**使用冗余内坐标有效优化分子几何结构**

讨论了从头算分子几何结构的优化。基于 30 个最小化实例和 15 个鞍点优化实例的比较，确定了坐标系、近似和精确 Hessians 以及步进控制的最有效组合。与使用冗余坐标相比，使用一组建议的额外冗余内坐标可以显着减少几何优化步骤的数量。针对最小和鞍点优化测试了各种更新方案，包括组合公式。给出了威尔逊**B**矩阵一阶和二阶导数的完整表达式，从而避免了需要通过有限差分方法进行计算。所提出的方案似乎是迄今为止最有效、最稳健和普遍适用的方案。

**1 引言**

在从头算量子化学方法对化学感兴趣的问题的许多应用中，一个重要的步骤是优化分子的几何构型，以确定极小点（对应于稳定的分子结构）或者鞍点（对应于分子的过渡态）。多年来，已经开发和改进了大量用于此类优化的方法。虽然有基于优化理论和数值分析的标准技术，但最有效的方法需要高度适应分子几何优化的特殊要求。对于坐标系的选择和拟牛顿理论中 Hessians 的选择尤其如此，这两者都会严重影响优化的性能。

在本文中，对从头算分子势能面的优化进行了回顾和分析，以建立最有效的最小值和鞍点锁定方案。基于 30个最小化实例和 15 个鞍点优化实例的比较，确定了坐标系、Hessian 和步进控制的最佳组合。引入了额外冗余内部坐标，并显示相对于使用标准冗余坐标显着减少了几何步骤的数量。此外，考虑各种 Hessian 更新用于极小点和鞍点优化，包括组合公式。所提出的优化方案在最小化和鞍点优化方面略有不同，似乎是迄今为止最有效和最稳健的方案

对于任何使用内坐标（例如冗余和额外冗余坐标）的优化，Wilson **B**矩阵起着重要作用。尤其是，对于笛卡尔坐标（其中，计算了能量、梯度和Hessian矩阵）和内坐标（其中，确定了优化的步长）之间的转换，Wilson **B**矩阵的推导是必须的。在本文中，这些推导的解析表达式是二阶的，避免了它们的有限差分计算。

Peng 等人引入了冗余内坐标1。我们在这里将这些和额外冗余坐标与Lindh 等人的模型Hessian 结合起来，2 并且注意到Eckert 等人通过使用具有自然内坐标的 Hessian 模型取得了良好的性能。3 以前的工作集中在最小化的性能上，尽管 Baker 和 Chan 提出了鞍点优化的基准。4 在这项工作中，我们对过渡态优化进行了基准测试并且也做了极小点优化的基准测试。

这里引入方案的计算尺度是*O*(*N*3)，即原子数的立方。对于通过传统从头算方法研究的中等大小的分子，与计算能量、梯度和可能的 Hessian 矩阵的成本相比，建立内坐标系和执行必要变换的成本可以忽略不计。然而，对于非常大的系统，建立坐标系和确定几何步长的成本可能成为计算中的瓶颈。已经开发了通过减少缩放来解决这个问题的方法。5, 6

在第II部分讨论了坐标系和Wilson **B** 矩阵的导数之后。我们在第 III部分中考虑了近似的 Hessians 及其更新。接下来，我们在第IV部分讨论了步长控制和收敛标准。然后，在第V部分我们对不同方案进行了详细比较，确定了坐标系、近似 Hessians 以及极小点优化和鞍点优化步长控制的最佳组合。结论在第VI部分给出。

**2 坐标系**

坐标系的选择对于几何优化的效率至关重要。因为梯度和Hessian矩阵通常在笛卡尔空间计算，直接的选择是笛卡尔坐标系。然而，许多研究表明，笛卡尔坐标系通常优于精心挑选的内坐标集合——即键长、键角和二面角组成的集合。内坐标系可分为冗余和非冗余。在非冗余内部坐标系中，坐标数等于内部自由度数 3 N - 6 ( 5 ) ，其中 N 是系统中的原子数；在冗余集中，坐标比自由度多。

Z 矩阵坐标是非冗余内坐标的典型例子。7 尽管使用它们用于优化的例子越来越少，1,8,9 但它们仍然可用于指定分子几何形状。使用 Z 矩阵坐标进行优化（尤其是对于环系统）的一个困难是决定包含哪些原始内坐标以及排除哪些内坐标。内坐标选择不当可能会严重降低优化器的性能。

