

Modyfikacja schematu SCPF obliczeń energii polaryzacji

Grzegorz Mazur

Zakład Metod Obliczeniowych Chemii



11 kwietnia 2006 roku

- 1 Energia polaryzacji
 - Po co?
 - Jak?
- 2 Klasyczny schemat SCPF
 - Algorytm
 - Analiza zbieżności
- 3 Proponowana modyfikacja
 - Czy da się to zrobić lepiej?
 - Metoda najszybszego spadku
 - Metoda sprzężonych residuów
 - Symetryzacja macierzy współczynników
 - Metoda sprzężonych residuów dla układów symetrycznych
 - Koszt obliczeniowy
- 4 Wyniki
 - Fulleren
- 5 Podsumowanie

Energia stanów z przeniesieniem ładunku

Energia stanów CT może być oszacowana przy pomocy wzoru Mullikena

$$\mathcal{E}_{CT} = I - A + C + P \quad (1)$$

gdzie

- $I - A$ różnica pomiędzy potencjałem jonizacji i powinowactwem elektronowym
- C energia oddziaływania kulombowskiego pary ładunków
- P energia polaryzacji kryształu przez parę ładunków (liczona względem energii polaryzacji przez nieoddziałujące ładunki)

Energia polaryzacji

Energia polaryzacji ośrodka

$$W = -\frac{1}{2} \sum_n E_n \mu_n \quad (2)$$

gdzie

μ_n wygenerowane momenty dipolowe

E_n pole zewnętrzne

Pole ładunków

Pole generowane przez układ ładunków w punkcie r_n wynosi

$$E_n = \sum_m U_{nm} q_m \quad (3)$$

gdzie q_m oznacza ładunek a tensor monopolowy U_{nm} dany jest przez

$$U_{nm} = \frac{1}{4\pi\epsilon_0} \nabla |r_n - r|_{r=r_m}^{-1} \quad (4)$$

Momenty dipolowe

Generowany moment dipolowy opisany jest przez

$$\mu_n = \alpha_n F_n \quad (5)$$

gdzie pole lokalne w punkcie r_n wynosi

$$F_n = E_n + \sum_{n' \neq n} T_{nn'} \mu_{n'} \quad (6)$$

a tensor dipolowy $T_{nn'}$ dany jest przez

$$T_{nn'} = \frac{1}{4\pi\epsilon_0} \nabla \nabla |r_n - r_{r=r'_n}|^{-1} \quad (7)$$

Sprzężenie

- Zbierając to wszystko razem dostajemy układ sprzężonych równań

$$\mu_n = \alpha_n E_n + \alpha_n \sum_{n' \neq n} T_{nn'} \mu_{n'} \quad (8)$$

- Możliwe podejścia
 - Jawne rozprzęgnięcie równań przez transformatę Fouriera (tylko dla układów periodycznych)
 - Zastosowanie metody typu samouzgodnionego pola

Szkic algorytmu

- Zaczynamy od jakichś (np zerowych) próbnych wartości momentów dipolowych
- Generujemy nowe momenty dipolowe zgodnie z

$$\mu_n^{(i)} = \alpha_n E_n + \alpha_n \sum_{n' \neq n} T_{nn'} \mu_{n'}^{(i-1)} \quad (9)$$

- Problem jest rozwiązany jeżeli

$$\mu_n^{(i+1)} \sim \mu_n^{(i)} \quad (10)$$

Wektor błędu

Mamy metodę iteracyjną

$$x_{(k)} = Bx_{(k-1)} + c \quad (11)$$

Wektor błędu

$$e_{(k)} = x_{(k)} - \hat{x} \quad (12)$$

mówi na ile i -ta iteracja jest różna od (dokładnego) rozwiązania

Ewolucja wektora błędu

$$e_{(k+1)} = x_{(k+1)} - \hat{x} = Bx_{(k)} + c - \hat{x} = Be_{(k)} = B^i e_{(0)} \quad (13)$$

Warunek zbieżności

Analiza spektralna

Oznaczając przez α_k współczynniki rozwinięcia $e_{(0)}$ w bazie wektorów własnych macierzy B otrzymujemy

$$e_{(k+1)} = B^i \sum_k \alpha_k v_k = \sum_k \lambda_k^i \alpha_k v_k \quad (14)$$

