

谈谈赝势基组的选用

文/Sobereva @[北京科音](#)

First release: 2017-Apr-18 Last update: 2022-Nov-12

1 前言

“算 xxx 元素该用什么赝势基组”是网上最常问的问题之一，本文说说该怎么选用赝势基组。本文主要面向 Gaussian 用户的情况，其它程序用户可以择情借鉴。看本文之前一定要非常仔细看《谈谈量子化学中基组的选择》(<http://sobereva.com/336>)，本文假定读者已经充分理解了那篇文章所讲的全部内容。如果你特别着急做计算，连把本文完整读完的耐心都没有，那就直接看第 4 节。

笔者之前还写过其它与赝势、赝势基组有关的文章。凡是以下文章里已讲过的内容在本文基本不会再重复，如果对赝势一点都不了解的话一定要先看看。

赝势的函数形式以及在量子化学程序中定义的方式 (<http://sobereva.com/188>)

在赝势下做波函数分析的一些说明 (<http://sobereva.com/156>)

详解 Gaussian 中混合基组、自定义基组和赝势基组的输入 (<http://sobereva.com/60>)

以下两篇文章与赝势基组也有关，若感兴趣可以看看

在线基组和赝势数据库一览 (<http://sobereva.com/309>)

用 NWChem 做 SODFT 在 DFT 计算中考虑旋轨耦合效应 (<http://sobereva.com/368>)

2 什么时候该用赝势基组？

先看看使用赝势基组计算带来的好处：

(1) 计算时不需要直接考虑内层电子，使计算量大幅降低，而且序号越靠后的元素这一点体现得越明显。用赝势基组计算一个很重原子，其省下来的计算量往往相当于全电子基组计算几个轻原子。

(2) 相对论效应从第四周期 (H, He 是第一周期，下同) 开始值得关注，但不考虑也问题不大，从第五周期开始则是绝对不能忽略。相对论效应主要来自于内层电子。拟合赝势的时候如果已考虑相对论效应，即使用相对论赝势，则基于赝势计算的时候就可以把相对论效应等效地体现出来，免得用 DKH 等方法考虑相对论效应了（用 DKH 等不仅导致额外增加耗时，多数程序还不支持其解析导数，导致做优化等任务极为困难）。

用赝势基组计算也有一些坏处：

(1) 由于内层电子没有明确表达，有可能影响计算精度。这一点看你具体用什么赝势和赝势基组以及计算哪些元素了，不能一概而论。对同一系列赝势，大核版本由于这个问题更明显，故精度大多不如小核版本好。

(2) 导致直接依赖于内核电子的性质无法计算。对于使用了赝势的原子，就没法计算其 NMR、超精细耦合、内核电子激发等依赖于内层电子的量了（不过对其它原子还是可以计算的）。

(3) 内层电子的缺失可能有碍一些波函数分析，但也有办法避免，详见《在赝势下做波函数分析的一些说明》中的讨论。

了解了以上信息，什么时候该用赝势基组心里就应该有数了。

前三周期元素：强烈不建议用赝势基组，因为此时使用赝势带来的上述两条优点很不明显，而带来的那些坏处却是实实在在的。而且赝势对前三周期可用的赝势基组比较有限，可选余地很小。

第四周期元素：用不用赝势基组皆可。一般我还是推荐用赝势基组，因为此时它带来的两条优点已经开始显现了，赝势基组可选择余地此时也多了。如果体系中只有很少量第四周期元素，比如配合物当中的过渡金属，并且想在写输入文件时省事或者希望所有元素用的基组都统一，那么都用全电子基组也完全可以。如果体系中有大量第四周期元素，比如计算较大的过渡金属团簇，那么强烈建议用赝势基组，耗时会节约甚巨。

第五周期及之后的元素：除非有特殊原因必须用全电子基组，否则强烈建议用赝势基组，它带来的那两条好处已经重要到必不可少。

3 常用的几种赝势基组

一般计算用得到的赝势基组就下面四类。除下面介绍的，在文献里还容易看到 CRENB 和 SBKJC 赝势，但如今

用得都不多，CRENB 虽然覆盖周期表很全面但是精度不算多好，SBKJC 主要也就是 GAMESS-US 用户才用，因为是其唯一自带的赝势，并没什么特别的优点。所以这俩赝势后文不提。还有个 LFK 赝势基组是在 SBKJC 对前六个周期主族（除 Ga、In、Tl）的定义基础上增加弥散和极化函数使计算极化率达到较好精度，如果需要赝势下计算极化率可以考虑。

