

计算化学概述

知识是实验的女儿。

——莱昂纳多·达·芬奇, 于彭谢里, 约 1492 年

摘要: 通过使用计算化学工具(分子力学、从头算、半经验和密度泛函方法及分子动力学), 您可以计算分子几何构型、速率和平衡、光谱, 以及其他物理性质。计算化学被广泛用于制药行业, 以探索潜在药物与生物分子的相互作用, 如用于将候选药物对接至酶的活性位点。它可用来研究材料科学中固体(如塑料)的特性, 以及实验室和工业中重要反应中的催化作用。计算化学不能取代实验, 因为实验仍然是自然真理的最后仲裁者。

1.1 您可以用计算化学做什么

在本章中, 我们将简要概述计算化学或分子建模的范围和方法。人们可以争论(有些人可能会吹毛求疵)这两个术语之间是否存在差异^[1]。追问这个问题可能没有意义, 我们把这两个术语都看作在计算机上研究化学问题的一套技术。通常用计算来研究的问题有以下几种。

分子几何构型: 分子的形状——键长、键角和二面角。

分子和过渡态的能量: 可显示哪个异构体在平衡状态时是有利的, 以及(来自过渡态和反应物能量)一个反应能有多快。

化学反应性: 了解电子集中在哪里(亲核位点)及它们想去哪里(亲电位点)有助于我们预测各种试剂攻击分子的位置。关于化学反应性的一个特别有用的应用是阐明催化剂可能的作用方式, 这有助于改进催化剂。

红外、紫外和核磁共振光谱: 这些都可以计算出来, 如果分子是未知的, 试图制备它的人知道该寻找什么。

底物与酶的相互作用: 了解分子如何适合酶的活性位点是设计更好的药物的一种方法。

物质的物理特性: 这取决于单个分子的特性及分子在块状材料中如何相互作用。例如, 聚合物(如塑料)的强度和熔点取决于分子的结合程度及分子间作用力的强弱。在材料科学领域工作的人会研究此类问题。

1.2 计算化学的工具

在研究上述问题时,可供计算化学家选择的方法有多种。可用的主要工具分为5大类。

(1) 基于分子模型的分子力学。分子模型是由弹簧(键)连接在一起的球(原子)的集合。如果我们知道正常的弹簧长度和它们之间的夹角,以及拉伸和弯曲弹簧所需要的能量,就能计算出给定的球和弹簧的集合,即给定分子的能量。改变几何构型直到找到最低能量,这使我们能够进行几何构型优化,即计算分子的几何构型。

分子力学的计算速度很快:一台普通的个人计算机可以在几秒钟内优化像类固醇(如胆固醇($C_{27}H_{46}O$))之类的相当大的分子。

(2) 基于薛定谔方程的从头算计算(ab initio,拉丁语“从头开始”,即“第一原理”)。它是现代物理学的基本方程之一,描述了分子中电子的行为和其他一些作用力。从头算方法求解分子的薛定谔方程,可给出分子的能量和波函数。波函数是一个数学函数,可用来计算电子分布(至少在理论上,还可以计算分子的其他任何信息)。根据电子分布,我们可以判断分子的极性,以及亲核试剂或亲电试剂会攻击分子的哪些部位。但对于具有多于一个(!)电子的任何分子,薛定谔方程无法精确求解。因此要使用近似值,近似值越少,从头算方法的水平就越高。无论其计算水平如何,从头算方法都仅仅是基于基本物理理论(量子力学),从这种意义上说是从“第一原理”开始的。

从头算的速度相对较慢:丙烷的几何构型和红外光谱(振动频率)可以在一台个人计算机上在几分钟之内以高水平计算,但一个相当大的分子(如类固醇)的几何构型优化,用合理的高水平可能至少需要几天的时间。当前,具有4 GB或更多随机存取存储器(random access memory, RAM)和1000 GB或更多磁盘空间的个人计算机是一种重要的计算工具,甚至在与高水平从头算相关的高要求任务方面能与UNIX计算机竞争。确实,人们现在很少听说“工作站”,曾经,这些机器的价格大约为15000美元,甚至更多^[2]。对于真正高要求的数字运算,个人还可以通过云计算访问超级计算机,即通过互联网访问远程站点的计算机来达成^[3]。

