

用重正交化 Lanczos 法求解大型非正交归一基稀疏矩阵的特征值问题

焦宝宝

Eigenvalue problems solved by reorthogonalization Lanczos method for the large non-orthonormal sparse matrix

Jiao Bao-Bao

引用信息 Citation: *Acta Physica Sinica*, 65, 192101 (2016) DOI: 10.7498/aps.65.192101

在线阅读 View online: <http://dx.doi.org/10.7498/aps.65.192101>

当期内容 View table of contents: <http://wulixb.iphy.ac.cn/CN/Y2016/V65/I19>

您可能感兴趣的其他文章

Articles you may be interested in

用重正交化Lanczos法求解大型非正交归一基 稀疏矩阵的特征值问题

焦宝宝[†]

(上海理工大学理学院, 上海 200093)

(2016年4月27日收到; 2016年7月5日收到修改稿)

在 ^{116}Sn 原子核壳模型结构下, 利用广义辛弱数截断多体空间得到的哈密顿矩阵是一个大型的实对称非正交归一基稀疏矩阵, 因此求解大型矩阵的能量特征值和能量特征向量是原子核物理上的一个重要问题. 为此, 利用重正交化Lanczos法与Cholesky分解法和Elementary transformation法相结合的方法, 实现了用内存较小的计算机求解大型实对称非正交归一基稀疏矩阵的特征值和特征向量. 用这种方法计算小型矩阵得到的特征值和精确值符合得较好, 且运用这个方法计算了 ^{116}Sn 壳模型截断后的大型非正交归一基稀疏矩阵的能量特征值, 得到的原子核低态能量与实验测量能量相吻合, 计算结果表明Lanczos法在Matlab编程和大型壳模型计算中的精确性和可行性. 此方法也有助于求解一些中质核或者重核的低态能量, 同时也有利于用内存稍大的计算机求解更大的非正交归一基矩阵的特征值问题.

关键词: 原子核壳模型, 特征值, 哈密顿量**PACS:** 21.60.Cs, 27.60.+j**DOI:** 10.7498/aps.65.192101

1 引言

原子的壳层结构是解释元素周期性的基础, 能很好地说明元素的物理、化学性质周期性源于原子内电子的壳层排布, 同时壳层模型也可以很好地解释许多核基态和低激发态的一些性质. 基于相互作用和模型空间来研究原子核壳模型结构^[1,2], 找到了原子核能量特征值与哈密顿元素之间的线性关系^[3,4]. 因为模型空间通常太大而不容易或者不能计算, 所以人们经常基于许多方法来截断^[5-9]壳模型空间, 然后用计算软件通过不同的对角化方法来找到近似的低态能量. 通常用这种方法研究幻数核的基态和低激发态, 但是对于非幻数核壳层模型的基态却不是一个很好的近似态.

许多理论研究和工程设计都涉及大型矩阵的特征值问题. 例如: 金融学^[10]、大型结构动力学^[11]和弹性结构的动力分析^[12]就要计算矩阵的部分特

征值和特征向量. 在大型原子核壳模型计算中, 感兴趣的往往是几个较低的能量特征值, 而我们遇到的问题是用内存较小的计算机通过Matlab编程软件计算出大型非正交归一基稀疏矩阵最小的几个特征值和特征向量. 因为在计算中选择的是非正交归一基, 所以会得到重叠矩阵和哈密顿矩阵, 这将在第2部分给出解释. 对于价键组态空间维数较小或者所占内存较小的实对称非正交归一基稀疏矩阵, 可用Matlab自带的函数来计算矩阵的特征值和特征向量. 在计算的过程中, 计算机需要将重叠矩阵和哈密顿矩阵读取到电脑的内存中去, 然后通过计算得到矩阵的特征值, 但是对于价键组态空间维数较大或者内存较大的矩阵, 就会出现内存不足而导致计算停止的问题. 而且也没人用内存较小的计算机来计算通过广义辛弱数^[13]截断 ^{116}Sn 壳模型得到大型非正交归一基稀疏哈密顿矩阵的能量特征值和特征向量, 更何况在计算较多价键组态的中质核或者重核时, 矩阵的维数或者存储内存会更

[†] 通信作者. E-mail: baobaojiao91@126.com

大, 尽管计算机在飞速的发展, 还是不能满足计算的要求. 因此, 寻找用内存较小的计算机来求解大型非正交归一基哈密顿矩阵能量特征值的方法是有意义的.