3.1 Lanl 系列赝势+标配赝势基组

Lanl (Los Alamos National Laboratory) 是 Hay 和 Wadt 在上世纪 80 年代提出的很常用的一种赝势。对第四周期元素并未考虑相对论效应，对更重的元素都考虑了相对论效应。

对 s 族和过渡金属，Lanl1 是大核赝势，Lanl2 是小核赝势。对 p 族二者等价，是大核赝势。都有 MB（极小基）和 DZ（双分裂价基）两种配套的赝势基组，最常用的是 Lanl2DZ。Lanl2DZdp 是 Lanl2DZ 基础上加了层 p 弥散和 d 极化。lanl1 以及 MB 都精度太烂，绝对不要用。

2008 年提出了 Lanl2TZ，是对 Lanl2DZ 重新进行收缩得到，算是准 3-zeta 基组。Lanl2TZ+ 是对第一周期金属增加一层 d 弥散。Lanl2TZ(f) 是对过渡金属增加一层 f 极化。

Lanl08 是对于过渡金属和主族元素在 Lanl2DZ 基础上都彻底去收缩化。Lanl08+ 是在 Lanl08 基础上对于第一周期过渡金属加上 d 弥散。Lanl08(d) 和 Lanl08(f) 分别是在 Lanl08 基础上对主族和过渡金属加上 d 和 f 极化。

Lanl2DZ 目前包含前六个周期绝大部分元素，镧系只支持 La 自己，锕系只支持 U、Np、Pu。其它变体支持的元素明显没有 Lanl2DZ 广泛，具体支持哪些元素可以在 BSE 网站（<https://www.basissetexchange.org>）上查。

Lanl2DZ 曾被用得非常广泛，即便是现在还用得很多。但**笔者强烈不推荐用 Lanl2DZ**，因为在如今看来其精度明显偏低。如今算过渡金属起码 lanl2TZ，其实比 Lanl2DZ 贵不了多少而精度大有改善，如果计算资源还富裕则建议升到 Lanl2TZ(f)。对于主族，lanl2DZ 也同样不理想，又没极化又是大核赝势，如今起码用 lanl08，条件稍好点就建议用 lanl08(d)。

3.2 Stuttgart 赝势+标配赝势基组

对于过渡金属和主族的大部分元素都是大核和小核版本都有，是计算镧系和锕系的不二之选，对于其它元素也都比较不错。此系列赝势几乎涵盖整个周期表，甚至 120 号元素都有。精度较好，可靠度高，不断在发展和完善。Gaussian 等程序自带的以及 BSE 上的 Stuttgart 赝势及其标配的赝势基组既不新也不全，如果比较讲究的话，建议去其作者官网上下载定义：<http://www.tc.uni-koeln.de/PP/clickpse.en.html>。

Stuttgart 官网上的每种元素往往有许多不同赝势版本，其命名在前面提过的《详解 Gaussian 中混合基组、自定义基组和赝势基组的输入》里有一些介绍，WB 和 DF 结尾的都是考虑相对论效应的。每种赝势有时候还提供了不同的标配的赝势基组，情况相当复杂，不同元素情况还不一样，三两句话说不清楚。这里只说 Gaussian 的情况。当你在 Gaussian 里直接用 SDD 关键词使用自带的 Stuttgart 赝势+配套基组时，其特征以及与 Lanl 系列赝势基组的大小对应关系如下

Sc~Ni: MDF 小核赝势，有 f 极化函数。尺寸略大于 Lanl2TZ(f)

Cu、Zn 及第二、三周期过渡金属元素: MWB 小核赝势，无 f 极化函数。尺寸略大于 Lanl2TZ

主族: 对 s 和 p 族分别是小核和大核 MWB 赝势，无 d 极化函数，尺寸介于 Lanl2DZ 与 Lanl08 之间

Stuttgart 作者主页上的上面提到的这些过渡金属的赝势的配套基组都有 2f,1g 极化函数（因此昂贵很多），其它元素的赝势基组定义和 Gaussian 直接用 SDD 关键词时多数情况一致。

lanl2DZ 的精度明显逊于 Gaussian 里直接写 SDD 时的情况。对过渡金属 Lanl2TZ(f) 和 SDD 精度相仿佛，对主族 lanl08 和 SDD 精度相仿佛。

