

УДК 539.192

МНОГОШАГОВЫЕ МЕТОДЫ ПЕРВОГО ПОРЯДКА В РЕШЕНИИ УРАВНЕНИЙ ТЕОРИИ СВЯЗАННЫХ КЛАСТЕРОВ**М.И. Бердник, В.В. Иванов**

Приведены результаты, описывающие особенности итерационного расчета волновой функции метода связанных кластеров с использованием оптимизационных подходов, основанных на информации о приближенных решениях предыдущих итераций. Рассмотрен метод прямого обращения в итерационном подпространстве (в английской аббревиатуре – DIIS), метод «тяжелого шарика» и алгоритмы, основанные на подходе Ю. Е. Нестерова. На примере полуэмпирических расчетов π -сопряженных систем продемонстрирована эффективность методов DIIS и «тяжелого шарика».

Ключевые слова: теория связанных кластеров, π -сопряженные системы.

Введение

Теория связанных кластеров (Coupled Cluster, CC) [1,2], со времени введения в квантовую химию в 1970 годах, привлекала внимание исследователей в связи с высокой точностью описания разнообразных молекулярных свойств. В последние годы, благодаря быстрому росту вычислительных мощностей, методы CC становятся все более популярными в исследовании систем и состояний, где значительны эффекты электронной корреляции (квазивырожденные состояния, высокосимметричные и сопряженные системы и *т.д.*). Основной вычислительной проблемой метода является необходимость итерационного решения нелинейных уравнений CC. Значительные вычислительные затраты при решении этих уравнений являются серьезным препятствием для широкого использования метода. В связи с этим представляет интерес сравнительная апробация различных градиентных методов нахождения волновой функции метода CC. В данной статье рассматривается несколько многошаговых методов первого порядка, *т.е.* методов, использующих информацию о приближенных решениях, полученных на предыдущих итерациях, не использующие, при этом, вычисления вторых производных. Последнее представляет собой определенные трудности в теории CC.

В данной работе, реализован полуэмпирический π -электронный метод CC с точным учетом лишь двукратных возбуждений (CCD). Следует отметить, что выбор полуэмпирической схемы не вносит существенных различий по сравнению с неэмпирическим (*ab initio*) подходом. Но π -электронные расчеты могут оказаться перспективными в проблеме исследования электронной структуры больших систем включающих сотни и даже тысячи тяжелых атомов. Сам метод CCD интересен тем, что является простейшим кластерным подходом эффективно учитывающим четырехкратные возбуждения. В последнее время интенсивно обсуждаются возможности подходов на основе CCD в описании типичных мультиреференсных проблем (квазивырожденные состояния, статические корреляции, и *т.д.*, см. например [3,4,5]).

Градиентные алгоритмы решения нелинейных уравнений CCD

Волновая функция метода CCD, в общем виде, может быть представлена следующим образом:

$$|\Psi_{\text{CCD}}\rangle = \exp(T_2)|0\rangle = \left(1 + T_2 + \frac{1}{2}T_2^2 + \dots\right)|0\rangle. \quad (1)$$

В этом выражении $|0\rangle$ означает референсный хартри-фоковский детерминант, а кластерный оператор T_2 , генерирующий суперпозицию двукратно-возбужденных конфигураций относительно $|0\rangle$, выражается через матрицу амплитуд t :

$$T_2 = \sum t_J |J\rangle. \quad (2)$$

Решения уравнений теории CCD сводятся к вычислению термов, которые включают линейную и нелинейную компоненты. С формальной точки зрения, эти уравнения, для t_J , могут быть найдены исходя из выражения:

$$\Delta_J(t) = \langle J | H-E | \Psi_{\text{CCD}} \rangle = \left(A_J + \sum_I B_{JI} t_I + \frac{1}{2} \sum_{L,K} C_{JKL} t_K t_L \right) / \Delta \varepsilon_J = 0, \quad (3)$$

где набор $\{A, B, C\}$ – включает одно- и двухэлектронные интегралы, а $\Delta \varepsilon_J$ – изменение одноэлектронных энергий МО которые соответствуют переходу $|0\rangle \rightarrow |J\rangle$. Детальные выражения для (3) см. например в [6].