壳模型考虑价键组态混合和剩余相互作用, 通过求解薛定谔方程得到原子核的波函数以及基态和激发态的能量等. 基于锡元素的单极最优化相互作用^[14], 通过文献^[13]的计算得到的哈密顿矩阵是一个大型的实对称非正交归一基稀疏矩阵, 所以就要求解大型矩阵的部分广义特征值和特征向量. 在 Matlab 编程计算中, 用 Cholesky 分解法按行(列)把重叠矩阵分解为一个上三角矩阵和一个下三角矩阵, 并把矩阵按行(列)保存为很多内存较小的文件并依次读取, 用 Elementary transformation 法按行(列)来求解出上三角矩阵和下三角矩阵的逆矩阵, 把逆矩阵按行(列)保存为许多内存较小的文件, 既克服了计算机内存不足的困难, 也便于下一步的读取与计算. 有了这两步计算和变换, 就把广义特征值问题转化为标准特征值问题了. 最后, 用重正交化的 Lanczos 方法得到大型非正交归一基稀疏矩阵最小的几个能量特征值和能量特征向量^[15]. 比较结果表明, 算法在保证计算可行性的同时, 也保证了计算的精确性, 而且这项工作不仅有助于求解一些中质核或者重核的低能态, 也有利于用内存稍大的计算机求解更大的非正交归一基矩阵的特征值问题.

2 广义辛弱数截断下的广义特征值问题

原子核壳模型能很好地描述轻质核及中质核的低态能量和电磁跃迁以及衰变等特性, 同时也给出了核子从一个轨道移动到另一个轨道需要的能量和变化的量子数, 而原子核壳层模型的关键是选择一个合适的哈密顿量. 原子核的态可以看成是模型空间所有可能的态的叠加, 壳模型使用单体和相互作用的哈密顿量, 通过求解壳模型空间的薛定谔方程得到原子核的波函数、基态能量和低激发态的能量等. 假定原子核有 N 个核子, 在非相对论近似下, 体系的薛定谔方程可以写为

$$\hat{H}V_{\alpha}(1, 2, \dots, N) = E_{\alpha}V_{\alpha}(1, 2, \dots, N), \quad (1)$$

其中 \hat{H} 是哈密顿量, E_{α} 是特征值, V_{α} 是波函数. 在忽略三体和三体以上的相互作用下, 体系就成了两体问题, 哈密顿量可以表示为

$$\hat{H} = \hat{H}_p + \hat{H}_z + \hat{H}_{pz}, \quad (2)$$

其中 \hat{H}_p 代表质子部分的哈密顿量, \hat{H}_z 代表中子部分的哈密顿量, \hat{H}_{pz} 代表质子-中子相互作用的哈密顿量.

由于哈密顿量具有空间转动不变性的特点, 所以系统的总角动量 \mathbf{J} 和 Z 轴投影的分量 M 都是守恒量. 在计算过程中, 利用体系的波函数构造了一组任意的基矢: $|u\rangle$ 和 $|w\rangle$ (非正交归一基), 因为构造的基矢只有较好的 M 而没有较好的 \mathbf{J} , 所以这种构造基矢的方法是 M -scheme. 还有一种构造基矢的方法叫 \mathbf{J} -scheme, 一般情况下, 由于 M -scheme 比 \mathbf{J} -scheme 更方便, 所以我们选择 M -scheme. 在原子核壳模型结构下利用广义辛弱数截断法计算时, 为了保证得到比较精确的结果, 在计算的过程中要包含较多的价键组态, 因此得到的矩阵维数一般会很大. 虽然辛弱数截断法有其计算的优点, 但是一般情况下辛弱数截断法得到的矩阵要比广义辛弱数截断法得到的矩阵大, 而且矩阵的维数也不好控制, 或者处理起来也比较浪费时间, 所以选择广义辛弱数截断壳模型空间得到大型的哈密顿矩阵.

只计算中子部分的哈密顿量作用在子空间上得到重叠矩阵和哈密顿矩阵. 重叠矩阵 S 的矩阵元写为: $S_{u,w} = \langle u | w \rangle \neq \delta_{uw}$, 由定义知道 $S_{uw} = S_{wu}^* = (S^+)_{uw}$, 所以 S 是一个厄米矩阵 ($S = S^+$), 维数是 $n \times n$, 同时 S 也是一个大型的实对称正定的稀疏矩阵. 另外, 假设原子核壳模型哈密顿矩阵是 H (H 是一个厄米矩阵, $H = H^+$), H 是一个维数为 $n \times n$ 的矩阵, 它的矩阵元表示为: $H_{u,w} = \langle u | H | w \rangle$.