Metoda jest zbieżna jeżeli

$$\max_k |\lambda_k| < 1 \quad (15)$$

Reality check

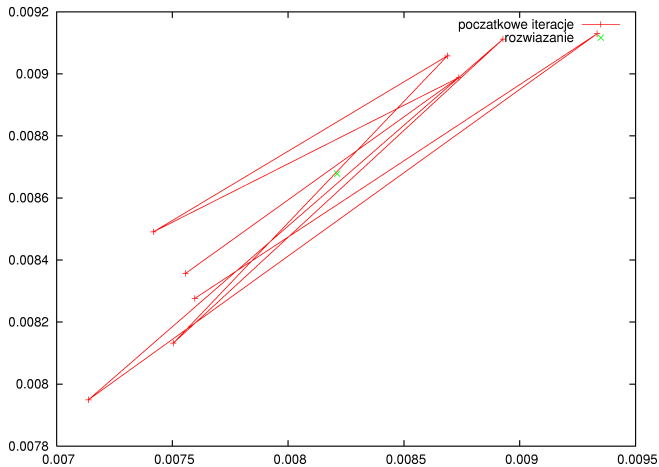
Tabela: Skrajne wartości własne macierzy B dla klastra fullerenu w zależności od rozmiaru klastra

Rozmiar klastra	λ_{min}	λ_{max}
3x3x3	-0.48267891	0.95545685
5x5x5	-0.51202888	1.02006632
7x7x7	-0.52105971	1.03999486

Wymuszanie zbieżności

- Jeżeli warunek $\max |\lambda_k| < 1$ nie jest spełniony, do uzbieźnienia konieczne jest stosowanie jakiegoś rodzaju dampingu
- Polegać to może na:
 - Nieuaktualnianiu niektórych (losowych) elementów w kolejnych iteracjach
 - Użyciu do kolejnej iteracji kombinacji liniowej wyników z bieżącej i poprzednich iteracji (bezwładność)
 - ...

Oscylacje



Wnioski

- SCPF nie zawsze jest procesem zbieżnym
- Stosowanie dampingu zwykle prowadzi do uzbieźnienia
- Oscylacje powodują, że liczba iteracji jest duża

Szukajcie a znajdziecie

- Szukamy metody, która
 - jest zbieżna
 - nie popada w oscylacje
- Gotowe rozwiązania do znalezienia w najbliższej dobrej książce do numerycznej algebry liniowej

Szukajcie a znajdziecie

- Szukamy metody, która
 - jest zbieżna
 - nie popada w oscylacje
- Gotowe rozwiązania do znalezienia w najbliższej dobrej książce do numerycznej algebry liniowej

Residuum

- Residuum

$$r = b - Ax \quad (16)$$

mówi na ile różny od dokładnego wynik daje wektor próbny

- Na residuum można też patrzeć jak na wektor błędu przetransformowany do przestrzeni rozwiązań

$$r = Ae \quad (17)$$

Prosta minimalizacja normy residuum I

Kierunek najszybszego spadku

$$\nabla(r^T r) = 2Ar \quad (18)$$

Schemat iteracyjny

Ewolucja w kierunku najszybszego spadku normy residuum daje schemat iteracyjny dla residuum

$$r_{(k+1)} = r_{(k)} - \alpha_k Ar_{(k)} \quad (19)$$

i wynikający z niego schemat dla wektorów próbnych

$$x_{(k+1)} = x_{(k)} + \alpha_k r_{(k)} \quad (20)$$

Prosta minimalizacja normy residuum II

Wyznaczenie parametru α

Przyrównując do zera pochodną

$$\frac{\partial}{\partial \alpha_k} r_{(k+1)}^T r_{(k+1)} = -2r_{(k)}^T A r_{(k)} + 2\alpha_k r_{(k)}^T A A r_{(k)} \quad (21)$$

dostajemy wzór na wartość parametru α w danej iteracji

$$\alpha_k = \frac{r_{(k)}^T A r_{(k)}}{r_{(k)}^T A A r_{(k)}} \quad (22)$$

Prosta minimalizacja normy residuum III

Ewolucja normy residuum

$$r_{(k+1)}^T r_{(k+1)} = r_{(k)}^T r_{(k)} - \frac{\left(r_{(k)}^T A r_{(k)}\right)^2}{r_{(k)}^T A A r_{(k)}} \leq r_{(k)}^T r_{(k)} \quad (23)$$

Zbieżność

- Metoda powinna być zbieżna dla nieosobliwych macierzy współczynników
- Nic nie zapobiega oscylacjom