В процессе итерационного решения уравнений (3) приходится многократно вычислять амплитудную матрицу t и матрицу Δ . На текущей, k -той, итерации такие матрицы мы будем обозначать как $t^{(k)}$ и $\Delta^{(k)} = \Delta(t^{(k)})$. Стандартный градиентный метод (**SGM**) [7] нахождения амплитудной матрицы, t , описывается как итерационный процесс:

$$t^{(k+1)} = t^{(k)} - \alpha^{(k)} \Delta^{(k)}, \quad (4)$$

остановка которого происходит при условии $|\Delta^{(k)}| \leq 10^{-10}$. В наших расчетах величина $|\Delta^{(k)}|$ соответствует евклидовой норме $\Delta^{(k)}$. В простейшем варианте **SGM**, который использовался в настоящей работе, параметр $\alpha^{(k)} = \alpha = \text{const}$. Отметим, что оптимизация величины $\alpha^{(k)}$ на каждом, k -том шаге, что соответствует методу наискорейшего спуска (Steepest Descent), представляет собой значительные дополнительные затраты в случае теории CC, поскольку требует вычисление вектора $\Delta(\Delta^{(k)})$.

Схема интерполяции **DIIS** (Direct Inverse in Iterative Subspace), применительно к задаче CCD, заключается в построении “подправленной” амплитудной матрицы t , которая на ℓ -той итерации выражается через аналогичные матрицы, полученные на ℓ -той и на предыдущих m итерациях:

$$t^{(\ell)} \leftarrow c_0 t^{(\ell)} + c_1 t^{(\ell-1)} + c_2 t^{(\ell-2)} + \dots + c_m t^{(\ell-m)}. \quad (5)$$

При этом, интерполяционные коэффициенты c_0, c_1, \dots, c_m минимизируют $|\Delta^{(\ell)}|$ с учетом условия

$$\sum_{i=0}^m c_i = 1. \quad (6)$$

Детальное описание метода **DIIS** можно найти в [8,9]. **DIIS** применительно к теории CC см. [10,11].

Метод «тяжелого шарика» (Heavy Ball, **Hb**) [7] интуитивно понятен исходя из механической аналогии. Процесс оптимизации, в этом случае, может быть интерпретирован как движение шарика по гиперповерхности к точке минимума. При этом, типичной, является ситуация, когда этот шарик, двигаясь по узкому желобу, периодически «ударяется» о стенки, что соответствует увеличению числа итераций. Для того чтобы избежать такие ситуации в методе **Hb**, к уравнениям **SGM** (4), добавляется «моментная» компонента:

$$t^{(k+1)} = t^{(k)} - \alpha \Delta^{(k)} + \beta (t^{(k)} - t^{(k-1)}). \quad (7)$$

Здесь фактор $t^{(k)} - t^{(k-1)}$ как бы подталкивает шарик в направлении предыдущего шага.

Идеология Ю. Е. Нестерова [12] (**NEST**) порождает группу подходов. В последнее время они интенсивно обсуждаются в научной литературе, см. [12,13]. В настоящей работе мы использовали два наиболее распространенных алгоритма [14].

Алгоритм **NEST1**: $y^{(0)} = t^{(0)}$, $\theta_0 \in (0,1)$,

for $k = 0, 1, \dots$, do

$$t^{(k+1)} = y^{(k)} - \alpha \Delta(y^{(k)});$$

$$\theta_{k+1} \in (0,1) \text{ from } \theta_{k+1}^2 = (1 - \theta_{k+1}) \theta_k^2 + q \theta_{k+1}; \beta_{k+1} = \theta_k (1 - \theta_k) / (\theta_k^2 + \theta_{k+1}); \quad (8)$$

$$y^{(k+1)} = t^{(k+1)} + \beta_{k+1} (t^{(k+1)} - t^{(k)});$$

end

Алгоритм **NEST2** соответствует неизменному, в процессе итерирования, параметру $\beta_k = \beta$:

$$\begin{aligned}
 & y^{(0)} = t^{(0)}; \\
 & \text{for } k = 0, 1, \dots, \text{do} \\
 & \quad t^{(k+1)} = y^{(k)} - \alpha \Delta(y^{(k)}); \\
 & \quad y^{(k+1)} = t^{(k+1)} + \beta(t^{(k+1)} - t^{(k)}); \\
 & \text{end}
 \end{aligned} \tag{9}$$

Расчетная часть

Для изучения сходимости методов многошаговой минимизации первого порядка мы использовали полуэмпирический π -электронный (гамильтониана Попла-Паризера-Парра, ППП, [15,16]) вариант теории CCD. Рассматривались две π -системы. Линейный полиен дека-1,3,5,7,9-пентаен, $C_{10}H_{12}$, и модельная планарная циклическая молекула [14]аннулен, $C_{14}H_{14}$.

