

# 谈谈量子化学中基组的选择

文/Sobereva @[北京科音](#)

First release: 2016-Jun-16    Last update: 2023-May-3

## 0 前言

量子化学中基组的选择是一门艺术，富有研究经验和量化理论直觉的研究者能够对所研究的问题、体系、要求的精度选择出性价比最高的基组，而初学者则几乎总是基组选得不合适，轻则糟蹋计算资源，重则结果不准确乃至根本没法用。本文目的是提供基组选用的一般性指导意见，但本文并不会把基组系统性地介绍一遍。想对基组有全面系统的了解，非常推荐参加笔者讲授的初级、基础量子化学培训班（见北京科音官网 [www.keinsci.com](http://www.keinsci.com) 的“科研培训”栏目）。自学的话看《基组入门资料小合集》（<http://bbs.keinsci.com/thread-2190-1-1.html>）。本文不怎么涉及赝势基组的选择问题，因为赝势基组的选择笔者另有其它文章：《谈谈赝势基组的选用》（<http://sobereva.com/373>）。

该用什么基组？琢磨这个问题前必须先明确以下要素，综合判断。问高手用什么基组的时候提供的以下信息越充分，高手也能给你越确切的答案（经常有人问我算个某某分子，用什么基组，这根本没法准确回答）

- a) 算什么体系，多少原子、都有什么元素
- b) 算什么问题
- c) 要求多高精度、用什么理论方法
- d) 能利用的计算资源如何（尤其是核数）、能接受最多多长的计算耗时
- e) 用什么程序

本文第 1 节先说通常怎么选择基组，第 2 节再对某些类型的问题专门给出选择建议。本文给出的基组选择建议是针对最常用的 Gaussian 而言的。本文提到的很多基组在 Gaussian 中没有自带，需要自行到 BSE（<https://www.basissetexchange.org>）基组数据库等地方去拷定义。

以下文章如果没看过建议看看，本文会涉及其中的内容：

谈谈弥散函数和“月份”基组 <http://sobereva.com/119>

给 ahlricks 的 def2 系列基组加弥散的方法 <http://sobereva.com/340>

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

谈谈能量的基组外推 <http://sobereva.com/172>

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>

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

浅谈为什么优化和振动分析不需要用大基组 <http://sobereva.com/387>

有些读者读完本文后，跟我说，我看某某文献，用的基组怎么和你推荐的不一样啊？怎么你说不能 xxx，但我看到文献里就是 xxx 了。我这里非常负责任地说：80% 以上文献里用的基组要么不合理，导致结果烂；要么用的基组不是最划算的，造成了计算量的无谓的增加。别以为发出去的文章都是内行高手、对基组非常了解的人写的和审的。凡是文献里和本文内容有冲突的，一律以本文为准。

## 1 一般问题的基组选择

### 1.1 不同方法计算时基组选用的差异

注意，不同方法对基组的要求是截然不同的，体现在以下两点：

(1) 对基函数角动量和 zeta 数的依赖性：

普通泛函的 DFT 及 HF 计算（下文直接说“DFT”时指代这两种情况）不需要太多高角动量基函数，而 s、p 角动量基函数对改进结果意义明显更大。因此对比如 B3LYP，用 3-zeta 基组 6-311G\* 相对于 2-zeta 基组 6-31G\* 改进明显，而 6-311G(3df,2pd) 比 6-311G(2d,p) 耗时高得多却不会有什么改进。有些人居然用 6-31G(2df,p) 这种基组做 DFT 计算，显然极为荒诞，连 3-zeta 还没达到，亟待改善的地方不改善，却把计算量增加对 DFT 结果影响较小的 f 极化上，一看就非常缺乏理论知识。

而后 HF、双杂化泛函、多参考方法（下文直接说“后 HF 方法”时指代这三种情况）对基组要求较高，在 DFT 的基础上还需要靠较多高角动量基函数才能充分描述电子动态相关作用。所以，对于 CCSD(T) 等高级别后 HF 方法用