此时, 就要求解广义特征值了. 广义特征值问题就是找到能量特征值 E , 能量特征向量 $V \neq 0$, 即

$$HV = SVE. \quad (3)$$

变换为

$$V = H^{-1}SVE, \quad (4)$$

其中, H^{-1} 表示矩阵 H 的逆. 为了更好地解决广义特征值问题^[16], 大多转化为标准特征值问题, 其

他的方法如广义 Givens 等^[17]. 在本文中, 转化为标准特征值问题时用了 Cholesky 分解法^[18-20] 和 Elementary transformation^[21] 法.

计算内存较小的矩阵, 用 Matlab 命令 chol(\mathbf{S}) 来完成 Cholesky 分解, 把重叠矩阵分解为下三角矩阵 \mathbf{L}_s 和上三角矩阵 \mathbf{L}_s^T 的乘积. 然而, 当遇到超出计算机内存的矩阵时, 就把重叠矩阵按行或列分为多个文件, 然后按行或列把矩阵分解为一个下三角矩阵和一个上三角矩阵. 在分解正定矩阵的过程中, 用 Matlab 编程来实现. 事实证明, Cholesky 分解法在数值稳定性方面完全可以与 Gauss 消去法相媲美, 而且它的运算量却比 Gauss 消去法少一半, 同时也减少了计算中的许多比较和判断. 根据 Cholesky 分解法, 可以一行一行地求解出下三角矩阵 \mathbf{L}_s (转置后可求得上三角矩阵 \mathbf{L}_s^T), 然后一行一行存储在一个较小的文件中, 这样就克服了电脑内存不足的问题.

在求解内存较小矩阵的逆矩阵时, 用 Matlab 命令 inv(\mathbf{L}_s) 来求解出下三角矩阵和上三角矩阵的逆矩阵. 然而, 当遇到超出计算机内存的矩阵时, 把矩阵按行或列分为多个文件, 用 Elementary transformation 法求解出矩阵的逆矩阵. 根据 Elementary transformation 法, 可一行一行地求解出下三角矩阵的逆矩阵 (转置后得到上三角矩阵 \mathbf{L}_s^T 的逆矩阵), 然后一行一行存储在较小的文件中, 就克服了电脑内存不足的问题.

\mathbf{S} 矩阵是一个实对称的正定矩阵, 用 Cholesky 分解法可把 \mathbf{S} 矩阵分解为一个下三角矩阵 \mathbf{L}_s 和一个上三角矩阵 \mathbf{L}_s^T (\mathbf{L}_s^T 代表矩阵 \mathbf{L}_s 的转置) 的乘积, 即

$$\mathbf{S} = \mathbf{L}_s \mathbf{L}_s^T. \quad (5)$$

方程 (4) 经过代换并且两边同时左乘 $\mathbf{L}_s \mathbf{L}_s^T$ 变为

$$\mathbf{L}_s \mathbf{L}_s^T \mathbf{V} = \mathbf{L}_s \mathbf{L}_s^T \mathbf{H}^{-1} \mathbf{L}_s \mathbf{L}_s^T \mathbf{V} \mathbf{E}. \quad (6)$$

假设 $\mathbf{y} = \mathbf{L}_s^T \mathbf{V}$, 再利用 Elementary transformation 法, 经过变换, 方程 (6) 转化为

$$\mathbf{L}_s^{-1} \mathbf{H} \mathbf{L}_s^{-T} \mathbf{y} = \mathbf{y} \mathbf{E}, \quad (7)$$

其中, \mathbf{L}_s^{-1} 代表矩阵 \mathbf{L}_s 的逆, \mathbf{L}_s^{-T} 代表矩阵 \mathbf{L}_s^T 的逆. 假设 $\mathbf{A} = \mathbf{L}_s^{-1} \mathbf{H} \mathbf{L}_s^{-T}$,

$$\mathbf{A} \mathbf{y} = \mathbf{y} \mathbf{E}, \quad (8)$$

其中 \mathbf{y} 代表特征向量. 这是一个大型矩阵的标准特征值问题. 在第 3 部分, 用 Matlab 编程的重正交化

的 Lanczos 方法计算得到矩阵的特征值和特征向量 \mathbf{E} 和 \mathbf{y} , 从而可得原矩阵的特征值和特征向量 \mathbf{E} 和 $\mathbf{V} = \mathbf{L}_s^{-T} \mathbf{y}$. 在第 4 部分通过比较说明这种方法是有意义的.

