前置因子为 1.57×10^{12} Hz。这个数值与较早前从一维 TST 所得到的结果(前置因子为 1.94×10^{12} Hz)是较为近似的。现在可以从图 6.5 中看到:在采用了多维 TST 之后,所能得到的多余信息并不多;特别是当读者回忆起:我们可以完全不经过任何计算而直接使用前置因子的估测值($10^{12} \sim 10^{13}$ Hz)。

与过渡态振动型式有关的最后一点是:能否用这些方法和数据证明所找到的过渡态是正确的?这似乎是一个迂腐而卖弄的问题,但是,在后文中就会明白:相比于寻找能量极小值,在 DFT 计算中确定过渡态要复杂得多。因此,如何能够知道这类计算到什么时候就已经完成了,这是很重要的。对于任何构型而言,在考察其是否为过渡态时,是否存在虚频是一个必要条件。也就是说,如果计算了一个原子构型的振型,并且发现多于一个(或者不存在)虚频,那么,可以非常确定地说:这个构型不是过渡态。此外,对于任意原子集合都可以确定其海塞矩阵,总的来说,在过渡态附近某个区域内就具有一个负的本征值。这也就意味着,还不能用仅仅存在一个虚频来证明所研究的构型就是过渡态。为了确保你所找到的构型一定是过渡态,必须满足以下两个方面:在正态振型中恰好存在一个虚频,并且该构型在能量面上是一个临界点。

6.3 寻找过渡态

在本章已经讨论的所有细节中,有两个需要完全记住的核心概念:

概念 1: DFT 计算可以用来确定涉及能垒的化学过程速率;

概念 2: 这类化学过程速率的一般形式是 $k \sim \exp(-\Delta E/k_B T)$, 其中活化能 ΔE 是能量最低点与该过程过渡态之间的能量差。

如果能够理解这两个概念,那么,在确定化学速率时最为关键的步骤就是找到能量极小值和过渡态。对于上述例子,可以利用体系的对称性来找到过渡态。然而,在许多情形下,这种建立在直觉上的方法就不再那么管用了。相应地,就需要能够在无法提前“知道”答案的前提下,使用数值方法找到过渡态,这是非常重要的。在第 3 章中,我们讨论了如何找到能量极小值,然而,当时所学习的那些迭代方法通常并不能用于寻找过渡态,因为在朝着最小值“下坡”移动时,这些迭代方法会给出一系列原子构型。为了搞明白为什么这些方法(一般来说)不能找到过渡态,可以考虑一下:如果从图 6.2 所示二维表面上邻近过渡态的一点开始采用某种迭代最小化方法,那将

会发生什么。

在确定化学反应速率时,找到过渡态是非常重要的,因此人们已经提出了许多数值方法来完成这一任务。在这些方法中具体选择何种方法,最需要考虑的是以下两点:①平面波 DFT 计算可以从能量面上快速地提供一阶导数信息,但是没有二阶导数信息。这意味着:所谓的模式追随法 (Mode - Following Methods) 通过追随能量最小值的低曲率方向来寻找过渡态,并不能较容易地应用在平面波计算中。这些方法常常是局域化基底量子化学计算中的方法选项,此时也能够得到二阶导数信息。②即便使用 DFT 方法 (或其他量子化学方法) 计算单个构型的能量,通常也需要相当可观的计算资源,因此确定过渡态的方法必须能够在少量的迭代步骤后就收敛于过渡态。

对于平面波 DFT 计算,在寻找过渡态时使用最普遍的方法就是微动弹性带 (Nudged Elastic Band, NEB) 方法。该方法是基于“状态链 (Chain - Of - States)”方法并对其修正后,由 Hannes Jónsson 及其合作者提出的。状态链计算的目的是在两个极小值之间确定最小能量路径 (Minimum Energy Path, MEP)。图 6.5 对这个思路进行了示意,其中的等高线就表示二维的能量面。在该图中,MEP 由连接反应物与产物极小值的光滑曲线所表示,并且这条线通过了过渡态。可以把图中标注为 0 ~ 7 的位置看作是所研究原子体系的不同坐标集合,我们把这些原子位置构型称为不同的图像。对于图中的每种图像,分别计算其能量,这样就需要进行 8 次独立的计算。

读者应该还记得,在任意构型中作用于原子上的力可以写作 $F = -\nabla E(\mathbf{r})$, 其中 \mathbf{r} 是原子体系的一组坐标,那么图 6.6 中的 8 种原子构型图像可以分为两组。图像 0 和 7 位于极小值上,因此这两种图像中的 $F = 0$ 。对于其他所有图像,作用在原子上的力都是非零的。最小能量路径可以用该路径上所有图像的力来进行定义:只有当该路径上任一图像中的力的指向都与该路径完全一致时,才能确定该路径就是 MEP。如果在图 6.6 中的图像 3 上画出由力所定义的矢量,可以看到这些图像并没有定义一个 MEP。

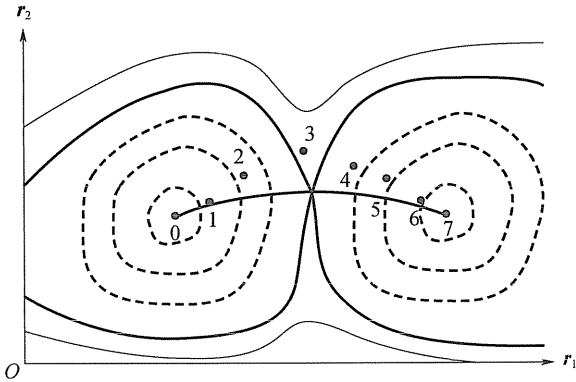


图 6.6 两个极小值点被过渡态所分隔开的二维能量面示意图。黑色等高线的能量值与过渡态相等。过渡态就是两个黑色等高线的交点。虚线(实线)表示能量值低于(高于)过渡态能量的等高线。用一条黑色曲线表示 MEP。阴影圆圈表示弹性带计算所使用的图像(images)位置

6.3.1 弹性带方法

首先介绍一下使用微动弹性带方法寻找能量最低点的思路。该方法是基于这样一个概念:对于沿着 MEP 分布的图像,能够定义一个处在两个极小值之间的最低能量路径;影像应该在空间上均匀地分布在这条路径上。对于一组影像 (r_1, r_2, \dots, r_p) ,可以定义一个目标函数来对这两个概念进行数学表达,即

$$M(r_1, r_2, \dots, r_p) = \sum_{i=1}^{p-1} E(r_i) + \sum_{i=1}^p \frac{K}{2} (r_i - r_{i-1})^2 \quad (6.16)$$

式中: $E(r_i)$ 为第*i*个图像的总能; K 定义了连接两个相邻图像的简谐弹簧(“弹性带”)刚度。目标函数没有包含图像0或*P*的能量,因为这两个图像已经在能量极小值保持固定。需要确定的一点是:对这个目标函数进行最小化,就会移动所有的图像,并使其更加接近于真正的 MEP。

图 6.7 显示了弹性带方法的一个问题:如果在目标函数中“拉伸”一个或多个弹簧的罚函数太小,图像就会倾向于沿下坡方向滑动远离过渡态,而我们原本希望用最高处的图像来近似得到过渡态的精确定位

置。通过找到一个合适的弹簧刚度常数,就可以减少这方面的困难,至少在理论上如此。

弹性带方法的第二个问题是所谓的“切角(Corner Cutting)”,如图 6.8 所示。图中显示了一组图像,虽然这些图像实现了上述目标函数的最小化,但并没有定义一个穿过了过渡态的途径。这个例子中的问题在于:真实的 MEP 是一条更长的途径,而由弹性带方法所确定的途径较短,从而导致对该过程活化能的高估。

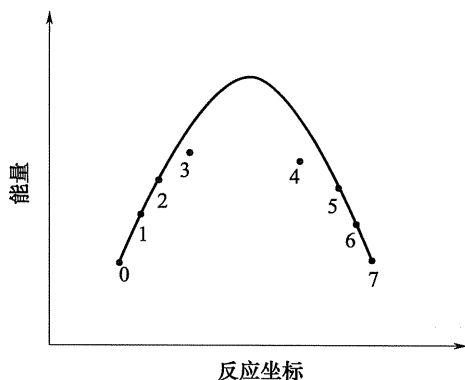


图 6.7 采用弹簧常数很小的弹性带法调整图 6.6 中图像后的典型计算结果。曲线表示在两个能量极小值间真实的 MEP 能量曲线

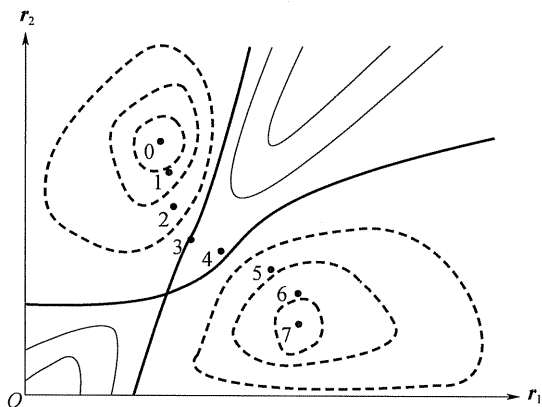


图 6.8 与图 6.6 相类似,但显示了从弹性带计算得到的具有“切角”特征的一组图像

微动弹性带方法

为了改进弹性带方法,人们设计了微动弹性带方法。对于该方法最好的介绍方式就是,该方法考虑了如何调整一组图像并使其向着 MEP 移动。对于每个图像,都可以使用 DFT 方法来计算体系上的作用力: $\mathbf{F}_i = -\nabla E(\mathbf{r}_i)$ 。根据所有图像的当前位置,我们可以估算出这些图像所定义路径的方向。估测这个方向的一个有效方法是:对于图像 i ,将路径方向($\hat{\boldsymbol{\tau}}_i$)定义为沿着与其相邻的两个影像($\mathbf{r}_{i+1} - \mathbf{r}_{i-1}$)连线方向的一个单位矢量。如果除路径方向以外,力的分量在其他方向上都为零的话,即 $\mathbf{F}_i^\perp = \mathbf{F}_i - (\mathbf{F}_i \cdot \hat{\boldsymbol{\tau}}_i) \hat{\boldsymbol{\tau}}_i = 0$,则这些影像就可以满足上述 MEP 的定义。这个描述说明了一个调整这些影像的简单办法,可以使其中的每个影像都沿着 \mathbf{F}_i^\perp 所确定的方向“下坡”。此时,如果还需要把影像之间的简谐弹簧也考虑进来,就需要在定义这个下坡方向时包含有弹簧力。为了做到这一点,一个有效的定义方法是

$$\mathbf{F}_{i,\text{update}} = \mathbf{F}_i^\perp + \mathbf{F}_{i,\text{spring}} = \mathbf{F}_i^\perp + K(|\mathbf{r}_{i+1} - \mathbf{r}_i| - |\mathbf{r}_i - \mathbf{r}_{i-1}|) \quad (6.17)$$

对于使用弹性带方法而言,这个定义是合适的。

我们在上面已经指出:对于由能量面所确定的力,我们只关心这些力在垂直于 MEP 方向上的分量。但对于 NEB 方法而言,所观测的就是那些由弹性带弹簧(Elastic Band Springs)所定义的力,而且只关心沿着 MEP 方向上这些力的分量。也就是说,我们希望弹簧力是唯一能使图像沿着途径均匀排列的力,并且不希望弹簧力把图像从 MEP 附近拉远、拉开。为了做到这一点,进行如下定义,即

$$\mathbf{F}_{i,\text{spring}}^\parallel = (\mathbf{F}_{i,\text{spring}} \cdot \hat{\boldsymbol{\tau}}_i) \hat{\boldsymbol{\tau}}_i \quad (6.18)$$

并且使用下式更新每个图像的位置,即

$$\mathbf{F}_{\text{update}} = \mathbf{F}_i^\perp + \mathbf{F}_{i,\text{spring}}^\parallel \quad (6.19)$$

如果计算中所有的图像都位于一个 MEP 上,则对于每个图像,上式中力的更新值均为零,则整个计算达到收敛。

虽然 NEB 方法的数学定义非常重要,但是可以抛开这些定义的细节,给出该方法的几个主要特点:

(1) NEB 计算的目的是确定一系列原子坐标(图像),从而在能量面上定义连接两个极小值的一个 MEP。

(2) NEB 方法通过使用力的投影方法找到 MEP,在该投影方法中,真实力(由势能所引发)垂直于该微动变形带,而弹簧力平行于该微动变形带。

(3) NEB 是一个迭代最小化方法,因此需要一个 MEP 的初始估值。初始估值与真实 MEP 之间的接近程度对 NEB 计算的收敛速度具有极大影响。

(4) 在表示两个极小值之间的路径时,所使用的图像数量越多,就能给出 MEP 更为精确的描述,但同时也会使计算成本增加,加重计算负担。在 NEB 计算中的每一个迭代过程中,必须对每一个图像进行一个 DFT 计算(除非是位于能量极小值处的端点)。

在基本 NEB 方法的基础上,还有几个变体。这些变体包括:根据当前图像更加精确地估算 MEP 方向的方法,以及沿着 MEP 使用变强度弹簧的技术,从而在靠近过渡态处提高图像密度。一个特别有用的思路是攀升图像微动弹性带方法(Climbing Image Nudged Elastic Band Method),其中对基本的 NEB 方法进行修改,使得在计算收敛后可以精确地得到位于过渡态的图像。读者应该在你自己所使用的计算软件包中,熟悉了解上述这些选项有哪些是你所能获取到的。

6.3.3 初始化 NEB 计算

对于进行 NEB 计算,有一个方面是特别重要的,因而必须对它进行详细讨论;这就是计算的初始化工作。和第 6.1 节和第 6.2 节一样,在这一部分仍然讨论同一个例子:Ag 原子在 Cu(100)上的扩散。为了与上文具有连续性和专一性,仍然使用 NEB 计算考察 8 个图像外加两个端点。在 DFT 计算中,使用具有 4 层原子和 8 个 Cu 原子的 Cu(100) 板块模型。顶部的 2 层 Cu 原子可以在任意方向弛豫,因此每个图像都有 $M = 17$ 个未约束(未固定)原子。这意味着 NEB 计算中的单个图像可以定义为

$$\mathbf{r}_i = (r_{1,i}, r_{2,i}, \dots, r_{3M,i}) \quad (6.20)$$

每个未固定 Cu 原子以及 Ag 原子都具有 3 个自由度。这里, r_1 、 r_2 和 r_3 是单个原子核的空间坐标。为了使用 8 个图像开始一个 NEB 计算, 必须分别确定每个图像中的 $3M$ 个坐标。

我们首先从这个计算的两端(r_0 和 r_9) 开始确定, 可以用标准的能量最小化计算得到这两个图像; 我们把 Ag 原子放置在一个四重对称位, 并允许该原子和表面原子弛豫。通常用来确定 NEB 计算中其他剩余图像的方法是: 在两点之间线性插值。也就是说, 为了在两个固定端点之间生成 8 个图像, 可以对图像 i 的 $3M$ 个坐标定义为

$$r_{j,i} = r_{j,0} + (r_{j,9} - r_{j,0}) \left(\frac{i}{9} \right) \quad (6.21)$$

线性插值所具有的最大优势就是非常易于使用——在使用该方法时, 几乎不需要任何必要工作就可以定义初始图像。我们选择图 6.5 和图 6.7 这两个示意图的另一部分原因就是为了表明图像的线性插值在某些问题上具有良好的近似性, 但对于另一些问题则不然。在图 6.5 中, 线性插值法可以对 MEP 给出一个合理的初始预测值。在图 6.7 中则相反, 真实 MEP 是非常弯曲的, 因此线性插值所给出的 MEP 估测值就较不精确。

可以作出如下推断: 对于我们所考察的 Ag 在 Cu(100) 上的例子, 线性插值法几乎是严格精确的, 因为在图 6.1 所示的俯视图中, MEP 实际上就是两个四重对称位之间的直线。然而, 这个推导并不正确, 因为其忽略了定义 MEP 所需考虑的其他自由度。具体到这个例子中, Ag 原子处于两个极小值时高于 Cu 表面的高度是完全一样的。这意味着在式(6.21)定义的所有图像中, 每一个图像里 Ag 原子都位于 Cu 表面上的相同高度。而实际上, Ag 原子沿着 MEP 在 Cu 表面上的高度是变化的, 可见: 即便是对于这个对称性很好的例子, 线性插值法也并没有给出精确的 MEP。

如图 6.9 所示, NEB 的计算结果可以描述 Ag 在 Cu(100) 上的跳动。在该计算中, NEB 方法进行了 27 次迭代才给出了收敛结果。对于任何一种迭代优化计算, 如果提高了初始预估解的质量, 就能减少达到收敛所需的迭代次数。对于沿着 MEP 构建初始图像的预估值, 下面

的做法值得一试:对于所描述的具体过程,可以根据物理或化学上的直觉来调整这些预估值。在本例中,我们猜测:相比于四重对称位,当 Ag 原子越过桥位时,应在表面上稍微更高一些。为了建立符合这一简单想法的初始图像,采用线性插值法确定原子构型,并且将 Ag 原子的高度在图像 2 和图像 7 中增大了 0.1\AA ,在图像 3 和图像 6 中增大了 0.15\AA ,在图像 4 和图像 5 中增大了 0.2\AA ,这些初始状态的能量如图 6.9 的空心菱形所示。以该初始估值重复进行的 NEB 计算在 24 次迭代后就给出了收敛结果。尽管这只是对计算工作一个相对不大的提高,但这仍然说明了一个概念:只要在简单的线性插值方法中投入少许时间和精力,就可以改善 NEB 计算的收敛速度。

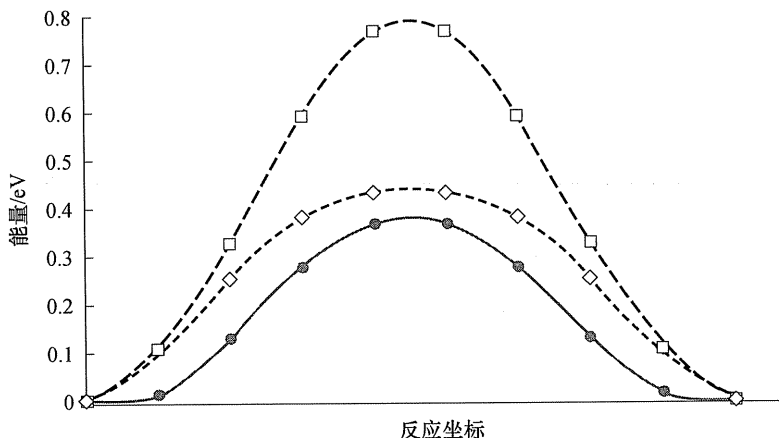


图 6.9 使用 NEB 方法计算 Ag 在 Cu(100) 扩散的能量曲线。空心方框对应于使用线性差值方法所定义的初始图像。空心菱形对应于采用线性插值法定义初始能量,但调整了 Ag 原子在表面上的高度(参见正文所述)。两种计算所得到的最终收敛状态如图中的阴影圆圈所示

比较图 6.9 所显示的由 NEB 计算得到的 MEP,以及图 6.3 所显示的根据对称性得到的 MEP,指出了我们上述 NEB 结果的一个缺点:在 NEB 计算得到的图像中,没有一个图像精确地处在过渡态。图像 4 和图像 5 都比所计算 MEP 的端点高了 0.35eV ,如果把这两个图像的结

果看作是过渡态,那就稍稍低估了实际的活化能。对于这个例子,通过使用奇数个图像,就可以非常简单地避免这个问题,因为从对称性上可以知道:过渡态一定处于反应坐标的中点。但是,我们并没有这么做,这是为了明确显示出在几乎所有实际计算中都很重要的一点:仅仅根据对称性并不能确定过渡态的位置。

在使用过渡态理论时,找到活化能的精确值是非常重要的(可以参见上述图 6.5 的讨论),因此在完成如上所述的 NEB 计算之后,还需要再做一些额外的工作才能找到实际过渡态。完成这个任务至少有两种方法:①已经开发出来的 NEB 的变体方法,不仅可以收敛得到沿着 MEP 分布的若干图像,而且也能够得到过渡态的一个图像。关于这种方法(攀升图像微动弹性带方法)的信息,可以参考本章末尾的扩展阅读。②建立在以下事实上:某些几何优化方法在能量面上收敛得到的临界点就是过渡态或鞍点,前提是如果初始预估值足够接近于该临界点的话。更简单的说,如果所给定的是一个非常接近于过渡态的几何构型,就可以用几何优化方法对该构型进行优化后得到过渡态。

使用 NEB 方法计算所得到 MEP 上的图像,可以作为得到过渡态近似几何构型的一个直接方法。对于 NEB 计算所得到的图像(例如图 6.9 中所示数据),拟合出一条光滑曲线,就可以明确地得到一个能量值最大点。对于上述这个简单例子,由于其对称性,其最大值恰好处于图像 4 和图像 5 之间的中点。如果在拟合曲线上最靠近最大值的这两个图像之间,在这两个图像的反应坐标上使用线性插值法新建一个原子构型,通常就能够得到非常近似于过渡态几何构型的一个预估值。以该原子构型为起点进行几何优化,很快就收敛到过渡态,所对应的活化能为 0.36eV。重复这个过程可以对过渡态构型进行进一步细化修正,对于那些无法使用对称性确定过渡态的计算,这更显得尤为重要。

6.4 寻找正确的过渡态

根据本章所讲述的方法,可以确定出位于两个状态之间最小能量路径上的过渡态几何构型,并且可以使用简谐过渡态理论计算出该过

程的速率。然而,仍然需要考虑一个目前尚未讨论过的问题:怎样才能知道所找到的过渡态就是“正确”的呢?用一个例子来说明这个问题是非常有用的。

对于某些现象而言,过程是非常重要的。例如,晶体生长和外延取向附生就是金属原子扩散通过一个金属表面的过程。如果扩散通过表面的原子与形成该晶体的原子是同种原子,该过程就称为自扩散。通常把自扩散看作是一个原子在表面顶端从一个极小值点跳动到另一个极小值点;在这种描述中,自扩散原子称为吸附原子。在图 6.10 中所示的过程就是反应路径,其中我们假设 Ag 原子是在 Cu(100) 表面上四处移动。我们也可以用品跳反应路径方式描述 Cu 原子在 Cu(100) 表面上的扩散。采用与上述 Ag/Cu(100) 计算相同的方法,可以对 Cu 原子在 Cu(100) 表面上的跳动进行 NEB 计算,预测得到的该过程活化能为 0.57eV。

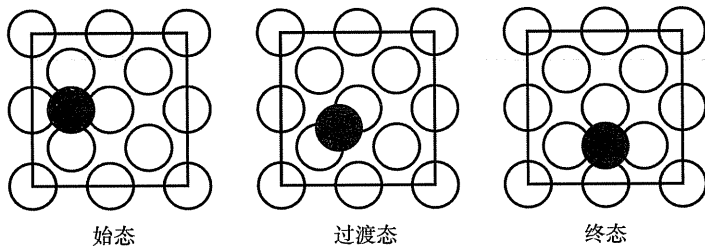


图 6.10 在 Cu(100) 上 Cu 吸附原子从孔穴位跳动到邻近孔穴位的示意图。左图表示跳动之前的吸附原子(灰色),中间图表示过渡态(TS),而右图表示跳动之后的吸附原子

然而,简单的跳动扩散机理并不是原子在 Cu(100) 上移动的唯一方式。图 6.11 显示了另一种可能的扩散过程,该过程称为交换扩散机理。在这种机理中,吸附原子(用深灰色表示)取代了一个原本的表面原子(用浅灰色表示),同时该表面原子跳动到表面上,变成一个吸附原子。这个过程的实际效果就是吸附原子在表面上的两个位置之间移动。这种机理最先是在某些特定 FCC 金属的(110)表面上发现的,但随后在包括 Al 和 Pt 等金属的(100)表面上也观察到了该现象。

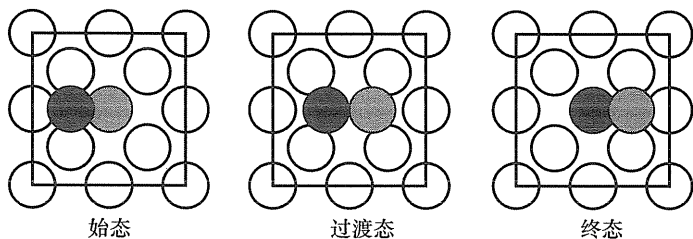


图 6.11 Cu(100) 表面自扩散的交换扩散机理示意图。左图表示处于孔穴位置的 Cu 吸附原子(深灰色), 以及一个浅灰色的表面原子。在中间图表示的过渡态中, 两个 Cu 原子都位于表面之上。右图表示交换之后的结果, 初始的吸附原子变为表面原子, 而初始的表面原子变为吸附原子

如果想要理解 Cu 吸附原子在 Cu(100) 上的扩散过程, 就需要比较图 6.10 和图 6.11 中两种扩散机理的扩散速率。关于交换跳动的发生可能性, 前面内容所讲述的 NEB 计算能告诉我们什么呢? 答案很简单: 没有半点说明! 因为在进行 NEB 计算时, 先在图 6.10 中的起点和顶点之间进行插值, 从而建立初始路径, 所得结果无法给出有关图 6.11 所示路径的任何提示, 就更不用说给出这些路径的详细信息了。可以用更加通用的语言描述为: NEB 计算是局域的, 而非全局的优化计算, 因此 NEB 计算可以得到特定过渡态的精确信息, 但该方法不能确定对应于同一个起始态是否还存在着其他过渡态。

为了弄清某个特定过程是否存在多个过渡态, 就必须对一系列可能的活化过程分别进行 NEB 计算。对于考察 Cu 在 Cu(100) 上交换跳动, 使用 NEB 方法是非常合理的。在如图 6.11 所示的始态和终态之间使用线性插值法建立了 7 个图像(外加端点), 对其进行计算。根据对称性, 这个例子的中心位置图像收敛于过渡态。对图 6.10 中的跳动机理进行了类似的计算, 这些结果都显示于图 6.12 中。这些计算所得出的一个主要结果就是: 交换扩散机理的活化能为 0.67 eV, 高于直接跳动机理的能垒(0.57 eV)。如果假设这两个过程的 Arrhenius 前置因子是近似相等的, 那这两个能垒就意味着: 在 300 K 下, 直接跳动大约要比交换跳动快 50 倍。相比于直接跳动而言, 虽然交换跳动中吸附原