Przestrzeń Kryłowa

Definicja

Dla danego wektora x i macierzy A przestrzeń Kryłowa $\mathcal{K}_k(A, x)$ to podprzestrzeń rozpięta przez działanie kolejnymi potęgami macierzy A na wektor x

$$\mathcal{K}_k(A, x) = \text{span}\{x, Ax, \dots, A^{k-1}x\} \quad (24)$$

Właściwości

Można pokazać, że dla nieosobliwej macierzy A mającej s różnych wartości własnych

$$\hat{x} = A^{-1}b \in \mathcal{K}_s(A, b) \quad (25)$$

Przestrzeń Kryłowa

Definicja

Dla danego wektora x i macierzy A przestrzeń Kryłowa $\mathcal{K}_k(A, x)$ to podprzestrzeń rozpięta przez działanie kolejnymi potęgami macierzy A na wektor x

$$\mathcal{K}_k(A, x) = \text{span}\{x, Ax, \dots, A^{k-1}x\} \quad (24)$$

Właściwości

Można pokazać, że dla nieosobliwej macierzy A mającej s różnych wartości własnych

$$\hat{x} = A^{-1}b \in \mathcal{K}_s(A, b) \quad (25)$$

Ortogonalność

Startując z początkowym wektorem próbnym $x_{(0)}$, kolejne iteracje $x_{(k)}$ generowane są w taki sposób aby minimalizować normę residuum zgodnie z

$$\|r_{(k)}\| = \min_{s \in AK_k(A, r_{(0)})} \|r_{(0)} - s\| \quad (26)$$

Ponieważ $r_{(0)}$ jest rzutowane na $AK_k(A, r_{(0)})$, dla każdego residuum zachodzi

$$r_{(k)} \perp AK_k(A, r_{(0)}) \quad (27)$$

Macierz współczynników

Równania na momenty dipolowe mogą być przedstawione w postaci

$$A\mu = b \quad (28)$$

gdzie

$$A = \begin{pmatrix} I & -\alpha_1 T_{12} & \dots & -\alpha_1 T_{1n} \\ -\alpha_2 T_{12} & I & & \\ \vdots & & \ddots & \vdots \\ -\alpha_n T_{1n} & -\alpha_n T_{2n} & \dots & I \end{pmatrix} \quad (29)$$

$$b = \begin{pmatrix} \alpha_1 E_1 \\ \alpha_2 E_2 \\ \vdots \\ \alpha_n E_n \end{pmatrix} \quad (30)$$

Nieosobliwy tensor polaryzowalności I

Jeżeli wszystkie tensory polaryzowalności α_i są odwracalne, możemy układ równań (28) przekształcić do postaci

$$\alpha \tilde{A} \mu = \alpha E \quad (31)$$

$$\tilde{A} = \begin{pmatrix} \alpha_1^{-1} & -T_{12} & \dots & -T_{1n} \\ -T_{12} & \alpha_2^{-1} & & \\ \vdots & & \ddots & \vdots \\ -T_{1n} & -T_{2n} & \dots & \alpha_n^{-1} \end{pmatrix} \quad (32)$$

$$\alpha = \begin{pmatrix} \alpha_1 & 0 & \dots & 0 \\ 0 & \alpha_2 & & \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & \alpha_n \end{pmatrix} \quad (33)$$

Nieosobliwy tensor polaryzowalności II

Dzięki tej transformacji momenty dipolowe możemy obliczyć rozwiązując układ równań

$$\tilde{A}\mu = E \quad (34)$$

zamiast oryginalnego układu (28)

$$A\mu = \alpha E \quad (35)$$

Kluczowa różnica pomiędzy tymi postaciami polega na tym, że w przeciwieństwie do macierzy A , macierz \tilde{A} jest symetryczna.

Ogólny przypadek

W ogólnym przypadku konieczne jest przetransformowanie układu równań (28) do układu osi głównych tensorów polaryzowalności. Wtedy

- równania na (równe zero) składowe momentów dipolowych związane z zerowymi składowymi polaryzowalności mogą zostać wyeliminowane z rozwiązywanego układu równań
- pozostałe równania mogą zostać sprowadzone do postaci symetrycznej

Uwzględnienie osobliwych tensorów polaryzowalności komplikuje implementację, ale nie zwiększa kosztu obliczeniowego.

Rekurencja Lanczosa

- Dla niesymetrycznej macierzy współczynników rozbudowa przestrzeni Kryłowa wymaga znajomości wszystkich dotychczasowych elementów
- W przypadku macierzy symetrycznej istnieje prosta trzelementowa rekurencja pozwalająca na rozbudowę przestrzeni Kryłowa bez trzymania wszystkich dotychczasowych elementów

Uproszczony szkic algorytmu

$$v_{(0)} = r_{(0)} \quad w_{(0)} = Av_{(0)} \quad \delta_0 = r_{(0)}^T AAr_{(0)} \quad (36)$$

$$\delta'_k = w_{(k)}^T w_{(k)} \quad (37)$$

$$\omega_k = \delta_k / \delta'_k \quad (38)$$

$$x_{(k+1)} = x_{(k)} + \omega_k v_{(k)} \quad (39)$$

$$r_{(k+1)} = r_{(k)} - \omega_k w_{(k)} \quad (40)$$

$$\delta_{k+1} = r_{(k+1)}^T AAr_{(k+1)} \quad (41)$$

$$\psi_k = -\delta_{k+1} / \delta_k \quad (42)$$

$$v_{(k+1)} = r_{(k+1)} - \psi_k v_{(k)} \quad (43)$$

$$w_{(k+1)} = Ar_{(k+1)} - \psi_k w_{(k)} \quad (44)$$

Implementacja

- Algorytm w przedstawionej postaci nie jest numerycznie stabilny
- MINRES: matematycznie równoważny i numerycznie stabilny