3 Lanczos 方法

Lanczos 法具有两端的特征值最先收敛的特点, 通过 Lanczos 迭代得到一个三对角矩阵. 三对角矩阵的特征值与大型非正交归一基稀疏矩阵的部分特征值一样, 通过迭代最小的特征值先收敛, 其他的特征值则会聚集在最小的特征值周围, 继续迭代求解出其他的特征值和特征向量. 当得到最小的特征值和特征向量时, 要继续迭代得到次小的特征值和特征向量, 需要把最小的特征向量从模型空间中剔除, 正交化是一个好的方法. 利用 Gram-Schmidt 正交法^[22,23], 可得到任意向量的正交向量. 此外, 在求解三对角矩阵时, 要得到一组正交的向量, 也需要用到 Gram-Schmidt 正交法. 因此, 在 Lanczos 方法中正交化是不可缺少的一步, 如果不能保证正交性, 计算中就会失去正交的 Lanczos 向量, 就不能得到正确的特征值和特征向量, 例如: 如果正交化不完全, 计算得到的特征值会重复出现. 通过使用 Gram-Schmidt 正交法, 用 Matlab 编程计算后由 $\tilde{\mathbf{q}}_i$ 向量得到正交化的向量 \mathbf{q}_i , 数学推导公式如下.

第一个向量就是它本身,

$$\mathbf{q}_1 = \tilde{\mathbf{q}}_1. \quad (9)$$

经过迭代可以求出第二个向量 $\tilde{\mathbf{q}}_2$, 计算可得第二个正交化向量 \mathbf{q}_2 ,

$$\mathbf{q}_2 = \tilde{\mathbf{q}}_2 - \mathbf{q}_1 \frac{\langle \mathbf{q}_1 | \tilde{\mathbf{q}}_2 \rangle}{\langle \mathbf{q}_1 | \mathbf{q}_1 \rangle}. \quad (10)$$

经过迭代可以求出第三个向量 $\tilde{\mathbf{q}}_3$, 计算可得第三个正交化向量 \mathbf{q}_3 ,

$$\mathbf{q}_3 = \tilde{\mathbf{q}}_3 - \mathbf{q}_2 \frac{\langle \mathbf{q}_2 | \tilde{\mathbf{q}}_3 \rangle}{\langle \mathbf{q}_2 | \mathbf{q}_2 \rangle} - \mathbf{q}_1 \frac{\langle \mathbf{q}_1 | \tilde{\mathbf{q}}_3 \rangle}{\langle \mathbf{q}_1 | \mathbf{q}_1 \rangle}. \quad (11)$$

利用这样的推导法, 用 Matlab 编程计算, 得到所有的正交向量.

3.1 Lanczos 迭代

用 Lanczos 法计算时需要一个较大的内存空间来存储大型矩阵, 但是由于计算机内存的限制, 需要把矩阵按行 (列) 分为多个较小的文件, 然后

在计算时一个一个地读取文件, 并且释放暂时不用的变量, 这样就克服了计算机内存不足的缺点. 求解大型矩阵的特征值和特征向量的大多数方法就是把矩阵三对角化, 即找到一个正交矩阵: $Q_n = [q_1, q_2, \dots, q_n]$, 使得大型稀疏矩阵 (维数是 $n \times n$) 变换成三对角矩阵 B_n [24], 矩阵方程为

$$Q_n^T A Q_n = B_n, \quad (12)$$

其中, Q_n^T 代表矩阵 Q_n 的转置. 要想充分利用大型稀疏矩阵的稀疏性, 就要直接计算矩阵 B_n 和 Q_n 元素的方法来实现矩阵的三对角化.

因为 Q_n 是正交矩阵, 可知 $AQ_n = Q_n B_n$. 由矩阵方程的第 i 列可知,

$$Aq_i = \beta_{i-1}q_{i-1} + \alpha_i q_i + \beta_i q_{i+1}. \quad (13)$$

计算可得, α_i 和 β_i , 即三对角矩阵的对角元,

$$\begin{aligned} \alpha_i &= q_i^T Aq_i, \\ \beta_i &= q_{i+1}^T Aq_i = \| (Aq_i - \beta_{i-1}q_{i-1} - \alpha_i q_i) \|_2, \end{aligned} \quad (14)$$

其中, q_{i+1}^T 代表矩阵 q_{i+1} 的转置. 反过来, 对于任意给定的随机初始向量 $q_1 \in R^n$, $\|q_1\|_2 = 1$, 由以上公式递推可以得到 q_i, α_i 和 β_i ,

$$\begin{cases} \alpha_1 = q_1^T Aq_1, \\ r_i = (Aq_i - \beta_{i-1}q_{i-1} - \alpha_i q_i) (\beta_0 q_0 = 0), \\ \beta_i = \|r_i\|_2, \\ q_{i+1} = r_i / \beta_i \quad (\text{其中 } \beta_i \neq 0), \\ \alpha_{i+1} = q_{i+1}^T Aq_{i+1}, \quad i = 1, 2, \dots, n-1. \end{cases} \quad (15)$$

这就是 Lanczos 迭代.