子在表面上的移动距离更远(参见图 6.10 和图 6.11),但上述结果表明:交换跳动对 Cu 吸附原子在 Cu(100)上扩散的实际贡献较小。

那么,我们现在是否已经对 Cu 吸附原子在 Cu(100)上扩散给出了完整描述?还没有,也许还有其他我们没有考虑到的活化过程,而且其扩散速率可能比直接跳动的扩散速率还要大得多。例如,如果发生多个原子参与的交换扩散,就可能造成某个吸附原子完全移动到另一个位置上,而这个位置要比图 6.11 所示“简单”交换扩散中偏离初始吸附原子更远。在任何情形下,都像这样考虑一系列可能情形是非常重要的,也就是说,必须使用物理或化学上的直觉来指导一个过程的计算工作。即使当我们已经采用 NEB 计算方法分析了一系列可能过程后,严格地说,也只能把这些计算中最低的活化能当作真实活化能的上限。

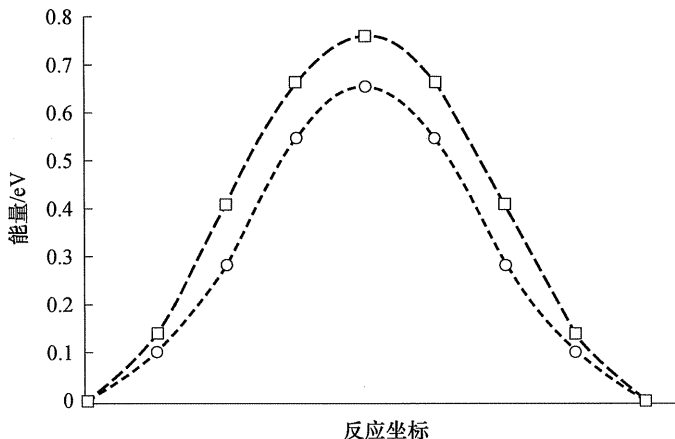


图 6.12 对于 Cu 在 Cu(100)上的自扩散,比较交换机理(方框)和吸附原子直接跳动机理(圆圈)的 NEB 结果

6.5 将单个速率组合为总动力学

现在已经学习了如何使用 DFT 方法来计算单个活化过程的速率。虽然这个信息极其有用,但仍无法用来充分完整地描述许多实际的物

理问题。在许多情形下,系统会随着时间演变,在极小值点之间经历许多次跳动。例如,在金属氧化物表面上,金属原子催化团簇的形成过程就涉及到表面上许多金属原子的跳动。这些团簇常常在氧化物表面的缺陷处形核,这个形核过程就是无缺陷表面区域以及邻近缺陷区域内原子跳动的综合结果。该问题的一个特点就是:它是原子在复杂能量表面上的长时间移动,而该能量表面是由我们感兴趣的许多个极小值所确定的。

下面给出一个简化的例子。把有关 Ag 在 Cu(100) 上跳动的例子进行扩展,将表面替换为被少量浓度的 Pd 原子随机掺杂后的 Cu(100)。为了与上述情形进行一个类比,可以将 Pd 原子看作是表面上的“缺陷”,图 6.13 示意了这个假想表面。如果仅仅根据每个位置周围的 4 个原子,对表面上的四重对称位进行分类,则只存在两类表面位置^①。根据同样的逻辑,表面上 Ag 原子的局部跳动可以使用 4 个不同速率(如图 6.13 所示的 k_1, \dots, k_4) 来描述。根据本章已经学习过的 DFT 方法,可以比较容易地确定其中的每个速率。

如果仅仅知道了图 6.13 中所列出的速率,仍然还有许多无法轻易回答的物理问题。例如,一旦 Ag 原子沉积在表面上,那么在它随后到达靠近 Pd 表面原子位置之前,它会在原先的位置上保持多长时间呢(在平均意义上)? 表面上的一个 Ag 原子在单位时间内会逗留过多少个不同的 Pd 表面位置呢? Ag 原子在该表面上总的扩散系数又是怎样的? 为了回答这些问题,需要使用一个工具来描述表面上一组 Ag 原子随时间的演变规律。

类似于这样的有关于体系动力学的问题,可以使用一个称作动态蒙特卡罗方法(kinetic Monte Carlo, kMC) 来描述。该方法的思路非常简单直接。对于当前给定的原子构型,如果知道了所有可能发生的过程及其速率,就可以随机选择一个与这些速率相一致的事件。通过重复这一过程,就可以模拟体系的时间演变。

^① 这个近似方法没有包括所有可能发生的情况。例如, Pd 原子可能是表面位置的次邻原子,一般说来,这也会影响该位置上吸附原子的能量。

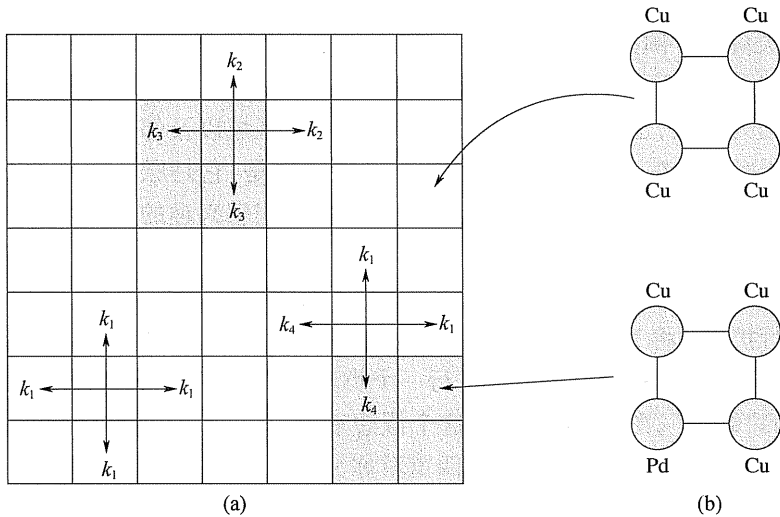


图 6.13 在 Pd 掺杂 Cu(100) 上 Ag 原子扩散的晶格模型示意图。右图是四个表面原子所定义的四重对称位。在左侧的表面示意图中包含有两个完全分隔开的 Pd 原子。左图中,两个 Pd 原子均位于灰色方格的中心

描述 kMC 方法的最简单方式就是对于上面所定义的体系, 给出一个该算法的具体例子。假设有一个大面积的表面, 该表面中的四重对称位可以分为两类(如图 6.13 所示), 并对其进行计算, 得到所有的相关速率。将所有速率中的最大值标记为 k_{\max} 。在初始时, 在表面上的随机位置上沉积 N 个 Ag 原子, 则可将该 kMC 算法描述为:

- (1) 随机选择 1 个 Ag 原子。
- (2) 随机选择跳动方向(向上、向下、向左或向右)。根据速率数据表, 选择与这种跳动相对应的速率 k_{hop} 。
- (3) 在 0 和 1 之间选择一个随机数 ε 。
- (4) 如果 $\varepsilon < k_{\max}/k_{\text{hop}}$, 使该原子在选定的方向上移动, 否则不移动该原子。
- (5) 无论上一步的输出结果是怎样的, 将时间增加一个变量值 $\Delta t = 1/(4Nk_{\max})$ 。

(6) 返回到步骤(1)。

通过比较和考察两个简单情形,可以验证:该算法能够生成一个符合物理事实的事件顺序。第一种情形是,如果表面是纯 Cu,那么每个跳动速率都具有相同的速率值 k ,这样就会一直进行算法第(4)步的跳动运算。此时,任何一个原子在两次跳动之间的平均时间就是 $\Delta t = \frac{1}{4}k$ 。

这意味着这些原子在表面上跳动的速率是符合公认准则的。另一个情形是,如果一个原子分别以不同的局部速率(k_{high} 和 k_{low})沿着两个方向跳出一个位置,则沿着较快与较慢方向的跳动速率之比为 $k_{\text{high}}/k_{\text{low}}$ 。相同的论断也可以用在两个原子吸附于不同的位置上。这些概念表明:不同的过程发生时都具有符合公认事实的速率比,且该过程的绝对速率也是符合事实的——换句话说,该算法定义了符合物理事实的动力学。实质上,kMC 模拟是随机猜测的,因此并不能期望两次模拟给出相同的事件顺序,即使二者都从相同的条件下开始。通常是必须对多个模拟的演变过程进行平均化,才能得到可信的平均信息。

我们应该知道,kMC 模拟有两方面的缺陷。第一个缺陷与算法的数值效率有关,如果速率列表中包含了很大范围的绝对速率值,那么对于最慢的速率而言,在该算法第(4)步中的比值 $k_{\text{high}}/k_{\text{low}}$ 就会很小。例如,如果最慢和最快速率的活化能分别是 1.0eV 和 0.2eV,且这两个过程的速率都具有相同的前置因子,则其比值在室温下为约 10^{-13} ,而在 300°C 下为约 10^{-7} 。这意味着该算法的效率很差;在经过几百万次迭代后,体系的状态并不会发生什么改变。若干年前就已经发展出了更加复杂的 kMC 算法,早就解决了这个问题,有关这些方法的文献可以参考本章最后的扩展阅读部分。

kMC 方法的第二个缺陷则更为严重。只有当该方法所使用的速率列表充分且完备时,kMC 结果才是符合事实的。具体说来,可以回顾一下有关 Ag 原子扩散的例子,如果表面上的 Ag 原子是隔离态的,则上述算法就已经是严格定义的了;但是,如果两个 Ag 原子处于两个相邻表面位置时,又会发生什么呢?需要在上述晶格模型中加入信息来定义在这种情形中会发生什么。从物理的观点来看,并不希望一个

Ag 原子跳动到已经被另一个 Ag 原子占据的位置上。当原子跳入的位置上已经有另一个原子时,只要简单地定义其速率为零就可以解决该问题。但是,一个 Ag 原子会以多快的速度跳动着远离邻近原子对呢?它们的速率很可能与表面上隔离态 Ag 原子的速率并不相同(并且更低),因为表面上的金属吸附原子通常更倾向于相互组合在一起。如果花些时间用 DFT 方法计算这些速率,就可以在上述 kMC 算法中把这些速率考虑进去。从图 6.13 可见,两个邻近 Ag 原子可以在一系列不同位置上共存。为了更加充分完备地来描述,就需要对其中的每一种情形分别计算其速率。类似地,当表面随机掺杂有 Pd 原子时,就存在表面层中两个 Pd 原子相互邻近的位置。为了解释这种情况(以及表面中更大的 Pd 原子团簇),就应当对每种可能的表面原子排列,分别考察其所有的表面过程速率。

即便是用 DFT 方法来处理来完成了上述所有过程,我们所建立的模型仍然是不完备的。在我们的 kMC 模拟中,随着时间延长,就可能在表面上形成 Ag 原子团簇。假设这些团簇中的每个 Ag 原子都位于表面的四重对称位置上,但是自然界中的实际情形并非严格符合该假设。对于表面上三维的微小金属团簇,在自然界中并不常见。在我们的 kMC 计算方案中对这个情形进行详细处理是非常困难的——必须考虑 Ag 原子在已存在的 Ag 团簇中,向上或向下攀升的所有可能过程,并计算它们的速率,然后还需要调整我们的算法允许其三维移动。虽然这些复杂问题并没有阻碍 kMC 模拟的进一步使用,但是这些问题的确意味着:应该始终根据该模型结构可能(或不可能)发生的现象,来对这些模拟结果进行解读。

6.6 量子效应及其他复杂问题

本章始终关注的都是与活化能垒有关的化学过程。只要该能垒的高度相对于 $k_B T$ 足够大,过渡态理论就能对总过程速率给出一个有益描述。但是,有些情形下并不适合用过过渡态理论,为了对此进行说明,需要在本节简述两个常见例子。

6.6.1 高温/低能垒

过渡态理论不能精确描述的第一种情形就是在高温下活化能近似于 $k_B T$ (甚至更小)。可用上述位于 Cu(100) 表面上的 Ag 原子为例来加以说明, 一个能够使 Ag 原子处于该表面上的方法是: 在包含由 Cu(100) 晶体的真空室中使 Ag 原子蒸发。Ag 原子从气相粘附在晶体表面这一过程不存在任何障碍。相比于气相, Ag 原子在能量上更倾向于存在于表面上, 而且可以经由能量单调降低的途径, 从气相移动到表面。上述有关过渡态的公式化阐述并不能描述这一情形。在化学动力学的研究中, 已经完全开发出来了适用于此类情形的理论——本章末尾的扩展阅读部分就这些内容给出部分信息。

6.6.2 量子隧穿

在极低的温度下, 还会产生过渡态理论另一种颇为不同的缺陷。建立过渡态理论的基本思想是: 在能量极小值之间, 体系得到能量并越过一个能垒时所发生的过渡转变。然而, 量子化学中还存在另一种过渡: 即便粒子能量比越过能垒时所需的能量低得多, 粒子也能穿过能垒; 这就是量子隧穿 (Quantum Tunneling) 现象。随着运动粒子的质量增加, 隧穿速率急剧降低, 因此对于与氢有关的过渡转变, 隧穿现象是非常重要的。

幸运的是, 从简谐过渡态理论出发, 可以较为简单地估测出量子隧穿对所考察的问题是否存在严重影响。应用多维过渡态理论 (式(6.15)), 需要找到体系在能量极小值 $A(\nu_1, \nu_1, \dots, \nu_N)$ 和过渡态 $(\nu_1^\ddagger, \dots, \nu_{N-1}^\ddagger)$ 位置上的振动频率。根据这些频率, 可以得到经过零点能校正的活化能为

$$\Delta E_{zp} = \left(E^\ddagger + \sum_{j=1}^{N-1} \frac{h\nu_j^\ddagger}{2} \right) - \left(E_A + \sum_{i=1}^N \frac{h\nu_i}{2} \right) \quad (6.22)$$

我们已经知道: 过渡态的一个虚频所具有的形式是 $i\omega_{1m}^\ddagger$ 。将这些量合并后, 就形成了横跨温度 (Crossover Temperature) T_c , 其表达式为

$$T_c = \frac{h\nu_{\text{lm}}^\dagger E_{\text{zp}}/k_B}{2\pi E_{\text{zp}} - h\nu_{\text{lm}}^\dagger \ln 2} \quad (6.23)$$

在该横跨温度以下,隧穿效应就会对总速率产生显著影响;而高于该温度,则可以忽略隧穿效应。Fermann 和 Auerbach 已经开发出了简谐过渡态理论的校正,可以根据经典理论所得到的信息,估测 T_c 温度以下的隧穿速率^[1]。

6.6.3 零点能

对于我们前面所使用的过渡态理论,最后一个复杂难点是:该理论建立在对体系能量经典描述的基础上。但是,正如我们在第 5.4 节所讨论的,应该使用经典的能量极小值和零点能校正,才能更加准确地定义原子构型的能量极小值。将这一思想引入到过渡态理论中也并不很困难。在考虑了零点能之后,式(6.15)调整后的完整表达式应该写为

$$k_{A \rightarrow B} = \frac{\prod_{i=1}^N f(h\nu_i/2k_B T)}{\prod_{j=1}^{N-1} f(h\nu_j^\dagger/2k_B T)} \frac{\nu_1 \times \nu_2 \times \cdots \times \nu_N}{\nu_1^\dagger \times \cdots \times \nu_{N-1}^\dagger} \exp\left(-\frac{\Delta E}{k_B T}\right) \quad (6.24)$$

式中: $f(x) = \sinh(x)/x$ 。其实,该表达式并不像它看起来那样难以理解。在高温下, $f(h\nu/2k_B T) \rightarrow 1$, 于是式(6.24)就恰好简化为我们之前所使用的那个经典表达式。在低温下,其极限结果则更为简单,即

$$k_{A \rightarrow B} = \frac{k_B T}{h} \exp\left(-\frac{\Delta E_{\text{zp}}}{k_B T}\right) \quad (6.25)$$

式中: ΔE_{zp} 是零点能校正后的活化能[由式(6.22)所定义]。在中间温度范围内,式(6.24)平滑地连接这两种极端情形。很多情形下并不涉及 H 原子,在经典结果和零点能校正结果之间的差值就非常小,也就不重要了。

根据式(6.25)可以得到一个有意思的结果:在这个 Arrhenius 表达式中,前置因子与体系的振动频率是彼此无关的。在第 6.1 节,我们已经提到这个前置因子可以用“通常典型的”振动频率($10^{12} \sim 10^{13} \text{ s}^{-1}$)来估算。式(6.25)也支持这个估算值:当 T 在 $100 \sim 500\text{K}$ 的范围变化

时, $k_B T/h$ 从 $2 \times 10^{12} \text{ s}^{-1}$ 变化为 $2 \times 10^{13} \text{ s}^{-1}$ 。

练 习

1. 当 Ag 原子在 Cu(111) 上两个相邻三重对称位置间扩散时, 使用 NEB 方法确定其活化能。在计算时需要注意, 两端点的能量并不严格相等, 因为 Cu(111) 表面上的 FCC 和 HCP 位置并不完全相同。计算该跳动过程的频率。

2. 确定一个电中性空位在体相 Si 中迁移时的活化能, 其模型是在体相 Si 材料含有 64 个原子的超晶胞中去掉一个 Si 原子。在开始进行 NEB 计算之前, 仔细思考一下: 在始态和终态之间, 如果某些原子移动并穿过了超晶胞边界, 这是否会产生什么影响。根据所计算的活化能, 以及过渡态前置因子的标准估算值, 预测电中性空位在 Si 中的跳动速率。

3. 氢原子可以容易地溶解到体相 Pd 中, 并在其中处于六重对称八面体或四重对称四面体间隙位置。如果体相 Pd 中的氢原子在八面体和四面体位置之间跳动, 分别计算其经典的和零点能校正后的活化能。在计算活化能时, 允许超晶胞中所有原子弛豫; 但是, 在估算振动频率时, 将所有金属原子约束固定。估测在什么温度以下, 隧穿效应就会对 H 原子在两个间隙位置间的跳动产生重要影响。

4. 计算 Pd(100) 表面上 Pd 吸附原子的扩散活化能, 其跳动过程分别为: 在表面两个四重对称位置之间的直接跳动, 以及与表面顶层的 Pd 原子进行协调置换。在开始计算之前, 考虑一下表面板块模型需要多大才能描述这两个过程。上述两个过程中, 预测室温下哪个过程主导了 Pt 吸附原子的扩展?

5. 对于初始化 NEB 计算而言, 线性差值并不总是有效的, 本练习就是为究其原因而给出一个例子。考察气相中的 HCN 分子, 该分子可以重新排列为 CNH。优化 HCN 和 CNH 的结构, 然后在这两个结构之间, 针对其键长采用线性差值法, 建立一些中间结构状态。为什么该过程中所定义的中间结构不具有化学合理性? 建立一系列具有化学合理

性的初始预估图像,并用这些图像进行 NEB 计算,估算该分子异构反应(Molecular Isomerization Reaction)的活化能。

参 考 文 献

- [1] Fermann J T, Auerbach S. Modeling proton mobility in acidic zeolite clusters: II. Room temperature tunneling effects from semiclassical theory. *J. Chem. Phys.* **112**,2000;6787.

扩 展 阅 读

化学动力学领域是非常广博的,并且已经相当完备。如果已知(或者可以计算得到)所有参加化学反应原子的全能量表面,就可以对那些不适用于简单过渡态理论的情形,使用复杂的反应速率理论给出精确的反应速率信息。该领域的经典教材是:Laidler K J. *Chemical Kinetics*. 3rd ed. Prentice Hall, New York, 1987. 该领域最近的著作是:Chorkendorff I, Niemantsverdriet J W. *Concepts of Modern Catalysis and Kinetics*. 2nd ed. Wiley - VCH, Weinheim, 2007. 当然,该领域还有其他许多著作。

为了弄清统计力学和 TST 之间的联系,请参阅:Chandler D. *Introduction to Modern Statistical Mechanics*. Oxford University Press, Oxford, UK, 1987 以及 Frenkel D, Smit B. *Understanding Molecular Simulation: From Algorithms to Applications*. 2nd ed. Academic, San Diego, 2002。

对于 NEB 方法的说明,以及与其他(确定过渡态时未使用海塞矩阵的)状态链方法之间的比较,请参阅:Sheppard D, Terrell R, Henkelman G. *J. Chem. Phys.* **128**, 2008:134106。关于攀升图像 NEB 方法,可以参阅:Henkelman G, Uberuaga B P, Jónsson H. *J. Chem. Phys.* **113**, 2000:9901 - 9904。有关于 NEB 方法的改进,其中包含了对 MEP 斜率的较好估算,请参阅:Henkelman G, Jónsson H. *J. Chem. Phys.* **113**, 2000: 9978 - 9985。

关于表面原子扩散的交换扩散机理,详细信息请参阅:Anderson M

L, D'Amato M J, Feibelman P J, et al. Phys. Rev. Lett. 90, 2003: 126102 和 Feibelman P J, Stumpf R. Phys. Rev. B. 59, 1999: 5892 – 5897。对于体相材料中原子扩散和掺杂质团簇化, 必须考虑多种机理, 有关内容请参阅: Uberuaga B P, Henkelman G, Jonsson H, et al. Phys. Stat. Sol. B. 233, 2002: 24 – 30。

我们所描述的动态蒙特卡罗 (kMC) 算法在推导和使用上都是简单而直接的, 但并不一定是现有各种类似算法中效率最高的。有关于不同 kMC 算法的特点和效率, 详细综述可参阅: Reese J S, Raimondeau S, Vlachos D G. J. Comput. Phys. 173, 2001: 302 – 321。

在 DFT 方法中合并使用动态蒙特卡罗方法, 对金属催化剂的活化位置进行建模的方法, 请参见综述: Neurock M. J. Catalysis. 216, 2003: 73 – 88。

附录 计算细节

本章的所有计算针对以下两个体系: 在 Cu(100) 表面上 Ag 吸附原子从孔穴位扩散到桥位; Cu 在 Cu(100) 上的自扩散。对于这些计算, 都使用了非对称的四层 Cu 板块模型, 底面的两层原子固定不动。按照 DFT 优化后的 Cu 晶格常数, 固定了超晶胞在表面晶面内的尺寸。该超晶胞是每层含有 8 个原子的 $c(4 \times 4)$ 表面单胞。真空层大于 28\AA 。所有计算都采用 PBE GGA 泛函, 截断能为 380eV 。在能量最小化和微动弹性带计算中, 采用共轭梯度算法使所有未约束离子发生弛豫。采用了宽度值为 0.1eV 的 Methfessel – Paxton Smearing 方法, 并在表面的法线方向上进行了偶极子校正。本章每节中具体计算的更多细节列于下面。

6.1 节 在 Cu(100) 孔穴位和桥位上 Ag 吸附原子的最小化运算中, 除了上述参数, 所使用的 k 点网格是 $4 \times 4 \times 1$ 。

6.2 节 Ag 吸附原子在 Cu(100) 孔穴位和桥位上的频率计算中, 使用 $4 \times 4 \times 1$ 的 k 点网格。使用有限差分计算确定频率, 其中所使用的位移为 0.04\AA 。

6.3 节 对于 Ag 吸附原子经由最小能量路径,从孔穴位经过桥位跳动到另一个孔穴位的微动弹性带计算,使用了 8 个图像外加两个端点,并对每个图像都使用了 $4 \times 4 \times 1$ 的 k 点网格。

6.4 节 对于 Cu 表面扩散的 NEB 计算,使用了 7 个图像,每个图像的 k 点网格都是 $2 \times 2 \times 1$ 。

第 7 章 基于 *Ab Initio* 热力学的平衡相图

在第 2 章中我们主要关注的计算任务就是使用 DFT 计算预测体相铜的晶体结构。一旦读者掌握了定义不同晶体结构的细节,并能用 DFT 方法对其进行能量最小化,那么这个任务就比较简单了。具体说来,就是在许多可能的晶体结构(简单立方、FCC 和 HCP)中,对体相 Cu 的能量进行最小化,然后比较每一种结构中单个原子的能量(能量/原子)。该计算给出了与实验观测一致的预测:铜具有 FCC 晶体结构。

更严格地说来,我们前面所做计算是当压强为零时对 Cu 晶体结构的预测。如何把这些结果扩展到更符合实际的条件下,这肯定是非常重要的! 在第 2.5 节中我们讲述了如何在非零压强下,通过计算材料内聚能与体积的函数关系,比较几种可能结构的相对稳定性。这一概念对许多应用都是非常重要的,比如地球物理学中所涉及的极高压强。在第 1.2 节中有关地球形成的少许说明就是该情形的一个具体例子。

想象一下,当你正在阅读此段文字时,就在你的旁边有一小块 Cu。我们能否用 DFT 计算预测这个材料的结构呢? 在第 2 章中的 DFT 计算与此时的 Cu 块之间,最主要区别就是:后者是由几十亿个气体分子(主要是 N_2 和 O_2)包围着,而这些分子很可能会与该金属材料发生反应。如果假设 N_2 不参与反应,就可以对上述讨论进行简化,也就是:Cu 在室温下被压强大致为 0.2 大气压的 O_2 包围着。这肯定有可能使铜氧化;一种常见的矿物氧化铜形式是 Cu_2O 。如果想要预测 Cu 在常温常压条件下的特性,就需要能够确定:在与气态 O_2 达到平衡时,Cu 和 Cu_2O 谁更加稳定?