缺乏充分高角动量基函数的 6-311G\*\*真是鲜花插在牛粪上，用高级后 HF 方法至少从 def2-TZVP 起步，即便低级后 HF 方法如 MP2 也得至少 def-TZVP 起步。

MCSCF 方法的重点不在于描述电子动态相关而是静态相关，对基组要求和 DFT 要求相同。

(2)耗时与基函数和 GTF（高斯型函数）数目的关系：

DFT 计算的耗时绝大部分用在双电子积分计算上，而双电子积分形式上正比于 GTF 数目的四次方，因此，体系的 GTF 数是最直接影响 DFT 计算耗时的。

后 HF 计算的耗时主要取决于基函数数目（也和体系电子数有关），因为这直接决定 Slater 行列式或组态函数的数目，以及 AO→MO 积分变换的花费，而在 SCF 迭代过程的双电子积分上的花费则相对次要。

鉴于此，在基函数数目相同情况下，通常 GTF 数越少的基组用在 DFT 上越吃香，比如 Pople、def2、pcseg 这些片段收缩基组属于此情况。Dunning 相关一致性基组（cc-pVnZ 系列）属于半广义收缩，GTF 数目相对较多，且此基组本身也是给后 HF 电子相关计算而优化参数的，所以用在 DFT 上性价比比较低（例如用 cc-pVTZ 就明显不如 def2-TZVPP 用合适），但很适合用在后 HF 计算上。

## 1.2 一般问题的基组选用建议

对于非双杂化泛函的 DFT 计算，根据精度要求，一般问题基组选用建议如下。符号对应于同一档内基组尺寸关系。

笑话级别：STO-3G （极小基）

玩具级别：3-21G （最烂 2-zeta 基组）

最低可接受级别：def2-SV(P)  $\approx$  6-31G\* < def2-SVP  $\approx$  6-31G\*\*  $\approx$  pcseg-1 （2-zeta 基组）

还行级别：6-311G\*\* < def-TZVP （一般 3-zeta 基组）

理想级别：def2-TZVP < def2-TZVPP  $\approx$  pcseg-2 （高档 3-zeta 基组）

无敌：def2-QZVP  $\approx$  pcseg-3 （4-zeta 基组。给 DFT 用很浪费）

对于后 HF、双杂化泛函、多参考方法计算，根据精度要求，一般问题基组选用建议如下。

最低可接受级别：def-TZVP （一般 3-zeta 基组）

较好级别：cc-pVTZ  $\approx$  def2-TZVPP （高档 3-zeta 基组）

高精度计算：cc-pVQZ  $\approx$  def2-QZVPP （4-zeta 基组）

无敌：cc-pV5Z （5-zeta 基组。很浪费且极昂贵，一般通过 TZ→QZ CBS 外推来达到这个档次）

上面说的“最低可接受级别”绝对不是指无脑地用这个级别就能拿来发文章！这就类似于穿裤衩和破背心上街，虽然不至于被警察抓走，但像样点的店都进不去。

学术界现在有共识认为 Pople 系列基组构建得不是很理想。对于 3-zeta 及以上档次的基组，对于 DFT 计算，强烈建议在同等基组尺寸下使用 def/def2、pcseg 系列基组代替 Pople 系列基组，以在差不多耗时的情况下获得更高精度。而对于后 HF 计算，Pople 基组更是极差的选择，强烈不建议用！Pople 基组超过 6-311+G(d,p)大小的版本基本没有实用价值，因为总有性价比更高的其它基组可以用。

单点能计算、几何优化（包括优化到过渡态）、振动分析是最常见的任务类型。几何优化和振动分析对基组的敏感度远低于单点能的计算，而这些任务的耗时又远高于单点计算，所以几何优化绝对用不着大基组，诸如用 6-31G\*\* 这样偏小的基组一般就可以，至多用到 6-311G\*\*或 def-TZVP 这样中等尺寸基组也就够了，用诸如 def2-TZVP 这样档次的基组已经属于明显浪费；等优化完，算单点能的时候再把基组提升一、两个档，是明智的做法。这属于量子化学研究者的最基本常识，完全不必担心有哪个内行人会对此有质疑，详见《浅谈为什么优化和振动分析不需要用大基组》（<http://sobereva.com/387>）。产生 IRC、优化过渡态对基组的要求与优化极小点相同，而且这三个任务用的基组应统一。