(3) 半经验计算。像从头算一样,它基于薛定谔方程。但是,在求解薛定谔方程时,它需要更多的近似值,在从头算方法中必须计算的非常复杂的积分在半经验计算中实际上并未计算。相反,半经验程序使用了一种积分库,该积分库通过寻找如几何构型或能量(生成热)等某些计算实体与实验值或当前高水平理论值的最佳拟合来编译。这种将实验值插入数学程序以获得最佳计算值的过程称为参数化。理论与实验的结合使该方法成为“半经验”方法:它基于薛定谔方程,但用实验(或高水平理论)值(经验意味着实验)进行参数化。当然,人们希望半经验计算能为尚未被程序参数化的分子提供良好的答案,而事实往往如此(分子力学也是参数化的)。

半经验计算的速度比分子力学慢,但比从头算快得多。半经验计算所需的时间大约是分子力学的100倍,而从头算所需时间大约是半经验计算的100~1000倍。在一台好电脑上,类固醇的半经验几何构型优化可能需要一分钟。

(4) 密度泛函计算(通常称为密度泛函理论(density functional theory, DFT),泛函是与函数相关的数学实体),像从头算和半经验计算一样,它基于薛定谔方程,但与其他两种方法

不同,密度泛函理论并不计算波函数,而是直接获得电子分布(电子密度函数)。

密度泛函计算的速度通常比从头算更快,但比半经验计算慢。密度泛函理论相对较新:化学上实用的密度泛函理论计算化学可以追溯到 20 世纪 80 年代,而从头算方法的“重要”计算化学则开始于 20 世纪 70 年代,半经验方法开始于 20 世纪 50 年代。

(5) 将运动定律应用于分子的分子动力学计算。由于在力场的影响下分子会改变形状或改变运动,因此,我们可以模拟酶在与底物结合改变形状时的运动,或者模拟一大群水分子围绕蛋白质分子的运动。这种面向生物化学的研究依赖于分子力学计算的力的作用下分子的运动。由于这不是一种电子结构的方法,因此,采用这种力场的分子动力学程序不能研究共价键的断裂和形成(与构象变化对比)。通过半经验、从头算或密度泛函方法产生的力场可以通过分子动力学研究化学反应。请勿将分子动力学(“运动”)与分子力学(分子的“力学”处理)混淆。

1.3 把所有这些放在一起

通常只用分子力学研究非常大的分子,因为其他方法(基于薛定谔方程的量子力学方法:半经验、从头算和密度泛函理论)将花费太长时间。具有不寻常结构的新分子最好通过从头算或密度泛函理论进行研究,因为分子力学或半经验方法中固有的参数化,使得它们对于那些与参数化中所用分子差异很大的分子而言是不可靠的。密度泛函理论比从头算和半经验的方法新,它的局限性和可能性不如其他方法明确。

蛋白质或 DNA 等大分子的结构分析通常是用分子力学来完成的。这些生物大分子的构象运动可以利用分子力学力场进行分子动力学研究。分子运动包括键的断裂和生成,可以利用分子动力学的半经验、从头算或密度泛函方法进行研究。大分子的关键部分,如酶的活性位点,可以用半经验方法或从头算方法来研究。比如,类固醇等中等大小的分子可以通过半经验计算进行研究,如果愿意投入时间,也可以通过从头算方法进行研究。当然,分子力学也可以用于这些研究,但要注意的是,分子力学并不能提供有关电子分布的信息,因此,与亲核或亲电行为有关的化学问题不能单靠分子力学来解决。

分子的能量可以用分子力学、半经验、从头算或密度泛函理论来计算。方法的选择在很大程度上取决于具体的问题。在很大程度上取决于电子分布的反应活性,通常必须用量子力学方法(半经验、从头算或密度泛函理论)来研究。光谱可以通过高水平从头算或密度泛函理论进行最可靠的计算,但通过半经验方法可以获得有用的结果,而一些分子力学程序可以计算出相当好的红外光谱(连接弹簧的球会振动!)