如果能够像预测纯铜的晶体结构那样,采用相同的方法预测 Cu 和 Cu_2O 的相对稳定性,这将是非常吸引人的。这个方法似乎是:先最

小化每种材料的能量,然后确定谁的能量最低。这个思路中存在有两个基本性的瑕疵:首先,我们如何对不同材料的能量进行标准化?当对体相 Cu 进行考察时,使用“能量/原子”是完全合理的;但是,当比较化学成分不同的几种材料时,这个标准化处理方法就不适合了。其次,这个方法没有包含任何有关 O_2 温度和压强的信息,而这正是我们力图所要描述的。如果当 O_2 压强从 0.2 大气压降低到 2×10^{-6} 大气压时,Cu 的平衡态就一定会发生改变,那么就需要使用某一方法描述这个过程。

如果我们只对体相铜及其氧化物感兴趣,就完全不需要求助于 DFT 计算。在冶金工程的许多具体实践中,体相金属及其氧化物的相对稳定性是极其重要的。正因如此,这类信息早已被广泛表征,并整理成图表。这些信息(以及金属硫化物的类似信息)已经被整理成所谓的 Ellingham 图表,而这些图表在许多地方都能找到。我们选取这些材料作为初始示例,是因为读者可能已经对这个情形具备了物理上的直觉。本章的主要内容就是:对应于上面所提出的物理问题,可以用 DFT 计算描述不同种类的相稳定性。在第 7.1 节我们主要讨论:如何对体相氧化物做到这一点。在第 7.2 节,我们考察一些 DFT 计算能够给出相稳定性信息的例子,这些信息在技术上也是很有意义的,但如果从实验上建立这些信息则是相当困难的。

7.1 体相金属氧化物的稳定性

在使用 DFT 计算描述相稳定性的初始示例中,我们仍将继续讨论体相金属氧化物。在本章中,我们所感兴趣的是金属 M 在温度 T 下,与分压为 P_{O_2} 的气态 O_2 平衡时的热力学稳定性。假设我们已知该金属及其氧化物的一系列可能晶体结构,那么对于我们所考察的 Cu 和 Cu_2O ,就可以均按其实验已知的晶体结构来进行计算。此外,实验表明 Ag 和 Ag_2O 分别与上述含铜材料具有相同的结构,因此我们同时考察了银和铜的氧化物。

在热力学上,我们希望知道:在特定条件下,哪种材料会使包含有气态 O_2 和该种固体的体系自由能最低。一个有效方法就是定义与每

种晶体结构相联系的巨势(Grand Potential)。对于含有 N_M 个金属原子和 N_O 个氧原子的金属氧化物,巨势可以定义为

$$\Omega(T, \mu_O, \mu_M) = E(N_M, N_O) - TS - \mu_O N_O - \mu_M N_M \quad (7.1)$$

式中: E 是金属氧化物的内能(Internal Energy); S 是材料的熵; μ_a 是元素 a 的化学势。如果有一系列不同材料($i = 1, 2, 3, \dots$),就可以用这个表达式定义每种材料的巨势 Ω_i 。

对于巨势中的内能,可以简单看作是該材料由 DFT 计算得到的总能。当我们把 DFT 能量标准化,使每个 DFT 计算所描述的材料都具有相同的金属原子个数,此时再来比较不同材料的巨势就显得非常合理了。这样等式(7.1)就可以重新写为

$$\Omega_i(T, \mu_O) = E_i - TS_i - \mu_O N_{O,i} - \Omega^M \quad (7.2)$$

式中: Ω^M 是一个附加常数,对于每一种材料都是相同的。可以直接根据气态 O_2 的温度和压强来确定氧的化学势。如果把 O_2 看作是理想气体,则有

$$\mu_{O_2} = \mu_{O_2}^o(T, p^o) + kT \ln(p_{O_2}/p_{O_2}^o) \quad (7.3)$$

式中:上标 o 表示参考压强,通常取为 1 大气压。分子态氧和原子态氧化学势之间的关系是

$$\mu_O = \frac{1}{2} \mu_{O_2} \quad (7.4)$$

等式(7.2)所定义的巨势具有一个显著而简单的特点:对于巨势最低的材料,就使气态 O_2 与該固体的组合模型具有最低自由能。换言之,一旦我们计算得到了如式(7.2)所示的巨势,则热力学稳定态就是具有最小巨势的那个状态。

当比较不同的晶态固体时,不同结构在内能上的差值,通常要比其熵的差值大得多。这个结果表明:对于所考虑的不同晶体结构,我们可以把式(7.2)中熵的贡献近似看作常数^①,从而使得我们所要计算的巨

① 如果在描述熵的贡献时,需要更大的精确度,可以计算每种材料的振动态密度(如第 5.5 节所述),从而计算得到由于晶格振动所导致的熵。关于这一思想的更多细节,请参见第 5.5 节中的相关参考文献。

势可以进一步表达为

$$\Omega_i(T, \mu_{\text{O}}) = E_i - \mu_{\text{O}} N_{\text{O},i} - \Omega_s^{\text{M}} \quad (7.5)$$

式中: Ω_s^{M} 是一个对于每种材料都相同的附加常数。因为我们只关心哪种材料具有最低的巨势,从而可以设 $\Omega_s^{\text{M}} = 0$,而这并不会对其通用性造成任何影响。如果现在来看一下等式(7.3)、(7.4)和(7.5),可以看到:对于找到金属氧化物的最稳定相这一任务,我们已经完成了其数学表达方式,而其中只涉及到所研究的金属和金属氧化物的 DFT 能量以及气态 O_2 的压强和温度。

将式(7.5)分别应用于 Cu 和 Ag,其结果如图 7.1 所示。该示意图显示了 M 和 M_2O 的相对稳定性。对于 $\text{M} = \text{Cu}$ 和 $\text{M} = \text{Ag}$,虽然其数值结果在具体细节上有所不同,但是二者的基本内容是完全一致的。图 7.1(a) 中虚线表示由式(7.5)所定义的巨势。对于纯金属 M,其斜率为零,此时 $N_{\text{O}} = 0$;而金属氧化物的斜率为负值。图 7.1(a) 中实线表示当氧的化学势变化时,具有最低巨势的结构。在化学势低于一定值时,体相金属在热力学上是更为优先的;当化学势足够高时, M_2O 更加优先。如果把氧的化学势用压强和温度来描述的话,在物理意义上也很容易理解这个结果。这就得到了如图 7.1(b) 所示的相图。该相

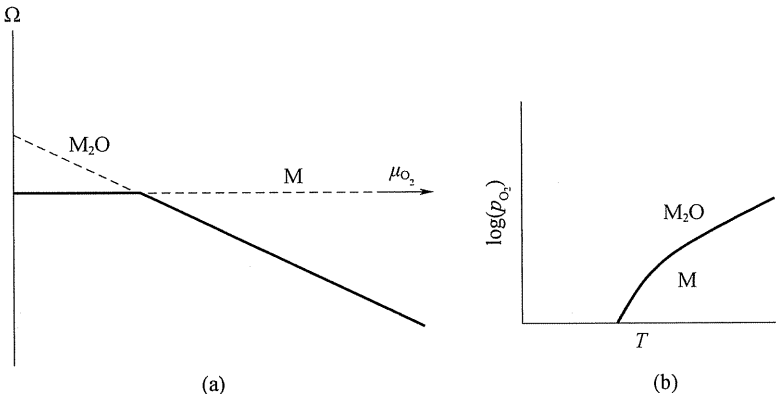


图 7.1 Cu 和 Ag 氧化过程示意图

(a) Cu 和 Ag 氧化过程的巨势示意图; (b) 由该巨势图所确定的相图。在相图中曲线所分割的两个区域中,金属或氧化物分别是这两个区域的热力学稳定材料

图也符合众所周知的金属氧化和还原趋势,即:将金属氧化物放置在足够高的温度下,或者是在足够低的 O_2 分压下,则其可以被还原为金属。图 7.1 的一个特点就是:在某些温度和 O_2 压强的组合条件下,图示的两种材料在热力学上都是有可能存在的。其实并不一定是这样。图 7.2 显示了上述巨势图的一个扩展情形,其中用粗线突出显示了已获确认的热力学稳定材料,并增加了第三种氧化物(化学计量式为 M_3O_2)的有关信息。在本例中,新添加氧化物的巨势线总是在至少一种材料的上方,这意味着 M_3O_2 根本不可能是热力学稳定的。这也就意味着:即使是在 M 和 M_2O 之外还考虑了 M_3O_2 ,图 7.1(b) 所示的相图仍然是正确的。

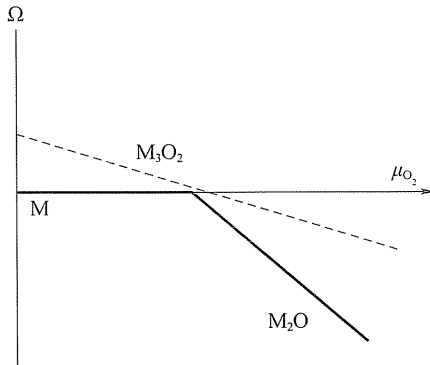


图 7.2 从图 7.1 扩展后的巨势图(添加了一种金属氧化物,但是该物质在任何条件下都不是热力学稳定的)

当使用这些方法时,必须谨记的一个重要忠告是:类似于图 7.1 和图 7.2 的示意图只能告诉我们材料的相对稳定性,而且是在我们已经进行了 DFT 计算的这些材料范围内。这一情形非常类似于第 2 章中我们“预测”晶体结构时所讲过的内容。如果我们的 DFT 计算没有包含某一个(或者多个)与真实材料有关的晶体结构,那么上述方法所给出的相图就不可能是完全完备的。

为了定量地描述图 7.1 中的相图,我们必须返回到式(7.3),并更

加完整地确定化学势。对于理想气体,化学势可以从统计力学中严格地推导得到。对于 O_2 的理想气体化学势,可以定义为

$$\mu_{O_2} = E_{O_2}^{\text{total}} + \tilde{\mu}_{O_2}(T, p^\circ) + kT \ln(p/p^\circ) \quad (7.6)$$

式中: $E_{O_2}^{\text{total}}$ 是 $T = 0K$ 时一个隔离态 O_2 分子的总能量。如果忽略零点能,可以用一个简单的 DFT 计算得到该隔离态分子的能量^①。式(7.6)中右侧第二项是在参考压强下,在 $T = 0K$ 和所研究温度下 O_2 的化学势差值。对于 O_2 和许多其他气体,可以使用 NIST - JANAF 热化学表(NIST - JANAF Thermochemical Tables)^② 中所列数据,估测该化学势差值。

采用这种方法所得到的 Ag 和 Cu 相图如图 7.3 所示。该相图所涵盖的 O_2 压强范围,从相当于实验条件的超高真空直到常压条件下。该图确认了我们可能早已从有关硬币和首饰的经验上知道的一个结果,即:Ag 远比 Cu 更抗氧化。需要记住的一点是:这个相图仅仅描述了每种材料的平衡态;对于达到平衡态的速率,它无法给出任何信息。在室温和常压下,Cu 转变为 Cu_2O 很慢,这是因为在 O_2 气氧化 Cu 的过程中存在较大的能垒。

图 7.3 中的结果并不严格精确,这一点应该并不令人奇怪。当 $p = 1$ 大气压时,对于 Ag 与 Ag_2O 具有相同稳定性的温度,在实验上测得的是约 460K;而我们的计算预测该温度是 655 K。出现偏差的原因是由于我们在热力学计算中采用了两个近似,以及 DFT 方法本身的不精确度。在我们的计算中,忽略了振动对固体自由能的贡献;正如上面所简述的那样,可以把这些影响包含进来,所付出的代价就是运行相对更多的计算。此外,更加重要的是,由于 DFT 本质上的不精确度,由 DFT 计算所得到的总能量不可避免地包含了系统误差。

① 对于 O_2 ,很重要的一点是在计算中允许自旋极化,因为 O_2 的基态是顺磁的。

② 在这些表中,理想气体的数据是根据隔离态分子的统计力学描述直接计算得到的。根据表格中的物理量,该化学势差值的定义是 $\tilde{\mu}_{O_2}(T, p^\circ) = [H^\circ(T) - H^\circ(T_r)] - TS(T) - [H^\circ(0) - H^\circ(T_r)]$,该表达式的推导思路是:理想气体分子的化学势等于理想气体的自由能,即 $G = H - TS$ 。

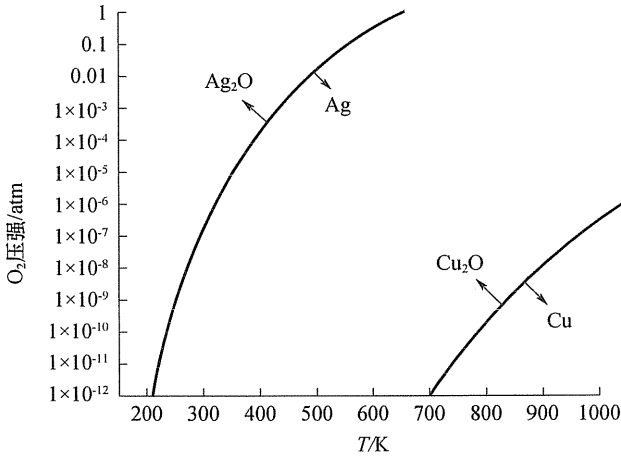


图 7.3 DFT 方法计算的体相 Cu 和 Ag 的氧化相图

7.1.1 含有无序化的例子——位形熵

在上述内容中所讨论的每种材料都是有序晶体,即:所研究的原子集合都是以一种方式排列的,从而形成晶体结构。这对于所有材料而言,显然并非都是如此。假设我们想要在上述可能的材料类别中,增加一种具有少数 O 空位的 Cu_2O 。具体说来,可以想象成一种含有 $2n$ 个 Cu 原子,并去掉 k 个氧原子的 Cu_2O 样品。则在该样品中存在有

$$N_{\text{config}} = \frac{n!}{(n-k)! k!} \quad (7.7)$$

种不同的 O 空位排列方式。对应于这些不同原子构型的熵,必定遵循熵的定义,即

$$S_{\text{config}} = k_B \ln N_{\text{config}} \quad (7.8)$$

这个物理量就称为位形熵 (Configurational Entropy)。根据 Stirling 近似,该熵值可以写为

$$S_{\text{config}} = -k_B n [(1-c) \ln(1-c) + c \ln c] \quad (7.9)$$

式中: $c = k/n$ 是 O 空位的浓度。

当我们推导上述巨势表达式[式(7.1)~式(7.5)]时,假设熵对每

种材料的贡献都是近似的,并对每个体系的自由能都进行了简化。如果我们比较有序材料(如不含空位的 Cu_2O)和无序材料(如含有空位的 Cu_2O),这个假设很显然就是不合理的,因为两种材料的位形熵存在较大差别。为了在分析相稳定性时把这一事实考虑进来,式(7.5)应该被替换为

$$\Omega_i(T, \mu_0) = E_i - TS_{\text{config},i} - \mu_0 N_{0,i} - \Omega_S^M \quad (7.10)$$

对于所有具有相同 Cu 原子数目的材料,式(7.10)的最后一项都是常数(如果对熵的贡献采用近似处理,则对于每一种材料而言,由于原子振动所引致的熵对能量的贡献都是近似的)。

在上文中,当我们比较 Cu 与 Cu_2O 时,使用巨势来确定二者中哪个材料更加稳定。现在我们可以着手处理一个稍有不同的问题,即:在某个具体压强和温度下,当 Cu_2O 与气态 O_2 处于平衡时,该材料中所应该存在的 O 空位浓度是多少? 为了回答这个问题,我们运行了两个 DFT 计算。首先,我们使用含有 $2n$ 个 Cu 原子和 n 个 O 原子的超晶胞,计算了无空位 Cu_2O 的能量 E ,并将其表示为 $E = nE_{\text{crystal}}$,其中: E_{crystal} 是符合化学计量比的材料中平均每个分子式的能量。其次:从该超晶胞中去掉一个 O 原子并重新计算其能量。如果在这些计算中所使用的超晶胞足够大的话,就可以把空位看作是与其周期性影像相互隔离开的,则第二个计算中的总能就可以写作是 $E = (n-1)E_{\text{crystal}} + E_{\text{vac}}$,其中: E_{vac} 是在含有空位时一个分子式的能量,而这个空位是由符合化学配比的材料包围着的。我们可以把这个表达式看作是 E_{vac} 的定义式,但是我们需要记得:只有当超晶胞足够大时,该定义才是准确的。在计算含有一个空位的超晶胞时,应当保持该超晶胞的晶格常数与化学计量比材料的数值相等,因为我们所感兴趣的问题是:极低浓度的空位在该原子位置上所导致的局部扭曲畸变,而并非改变了该材料的整体晶格结构。

一旦我们以上述方法确定了 E_{crystal} 和 E_{vac} ,就可以对具有 N 个分子式和 m 个空位的材料,写出其能量为

$$\frac{E(N, m)}{N} \approx (1 - c)E_{\text{crystal}} + cE_{\text{vac}}$$

式中： $c = m/N$ 。如果在式(7.10)的巨势表达式中使用这个能量，就得到了含有空位浓度的巨势定义。像以前一样，我们想要找到能够最小化巨势的材料，只要找到满足 $\partial\Omega/\partial c = 0$ 的 c 值，也就实现了这个目的。当 $c \ll 1$ 时，就能引致一个简单结果，即

$$c = \exp[-\beta(E_{\text{vac}} - E_{\text{crystal}} + \mu_{\text{O}})] \quad (7.11)$$

图 7.4 显示了对于三种不同的 O_2 压强，以这种方式计算的空位浓度。因为氧化学势受压强的直接影响[参见式(7.6)]，提高氧的压强就会极大地降低空位浓度。图 7.4 中的结果证实了我们在上述计算中所做的两个近似：①在相应的实际物理条件下，空位浓度始终是非常小的，因此在 DFT 建模计算中，将空位表示为一个较大区域的无空位材料中包围着一个空位，这一假设是恰当的。②在图 7.3 中纯金属与金属氧化物同样稳定的条件下，氧化物中的空位浓度也是非常低的。这一结果意味着：在建立图 7.3 相图时，所依据的计算中忽略了空位存在，这一假设也是合理的。

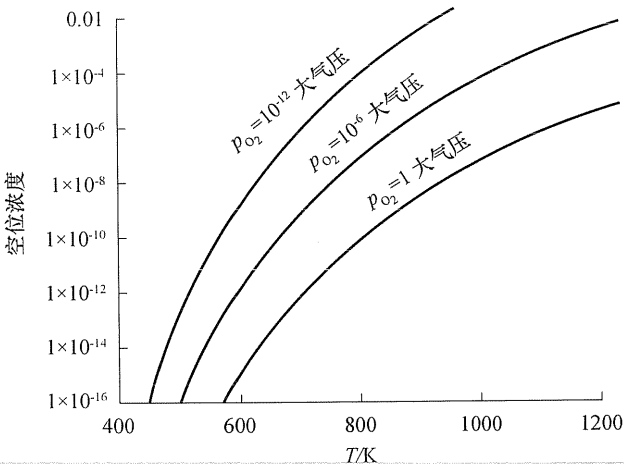


图 7.4 在三个不同氧压强下计算 Cu_2O 中的氧空位浓度。两个较低压强的虚线所表示的是：相比于图 7.3 中所确定的 $\text{Cu}, \text{Cu}_2\text{O}$ 在该温度区间处于热力学不稳定态

7.2 金属和金属氧化物表面的稳定性

在第 7.1 节中所考察的关于 Ag 氧化的相图,以一个材料为例说明了从头算热力学这一概念,而该材料所具有的特性已经能够用实验手段观测。现在我们转向另一个相关问题。对于该问题而言,使用实验手段来建立相图是很具挑战性的。这个问题就是:考察金属和金属氧化物表面的稳定性。下面先说明求解该问题的动机,众所周知,Ag 是乙烯环氧化反应的商业催化剂^①,其反应条件为:O₂压强近似于大气压,温度为 200 ~ 300℃。体相 Ag 的相图表明该催化剂在这些条件下(以及更低的 O₂压强下)的稳定相是金属态 Ag。然而,如果在超高真空条件下使用洁净的 Ag 来试图进行该反应,就会发现其表面实际呈现出催化惰性^[1]。这暗示着:在工业条件下,活化态的 Ag 催化剂表面可能并不是金属态 Ag。那么,活化态催化剂表面实质上又是什么呢?

使用从头算热力学方法,可以较容易地探索下列可能性:Ag 催化剂真实表面是否在某种程度上涉及到吸附氧,如果建立了备选表面结构的目录表,并使用 DFT 方法计算了每种结构的能量,那么就可以用第 7.1 节中有关体相相图的同样方法,预测得到一个表面相图。图 7.5 中的示例是 Li、Stampfl 和 Scheffler 采用这种方法,计算氧与 Ag(111)表面相互作用的一个相图^[1]。该相图包含了 Ag(111)上 O 原子的表面覆盖率低于一个单原子层(ML)的结果,其结构类似于第 4 章中所考察的层状重叠结构。这些计算中也考虑了几个更为复杂的结构,其中不仅有 O 原子处于表面上的,而且有位于表面原子以下的。其中的两种结构显示于图 7.6 中。最终发现,具有多层 O 原子的结构完全没有出现在最后的相图上;这种表面结构在任意温度/压强的组合条件下,都不具有优先性(该结果在性质上类似于 M₃O₂在图 7.2 中的情形)。然而,第二种结构被发现在整个相图中都具有重要作用。该结

① 仅在美国,每年就有超过 9 百万磅的乙烯氧化物是用这个工艺方法生产的。

构称为一个表面氧化物,因为其最外层呈现为氧化物形式,而其内部体相是纯金属。从图 7.5 中可见,在大多数温度下,对于这种表面氧化物结构而言,存在一个跨越了若干数量级的压强范围,且相比于体相氧化物或者洁净金属表面,这种表面氧化物结构是更加稳定的。

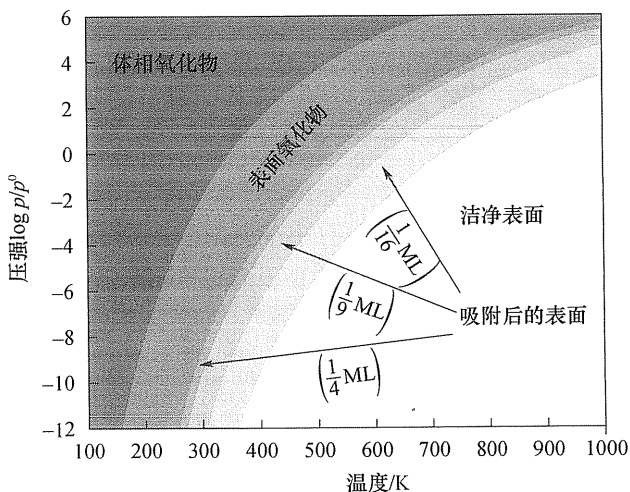


图 7.5 对于 O_2 与 $Ag(111)$ 相互作用的从头算相图。[重印许可来自于 Li W X, Stampfl C, and Scheffler M, Insights into the Function of Silver as an Oxidation Catalyst by *ab initio* Thermodynamics, Phys. Rev. B68. 2003: 165412 (Copyright 2003 by the American Physical Society)]

以上示例说明了:一旦已经确定了所有备选结构的目录列表,构建一个相图就颇为简单了。但与此同时,图 7.6 中表面氧化物结构的复杂性也可以作为一个非常好的例子,用以说明下面这一事实及其原因:形成合理的备选结构目录表常常远非那么容易。对于图 7.6 所示的这个有序表面相结构而言,依据的是 Li、Stampfl 和 Scheffler^[1] 计算时所能得到的最好实验数据。然而,更多实验和 DFT 计算已经表明:“真实的”表面氧化物结构在某种程度上与图 7.6 所示有所不同;而且也存在具有类似稳定性的其他表面氧化物相^[2]。

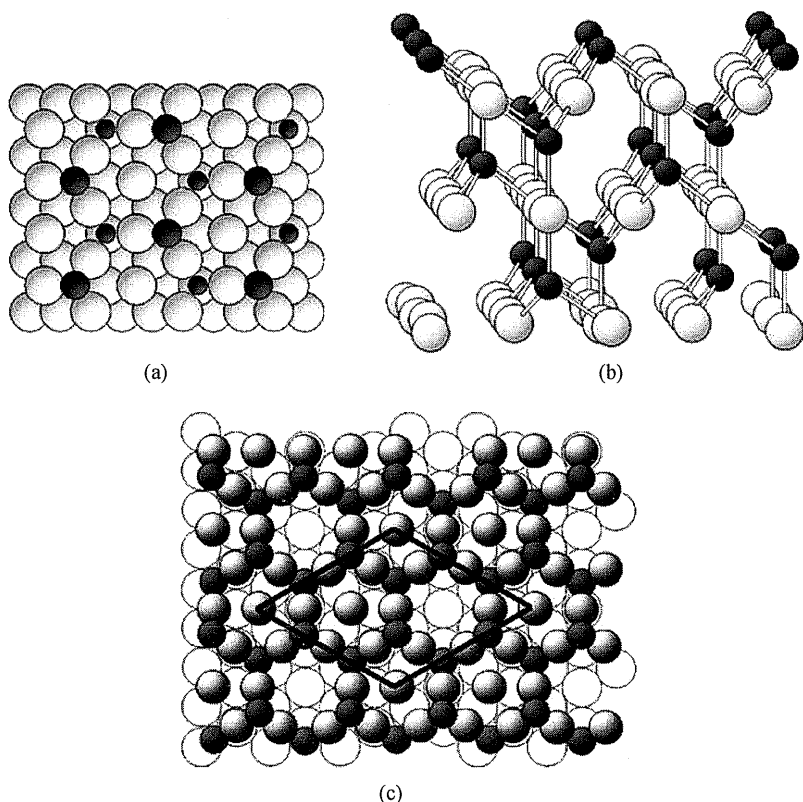


图 7.6 在建立 $O_2/Ag(111)$ 相图时, Li、Stampfl 和 Scheffler^[1]

所考察的几种结构

(a) 具有表面和次表面 O 原子(分别是较大和较小的深色小球)结构的俯视图; (b) 同一结构的侧视图; (c) (4×4) 表面氧化物的俯视图, 在 $Ag(111)$ 基底(空心圆圈)上, 具有复杂 Ag 原子(浅色小球)和 O 原子(深色小球)排布, DFT 计算预测该结构在特定温度及压强下是更为优先的(重印许可来自于图 7.5 所引证的文献)。

7.3 多元化学势和偶联化学反应

在上面的所有例子中, 所研究的相图是作为温度和单个气体物质 O_2 的函数。可以肯定的是, 许多情形下, 相图与一种以上物质的化学

势有关。这里只给出一个例子:从燃料中去除硫通常使用加氢脱硫(HDS)方法。该工艺过程中,过量的 H_2 通过硫化物催化剂与燃料发生反应,形成 H_2S 。该反应中催化剂的结构是两个量的函数:一个是氢化学势 μ_H ,另一个是硫化学势 μ_S 。尽管这超出了本书的讲述范围,但是对于确定类似于这种情形的相图,仍然可以用本章所介绍的方法作为基础,使用 DFT 方法计算所有的备选结构。本章最后的扩展阅读部分列出了这种计算的几个例子。

对于使用 DFT 计算(或者任何其他方法)确定相图,必须贯彻的最后一个忠告是:并不是所有在物理上有意义的现象都会以平衡态发生。对于在一个开放体系中发生的化学反应,如催化剂的实际应用,可以使系统处于稳定的反应速率,但并不是处于热力学平衡状态。为了对这类情形进行详细分析,必须针对控制该体系的微观过程,收集有关速率信息。扩展阅读部分给出了一个近期的例子,其中把 DFT 计算与动态蒙特卡罗计算相结合来处理该问题。

练 习

1. 在合适的温度和压强条件下,锂的氢化物(LiH)分解为固体 Li 和气态 H_2 。固体 Li 是 BCC 晶体,而 LiH 具有 NaCl 结构。对 LiH 和 Li 建立一个从头算相图,且相图是关于温度和 H_2 压强的函数。

2. 将上一题中对 LiH 的计算进行扩展,估算 LiH 中的 H 空位浓度。估算该浓度的条件是在室温下,且 $p_{H_2} = 1$ 大气压;然后,在有可能发生的超高真空 $p_{H_2} = 10^{-9}$ 大气压条件下重复该计算。

3. 推导式(7.11)。

4. 对于吸附在 $Cu(100)$ 上的 H 原子,建立一个关于温度和 H_2 压强函数的表面相图。方法是比较含有 H 的有序层状重叠结构的相对稳定性,其覆盖度分别为 1、1/2、1/4 和 1/9 倍的单原子层。此外,还可以考虑 H 原子处于最上层表面原子下方间隙位置时的情形。

5. 扩展第 7.1.1 节的概念,用以估算处于体相 Cu 八面体位置上间隙 H 原子的浓度,其中体相 Cu 的条件与上一练习相同。

参考文献

- [1] Li W - X, Stampfl C, Scheffler M. Insights into the Function of Silver as an Oxidation Catalyst by *ab initio* Thermodynamics. Phys. Rev. B68, 2003:165412.
- [2] Schnadt J, et al. Revisiting the Structure of the (4×4) Surface Oxide on Ag(111). Phys. Rev. Lett. 96, 2006:146101.