对于需要弥散函数的情况勿忘加弥散函数！什么情况该加在前述的《谈谈弥散函数和“月份”基组》中已经明确说了，后文也有提到。

## 2 对不同任务基组的选择

### 2.1 激发态计算

对激发能、电子光谱计算（吸收、荧光、ECD 光谱），适合用的基组在《乱谈激发态的计算方法》（<http://sobereva.com/265>）当中已经写了，这里不再累述。

## 2.2 弱相互作用

用什么基组和弱相互作用类型有一定关系。弱相互作用本质分为两大类：

色散作用： $\pi$ - $\pi$  堆积、范德华吸引

静电吸引主导，少量色散吸引参与：氢键、二氢键、卤键、 $\pi$ -氢键、碳键、硫键、磷键等

色散作用起的比重越大，越需要弥散函数、考虑 BSSE 问题。

弱相互作用能计算的基组建议：

底线：对 DFT 用 6-31+G\*\*或 ma-def2-SVP；对后 HF 用 jun-cc-pVDZ

还成：对 DFT 用 6-311+G\*\*；对后 HF 用 may-cc-pVTZ

不错：对 DFT 用 ma-def2-TZVP；对后 HF 用 jun-cc-pVTZ

较理想：jul-cc-pVTZ

极理想：aug-cc-pVQZ

完美：CBS

对于色散作用占主导的弱相互作用，如果计算能力尚有余裕，相互作用能计算时建议用 counterpoise 方式处理 BSSE 问题，所耗时间会是复合物单点计算时间的两倍多。

含弱相互作用体系的几何优化的基组建议：

色散主导的弱相互作用：may-cc-pVTZ 足够理想，6-311+G\*\*一般够用，6-311G\*\*或 def-TZVP 也能接受

静电主导的弱相互作用：def-TZVP 足够理想，6-31G\*\*或 def2-SVP 也可以接受

## 2.3 NMR 计算

算 NMR 强烈建议用 pcSseg 系列基组 (JCTC,11,132)，专门给 NMR 计算优化。pcSseg-1 只比 6-31G\*\*大一丁点，NMR 计算结果却和比它大得多的 def2-TZVP、cc-pVTZ 差不多，已经算是颇理想了。pcSseg-0 大小和 3-21G 相仿佛，精度虽明显不如 pcSseg-1，但是却比 6-31G\*\*整体还好一点。如果计算量还有富余可以提升到 pcSseg-2，不过比 pcSseg-1 大不少但结果改进有限，不算很划算。如果非要用常见的基组算 NMR，可以用 def2-TZVP，算不动就降到 def2-SVP。IGLO-II 算 NMR 也不错，不过推荐程度低于 pcSseg 系列。不建议用 pople 系列基组算 NMR，非要用起码得 6-311G(2d,p)。

如果研究 NMR 峰的分裂需要算自旋-自旋耦合常数 J，算这个绝对不能用一般的基组，必须用专用的在内核附近有很紧的（指数很大的）基函数的基组才行。强烈建议用 pcJ 系列，用其中 pcJ-1 结果就很不错了。IGLO-III 基组也很适合计算 J。

一些人在 NMR 计算时加弥散，这是愚蠢的。加弥散对 NMR 的计算精度几乎无任何改进。

算 NMR、J 值时，被考察的原子绝对不能用赝势基组，否则根本没法表现内层电子对外磁场的屏蔽。但对相邻原子用赝势基组则完全没关系。

## 2.4 势垒计算

势垒计算的基组按照 1.2 节推荐的就可以。如果已用到了较好 3-zeta 级别基组，计算量还有富裕，想用来进一步提升精度，那么建议加上最低程度的弥散（对重原子加 s,p 角动量的弥散），会令结果有定量改进。值得一提的是 def2 系列加弥散对势垒计算结果的改进没有 pople 基组那么大。