现记: $Q_m = [q_1, q_2, \dots, q_m]$, 由 Q_m 可得三对角矩阵 B , 其中 Q_m 是正交向量.

$$B = \begin{pmatrix} \alpha_1 & \beta_1 & & & & \\ \beta_1 & \alpha_2 & \beta_2 & & & \\ & & \dots & \dots & & \\ & & & \beta_{m-2} & \alpha_{m-1} & \beta_{m-1} \\ & & & & \beta_{m-1} & \alpha_m \end{pmatrix}. \quad (16)$$

由文献 [24] 可知, (15) 式的矩阵形式为: $AQ_m = Q_m B + r_m e_m^T$, 而矩阵 B 就是要得到的维数较小的三对角矩阵.

A 是一个大型的实对称稀疏矩阵, 矩阵的维数是 $n \times n$, 由于电脑内存的限制, 所以把它按行 (列)

分为许多文件, $A = [A_1; A_2; \dots; A_g]$, 以便于电脑依次读取矩阵文件. 在编程计算中会得到一系列的非正交归一向量, 通过 Gram-Schmidt 正交法转化为正交归一向量, 再经过计算就可以得到三对角矩阵. 但是在求解三对角矩阵时, 有这样一个公式: $\alpha = q^T Aq$, 假设

$$v = Aq, \quad (17)$$

计算时需要注意: 尽管矩阵 A 已经按行 (列) 分为许多文件了, 但是在此处计算时仍然不能用 A 直接乘以 q , 因为矩阵 A 已经失去了稀疏性, 需要占很大的内存空间甚至计算机硬盘都不能存储了. 所以要把 A 用其他形式替换, 如 $Aq = L_s^{-1} H L_s^{-T} q$, 从右向左依次计算, 因为 q 是一个列向量, 每次计算后都是一个列向量, 这样就避免了由于电脑内存的限制而导致计算停止的问题.

经过编程计算, 得到对称三对角矩阵 B , B 的维数是 $m \times m$. 在一定条件下矩阵 B 的某些特征值是原矩阵 A 的某些特征值的很好的近似. 由文献 [24] 中的 Kaniel-Paige-Saad 理论可知, 随着 m 的增加, B 的两端的特征值会非常快地收敛到 A 的两端的特征值, 而我们恰好要计算矩阵最小的几个特征值. 在实际计算中也发现, 随着 m 的增加, 矩阵 B 的最小的几个特征值将会在迭代次数较少的情况下与矩阵 A 的最小的几个特征值达到很好的近似, 即特征值的逼近程度随着 m 的增加越来越好; 另外, 为了防止向量在迭代过程中失去正交性, 所以每次迭代的初始向量都要进行重正交化. 然而, m 并不是越大越好, 因为我们要求解的是大型对称稀疏矩阵的最小的几个特征值, m 太大, 读取大型矩阵的次数会更多, 延长计算时间; m 太小, 又不能得到所需正确特征值的个数. 所以合适地选取 m 的值, 得到对称三对角矩阵 B , 计算求得矩阵 B 的特征值, 再经过收敛迭代, 得到原矩阵的近似特征值. 下文中的算例也说明了这种方法的精确性和可靠性. 接下来我们对迭代程序进行简要叙述.

3.2 收敛迭代

B 是维数较小的对称三对角矩阵, 而求解三对角矩阵的特征值和特征向量是一个标准的特征值问题 [25], 所以运用 Matlab 命令, $[Y, E_m] = eig(B)$, 可得矩阵的特征值, 表达式为

$$BY = Y E_m, \quad (18)$$

其中 E_m 是矩阵 B 的特征值, Y 是特征向量, 经过多次迭代, 达到收敛精度, 迭代停止, 得到矩阵 A 的特征值 E 和特征向量 $y(y = QY)$, 如算法 1 所示, k 是想要得到的最小的几个特征值的个数, $1 < iter < 300$ 作为允许迭代的次数.

算法 1 (收敛迭代)