将一个分子对接至酶的活性位点以了解它是如何匹配的,这是计算化学的一个极其重要的应用。可以使用鼠标或某种操纵杆来操纵底物,然后尝试将其安装(对接)到活性部位,但自动对接现已标准化。由于涉及大分子,这项工作通常用分子力学来完成,尽管可以通过一种量子力学方法来研究大的生物分子的选定部分。这种对接实验的结果可指导人们设计更好的药物,例如,与指定酶更好地相互作用而不被其他酶识别的分子。

在材料科学中,计算化学在研究材料的性质方面很有价值。半导体、超导体、塑料、陶瓷都采用计算化学进行了研究。最近的一个独创性的发展是一种开发具有可计算特性的材料

的程序,如果它能兑现承诺,将会非常有潜力^[4]。这类研究往往涉及固体物理学的知识,而且有些专业化。在不那么功利的情况下,也有人在这门科学的帮助下研究了具有艺术价值的文物^[5]。

计算化学相当便宜,与实验相比速度很快,而且对环境安全(尽管近十年来,计算机的大量使用引起了人们对能源消耗^[6]和废弃机器处置^[7]的关注)。但它不能取代实验,实验仍然是自然真理的最后仲裁者。此外,要制备某种东西(新药、新材料),就必须进入实验室。而且,需要注意的是尽管计算能力很强^[8],但人们应小心,不要超越其有效范围:在极端情况下,用泡利的话来说,您可能“连错误都算不上”^[9]。计算在某些方面已经变得如此可靠,以至于越来越多的科学家在开始实验项目之前就会使用它,可能有一天,您只有说明您在多大程度上通过计算探索了该提案的可行性,才能获得某种实验工作的资助。

1.4 计算化学的哲学

计算化学是一种观点的顶点(迄今为止)。这种观点认为化学最好被理解为原子和分子行为的表现,且这些原子和分子都是真实的实体,而不仅仅是便捷的智力模型^[10]。它是对迄今为止在有机化学结构式中得到了最大胆表达的 trends 的物理和数学肯定^[11],并且明确否定了直到最近才流行的说法^[12],即科学是一种“范式”的游戏^[13]。

在计算化学中,我们认为我们正在模拟真实物理实体的行为,尽管是借助于智力模型。随着模型的改进,它们可以更准确地反映现实世界中原子和分子的行为。

1.5 总结

计算化学使人们能够计算分子的几何构型、反应性、光谱和其他性质。它采用了以下几类工具。

分子力学:基于分子的球和弹簧模型。

从头算方法:基于薛定谔方程的近似解,无须拟合实验。

半经验方法:基于薛定谔方程的近似解,须对实验进行拟合(即使用参数化)。

密度泛函理论(DFT)方法:基于薛定谔方程的近似解,绕开了作为从头算和半经验方法的主要特征的波函数。

分子动力学方法:研究运动中的分子。

利用从头算和更快的密度泛函理论可以研究具有理论价值的新型分子,前提是分子的相对分子质量不太大。半经验方法的速度比从头算或密度泛函理论快得多,可以很容易地将其应用于相当大的分子(如胆固醇($C_{27}H_{46}O$)和更大的分子),而分子力学能用于计算非常大的分子(如蛋白质和核酸)的几何构型和能量,尽管分子力学并未提供有关电子性质的信息。计算化学被广泛用于制药行业,以探索潜在药物与生物分子的相互作用,例如,将候选药物与酶的活性位点对接。在材料科学中,它也被用于研究固体(如塑料)的性质。

较容易的问题

1. 计算化学是什么意思?
2. 计算化学可以回答哪些问题?
3. 列出计算化学家可用的主要工具。概述每个工具的特征(每个工具几句话)。
4. 一般来说,哪种是最快的计算化学方法(工具),哪种是最慢的?
5. 为什么计算化学在工业中 useful?
6. 从化学家的角度来看,薛定谔方程基本上描述了什么?
7. 我们可以得到薛定谔方程精确解的分子种类有什么限制?
8. 什么是参数化?
9. 计算化学相对于“湿化学”有什么优势?
10. 为什么计算化学不能取代“湿化学”?

