

ROZDZIAŁ 7. SUPLEMENT 1

METODA DYFUZYJNA

Opisana metoda¹ jest ogólna, można ją stosować w każdej dziedzinie, w której pojawiają się problemy minimalizacji funkcji o wielu minimach. Jest to metoda przybliżona, nie dająca gwarancji, że znalezione minimum jest minimum globalnym².

Podstawą metody jest idea deformacji hiperpowierzchni energii potencjalnej V

w taki sposób, aby najpierw (przy mniejszym parametrze deformacji) zaczęły znikać płytkie studnie potencjału, podczas gdy inne studnie rosłyby ich kosztem. Przy dalszej deformacji (większy parametr deformacji) zaczęłyby znikać głębsze minima. W tej metodzie zakłada się, że *jedyna* studnia, która pozostaje przy dużej deformacji, pochodzi od najgłębszego minimum (czyli globalnego) oryginalnej hiperpowierzchni energii potencjalnej.

Podczas takiej deformacji studnie potencjału zmniejszają stopniowo swoją głębokość, zmieniają swój kształt i położenie. To powoduje, że położenie jedynego minimum, uzyskanego przy dużej deformacji, nie odpowiada dokładnie pozycji minimum globalnego (jeśli nasze założenie o tym, że przeżyje minimum globalne, jest słuszne).

Procedura powrotu

Pozycję minimum globalnego pierwotnej funkcji staramy się odtworzyć, postępując w następujący sposób. Minimalizując energię przy dostatecznie dużej wartości parametru deformacji, otrzymujemy położenie jedynego minimum. To położenie stanowi dla nas punkt startowy w *procedurze powrotu*. Procedura powrotu polega na tym, że zmniejszamy o pewną małą wartość parametr deformacji i minimalizujemy energię przy *nowej* wartości parametru, startując jednak z położenia w przestrzeni położenia jąder określonego przy użyciu *starej* (większej) wartości tego parametru. Uzyskujemy nową konfigurację jąder. Powtarzając ten krok wielokrotnie, wykreślamy trajektorię w przestrzeni położenia jąder jako funkcję zmniejszanego parametru deformacji. Na końcu parametr deformacji jest równy zero i otrzymujemy położenie końcowe wspomnianej trajektorii. Jest to punkt startowy minimalizacji wyjściowej (niezdeformowanej) funkcji,

¹Na tę ideę wpadłem, gdy leżałem chory na grype w łóżku i przemyślałem, jakby tu dobrać się do minimum globalnego. Koc przede mną miał liczne minima lokalne i jedno globalne. Gdy zacząłem naciągać koc, najpierw zaczęły znikać małe dolinki, a w końcu zostało jedno minimum — ostatni ślad minimum globalnego. Wtedy bardzo łatwo mogłem tam trafić, zjeżdżając z dowolnego punktu koca.

W książce „*The Character of Physical Law*” (Cox and Wyman Ltd, London 1965) Richard Feynman napisał słowa, które bardzo pasują do mojej przygody z deformacją funkcji: „*Moja idea wydawała mi się tak logiczna i tak elegancka, że zakochałem się w niej bez pamięci. A przecież zakochać się bez pamięci w dziewczynie można tylko wtedy, gdy jeszcze jej dobrze nie znasz i nie widzisz wszystkich jej wad. Wady zobaczysz później, ale miłość jest już dostatecznie silna, żeby cię zatrzymać*” (tłum. L. P.).

²Nie jest to szczególna cecha tej metody — taka metoda nie istnieje.

w której osiągamy dno lokalnego basenu V . Mamy nadzieję, że uzyskane minimum jest minimum globalnym. W jakim stopniu założenie to jest słuszne, pokażą przykłady numeryczne.

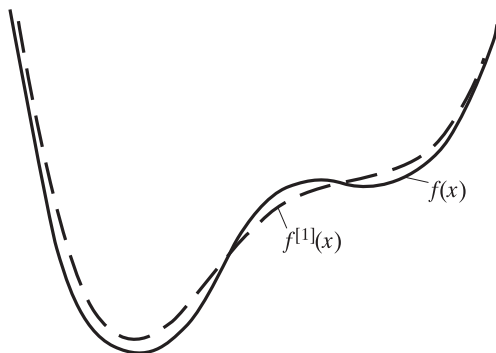
Jak deformować funkcję

Po tym opisie głównej idei metody przystąpmy do jej praktycznego sformułowania. Łatwo wygładzić koc, zasadniczą rzeczą jest pokazanie, w *jaki sposób wygładzić pierwowrotną funkcję*³.