扩展阅读

作为理论基础的统计热力学,在有关分子统计力学的很多教科书中都有述及。一个知名的例子就是:McQuarrie D A. Statistical Mechanics. University Science Books, Sausalito, CA, 2000。

NIST - JANAF 热力学表(NIST - JANAF Thermodynamics Tables)可以从 National Institute of Standards 的网页上获取:www.nist.gov/srd/monogr.htm。

在第 7.1.1 节中我们对空位的讨论只关注了最简单的情形,其中空位是由去掉一个原子及其所有电子所形成的。在某些半导体材料中,带电缺陷要比未带电缺陷更加稳定。用 DFT 方法描述该情形的一个很好例子是:Van de Walle C G, Neugabauer J. Universal Alignment of Hydrogen Levels in Semiconductors, Insulators, and Solutions. Nature 423, 2003:626。

在其他例子中,也对空位进行了比我们的讨论更加复杂的全面描述,参见以下文献:

Ozolins V, Sadigh B, Asta M. Effects of Vibrational Entropy on the Al - Si Phase Diagram. J. Phys. :Condens. Matt. 17, 2005:2197.

Lu G, Kaxiras E. Hydrogen Embrittlement in Aluminum: The Crucial Role of Vacancies. Phys. Rev. Lett. 94, 2005:155501.

Forst C J, Slycke J, Van Vliet K J, et al. Point Defect Concentrations in Metastable Fe - C Alloys. Phys. Rev. Lett. 96, 2006:175501.

对于与多元化学势有关的相图例子,参见以下文献:

Cristol S, Paul J F, Payen E, et al. Theoretical Study of the MoS₂ (100) Surface: A Chemical Potential Analysis of Sulfur and Hydrogen Coverage. 2. Effect of the Total Pressure on Surface Stability, J. Phys. Chem. B 106, 2002:5659.

Stampfl C, Freeman A J. Stable and Metastable Structures of the Multi-phase Tantalum Nitride System. Phys. Rev. B71, 2005:024111.

上述文献的第二个, 其中一个物种的化学势不是相对于气态来定义的。

偶联化学反应可以影响体系的观测状态, 这一情形的详细信息可以参见下列文献:

Reuter K, Scheffler M. First - Principles Kinetic Monte Carlo Simulations for Heterogeneous Catalysis: Application to the CO Oxidation at RuO₂ (110). Phys. Rev. B73, 2006:045433.

附录 计算细节

第 7.1 节的所有计算都使用了 PW91 GGA 泛函。对于体相 Ag 和 Cu, 使用含有 4 个原子的立方超晶胞模型; 而对于体相 Ag₂O 和 Cu₂O, 则使用含有 6 个原子的立方超晶胞模型。对于每种体相材料, 使用 Monkhorst - Pack 方法对倒易空间取样, k 点网格为 $10 \times 10 \times 10$ 。对 O₂ 的计算使用了含有一个分子的 $10 \times 10 \times 10 \text{ \AA}$ 的超晶胞, k 点网格为 $3 \times 3 \times 3$ 。在 O₂ 的计算中, 使用了自旋极化, 但自旋对所考察的体相材料没有任何影响。所有计算的截断能均为 396eV。

第 7.1.1 节中的计算考察了 Cu₂O 中的空位, 使用了含有 16 个化合物分子式的立方超晶胞, k 点网格为 $5 \times 5 \times 5$ 。这些计算中的所有其他细节与第 7.1 节中对体相材料的计算相同。

第 8 章 电子结构和磁特性

在前面的章节中,我们主要讨论了与电子结构关系不大的物理特性;这之前只在一个地方显示出电子的重要性,即:只有懂得电子基态,才能准确理解材料的能量。但在实际上,可以列出很多与材料电子结构有极重要关系的物理性质。关于这些特性,其中的两个例子就是:体相材料分类为金属、半导体或绝缘体;磁特性的存在及其特征。在本章中,我们将主要考察:如何根据 DFT 计算获取这两个问题的有关信息。

8.1 电子态密度

用于描述材料电子态的一个基本物理量就是电子态密度 (Density Of States, DOS), 即

$\rho(E) dE$ = 在能级区间 $(E, E + dE)$ 上的电子态数目

回忆一下第 3 章,有关平面波 DFT 计算的基本概念,就是以函数形式 $\exp(\mathbf{ik} \cdot \mathbf{r})$ 表达电荷密度。与该式中平面波所对应的电子能量为 $E = (\hbar k)^2/2m$ 。因此,一旦完成了一个 DFT 计算,就可以通过在 k 空间中对所得到的电荷密度进行积分,进而得到其电子 DOS。下面我们给出一系列例子来说明这类计算,在阅读这些例子之前,最好简单回顾一下第 3.1 节中与 k 空间有关的基本概念,以便对这些概念有清晰明确的理解。

第一个例子就是体相 Ag 的 DOS,如图 8.1 所示。关于该数值算的细节,需要注意的几点是:

(1) 使用了较大数量的 k 点,至少要比精确计算 Ag 总能(如图 3.2)所需要的 k 点数目更多。使用更多 k 点数目来计算 DOS 是必须的,因为上面已经说过,DOS 的细节信息都是由 k 空间的积分所得到的。

(2) 在图 8.1 中对能量作图时,并没有用绝对能量,而用与 Fermi

能级 E_f 的相对值。Fermi 能级就是最高已占满电子态的能量^①。简单观察图 8.1 中就可以看到:在 Fermi 能级上 Ag 的 DOS 是非零的。对于金属最为简单的定义就是在 Fermi 能级上具有非零 DOS 的那些材料。思考该定义的一个有效方法就是考虑如下问题:当在材料上作用一个电场时,会发生什么呢? 相比于没有电场时的电子能级,电场会把这些电子加速到更高的能级上。在金属中,存在着仅仅高于 Fermi 能级的电子态,它们可以被这些加速后的电子所占据,而其所造成的结果就是:金属材料可以很容易地导通电流。图 8.1 中的结果符合我们对 Ag 的认知——它是金属。

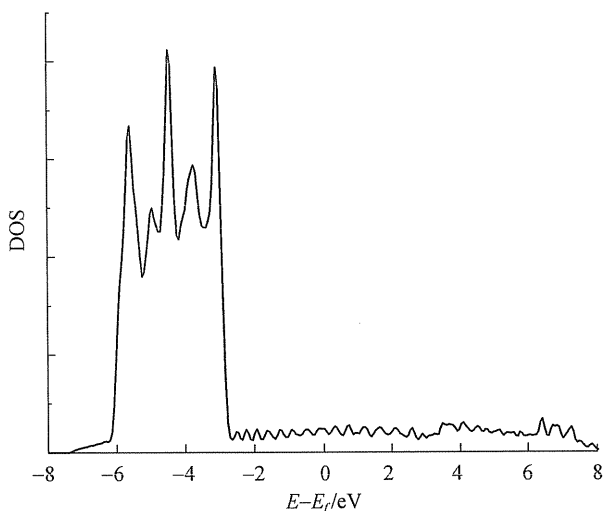


图 8.1 体相 Ag 的电子 DOS(使用含有两个原子的超晶胞,且 k 空间的取样为 $24 \times 24 \times 24$ 个 k 点)

(3) 数值细节是非常重要的,且其起因于 Ag 是金属这一事实,以及从 k 空间积分所得的 DOS。我们已经在第 3.1.4 节中讲述了,在 k

^① 更准确地说来,是在 $T=0K$ 下的最高已占满态。在非零温度下,电子的热激发会导致高于 Fermi 能级的某些电子态布居分布。

空间中对金属进行积分运算为什么会具有一定的数值困难。在图 8.1 中的结果中,使用了 Methfessel – Paxton Smearing 方法,用以提高在 k 空间积分的数值精度。

图 8.2 显示了体相 Pt 电子 DOS 的类似计算。与 Ag 的计算一样, Pt 的 DOS 在 Fermi 能级上是非零的,意味着 Pt 也是金属。比较 Ag 和 Pt 两的 DOS,可以得到两个定性的结果:①相比于 Pt, Ag 的电子态集中于较小的能级范围内;②对于高于 Fermi 能级的未占满态而言, Ag 的密度相比于 Pt 较低。这两个结果都可以大致对应于一个物理实际:Pt 通常比 Ag 具有更大的化学活泼性。

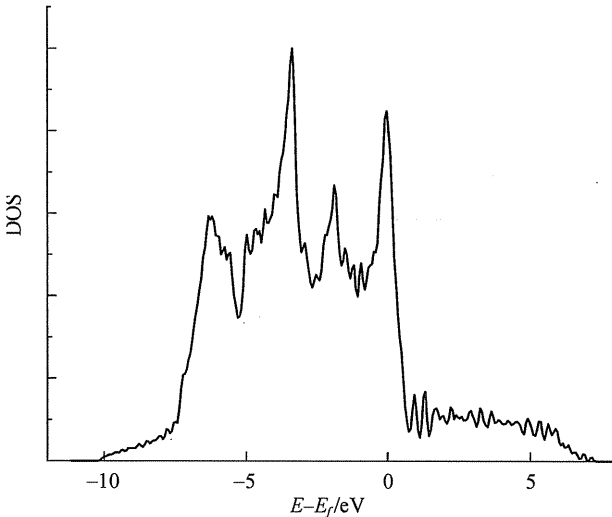


图 8.2 类似于图 8.1,但计算对象是 Pt

我们的下一个例子是完全不同的另一种材料:硅。图 8.3 显示了体相 Si 的 DOS,该 DFT 计算使用含有两个原子的体相 Si 超晶胞,采用 $12 \times 12 \times 12$ 个 k 点。如图 8.3 所示,其 DOS 可以分为两个区域,即价带和导带。价带是所有已占满电子态的集合,而所有导带上的电子态都是未占满的(在 $T = 0$ K 下)。在价带和导带之间的分隔区域内,则根本没有任何电子态;这就是带隙(或能隙, Band Gap)。根据上述有

关金属导电性的定性描述,考察一下具有带隙的材料,显然可以得到:对这些材料外加电场后,不会像在金属上那样容易地产生电子导电性。具有带隙的材料可以分为两类:如果带隙“很小”,就是半导体;如果带隙“很宽”,就是绝缘体。这两种材料之间的区别具有某种程度上的随意性,但是,当带隙大于约 3eV 时,通常都将其认为是宽带隙材料。

需要指出的是:虽然图 8.3 中的 DFT 计算正确预测了 Si 存在着带隙,但 DFT 方法并没有精确地预测出该带隙的宽度。上述计算预测其带隙为 0.67eV ,但实验上所观测的 Si 带隙为 1.1eV ^①。

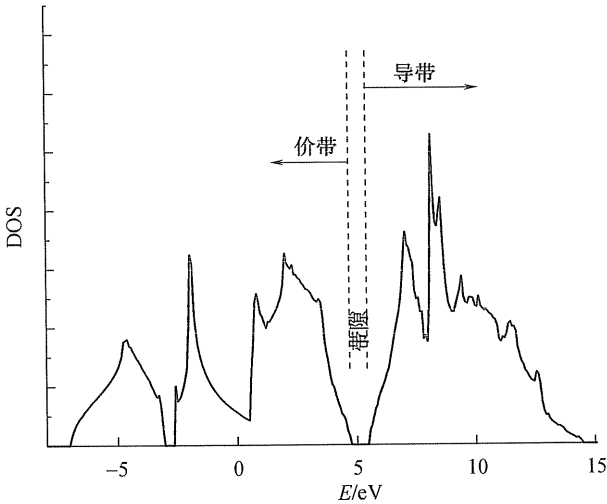


图 8.3 计算得到体相 Si 的电子 DOS

上文中提到,必须采用特殊方法才能在金属材料的 k 空间中进行精确积分。需要着重指出的是:对金属材料所使用的 Smearing 方法不能用于半导体和绝缘材料,这一点很重要。在图 8.3 的 Si 计算中,使用了 Blöchl 四面体方法(参见第 3.1.4 节)在 k 空间中插值,而并没有使用 Smearing 方法。

^① DFT 方法对带隙值的低估是一个系统性的难题。大量研究工作试图对其进行修正,从而提高带隙预测的精度(参见扩展阅读)。

由图 8.3 所得到的最后一个结果是: DOS 曲线中有很多地方的斜率是不连续变化的。这就是所谓的 van Hove 奇异点。

半导体中带隙的特点常常决定了这些材料在实际应用中的适用性。这里举一个例子: Si 是非常重要的太阳能电池材料。之所以在这一领域使用 Si 材料,其基本原理是光子可以把 Si 的电子从价带激发到导带中。一方面,在价带中所形成的未占满电子态称为空穴,因此该过程产生了一个电子-空穴对。如果电子和空穴能从物理上分隔开来,那么它们就能形成净电流。另一方面,如果电子和空穴在二者分隔开之前就重新又结合起来,就不会形成电流。Si 太阳能电池中所存在的金属杂质能够增加重新结合率。该效应如图 8.4 所示比较了体相 Si 的 DOS,以及在较大 Si 超晶胞中含有单个 Au 原子杂质的 DOS。在后一个超晶胞中,一个 Si 原子被一个 Au 原子所替换;将超晶胞中所有原子位置进行弛豫,直至体系能量达到最低。该超晶胞的晶格常数固定于体相 Si 的数值,这是因为真实材料中所可能存在的金属杂质密度极低,不会显著改变 Si 的晶格常数。

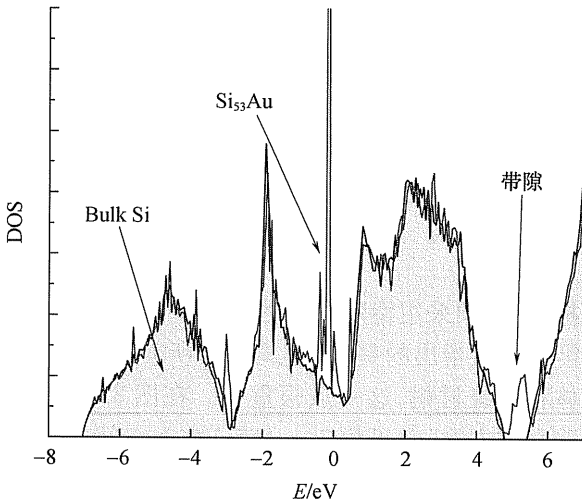


图 8.4 体相 Si(黑色线)和体相 Si_{53}Au (灰色线)的电子 DOS(两个 DFT 计算所用超晶胞均含有 54 个原子,且 k 空间的取样为 $12 \times 12 \times 12$ 个 k 点)

图 8.4 的最主要特点就是 Au 掺杂后的 Si 在纯 Si 的带隙区域中也存在有电子态。这些新的电子态意味着：相比于纯 Si，在掺杂后的材料中电子 - 空穴对可以更快地重新结合。也就是说，在晶态 Si 中所存在的 Au 杂质，会对其作为太阳能电池材料的性能，产生严重的负面影响。并不是只有 Au 的掺杂才会产生这一问题，已知有许多种类的金属掺杂都会导致类似问题。有关这方面的信息，可以参见本章末扩展阅读中所给出的相关文献。

下面给出一个宽带隙材料的例子，图 8.5 显示了体相石英 (SiO_2) 计算得到的 DOS。如图 8.5 所示，该计算得到的带隙为 5.9eV。前文中已经说过 DFT 方法通常低估带隙值，因此石英肯定是一种绝缘体。与 Si 不同，石英的价带包含有几个彼此分开的能级区域，而它们之间都存在不同的带隙。对于这些能带，我们将在 8.2 节讨论局域态密度时再来进行说明。

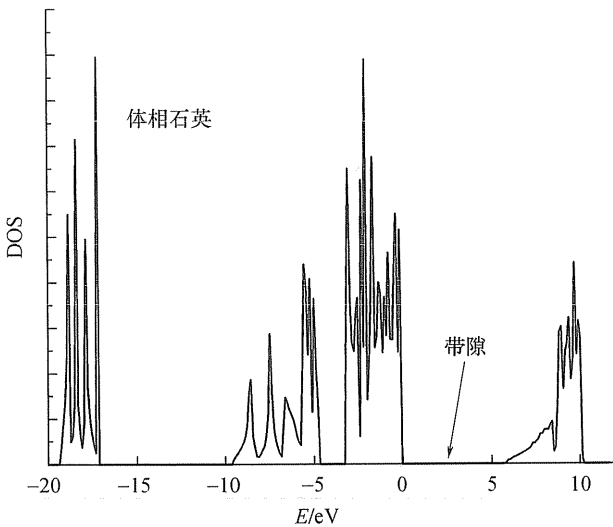


图 8.5 采用 DFT 方法计算得到的体相石英电子 DOS (所用超晶胞含有 9 个原子, k 空间的取样为 $9 \times 9 \times 9$ 个 k 点。定义能级时,使价带顶对应于 0eV)

作为体相材料 DOS 的最后一个例子,我们来考察一下第 7 章中所讨论过的一种氧化物,即 Ag_2O 。使用 GGA 泛函所计算得到的 DOS 如图 8.6 所示,在 Fermi 能级以上存在一个很小的、非零的电子态分布,这也就预测了 Ag_2O 是金属性的。该预测结果与计算所用的数值方法细节无关;无论在 k 空间中使用或不使用 Smearing,以及使用 LDA 或者 GGA 泛函,该预测结果都是一样的。但是,这个预测结果并不符合实验观测,实验结果表明 Ag_2O 是带隙为约 1.3eV 的半导体材料。在这个例子中,标准的 DFT 计算无法正确地描述 Fermi 能级附近的电子态。类似问题也会发生在许多其他的窄带隙半导体上,如 Ge、InN 和 ScN。需要指出的是:虽然 DFT 方法无法描述 Ag_2O 的半导体特性,但这并不意味着 DFT 方法不能描述该材料的其他所有物理特性。图 8.6 中所涉及的 DFT 计算预测 O - Ag 和 Ag - Ag 距离分别为 2.10Å 和 3.43Å;这两个数值分别与实验值 2.04 ~ 2.04Å 和 3.34 ~ 3.35Å 具有良好的 consistency。

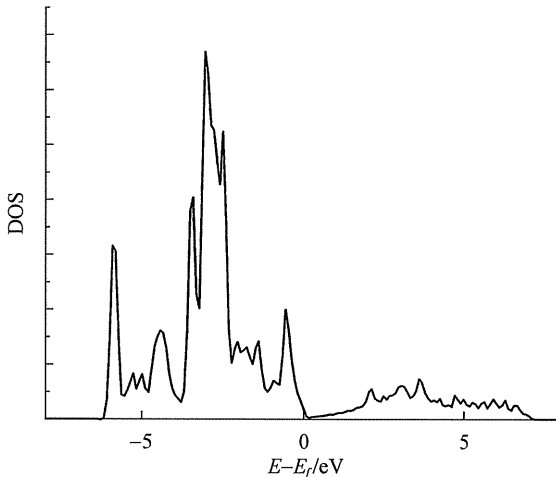


图 8.6 采用 DFT 方法计算 Ag_2O 的 DOS(所用超晶胞含有 6 个原子, k 点为 $20 \times 20 \times 20$)

电子 DOS 以一种简单形式凝练了倒易空间所有可能位置上的电子态特点。此外,也可以对材料能带结构进行考察,而且常常能从更加细微的观察角度观察材料的电子结构。能带结构所表示的是:沿着倒易空间中的一系列线(这些线通常开始并结束于 Γ 点,从而形成一个闭环),表示出所有电子态可能具有的能级值。能带结构图通常可以经过 DFT 计算得到。在这类计算中,需要特别注意 k 点的放置方法,因为计算电子态必须在倒易空间中对应于能带结构图的特殊方向上,使用一系列紧密间隔的 k 点。

8.2 局域态密度和原子电荷

为了解释材料的电子结构,常常需要明确:在特定原子周围,什么电子态是更重要的。能够实现此目的的标准方法就是局域态密度(Local Density Of States, LDOS),可以将其定义为:对于出现在原子核周围特定体积内的那些电子态,按其在全电荷密度中所占比例确定比重系数,并据此系数进行加权,从而得到其在指定能级上加权后的电子态数量。通常把原子核周围的体积取作球形,为了便于计算 LDOS,我们必须确定每一个原子的有效半径。对于这个定义,无法给出完全确切的表述。如果所使用的半径太小,就会丢失那些真正与原子核有关的电子态信息;反之,如果半径太大,LDOS 就会包含来自于其他原子的贡献。

图 8.7 和图 8.8 显示了体相石英中 O 和 Si 原子的局域 DOS。这些计算中,O 和 Si 的半径分别设为 1.09\AA 和 1.0\AA 。图 8.5 所示为该材料的总 DOS。每一个 LDOS 都分解为电子结构中的 s 和 p 轨道的两部分贡献。从图 8.7 和图 8.8 可以看到,在价带顶端区域,主要来自于 O 原子的 p 电子态,少部分来自于 Si 的 p 电子态。在低于价带顶端 $5 \sim 10\text{eV}$ 的能带范围内,主要是 O 原子 p 电子态和 Si 原子 s, p 电子态的混合。对于距离价带顶最远处的最低能带,是 Si 和 O 原子 s 电子态的混合。

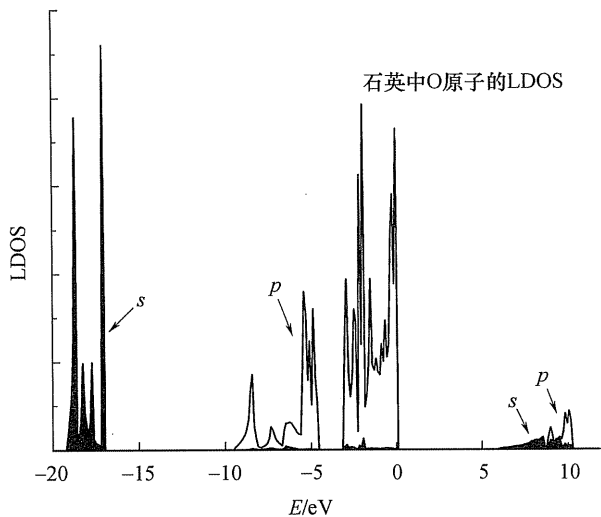


图 8.7 在体相石英中 O 原子的局域 DOS, 该计算与图 8.5 中所描述的 DFT 计算相同, 所用 O 原子半径为 1.09 Å

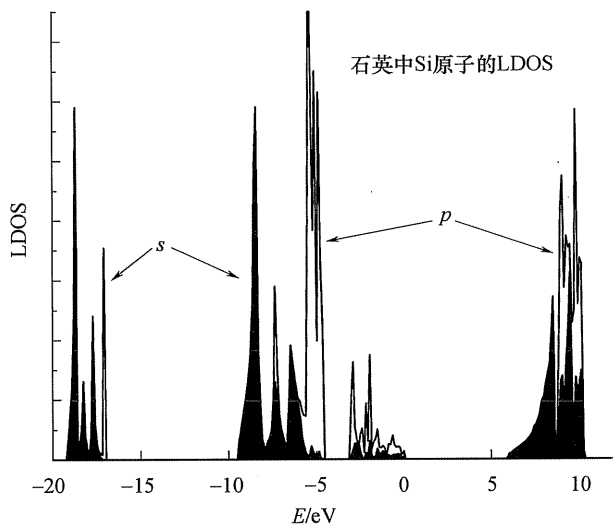


图 8.8 在体相石英中 Si 原子的局域 DOS, 该计算与图 8.5 中所描述的 DFT 计算相同, 所用 Si 原子半径为 1.0 Å

将体相材料或分子中的原子看作是带有一定的静电荷,这通常可以带来很多方便。以离子材料 NaCl 为例,通常把 +1 和 -1 电荷(单位为电子伏特)分别与 Na 和 Cl 原子相联系起来。定义 LDOS 时,关于所使用体积的描述是含糊不清的,这也就说明了:为什么根据所计算的电荷密度来完成该任务其实并不简单。对于计算化学而言,这是一个早就已经普遍存在的问题;并非仅仅只有平面波 DFT 计算是这样,对于所有已被尝试过的其他可能方法也是如此。许多最为普遍使用的方法,如 Mulliken 分析和 ChelpG 方法,只能用于完善定义的局域函数基底方法中,因此它们通常不能用于平面波 DFT 计算。能够在平面波计算中应用的一个精致方法就是 Bader 分解(Bader Decomposition),该方法使用三维电荷密度中的驻点,在不同原子之间分隔电子。需要指出的是:对于以上所提到的各种不同方法,如果它们都能用在同一种量子化学方法中,那这些方法也未必能得到完全相同的结果。因此,在计算中,只要是对原子电荷进行了分配,就应该说明分配电荷所使用的方法;当对该方法的结果进行解释时,也需要始终注意这些电荷分配方法的含糊性。