较难的问题

讨论以下内容,并论证您的结论。

1. 在电子计算机问世之前,有没有计算化学?
2. “常规”物理化学,如动力学、热力学、光谱学和电化学的研究,可以被视为一种计算化学吗?
3. 最经常被计算的分子性质是几何构型、能量(与其他异构体相比)和光谱。为什么计算诸如熔点和密度之类的“简单”性质更具挑战性?(提示:分子 X 和物质 X 之间是否存在差异?)
4. 一个分子的几何构型和能量(与其他异构体相比)通常可以通过球和弹簧模型(分子力学)被准确地计算出来,这是否令人惊讶?
5. 您认为分子力学无法计算哪些性质?
6. 第一原理(从头算)的计算是否应该比那些利用实验数据(半经验的)的计算更可取?
7. 实验和计算都可能给出错误的答案。那么,为什么实验有最终决定权?
8. 考虑将潜在的药物分子 X 与酶的活性位点对接:影响 X“对接”程度的一个因素显然是 X 的形状。您能想到另一个因素吗?(提示:分子由原子核和电子组成。)
9. 近年来,组合化学技术已被用于快速合成各种相关化合物,然后对其进行药理活性测试。与借助计算化学研究分子如何与酶相互作用的“合理设计”方法相比,这种寻找候选药物方法的优缺点是什么?
10. 想出一些可以通过计算来研究的不寻常的分子,是什么让您的分子与众不同?

参考文献

1. For example, summary of a discussion on the Computational Chemistry List (CCL), at www.chem.yorku.ca/profs/renef/whatiscc.html. Accessed 22 Sept 2014
2. Schaefer HF III (2001) The cost-effectiveness of PCs. *Theochem* 573:129
3. (a) Fox A (2011) Cloud computing-what's in it for me as a scientist? *Science* 331:406; (b) Mullin R (2009) *Chem Eng News*. May 25, 10
4. (a) Cerquera TFT et al (2015) *J Chem Theory Comput* 11:3955; (b) Jacoby M (2015) *Chem Eng News*, December 30, 8
5. Fantacci S, Amat A (2010) Computational chemistry, art, and our cultural heritage. *Acc Chem Res* 43:802
6. (a) McKenna P (2006) The waste at the heart of the web. *New Sci* 192(2582):24; (b) Keipert K, Mitra G, Sunriyal V, Leang SS, Sosokina M (2015) Energy-Efficient Computational Chemistry: Comparison of run times and energy consumption for two kinds of computer architecture (ARM-, i.e. RISC-based and x86) and three families of calculations. *J Chem Theory Comput* 11:5055
7. *Environmental Industry News* (2008) Old computer equipment can now be disposed in a way that is safe to both human health and the environment thanks to a new initiative launched today at a United Nations meeting on hazardous waste that wrapped up in Bali, Indonesia, 4 Nov 2008
8. E.g. Cheng G-J, Zhang X, Chung LW, Xu L, Wu Y-D (2015) *J Am Chem Soc* 137:1706
9. Peierls R (1960) Pauli's words: the physicist Rudolf Peierls reported that Pauli used these (the German equivalents) in reference to the work of a third party. *Biograph Mem Fellows R Soc* 5:186; Plata RE, Singleton DA (2015) "Wolfgang Pauli, 1900–1958." The critical paper which invokes them. *JACS* 137:3811
10. The physical chemist Wilhelm Ostwald (Nobel Prize 1909) was a disciple of the philosopher Ernst Mach. Like Mach, Ostwald attacked the notion of the reality of atoms and molecules ("Nobel laureates in chemistry, 1901–1992", James LK (ed) American Chemical Society and the Chemical Heritage Foundation, Washington, DC, 1993) and it was only the work of Jean Perrin, published in 1913, that finally convinced him, perhaps the last eminent holdout against the atomic theory, that these entities really existed (Perrin showed that the number of tiny particles suspended in water dropped off with height exactly as predicted in 1905 by Einstein, who had derived an equation assuming the existence of atoms). Ostwald's philosophical outlook stands in contrast to that of another outstanding physical chemist, Johannes van der Waals, who staunchly defended the atomic/molecular theory and was outraged by the Machian positivism of people like Ostwald. See Ya Kipnis A, Yavelov BF, Powlinson JS (1996) *Van der Waals and molecular science*. Oxford University Press, New York. For the opposition to and acceptance of atoms in physics see: Lindley D (2001) *Boltzmann's atom. The great debate that launched a revolution in physics*. Free Press, New York; and Cercignani C (1998) *Ludwig Boltzmann: the man who trusted atoms*. Oxford University Press, New York, 1998. Of course, to anyone who knew anything about organic chemistry, the existence of atoms was in little doubt by 1910, since that science had by that time achieved significant success in the field of synthesis, and a rational synthesis is predicated on assembling atoms in a definite way
11. For accounts of the history of the development of structural formulas see Nye MJ (1993) *From chemical philosophy to theoretical chemistry*. University of California Press; Russell CA (1996) *Edward Frankland: chemistry, controversy and conspiracy in Victorian England*. Cambridge University Press, Cambridge