Sformułowanie teorii podamy wpierw dla funkcji jednej zmiennej, a potem uogólnimy otrzymane rezultaty na przypadek wielu zmiennych. Weźmy funkcję $f(x)$ i wykonajmy na niej prostą operację:

$$f^{[1]}(x) = f(x) + \beta f''(x) \quad \text{dla} \quad \beta > 0. \quad (1)$$

Zwróćmy uwagę, że transformacja $f \rightarrow f^{[1]}$ *destabilizuje* (rys.1) *minima funkcji* f . Istotnie, pod wpływem transformacji 1 punkty przegięcia nie zmieniają swoich wartości (bo $f'' = 0$), podczas gdy te części krzywej f , które są wklęsłe (wypukłe), ulegają podwyższeniu (obniżeniu).



Rys.1. Destabilizacja basenów funkcji przez dodanie jej drugiej pochodnej. Krzywa przerywana $f^{[1]}(x)$ powstała przez dodanie do krzywej ciągłej $f(x)$ zaburzenia $\beta f''(x)$. Nowa krzywa ma już tylko jeden basen

Destabilizacja jest tym pewniejsza⁴, im mniejsze jest β . Jest to spełnione, gdy $\beta = t/N$, gdzie $t > 0$ jest pewnym parametrem, zwanym parametrem deformacji, a N jest bardzo duże. Gdy N dąży do nieskończoności i procedura jest zbieżna, otrzymuje się w granicy funkcję $F(x, t)$:

$$F(x, t) = \lim_{N \rightarrow \infty} \left(1 + \frac{t}{N} \frac{d^2}{dx^2} \right)^N f(x) = \exp \left(t \frac{d^2}{dx^2} \right) f(x), \quad (2)$$

przy czym $\exp(\hat{A})$ dla operatora \hat{A} oznacza szereg Taylora $1 + A + \frac{1}{2!}A^2 + \frac{1}{3!}A^3 + \dots$

³Tego jeszcze podczas grypy nie wiedziałem. W roku 1987 byłem na Uniwersytecie w Namur w Belgii i kupiłem dla syna komputer, którym wieczorem się bawiłem. Podczas zabawy z wykresami funkcji zobaczyłem to, co jest opisane poniżej we wzorach (1) i (2).

⁴ β nie może być zbyt duże. W skrajnym przypadku prowadziłoby to do zastąpienia funkcji f funkcją f'' , a to jest niecelowe, bo ta ostatnia *nie jest* wygładzoną funkcją f .

Operator

$$\hat{T}(t) = \exp\left(t \frac{d^2}{dx^2}\right) \quad (3)$$

ma kilka interesujących właściwości. Po pierwsze, jego funkcjami własnymi są funkcje $\sin \omega x$ i $\cos \omega x$:

$$\hat{T}(t) \sin \omega x = a(\omega, t) \sin \omega x, \quad (4)$$

$$\hat{T}(t) \cos \omega x = a(\omega, t) \cos \omega x. \quad (5)$$

Najciekawsza jest wartość własna $a(\omega, t)$:

$$a(\omega, t) = \exp(-\omega^2 t). \quad (6)$$

Istotnie,

$$\begin{aligned} \hat{T}(t) \sin \omega x &= \exp\left(t \frac{d^2}{dx^2}\right) \sin \omega x = \\ &= \left[1 + t \frac{d^2}{dx^2} + \frac{1}{2!} \left(t \frac{d^2}{dx^2}\right)^2 + \frac{1}{3!} \left(t \frac{d^2}{dx^2}\right)^3 + \dots \right] \sin \omega x = \\ &= \sin \omega x - t\omega^2 \sin \omega x + \frac{1}{2!} t^2 \omega^4 \sin \omega x + \dots = \\ &= \sin \omega x \left(1 - t\omega^2 + \frac{1}{2!} t^2 \omega^4 - \frac{1}{3!} t^3 \omega^6 + \dots \right) = \\ &= \sin \omega x \exp(-\omega^2 t). \end{aligned}$$

Podobny wynik uzyskalibyśmy dla $\cos \omega x$.