Analiza zbieżności

- Jeżeli założymy że
 - macierz współczynników jest nieosobliwa
 - obliczenia są wykonywane w dokładnej arytmetycemetoda sprzężonych residuów powinna być zbieżna w co najwyżej n krokach, gdzie n jest rozmiarem rozwiązywanego układu
- Zwykle zbieżność jest o wiele lepsza. Konkretnie, dla wskaźnika uwarunkowania κ macierzy A wynoszącego

$$\kappa = \frac{\min |\lambda|}{\max |\lambda|} \quad (45)$$

zachodzi

$$\frac{\|r(k)\|}{\|r(0)\|} \leq \left(\frac{1 - \sqrt{\kappa}}{1 + \sqrt{\kappa}} \right)^k \quad (46)$$

Porównanie kosztu omawianych metod

Tabela: Liczba operacji na iterację

Metoda	Mv	dot	axpy
SCPF	1	0	1
MINRES	1	2	5

Tabela: Oszacowanie względnej zmiany normy residuum z liczbą iteracji

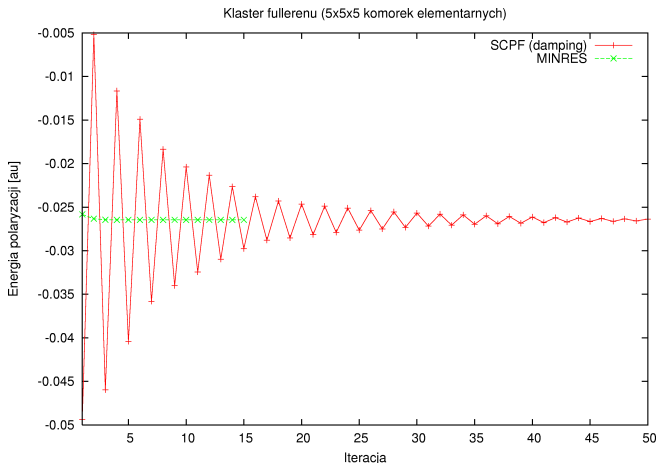
Metoda	Liczba iteracji
SCPF	$\kappa\rho(B)^k$
MINRES	$\left(\frac{1-\sqrt{\kappa}}{1+\sqrt{\kappa}}\right)^k$

Porównanie zbieżności energii

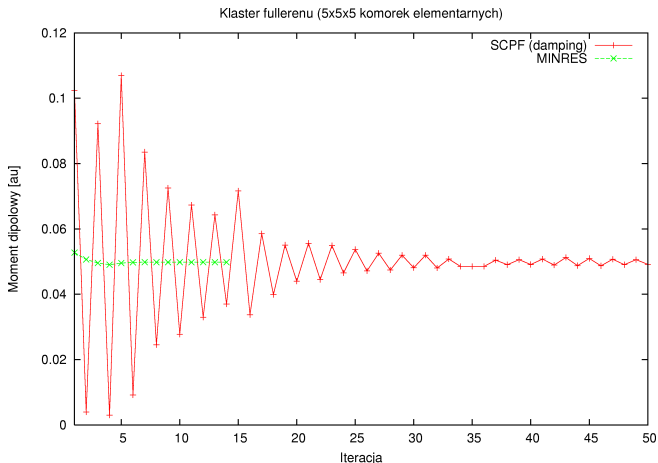
Tabela: Liczba iteracji konieczna do uzbieźnienia energii polaryzacji klastra fullerenu w zależności od rozmiaru klastra

Rozmiar klastra	SCPF	SCPF (damping=0.05)	MINRES
3x3x3	302	92	12
5x5x5	–	153	14
7x7x7	–	208	13
9x9x9	–	230	13
11x11x11	–	246	13

Porównanie zbieżności energii



Porównanie zbieżności momentów dipolowych



Wnioski

- Zastosowanie algorytmu MINRES w schemacie SCPF:
 - pozwala uniknąć problemów ze zbieżnością
 - zmniejsza liczbę iteracji o rząd wielkości
- Złożoność obliczeniowa pojedynczej iteracji w obu podejściach, zarówno formalna jak i praktyczna, jest taka sama. W rezultacie ograniczenie liczby iteracji bezpośrednio przekłada się na skrócenie czasu obliczeń