自然内坐标通过形成原始内部的局部线性组合来解决这个问题。10,11 尽管规则和特殊情况的数量非常广泛，但可以通过应用一组规则来自动设置坐标。然而，这些坐标的性能非常好。3,9 在这项研究中，我们关注冗余内坐标1 和（非冗余的）离域内坐标。12 在下文中，我们首先在第二节II A 和 II B中考察冗余内坐标及其导数；接下来，我们在第 2 节的II C中考察离域内坐标。

**A. 冗余内坐标**

冗余内坐标在参考文献1中定义。我们将在这里简要回顾定义并考虑有关选择坐标系的一些要点。

**1. 设置分子的冗余内坐标**

建立冗余内坐标的关键是确定哪些原子相互键合以及每个键的性质。键的分配取决于共价键的可用值（列表）和作为原子序数函数的范德华半径的值。

我们把常规键分配给原子间距离小于或等于其各自共价半径之和的 1.3 倍的所有原子对。一旦常规键就位，我们就会识别出孤立的、不连贯的片段。如果找到两个（或更多）这样的片段，属于不同片段的两个原子之间的最短键定义了片段间键坐标。小于 2 Å 或小于 1.3 倍的片段间键坐标的其他片段间距离构成辅助片段间键。最后，我们来检查氢键。通过检查氢原子与小的电负性原子 X的距离，其中 X = N、O、F、P、S、Cl。如果这个氢原子到另一个小的电负性原子Y的距离大于H和Y的共价半径之和，但小于它们的范德华半径之和的0.9倍，如果角度 X–H…Y 大于 90°，那么在 H 和 Y 之间指定氢键。

我们还提出，可以通过将辅助键分配给原子之间的距离小于其共价半径总和的 2.5 倍的所有原子对来创建一个额外的坐标系统。这将在具有共同相邻的原子之间生成键坐标 - 例如，跨键角。稍后我们将回到这个替代坐标系。

一旦键合到位，就很容易生成弯曲坐标。键角分配给所有三个原子 A、B 和 C 的集合，其中 A 与 B 键合，B 与 C 键合。仅允许常规键、片段间键和氢键产生键角——辅助键和辅助片段间键不产生角度。角度接近线性 (∠A B C >175° ) 的弯曲坐标需要特别注意。在这些情况下，分配第二个正交弯曲坐标，确保线性结构稳定。

二面角（弯曲坐标）以与角度相同的方式分配。包括所有四个原子 A、B、C 和 D 的集合，其中 A 与 B 键合，B 与 C 键合，C 与 D 键合，前提是 ∠A B C≠ 180° 和 ∠ B CD≠ 180° 。同样，只有常规键、片段间键和氢键会产生扭曲坐标——辅助键和辅助片段间键不会。

如果一个分子包含四个或更多原子并且使用上述方案没有找到二面角，则进行额外的搜索以确保添加了平面外弯曲。尝试四个原子的任意组合，直到找到明确定义的二面角（如果分子是线性的，则找不到）。仅添加一个二面角可能会破坏分子对称性。因此，还添加了所选二面角的所有定义明确的排列（ ∠ A B D C ， ∠ B C D A 等），最多生成 12 个独立的扭转坐标。

**2. Wilson B矩阵**

Wilson 的 B 矩阵给出了冗余内坐标和笛卡尔坐标之间的关系。 13 它的矩阵元由下式给出



其中*qi*是内坐标，*xj* 是原子的笛卡尔位移坐标。因此 B 矩阵是一个长方阵；行数等于内坐标数，列数等于笛卡尔坐标数。 B 矩阵中矩阵元的显式表达式在II B部分给出。

笛卡尔坐标的微小变化转化为内坐标的微小变化：



逆变换由下式给出：



这里（或广义逆）引入了一个矩形矩阵的伪逆。任何矩形矩阵都存在广义逆，无论它是否是秩不足的。一般来说，它是从 **B** 的奇异值分解中获得的；对于一个方形非奇异矩阵，它可以简化为标准逆矩阵。

广义逆的性质如下： 如果线性方程 Eq（2）是超定的，那么等式Eq（3）中的伪逆阵**B**+返回使残差范数最小化的解δ**x**；另一方面，如果系统是欠定的，则伪逆最小化解的范数。此外，根据广义逆理论，可以得出矩阵