Ze wzorów (4) i (5) wynika, że operator $\hat{T}(t)$ dla $t > 0$ powoduje zmniejszenie amplitudy funkcji sinus i cosinus. Ważne jest to, że to spłaszczenie jest tym większe, im wyższej częstotliwości ω odpowiada funkcja i im większy jest parametr deformacji t . Konsekwencją tego faktu jest to, że jeśli operator $\hat{T}(t)$ zadziała na funkcję $f(x)$ przedstawioną w postaci szeregu Fouriera, to przy wzroście parametru t najszybciej będą znikać składowe o wysokiej częstotliwości. \hat{T} jest więc „mordercą wysokich częstotliwości”.

Równanie dyfuzji

Bardzo łatwo przekonać się, że jeśli szereg Taylora w równ. (2) jest zbieżny, to:

funkcja $F(x, t)$ spełnia równanie dyfuzji

$$\frac{\partial^2 F}{\partial x^2} = \frac{\partial F}{\partial t}, \quad (7)$$

w którym parametr t przyjmuje znaczenie *czasu*⁵, a warunkiem brzegowym jest⁶ $F(x, 0) = f(x)$.

Alternatywne postacie transformacji

Jest kilka alternatywnych postaci funkcji⁷ $F(x, t)$. Wyraża się ona

- albo jako rezultat działania operatora \hat{T} na funkcję f ,
- albo jako rozwiązanie równania dyfuzji,
- albo w bardzo przydatnej postaci jako *splot f z funkcją gaussowską*:

$$F(x, t) = \frac{1}{\sqrt{4\pi t}} \int_{-\infty}^{+\infty} f(y) \exp\left[-\frac{(x-y)^2}{4t}\right] dy. \quad (8)$$

Jeśli wpatrzeć się w funkcję podcałkową, to można zauważyć, że wygładzanie funkcji f (rezultat to F) polega na lokalnym *uśrednianiu* jej wartości w każdym punkcie x z wagą gaussowską, przy czym im większe t , tym szerszy zasięg uśredniania. Gdy $t = 0$, funkcja gaussowska przechodzi w funkcję δ Diraca, co oznacza brak uśredniania, $F(x, 0) = f(x)$.

⁵Jest to nieoczekiwane. Problem jest czysto matematyczny i nagle pojawia się w nim czas!

⁶Równanie (7) sprawdzimy w następujący sposób:

$$F(x, t) = \exp\left(t \frac{\partial^2}{\partial x^2}\right) f(x) = \left[1 + t \frac{\partial^2}{\partial x^2} + \frac{1}{2!} \left(t \frac{\partial^2}{\partial x^2}\right)^2 + \frac{1}{3!} \left(t \frac{\partial^2}{\partial x^2}\right)^3 + \dots\right] f(x)$$

Prawa strona równ. (7):

$$\begin{aligned} \frac{\partial F}{\partial t} &= \left[0 + \frac{\partial^2}{\partial x^2} + \frac{1}{2!} 2t \left(\frac{\partial^2}{\partial x^2}\right)^2 + \frac{1}{3!} 3t^2 \left(\frac{\partial^2}{\partial x^2}\right)^3 + \dots\right] f(x) = \\ &= \left[1 + t \frac{\partial^2}{\partial x^2} + \frac{1}{2} t^2 \left(\frac{\partial^2}{\partial x^2}\right)^2 + \dots\right] \frac{\partial^2 f}{\partial x^2} = \exp\left(t \frac{\partial^2}{\partial x^2}\right) \frac{\partial^2 f}{\partial x^2}. \end{aligned}$$

Lewa strona równ. (7):

$$\begin{aligned} \frac{\partial^2 F}{\partial x^2} &= \frac{\partial}{\partial x} \frac{\partial}{\partial x} \left[1 + t \frac{\partial^2}{\partial x^2} + \frac{1}{2!} \left(t \frac{\partial^2}{\partial x^2}\right)^2 + \frac{1}{3!} \left(t \frac{\partial^2}{\partial x^2}\right)^3 + \dots\right] f(x) = \\ &= \left[\frac{\partial^2}{\partial x^2} + t \frac{\partial^4}{\partial x^4} + \frac{1}{2!} \left(t \frac{\partial^2}{\partial x^2}\right)^3 + \frac{1}{3!} \left(t \frac{\partial^2}{\partial x^2}\right)^4 + \dots\right] f(x) = \\ &= \left[1 + \left(t \frac{\partial^2}{\partial x^2}\right) + \frac{1}{2!} t^2 \left(\frac{\partial^2}{\partial x^2}\right)^2 + \dots\right] \frac{\partial^2 f}{\partial x^2} = \exp\left(t \frac{\partial^2}{\partial x^2}\right) \frac{\partial^2 f}{\partial x^2}. \end{aligned}$$

A więc lewa strona równ. (7) równa się stronie prawej, c.b.d.o.