如果没特殊情况，用 Stuttgart 赝势基组时就在 Gaussian 里写 SDD 就行了，有必要特殊考虑的主要也就两种情况：

(1) 计算稀土元素的团簇，用默认的 Stuttgart 小核赝势耗时太高，SCF 收敛比用大核赝势也往往困难得多得多，这时候需要换成大核版本。对于 d 族过渡金属只有小核版本，所以没得选，ds 族倒是也有大核的，不过意义不大

(2) 计算主族时，为了结果更好，或者能更好地做波函数分析，可改用小核版本（没内置，需要自行到 Stuttgart 网站上下）。

3.3 cc-pVnZ-PP 系列赝势基组

上述 lanl 赝势基组和 Stuttgart 赝势基组（指 Gaussian 自带的版本）对于 DFT 计算是很合适的，但对于后 HF 计算，尤其是 CCSD(T) 这样高档次后 HF 方法就显得比较 low 了。

cc-pVnZ-PP (n=D,T,Q,5) 系列是 Dunning 相关一致性基组的赝势版本，是最适合后 HF 计算的赝势基组，很适合搭配 cc-pVnZ 给轻原子使用。结合的是 Stuttgart 小核 MDF 赝势。此系列在不断完善中，目前对从 Cu 到 Rn 的 d,ds,p 族元素、从 K 开始的碱金属、从 Ca 开始的碱土金属、Th/Pa/U 有定义。其中有些在 BSE 上查不到，在作者的主页上可以下载最新最全版本：<http://tyr0.chem.wsu.edu/~kipeters/basis.html>。由于此系列赝势基组高角动量函数多，收缩度高，用在 DFT 计算上较昂贵，不值得。

cc-pVnZ-PP 标准的带弥散函数的版本是 aug-cc-pVnZ-PP，读过《谈谈弥散函数和“月份”基组》(<http://sobereva.com/119>) 一文的人肯定知道，aug 包含的大量高角动量弥散函数意义不大。对于 aug-cc-pVnZ-PP 系列，也完全可以简化成月份基组的版本来节约时间。

另外还有 SDB-cc-pVnZ-PP (n=T,Q) 系列赝势基组，结合 Stuttgart 大核 MWB 赝势，力图达到全电子 cc-pVnZ (n=T,Q) 的质量。仅对于 Ga~Kr 和 In~Xe 有定义。精度略逊于 cc-pVnZ-PP，但由于是大核所以更便宜。

3.4 def2 系列赝势基组

def2 系列基组从 def2-SV(P)、def2-SVP、def2-TZVP、def2-TZVPP、def2-QZVP、def2-QZVPP 精度逐级增高。对前四周期是全电子基组，从第五周期开始是赝势基组，结合的是 Stuttgart 小核赝势。def2 系列一大优点是除了镧系以及最后一个周期外覆盖了周期表所有元素，因此对绝大多数体系可以只用 def2 而无需混合基组。此系列基组既适合 DFT 也适合后 HF 计算，但对于后者起码得用 def2-TZVP。Gaussian 直接内置了 def2 系列，不过其镧系的定义没有内置（至少对于 G16 B.01 而言仍是如此），镧系的定义得从 BSE 或 Turbomole 基组库 (<http://www.cosmologic-services.de/basis-sets/basissets.php>) 里获得。

def2 系列赝势基组虽没有标配的弥散函数，但需要弥散函数时可以按此文方式很容易地加上：《给 ahlich 的 def2 系列基组加弥散的方法》(<http://sobereva.com/340>)。

搞 def2 的 Weigend 等人还搞了 dhf 系列赝势基组，只包含第五、六周期，主要是把 def2 对 s、d 族元素用的 Stuttgart 赝势从 MWB 改为了 MDF 并重新优化了赝势基组，使得精度有一定提升，并且 dhf 原文中的数据表明 dhf 系列比同级别 cc-pVnZ-PP 结果更好。不过 dhf 没有被 Gaussian 等主程序内置，得去 Turbomole 基组库里拷定义。

def2 之前还有 def 系列基组，同样从第五周期开始是赝势基组。除 def2 目前支持的元素外还多定义了镧系，但是对镧系镧系只有 SVP 和 TZVP 两个级别，而且大小基本一样。

cc-pVnZ-PP 系列、def2/dhf 系列赝势基组都是基于 Stuttgart 赝势，比起 Stuttgart 赝势标配的赝势基组的关键好处就是基组选择余地大，可以根据自己要求的精度来选。