构成到 **B** 范围的投影矩阵。应用于任何物理或非物理的内部位移坐标δ**q** 集，它会产生一个唯一的向量



这在物理上是有效的，因为它与分子笛卡尔坐标的变化一致。

**3. 笛卡尔和冗余内坐标之间的转换**

根据 Wilson B 矩阵及其伪逆矩阵，我们可以通过以下方式建立梯度和 Hessians矩阵在内坐标系和笛卡尔坐标系之间的关系：





此处K矩阵由下式给出：



通过将广义逆应用于Eq 6和7，得到相反的变换：





显然，为了变换 Hessian，**B** 的矩阵元必须相对于笛卡尔坐标微分。在IIB部分将给出导数。

**4. 梯度和Hessian的投影**

在冗余坐标系中，必须注意确保梯度在物理上是有效的——也就是说，它对应于内部坐标的有效变化。在数学上，一个有效的梯度必须属于 **B** 的范围。一个无效的梯度，用Eq. (4) 可以通过投影生成一个有效的梯度，对于Hessian来说同样如此。





为了确保有效的梯度和 Hessians，这些投影可以在优化过程中的任何时候进行。尤其是，必须对不精确的初始 Hessians 和更新的 Hessians 以及插值梯度（产生于线性搜索法）时执行此投影。在显式计算梯度和 Hessians 时，也可以执行投影以消除在坐标变换过程中引入的任何数值噪声。

**5. 优化过程中步子的变换**

一旦确定了内坐标中的步长，就必须将其转换为笛卡尔坐标——计算能量、梯度和 Hessians 的参考系。然而，由于笛卡尔坐标是直线的，而内坐标是曲线的，所以没有简单的有限位移变换。相反，几何步骤的转换必须迭代完成。

考虑一个具有初始笛卡尔坐标 **x**0 和初始内坐标 **q**0 的系统。对于内坐标中的步长 **s̃**q，新笛卡尔坐标的第一个估计值由下式给出：



从新的笛卡尔坐标**x**1出发来确定一组更新的内坐标**q**1。接下来，内坐标的需求和实际变化之间的差用下式计算：



其中 *k* = 1 是在第一次迭代中，注意从角度中删除任何 360° 的倍数。然后通过以与初始步骤相同的方式转换上式中的差值来细化笛卡尔位移。

方程（14）和（15）定义了我们的迭代过程。当笛卡尔坐标中的均方根变化小于10-6时，如果此均方根变化与前一次迭代的差异小于10-12时，或者如果迭代次数超过 25次，则宣布收敛。在细化期间，将 与  进行比较。在极少数情况下， ，这时我们恢复到Eq.（13）中的初猜**x**1；否则，用于下一次优化迭代的就是**x***k*了。

**B. 冗余内坐标的导数**

**1. 键长**

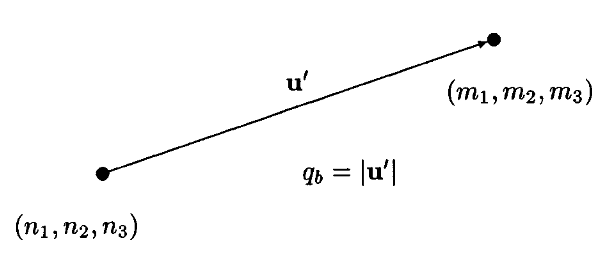


图1 两个原子间的键

我们首先考虑与键伸缩振动相对应的冗余内坐标*qb*，有时称为拉伸坐标。图1显示了两个原子m和n之间的键，笛卡尔坐标为（m 1，m 2，m 3）和（n 1，n 2，n 3）。键向量由下式给出：