В так называемой «мягкой» параметризации гамильтониана ППП, стандартный резонансный интеграл пары связанных углеродных атомов равен $b_0 = -2.274$ эВ, одноцентровый кулоновский интеграл $\Gamma_0 = 11.13$ эВ. Двухцентровые кулоновские интегралы оценивались по известной формуле Оно [17]. В расчетах использовалась идеализированная геометрия: длины всех $-C-C-$ связей равны 1.4 \AA , углы в линейном *trans*-полиене – 120° . Циклополиен $C_{14}H_{14}$ описывался как правильный многоугольник. При расчете $C_{10}H_{12}$ вводилось альтернирование резонансных интегралов для двойных (+) и одинарных (–) связей: $b_{\pm} = b_0(1 \pm 0.1)$. В расчетах циклической системы $C_{14}H_{14}$ резонансный интеграл уменьшался от стандартного b_0 , до величины -1.3 эВ. В этом случае сходимость итерационного процесса обычных методов, ухудшается в следствии квазивырождения молекулярных орбиталей.

В методе **DHS** скорость нахождения решения определяется выбором параметра α и числом шагов интерполяции $n_{DHS} = m + 1$, см. (5). В табл. 1 и 2 показано, как изменяется скорость сходимости (число итераций) до требуемой точности, $|\Delta| \sim 10^{-10}$, при изменении параметра α , для метода **DHS** с различной размерностью вектора (5), n_{DHS} .

Приведенные таблицы говорят об относительной независимости кол-ва итераций от выбора параметра α . Такая независимость становится заметнее при увеличении n_{DHS} . Так, при $n_{DHS} = 8$ α практически не влияет на число итераций. Это обстоятельство является ценным свойством **DHS**.

При переходе ко второй системе, $C_{14}H_{14}$, с разрушенными связями, $b = -1.4$ эВ, сходимость при малых n_{DHS} заметно ухудшается, что вполне характерно для квазивырожденных ситуаций. Однако, при $n_{DHS} = 8$, характеристики сходимости для двух систем идентичны!

Таблица 1. Количество итерационных шагов метода **DHS** для достижения сходимости по критерию $|\Delta| \sim 10^{-10}$ при различных выборах параметра α . Система $C_{10}H_{12}$.

α	$n_{DHS} = 2$	$n_{DHS} = 3$	$n_{DHS} = 4$	$n_{DHS} = 8$
0.7	69	48	47	32
0.8	61	44	40	31
0.9	52	52	45	32
1.0	62	54	42	32
1.1	67	51	46	33
1.2	67	49	46	33
1.3	66	50	46	34

Таблица 2. Количество итерационных шагов метода **DIIS** для достижения сходимости по критерию $|\Delta| \sim 10^{-10}$ при различных выборах параметра α . Система $C_{14}H_{14}$ ($b = -1.4$ эВ).

α	$n_{DIIS} = 2$	$n_{DIIS} = 3$	$n_{DIIS} = 4$	$n_{DIIS} = 8$
0.7	118	100	63	36
0.8	104	108	54	32
0.9	92	72	75	32
1.0	136	76	73	32
1.1	187	76	59	34

К сожалению, при работе с большими системами, сохранение матриц в оперативной памяти ЭВМ становится невозможным, а затраты на считывание и перезапись на жесткий диск будут существенно сказываться на общем времени счета. Поэтому практический интерес представляют методы, удерживающие как можно меньше данных в оперативной памяти. Таковым является метод **Hb**, сходимость которого, определяется выбором двух параметров. Параметр α задает величину шага в направлении градиента, а β – релаксационный параметра см. (7).

Результаты расчета двух π -систем методом **Hb** сведены в табл. 3 и в табл. 4. Из приведенных данных видно, что введение небольшой величины для «релаксационного» параметра, скорость сходимости может улучшиться даже при значительной величине α (> 1.3 для линейного полиена). Очевидно, что ситуации, когда $\beta = 0$, соответствует **SGM**. Введение релаксационного параметра стабилизирует сходимость итерационной процедуры, позволяя ей оставаться постоянной в широком интервале изменений α . Это большое преимущество метода **Hb** по сравнению с методом Нестерова, который, как будет показано ниже, расходится при неудачном выборе параметров итерационного процесса. Следует отметить, однако, что число итераций существенно отличается для $C_{10}H_{12}$ и для $C_{14}H_{14}$ (при $b = -1.4$ эВ) в методе **Hb** в отличие от **DIIS** с $n_{DIIS} = 8$. Тем не менее, для $C_{14}H_{14}$ ($b = -1.4$ эВ), метод **Hb** гарантирует нахождение CCD решения за 74-75 итераций.