⁷Postaci te różnią się domenami funkcji, do których procedury wygładzania mogą być aplikowane.

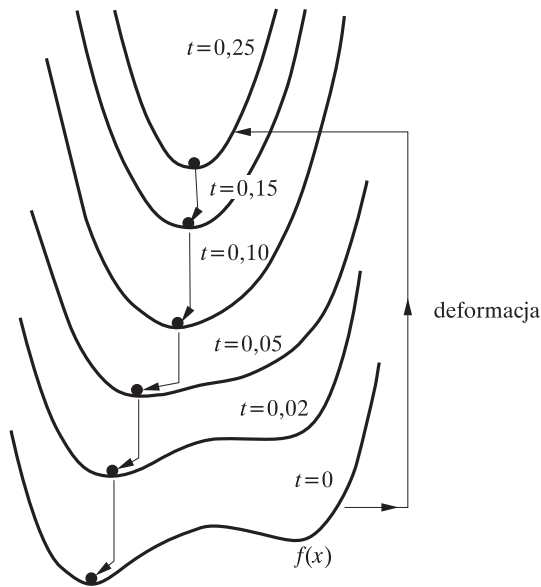
Funkcja $F(x, t)$ przyjęła, dzięki interpretacji poprzez równanie dyfuzji, znaczenie *temperatury* (gdy rozpatrujemy przepływ ciepła) lub znaczenie stężenia (gdy rozpatrujemy transport substancji w ośrodku). Ta interpretacja fizyczna sugeruje od razu, że w pewnych przypadkach minimum globalnego tą metodą znaleźć się nie da. Na przykład jeśli minimum globalne odpowiada wąskiej studni, to może się zdarzyć, że przy wzroście czasu ta właśnie studnia potencjału znika, podczas gdy inna, może nie tak głęboka, ale dostatecznie szeroka, pozostaje. Można jednak zadać pytanie, czy chemika powinny interesować wąskie studnie potencjału (nawet jeśli są głębokie). Istotnie, szerokie studnie (dużo stanów oscylacyjnych) są faworyzowane entropowo, co powoduje obniżenie energii swobodnej.

Ograniczmy się do zastosowania przedstawionej metody do dwóch przykładów nie związanych z chemią.

Przykład 1. Funkcja $f(x) = x^4 + ax^3 + bx^2$, $a = 2$, $b = 0,9$ ma dwa minima: lokalne dla $x = 0$ i globalne dla $x = -1,085$ oraz jedno maksimum dla $x = -0,145$. Zastosowanie operatora $\hat{T}(t)$ do funkcji f daje:

$$\hat{T}(t)f(x) = x^4 + ax^3 + (12t + b)x^2 + 6tax + 12t^2 + 2bt, \quad (9)$$

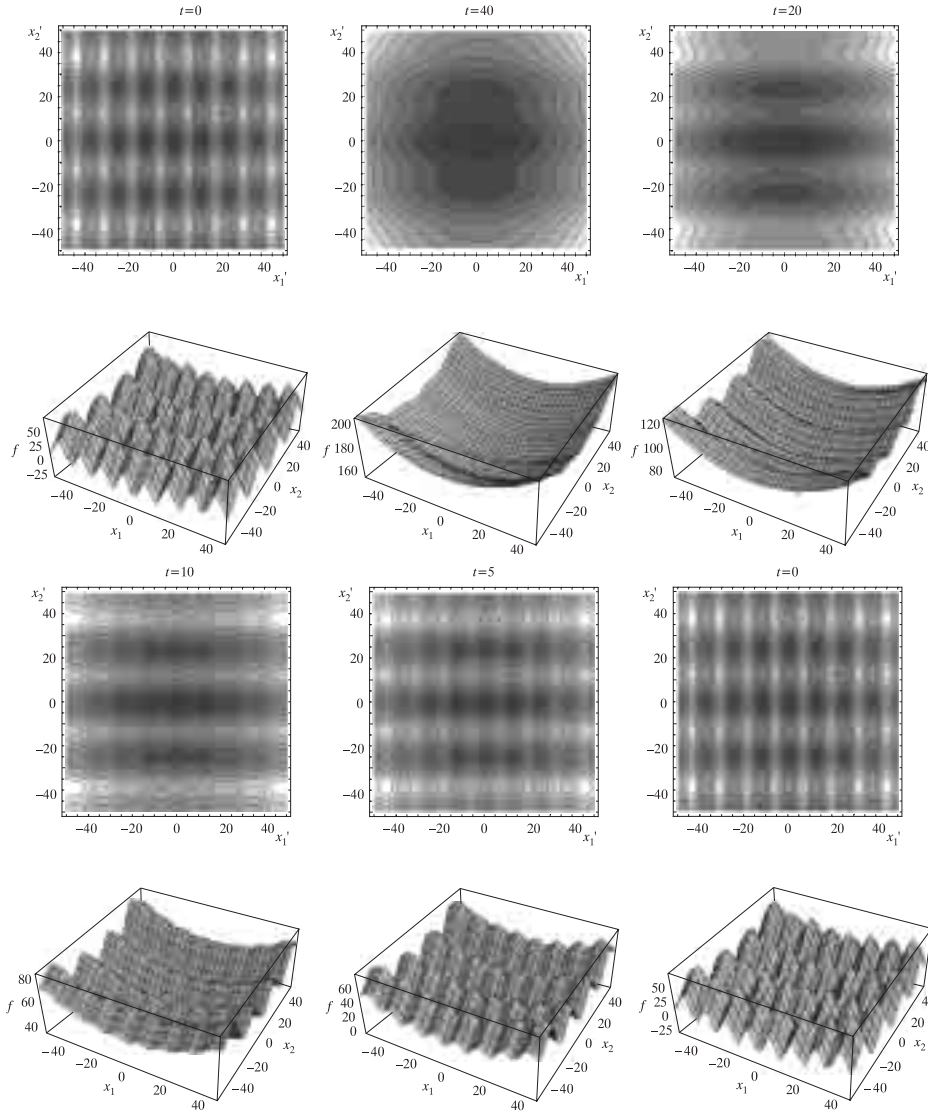
co oznacza, że ze wzrostem t wykres funkcji deformuje się i pędzi do góry z prędkością początkową $2b$ i przyśpieszeniem 24. Gdy t jest bardzo duże, wykres upodabnia się do paraboli $12tx^2 + 6tax$, która wykazuje minimum przy $-a/4$. Z tego minimum poprzez procedurę powrotu osiągamy minimum przy $x = -1,085$, czyli minimum globalne (rys.2).



Rys 2. Deformacja oryginalnego potencjału $f(x)$ przez operator $\hat{T}(t)$ i procedura powrotu. Deformacja przez operator $\hat{T}(t = 0,25)$ prowadzi do krzywej z jednym minimum osiągalnym z *dowolnego* punktu przestrzeni przez prostą minimalizację. Następnie stosowana jest procedura powrotu (pokazana strzałkami skierowanymi do dołu) poprzez sekwencję zdeformowanych krzywych: $\hat{T}(0,15)f(x)$, $\hat{T}(0,10)f(x)$, $\hat{T}(0,05)f(x)$, $\hat{T}(0,02)f(x)$ i na koniec $\hat{T}(0)f(x) = f(x)$. Po każdym kroku procedury dokonywana jest minimalizacja symbolizowana na rysunku kulką staczającą się z pozycji minimum krzywej górnej i osiągnącą pozycję minimum krzywej dolnej. W ostatnim kroku znajdujemy minimum globalne

Przykład 2. W literaturze dotyczącej globalnej minimalizacji poszukiwano minimum globalnego tzw. funkcji Bohachevskiego (rys.3):

$$f(x_1, x_2) = ax_1^2 + bx_2^2 + c \cos \gamma x_1 + d \cos \delta x_2 \quad a, b > 0; \quad c, d < 0. \quad (10)$$