算势垒不要在 2-zeta 基组上加弥散，把加弥散的精力用在增加 zeta 数上会令精度明显提升得更多。

## 2.5 偶极矩、极化率、超极化率

偶极矩、极化率、第一超极化率分别是能量对外电场的一阶、二阶、三阶导数，不了解的话看《使用 Multiwfn 分析 Gaussian 的极化率、第二极化率的输出》(<http://sobereva.com/231>)。算这三个必须带弥散函数，否则结果烂得根本没法用。而且随着导数阶数的增加，对弥散函数要求越来越高。

算偶极矩：用 def2-TZVPD 的话在基组层面就没什么误差了，用 aug-cc-pVTZ 也很好但昂贵得多。如果 def2-TZVPD 用不动就降到 aug-cc-pVDZ，如果还算不动就降到 def2-SVPD，还算不动就干脆放弃或者用 ORCA 程序开 RI 来算。Pople 基组算偶极矩极烂，绝对不能用。

算极化率：根据笔者的测试(待发表)，对此问题最划算的基组从小到大排序为 ZPOL, jul-cc-pVDZ, aug-cc-pVDZ, POL, aug-cc-pVTZ(-f,-d), LPol-ds。其中 ZPOL、POL、LPol-ds 都是 Sadlej 提出的，参见《各种 Sadlej 基组的 Gaussian 格式的定义》(<http://sobereva.com/345>)。其中 aug-cc-pVTZ(-f,-d)是把 aug-cc-pVTZ 的重原子的 f 极化和轻原子的 d 极化砍掉得到的。另外，虽然诸如 def2-SVPD 这样 def2 系列基组带 D 后缀的基组也是给极化率计算优化的，但是

在我的测试中不如上述基组划算，但如果想用也可以。如果用到赝势，建议用 LFK 赝势基组，是在 SBKJC 赝势基组基础上修改得到的，以使得极化率计算精度和全电子 POL 基组相仿佛。

算第一超极化率：根据笔者的测试（待发表），对此问题最划算的基组从小到大排序为 aug-cc-pVDZ, POL, aug-cc-pVTZ(-f,-d), LPol-ds, LPol-fs。微型体系极高精度第一超极化率计算可以用 t-aug-cc-pVTZ 或 aug-pcseg-3，二者大小相近，基函数数目约为 LPol-ds 的两倍。如果你不会产生 t-aug-cc-pVTZ 的定义，可以用这个工具《给基组以 even-tempered 方式增加弥散函数的工具 adddiffuse》(<http://sobereva.com/347>)。

## 2.8 振动频率以及与振动有关的谱的计算

对于计算分子振动频率、红外光谱、VCD 谱，对大体系，这类任务基组用 6-31G\* 其实就够了，最多用到 def-TZVP 也足矣，不能再提高了，再提高纯属白浪费时间。通常来说都这类任务都是基于谐振近似算的，一般都要用频率校正因子来修正谐振近似和理论方法的系统性误差，见《谈谈谐振频率校正因子》(<http://sobereva.com/221>)。经过频率校正因子修正后，用小基组、中等基组、大基组其实结果根本没什么差别。用大基组算这类任务都是菜鸟所为。

虽然 Raman、ROA 谱也是振动谱，但是计算它们需要有弥散函数才能得到较好结果。因为它们在计算时都要涉及极化率对振动坐标的导数，而极化率要算准是需要有弥散函数的。J. Chem. Theory Comput., 7, 3323 (2011) 提出计算 Raman、ROA 的两步方法，在优化和振动分析时不用弥散函数（低至 6-31G\* 高至 def-TZVP 就够），但是计算极化率导数的时候再考虑弥散函数（建议  $\geq$  aug-cc-pVDZ），这样比整个过程都用弥散函数更便宜而且精度没什么损失。Gaussian 支持这种两步方法。（由于 ROA、Raman 是振动谱，所以也应考虑频率校正因子）