8.3 磁性

磁性是电子非零自旋所造成的直接结果。在反磁性材料中,每个电子态独由两个电子所占据,一个向上自旋,另一个向下自旋,该情形如图 8.9(a)中对周期性材料的示意性描述。然而,在磁性材料中,存在着只含有一个电子的电子态。当存在未配对电子时,它们可以形成多种有序化方式,而每一种有序化方式都定义了一个不同的磁性状态。两个最常见的磁性状态如图 8.9(b)和(c)所示,分别是铁磁态(其中所有的电子自旋都指向同一个方向)以及反铁磁态(其中相邻原子的电子自旋方向交替变化)。也有可能出现更加难以理解的自旋有序化,其中一个例子如图 8.9(d)所示。每个原子上的平均电子自旋就是所谓的磁矩(Magnetic Moment)。图 8.9(b)所示铁磁态的磁矩为 1,而图 8.9 中其他所有情形的磁矩均为零。

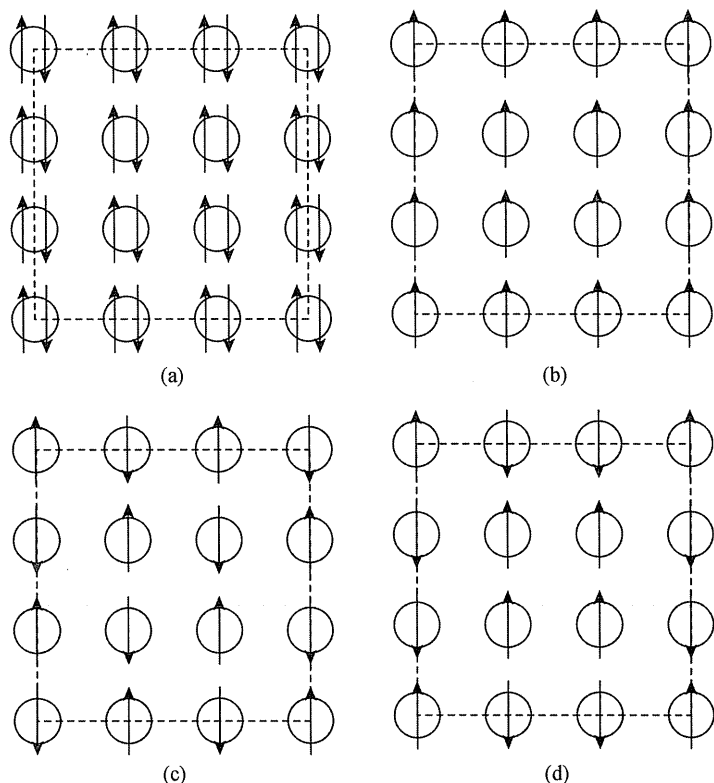


图 8.9 在二维周期性材料中自旋态的示意图

(a) 每个原子的所有电子都是成对的; (b) 铁磁态电子磁态示意图;

(c) 反铁磁态电子磁态示意图; (d) 更复杂的电子磁态示意图。

在 DFT 计算中通常没有明确地考虑电子自旋, 其中一个简单原因就是: 这一近似(不考虑自旋)能使计算速度更快。然而, 当材料中的自旋效应非常重要时, 就必须把自旋考虑进来。例如: Fe 是人所共知的具有磁特性的一种金属^①。图 8.10 显示了 BCC 晶体结构下体相 Fe 的能量, 计算方法为两种: 一种没有考虑自旋极化; 另一种考虑了铁磁

① 根据铁的磁特性制成指南针, 已经使用了 1000 年左右了。

性自旋有序化。二者的区别是非常明显的；电子自旋显著地降低了能量，并且预测得到的平衡晶格常数也增大了约 0.1 \AA 。

对于考虑了自旋的 DFT 计算,图 8.9 显示了一个非常重要的微妙之处,即:电子自旋可能会形成许多不同的磁性有序化。在单个 DFT 计算中的电子自旋优化,类似于在特定晶体结构下的原子位置优化;在计算开始时,就确定了自旋的初始化近似,随后的计算中只是找到该初始自旋近似所对应的极小值。常常把铁磁态用作初始近似,除非另有说明。此时,一个关键问题就是:该方法只能给出一个能量表面上的局部最小值,而该能量表面是由所有可能的自旋有序化来定义的。确定该能量表面上与全局最小值所对应的自旋有序化是更加困难的;这一问题可以完全类比于:在所有可能的晶体结构中,预测某个材料的最优晶体结构。如同使用 DFT 方法探讨不同晶体结构那样,当探讨磁性有序化时,考察多于一种的自旋有序化初始近似,是非常明智的。例如,图 8.10 显示了根据一系列计算得到反铁磁有序化 FCC Fe 的能量,这些计算中,使用反铁磁有序化作为自旋态的初始化近似。反铁磁态的

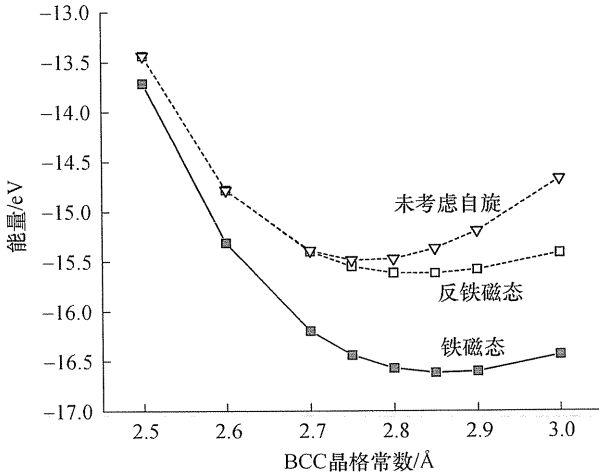


图 8.10 根据 DFT 计算使用不同自旋态得到的 BCC Fe 能量与 Fe 晶格常数间函数关系。标注了“*No spin*”的结果在计算时没有使用自旋极化

能量显著高于铁磁态的能量,这与实验事实是一致的:Fe 是铁磁性的。需要注意的是,对于是否存在比铁磁性或反铁磁性有序化更加复杂的有序化自旋态,上述计算并没有提供有关任何信息。

我们必须强调的是,在这个关于磁性极其简略的介绍中,我们仅仅对这个深厚主题进行了非常肤浅地讨论。因为磁性是大多数数字数据存储技术的核心问题,与磁性有关的很多现象早已成为世界范围近几十年的重点研究主题。扩展阅读部分列出了一些文献,可以为这一主题提供一个切入点。

练 习

1. 对体相 Ag 计算电子 DOS 和总能,计算中使用不同数量的 k 点在 k 空间中取样。本章中与图 8.1 有关的一条结论是:相比于总能计算,在 DOS 计算中需要更多 k 点才能得到收敛良好的结果。你的计算结果是否符合该结论? 定性地说明其发生原因。

2. 在结构优化后,计算体相 Ag_2O 的电子 DOS。根据计算结果,检验“DFT 计算预测该材料是金属性的”这一论断是否正确。

3. 画出类似于图 8.9(d) 的三种自旋有序化,但三者均具有非零总磁矩。

4. Co、MnBi 和 MnF_2 全都是磁性材料。明确它们为何是磁性(或反铁磁性)的。每种材料的磁性状态对其晶格常数具有多大影响? MnBi 是 NiAs 结构(空间群 $R\bar{3}m$),而 MnF_2 是金红石(TiO_2)结构。

5. 对于双原子分子 N_2 、 O_2 和 CO,通过计算确定自旋是否会影响其结构及能量。

6. 图 8.10 所示 DFT 计算对于 $T = 0$ K 下的体相 Fe 是有效的。当温度高于所谓的 Curie 温度(T_c)时,Fe 就不再是铁磁性的,这是由于发生了自旋有序化的热效应。检索找到能够说明下述内容的文献:周期表中哪些金属是铁磁性的,而它们的 Curie 温度又是多少?

扩展阅读

关于硅太阳能电池中金属杂质作用的更多详细信息,以及材料工程师对降低这些杂质影响的相关工作,可以参见 Dubois S, Palais O, Pasquinelli M, et al. *J. Appl. Phys.* 100, 2006:123502, 以及 T Buonassisi, Istratov A A, Marcus M A, et al. *Nature Materials* 4, 2005:676。

有关于为什么 DFT 方法不能精确预测带隙的解释,请参见: Perdew J P, Levy M. *Phys. Rev. Lett.* 51, 1983:1884 和 Sham L J, Schluter M. *Phys. Rev. Lett.* 51, 1983:1888。对于一些校正方法的描述,请参见: Lany S, Zunger A. *Phys. Rev. B* 78, 2008:235104。

对于 DFT 计算在预测 Ag_2O 半导体特征时的缺陷,以及应用“post - GGA”方法解决该难题的详细信息,可以参见: Li W - X, Stampfl C, Scheffler M. *Phys. Rev. B* 68, 2003:165412。

对于使用电荷分解法分配原子电荷,如果想要了解更多,可以参见: Henkelman G, Arnaldsson A, Jonsson H. *Comp. Mater. Sci.* 36, 2006:354。

磁性的基本原理几乎在每个固体物理教材中都有所描述。其中的一个经典参考书是: Kittel C. *Introduction to Solid State Physics*. Wiley, New York, 1976。对磁性现象进行更深入讨论的几本著作是: Getzlaff M. *Fundamentals of Magnetism*. Springer, Berlin, 2007, Mohn P. *Magnetism in the Solid State: An Introduction*. Springer, Berlin, 2005, 以及 Morrish A H. *The Physical Principles of Magnetism*. Wiley IEEE Press, New York, 2001。

附录 计算细节

本章所有结果都来自于使用 PW91 GGA 泛函的计算,且所有计算都使用了 Monkhorst - Pack 方法对 k 空间进行取样。

8.1 节 对于体相 Ag、Pt、Si、石英和 Ag_2O 的计算,所用超晶胞具

有 DFT 优化后的晶格参数,且截断能分别为 396eV、239eV、151eV、396eV 和 396eV。对于 Ag、Pt、Si₃₃Au、石英和 Ag₂O,超晶胞尺寸和 *k* 点数量均如上所述。在图 8.3 体相 Si 的结果中,使用的是含有 2 个原子的超晶胞,12 × 12 × 12 个 *k* 点。在图 8.4 的结果中,计算使用的是含有 54 个原子的超晶胞,截断能为 181eV。

8.3 节 对于 BCC Fe 的所有计算,均使用含有两个原子的立方超晶胞,5 × 5 × 5 个 *k* 点,以及 270eV 的截断能。

第 9 章 从头算分子动力学

我们周围材料所必定具有的一个特点就是：它们的原子始终在运动。已经有大量文献讲述如何把材料中的原子运动描述为时间的函数，而这也正是描述某些实际体系性质时的一个前提条件。在本章中，我们讲述分子动力学方法 (Molecular Dynamics, MD)，这类计算工具能够实现跟踪运动原子的轨迹。在第 9.1 节这一开始部分，我们暂时从量子力学后退出来，描述一下如何用经典力学进行 MD 模拟计算。在这个“迂回”中介绍几个概念，而这些概念对于在 DFT 计算中使用 MD 方法，是非常重要的。第 9.2 节介绍了可以如何使用 DFT 计算所得信息进行 MD 模拟的方法。第 9.3 节给出了使用从头算 MD 方法的两个例子，一个例子与生成无序体相材料有关，而另一个例子是用 MD 方法探索复杂的能量表面。

9.1 经典分子动力学

9.1.1 恒定能量的分子动力学

为了能够理解从头算分子动力学 (*ab initio* MD) 的主要概念，首先需要回顾一些经典力学中的概念。经典 MD 是一个已经发展完善的方法，该方法广泛用于许多不同种类的计算化学和材料模拟中。本章末的扩展阅读部分所包含的文献，可以作为进入该领域的详细切入点。

我们所考虑的情形是：在体积 V 中的 N 个原子，而且我们所感兴趣的是理解这些原子的动力学。为了确定原子集合在任意时刻的构型，我们需要确定 $3N$ 个位置 $\{r_1, \dots, r_{3N}\}$ ，以及 $3N$ 个速度 $\{v_1, \dots,$

v_{3N} 。对于描述该体系的总体状态,有两个非常有用的物理量。一个是总动能,即

$$K = \frac{1}{2} \sum_{i=1}^{3N} m_i v_i^2 \quad (9.1)$$

式中: m_i 是对应于第 i 个坐标的原子质量。而另一个物理量就是总势能,即

$$U = U(r_1, \dots, r_{3N}) \quad (9.2)$$

因为我们在经典力学的框架内来处理这些原子运动,所以可以把牛顿运动定律应用于这些原子上,即

$$F_i = m_i a_i = m_i \frac{dv_i}{dt} \quad (9.3)$$

式中: F_i 和 a_i 分别是作用在第 i 个坐标上的力和加速度; t 是时间。其中,力与总势能的导数存在关系,即

$$F_i = - \frac{\partial U}{\partial r_i} \quad (9.4)$$

这些关系式定义了原子的运动方程,该方程可以写作是 $6N$ 个一阶常微分方程,即

$$\begin{cases} \frac{dr_i}{dt} = v_i \\ \frac{dv_i}{dt} = - \frac{1}{m_i} \frac{\partial U(r_1, \dots, r_{3N})}{\partial r_i} \end{cases} \quad (9.5)$$

这些方程的一个重要特征就是它们都符合能量守恒,即 $E = K + U$ 不随着时间变化而改变。如果用统计力学语言来描述,就是这些原子是在微正则系综(Microcanonical Ensemble)中运动。所谓微正则系综就是具有固定 N 、 V 、 E 的一组可能状态。在这个系综中,能量并不是唯一守恒的量,粒子的总动量也是守恒的。对于一个隔离的粒子集合,总的角动量也是守恒的,但该物理量在具有周期性边界条件的体系中则不是守恒的。

如果想要把这些体系性质对应联系到某个物理情形上,就需要我们能够表征该体系的温度 T 。对于在温度 T 下处于平衡的宏观原子集合,原子的速度分布符合 Maxwell - Boltzmann 分布关系。该分布的一

个主要特点就是每个自由度的平均动能为

$$\frac{1}{2} m \overline{(v^2)} = \frac{k_B T}{2} \quad (9.6)$$

在分子动力学学中,可以把该关系反过来用于确定温度,即

$$\frac{k_B T_{\text{MD}}}{2} \equiv \frac{1}{6N} \sum_{i=1}^{3N} m_i v_i^2 \quad (9.7)$$

需要说明的是,由于在上述运动方程中动能 K 并不守恒,因此在微正则系综 MD 模拟中,所观测到的温度一定是随时间波动的。

基本上说来,对于具有物理意义的体系,上述运动方程都是非常复杂的,因而无法得到其解析解。所以,能够对这些方程进行数值积分,以便追踪原子的动力学过程,就显得非常重要。一个简单的方法就是使用泰勒展开式,即

$$\begin{aligned} r_i(t + \Delta t) = r_i(t) + \frac{dr_i(t)}{dt} \Delta t + \frac{1}{2} \frac{d^2 r_i(t)}{dt^2} \Delta t^2 + \\ \frac{1}{6} \frac{d^3 r_i(t)}{dt^3} \Delta t^3 + O(\Delta t^4) \end{aligned} \quad (9.8)$$

对上式右侧的导数用相应的名称符号来表示^①,则上式可以改写为

$$\begin{aligned} r_i(t + \Delta t) = r_i(t) + v_i(t) \Delta t + \frac{1}{2} a_i(t) \Delta t^2 + \\ \frac{1}{6} \frac{d^3 r_i(t)}{dt^3} \Delta t^3 + O(\Delta t^4) \end{aligned} \quad (9.9)$$

如果分别用正的或负的时间步长来表示式(9.9),并取这两个表示形式的差,则可以看到

$$r_i(t + \Delta t) \approx 2r_i(t) - r_i(t - \Delta t) + \frac{F_i(t)}{m_i} \Delta t^2 \quad (9.10)$$

这就是所谓的 Verlet 算法。假如时间步长 Δt 足够小,该算法就能够对式(9.5)所定义的轨迹给出一个精确近似。

我们在上述运动方程的介绍中,是从牛顿经典力学的观点出发,并

① 三阶导数项(也就是加速度的一阶导数)也称为觉克(jerk)。

使用了著名的牛顿方程 $F = ma$ 。需要明确的是:对于经典动力学中运动方程的定义,这不是唯一(甚至也不是最好)的方法。另一个非常有效的方法是定义一个所谓的拉格朗日量(Lagrangian) L ,其表达式中含有动能和势能,即

$$L = K - U$$

$$= \frac{1}{2} \sum_{i=1}^{3N} m_i v_i^2 - U(r_1, \dots, r_{3N}) \quad (9.11)$$

对于每个坐标的运动方程,其拉格朗日量为

$$\frac{d}{dt} \left(\frac{\partial L}{\partial v_i} \right) = \frac{\partial L}{\partial r_i} \quad (9.12)$$

可以简单推导并验证,该方法所定义的运动方程,与上述牛顿力学方法所定义的完全相同。

9.1.2 正则系综的分子动力学

在分子动力学的许多应用中,都需要将 MD 结果和实验观测进行比较。在通常的实验条件下,所研究材料中的分子能够与其环境交换热量。在这种情形下,这些原子就处于正则系综(Canonical Ensemble)中,该系综的 N 、 V 和 T 都是常数。对于使用上述微正则 MD 方法模拟一个正则系综,有许多方法都可以实现。一个最简洁的方法是由 Nosé 最先提出的,他开始在微正则系综中使用拉格朗日量[式(9.11)],从而形成了增广拉格朗日量(Extended Lagrangian),即

$$L = \frac{1}{2} \sum_{i=1}^{3N} m_i s^2 v_i^2 - U(r_1, \dots, r_{3N}) + \frac{Q}{2} \left(\frac{ds}{dt} \right)^2 - g k_B T \ln(s) \quad (9.13)$$

请注意,如果 $s(t) \equiv 1$,则该增广拉格朗日量就正好简化为式(9.11)。我们可以使用式(9.12)来确定与增广拉格朗日量所对应的运动方程。Hoover 使用与上述增广拉格朗日量稍有不同的变量,给出了这些方程的一种方便表达形式,即

$$\begin{cases} \frac{dr_i}{dt} = v_i \\ \frac{dv_i}{dt} = -\frac{1}{m_i} \frac{\partial U(r_1, \dots, r_{3N})}{\partial r_i} - \frac{\xi}{m_i} v_i \\ \frac{d\xi}{dt} = \frac{1}{Q} \left[\sum_{i=1}^{3N} m_i v_i^2 - 3 N k_B T \right] \\ \frac{d \ln s}{dt} = \xi \end{cases} \quad (9.14)$$

这些方程并不像它们看上去那样难以理解。第一个和第二个方程与式(9.5)中所看到的是一样的,只是在第二个方程中增加了“摩擦力”项,它要么使速度增大,要么使其减小,而增大或减小取决于 ξ 的符号。第三个方程决定了 ξ 的符号和数量级。如果我们按如下方式重写式(9.7),就会更清楚这个方程的含义,即

$$\frac{d\xi}{dt} = \frac{3 N k_B}{Q} [T_{MD} - T] \quad (9.15)$$

这个方程的作用就是一个反馈控制,从而保持模拟过程中原子的瞬时温度 T_{MD} 接近于所需要达到的温度 T 。如果瞬时温度太高,则由式(9.15)平滑地调整 ξ ,从而调整所有原子的速度并降低其平均动能。参数 Q 确定了在温度差($T_{MD} - T$)和 ξ 之间的反馈变化的速度。

正如我们在式(9.8)中所做的那样,通过应用泰勒展开式可以对Verlet算法进行进一步推导,从而使这些公式能够被数值积分。这个控制温度的方法称为Nosé - Hoover恒温器。

9.1.3 分子动力学的实际应用方面

为了明确经典MD计算方法能够处理哪些问题,需要掌握经典MD计算的两方面内容。通过考察能量恒定MD计算[式(9.10)]的Verlet算法,可以对这两个概念同时进行说明。

首先,需要思考的是该算法中每步计算所涉及到的计算工作量。在每个时间步的计算中,必须确定作用在每个自由度上的力 $F_i(t)$,这个值(大致)相当于在每个时间步上计算总势能 $U = U(r_1, \dots, r_{3N})$ 。即

使这个势能是坐标的简单函数,例如仅由原子间距所决定的若干项连加和,这个计算所涉及的工作量也会随着体系尺寸增大而急剧增加。在任何 MD 模拟中,对力进行估算时所耗费的计算成本占据最主要的部分,因此必须非常慎重地选定一种效率较高的数值方法进行力的估算。

其次,用泰勒展开方法对 Verlet 算法所推导得到的表达式中,突出了这样一个事实:只有当时间步长 Δt 足够小时,该算法才是精确的。为了估算一下这些时间步长应该是多小,我们注意到:对于几乎所有的原子或分子体系,均存在频率为 10^{13} s^{-1} 数量级的特征振动。这意味着一次振动需要约 100 fs。如果想要使数值上形成的 MD 轨迹能够精确描述这类运动,就必须在这个时间周期内选用更多数量的时间步长。上述简单推理表明:MD 模拟通常所用的时间步长应该小于 10 fs。对于一组原子的动力学过程,如果给出 1ns 的跟踪轨迹,在使用 10 fs 时间步长时就需要 10^5 个 MD 步骤,这意味着需要投入相当大的计算资源。

9.2 从头算分子动力学

上述有关经典 MD 的解释,部分意义上是为了强调:如果已知了以原子坐标函数表达的原子集合势能,即 $U = U(r_1, \dots, r_{3N})$,就可以对原子集合的动力学过程进行描述。也许读者已经想到,一个很自然的 DFT 计算使用方式可能就是:先使用 DFT 方法计算 $U = U(r_1, \dots, r_{3N})$,再用以进行分子动力学模拟。换言之,所研究体系的势能可以使用量子力学“飞速地”计算,这就是从头算 MD 的基本概念。该方法的拉格朗日量可以写为

$$L = K - U = \frac{1}{2} \sum_{i=1}^{3N} m_i v_i^2 - E[\varphi(r_1, \dots, r_{3N})] \quad (9.16)$$

式中: $\varphi(r_1, \dots, r_{3N})$ 表示在体系的电子基态下一整套 Kohn - Sham 单电子方程波函数(详见第 1.4 节)。该拉格朗日量表示出应该以某一个顺序完成所有计算:首先,计算基态能量;最后进行一步 MD,原子核位置向前移动;接着计算新的基态能量;最后按此顺序从头开始。对于经

典力学所定义的,沿着轨迹向前移动原子核位置的任何方法,如果引发该位移的力是由 DFT 方法计算得到的,那我们就把这些方法都称为从头算 MD 方法。因为在 MD 中所能采用的时间步长是是有限的,所以对于在实际物理问题中运用 MD 方法的可行性而言,找到效率较高的计算方法来完成 MD 运算是极其重要的。

对于把从头算 MD 从简单有趣的想法改变为强大而实用的计算方法,一个关键性的突破是由 Car 和 Parrinello 所完成的(参见扩展阅读)。他们介绍了一个算法,通过一个增广拉格朗日量,把两个分开的任务(跟踪原子核的运动,以及在给定原子核位置下找到电子基态)联合到一起进行处理。该方法中的核心思路就是:使用分子动力学,定义能够同时跟踪原子核和电子自由度的运动方程。Car 和 Parrinello 的增广拉格朗日量建立得非常巧妙,其使用了类似于式(9.16)的原子核运动方程,并引入了电子自由度作为虚拟动态变量。该增广拉格朗日量为

$$L = \frac{1}{2} \sum_{i=1}^{3N} m_i v_i^2 - E \left[\varphi(r_1, \dots, r_{3N}) \right] + \frac{1}{2} \sum_j 2\mu \int dr \left| \dot{\psi}_j(r) \right|^2 + L_{\text{ortho}} \quad (9.17)$$

上式右侧的前面两项与式(9.16)相同,而后两项引入了虚拟自由度。具有动能形式的第三项引入了一个虚拟质量 μ ,而最后一项则保证单电子波函数相互正交。当用动力学等式(9.17)中的速度表示温度,并按比例外推到 $T \rightarrow 0$ 时,就达到最小值 E 所对应的平衡态,并且拉格朗日量描述了势能表面上一个真实的物理体系。这种方法称为 Car - Parrinello 分子动力学方法(Car - Parrinello Molecular Dynamics, CPMD)。

在 CPMD 计算中,对原子核和电子自由度同时进行了扩展,所以在每个时间步所计算的总能并不严格对应于真实 Born - Oppenheimer 势能表面上的相应原子核位置。这一概念如图 9.1 所示,性地显示了在 CPMD 模拟过程中所观测到的瞬时电子能量。还需要清楚的是:在 CPMD 计算中电子自由度的动力学变化,在物理上不能将其解释成电子的动力学过程;电子的运动方程仅仅是一种数学模型,其目的就是以一种高效率的数值方式,生成原子核的动态变化。

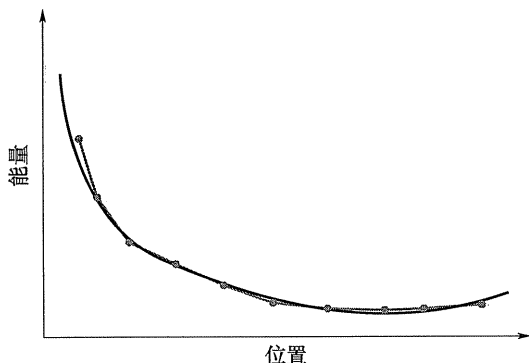


图 9.1 沿着单个空间坐标方向的 CPMD 轨迹示意图(平滑曲线表示精确的电子基态能量,而圆形符号表示在 MD 模拟的每个时间步上 CPMD 计算的电子能量)