Таблица 3. Количество итерационных шагов метода **Hb** для достижения сходимости при различных выборах параметра α и β . Система $C_{10}H_{12}$.

α	β				
	0	0.1	0.2	0.3	0.4
1.6	нс*	нс	нс	57	54
1.5	нс	нс	107	42	54
1.4	нс	206	37	42	55
1.3	627	62	33	42	55
1.2	92	56	39	43	55
1	84	71	55	43	56
0.9	94	81	65	43	56
0.8	108	93	76	53	56

* нс – нет сходимости итерационной процедуры.

Расчеты в рамках метода Нестерова (алгоритмы **NEST1** и **NEST2**) проявили значительную зависимость от значений параметров, вплоть до полного отсутствия сходимости при небольшом варьировании параметров. В табл. 5 и 6 продемонстрированы изменения скорости сходимости в методах Нестерова при изменении параметров итерационной процедуры. Представленные данные говорят о значительном разбросе числа итераций при разных значениях «моментных» параметров (q для **NEST1** и β для **NEST2**).

Задача нахождения решения в методе CCD чрезвычайно чувствительна к степени квазивырождения. Для циклической системы $C_{14}H_{14}$, при достаточно маленьких значениях резонансных интегралов ($|b| < 1.3$ эВ) CCD решение вообще не удалось получить. При значениях резонансного интеграла ($|b| \sim 1.3$ эВ) все методы примерно с одинаковой скоростью сходились до точ-

ности $|\Delta| \sim 10^{-8}$. При дальнейшем проведении итерационной процедуры методы **Hb** и **NEST** расходились, но метод **DHS** ($n_{\text{DHS}} = 2$) с заметным уменьшением скорости сходимости и с появлением осцилляций сошелся к искомому решению. При этом в **DHS** 88 итераций было потрачено на достижение точности $|\Delta| \sim 10^{-8}$ и еще 184 на достижение точности $|\Delta| \sim 10^{-10}$. При решении данной задачи методом **DHS** с большей размерностью эта проблема исчезала. Так, в случае $n_{\text{DHS}} = 8$ метод эффективно сходится к точности $|\Delta| \sim 10^{-10}$ всего за 35 итераций!

Таблица 4. Количество итерационных шагов метода **Hb** для достижения сходимости при различных выборах параметра α и β . Система $C_{14}H_{14}$ ($b = -1.4$ эВ).

α	β			
	0.3	0.4	0.5	0.6
1	195	156	105	98
1.1	175	138	83	99
1.3	144	110	74	99
1.4	132	98	74	100
1.5	121	88	74	99
1.6	111	78	75	100
1.7	103	67	75	102
1.8	261	58	76	101
1.9	нс	86	76	104

Таблица 5. Количество итерационных шагов метода **NEST1** для достижения сходимости по критерию $|\Delta^{(k)}| \sim 10^{-10}$ при различных выборах параметра α и q . Система $C_{14}H_{14}$ ($b = -1.4$ эВ).

α	q			
	0.04	0.05	0.06	0.1
0.6	110	134	161	222
0.7	102	95	127	186
0.8	102	97	90	158
0.9	89	87	87	136
1	701	419	308	165

Таблица 6. Количество итерационных шагов метода **NEST2** для достижения сходимости по критерию $|\Delta^{(k)}| \sim 10^{-10}$ при различных выборах параметра α и β . Система $C_{14}H_{14}$ ($b = -1.4$ эВ).

α	β							
	0.4	0.5	0.55	0.57	0.58	0.59	0.6	0.62
0.8	215	167	140	127	119	111	99	92
0.9	188	145	118	104	93	84	86	89
1	203	152	203	234	253	274	300	367

В целом, сравнение различных итерационных методов на примере расчета полиена $C_{10}H_{12}$ представлено графически на Рис. 1.