## 2.9 全电子相对论计算

全电子相对论计算要用到 DKH、ZORA、X2C 等考虑相对论效应的哈密顿，此时必须要用专门给相对论计算优化收缩方式的基组，比如 cc-pVnZ-DK 系列、SARC 等等，或者索性用完全没收缩的基组如 UGBS。哪些可以用在此文开头简单提及了《在赝势下做波函数分析的一些说明》(<http://sobereva.com/156>)。1.2 节提到的那些普通基组绝对不能拿来全电子相对论计算。

一些初学者对全电子相对论计算一无所知，居然用了赝势基组还用诸如 DKH2 相对论哈密顿，此时结果完全错误！用相对论哈密顿时必须是全电子基组。

不过，如果不考虑相对论效应中的标量部分而只考虑旋轨耦合部分的话可以照常用普通全电子基组和赝势。

## 2.10 波函数分析

对于基于实空间函数的波函数分析，比如讨论电子密度、ELF、静电势，做 AIM、RDG 分析之类的，基组用 6-31G\* 一般就足够得到定性正确的结果，用到 6-311G\*\* 或 def-TZVP 就已经绰绰有余了，如果是阴离子体系再加层 s,p 角动量弥散函数也足够了，再增大基组没什么用处，因为这些实空间函数随基组增大很快就会收敛。NBO 分析对基组的要求与此相同。加弥散函数对实空间函数的分布几乎没有任何影响，因此加了只会白浪费计算时间，但也不会令结果更糟。

对于 CDA 分析、Mayer 键级、Wiberg 键级、Mulliken 分析、SCPA 分析，还有 Multiwfn 里基于 Mulliken/SCPA 绘制 PDOS、OPDOS 图之类的直接基于基函数做的分析，基组用大了不会改进结果，6-311G\*\* 下的结果绝不比 6-31G\*\* 下的更好。若把基组继续提升到 4-zeta，结果只会变差。如果基组带了弥散函数，那么结果根本没法用！根本没法用！（很重要所以说三遍）绝对不能用带有弥散函数的基组做这些分析，因为弥散函数完全没有化学意义，因此直接基于基函数来划分原子空间的这些方法结果会一塌糊涂。我无数次看到文献中，包括一些 IF 很高的期刊的文献中居然带着弥散函数做这些分析，作者完全没有波函数分析的常识。关于弥散函数对这些分析带来的害处，看笔者曾经写的《原子电荷计算方法的对比》(<http://www.whxb.pku.edu.cn/CN/abstract/abstract27818.shtml>)、《广义化的电荷分解分析(GCDA)方法》(<http://dx.doi.org/10.12677/japc.2015.44013>)、《分子轨道成分的计算》(化学学报, 69, 2393)。

做波函数分析如果涉及到赝势，一定要用小核赝势，这里有详述：《在赝势下做波函数分析的一些说明》(<http://sobereva.com/156>)

## 2.11 溶剂化能（溶解自由能）

利用隐式溶剂模型算溶剂化能一定要使用溶剂模型专门参数化的级别，比如 SMD 参数化最好的级别是 M05-2X/6-31G\*，所以建议在 M05-2X/6-31G\* 级别结合 SMD 算溶解自由能。而基组用得更大，比如 aug-cc-pVTZ，结果只会变糟。关于这点在此文有详述：《谈谈隐式溶剂模型下溶解自由能和体系自由能的计算》(<http://sobereva.com/327>)。