Car - Parrinello 方法最初不仅用于从头算分子动力学,也用于确定 Kohn - Sham 基态(因为相比于同时期的其他方法,该技术能够更快地完成矩阵的对角化)。这一情形直到 20 世纪 90 年代才发生改变,许多基于迭代线性代数求解 Kohn - Sham 方程的高效数值方法被相继提出并获得广泛应用。只要恰当地使用前一个 MD 步骤所得到的电子信息,可以为原子核位置变更后的基态提供良好的初始近似,从而使得在接下来的时间步内,能量和力能够被更高效地计算。基于这种途径的 MD 方法常称作 Born - Oppenheimer 分子动力学(BOMD),因为这些方法直接在 Born - Oppenheimer 势能表面上考察了原子核的动态变化。尽管 Car 和 Parrinello 的增广拉格朗日方法仍然深具影响,但是在每个时间步的计算上,基于 Kohn - Sham 方程直接最小化的方法现在已经获得更广泛的使用。

总结以上有关从头算 MD 的简单综述,可以得出结论:式(9.16)所确定的动态过程定义了微正则系综。也就是说,由其拉格朗日量所定义的轨迹会使体系的总能守恒。与经典 MD 模拟相类似,通常更有意义的从头算 MD 是计算在某个固定温度下动态过程所对应的轨迹。而普遍采用的有效方法就是在拉格朗日量中增加一个附加项,从而能够使用第 9.1.2 节介绍的 Nosé - Hoover 恒温器在正则系综(恒定的 N 、 V 和 T) 中完成计算。

9.3 从头算分子动力学的应用

9.3.1 考察结构复杂材料:液体和非晶相

如果回头去看本书在这之前的所有章节,可以注意到,对于体相材料的计算,几乎所有讨论过的例子都是针对晶体材料的。一个实际原因就是在晶体材料中定义原子的位置更为简单直接;但是实际上肯定存在许多对物质非晶态(包括液体和无定形材料)更感兴趣的情形。使用 DFT 方法考察这些无序态是深具挑战性的,因为首先必须要找到材料中原子的组构方式,从而用 DFT 方法对这些材料特性进行某些处理。从头算 MD 提供了处理该问题的一个强有力工具。

我们采用 Lewis、De Vita 和 Car 对磷化铟(InP)的详细研究结果^[1],来说明从头算 MD 对液体和非晶相的应用。该项工作的出发点或动机是在微电子电路制造中使用无定形 InP。建立液体和无定形材料计算模型的思路,就是模拟实验上获得这些材料所采用的方法:先把一个试样加热到非常高的温度(形成液体),然后快速淬火降温(形成无定形材料)。在 Lewis 等人所做的计算中,使用 64 个原子的超晶胞(32 个 In 原子和 32 个 P 原子),并在 k 空间的 Γ 点取样;分子动力学模拟所使用的时间步长为 0.25 fs。该计算中,以平衡态晶体结构(闪锌矿结构)为起始态;该计算中温度的历史记录如图 9.2 所示。请注意:该图对应着 360000 个 MD 步骤。尽管这已经是一个非常耗时的计算,但是模拟中的淬火降温速率仍然是约 10^{13} K/s;相比于实际上处理无定形材料时的典型降温速率,这个模拟的降温速率仍然很大。在该计算中,调整超晶胞尺寸使其液体密度比晶态密度大 1.07 倍,而非晶态密度与晶态密度相等。

如图(9.3)所示,在该模拟的淬火降温部分,InP 体系的能量可以看作是温度的函数。观察能量对于温度的变化斜率,在大约 900K 时具有液态-玻璃转变特征。无定形相的最终结构是能量的极小值点,因为温度淬火降温到了 0K,但相比于晶态结构,该结构的能量大致要高出约 0.24eV。

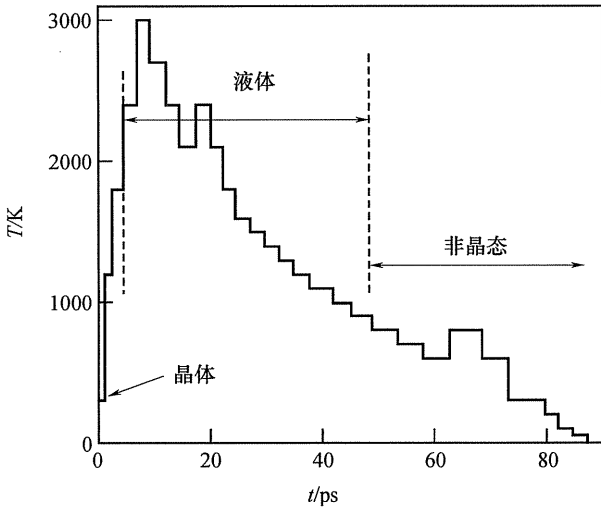


图 9.2 对 InP 进行 CPMD 模拟的温度历史记录,表示了晶体变为液体的熔化过程和液体变为非晶固体的淬火过程。[重印许可来自于 Lewis L J, Vita A De, Car R, Structure and Electronic Properties of Amorphous Indium Phosphide from First Principles, Phys. Rev. B 57, 1998:1594 (Copyright 1994 by the American Physical Society).]

一旦用 MD 模拟生成了液态和非晶 InP 的结构,就可以考察这些材料的各种不同性质。这里只给出一个例子,无序化材料的结构常用径向分布函数来表征,即

$$g(r) = \frac{\rho(r)}{4\pi r^2 \bar{\rho}} \quad (9.18)$$

式中: $\rho(r)$ 是在材料中距任一原子的半径为 r 的球形薄壳内,所发现原子的平均密度; $\bar{\rho}$ 是整个材料的平均密度。对于非常小的 r 值,则 $g(r) \rightarrow 0$,因为原子不能彼此重叠。对于较大的 r 值,则 $g(r) \rightarrow 1$,因为在无序材料中彼此分隔开的原子之间,并不会发生相互影响。对于液态和非晶态的 InP, Lewis 等人计算的分布函数如图 9.4 所示。正如我们所想象的那样,相比于液态,非晶态材料具有更多的结构。非晶态材料的一个重要特点就是 P-P 分布曲线中靠近 $r = 2.2 \text{ \AA}$ 处的峰值,它显示出非晶态材料中存在 P-P 键,而这个键在晶态材料中并不存在。

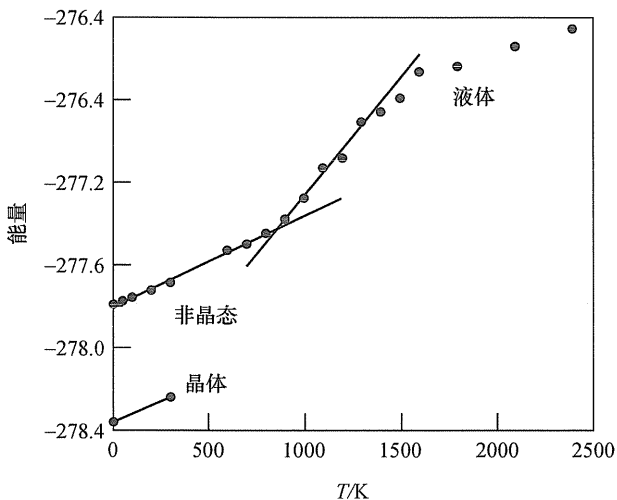


图 9.3 在从液态变化为非晶态的淬火过程中 64 个原子 InP 超晶胞的能量。图中也显示了晶态 InP 的能量。(重印许可来自于图 9.2 所引用的文献)

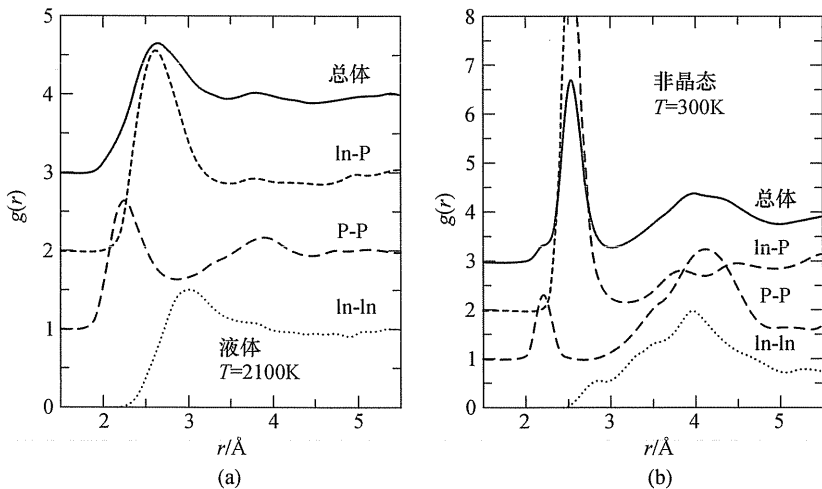


图 9.4 液态和非晶态 InP 的径向分布函数。在每个小图的竖直方向上,都分别显示了 In - In、P - P、In - P 原子对,以及全体原子的径向分布曲线 (重印许可来自于图 9.2 所引用的文献)

(a) 液态;(b)非晶态。

9.3.2 考察复杂能量表面

在第2章有关DFT计算的例子中,我们给出的第一个例子就是通过比较几种可能晶体结构的能量,推断简单金属的晶体结构。现在,我们来考虑一个与此相关的问题:预测金属纳米颗粒的结构。在多相催化中,金属纳米颗粒是极为重要的,而且近年来金属纳米颗粒也广泛用于其他科技用途。因此,这就存在着需要对其特性进行深入理解的强大动力。对于用DFT方法(或任何其他理论方法)研究纳米颗粒,一个合适的起始点就是确定纳米颗粒的原子结构;而这一问题的主要特点就是:可能存在较大数量的不同极值点,且这些极值点具有完全不同的结构。如果知道了这些纳米颗粒的初始几何构型,DFT计算方法非常适合于优化这些结构,但是只有在能够生成一系列初始结构的条件下,这样的计算才是有意义的。

对于如何生成纳米颗粒的一系列可能结构,其中一个方法是根据物理直觉或寻遍相关文献,提出看似合理的结构。这个方法完全类似于我们对体相材料考察几种已知晶体结构的方法。但是这种方法的严重局限是随着原子数量的增加,预期所能得到的原子排列数量会急速增长。另一个局限是采用该方法进行的计算,本质上会偏向那些“熟悉的”结构。如果最稳定的纳米颗粒具有我们完全没想到的形式,那么这个最稳定结构就很难用这个方法发现。以上几点就是那些具有复杂能量表面(且具有很多极小值点)物理问题的特点。

从头算MD提供了能够对复杂能量表面进行补充搜索的有效方法,而该方法所建立的思路类似于“围捕嫌犯”。在这类MD应用中,计算目的不一定是生成一个完全处于Born - Oppenheimer势能面上的轨迹。相反地,在高温模拟中可以使用较大的时间步长,其目的就是快速生成一个(克服两个能量极小值之间能垒的)近似轨迹。

在开始这类计算之前,需要进行如下大致估计:在感兴趣的事件发生之前,我们需要等待多久。如果我们想要看到的事件可以被描述为跳过一个能垒,我们就可以使用过渡态理论,把该过程的速率近似为 $k = \nu \exp(-\Delta E/k_B T)$,而观察到这一事件所用的平均时间是 $\tau = 1/k$ 。如果我们所感兴趣的是1eV能垒所分隔开的两极小值间的跳越,且 ν 的预估值为标准值

10^{13} s^{-1} , 则 τ 在 1000K 下为 10^{-8} s , 而在 2000K 下为 10^{-11} s 。这意味着对于一个 MD 模拟(其中时间步的测量单位为飞秒), 如果在 1000K 下运算则肯定是毫无结果的, 而在 2000 K 下运算则会产生有意义的结果。

关于使用从头算 MD 方法考察势能面, 给出一个例子: 使用该方法考察 13 个 Pt 原子纳米团簇的结构^①。图 9.5 显示了“随手”建立的三种高对称性 Pt₁₃ 纳米团簇, 并使用 DFT 方法进行了优化。大致呈球形的团簇具有 O_h 对称关系, 且本质上就是体相 FCC 晶体结构中含有 13 个原子的一个片段。从直觉上来看, 这种结构似乎应该是 13 个 Pt 原子的优先结构, 但是, 计算结果表明并非如此。在图 9.5 所示的计算中, 相比于最低能量构型, 呈 O_h 对称关系的团簇其能量要高出 2.5eV。

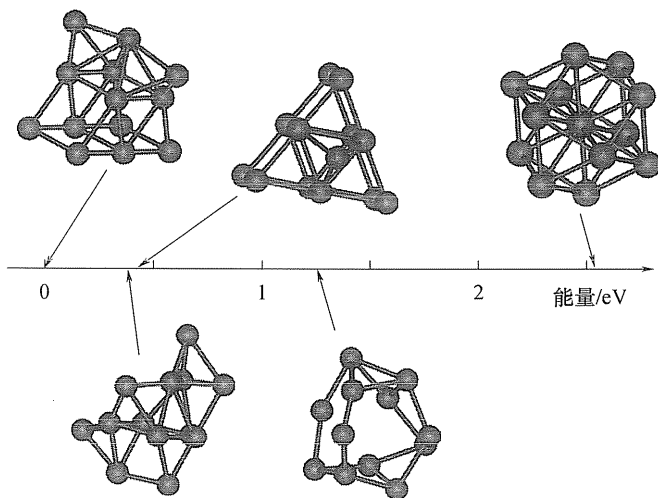


图 9.5 采用 GGA DFT 方法确定的 Pt₁₃ 团簇在其势能面极小值点的 5 种构型。图中上侧的 3 种团簇是基于对称性考虑“随手”建立的; 下侧的两种团簇是使用从头算 MD 方法(如正文所述)得到的。每一种团簇的能量值都是该结构相对于能量最低团簇的能量差。

① 该例取自于 Wang L L, Johnson D D. Density Functional Study of Structural Trends for Late Transition Metal 13 Atom Clusters. Phys. Rev. B 75, 2007: 235405 中对过渡金属纳米团簇的广泛研究。

为了寻找处于能量极小值上的更多 Pt_{13} 构型,我们使用从头算 MD 方法,在 2000K 下使用 20fs 时间步长,以 O_h 对称关系的团簇作为初始条件,进行了时间总长为 3ps 的模拟。在该计算中,使用 Nosé - Hoover 恒温器来控制温度。该计算中,纳米团簇在每一步 MD 模拟完成后的能量(不包括动能)如图 9.6 所示。该图显示出:在该计算的 MD 轨迹中,纳米团簇尝试了很宽能量范围内的不同原子构型。为了找到 Pt_{13} 势能面上的极小值,对 MD 模拟过程中所观察到的较低能量构型,使用标准的能量最小化方法进行了进一步优化。对图 9.6 轨迹中 47、95 和 147 步完成后的原子构型,分别进行上述计算后,又找到了两个不同的极小值点。这两个极小值如图 9.5 所示。这两个结构全都是非对称性的;因此,“随手”生成这两个模型在本质上是不可可能的。更为重要的是,二者的能量更接近于能量最低的对称性团簇。在 Wang 和 Johnson 对该体系的全面研究中,他们确认了 18 种不同的极小值点,且其能量均处于图 9.5 所示范围。其中:12 种结构是基于对称性结构生成的,而其余的 6 种均是使用从头算 MD 方法找到的。

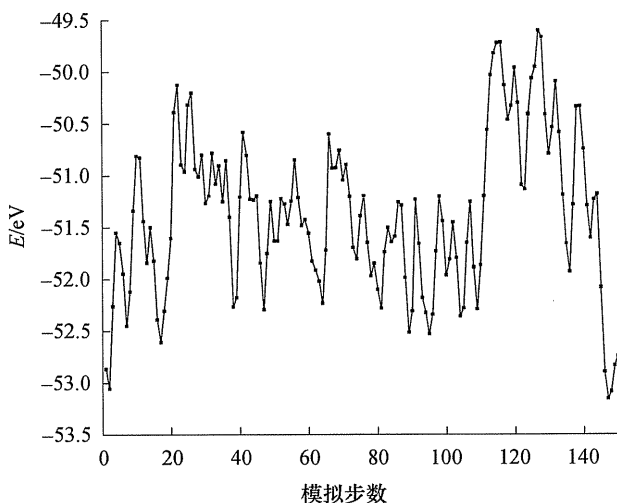


图 9.6 采用从头算 MD 方法模拟的 Pt_{13} 纳米团簇的能量

尽管采用上述方法,使用从头算 MD 可以生成其他方法难以找到的极小值点,但这并不能保证这类计算能够找到所有物理上的相关最小值。当然,这个警告也适用于以下情形:基于结构对称性(或者根据预期可能存在的结构列表)来寻找能量极小值。在对 MD 模拟所确认的结构进行解释时,必须记得这个警告。总之,从头算 MD 可以成为探索复杂势能面的强大工具。

练 习

1. 在第 6 章的一个练习中,估算了体相 Si 中电中性空穴跳动时的活化能。对于该体系,使用同一个超晶胞进行从头算 MD 模拟,只在 Γ 点上对倒易空间取样。首先,确定需要使用多大范围的时间步长,才能使用能量恒定的 MD 方法,生成体系总能(电子能量加动能)严格守恒的轨迹。观察在模拟过程中空穴是否跳动,以及该结果与你(根据前面对该过程所估算的活化能)所想象的是否一致。

2. 计算图 9.5 中所示 P_{13} 的最低和最高能量结构。对于该纳米团簇由 Wang 和 Johnson 所确认的对称结构,计算至少其中一种结构的能量。使用从头算 MD 方法,以能量最低的对称结构作为起始点,至少找到 P_{13} 的一个非对称性极小值。

参 考 文 献

- [1] Lewis L J, Vita A De, Car R. Structure and Electronic Properties of Amorphous Indium Phosphide from First Principles. *Phys. Rev. B* 57, 1998:1594.

扩 展 阅 读

关于了解经典 MD 模拟中的算法,两个极好的参考文献是:Allen M P, Tildesley D J. *Computer Simulation of Liquids*. Clarendon, Oxford, UK, 1987 以及 Frenkel D, Smit B. *Understanding Molecular Simulations*:

From Algorithms to Applications. 2nd ed. Academic, San Diego, 2002。

关于从头算 MD (*ab initio* MD) 的算法, 更多详细信息可以参见: Payne M C, Teter M P, Allan D C, et al. Rev. Mod. Phys. 64, 1992: 1045 以及 Martin R M. Electronic Structure: Basic Theory and Practical Methods. Cambridge University Press, Cambridge, 2004。

使用 DFT 方法, 对从头算 MD 的算法进行改进, 现在仍是一个活跃的研究领域。有关这方面的著作, 请参见: Kühne T D, Krack M, Mohamed F R, et al. Phys. Rev. Lett. 98, 2007: 066401。此外, 也有讲述使用高级量子化学技术来进行 MD 计算的类似著作, 例如: J M Herbert, Head - Gordon M. Phys. Chem. Chem. Phys. 7, 2005: 3629。

在大量的相关文献中, 仅以下列研究为例, 给出几个从头算 MD 方法的应用。对于模拟过冷液体合金的讨论, 请参见: Widom M, Ganesh P, Kazimirov S, et al. J. Phys. : Condens. Matter, 20, 2008: 114114。分别用 NEB 和 MD 计算表征固体中的原子迁移, 并进行内容比较, 请参见: Wang Z C, Gu T K, Kadohira T, et al. J. Chem. Phys. 128, 2008: 014704。一个对 Si 中液 - 液相变的研究, 参见: Jakse N, Pasturel A. Phys. Rev. Lett. 99, 2007: 205702。用 MD 方法生成的原子构型计算系综平均特性 (具体来说, 就是体相 GaAs 的光学响应), 参见: Ibrahim Z A, Shkrebtii A I, Lee M J G, et al. Phys. Rev. B 77, 2008: 125218。对于模拟计算许多不同液体金属的径向分布函数, 并讨论了 MD 计算 (其中采用了 GGA 泛函) 的精度, 参见: G Kresse. J. Non - crystalline Solids 312 - 314, 2002: 52。

附录 计算细节

所有对 Pt₁₃ 团簇的计算都使用了 PW91 GGA 泛函, 截断能均为 191eV。每个计算中, 都把 Pt₁₃ 团簇放置在边长为 20Å 的立方超晶胞中。只在 Γ 点上对倒易空间取样, 而且使用了 Methfessel - Paxton 方法, Smearing 宽度为 0.1eV。在几何优化中, 所有的原子均可以自由弛豫, 直到作用在每个原子上的力小于 0.03eV/Å。

第 10 章 精确度及“标准”计算以外的其他方法

10.1 DFT 计算有多精确？

本书始终关注的都是平面波 DF 计算,每当开始进行这类计算时,肯定会关心其计算精度有多高。归根结底,电子结构计算的目的就是预测材料的性质,并且这些预测值能够与已经(或尚未)完成的实验测试结果进行比较。每当你在任何一种学术场合介绍 DFT 计算结果时,都可能会被听众问道:“DFT 计算有多精确?”

是的,DFT 计算有多精确?这是需要弄明白的一个极其重要的问题。尽管这个问题简单明了,但并没有被很好地提出。精确度这个概念包括了许多方面,每个方面都需要单独地加以思考,尤其需要说明的是应该把物理精确性(Physical Accuracy)和数值精确性(Numerical Accuracy)分开。当提到物理精确性时,目的是想要知道:对于某个 DFT 计算所预测的物理特性而言,相对于该特性以完美实验测试(常常是假设的)测得的实际值,预测值有多精确。与此相反,数值精确度所评估的是:对于由 Kohn - Sham (KS) 方程定义的数学问题,某个计算是否提供了收敛良好的数值解。如果你将要进行 DFT 计算,那么你日复一日的大部分精力都需要集中在数值精确度的问题上。然而,在把你的计算结果与更大的科学结论之间联系起来的时候,各方面的物理精确度就是最重要的了。图 10.1 描绘了这两方面的精确度。

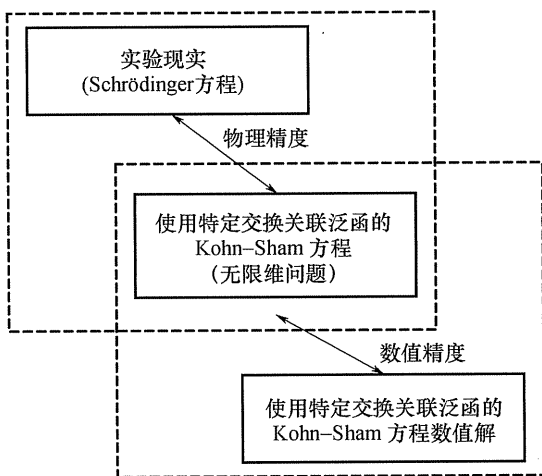


图 10.1 物理精度和数值精度区别的示意图

我们首先来讨论数值精度问题。对于任何 DFT 技术,其核心本质就是对于位于确定位置上的一组原子核,计算它们所对应的电子基态能量。从数学上来看,我们在平面波计算中的目的就是对于空间周期性的原子核集合,使用特定的交换关联泛函,求解其 KS 方程 [式(1.5)]。严格说来,这是一个无限维度的问题,因为 KS 方程中的单电子波函数是空间坐标的连续函数。由于我们必须在有限时间内,使用有限的计算资源,在计算机上找到解,所以我们就必须寻找一种数值近似解^①。为了使计算更易于在计算机上处理,会采用一些近似方法,而要评估某个计算的数值精度,最好是分别考察这些近似方法。对于平面波 DFT 计算,这些近似包括了:用来对 k 空间取样的 k 点数量和密度,以及用来定义平面波基底大小的截断能。这两个选项的涵义是较容易理解的:采用足够多的 k 点和足够高的截断能,可以得到收敛于数学真实解的数值解。即使当 k 点数量和截断能已经固定时,仍然必须采用迭代方式求解 KS 方程。于是该迭代过程的输出结果将向着方

^① 当然,并非只有 DFT 计算才是如此;对于任何无法解析求解的积分或微分方程,都存在类似问题。

程自洽解不断收敛,直到满足特定的收敛标准。这些内容早已在第3章进行了详细说明。需要着重说明的是:提高数值精度的级别,就不可避免地要付出计算时间更长的代价。对于所研究的许多实际问题,必须明智地在这些竞争性需求之间取得平衡。

如果我们把这数学问题总结为不断寻找原子核和电子的基态能量极小值,而这个极小值又是由KS方程和选定的交换关联泛函所确定的,那么就可以展现出这个颇为微妙的数值精度所具有的重要性。为了求解这个问题,先在原子核的当前位置下,找到KS方程的数值解,然后使用迭代优化方法更新原子核的位置。这个数值问题的常见特征已经在第3.3.1节进行了总结。这些计算最重要的特点也许就是不断改变原子核位置来寻找总能的极小值点。虽然这些计算可以给出可靠的结果,但是它们无法确保找到能量的全局最小值点。

在许多情形下,可以用基态能量定义一个从属性的物理量,而它要比能量本身的意义更大。对于这些物理量,也存在数值精度的问题。这里可以给出一个具体例子,读者可以回顾第5章振动频率的计算(特别是图5.1),为了从数值上确定振动频率,对总能的导数使用了有限差分近似。进一步来说,在极限条件下,当有限差分近似所用步长消失为零时,这个近似就变为精确相等。这意味着步长太大就会导致不精确的结果。但是,如果所用的步长太小,则又会由另一个原因引起数值的不精确性,即由KS方程迭代求解过程所引发的有限精度。这个例子也从更加普遍的角度上说明:为了控制计算中所得结果的数值精度,明确和理解数值不精确性的来源是很重要的。

根据数值精度的讨论,可以引出以下几个描述DFT结果时所应遵循的准则:

(1) 详细说明所使用的交换关联泛函,从而精确定义所求解的数学问题。由于存在有许多不同的泛函,因此,只陈述泛函种类(LDA、GGA等)是不够的,必须准确而具体地指出所用泛函。