12. (a) An assertion of the some adherents of the “postmodernist” school of social studies; see Gross P, Levitt N (1994) *The academic left and its quarrels with science*. John Hopkins University Press, Baltimore; (b) For an account of the exposure of the intellectual vacuity of some members of this school by physicist Alan Sokal’s hoax see Gardner M (1996) *Skeptical Inquirer* 1996, 20(6):14
13. (a) A trendy word popularized by the late Thomas Kuhn in his book– Kuhn TS (1970) *The structure of scientific revolutions*. University of Chicago Press, Chicago. For a trenchant comment on Kuhn, see ref. [12b]; (b) For a kinder perspective on Kuhn, see Weinberg S (2001) *Facing up*. Harvard University Press, Cambridge, MA, chapter 17

势能面的概念

一切都应该尽可能简单,而不是比较简单。

——阿尔伯特·爱因斯坦,但这些精确的词,
或者说德语中的对应词,并没有出现在他的作品集中(在线可得)

摘要: 势能面(potential energy surface, PES)是计算化学中的一个核心概念。势能面是一个分子(或分子集合)的能量与其几何构型之间的数学或图形关系。波恩-奥本海默近似认为,与电子相比,分子中的原子核基本是静止的。这是计算化学的基石之一,因为它使分子形状(几何构型)的概念变得有意义,使势能面的概念成为可能,并通过允许我们专注于电子能,并在以后加入核排斥能而简化了薛定谔方程对分子的应用。势能面在适用的分子计算中非常重要,这将在第5章中进行详细阐述。本章阐述了几何构型优化及过渡态的性质。