## 2.12 周期性计算

对于 Gaussian、Crystal 这样基于高斯基函数的程序，算孤立体系常用的 6-31G\*\* 基组也适合有机周期性体系计算。但很多专门给分子计算的基组不宜直接用在周期性计算上，如 6-311G\* 计算 Si。因为这些基组往往有指数较小的呈一定弥散特征的基函数，对于描述原子轨道的尾巴特征有益，但在周期性体系中并不需要描述这点，还可能造成数值问题，故需要将这些基函数的指数人为调大或去掉。一些基组有冗余的基函数也应适当去掉。本身带弥散函数的基组更是绝对不能用于周期性计算的，否则麻烦极多。

适合周期性计算的高斯函数基组可以从 Crystal 网站 [https://www.crystal.unito.it/basis\\_sets.html](https://www.crystal.unito.it/basis_sets.html) 获得，但需要改格式：其中的 pob-TZVP 是基于 def2-TZVP 修改的适合固体计算的版本，比用其它的 3-zeta 基组可以得到更好结果。后来 2019 年提出的 pob-DZVP-rev2 和 pob-TZVP-rev2 在 SCF 稳定性上比 pob 初代有了很大提升，而且改进了晶格常数计算结果。pob 系列基组也可以从这里获得定义：

[https://www.chemiebn.uni-bonn.de/pctc/mulliken-center/software/ssc\\_basis/basis](https://www.chemiebn.uni-bonn.de/pctc/mulliken-center/software/ssc_basis/basis)。

## 2.13 超精细耦合

Fermi contact 是超精细耦合常数的主要来源，需要算原子核处的自旋密度，建议用专为 DFT 计算超精细耦合优化的 EPR-II、EPR-III 基组（Gaussian 内置，仅对 H, B, C, N, O, F 有定义），或适合计算核自旋-核自旋耦合常数的 IGLO-III、pc-J 系列基组计算。这类基组都有比较紧的 s 函数，能够比较好地描述原子核处的自旋密度。

## 2.14 显式相关计算

应当用专门给显式相关计算优化的基组，最知名的是 cc-pVnZ-F12 系列。

## 2.15 涉及到密度拟合加速的计算

一般建议用有专门配套的密度拟合基组的基组，诸如 def2 系列。而没有配套的密度拟合基组的基组，如 pople 系列，使用密度拟合计算时就得让程序自动产生密度拟合基组，这样产生的比标配的耗时会更高，因此加速效果没那么好。借用相似级别或更高级别的基组的密度拟合基组也可以，但终究不完美。关于密度拟合加速详见《大体系弱相互作用计算的解决之道》(<http://sobereva.com/214>)。

# 3 基组相关的若干问题

## 3.1 什么时候用赝势？

赝势的主要用处在于：(1)将化学上不感兴趣的内层电子以等效的势场描述，从而不需要将内层电子显式地表达出来，大大节约了计算量 (2)可以等效地体现标量相对论效应。从第四周期相对论效应开始显现，但不考虑也问题不大；而对第五周期及之后相对论效应是不可忽视的，不考虑甚至结果定性错误。

赝势和赝势基组是搭配使用的。一般来说，从第四周期开始就推荐使用赝势基组，这样不仅显著节约了耗时，还可以等效纳入相对论效应改进结果（前提是使用的是相对论赝势 RECP，即当初拟合赝势时已考虑了相对论效应）。有兴趣可以看《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>) 和前述的《在赝势下做波函数分析的一些说明》，赝势基组的选择在前述的《谈谈赝势基组的选用》中有详细讨论。

## 3.2 基组包含哪些元素？

只有极少数基组，如 UGBS、WTBS 是几乎覆盖整个周期表的，大多数基组只对周期表中的一部分有定义。基组定义了哪些元素，可以自行去 BSE 网站上，点击相应基组查看。

## 3.3 CBS 是啥意思？

CBS=complete basis set（完备基组）。用 CBS 相当于用了无穷大基组，此时基组不完备性导致的结果误差为 0。文献中经常有 CCSD(T)/CBS 这种写法，当然实际计算用的基组不可能无穷大，CBS 下的结果是通过有限大小基组下的结果外推来达到的，具体见前述的《谈谈能量的基组外推》。