(2) 说明在进行计算时所使用的数值细节。有关这些细节,我们这里给出一个并不完全详尽的列表:超晶胞的尺寸和形状, k 点的数量和位置,迭代求解KS方程的方法和收敛标准,以及原子核位置的优化

方法和收敛标准。关于是否列出这些细节内容,一个有效的度量标准就是:根据这些信息,应该使远隔重洋的某个人能够再现你的计算,并精确地得到相同的数值结果。

(3) 如果你要比较数值近似等级不同的计算结果,就必须非常小心。例如,对于使用了不同 k 点数量的计算,如果计算它们之间的能量差,就会在结果中引入系统误差。对于所有的计算都使用相同的数值方法,就可以避免这一问题,无论何时你都应该这样做。当这一问题无法避免时(例如,当你把自己的计算结果与已发表结果进行比较时,并且已发表结果中的数值方法对你而言是无法获得或无法实现的),就需要清楚地说明这两个计算之间的区别。

(4) 当能够把你的计算结果和实验观测之间进行明确对比时,就应该进行这个比较。

在表述你的计算工作时,你对上述准则的陈述方式应该随着场合不同而有所变化。在科技论文中,几乎没有任何理由可以不对准则(1)和(2)进行详细说明,至少在(几乎所有期刊都免费提供的)附录信息中都可以做到。但是,当出席一个会议或座谈会时,在作报告时请不要按准则(2)进行说明,因为在房间前面读一长串列在幻灯片上的数值计算细节,可能需要从房间后排使用望远镜才能看清。除非听众非常重要,房间里绝大部分人主要关心的是你计算结果的物理涵义,而且你需要让他们确信:在计算工作中你已经足够细心谨慎,因而所得到的是足够可靠的信息!在这种情形下,一个好的策略就是:准备一个包含所有数值细节的幻灯片,从而既可以非常简略地讲解,也可以在回答有关计算细节的提问时作为附加幻灯片来使用。

现在我们来看看物理精确度。正如“精确度”不应该被看作是单独的某一方面,“物理精度”也是一个非常模糊的概念,无法非常清晰地对其进行考察。如果问 DFT 预测对所研究的具体特性有多精确,答案有可能会更好。这是因为,DFT 结果可能对某些物理特性是相当准确的,但是对相同材料的其他物理特性则相对较不精确。举例说明,平面波 DFT 计算精确地预测了晶态硅的几何结构,并可以准确地将其划分为半导体材料,但是正如我们在第 8 章所看到的那样,DFT 计算极大

地高估了硅的带隙。

物理精确度的第二个重要方面,具有明显的社会特点:为了进行一个有意义的科学判断,需要对所研究的特性进行多精确的预测?对这一重要问题的答案可能会随着学科领域(甚至学科分支)的不同而显著变化,并且主要决定于这些计算所得信息被用来做什么。答案也会随着时间变化;该技术在 5、10 或 20 年以前的研究方式常常会被新的研究方法所取代。相比于实验领域,在材料计算及模拟方面由于出现新方法而引致的思想演变可能会越来越快,尽管这种演变对两个领域都很重要。

在许多应用领域中,可以用计算来考察具有实际意义的特性,从而在一系列不同材料中进行筛选。如果 DFT 计算(或其他任何种类的理论方法)能够可靠地预测出某个特性的数值结果图线,且材料 A 显著要比材料 B、C、D 或 E 更好,那么这些计算就极有意义。换句话说,在不同种类的材料中,如果可以精确预测出某个特性所具有的变化趋势,这也许要比预测某个材料特性的具体数值更加重要。但是,在某些情形下,对于单种材料的某个具体性质而言,预测的精确性也许是更加重要的。例如,对于在复杂材料的振动谱中所观察到的振型,如果使用计算作为辅助来确定其振动型式,就很容易理解为什么我们期望计算结果应该位于 0.01 cm^{-1} 、 1 cm^{-1} 或 100 cm^{-1} 这些真实振型的范围内。

为了对物理精确度的第二方面进行有目的地判定,就必须熟悉你所感兴趣研究领域内的最新文献(包括计算和实验两方面的)。在这样做的时候,应该对你所找到的信息进行价值判断,判定这些信息到底有多“好”。如果你没能对某个领域内哪些研究相比于其他研究是更有价值的这一决定,那么你很可能会有一个困难时期,即需要对如何开始你自己的计算做出一个正确决策^①。

在考虑把 DFT 计算应用在任何应用科学或工程领域时,应该对以下问题做出如实评估:在计算所能得到的物理特性与所感兴趣的应用

^① 正如电影 *Spinal Tap* 中一个虚构的摇滚音乐家所说的,“我相信几乎所有我读到的事情,而且我认为:相比于那些不相信任何事情的人,正是这一点让我成为更有选择性的人。”

之间,到底存在怎样的关系。一个有效方法就是列出可能影响一个“真实”材料性能的物理现象,而 DFT 计算中实际上并没有包含这个“真实”材料。我们可以肯定:你自己在这个过程中进行严格谨慎地仔细检查,远远好过于请其他持怀疑态度的实验者为你完成这一过程(这样做的通常结果就是他们所得结论正是你所谈到的内容,抑或是请求你的资助)!

例如,假设我们需要使用 DFT 计算来探索某些大体积化合物多相催化的实际工作。已经有许多科学工作者对该任务进行了大量研究。对于“标准的”DFT 计算,就是考察主要的化学中间体在扩展金属表面上的吸附和反应,其中金属表面就是由气相区(用真空近似代替)所包围的催化剂。对严格控制的实验和计算结果之间进行仔细比较,可以得出以下结论:DFT 计算可以对该情形进行许多强大的、定量的详细研究。但是,工业上的多相催化通常是氧化物支撑的金属(或半金属)纳米颗粒,且具有辅助性的局部和微观结构。催化剂也常常掺杂“促进剂(Promoters)”,例如经验上已经发现碱金属原子可以改善催化剂性能。同时,催化剂是在中压到高压下工作,且处于复杂的气体混合物中(或者是处于溶液中)。此外,我们所期望的反应通常只是很多可能反应中的某一个反应,而其余反应所产生的是我们并不期望得到的副产物。在经过明智的建模和参数选择后,完全可以使用计算来考察这些真实材料中的某些(并非全部)复杂性(而且肯定无法同时考察这些复杂性)。这并不意味着计算在这个领域中没有价值^①,反而显示出我们需要学会的一项重要技巧:充分理解在哪些环境中能够使用这些计算。

当考虑 DFT 计算(也可以笼统地说成是量子化学)的物理精度时,前面的讨论突出了一个几乎总是普遍存在的矛盾。一方面,对于严格定义的问题,通常总是想要从较不精确的计算方法换成更加精确的方法,而其目的就是预测“真实的”结果。这个前进过程几乎不可避免地会产生以下结果:随着所采用理论等级的提高,所需要的计算工作量也

^① Karl Popper,著名的科学哲人曾经说过:“科学可以被描述为有条理的过于简化的艺术。”

会快速增加。而另一方面,科学上的关注点越来越宽泛,这也就意味着:存在一个强烈需求,总是想要考察由更大和更复杂原子集合所定义的体系。但是,增大体系尺寸也就会增加计算负担,表现为需要更多的计算资源。采用计算方法所得到的重要科学贡献,肯定来自于符合以下条件的研究工作:充分理解本领域内的最新知识,能够明确与所研究问题密切相关的物理现象,了解所使用计算方法的能力和限制,从而能够谨慎地平衡上述所有需求。

10.2 选择一个泛函

图 10.1 中着重说明:只有确定了所用的交换关联泛函,才能用密度泛函理论描述一个具体问题。在选择使用交换关联泛函之前,需要懂得各类常用泛函之间的相似点与不同点。John Perdew 及合作者^[1]给出了一个很有用的泛函分类方法,而所参照的就是圣经《创世纪》28 章中有关 Jacob 的记载,其中 Jacob 说“…有个梦…在地面上竖起了一个通往天堂的梯子…上帝的天使们沿着梯子在上上下下”。图 10.2 显示了用 Jacob 的梯子来描述 DFT 泛函的分类。梯子上较高处的横木代表包含了更多物理信息的泛函,且沿着梯子不断向上就通往求解 Schrödinger 方程时不用任何近似的“完美方法”。为了进一步完善这个隐喻,可以说:人们应用 DFT 计算就如同天使们在各种合适的泛函之间来回移动。

尽管 Jacob 梯子是一种很好的泛函分类方法,但其他分类方法也很重要,因为这些方法明确区分了许多常用泛函的特点。这些分类方法中最重要的一种也许就是把泛函分为非经验的和经验两类。建立非经验泛函时,就已经满足了特定的限制条件,从而正好可以利用真实 Kohn - Sham 泛函已知的限制条件。对于图 10.2 中较高横木所对应的非经验泛函,考虑了更多的组成要素,同时也满足更多的限制条件。对于非经验泛函,可以合理地作出以下预测(总的说来):在满足更多限制条件的情形下,精确度也将会提高。经验泛函则相反,由于经验泛函是选择实验(或从头算)数据经过拟合得到的,所以含有许多在拟合过

程中所引入(或调整后)的参数。经验泛函也许(并且常常)会违背真实 Kohn - Sham 泛函的已知特点。这并非意味着经验泛函是不精确的;对于特定问题而言,这些泛函也能生成非常精确的结果,尤其是本书练习题所包含的那些相关体系和问题。然而,从图 10.2 中较低的横木到较高的横木时,也并非总如我们所预期的那样,即并不一定总会得到更为精确的结果,特别是当较高处的横木是一个经验泛函时就更是如此。

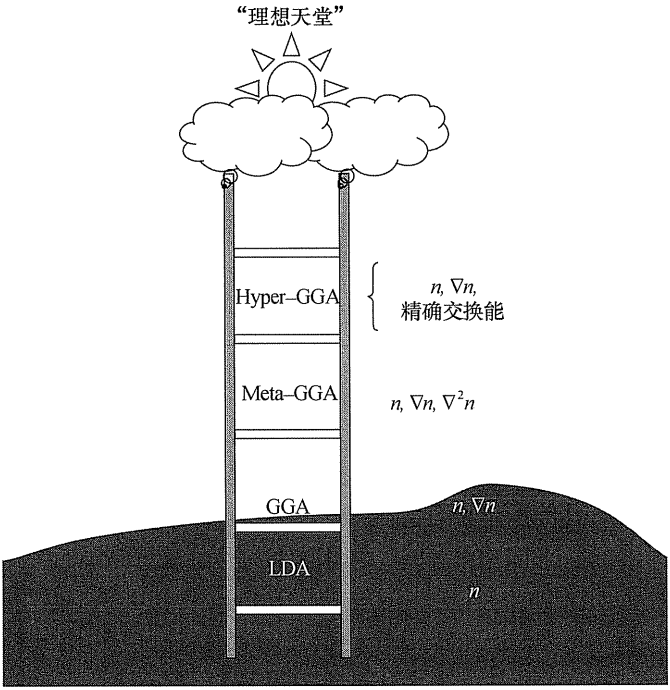


图 10.2 Perdew 使用“Jacob 梯子”对 DFT 泛函进行分类的示意图,梯子的横木从地面一级一级级向上到达“天堂”。对于每个横木所代表的泛函,其中所包含的物理因素如右侧所示

在本节的其余部分,我们依次考察图 10.2 所显示的每种不同方法,并对平面波 DFT 计算中使用最普遍的一些泛函进行简要综述。对真实 Kohn - Sham 泛函最简单的近似就是局域密度近似(Local Density Approximation, LDA)。在 LDA 中, Kohn - Sham 方程[式(1.5)]的局域

交换关联势被定义为空间内的均匀电子气交换势,其密度与局域电荷密度相同,即

$$V_{XC}^{LDA}(\mathbf{r}) = V_{XC}^{\text{electron gas}}[n(\mathbf{r})] \quad (10.1)$$

我们已经知道,均匀电子气交换关联泛函对所有的电荷密度 n 都具有较高精确度。在某些领域内,这些结果是经由详细的量子蒙特卡罗计算所确定的,这些密集计算方法能收敛到 Schrödinger 方程的精确解。实际计算中的 LDA 泛函使用了连续函数,可以精确地拟合 $V_{XC}^{\text{electron gas}}(n)$ 的已知数值。当这种方法刚刚被开发出来的时候,还开发出了另外几种不同泛函,它们实质上都是等效的,因为它们全部都精确地描述了 $V_{XC}^{\text{electron gas}}(n)$ 。真实 Kohn - Sham 泛函一个最基本的局限就是均匀电荷密度,可以将其简单描述为:在均匀电荷密度的极限条件下,交换关联能应该正好等于该密度下均匀电子气的能量。在建立泛函时,LDA 泛函正好就是均匀密度极限下的形式。因为许多体相材料中价电荷密度的变化都很缓慢,因此对于体相固体的任何泛函而言,都是满足这一限制条件的,这也是该类泛函能得以成功运用的关键原因。但是,对于原子团和分子中的电荷密度,总的说来并不是变化缓慢的,因此 LDA 并不能很好地预测原子团和分子的性质。

对于 Kohn - Sham 泛函的下一个近似就是广义梯度近似 (Generalized Gradient Approximation, GGA)。GGA 所包含的物理概念很简单;实际上的电荷密度并不是均匀的,因此在电荷密度中包含一些空间变化信息,所建立的泛函就能够更加灵活地描述实际材料。对于缓慢变化的电荷密度,基于 GGA 泛函的方程才是有效的。在 GGA 泛函中,交换关联泛函是用局域电荷密度和电荷密度梯度来表达的,即

$$V_{XC}^{GGA}(\mathbf{r}) = V_{XC}[n(\mathbf{r}), \nabla n(\mathbf{r})] \quad (10.2)$$

非经验的 GGA 泛函满足均匀密度极限。此外,它们还满足交换关联空穴 (Exchange - Correlation Hole) 几个确切已知的特性。满足这些特点的两个常用非经验泛函是 Perdew - Wang 91 (PW91) 泛函和 Perdew - Burke - Ernzerhof (PBE) 泛函。因为 GGA 泛函比 LDA 泛函包含有更多的物理因素,通常认为非经验的 GGA 泛函应该比 LDA 泛函更精确。虽然大多数时候这都是正确的,但也有例外的时候,其中一个例

子就是在计算过渡金属及其氧化物表面能的时候^[2]。

Jacob 梯子的第三根横木是 meta - GGA 泛函, 该泛函包含了来自于 $n(\mathbf{r})$ 、 $\nabla n(\mathbf{r})$ 和 $\nabla^2 n(\mathbf{r})$ 的信息。在实际计算中, Kohn - Sham 轨道的动能密度为

$$\tau(\mathbf{r}) = \frac{1}{2} \sum_{\text{occupied states}} |\nabla \varphi_i(\mathbf{r})|^2 \quad (10.3)$$

其中包含了与电荷密度拉普拉斯算子相同的信息, 并且使用这些物理量也带来了许多优点; 因此在 meta - GGA 中将用 $\tau(\mathbf{r})$ 代替 $\nabla^2 n(\mathbf{r})$ 。Tao - Perdew - Staroverov - Scuseria (TPSS) 就是一个非经验的 meta - GGA 泛函。

图 10.2 中梯子的第四根横木是很重要的, 因为在量子化学使用局域化基底的计算中, 最为普遍使用的泛函就属于这个等级。精确的交换能可以由交换能密度得到, 且可以用 Kohn - Sham 轨道的形式写为

$$E^{\text{exchange}}(\mathbf{r}) = \frac{1}{2n(\mathbf{r})} \int d^3 r' \frac{\left| \sum_{\text{occupied states}} \varphi_i^*(\mathbf{r}') \varphi_i(\mathbf{r}) \right|^2}{|\mathbf{r} - \mathbf{r}'|} \quad (10.4)$$

这个物理量的一个主要特征就是: 它是非局域的(Nonlocal); 换言之, 基于该物理量所建立的泛函不能对某个特定空间位置进行估值, 除非所有空间位置上的电荷密度都是已知的。如果回顾一下第 1 章中的 Kohn - Sham 方程, 就会明白: 在交换关联泛函中引入这个非局域量就会形成新的数值难题, 而在使用局域化函数时根本不存在类似难题。如果一个泛函中包含了来自于 GGA 的精确交换能, 就将该泛函归类为 hyper - GGA。

hyper - GGA 泛函使用精确交换能和 GGA 交换泛函的混合形式来描述交换能。目前这类泛函中最广泛使用的是 B3LYP 泛函, 它可以写为

$$V_{\text{XC}}^{\text{B3LYP}} = V_{\text{XC}}^{\text{LDA}} + \alpha_1 (E^{\text{exchange}} - V_X^{\text{LDA}}) + \alpha_2 (V_X^{\text{GGA}} - V_X^{\text{LDA}}) + \alpha_3 (V_C^{\text{GGA}} - V_C^{\text{LDA}}) \quad (10.5)$$

式中: V_X^{GGA} 是 Becke 88 交换泛函; V_C^{GGA} 是 Lee - Yang - Parr 关联泛函; α_1 、 α_2 和 α_3 分别是三个数值参数。这三个数值参数都是从经验上选定

的,从而对于较大分子测试集的特性而言[如键长、形成能(Formation Energies),等],可以使该泛函的性能得以进一步优化。B3LYP 泛函以及其他类似泛函已经在预测小分子特性方面取得了极大成功。需要说明的是,B3LYP 泛函并不满足均匀密度极限条件。因此,当使用该泛函对体相材料(尤其是金属)进行预测时,不太可能使该计算运行得特别好(事实上也确实如此)。

图 10.2 所示的泛函分类并没有把能够所有泛函种类都包括进去。任何一种合并了精确交换能的泛函称为混合泛函(Hybrid Functional)。因此,上文中提到的 hyper-GGA 泛函也可以称作 hybrid-GGA 方法。此外,也存在一类 Hybrid-meta-GGA 泛函——这些泛函把精确交换能和 meta-GGA 泛函相结合;图 10.2 中没有显示此类别的泛函。

如果按照图 10.2 的表面含义,可以把该图简单解读为:横木位置上较高的泛函总是优于较低横木处的泛函。其实并不一定是这样的!升高到更高横木时,就在泛函中包含了更多的物理因素。但是,每一行都会涉及到 DFT 结果和自然界真实结果之间的系统误差。对于所考察的某些特性而言,当从梯子的前几级向上移动时,这些误差并不是总会变得更小。例如,Litimein 等人使用不同的非经验泛函考察了几种 III-V 氮化物³;对于该研究中所考察的物理特性,用 TPSS 泛函(上面提到的非经验 meta-GGA 泛函)来计算是最合适的。此外,该研究结果还表明:LDA 计算结果通常比使用 PBE-GGA 泛函的计算更加精确。

当考虑如何把图 10.2 应用于经验泛函时,尤其需要在实际操作中更加地谨慎和细心。从经验泛函的开发设计上来说,开发这些泛函时使用了一些测试集(Test Set)。对于在化学上类似于这些测试集的物理问题,经验泛函可以极好地适用于这些情形。然而,对于与这些测试集存在很大不同的物理问题,在应用经验泛函时就需要非常小心。这里可以给出一个例子:对于几种矿石硫化物的晶体和电子结构计算,Von Oertzen 和合作者比较了 LDA、PBE-GGA 和 B3LYP-hyper-GGA 三种泛函^[4],这三种泛函所预测硫化物晶格参数相比于实验值的误差都在 1%~2% 之内。对于 PbS,使用 LDA、PBE-GGA 和 B3LYP 所得

到的能隙,相比于实验观测值的误差分别为 -0.1eV 、 $+0.1\text{eV}$ 和 $+0.7\text{eV}$ 。对于 FeS_2 ,两种非经验泛函低估了带隙值大约 0.5eV ,而 B3LYP 给出 2.2eV 的带隙则太大。

对于基于局域化基组的和基于平面波的 DFT 计算,引入混合泛函后就使两者之间形成了一个明显区别。由于在平面波基底中求解 Kohn - Sham 方程的数值细节引入了精确交换能的非局域性,从而大大增加了求解这些方程时的数值计算工作量。这个问题在使用局域化基组时则更加严重。因此,对于化学领域所使用的局域化基底计算程序,几乎所有泛函都只是包括了部分的精确交换能;然而,这些泛函目前还很难应用于具有周期性边界条件(或空间扩展材料)的超晶胞。在筛选混合泛函(Screened Hybrid Functionals)的开发工作中,正在出现了令人激动的进展;该方法中,交换作用可以分为两个组成部分:一个是长程的;一个是短程的,且仅在短程的部分应用了一定程度的精确交换能。HSE 泛函正是基于该方法,以 PBE 泛函为原型,并只在其短程部分混入了一部分精确交换能。本章末的扩展阅读部分列出了几个有关该领域更多信息的参考文献。

对以上交换关联泛函的简要说明进行总结,我们重申上述观点:你必须花时间理解当前在你所感兴趣研究领域内的最新文献,才能判断出哪个泛函适用于你自己的工作。

10.3 物理精度的例子

10.3.1 分子体系的基准计算——能量和几何构型

对于分子体系量子化学中使用局域化基底的 DFT 计算(或者其他的量子化学方法),与其物理精度有关的文献是非常多的。Sousa、Fernandes 和 Ramos 最近评述了这方面的文献^[5],讨论了几十个基准性的研究,其中比较了多种 DFT 泛函的性能表现。表 10.1 总结了在用 B3LYP 泛函进行计算时的常见结果(该泛函是局域化基底计算最常用的泛函)。该表把许多人所完成的大量计算凝练成了几行数据,因此

为了更详细地理解与这些计算相关的细节,建议读者直接参考 Sousa、Fernandes 和 Ramos^[5]所引用的相关文献。

表 10.1 B3LYP DFT 计算在小分子测试集上的常见精度^a

物理特性	典型的平均绝对误差	测试集中的典型数值
键长 (Bond Length)	0.01 Å	1.1 ~ 1.5 Å
键角 (Bond Angle)	1°	110°
能垒高 (Barrier Height)	3 ~ 4 kcal/mol	12 ~ 25 kcal/mol
原子化能量 (Atomization Energies)	1 ~ 2 kcal/mol (对于金属二聚体为 10 ~ 20 kcal/mol)	500 kcal/mol (对于金属二聚体为 58 kcal/mol)
结合能 (键能) (Binding Energies)	5 ~ 10 kcal/mol	60 ~ 80 kcal/mol
形成热 (Heats Of Formation)	5 ~ 20 kcal/mol	50 kcal/mol
氢键强度 (Hydrogen Bond Strengths)	0.5 ~ 1 kcal/mol	8 kcal/mol

^a 摘自 Sousa、Fernandes 和 Ramos 的文献综述^[5]。原子化能量、结合能和形成热的数据是平均到每个分子上

当考虑表 10.1 所列数值时,需要记住以下几点:结果的精确度是由平均绝对误差所表征的,换言之,就是对实验值和测试值之间差值的绝对值取平均值。在每个测试集中,一定存在某些值的误差显著高于平均值,并且它们通常并不是随机分布于整个测试集上的。也就是说,如果过分密切关注平均绝对误差,就会忽略掉 B3LYP(和其他 DFT 计算)精度中不同分子之间所存在的规律趋势。表 10.1 中所表述的原子化能量就是这方面的一个简单例子,对于金属二聚体的结果,其精确性明显低于不含金属的小分子。

表 10.1 中结果的第二方面就是:在所有测试过的 DFT 泛函中,B3LYP 的表现并不一定是最好的。在 Sousa、Fernandes 和 Ramos^[5]所列出的 55 个基准研究中,B3LYP 泛函仅在 4 个研究中具有最好的性能

表现(以平均绝对误差所表征的)。在至少一个研究中,B3LYP 在所有测试过的泛函中表现最差。而且,在这些基准研究中,并不存在某个泛函对所考察的所有(或是绝大多数)性质总是“最优的”。然而,作为一种为局域化基底计算所精心开发的 DFT 泛函,可以预期:B3LYP 的结果在精度等级上是具有代表性的。

10.3.2 分子体系的基准计算——振动频率

所有量子化学方法(不仅仅限于 DFT)计算的谐振动频率(Harmonic Vibrational Frequencies)往往高估于实验值。出现高估的一部分原因是由于实际振动的非简谐性,另一部分原因是由于 Schrödinger 方程解的非精确性本质。可以对振动频率计算出非简谐修正值,其过程是:使用超过二阶(简谐)项以上的泰勒展开式,在能量极小值附近拟合势能面。但是该方法很少使用。在分子体系量子化学计算中,常用的方法是:假如在大量计算实例中都使用相同等级的理论,则简谐频率计算值和实验观测频率之间的比值变化很小。这一结果意味着:当与实验值进行比较时,可以将这些计算所得的振动频率值乘以一个比例因子,这样的计算数据解释方式是较合理的。对于使用局域化基底的计算,Merrick 及合作者已经给出了一个大量比例因子的列表,且适用于许多不同的理论级别,包括多种 GGA 和 hyper-GGA 泛函^[6]。这些比例因子通常处于 0.95 和 0.99 之间;换言之,在这些计算中振动频率只被高估了 1%~5%。对于平面波 DFT 计算频率和实验数据之间的类似比较,尽管还没有进行与上述局域化基底计算同等详细的研究,但对于平面波计算而言,上述 1%~5%的精确度范围也是较常见的。

10.3.3 晶体结构和内聚能

在以下的两个部分中,我们说明如何用一系列计算考察平面波 DFT 计算的物理精度。平面波计算的一个主要特点就是:它们可以应用于体相材料,以及在计算上无法使用分子体系量子化学局域化方法的其他情形。为了对平面波方法的性能进行评价,必须把其计算结果与精确的实验数据相互比较。分子体系量子化学计算的基准测试工

作之所以会如此成功,其中一个原因就是:对于小分子而言,已经有非常大量的高精度实验数据。但是,对于平面波 DFT 计算中的某个特性,并不总是存在类似规模的实验数据集,而这就限制了对平面波 DFT 预测值和实验值(其中的实验信息应该得自于大量不同材料)进行比较。当然,目前已经与单独实验测试之间进行了成百上千种个体比较。如果读者遵循我们的建议,并且逐渐熟悉了所研究领域的最新文献,那么你就能找到类似例子。下面,我们给出几个例子,其研究工作就是在平面波 DFT 计算和系统的实验数据集之间进行比较。