2.1 观点

我们开始用势能面更详细地研究计算化学,因为这是该主题的核心。许多在数学上可能具有挑战性的重要概念都可以通过势能面提供的洞察力直观地掌握^[1]。

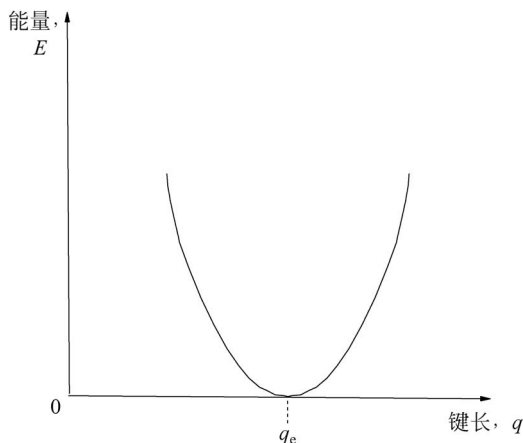


图 2.1 双原子分子的势能面。当键长 q 被拉伸或压缩远离平衡位置 q_e 时, 势能增加。这里选择 q_e (键长的零变形) 处的势能作为能量零点

以双原子分子 AB 为例。在某些方面,分子的行为类似于由弹簧(化学键)连接在一起的球(原子)。事实上,这个简单图形是第 3 章讨论的分子力学中重要方法的基础。当我们拿起正常几何构型(平衡几何构型)的双原子分子的球-弹簧宏观模型,抓住“原子”并通过拉伸或压缩“键”使模型变形时,就会增加分子模型的势能(图 2.1)。因为我们用力将弹簧移动了一段距离以使其变形——即对弹簧做功,所以,根据定义,被拉伸或压缩的弹簧具有能量。当我们将模型保持在新的几何构型时,它是静止的,因此,该能量不是动能,而默认是势能(取决于位置)。势能与键长的关系图是势能面的一个示例。线是一维的“表面”。我们将很快看到更熟悉的二维表面的示例,而不是图 2.1 中的曲线。

真实分子的行为与上述宏观模型相似,但在两个方面不同。

(1) 它们在平衡键长附近不停地振动(正如我们由海森堡不确定性原理预计的那样,一个静止分子将具有精确定义的动量和位置),因此它们始终具有动能(T)和/或势能(V):当键长达到平衡长度时, $V=0$,而在振幅极限时, $T=0$;在其他位置, T 和 V 都不为零。事实上,分子从来都不会处于动能为零的静止态(它始终具有零点能(ZPE)或零点振动能(ZPVE),2.5 节),通常在势能-键长图上绘制一系列高于曲线底部的直线(图 2.2),以表示分子可能具有的振动能量(它可以占据的振动能级)。分子从不位于曲线的底部,而是占据其中一个振动能级;在一个分子集合中,能级根据其间距和温度分布^[2]。我们通常会忽略振动能级,认为分子停留在实际势能曲线或表面上。

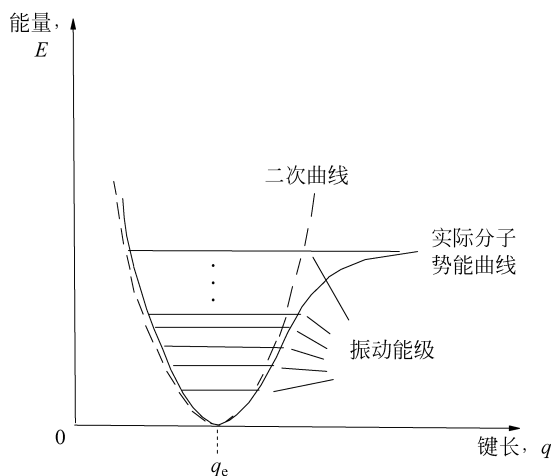


图 2.2 实际分子不会位于势能曲线的底部,而是占据振动能级。而且,只有在平衡键长 q_e 附近,二次曲线才接近实际势能曲线

(2) 在平衡键长 q_e 附近,宏观球-弹簧模型或真实分子的势能-键长曲线可用一个二次方程很好地描述,即简单谐振子($E = (1/2)k(q - q_e)^2$, 其中 k 是刚度系数)。然而,当远离 q_e 时,势能就偏离二次(q^2)曲线(图 2.2)。也就是说,由这种非简谐性所代表的与分子实际的偏差,在远离平衡几何构型的地方变得越来越重要。

图 2.1 表示 E 与 q 的二维图中的一维势能面。双原子分子 AB 仅有一个几何构型参数可以改变,即键长 q_{AB} 。假设有一个具有多个几何构型参数的分子,如水:几何构型由两个键长和一个键角确定。如果允许两个键长相等,即如果将分子限制为 C_{2v} 对称(两个对称面

和一个二重对称轴,见 2.6 节),则该三原子分子的势能面为 E 与两个几何构型参数的关系图, $q_1 = \text{O—H}$ 键长, $q_2 = \text{H—O—H}$ 键角(图 2.3)。图 2.3 表示三维图中的二维势能面(法向表面是二维对象)。我们可以绘制一个实际的三维模型,该模型是 E 与 q_1 和 q_2 的三维图形。