## 3.4 基组怎么获得？

由于绝大多数量子化学程序用的基函数都是高斯函数，数学形式是一致的，因此高斯函数基组可以在任何基于高斯函数的量子化学程序中使用。基本上所有量子化学程序都有自己的基组库，包含主流的基组，可以直接使用。

但是内置的基组往往不够全面，往往需要从其它地方获得基组定义来在程序中使用。可以从其它程序基组库中借来，或者根据基组原文献中的数据自行手写，或者从一些基组开发者的网站上获取。最常规的获得外部基组的做法是通过 BSE 基组数据库，这是基组最全的公开的基组数据库（极个别基组对于个别元素的数据可能有错误）。详

见前述的《在线基组和赝势数据库一览》。

### 3.5 关于混合基组

混合基组是指计算时有的原子用某种基组，而有的原子用另一种基组。适当使用混合基组可以令计算耗时大为降低，而精度损失不多，建议大家充分利用这种做法。例如，计算一个大体系的化学反应过程，反应部分只涉及其中一个区域，那么可以只给这个区域的原子用大基组（如 def2-TZVP）而其余部分用中/小基组（如 6-31G\*乃至 3-21G）。再比如，若体系有的地方局部带大量的负电荷，那么没必要对整个体系所有原子都加弥散函数，只给那部分的原子加上弥散函数就足够了。计算配合物时通常给过渡金属原子用赝势基组而轻原子用全电子基组，这实际上也属于使用了混合基组。

Gaussian 中用混合基组的方法见前述的《详解 Gaussian 中混合基组、自定义基组和赝势基组的输入》。

### 3.6 BSSE 问题

高斯函数基组实际用的时候基本上都是以原子为中心的，某原子带的基函数会侵入到它附近原子的空间内，这导致通过整体减片段来求能量差的时候会有 BSSE（基组重叠误差）问题，算相互作用能的时候会高估作用强度。解决 BSSE 问题有 counterpoise、gcp 等方法，特别强调一下用 counterpoise 的时候只适合在计算能量的时候，强烈不建议在优化和振动分析的时候使用，否则无解析导数，对大体系慢得吐血，且一般也带来不了什么可察觉得到的改进。详见《谈谈 BSSE 校正与 Gaussian 对它的处理》(<http://sobereva.com/46>)。

### 3.7 给氢加不加极化函数、加不加弥散函数？

和氢关系密切的计算，比如计算氢键、氢转移反应、质子化能、氢的 NMR 谱等等才很有必要给氢加极化函数，其它情况不给氢加极化也不会带来多大误差。

给氢加弥散函数通常来说意义很低，在《谈谈弥散函数和“月份”基组》(<http://sobereva.com/119>)中特意专门做了说明，这里不再累述。

### 3.8 哪些任务用的基组必须一致？

很多初学者，道听途说网上一一些乱七八糟说法，对概念稀里糊涂，经常问某某计算用的基组是否必须得和优化用的基组一致之类。实际上，只有走 IRC、做振动分析的时候用的级别（显然包括所用的基组）必须和几何优化（包括优化过渡态）时候的一致，除此之外，任何常见的任务之间都绝对不要求所用基组一致!!!

### 3.9 关于 Pople 系列基组加极化函数

Pople 系列基组的一个特点是自定义强，对 6-31G 和 6-311G 后面极化函数加多少、加什么角动量的可以自己随意控制，对重原子可以是 d,df,2d,2df,3d,3df,3d2f, 对轻原子可以是 p,pd,2p,2pd,3p,3pd,3p2d。很多人乱加，不注意各角动量的平衡性，比如说按这个顺序逐渐增加试图改进结果是合适的：(d,p) (2d,p) (2df,2p) (2df,2pd) (3d2f,2pd) (3d2f,3p2d)。但是如果使用诸如(3d,3p2d)，就很不合适了，对重原子 d 极化都那么多了怎么一个 f 都没有？而且氢的极化给那么多，重原子的极化显得也偏少。另外，对 6-31G 这个 2-zeta 基组，给个(d,p)或至多(2d,p)也就够了，如果用 6-31G(2df,2pd)就很蠢了，把闲资源用在加那么多极化函数上带给结果的改进远远不如先把基组升到 3-zeta！