- 1) $kk = 1$;
- 2) $yn = q$ (q 是一个随机向量, 归一化后作为初始迭代向量);
- 3) 用 Lanczos 迭代法得到三对角矩 B ;
- 4) 用 Matlab 命令 $[Y, E_m] = eig(B)$ (得出最小的特征值 $E(1)$ 和特征向量 $Y(1)$);
- 5) $stavec = Q(:, 1 : m) * Y(1)$;
- 6) $yn(:, 1) = stavec$ (归一化的 $stavec$ 作为下一次迭代的初始向量);
- 7) $Err = E(1, iter) - E(1, iter - 1)$;
- 8) $Err < 10^{-10}$ (如果达到迭代停止的精度, 跳出循环, 继续下一个特征值的计算; 否则, 继续迭代);
- 9) $y_1 = Q(:, 1 : m) * Y(1)$ (矩阵最小的特征向量);
- 10) $EN = E_1$ (矩阵最小的特征值);
- 11) $yn(:, 1) = y_1$;
- 12) $yn = [y_1, q]$ (q 是一个新的随机向量, 与 Y_1 正交归一化后作为求解第二个特征值的初始向量);
- 13) for $kk = 2 : k$;
- 14) 重复程序 3)—8) (注意: $stavec$ 向量与前 $kk - 1$ 个特征向量正交归一化后作为下一次迭代的初始向量);
- 15) $y_{kk} = Q(:, 1 : m) * Y(kk)$ (得到 E_{kk} 和 Y_{kk});
- 16) $EN = [EN; E_{kk}]$;
- 17) $yn = [y_1, \dots, y_{kk}, q]$; (q 是新的随机向量, 与前 kk 个特征向量正交归一化后作为下一个特征值计算的初始向量);
- 18) end;
- 19) $VN = L_s^T * yn(:, 1 : k)$.

根据算法 1, 得到大型非正交归一基稀疏矩阵最小的 k 个广义特征值 EN 和广义特征向量 VN . 运用这个方法, 在一定程度上克服了电脑内存不足的问题, 而且在计算机内存限制的情况下计算出较为精确的特征值.

4 计算结果与讨论

利用重正交化 Lanczos 方法与 Cholesky 分解法和 Elementary transformation 法相结合的方法, 计算了其他原子核壳模型的价键组态维数是 12022×12022 的非正交归一基稀疏矩阵 (包含哈密顿矩阵和重叠矩阵) 的特征值和特征向量, 且与计算机得出的精确值进行了对比. 在表 1 中列出了最小的 10 个计算值和精确值. 因为计算机在计算时也是选择了一个随机的向量开始迭代的, 所以每次得到的特征值的精确度也不同, 经过分析和对比, 选择小数点后十位作为精确值.

表 1 最小的 10 个计算值和精确值

Table 1. The 10 smallest eigenvalues with calculated and exact.

Values	Exact values	Calculated values
1	24. 7434570668	24. 7434570667
2	26. 0686615937	26. 0686615938
3	28. 2398305237	28. 2398305237
4	28. 4920457030	28. 4920457030
5	28. 7017471744	28. 7017471744
6	28. 9312301581	28. 9312301581
7	29. 0616260322	29. 0616260322
8	29. 0959693096	29. 0959693096
9	29. 3497754458	29. 3497754458
10	29. 5307621113	29. 5307621113

由表 1 可以看出, 计算值和精确值仅仅在 10^{-10} 数量级上有差异, 对比结果说明了这种方法的精确性和可靠性. 既然用这种方法计算出的特征值比较准确, 因此就运用这种方法来计算大型的非正交归一基矩阵的特征值和特征向量.

在 ^{116}Sn 原子核壳模型下利用广义辛弱数截断多体空间得到的哈密顿矩阵的维数或者矩阵所占内存的大小是随着 s ($s=2, 4, 6, 8$; s 代表未配对的粒子数, 即广义辛弱数) 的增大而增大, 能量计算值的精确度也是随着 s 的增大而提高的, 另外在计算时可以看出, 基态能量的收敛比其他激发态的要快. 本文中计算的是 $s = 8$ 的广义辛弱数截断 ^{116}Sn 壳模型空间的哈密顿量的能量特征值问题, 截断得到的哈密顿矩阵的内存是 27.8 G, 维数是 646430×646430 . 如果计算中质核或者重核时, 矩阵的维数会很容易达到 10^8 或者更大, 超出了目前的计算能力, 所以寻找一种用内存较小的计算机

求解更大的非正交归一基矩阵的特征值的方法是有意义的. 本文中用内存仅为3.6 G的计算机, 通过运用重正交化 Lanczos 法与 Cholesky 分解法和 Elementary transformation 法相结合的方法, 得到了 ^{116}Sn 原子核壳模型空间截断下的内存为 27.8 G 的矩阵的部分能量特征值, 即 ^{116}Sn 原子核几个低态的能量. 表 2 列出了几个计算值与实验测量值的低态能量.

表 2 ^{116}Sn 几个最低能态的计算值和实验测量值
Table 2. The calculation and experiment value of ^{116}Sn for some smallest levels.