Из приведенных зависимостей можно заключить, что сходимость итерационных процедур **NEST2** (зависимость метода **NEST1**, при оптимально подобранных параметрах, практически идентична получаемой в методе **NEST2**) и **Hb**, с оптимально подобранными параметрами, в случае расчета «хороших», невырожденных систем, оказывается даже выше, чем для **DHS** ($n_{\text{DHS}} = 2$). Тем не менее, **DHS** оказался наиболее надежным методом расчета, гарантирующим решение задачи даже в трудных ситуациях. Отметим, однако, и тот факт, что для метода **DHS** $n_{\text{DHS}} > 2$ требуется значительный объем операций ввода-вывода, множество вычислений скалярных произведений векторов большой размерности и обращение матрицы. Последняя опера-

ция может быть чревата дополнительными проблемами, связанными с вырожденностью матрицы, в ситуации, когда задача близка к решению. В этом случае приходится применять псевдообращение.

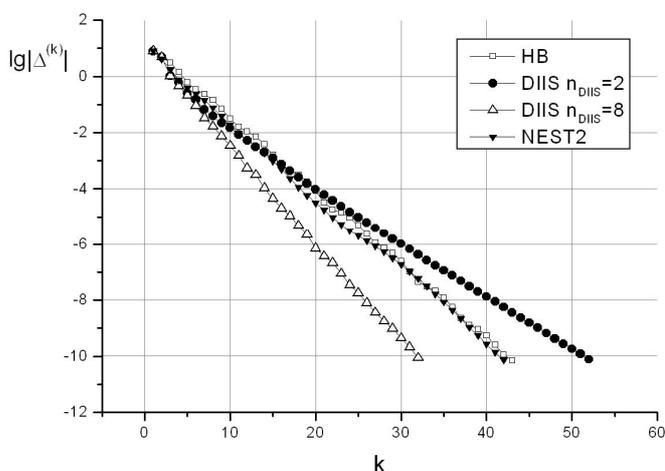


Рисунок 1. Зависимость $\lg|\Delta^{(k)}|$ от номера итерации (полиен $C_{10}H_{12}$)

Заключение

В одной из лучших, в методическом отношении, современных книг по теории оптимизации автор (Ю. Е. Нестеров) писал «Здоровый скептицизм должен привести нас к следующей догадке: Задачи нелинейной оптимизации, в их самой общей форме, являются *численно неразрешимыми*» [14]*. Нельзя не согласиться с этим. Тем не менее, материалы настоящей заметки, позволяют сделать скорее оптимистическое заключение для конкретных нелинейных задач. Даже в достаточно сложных ситуациях, в присутствии квазивырождения, искомые СС-решения могут быть успешно найдены. Метод **DIIS** демонстрирует превосходные характеристики сходимости даже в варианте с малой размерностью вектора интерполяционного разложения (5), при умеренных вычислительных затратах. Метод **Hb**, являясь незначительной модификацией простейшего градиентного алгоритма, также проявляет эффективность при расчете квазивырожденных систем. Необходимо отметить и перспективу применения **Hb** в построении алгоритма **DIIS**.

Литература

1. Coester F., Kümmel H. // Nucl. Phys. – 1960. – V. 17. – P. 477-485.
2. Kümmel H. // Theor. Chim. Acta. – 1991. – V. 80. – P. 81-89.
3. Ivanov V. V. Adamowicz L. // J. Chem. Phys. – 2000. – V. 112, №21. – P. 9258-9268.
4. Henderson T. M., Bulik I. W., Stein T., Scuseria G. E. // J. Chem. Phys. – 2014. – V. 141. – P. 244104.
5. Bulik I. W., Henderson T. M., Scuseria G. E. // J. Chem. Theory Comput. – 2015. – V. 11. – P. 3171-3179.
6. Bartlett R. J. // Ann. Rev. Phys. Chem. – 1981. – V. 32. – P. 359-401.
7. Поляк Б. Т. Введение в оптимизацию. М.: Наука, 1983. – 384 с.
8. Pulay P. // J. Comput. Chem. – 1982. – V. 3. – P. 556-560.
9. Cancés E., Le Bris C. // Int. J. Quant. Chem. – 2000. – V. – 79. – P. 82-90.
10. Scuseria G. E., Lee T. J., Schaefer III H. // Chem. Phys. Letter. – 1986. – V. 130, № 3. – P. 236-239.
11. Иванов В. В. // Вісник Харківського національного університету. 2002. – № 549. Хімія, Вип. 8 (31). – С.21-23.
12. Нестеров Ю. Е. // Ж. вычисл. матем. и матем. физ. – 1984. – Т. 24, № 7. – С. 1090-1093.