另外，pople 基组主要适合做粗鄙或中庸的 HF/DFT 计算，哪怕一般精度的后 HF 都不适合 pople 基组（适合用 dunning 相关一致性基组）。6-31G\*、6-31G\*\*、6-311G\*\*这样的对 DFT 可以用，但是再想提高精度，那就强烈不建议用 pople 基组了，自行给它增加更多的极化函数远不如直接用人家专门搭配好的基组（def2、pc 系列）。比如 def2-TZVP 和 6-311G(2df,2p)耗时相仿佛，用前者明显比用后者结果更好。

顺便再说一下，**6-31G、6-311G 不加极化函数根本没法用！**6-31G\*相比 6-31G、6-311G\*相比 6-311G 精度一个天上一个地下，虽然加了极化耗时高了些，但这是超划算的。频繁看到初学者居然直接就敢用秃秃秃的 6-31G、6-311G，明显是不懂极化函数有多么必不可少。**用 6-31G、6-311G 暴露出研究者在基组选用常识上的极度无知**，如果确实连 6-31G\*都用不起，索性降到 3-21G，这体现出因计算能力不足而迫不得已，并不丢人。

### 3.10 乱用弥散函数的问题

什么时候该用弥散函数在前述的《谈谈弥散函数和“月份”基组》当中专门说了，本文前面也提到了很多次什么时候该加什么时候不该加弥散函数。老夫长期在网上回答计算化学问题，看到别人乱用弥散函数的情形真是太多了。再次强调一下，如果你已经仔细看过上文以及《谈谈弥散函数和“月份”基组》了，**碰到一个问题不确定该不该加弥散函数，那么答案就是：别加弥散函数!!!**根据洒家网上长期答疑经验，不该加弥散函数的时候加了弥散函数的情况，比该加弥散函数却没加的情况多至少 10 倍！

不该加弥散函数的时候加弥散几乎对结果完全没改进，还导致许多问题，尤其是计算量暴增、难收敛。要知道，同样是加一层 sp 基函数，如果加的是弥散函数，远比增加一级 zeta 数耗时多得多！即曰，对重原子，相对于 6-31G\*，增大到 6-31+G\* 的耗时显著高于增大到 6-311G\*（体系越大这个现象越明显）。因为弥散函数指数小，延展范围广，量化程序不容易靠积分屏蔽（integral screening）策略来忽略掉对结果贡献很小的积分来节约计算量，而且本身弥散函数也容易导致收敛需要更多圈数。而对于不是非得有弥散函数不可的情况，6-311G\* 比 6-31+G\* 结果好到不知哪里去了。即便是对于弥散函数较有用的问题，比如色散主导的弱相互作用体系优化，考虑到计算量以及任务类型，我也推荐用 6-311G\* 而非 6-31+G\*，因为前者本身此时结果也不比后者差，虽然它没有弥散函数，但是 3-zeta 最外层本身就体现出一点弥散的特征，也基本够了，而且 3-zeta 对价层电子的描述准确度也比 2-zeta 充分得多。

此外，对于后 HF 方法，乱加弥散函数，还不知道用足够的高角动量极化函数，结果更是荒诞的！JACS, 128, 9342 (2006) 里专门报道过，用 MP2 优化苯，用诸如 6-31++G\*、6-311+G\*、6-311++G\*\* 等，优化后的平面结构的苯居然有虚频，换句话说，在这样的级别下苯的极小点结构是弯的！在 <http://bbs.keinsci.com/thread-12653-1-1.html> 这个帖子里，有人用 MP2/6-311++G\*\* 结合 SDD 去优化碘代乙炔，结果没虚频的结构竟是弯的，也是类似原因。可见缺乏计算级别选用常识的话，连看似最简单的问题都可能给出完全错误的结果。