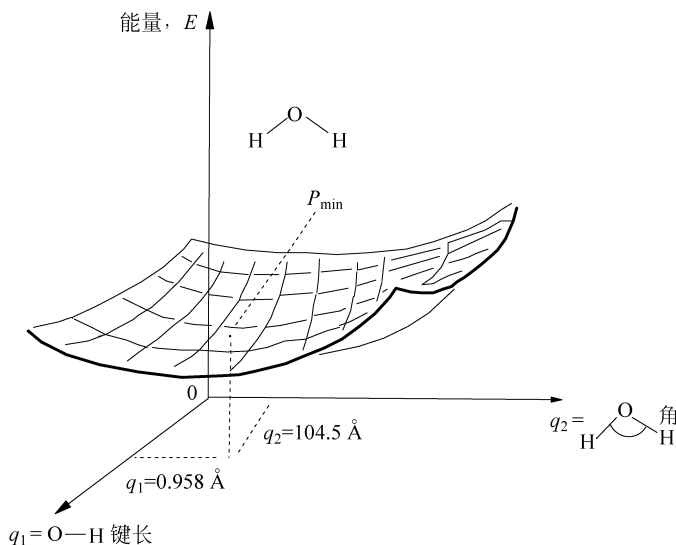


图 2.3 H_2O 的势能面。点 $P_{\min}$ 对应于三原子分子的最小能量几何构型,即水分子的平衡几何构型

我们可以抛开水分子,考虑一个对称性较低的三原子分子,如 HOF(次氟酸)。它有 3 个几何构型参数,即 H—O 和 O—F 键长及 H—O—F 键角。要为 HOF 构造一个类似于 H_2O 的笛卡尔势能面图,我们需要绘制 E 与 $q_1 = \text{H—O}$ 、 $q_2 = \text{O—F}$ 和 $q_3 = \text{H—O—F}$ 角的关系图。我们将需要 4 个相互垂直的轴(对应 E 、 q_1 、 q_2 、 q_3 ,图 2.4)。由于此四维图无法在三维空间中构建,因此我们不能准确地绘制它。HOF 势能面是四维空间中一个二维以上的三维“表面”:它是一个超曲面,势能面有时也称为势能超曲面。尽管绘制超曲面存在问题,但我们仍可以将方程 $E = f(q_1, q_2, q_3)$ 定义为 HOF 的势能面,其中, f 是描述 E 如何随 q 变化的函数,并用数学方法处理超曲面。例如,在图 2.1 的 AB 双原子分子势能面(一条线)中,最小势能的几何构型是 $dE/dq = 0$ 的点。在 H_2O 势能面(图 2.3)中,最小势能的几何构型由点 $P_{\min}$ 定义,对应于 q_1 和 q_2 的平衡值。在这一点上, $dE/dq_1 = dE/dq_2 = 0$ 。尽管无法如实地通过图形绘制超曲面,但对于计算化学家来说,建立对它们的直观理解非常有用。这可以通过示意图图 2.1 和图 2.3 来获得。图 2.1 和图 2.3 中的线或二维平面实际上用了多维图中的一个切片。可以通过类比来理解:图 2.5 展示了如何用水的三维图制作二维切片。切片可以保持两个几何构型参数中的任意一个不变,也可以同时包含这两个参数,从而给出一个图,其中的几何轴是一个以上几何参数的组合。类似地,我们可以获取 HOF 的超曲面(图 2.6),甚至更复杂分子的一个三维切片,并用 E 与 q_1 、 q_2 的关系图来表示势能面。我们甚至可以使用一个简单的二维图,其中, q 表示一个、两个或所有几何构型参数。我们将看到,这些二维图,特别是三维图,保留了数学上严谨但无法可视化的 $E = f(q_1, q_2, \dots, q_n)$ 的 n 维超曲面的定性,甚至是定量特征。