Levels	Experiment values	Calculated values	$ E_c - E_e $ MeV
	E_e/MeV	E_c/MeV	
1	0	0	0
2	1.293560	1.250522	0.043038
3	1.756864	1.685395	0.071469
4	2.027480	2.170830	0.143350
5	2.112323	2.341860	0.229537
6	2.225379	2.643817	0.418438
7	2.266159	2.645112	0.378953
8	2.365975	2.706631	0.340656
9	2.390879	2.708133	0.317254
10	2.529202	2.750437	0.221235

实验测量值来源于 NNDC 网站, 网址: <http://www.nndc.bnl.gov/chart/getdataset.jsp?nucleus=116SN>.

本文中计算了 ^{116}Sn (只有价中子活跃而质子不活跃, 且质子数是幻数的偶-偶核) 原子核的基态能量和最低的 9 个激发态的能量. 由表 2 可知, 计算值与实验测量值符合得尚好, 较高能态的计算值和实验测量值的差值是几百 keV, 而较低的几个能态与实验测量值只差几十 keV, 所以说这项工作是有意义的. 理论上的解释也能说明符合较好的原因, 因为那些较低能量的价键组态占了波函数大部分甚至全部的贡献, 所以当给出一个比较合理的截断之后舍去高能量部分并不影响波函数, 使得截断计算能够得出比较好的结果. 计算结果同时也证实了目前的广义辛弱数法重要的一点, 即当辛弱数 ($\text{Sn}: s = 8$) 较大时, 低能态会达到较好的收敛而重现. 另外, 在计算时发现, 基态的收敛比其他激发态收敛得要快.

广义辛弱数截断后是在子空间中计算的, 降低了矩阵的维数和复杂程度, 从而可以较容易地求解出原子核较低的能级, 而且得到的能级与实验测量

值也较吻合. 然而, 有时在较大的广义辛弱数截断下, 矩阵的维数仍然会非常大, 尤其对于那些质子和中子都活跃的原子核. 这时可能需要额外的截断或者限制未配对的粒子数. 对于给定了广义辛弱数 s 的情况下, 子空间的维数以及算法的复杂程度与价粒子数无关, 这与标准壳模型和辛弱数截断法形成鲜明的对比. 虽然现在只运用于偶-偶核的两体哈密顿量中, 但是有些理论推导和相关的计算也能很好地适用于奇粒子体系或者三体哈密顿量中.

原子核的能级是了解原子核结构和其他特性的前提. 本文的计算可得到相应的原子核的能级, 所以这项工作有助于更深入、透彻地了解原子核的结构. 同时, 也可以用此方法来求解一些中质核或者重核的能级.

5 结 论

本文将重正交化 Lanczos 方法与 Cholesky 分解法和 Elementary transformation 法相结合, 把广义特征值问题转化为标准特征值问题, 把大型的标准特征值问题转化为小型的三对角矩阵的标准特征值问题, 得到一种用内存较小的计算机求解大型实对称非正交归一基矩阵的特征值和特征向量的方法, 而且此方法具有较高的收敛性. 用这种方法计算得到小型矩阵的特征值与精确值相比较, 结果表明这种方法具有较高的精确性.

根据此方法, 求解了在 ^{116}Sn 原子核壳模型结构下利用广义辛弱数 ($s = 8$) 截断多体空间得到的哈密顿矩阵的能量特征值, 与实验测量能量值的比较结果说明了这一方法的有效性和可行性. 此外, 该方法不仅有助于求解某些中质核和重核的能态, 而且对于用内存稍大的计算机计算更大的非正交归一基矩阵的特征值都有实际意义.

作者非常感谢在上海理工大学理学院贾力源老师的帮助下得到重叠矩阵和哈密顿矩阵.

参考文献

- [1] Ring P, Schuck P 1980 *The Nuclear Many-body Problem* (Berlin: Springer-Verlag) pp36-95
- [2] Shen J J, Zhao Y M 2009 *Sci. China: Ser. G* **52** 1477
- [3] Shen J J, Arima A, Zhao Y M, Yoshinaga N 2008 *Phys. Rev. C* **78** 044305