Milman 及合作者用 PW91 GGA 计算,优化了 148 种无机晶体和分子化合物的结构,其中包含了周期表中 81 种不同的元素^[7]。这些计算包括氧化物、氮化物、碳化物、碱土金属、过渡金属以及半导体。总的来说,相比于实验值,晶态固体的晶格参数和分子类物质的键长均具有 2% 以内的误差,二者都在实验数据的基础上出现了正偏差和负偏差。在 148 个例子中,只有 5 个例子的误差(相对于实验数据)大于 3%。类似的精度等级也出现在 Alapati 等人对轻金属氢化物和硼化物大量晶体结构的 PW91 GGA 计算中^[8]。从这些结果可以得到的总观测结论,与局域化基底计算(参见表 10.1)是一致的,即 DFT 计算能够对大量不同种类的材料,可靠地给出其几何结构的精确信息。

对于元素周期表其中 4 列的 11 种单质固体,Philipsen 和 Baerends 使用 PW91 GGA 计算方法,分别考察了内聚能和体模量^[9]。在他们对这些材料的研究结果中,内聚能和体模量的平均绝对误差分别为 0.35eV 和 0.15Mbar。相比于 LDA 计算的相应结果,上述误差较小(在平均值上);但是,对于 LDA 的某些特定结果,要比 GGA 计算更符合实验值。Korhonen 及合作者采用 PW91 GGA 计算研究了 12 种金属单质中的空穴形成能^[10]。在不同的实验测试值与部分(虽然肯定并非全部)金属的预测值之间,该物理量存在着 $>0.3\text{eV}$ 的差值。如果使用实验值的平均值作为比较基准,12 种金属计算值的平均绝对误差是 0.30eV。

10.3.4 吸附能和成键强度

对于预测金属表面上小分子吸附能这一方面,许多研究团队都对此进行了系统研究,使用不同的 DFT 泛函并比较其表现。Hammer、Hansen 和 Nørskov 使用 5 种不同泛函,分别考察了 CO 吸附在 5 个金属表面,以及 NO 吸附在 4 个金属表面^[11]。总的说来,DFT 高估了这些吸附能,使用 LDA 泛函时要比使用各种 GGA 泛函所得结果更加糟糕。PW91 GGA 泛函所给出的 RMS 计算值和实验值之间偏差为:对 CO 为 0.78eV;对 NO 为 0.52eV。当使用 PBE GGA 泛函时,上述 RMS 偏差分别稍稍降低到 0.67eV 和 0.43eV。Hammer 等人介绍了一个 PBE 泛函的修正版,就是现在所谓的 RPBE GGA 泛函。相比于前面那些泛函,该泛函能够显著改善对 CO 和 NO 吸附能的预测值,RMS 偏差分别为 0.37eV 和 0.22eV^[12]。通常认为:相比于 PW91 和 PBE 泛函,RPBE 泛函能够改善对化学吸附的描述。

对于金属表面上 CO 吸附能的 DFT 预测值和实验数据,Abild - Pedersen 和 Andersson^[13]给出了一个更为广泛的比较,其中包括了 16 种不同金属表面的吸附数据。与前面 Hammer 等人的比较有所不同的是:通过在某个可能的特定表面上进行多个实验,该研究还包含了有关实验测试吸附能时不确定因素的信息。对于大多数表面,估测得到这些不确定性的数量级为 0.1eV。对于具有多个实验结果的情形,在比较 DFT 结果时,所使用的实验值是这些实验结果的平均值。对于使用 PW91 GGA 泛函的计算,在 DFT 结果和实验测试吸附能之间的平均绝对偏差是 0.48eV;当平均偏差为 -0.46eV 时,意味着在所有计算中,该泛函都对 CO 产生了过束缚(Overbind)。使用 RPBE GGA 泛函的类似计算给出的平均偏差为 -0.12eV,且平均绝对偏差为 0.25eV。Abild - Pedersen 和 Andersson 也对 CO 的吸附能修正值进行了测试,其方法是在 DFT 泛函计算时考虑了 CO 的分子轨道特性。把这一修正应用于 RPBE GGA 计算结果,则 CO 总吸附能的偏差为 0.06eV。这是一个令人鼓舞的信号:当预测分子吸附能时,使用 GGA 泛函所带来的相关系统误差是可以降低的;但是,该方法仍未发展完善到能普遍适用于

所有的分子吸附问题。

对于 DFT 方法在计算吸附能时的系统不精确性,一个概括性的解释就是:这些计算必然涉及到在化学非近似态间的比较,比如:比较一个隔离态分子和一个结合在表面上的分子。对于这类情形,一个更为简单的例子就是比较简单分子的成键强度,这里给出一个具体例子: Fuchs 等人¹⁴用不同的 DFT 泛函预测了 N_2 的键能(作为一个更大范围研究的一部分)。他们发现, LDA 泛函过束缚了 N_2 (1eV)。当使用 PBE 和 RPBE GGA 泛函时,这个过束缚效应可以分别减小到 0.5eV 和 0.2eV。并不是只有平面波 DFT 计算会在预测键能时出现不精确性,如表 10.1 所示,在使用局域化基底 DFT 方法计算分子能量时也存在类似现象。

对于简单分子,结合能的高估会影响这些物质的从头算热力学计算。在第 7 章中,我们使用从头算热力学方法考察了体相氧化物材料的稳定性。为了定量地预测体相处于平衡时的条件,就必须使用式(7.6)计算 O_2 的化学势,其中包含了来自于 O_2 总能贡献。对于这类总能计算中 DFT 不精确度的一个校正方法就是使用所能够得到的实验数据。如果使用了这一种方法,一定要对此清楚地进行说明。

10.4 改善电子关联效应处理方式的 DFT + X 方法

在第 10.2 节中有关泛函的分级讨论中,突出强调了一个结果:即使当电子集合的交换能(理论上)能够被准确估测时,也无法准确估测其关联能。因此,很自然地就需要特别考虑一下:对于那些主要受电子关联作用影响的物理特性而言, DFT 计算对这些物理特性的可靠性是怎样的。这类问题中,有两个重要例子,分别来自于两个物理上看似完全不同的情况:原子之间的弱长程作用,以及强关联材料的电子结构。下面,我们将就这两个主题进行介绍,并简单描述可以将标准 DFT 进行扩展后用以处理上述情形的“附加项(Add - Ons)”。

10.4.1 色散相互作用和 DFT + D

在我们的日常生活中,色散相互作用(也称为范德华作用或伦敦力)具有难以置信的重要作用。不妨考虑一下我们所依赖的交通燃料:汽油或柴油。全球存在着庞大的交通基础设施,每天都在运送并使用这些燃料,而这些都基于一个事实:这些燃料是液体。非极性分子(如己烷)能够很容易地形成液体,这一事实标志着在己烷分子之间存在净吸引作用。这些吸引作用就是直接由电子关联作用所导致的。

对于电子关联作用和原子长程力之间的关系,最早在 1930 年代就由 London 进行了考察。他认为:在一个原子或非极性分子周围,尽管平均时间内的电荷密度没有偶极矩,电子的振动也会导致电荷密度的变形,从而引发一个瞬间偶极矩。这个瞬时偶极矩可以致使其他原子或分子的电荷密度发生扭曲,从而在其他原子或分子上产生短暂的偶极矩。两个偶极矩的存在就会形成一个总的相互作用。London 给出了相距较大距离的两个球形对称原子之间,该相互作用的一般表达式为

$$V^{\text{dispersion}} = -\frac{C}{r^6} \quad (10.6)$$

式中: r 是原子间的距离; C 是物理常数的集合^①。

能够说明色散相互作用的一个简单的实际例子就是稀有气体原子(如 He、Ne 和 Ar)的二聚体。众所周知,这些原子在化学上都很不活泼,但事实是这些气体可以在足够低的温度下液化,这清楚地表明了稀有气体原子之间存在吸引作用。Zhao 和 Truhlar 使用局域化基底计算,采用许多 meta - GGA 泛函描述了 He - He、He - Ne、He - Ar 和 Ne - Ar 二聚体^[15]。实验上测得这些二聚体的相互作用能所处范围为 0.04 ~ 0.13 kcal/mol,且平衡态下两个原子的间距所处范围是从 3.03 Å 到 3.49 Å。虽然二聚体肯定存在势能最小值,但是当使用 B3LYP 泛函

① 这些物理常数都包括一些相同信息:表征材料对于光波长的折射率色散关系,因此这些相互作用才称为色散相互作用。

时,能否找到该最小值也取决于该计算的细节信息。在没有进行平衡校正(Counterpoise Corrections)的计算中,平衡原子间距预测值的平均绝对误差为 0.13 \AA ,而相互作用能的平均绝对误差为 0.24 kcal/mol 。

另一个表明了 DFT 计算对色散相互作用具有局限性的简单例子是分子晶体的晶体结构。Neumann 和 Perrin 使用平面波 DFT 方法,采用 PW91 GGA 泛函考察了 20 种分子晶体的晶体结构^[16]。平均说来,用 DFT 方法预测的单体体积要比实验值大 20%。例如:晶态丁烷(C_4H_{10})的单体体积就高估了 27%;甲醇(CH_3OH)和对苯二甲酸($\text{C}_6\text{H}_4(\text{COOH})_2$)的晶态体积分别被高估了 14% 和 25%。这些结果与上面在第 10.3.3 节所讨论的晶态固体截然不同。在分子晶体计算中出现这个问题的原因是:在这些材料中,色散相互作用对分子间力具有重要贡献。

从概念上来说,对于 DFT 在处理色散相互作用力的缺陷,一个简单的补救方法就是在材料中每对原子之间的总能上添加一个与色散有关的贡献量。这一思想已经在局域化基底方法中开发出来,即所谓的 DFT-D 方法。在 DFT-D 计算中,原子集合总能可以用 DFT 方法计算的 E_{DFT} 增广为

$$E_{\text{DFT-D}} = E_{\text{DFT}} - S \sum_{i \neq j} \frac{C_{ij}}{r_{ij}^6} f_{\text{damp}}(r_{ij}) \quad (10.7)$$

式中: r_{ij} 是原子 i 和 j 的间距; C_{ij} 是原子 i 和 j 的色散系数,该物理量可以直接根据每个原子的属性列表计算出来; $f_{\text{damp}}(r_{ij})$ 是一个阻尼函数,用以避免色散项在很小距离时不符合实际的行为。该表达式中唯一的经验参数 S 是一个完全一致地应用于所有原子对的比例因子。在 DFT-D 的实际应用中,对于每种泛函需要事先分别估算该比例因子,其方法是:对于色散相互作用具有重要影响的一组分子复合物,不断优化该比例因子的数值。在平面波 DFT 计算中应用 DFT-D 的思想也不存在任何根本性障碍。在上面提到的 Neumann 和 Perrin 的研究工作中,他们显示出:对平面波 DFT 计算的力添加一个色散校正项,可以使计算所得分子晶体单体体积相比于实验结果的误差大致处于 1% 以内。在这些计算中,把平面波 DFT 程序和另一个用于优化晶体结构的

单独程序结合在了一起。目前,据我们所知,还没有任何平面波计算程序将色散校正作为一个“现成(Off The Shelf)”选项包含在内。

10.4.2 自作用误差、强关联电子体系和 DFT + U

粗略地说来,色散相互作用是与空间上彼此分离的相互作用电子密切相关的。DFT 的另一个系统问题是由电子自身相互作用所导致的,而这是不符合物理实际的。为了弄明白自作用误差的来源,需要来看看 Kohn - Sham 方程。在 KS 表达式中,通过求解一系列单电子方程来计算能量,这些方程为

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r}) + V_H(\mathbf{r}) + V_{xc}(\mathbf{r}) \right] \psi_i(\mathbf{r}) = \varepsilon_i \psi_i(\mathbf{r}) \quad (10.8)$$

式中:左侧第一项描述了电子的动能; V 是单个电子与原子核相互作用的势能; V_H 是 Hartree 电子 - 电子排斥势能; V_{xc} 是交换关联势能。该方法把电子 - 电子相互作用分为两部分:一个是经典部分,即由 Hartree 项所定义;而所有其他内容,全部都包含到交换关联能项。Hartree 势能描述了电子和体系总电荷密度之间的 Coulomb 排斥作用,即

$$V_H(\mathbf{r}) = e^2 \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3 r' \quad (10.9)$$

这个势能中包含了在电子和其自身之间一个不符合实际的排斥作用(因为该电子对总电荷密度是有贡献的)。与这个不符合实际的相互作用有关的能量就是自相互作用能。

在 Hartree - Fock(HF)方法中,Hartree 势能中这个虚拟的自相互作用能,恰好被交换作用对能量的贡献给抵消掉了。如果我们能知道精确的 Kohn - Sham 泛函,也可以在 DFT 中做到这一点。然而,在任何近似的 DFT 泛函中,都会由于自相互作用能没有被完全抵消而形成系统误差。

自相互作用误差的物理涵义可以非常容易地用 H_2^+ 这个单电子体系加以说明。在这个例子中,本来完全不应该有电子 - 电子相互作用,但该体系的 Hartree 势能仍然是非零的。由于自相互作用,DFT 计算无法对 H_2^+ 解离给出定性上的正确描述。从更普遍的意义上说,使用近

似的交换关联泛函时,自相互作用误差引发了高度局域化的 Kohn - Sham 轨道,从而无法正确地进行非平衡化。为了使自相互作用最小化,未配对电子倾向于在空间上去局域化。当存在电子局域化很强的电子态时,这些电子态就称为是强关联的。这类材料中最为人所知的就是 铜系元素和各种过渡金属氧化物,它们都含有部分充满的 d 或 f 电子壳层。对于某些 $3d$ 过渡金属,自相互作用误差可以导致 DFT 生成一个不正确的 $d^{n-1}s^1$ 构型,而不是 $d^n s^2$ 构型,从而在特定材料的内聚能计算中出现误差。

自相互作用误差 (Self - Interaction Error, SIE) 也会对计算固体中的缺陷产生影响。一个例子就是 Pacchioni 对 SiO_2 中 Al 置换缺陷的研究^[17],采用(无 SIE 的)HF 计算方法生成了 $[\text{AlO}_4]^\circ$ 中心的自旋密度图,对照于实验数据,该图正确地显示了空穴被困在靠近 Al 的一个 O 原子未成键的 $2p$ 轨道。对于使用某些泛函的 DFT 计算,会出现一个不正确现象:将空穴扩展到了邻近 Al 的所有 4 个 O 原子上。使用混合泛函时可以减少这种效应,因为这些泛函中所包含的精确交换能,在一定程度上修正了非混合泛函中的自相互作用误差。关于缺陷的自相互作用误差,另一个例子是:使用标准 DFT 泛函时,磁性绝缘体(例如 NiO)被描述为具有金属性^[18]。由于混合交换泛函(部分地)使用了精确交换能,它们可在一定程度上修正自相互作用误差。正如 10.2 节所讨论的,混合泛函目前正逐渐成为平面波计算软件中的一个选项。

在 HF 计算中恰好抵消了自相互作用误差,这一事实表明:如果把局域态的类 HF 方法和“其他所有情形”的 DFT 方法结合起来,也许是对电子强关联材料的一个可行方法。这一思想就是称为 DFT + U 这类方法的开发动机之所在。应用该方法的常见方式就是引入了一个对 DFT 能量的修正,即通过引入一个单数值参数 ($U - J$) 修正了电子自相互作用,其中 U 和 J 分别与自相互作用的不同方面有关。目前,使用 DFT + U 的数值工具已经在平面波计算软件中获得了广泛应用。

很明显,使用 DFT + U 方法需要先确定 $U - J$ 的具体数值。有两个常用方法可以解决这一问题,而它们各自都有优点和缺点。第一种方法是:先取得某些相关材料(通常是理想晶体)的一个已知特性,然后

对于给定的泛函,确定 $U - J$ 的哪个数值会对该特性给出最为接近的结果。第二个方法是:对一个测试体系使用其他种类的从头算方法,且对于该测试体系,可以用这些计算来估算 $U - J$ 。该方法通常给出的是 $U - J$ 的取值范围,而不是其具体数值。因为在指定 $U - J$ 数值时存在着不明确性,所以当运行 DFT + U 计算时,一定需要先确定计算结果对于这个参数值的敏感性。

10.5 线性比例方法的更大尺寸体系和经典力场

不难想象这样一些例子,如果能够对几千、几万个原子,甚至更多个原子进行定量计算,这将会是极有意义的。根据现在所掌握的 DFT 计算经验,读者应该可以得出以下结论:使用“标准”方法几乎不可能达到这个目标,因为这些计算所需要的计算工作量会随着体系尺寸的增大而急速增加。更具体地说,使用本书始终讨论的 DFT 方法,计算一组原子^①的总能所需要的计算时间为

$$T = cN^3 \quad (10.10)$$

式中: N 是原子个数; c 是与计算方法的各种细节有关的前置因子。这个比例表达式是反映大体系行为的一个渐进结果。当把我们的体系双倍扩展后,如果要使计算时间不变,那么唯一的方法就是使计算机的计算速度大致增加一个数量级。

对于式(10.10)所示比例关系,人们希望能够开发出具有更好比例关系的计算方法,而且已经付出了许多艰苦努力。该工作主要集中于所谓的线性比例方法(Linear Scaling Methods):对于单独的能量计算,其所需时间的比例关系为

$$T = c'N \quad (10.11)$$

相比于式(10.10),服从式(10.11)的新方法具有非常明显的优势,似乎只要这些新方法一出现就会把旧方法完全清除掉。但是,实际

^① 一个计算的所有相关时间不仅仅包括单独的一次能量计算,也包括正确表述所研究物理问题时所需要进行的许多次能量计算。

情形却并非如此简单。把线性比例方法和标准方法相互比较,总的特点就是:当使用相同的理论级别时有 $c/c' < 1$ 。也就是说,只有当体系尺寸足够大时,线性比例方法才会比标准计算更快。对于目前能够用于体相系统的线性比例计算程序,当对含有几百个原子以上的体系进行计算时,线性比例方法才会变得更为有利;但是,这个原子数目是随着所研究体系而变化的。这也就意味着:尽管已经开始出现使用该方法处理具体问题的一些计算实例,但是线性比例方法(仍然)还没有取代标准计算方法。然而不可否认的是,线性比例计算终究会在未来获得更加广泛的应用。

对于涉及到很多原子的计算,另一个可选的方法是使用基于经典力场的方法。这类涉及到几十万甚至上百万个原子的计算现在已经是较为常规的了。当使用 MD 方法进行经典模拟时,所能达到的总时间尺度是有限制的,其原因已经在第 9 章讨论过了,尽管这些时间尺度相比于使用从头算 MD 方法而言可能要大几个数量级,可以使用很多种复杂的蒙特卡罗抽样技术,对许多无法仅用 MD 方法获得的物理特性进行抽样。DFT 和其他种类的量子化学计算可以为经典力场方法提供数据,用以开发该方法所使用的原子间力场,从而对经典力场计算的发展提供重要的支持。

10.6 结论

关于平面波 DFT 计算在描述真实材料时的多方面应用,我们希望本书能为读者提供一个简短有效的入门介绍。对于任何一个活跃的科学领域,将你对任何方法的理解总结为“完整而彻底”的结论都是非常危险的,对于本书所陈述的主题,这一点也不例外。我们奉劝读者在运行 DFT 计算时,广泛地进行阅读并批判性地加以思考,或者是与合作者一起对计算结果加以解释。我们还奉劝读者在学习当前原子尺度材料模拟的核心方法——平面波 DFT 方法的同时,也学习一些其他方法;例如在第 1.6 节简要介绍的量子化学方法,以及在本章前面几节所接触到的经典方法。对于目前的许多科学问题,材料的精细模拟都扮

演着不可或缺的作用,只有当读者对所有这些方法的优势和局限都已经了然于胸时,才能够用严格而挑剔的方式对这些科学问题进行建模评估,并做出积极贡献。

参 考 文 献

- [1] Perdew J P, Ruzsinszky A, Tao J M, et al. Prescription for the Design and Selection of Density Functional Approximations: More Constraint Satisfaction with Fewer Fits, *J. Chem. Phys.* 123, 2005;062201.
- [2] Alfé D, Gillan M J. The Energetics of Oxide Surfaces by Quantum Monte Carlo. *J. Phys. : Condens. Matt.* 18, 2006;L435.
- [3] Litimein F, Bouhafas B, Nouet G, Ruterana P, Meta – GGA Calculation of the Electronic Structure of Group III V Nitrides, *Phys. Stat. Sol. B* 243, 2006;1577.
- [4] von Oertzen G U, Jones R T, Gerson A R, Electronic and Optical Properties of Fe, Zn and Pb Sulfides, *Phys. Chem. Minerals* 32, 2005;255.
- [5] Sousa S F, Fernandes P A, Ramos M J, General Performance of Density Functionals. *J. Phys. Chem. A* 111, 2007;10439.
- [6] Merrick J P, Moran D, Radom L, An Evaluation of Harmonic Vibrational Frequency Scale Factors. *J. Phys. Chem. A* 111, 2007;11683.
- [7] Milman V, Winkler B, White J A, et al. Electronic Structure, Properties, and Phase Stability of Inorganic Crystals: A Pseudopotential Plane – Wave Study. *Int. J. Quantum Chem.* 77, 2000;895.
- [8] Alapati S V, Johnson J K, Sholl D S. Using First Principles Calculations to Identify New Destabilized Metal Hydride Reactions for Reversible Hydrogen Storage. *Phys. Chem. Chem. Phys.* 9, 2007;1438.
- [9] Philipsen P H T, Baerends E J. Relativistic Calculations to Assess the Ability of the Generalized Gradient Approximation to Reproduce Trends in Cohesive Properties of Solids. *Phys. Rev. B* 61, 2000;1773.
- [10] Korhonen T, Puska M J, Nieminen R M. Vacancy – Formation Energies for fcc and bcc Transition Metals. *Phys. Rev. B* 51, 1995;9526.
- [11] Hammer B, Hansen L B, Nørskov J K. Improved Adsorption Energetic within Density – Functional Theory Using Revised Perdew Burke Ernzerhof Functionals. *Phys. Rev. B* 59, 1999;7413.
- [12] Zhang Y K, Yang W T. Comment on Gradient Approximation Made Simple. *Phys. Rev.*

Let. 80, 1998:890.

- [13] Abild – Pedersen F, Andersson M P. CO Adsorption Energies on Metals with Correction for High Coordination Adsorption Sites A Density Functional Study. *Surf. Sci.* 601, 2007:1747.
- [14] Fuchs M, Da Silva J L F, Stampfl C, et al. Cohesive Properties of Group – III Nitrides; A Comparative Study of AllElectron and Pseudopotential Calculations Using the Generalized Gradient Approximation. *Phys. Rev. B* 65, 2002:245212.
- [15] Zhao Y, Truhlar D G. Hybrid Meta Density Functional Theory Methods for Thermochemistry, Thermochemical Kinetics, and Noncovalent Interactions: The MPW1B95 and MPWB1K Models and Comparative Assessments for Hydrogen Bonding and van der Waals Interactions. *J. Phys. Chem. A* 108, 2004:6908.
- [16] Neumann M A, Perrin M – A. Energy Ranking of Molecular Crystals Using Density Functional Theory Calculations and an Empirical van der Waals Correction. *J. Phys. Chem. B* 109, 2005:15531.
- [17] Pacchioni G. Modeling Doped and Defective Oxides in Catalysis with Density Functional Theory Methods; Room for Improvements. *J. Chem. Phys.* 128, 2008:182505.
- [18] Moreira I P R, Illas F, Martin R L. Effect of Fock Exchange on the Electronic Structure and Magnetic Coupling in NiO. *Phys. Rev. B* 65, 2002:155102.

扩展阅读

我们对 DFT 泛函的讨论很大程度受到了文献 Perdew J P, Ruzsinszky A, Tao J M, et al. *J. Chem. Phys.* 123, 2005:062201 的影响。

对于平面波计算中使用混合泛函的近期文献, 一个切入点是: Da Silva J L F, Ganduglia – Pirovano M V, Sauer J, et al. *Phys. Rev. B* 75, 2007:045121, 以及 Paier J, Marsman M, Hummer K, et al. *J. Chem. Phys.* 124, 2006:154709。

有关非极性原子或分子 London 力定律(London's force law)的推导过程, 更多信息可参见: Prausnitz J M, Lichtenthaler R N, De Azevedo E G. *Molecular Thermodynamics of Fluid – Phase Equilibria*. 3rd ed. Prentice Hall, Upper Saddle River, NJ, 1999。

在基于局域化基底的分子体系计算中, 包含了色散相互作用的 DFT – D 可以参见: Schwabe T, Grimme S. *Acc. Chem. Res.* 41, 2008:

569, 以及 Morgado C, Vincent M A, Hillier I H, et al. *Phys. Chem. Chem. Phys.* 9, 2007:448。

DFT + U 计算以及确定参数($U + J$)的几个方法可以参见上述已引用的 Da Silva 等人的研究, 以及 Mosey N J, Carter E A, *Phys. Rev. B* 76, 2007:155123。有关该领域的进展和挑战可参考: Huang P, Carter E A. *Ann. Rev. Phys. Chem.* 59, 2008:261。

对于与线性比例法有关的挑战和方法, 可参考一个全面的综述文献: Goedecker S. *Rev. Mod. Phys.* 71, 1999:1085。如果要学习正处于不断发展中的, 能够适用于体相材料的线性比例计算程序, 一个合适的起始点是: Artacho E, et al. *J. Phys. : Condens. Mat.* 20, 2008:064208, 以及 Skylaris C K, Haynes P D, Mostofi A A, et al. *J. Phys. : Condens. Mat.* 20, 2008:064209。

正如第 9 章所描述的那样, 关于经典模拟的两个代表性文献是: Allen M P, Tildesley D J. *Computer Simulation of Liquids*. Clarendon Press, Oxford, UK, 1987, 以及 Frenkel D, Smit B. *Understanding Molecular Simulations: From Algorithms to Applications*. 2nd ed. Academic, San Diego, 2002。使用 DFT 计算对经典模拟配置力场参数的例子, 请参见: van Duin A C T, Merinov B V, Jang S S, et al. *J. Phys. Chem. A* 112, 2008:3133。