Rys.3. Deformacja funkcji Bohachevskiego $f(x_1, x_2) = 0,01(x_1^2 + ax_2^2) - 20 \cos \frac{1}{2}x_1 - 20 \cos \frac{1}{4}x_2$ przez operator $\hat{T}(t)$ i procedura powrotu. Deformacja zastosowana przez operator $\hat{T}(0, 40)$ prowadzi do powierzchni z jednym minimum osiągalnym z *dowolnego* punktu przestrzeni przez prostą minimalizację. Następnie stosowana jest procedura powrotu poprzez sekwencję zdeformowanych powierzchni: $\hat{T}(20)f(x_1, x_2)$, $\hat{T}(10)f(x_1, x_2)$, $\hat{T}(5)f(x_1, x_2)$ i na koniec $\hat{T}(0)f(x_1, x_2) = f(x_1, x_2)$. Po każdym kroku procedury dokonywana jest minimalizacja, startująca z minimum poprzedniej powierzchni i osiągająca pozycję minimum powierzchni następnej. W ostatnim kroku znajdujemy minimum globalne funkcji Bohachevskiego

Metoda dyfuzyjna osiąga rozwiązanie w następujący sposób. Po zadziałaniu na funkcję f operatorem $\hat{T}(t)$ otrzymujemy funkcję $F(x_1, x_2, t)$:

$$F(x_1, x_2, t) = a(x_1^2 + 2t) + b(x_2^2 + 2t) + c \exp(-\gamma^2 t) \cos \gamma x_1 + d \exp(-\delta^2 t) \cos \delta x_2. \quad (11)$$

Gdy t dąży do nieskończoności, człony trzeci i czwarty po prawej stronie równ. (11) dążą do zera. Pozostałe wyrazy przedstawiają parabolę dwuwymiarową o niezmienniej pozycji minimum [w punkcie $(0, 0)$]. Ten właśnie punkt odpowiada położeniu minimum globalnego i zastosowanie procedury powrotu (w tym przypadku banalnej) wskazałoby ten właśnie punkt. Bohachevsky i inni⁸ otrzymują ten sam rezultat, stosując pewną udoskonaloną procedurę stochastyczną.

Te i podobne przykłady dla funkcji w wielu wymiarach, a także przykłady z klastrami atomowymi wskazują, że w wielu przypadkach metodą dyfuzyjną znajduje się minimum globalne. Czasem zamiast minimum globalnego metoda wskazuje szeroką studnię potencjału o niskiej, choć nie najniższej energii.

Reguły ewolucji

Idea metody dyfuzyjnej⁹ została później związana z równaniami ewolucji pewnych wielkości fizycznych — równaniem Schrödingera ewolucji funkcji falowej w *czasie*¹⁰, równaniem Blocha ewolucji rozkładu kanonicznego przy zmianie *temperatury*¹¹, równaniami Smoluchowskiego i Fokkera–Plancka ewolucji pewnego rozkładu cząstek w *czasie*¹² i zasadą wariacyjną Gibbsa dla energii swobodnej¹³.

Równania ewolucji były potrzebne, aby początkowo zaproponowane położenie pewnej *próbnej funkcji gaussowskiej* zaczęło się zmieniać w kierunku wskazywanym przez określone równanie fizyki *i osiągnęło na końcu pozycję odpowiadającą minimum globalnemu potencjału $V(x)$.*

Zwracało uwagę to, że mimo rozpatrywania szerokiej gamy zjawisk *otrzymywane równania na trajektorię pozycji i szerokości funkcji gaussowskiej były podobne do siebie!* Te obserwacje doprowadziły do uogólnienia, wyjaśniającego w pewnym stopniu, dlaczego przyroda nie ma wielkich problemów z globalną optymalizacją¹⁴. Oto w skrócie ta idea zastosowana do ewolucji położenia i szerokości funkcji gaussowskiej.

⁸I. O. Bohachevsky, M. E. Johnson, M. L. Stein, *Technometrics*, **28** (1986) 209. Trzeba przyznać, że funkcja Bohachevskiego jest szczególnie łatwym łupem metody dyfuzyjnej. Jest tak jeszcze w przypadku kilku podobnych funkcji. Metoda natrafia jednak na poważne trudności przy zagadnieniach o praktycznym znaczeniu w chemii (np. dużych molekułach wieloatomowych).

⁹L. Piel, J. Kostrowicki, H. Scheraga, *J. Phys. Chem.*, **93** (1989) 3339.

¹⁰P. Amara, D. Hsu, J. E. Straub, *J. Phys. Chem.*, **97** (1993) 6715.

¹¹J. Ma, J. E. Straub, *J. Chem. Phys.*, **101** (1994) 533.

¹²J. E. Straub, J. Ma, P. Amara, *J. Chem. Phys.*, **103** (1995) 1574.