- [4] Zhang L H, Shen J J, Lei Y, Zhao Y M 2008 *Int. J. Mod. Phys. E* **17** 342
- [5] Jia L Y 2013 *Phy. Rev. C* **88** 044303
- [6] Pittel S, Sandulescu N 2006 *Phys. Rev. C* **73** 014301
- [7] Thakur B, Pittel S, Sandulescu N 2008 *Phys. Rev. C* **78** 041303
- [8] Papenbrock T, Dean D J 2005 *J. Phys. G: Nucl. Part. Phys.* **31** S1377
- [9] Kruse M K G, Jurgenson E D, Navratil P, Barrett B R, Ormand W E 2013 *Phys. Rev. C* **87** 044301
- [10] Han H, Wu L Y, Song N N 2014 *Acta Phys. Sin.* **63** 138901 (in Chinese) [韩华, 吴翎燕, 宋宁宁 2014 物理学报 **63** 138901]
- [11] Li S, Wang B, Hu J Z 2003 *Appl. Math. Mech.* **24** 92
- [12] Morris N F 1990 *J. Struct. Eng.* **116** 2049
- [13] Jia L Y 2015 *J. Phys. G: Nucl. Part. Phys.* **42** 115105
- [14] Qi C, Xu Z X 2012 *Phys. Rev. C* **86** 044323
- [15] Simon H D 1984 *Math. Comput.* **42** 115
- [16] Zhao X H, Chen F W, Wu J, Zhou Q L 2008 *Acta Phys. Chim. Sin.* **24** 823 (in Chinese) [赵小红, 陈飞武, 吴健, 周巧龙 2008 物理化学学报 **24** 823]
- [17] Cao Z H 1980 *Eigenvalue Problems of Matrices* (Shanghai: Shanghai Scientific and Technical Publishers) pp212–220 (in Chinese) [曹志浩 1980 矩阵特征值问题 (上海: 上海科学技术出版社) 第212—220页]
- [18] Harbrecht H, Peters M, Schneider R 2012 *Appl. Numer. Math.* **62** 428
- [19] D’azevedo E, Dongarra J 2000 *Pract. Exper.* **12** 1481
- [20] Schweizer S, Kussmann J, Doser B, Ochsenfeld C 2008 *J. Comput. Chem.* **29** 1004
- [21] Liu D, Gabrielli L H, Lipson M, Johnson S G 2013 *Opt. Exp.* **21** 12
- [22] Giraud L, Langou J 2005 *Comput. Math. Appl.* **50** 1069
- [23] Hoffmanm W, Amsterdam 1989 *Computing* **41** 335
- [24] Xu S F 1995 *Theory and Method of Matrix Calculation* (Beijing: Peking University Publishers) pp307–319 (in Chinese) [徐树方 1995 矩阵计算的理论与方法 (北京: 北京大学出版社) 第307—319页]
- [25] Qiu Z, Wang X 2005 *J. Sound Vib.* **282** 381

Eigenvalue problems solved by reorthogonalization Lanczos method for the large non-orthonormal sparse matrix

Jiao Bao-Bao[†]

(Department of Physics, University of Shanghai for Science and Technology, Shanghai 200093, China)

(Received 27 April 2016; revised manuscript received 5 July 2016)

Abstract

Using shell model to calculate the nuclear systems in a large model space is an important method in the field of nuclear physics. On the basis of the nuclear shell model, a large symmetric non-orthonormal sparse Hamiltonian matrix is generated when adopting the generalized seniority method to truncate the many-body space. Calculating the energy eigenvalues and energy eigenvectors of the large symmetric non-orthonormal sparse Hamiltonian matrix is of indispensable steps before energies of nucleus are further calculated. In the mean time, some low-lying energy eigenvalues are always the focus of attention on the occasion of large scale shell model calculation. In this paper, by combining reorthogonalization Lanczos method with Cholesky decomposition method and Elementary transformation method, converting the generalized eigenvalue problems into the standard eigenvalue problems, and transforming the large standard eigenvalue problems into the small standard eigenvalue problems, we successfully calculate the eigenvalues and eigenvectors of large non-orthonormal sparse matrices with the help of computers with limited memory. The values obtained by using this method to calculate the small matrix agree with the exact values, which demonstrates that this method is accurate and can be used to calculate the energy eigenvalues and energy eigenvectors of large symmetric non-orthonormal sparse matrix. We take ^{116}Sn ($s = 8$, the number of unpaired particles, namely the generalized seniority) as an example in which there are active valence neutrons but inert protons at the magic number, and calculate ten of its lowest energy eigenvalues. Through calculation, we find that among these low-lying energy eigenvalues, the lowest energy eigenvalue converges fastest. A comparison between the calculation values and the experiment values shows that the difference between the calculated high-lying energy eigenvalue and its corresponding experimental one arrives at hundreds of keV, while for the low-lying energy eigenvalue, its calculation value can reach an accuracy of a few tens of keV. The results demonstrate that the Lanczos method is feasible in Matlab programming and shell model calculations. The significance of this research lies in the fact that this method will not only greatly help to calculate and obtain the low-lying energy eigenvalues of some medium-mass and heavy nuclei, but also possess great importance in calculating partial eigenvalues involved in large matrices in other theoretical researches and engineering designs.

Keywords: nuclear shell model, eigenvalues, Hamiltonian

PACS: 21.60.Cs, 27.60.+j

DOI: 10.7498/aps.65.192101

[†] Corresponding author. E-mail: baobaojiao91@126.com