并且相关的键长表示为。微分方程Eq. (16) 在笛卡尔坐标下，我们得到



我们在式中引入符号因子



且，具有分量 (u 1 , u 2 , u 3) 的归一化键向量如下



因此，除了符号因子，拉伸坐标方程Eq. (17)的导数完全由归一化键向量 **u** 给出。进一步微分产生键拉伸坐标的二阶导数



式中*λu*是方程Eq. (16)的键长。

**2. 键角**

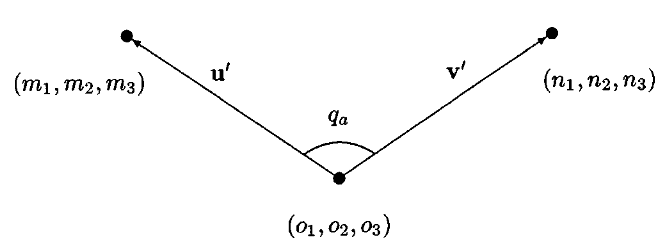


图2 两个键之间的键角

接下来，我们考虑图 2 中从 o 到 m 和从 o 到 n 的向量之间的键角或弯曲坐标 *q a*。键向量 **u’** 和 **v’** 由下式给出：





并且相关的归一化键向量和键长分别由 **u** 和 **v** 以及*λu*和 *λv* 表示。就这些实体而言，的键角由下式给出：



键角的导数比键长的导数更复杂，部分原因是中心原子 o 与末端原子 m 和 n 不等价。为了以紧凑的方式表达这些导数，我们引入垂直于 **u** 和 **v** 的向量 **w’** ，写作：



其中线性情况（*qa* = 180° ）被明确处理。向量[1,-1,1]和[-1,1,1]是任意选择的，以确保生成垂直向量。相关的归一化坐标表示为**w**。

键角的一阶导数对应于笛卡尔坐标的一阶导数，现在可以用以下方式表示，对所有角度都是有效的：



接下来，使用关系



对于键角的二阶导数，我们得出以下表达式：



其中最后一项包含一阶导数。请注意，对于线性情况 *qa* = 180°，由于分母中的因子 sin(*qa*)，二阶导数未定义；然后将导数的分量简单地设置为零。

**3. 二面角**

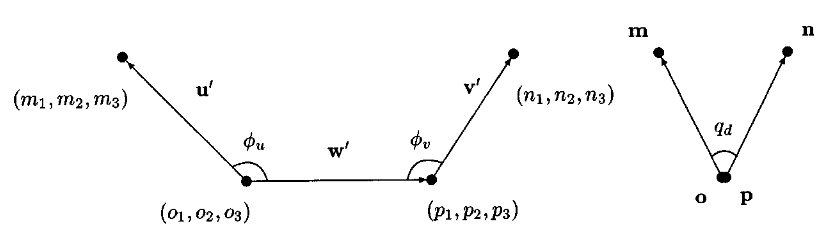


图3 四个原子之间的二面角。右边是沿着 **o** 和 **p** 之间的键为中心的侧视图。

在图 3 中，二面角 *qd* 被引入为由三个点 **m**、**o** 和 **p** 定义的平面与由 **n**、**p** 和 **o** 定义的平面之间的角度。通常，*qd* 被限制在之间。沿着从 **o** 到 **p** 的向量看，如果**o**和**m**之间的向量必须顺时针旋转（角度小于或等于*π*）以与从**p**到**n**的向量重合，则称二面角为正。

为了计算二面角及其导数，我们引入了三个键向量







相应的键长是*λu*、*λv*和 *λw*，而归一化的键向量表示为 **u** 、 **v** 和 **w**。二面角坐标的表达式（扭转）现在变成



图3中键角*ϕu*和*ϕv*满足下面关系式





如果键角 *ϕu*和*ϕv*中的至少一个等于 0° 或 180°，则二面角 *qd* 是不确定的，这意味着原子 **m**、**o** 和 **p** 或原子 **o**、**p** 和 **n** 位于同一条线上。

微分后，我们发现二面角 *qd*的一阶和二阶导数由下面的关系式给出





其中，置换算子*Pij*交换了指数*i*和*j*。二阶导数的表达式可能看起来大得笨拙，但进一步检查表明，大多数项都是由相同的组件构成的。由于方程式Eq. 35中的最后两项，只有当 *i* 和 *j* 指向不同的笛卡尔方向时，*k* 才唯一地定义为第三个笛卡尔分量。由于末端原子 **m** 和 **n** 之间没有耦合，我们得出结论



这也是方程式Eq. 35的一个特例。

**C. 离域内坐标**

一个非冗余的内坐标系可以很容易地从具有非零特征值的 **B** 矩阵的特征向量构造。这些离域内坐标是原始内坐标的线性组合，通常这些组合是非局域的，因此称为“离域”。

特征向量给出了离域和冗余内坐标之间的关系。通过分解冗余**B**矩阵，我们可以构造一个非冗余**B**矩阵，其行为与冗余**B**矩阵基本相同，根据笛卡尔坐标确定非定域坐标。这里导出的变换方程也适用于非冗余**B**矩阵，可以看作是冗余矩阵的一个特例。