4 赝势基组的选择建议

本文一开始提到的，也是初学者老问的“算 xxx 元素该用什么赝势基组”这个问题，根本就是个 unanswerable 的问题，提供的前提太少！要算的元素只是要被考虑的诸多因素之一而已。假设只是计算常规问题，比如几何优化、振动分析、计算反应能量/势垒、价层电子激发之类，那么根据你的计算条件的好坏（显然也要考虑体系大小、有多少原子要用赝势基组），以及所用的理论方法，赝势基组有以下推荐。当然，推荐的赝势基组能用的前提是它对你要算的元素有定义，哪些赝势基组能用于哪些元素前面介绍过了。

DFT 计算：

计算条件差：Lanl2DZ

计算条件一般：过渡金属用 lanl2TZ，主族用 lanl08(d)

计算条件不错：过渡金属用 lanl2TZ(f) 或 SDD，主族用 lanl08(d)

计算条件很好：def2-TZVP

计算条件无限好：def2-QZVP（DFT 计算用 QZ 级别其实很浪费）

后 HF 或多参考方法计算（对于 cc-pVnZ-PP 不支持的元素只能用 def2 了）：

计算条件奇差无比：cc-pVDZ-PP 或 def2-SVP

计算条件一般：过渡金属用 Lanl2TZ(f) 或 SDD，主族用 lanl08(d)

计算条件不错：cc-pVTZ-PP 或 def2-TZVPP

计算条件很好：cc-pVQZ-PP 或 def2-QZVPP

def2 对第四周期还是全电子基组，cc-pVnZ-PP 从 Cu 才开始有定义。因此后 HF 计算时非要对 Cu 之前的用赝势基组的话那就只能靠 lanl 或 SDD 了。

提示 1: 对前三周期不要用赝势基组。并且能接受的基组底限是 6-31G*。

提示 2: 做波函数分析时一定要用小核赝势。

提示 3: 铜系可以用 Stuttgart 赝势+标配基组也可以用 def2 系列赝势基组。铜系一般用 Stuttgart 赝势+标配基组或 def-TZVP 赝势基组, 对 Th、Pa、U 也可以用 cc-pVnZ-PP 系列。

如果研究的问题需要基组含有弥散函数才能描述好 (见《谈谈弥散函数和“月份”基组》), 赝势基组显然必须带弥散函数。cc-pVnZ-PP 前面带上 aug 即可, def2 怎么加弥散前面也说了, lanl 系列赝势基组直接也有带弥散的版本。而对于 SDD 标配的赝势基组, 可自行用 even-tempered 方式加弥散 (在《给 ahlich 的 def2 系列基组加弥散的方法》里面说过怎么做) 或者把 lanl 的带弥散版本的弥散函数借过来也可以。

5 关于计算配合物时赝势基组与全电子基组的匹配问题

老有人问计算过渡金属配合物时配体原子用 xxx 全电子基组, 过渡金属原子用 yyy 赝势基组合不合适, 匹配不匹配什么的。我感觉很多人对匹配问题太敏感了, 纯粹想多了。在我看来, 如果按照基组的价层分裂数 (zeta) 来划分, 给过渡金属用的赝势基组的档次约等于配体的全电子基组, 或者前者比后者高一个档, 都是完全可以的。比如说, DFT 计算二茂铁, 可以 Fe 用 lanl2TZ(f) 或 SDD, 茂环原子用 6-31G* 或者 6-311G*。我主张过渡金属用的赝势基组的档次比配体更高一点, 毕竟过渡金属在配合物里面占的比例通常极少, 基组大点不至于耗时增加多少, 而它又处在体系的中心地位, 重要性最高, 所以给过渡金属用好点的赝势基组, 是明显的好钢用在刀刃上的行为。因此, 比如 Fe 用较烂的 lanl2DZ, 而茂环用较好的 TZVP 就严重不合理了, 明显主次不分。不过, 我也不建议过渡金属用的赝势基组高过轻原子太多, 比如算个顺铂, Pt 用 def2-TZVPP, 配体用 6-31G*, 这差距就过于悬殊, 配体的烂基组就成为了制约计算精度的严重瓶颈 (而且这种基组尺寸的严重不平衡, 对于 Mulliken 等直接基于基函数做的波函数分析也有不良影响)。