* Цитируется по русскоязычному изданию.

13. Nesterov Yu. E., Nemirovskii A. S. Interior-Point Polynomial Algorithms in Convex Programming. SIAM, Philadelphia, 1994. – 405 p.
14. Nesterov Yu. E. Introductory lectures on convex optimization: a basic course New York: Springer science+Business media.–2004.–235 p.
15. Pariser R., Parr R. G. // J. Chem. Phys. – 1953. – V. 21. – P. 466-471.
16. Pople J. A. // Trans. Faraday Soc. – 1953. – V.49. – P. 1375-1385.
17. Onho K. // Theor. Chim. Acta. – 1964. – V. 2. – P. 219-227.

References

1. Coester F., Kümmel H. // Nucl. Phys. – 1960. – V. 17. – P. 477-485.
2. Kümmel H. // Theor. Chim. Acta. –1991. – V. 80. – P.81-89.
3. Ivanov V. V. Adamowicz L. // J. Chem. Phys. – 2000. – V. 112, №21. – P. 9258-9268.
4. Henderson T. M., Bulik I. W., Stein T., Scuseria G. E. // J. Chem. Phys. – 2014. – V. 141. – P. 244104.
5. Bulik I. W., Henderson T. M., Scuseria G. E. // J. Chem. Theory Comput. – 2015. – V. 11. – P. 3171-3179.
6. Bartlett R. J. // Ann. Rev. Phys. Chem. – 1981. – V. 32. – P. 359-401.
7. Polyak B. T. Vvedenie v optimizaciyu. M.: Nauka, 1983. – 384 p. (in Russian)
8. Pulay P. // J. Comput. Chem. – 1982. – V.3. – P. 556-560.
9. Cancés E., Le Bris C. // Int. J. Quant. Chem. – 2000. – V. – 79. – P. 82-90.
10. Scuseria G. E., Lee T. J., Schaefer III H. // Chem. Phys. Letter. – 1986. – V. 130, № 3. – P. 236-239.
11. Ivanov V. V. // Visn. Hark. Nac. Univ. 2002. – № 549. Ser. Him. Issue 8 (31). – P. 21-23. (in Russian)
12. Nesterov Yu. E. // J. Vy'chisl. Matem. i Matem. Fiz. – 1984. – V. 24, № 7. – P. 1090-1093. (in Russian)
13. Nesterov Yu. E., Nemirovskii A. S. Interior-Point Polynomial Algorithms in Convex Programming. SIAM, Philadelphia, 1994. – 405 p.
14. Nesterov Yu. E. Introductory lectures on convex optimization: a basic course // New York: Springer science+Business media. – 2004. – 235 p.
15. Pariser R., Parr R. G. // J. Chem. Phys. – 1953. – V. 21. – P. 466-471.
16. Pople J. A. // Trans. Faraday Soc. – 1953. – 49. – P. 1375-1385.
17. Onho K. // Theor. Chim. Acta. – 1964. – V. 2. – P. 219-227.

Поступила до редакції 27 серпня 2015 г.

М.І. Бердник, В.В. Иванов. Багатокрокові методи першого порядку в розв'язку рівнянь теорії зв'язаних кластерів.

Наведено результати, що описують особливості ітераційного розрахунку хвильової функції методу зв'язаних кластерів з використанням оптимізаційних підходів, заснованих на інформації про наближені розв'язки попередніх ітерацій. Розглянуто метод прямого обернення ітераційного підпростору (в англійській аббревіатурі – DIIS), метод важкої кульки і алгоритми, засновані на підході Ю. А. Нестерова. На прикладі напівемпіричних розрахунків π -спряжених систем продемонстровано ефективність методів DIIS і важкої кульки.

Ключові слова: теорія зв'язаних кластерів, π -спряжені системи.

M. I. Berdnyk, V. V. Ivanov. Solution of coupled cluster equations with usage of first order multistep methods.

Results describing features of iterative calculation of coupled cluster method wave function with usage of optimization approaches based on information stored from approximate solutions obtained from previous iterations are presented. The direct inversion in the iterative subspace technique (DIIS), heavy ball method, and techniques based on Yu. E. Nesterov approach were considered. As an example the effectiveness of DIIS and heavy ball methods in semi-empirical calculations of π -conjugated systems was described.

Key words: coupled cluster theory, π -conjugated systems.

Kharkov University Bulletin. Chemical Series. Issue 25 (48), 2015