¹³S. Schelstraete i H. Verschelde, *J. Chem. Phys.*, **106** (1997) 1556, wykorzystali do globalnej optymalizacji zasadę wariacyjną Gibbsa mówiącą, że entalpia swobodna (energia swobodna Gibbsa) to minimum pewnego funkcjonału w zbiorze znormalizowanych rozkładów gęstości $\rho(x)$, który autorzy aproksymowali funkcją gaussowską o zmiennej pozycji i szerokości. Startując z szerokiej funkcji gaussowskiej o dowolnym położeniu, dostaje się (podczas minimalizacji funkcjonału energii swobodnej) trajektorię jej położenia i trajektorię jej szerokości. Obie trajektorie zawierają funkcję $F(x, t)$, z potencjałem $V(x)$ jako odpowiednikiem $f(x)$. Oprócz jednak wyrażenia na F występuje tam człon entropowy. Związanie metody dyfuzyjnej z entropią i entalpią swobodną jest wielkim osiągnięciem autorów.

¹⁴L. Piel, *Collect. Czech. Chem. Commun.*, **63** (1998) 1368.

Załóżmy, że postęp optymalizacji kontroluje parametr α , który zmienia się od 0 do wartości maksymalnej $\alpha_{\max} > 0$ w sposób „ziarnisty”, tzn. tyka „zegar optymalizacji”. Za każdym tyknięciem α zwiększa się o $d\alpha = \alpha_{\max}/N$, gdzie N jest liczbą kroków optymalizacji, i zmienia się pozycja x_0 oraz „szerokość” mierzona parametrem t funkcji gaussowskiej. O ile się zmieniają te dwie wielkości? Okazuje się, że w różnych wspomnianych procedurach różnie, ale według bardzo podobnych równań różniczkowych:

$$\frac{\partial x_0}{\partial \alpha} = -A \left(\frac{\partial F(x, t)}{\partial x} \right)_{x=x_0}, \quad (12)$$

$$\frac{\partial t}{\partial \alpha} = B - C \left(\frac{\partial^2 F(x, t)}{\partial x^2} \right)_{x=x_0}, \quad (13)$$

w których $A, B, C \geq 0$!

W skrócie można powiedzieć, że strategia stojąca za wspomnianymi równaniami ewolucji jest następująca:

- Należy używać nie potencjału V , lecz V *uśrednionego* przez spłot $f = V$ z funkcją gaussowską według wzoru (8), czyli $F(x, t)$. Oba równania mówią, że parametr uśrednienia, czyli „czas” t , jest funkcją położenia x_0 funkcji gaussowskiej.
- Jak wynika z równ. (12), po starcie z jakiegoś położenia x_0 trzeba zawsze *iść w dół* uśrednionego potencjału $F(x, t)$, czyli kierunek marszu to $-\left(\frac{\partial F}{\partial x}\right)_{x=x_0}$.
- Jak wynika z równ. (13), uśrednienie przez spłot z funkcją gaussowską o szerokości t staje się mniej istotne dla x_0 w pobliżu minimów [dodatnia wartość krzywizny krzywej, czyli drugiej pochodnej we wzorze (13) powoduje zmniejszenie t] i bardzo istotne dla x_0 w pobliżu maksimów (ujemna wartość tej drugiej pochodnej powoduje zwiększenie t). Dzięki ostatniej właściwości to maksimum może zacząć znikać, a wtedy przekroczy barierę dzielącą dwa minima i możemy dotrzeć do minimum niższego!
- Z równania (13) wynika, że jeśli znajdujemy się na płaskowyzu lub w basenie o małej krzywiznie krzywej, to zakres uśrednienia t zaczyna się zwiększać¹⁵. Oznacza to, że, podczas zwiększania t , uśrednienie może objąć pozycję jakiegoś dalekiego minimum, a wtedy $-\left(\frac{\partial F}{\partial x}\right)_{x=x_0}$ może nas do tego minimum zaprowadzić przez równ. (12).

Wydaje się, że ten algorytm nie robi tylko jednej rzeczy, którą robi przyroda: nie potrafi podzielić jednego zwartego rozkładu na kilka takich rozkładów. Ta modyfikacja została wprowadzona w metodzie dyfuzyjnej przez Davida Shallowaya¹⁶.

¹⁵Jest to „polowanie” na dalekie minima.

¹⁶D. Shalloway, w „*Recent Advances in Global Optimization*”, red. A. Floudas i P. M. Pardalos, Princeton University, Princeton 1992.